Andrew Emili · Jack Greenblatt
Shoshana Wodak *Editors*

Systems Analysis of Chromatin-Related Protein Complexes in Cancer

Systems Analysis of Chromatin-Related Protein Complexes in Cancer

Andrew Emili • Jack Greenblatt
Shoshana Wodak
Editors

Systems Analysis of Chromatin-Related Protein Complexes in Cancer

Editors
Andrew Emili
Donnelly Centre for Cellular
and Biomolecular Research
Banting and Best Department of Medical
Research
University of Toronto
Toronto, ON, Canada

Jack Greenblatt
Terrence Donnelly Centre for Cellular
and Biomolecular Research
University of Toronto
Toronto, ON, Canada

Shoshana Wodak
Molecular Structure and Function Program
The Hospital for Sick Children
Toronto, ON, Canada

ISBN 978-1-4614-7930-7 ISBN 978-1-4614-7931-4 (eBook)
DOI 10.1007/978-1-4614-7931-4
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013944686

© Springer Science+Business Media New York 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The chromatin fibres packaging our genes consist of the DNA double helix wrapped around histone proteins to form nucleosomes. Chromatin structure and function are modified by a highly specialized protein machinery that is responsible for epigenetic memory, which is defined as changes in gene accessibility and activity states that are stably inherited yet do not involve alterations in the DNA sequence of an organism. These include reversible covalent histone modifications generated enzymatically, for example site-specific methylation and demethylation marks, in a manner that alters the intrinsic stability of nucleosomes and compaction and accessibility of chromatin to the transcription apparatus. By creating a combinatorial "histone code", which is subsequently differentially recognized by effector chromatin protein complexes that physically associate with particular modified nucleosomes, chromatin can be dynamically altered to direct chromatin-based processes such as transcription and DNA repair in a locus-dependent contextual manner.

Most, if not all, chromatin modifying and remodelling enzymes are components of multi-subunit protein complexes, which often include physically associated non-coding RNAs. Controlled expression, assembly and activity of these chromatin protein complexes are essential for the proper execution of normal cellular behaviours, such as stem cell pluripotency or lineage-specific cellular differentiation, and when disrupted can cause the emergence of transformed phenotypes. Therefore, one key goal of the rapidly emerging field of epigenetics is to understand the molecular composition, biochemical regulation and physiological roles of these chromatin protein complexes, particularly in the accurate control of gene expression, DNA repair, genome stability, chromosome compaction and segregation. All of these are vital processes for normal human development, including stem cell renewal and the formation of cell types and tissues.

To elucidate the properties and biological roles of chromatin modifying complexes and the mechanisms underlying disruptions in chromosome structure and gene expression patterns that occur when chromatin protein interaction networks become perturbed in pathological states, the field of epigenetics draws on increasingly sophisticated analytical strategies. As a result, the technical expertise and

clinical experience needed to address the range and complexity of the biological problems involved can be overwhelming to the novice and expert alike, particularly as many scientists now entering the epigenetics field are trained in either traditional molecular biology, cancer biology, structural biology, drug discovery or high-throughput "omic" sciences, but are rarely familiar with all the relevant domains. This book was conceived to address this gap.

In this volume, leading international experts discuss recent progress in the application of both traditional and cutting-edge methods to explore the unique biology and complicated biochemistry of chromatin protein complexes, including the identification, functional evaluation and biomedical assessment of particular chromatin protein complexes in different epigenetic systems ranging from stem cell development to human cancer: *Bremner*, *Emili*, *Greenblatt* and *Wodak* review current knowledge of *chromatin protein networks and systems* in human; *Cagney*, *Coulombe*, *Figeys*, *Garcia*, *Moffat*, *Vermeulen* and *Washburn* describe *systematic interaction mapping efforts* aimed at documenting the networks of physical and functional interactions that occur among the components of the chromatin-related protein machinery; *Zhang* and *Mitchell* explore *global regulome studies* aimed at understanding the mechanistic aspects controlling chromatin states in normal and diseased states; *Copeland* and *Knapp* review progress in the *discovery and development of epigenetic drugs and small molecule chemical probes*; and, finally, *Stein* and *Wodak* illustrate how *disruption of chromatin protein complexes* is linked to common polygenic diseases, notably cancer, with major economic and social impact. Taken together, this collection of 16 invited chapters provides a holistic and authoritative overview into efforts to "crack" the epigenetic code, one of the most important challenges in biomedical research today.

Toronto, ON, Canada

Shoshana Wodak, Ph.D.
Jack Greenblatt, Ph.D.
Andrew Emili, Ph.D.

Contents

Networks of Histone Demethylases and Their Relevance to the Regulation of Chromatin Structure and Dynamics 1
Edyta Marcon, Sandra Smiley, Andrei L. Turinsky, and Jack Greenblatt

Histone Methyltransferase Complexes in Transcription, Development, and Cancer 33
Jonathan B. Olsen, Jack Greenblatt, and Andrew Emili

Chromatin Biology and Cancer Linked Through Protein–Protein Interactions 49
Shuye Pu, Andrei L. Turinsky, and Shoshana Wodak

Familial and Somatic Mutations of Histone-Modifying Enzymes in Cancer 65
Chenyi Liu and Lincoln Stein

ncRNA–Protein Interactions in Development and Disease from the Perspective of High-Throughput Studies 87
Dorothy Yanling Zhao, Yue Li, Jack Greenblatt, and Zhaolei Zhang

Identification of Chromatin-Binding Protein Complexes 117
Alain Doucet and Daniel Figeys

Exploring Chromatin Readers Using High-Accuracy Quantitative Mass Spectrometry-Based Proteomics 133
Arne H. Smits and Michiel Vermeulen

Proteomic Interrogation of Human Chromatin Protein States 149
Natarajan V. Bhanu and Benjamin A. Garcia

Quantitative Proteomics Characterization of Chromatin-Remodeling Complexes in Health and Disease 177
Mahadevan Lakshminarasimhan and Michael P. Washburn

Proteome Analysis of Chromatin Complexes in Differentiating Stem Cells 197
Ariane Watson and Gerard Cagney

Structural Genomics and Drug Discovery for Chromatin-Related Protein Complexes Involved in Histone Tail Recognition 211
Panagis Filippakopoulos and Stefan Knapp

Proteomic Analysis Reveals a Role for the GTPase RPAP4/GPN1 and the Cochaperone RPAP3 in Biogenesis of All Three Nuclear RNA Polymerases 227
Diane Forget, Philippe Cloutier, Céline Domecq, and Benoit Coulombe

Targeting Chromatin Modifying Enzymes in Anticancer Drug Discovery 239
Victoria M. Richon, Mikel P. Moyer, and Robert A. Copeland

Genome Organization in Cancer Cells 257
Harry Yu Zhou and Jennifer A. Mitchell

Histone Chaperones, Epigenetics, and Cancer 277
Mina Rafiei and Rod Bremner

Evaluating Chromatin Regulators in Cancer Cell Lines Through RNAi-Glasses 297
Troy Ketela, Kevin R. Brown, and Jason Moffat

Index 309

Contributors

Natarajan V. Bhanu, Ph.D. Department of Biochemistry and Biophysics, Epigenetics Program, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Rod Bremner, Ph.D. Department of Ophthalmology and Vision Sciences, Toronto Western Hospital, Toronto, ON, Canada

Kevin R. Brown, Ph.D. Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Gerard Cagney, Ph.D., B.Sc. University College Dublin, The Conway Institute, School of Biomolecular and Biomedical Science, Belfield, Dublin, Ireland

Philippe Cloutier Laboratory of Gene Transcription and Proteomics, Institut de recherches cliniques de Montréal, Montréal, QC, Canada

Robert A. Copeland, Ph.D. R&D Department, Epizyme, Inc., Cambridge, MA, USA

Benoit Coulombe Laboratory of Gene Transcription and Proteomics, Institut de recherches cliniques de Montréal, Montréal, QC, Canada

Céline Domecq Laboratory of Gene Transcription and Proteomics, Institut de recherches cliniques de Montréal, Montréal, QC, Canada

Alain Doucet, Ph.D. Department of Biochemistry Microbiology and Immunology, Ottawa Institute of Systems Biology, University of Ottawa, Ottawa, ON, Canada

Andrew Emili, Ph.D. Donnelly Centre for Cellular and Biomolecular Research, Banting and Best Department of Medical Research, University of Toronto, Toronto, ON, Canada

Daniel Figeys, Ph.D. Department of Biochemistry Microbiology and Immunology, Ottawa Institute of Systems Biology, University of Ottawa, Ottawa, ON, Canada

Panagis Filippakopoulos, Ph.D. Nuffield Department of Clinical Medicine, Medicine Division, Oxford University, Oxfordshire, Oxford, UK

Diane Forget Laboratory of Gene Transcription and Proteomics, Institut de recherches cliniques de Montréal, Montréal, QC, Canada

Benjamin A. Garcia, Ph.D. Epigenetics Program, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Jack Greenblatt, Ph.D. Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Troy Ketela, Ph.D. Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Stefan Knapp, Ph.D. Nuffield Department of Clinical Medicine, Medicine Division, Oxford University, Oxford, Oxfordshire, UK

Mahadevan Lakshminarasimhan, Ph.D. Stowers Institute for Medical Research, Kansas City, MO, USA

Yue Li, B.Sc., M.Sc., Ph.D. Candidate Department of Computer Science, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Chenyi Liu, M.Sc., Ph.D. Candidate Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

Edyta Marcon, Ph.D. Banting and Best Department of Medical Research, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Jennifer A. Mitchell, Ph.D. Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada

Jason Moffat, Ph.D. Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Banting and Best Department of Medical Research and Department of Molecular Genetics, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Mikel P. Moyer, Ph.D. R & D Department, Epizyme, Inc., Cambridge, MA, USA

Jonathan B. Olsen, B.Sc., M.Sc.A Department of Molecular Genetics, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Shuye Pu, M.D., M.Sc. Molecular Structure and Function Program, The Hospital for Sick Children, Toronto, ON, Canada

Mina Rafiei, M.Sc. Laboratory of Medicine and Pathobiology, Genetics and Development Division, University of Toronto, Toronto, ON, Canada

Victoria M. Richon, Ph.D. Oncology Discovery and Preclinical Sciences, Global Oncology Division, Sanofi, Cambridge, MA, USA

Sandra Smiley Molecular Genetics, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Arne H. Smits, M.Sc. Department of Molecular Cancer Research, University Medical Center Utrecht, Utrecht, The Netherlands

Lincoln Stein, MD, Ph.D. Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

Informatics and Bio-Computing, Ontario Institute for Cancer Research, Toronto, ON, Canada

Bioinformatics and Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

Andrei L. Turinsky, Ph.D., M.Sc. Research Institute, Molecular Structure and Function Program, Hospital for Sick Children, Toronto, ON, Canada

Michiel Vermeulen, Ph.D. Department of Molecular Cancer Research, University Medical Center Utrecht, Utrecht, The Netherlands

Michael P. Washburn, Ph.D. Stowers Institute for Medical Research, Kansas City, MO, USA

Ariane Watson, B.Sc. University College Dublin, The Conway Institute, School of Biomolecular and Biomedical Science, Belfield, Dublin, Ireland

Shoshana Wodak, Ph.D. Molecular Structure and Function Program, The Hospital for Sick Children, Toronto, ON, Canada

Dorothy Yanling Zhao, B.Sc., M.Sc., Ph.D. Candidate Department of Molecular Genetics, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Zhaolei Zhang, B.Sc., Ph.D. Banting and Best Department of Medical Research, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Harry Yu Zhou, B.Sc. Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada

Networks of Histone Demethylases and Their Relevance to the Regulation of Chromatin Structure and Dynamics

Edyta Marcon, Sandra Smiley, Andrei L. Turinsky, and Jack Greenblatt

Abstract Regulation of chromatin structure and dynamics is crucial for gene expression, chromosome segregation, DNA replication, and DNA repair, effectively controlling all cellular processes. Such regulation is achieved by a multitude of chromatin modifying enzymes that cause changes in DNA accessibility. Modifications imposed by these enzymes include DNA methylation and various histone modifications. Often, an interplay of multiple mechanisms is necessary to properly regulate chromatin dynamics, leading to a defined functional outcome. Of the various histone modifications, methylation is the most complex and is reversed by the KDM1 and JMJC families of histone demethylases. Collectively, these enzymes can reverse all three histone methylation states, often acting on the same substrates, and yet having different functional outcomes. Thus far, substrates have been identified for 26/32 (80 %) of all known histone demethylases, but functional

E. Marcon, Ph.D. (✉)
Banting and Best Department of Medical Research, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Rm 930, 160 College Street, Toronto, ON, Canada, M5S 3E1
e-mail: edyta.marcon@utoronto.ca

S. Smiley
Molecular Genetics, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Rm 930, 160 College Street, Toronto, ON, Canada, M5S 3E1
e-mail: sandra.smiley@utoronto.ca

A.L. Turinsky, Ph.D, M.Sc.
Research Institute, Molecular Structure and Function Program, Hospital for Sick Children, Toronto, ON, Canada
e-mail: turinsky@sickkids.ca

J. Greenblatt, Ph.D.
Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, 160 College St., Rm 940, Toronto, ON, Canada, M5S 3E1
e-mail: jack.greenblatt@utoronto.ca

A. Emili et al. (eds.), *Systems Analysis of Chromatin-Related Protein Complexes in Cancer*, DOI 10.1007/978-1-4614-7931-4_1,
© Springer Science+Business Media New York 2014

studies have lagged behind. Analysis of protein–protein interactions has greatly contributed to our understanding of the roles some of these proteins play in the regulation of chromatin structure and dynamics, sometimes explaining genetic associations previously established between aberrant expression of histone demethylases and certain human disorders. Here, we will discuss our current understanding of histone demethylases, emphasizing protein complexes, and their contributions to the function of histone demethylases, their connections to various human disorders and multiple types of cancer.

Keywords Histone demethylases • PHD (plant homeodomain) • Tudor • Chromo • JMJC (Jumonji C-terminal domain) • JMJN (Jumonji N-terminal domain) • Methylation • Demethylation • Histones

Chromatin Structure and Function

The regulation of chromatin structure and dynamics serves as the foundation of epigenetics. Regions of open, more accessible chromatin and closed, more condensed chromatin are often interspersed and highly dynamic, allowing the local chromatin structure to be fine tuned in response to cell type, cell cycle stage, DNA damage, cellular signals, and external stimuli.

Regulation of Chromatin Conformation

Chromatin regulation is controlled by two types of covalent modifications: DNA methylation and posttranslational modification of histones. The 5-methylcytosine mark imposed on specific CpG motifs in DNA is not inherently repressive; instead, these sites are recognized and bound by methyl-CpG-binding proteins, which recruit histone-modifying corepressor complexes to create closed chromatin and inhibit transcription.

Posttranslational histone modifications, primarily on their N-terminal tails, either directly affect interactions between histones and DNA or act as docking sites for other effector modules. For example, histone acetylation is mostly correlated with activation of gene expression, while histone methylation can be either activating or repressing, depending upon the context. To further complicate matters, combinations of marks on the same or adjacent histone tails create a "histone code" that fine tunes a precise chromatin state [1]. As well, histone variants that replace canonical histones in certain regions or situations also impact chromatin structure. Histone modifications and the exchange of histone variants are thought to be more dynamic than DNA methylation.

The human genome encodes large families of DNA methyltransferases and histone-modifying acetyltransferases, deacetylases, methylases, demethylases,

ubiquitinating and deubiquitinating enzymes, and kinases and phosphatases. Other protein families with over 200 members containing bromodomains, Tudor domains, chromodomains, methyl-binding domains, and/or PHD domains recognize particular patterns of histone modifications to recruit chromatin remodeling enzymes and influence chromatin function [2].

Studying Chromatin Dynamics

The study of chromatin modifying enzymes has progressively moved from studying individual components towards elucidating protein complexes and pathways, as it became clear that protein partners play key roles in altering enzymatic activity, localization, and, consequently, protein function. Royer et al. [3] recently compared 13 datasets of yeast protein–protein interactions obtained through various methods (e.g., yeast two-hybrid (Y2H) and affinity purifications/mass spectrometry (AP–MS)) and concluded that AP–MS is a robust and highly accurate high-throughput method for the identification of protein complexes. Genome-wide AP–MS studies have already been performed for *E. coli* and yeast but are more difficult for mammals due to the larger genome size and technical difficulties [4–6]. Therefore, mammalian AP–MS studies have focused on subsets of proteins (e.g., mitotic complexes, deubiquitinating enzymes, and nuclear receptor coregulators), but even from this limited subset it is easy to see that studies of protein complexes have significantly contributed to the functional characterization of mammalian proteins [7–9].

Over the last few years, a project has been underway in the Greenblatt lab to define protein complexes containing known and bioinformatically predicted human chromatin-related proteins to further our understanding of epigenetic mechanisms [10–12]. Our ultimate goal is to produce a database of a comprehensive interactome of chromatin-associated proteins to facilitate further investigations of the regulation of chromatin structure, dynamics, and function. In this chapter, we focus mainly on histone demethylases in the regulation of chromatin function, emphasizing protein complexes and their roles in guiding various cellular processes, as well as their implications for human health and disease.

Roles of Histone Demethylases in Chromatin Structure and Function

Histone methylation is mediated by a family of 59 histone methyltransferases described elsewhere in this book. Methylation is one of the most complex of histone marks and has been demonstrated to occur on specific arginine and lysine residues on three of the four canonical histones (H3, H4, and H2B) [13]. Arginine residues can be monomethylated or dimethylated symmetrically or asymmetrically, whereas

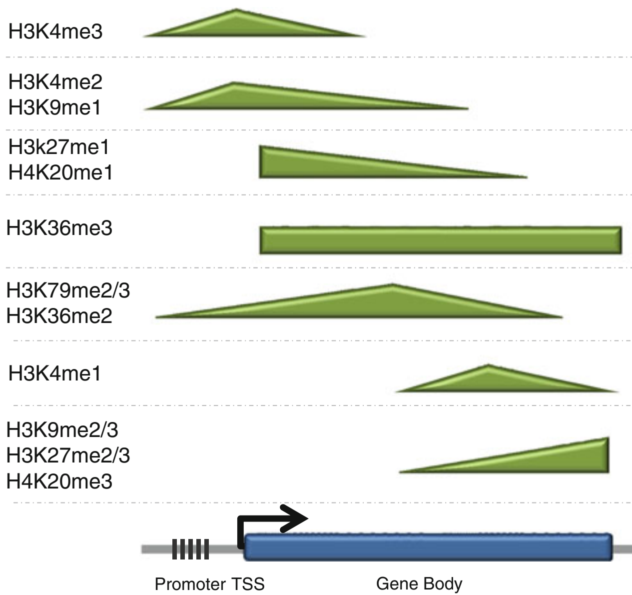

Fig. 1 Histone methylation pattern most typically found on actively transcribed genes. H3K4me3, H3K4me2, and H3K9me1 are present in the promoter region, peaking at the transcription start site (TSS). H3K4me3 falls off quickly following the TSS, while H3K4me2 and H3K9me1 persist further towards the 3′ end of the gene. H3K27me1 and H4K20me1 are absent in the upstream (promoter) region, begin at the TSS and then tail off towards the 3′ end of the gene. H3K36me3 is also absent from the promoter region but is found along the downstream region transcribed by RNA polymerase II. H3K4me1, H3K9me2/3, H3K27me2/3, and H4K20me3 are all present in the 3′ end of the gene

lysine residues can accept up to three methyl groups, resulting in mono, di, or trimethylated states. Because methylation does not affect the charge on lysine or arginine residues, it does not alter chromatin structure directly, but instead, serves as a platform for the recruitment of other protein complexes. Methylation can facilitate either transcriptional activation or repression, depending on the cellular context and the specific nature of the modification. In general, methylation of histone 3 on lysine 4 (H3K4) or lysine 36 (H3K36) is a mark of active transcription, whereas methylation of either lysine 9 or lysine 27 (H3K9, H3K27) is associated with transcriptional silencing. These marks tend to be associated with specific genomic regions such as enhancers, transcription start sites, gene bodies, or 3′ UTRs (Fig. 1). Besides transcriptional control of gene expression, histone methylation is also involved in the regulation of other processes such as DNA replication and repair. For example, methylation of lysine 79 (H3K79) less studied than other methylation marks, is thought to have a role in the establishment of chromatin boundaries and DNA repair (Fig. 2) [14, 15].

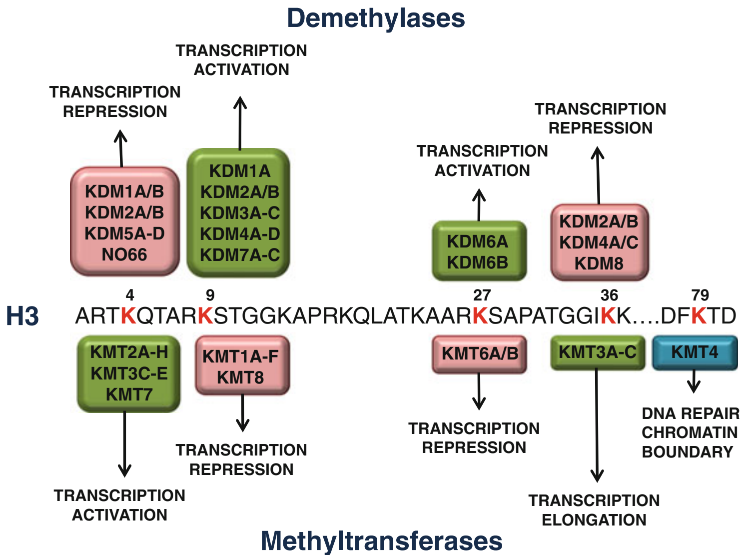

Fig. 2 Methylation of lysine residues on the histone H3 tail can produce a variety of outcomes. Shown are the major methylated residues, the enzymes required to deposit or remove those marks and their corresponding cellular functions. Some KDMs can remove more than one type of histone mark with the specificity provided by the cellular environment, and may exert either activating or repressing effects (e.g., KDM4A demethylates both H3K9 and H3K36 with opposing outcomes) depending on the context. To date, no histone demethylase has a demonstrated role in removing H3K79 methylation

There are two distinct families of histone demethylases in mammalian cells, amine oxidases and oxygenases that are largely conserved throughout the eukaryotic kingdom. The human amine oxidase family has only two members (KDM1A/LSD1 and KDM1B/LSD2) and removes methyl groups in a flavine adenine dinucleotide-dependent reaction. They can demethylate mono and dimethyl lysines, but not trimethyl lysines, on histones and certain nonhistone proteins. In contrast, the oxygenase family members are capable of removing all three histone lysine methylation states in a Fe (II)- and α-ketogluterate-dependent process. The oxygenase family is referred to as the JUMONJI family due to the presence of a C-terminal catalytic JUMONJI domain (JMJC). The 32 known JUMONJI family members in the human genome are divided into groups based on sequence homologies and structural similarities (Fig. 3). Besides the JMJC domain, they can also possess ARID, Tudor, PHD, FBOX, and zinc-finger domains, as well as tetratricopeptide repeats and N-terminal JUMONJI (JMJN) domains. The ARID domain (AT-rich interaction domain) is a DNA-binding domain, whereas both the PHD (Plant Homeodomain) and Tudor domains recognize methylated histone residues, serving as recruitment platforms for other proteins. FBOX domains and tetratricopeptide repeats serve as protein interaction scaffolds, whereas zinc finger domains are versatile and can bind

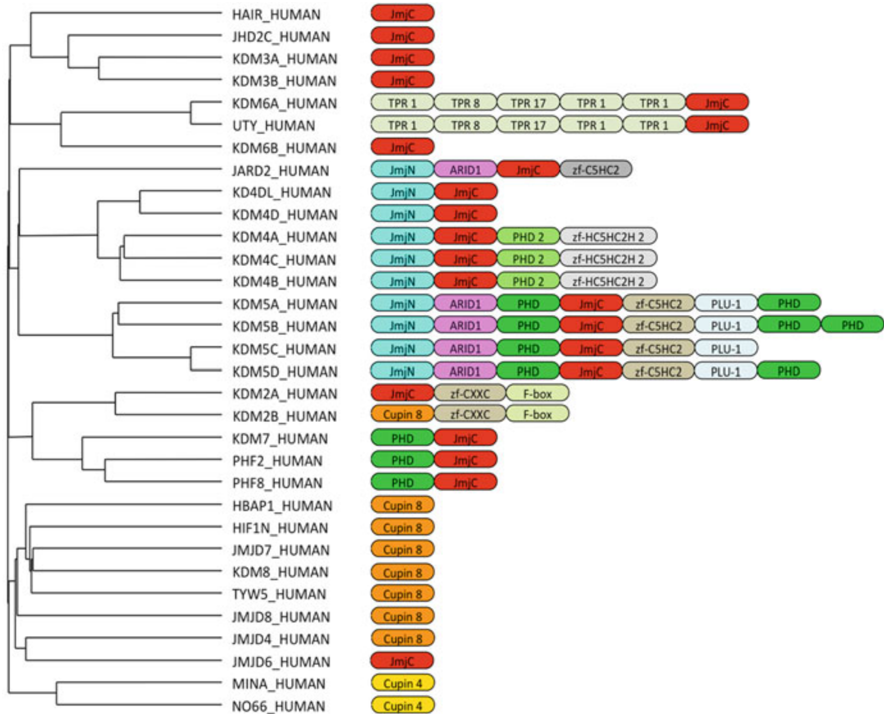

Fig. 3 Alignment of all the KDMs from the human genome. The human genome has 32 KDMs based on the presence of the JMJC/cupin domain. The main feature of these proteins is that they use Fe(II) ions and a-ketogluterate as cofactors to remove methyl groups from histones. Some of these proteins also have the ability to hydroxylate protein substrates. In addition to the JMJC/cupin domain most possess other domains involved in guiding the proteins to their substrates, providing specificity. *JMJC* Jumonji C-terminal domain, *JMJN* Jumonji N-terminal domain, *ARID1* AT-rich interaction domain, *PHD* Plant homeodomain, *TPR* tetratricopeptide repeat, *ZF* zinc finger domain, F-BOX F-box sequence motif

DNA, RNA, or other proteins. While it is clear that JMJC domains possess catalytic activity, to date none of the JMJN domains has been demonstrated to have enzymatic activity. Recent evidence suggests that the JMJN domain in GIS1, the yeast homolog of human KDM4A physically interacts with the JMJC domain to control protein stability and transcriptional regulation [16]. The presence of these additional domains modifies the specificity of this group of enzymes. While members of the same group often act on similar substrates, they often differ in their temporal and spatial expression patterns, leading to different roles in the control of chromatin structure and transcriptional output.

Most JUMONJI proteins are able to remove methyl groups from particular methylated lysines on histones, but some are specific for arginine residues, nonhistone proteins (JMJD6), or act primarily as recruitment factors [17–20]. Several JUMONJI family members have not been investigated in detail, and their demethylase activities, substrates, and functional roles have not been described. Here, we review the roles of histone demethylases in the regulation of chromatin structure and dynamics, as well as their relationships to cellular processes and human diseases, with a specific focus on the data obtained through AP–MS analysis of protein complexes (Table 1).

KDM1 Family

KDM1A and KDM1B are highly homologous, act on similar substrates, and have multiple cellular roles governed largely by the compositions of the complexes in which they reside. KDM1A can have diametrically opposite roles in that it demethylates two different lysine residues, one of which is a mark of active transcription (H3K4me1/me2), the other associated with transcriptional repression (H3K9me1/2) [21, 22]. Formation of a ligand-dependent complex between KDM1A with either androgen receptor (AR) or estrogen receptor alpha (ERα) leads to H3K9me1/2 demethylation at their respective target genes [22, 23]. This demethylation event produces hydrogen peroxide, leading to local DNA damage that, in turn, recruits DNA repair enzymes, themselves important for the transcriptional activation process at these loci. In contrast, transcriptional repression by KDM1A is mediated by its association with the corepressor complex, CoREST, with KDM1A demethylating the methylated H3K4, leading to transcriptional repression of CoREST target genes [22, 24–26]. Recently, KDM1A has also been shown to associate with another complex involved in transcriptional repression, the SWI/SNF chromatin-remodeling complex known as NURD [27]. Here, the metastasis-associated (MTA) protein plays a functional role analogous to that of CoREST, directing the demethylation of NURD target genes [27]. An added layer of regulation can be achieved through cross-talk with other histone modifications. For example, phosphorylation of histone H3 can alter the substrate specificity of KDM1A, inhibiting KDM1A-dependent demethylation of H3K4 [28].

Very recently, SFMBT1 was identified as another KDM1A interactor. SFMBT1 is a member of the MBT family of proteins. MBT domains bind mono and dimethylated lysine residues on histones and have been postulated to be components of polycomb repressive complexes (PRC) [29]. SFMBT1 was shown previously to function as a transcriptional repressor and this interaction suggests that SFMBT1 may exert its repressor activity through KDM1A [30]. It is clear from these results that KDM1A function is highly modulated by interactions with other proteins and complexes.

Table 1 Human histone demethylases, their substrates, postulated functions, and their associations with human disease

Name	Alternate name	Substrate	Function	Disease link
KDM1A	LSD1	H3K4ME1/2, H3K9ME1/2	Regulator of hematopoiesis	Breast and prostate cancer
KDM1B	LSD2	H3K4ME1/2	Genomic imprinting	–
KDM2A	JHDM1A	H3K36ME2, H3K4ME2/3	Cell proliferation, senescence, rRNA tx, genomic stability	Prostate cancer
KDM2B	JHDM1B	H3K36ME2, H3K4ME2/3	Cell proliferation and senescence	Liver and breast cancer
KDM3A	JMJD1A	H3K9ME1/2	Metabolism, spermatogenesis	Kidney, renal, bladder, lung, colon, and prostate cancer
KDM3B	JMJD1B	H3K9ME1/2	Cell proliferation	Leukemia, colorectal cancer
KDM3C	JMJD1C	H3K9ME1/2	ES cells differentiation	Gastric cancer
KDM4A	JMJD2A	H3K9ME2/3,H3K36ME2/3	Cell proliferation and replication, DNA damage	Breast cancer
KDM4B	JMJD2B	H3K9ME2/3,H3K36ME2/3	DNA damage, genomic stability	Breast and colorectal cancer
KDM4C	JMJD2C	H3K9ME2/3,H3K36ME2/3	ES cell renewal, development	Squamous cell carcinoma, prostate cancer and obesity
KDM4D	JMJD2D	H3K9ME2/3,H3K36ME2/3	Cell proliferation and survival, spermatogenesis	–
KDM4DL		–	–	–
KDM5A	JARID1A	H3K4ME2/3	Proliferation and DNA repair	Leukemia, arthritis
KDM5B	JARID1B	H3K4ME2/3	Cell cycle	Breast cancer
KDM5C	JARID1C	H3K4ME2/3	Neuronal differentiation	X-lined mental retardation

KDM5D	JARID1D	H3K4ME2/3	Development	Prostate cancer
KDM6A	UTX	H3K27ME/2/3	ES cells differentiation	Leukemia
KDM6B	JMJD3	H3K27ME/2/3	ES cells differentiation	Prostate cancer
KDM6C	UTY	–	–	Immunodeficiency
KDM7A	JHDM1D	H3K9ME1/2, H3K27ME1/2, H4K20me1	Neuronal differentiation and brain development	–
KDM7B	PHF8	H3K9ME1/2, H4K20ME1	Cell cycle progression and brain development	Prostate cancer
KDM7C	PHF2	H3K9ME1/2	Liver and rDNA transcription	Esophageal squamous cell carcinoma
KDM8	JMJD5	H3K36ME2	Cell cycle progression	–
JARID2		–	ES cells differentiation	T-cell acute leukemia
HR		–	–	Hair deficiency, loss
HSPBAP1		–	Stress response	Epilepsy
HIF1AN		ASPARAGINE HYDROXYLASE	Hypoxia response	Breast cancer, renal carcinoma
MINA	MINA53		Cell proliferation	Esophageal squamous cell carcinoma
NO66		H3K36ME2, H3K4ME3	Osteoblast differentiation	–
JMJD4		–	–	–
JMJD6	PTDSR	–	RNA splicing	–
JMJD7		–	–	–
JMJD8		–	–	–
TYW5		RNA HYDROXYLASE		–

Far less has been reported about the functional significance of the other amine oxidase family member, KDM1B. Like KDM1A, KDM1B can demethylate H3K4me1/2 and H3K9me1/2, again suggesting a dual role in both transcriptional repression and activation [31]. KDM1B has been shown to interact with RNA polymerase II elongation factors and demethylates H3K4me2 during transcription elongation. This leads to repression of target genes and plays a role in the establishment of genomic imprinting [32, 33]. In contrast, an interaction with NF-kB directs KDM1B to NF-kB target genes where it demethylates H3K9me2, thereby relieving the H3K9 silencing mark and leading to transcriptional activation [34]. Interestingly, KDM1B interacts with components of the DNA replication machinery and proteins involved in DNA repair, as well as nucleosome remodeling and histone modification complexes, indicating that KDM1B may be a member of multiple pathways [32]. KDM1B can also repress transcription through its N-terminal zinc finger domain independently of its histone demethylase activity [35].

JUMONJI Family

KDM2 Group

The KDM2 group is characterized by the presence of ARID, PHD, and FBOX domains in addition to the JMJC domain and contains two group members in mammals: KDM2A, the first JMJC family member to be identified; and its close homolog, KDM2B [36, 37]. Both KDM2A and KDM2B are able to demethylate H3K36me2 and H3K4me2/3, although the preference of KDM2B for either of these modifications is controversial. Nevertheless, both repress RNA polymerase I (RNA POLI) and RNA Polymerase II (RNA POLII) transcription and associate with the repressive polycomb group complexes (PRC) [36, 38, 39].

Despite high homology and similar substrate specificity of KDM2A and KDM2B, the two proteins have distinct cellular localizations and functions. KDM2A is found in regions of constitutive heterochromatin, including the pericentromeric satellite repeats. This localization is likely mediated through KDM2A binding to the repressive histone methylation mark, H3K9me3, and occurs preferentially in CpG islands that lack DNA methylation, an example of cross talk between histone and DNA marks [38, 40]. All three heterochromatin binding proteins (HP1) physically associate with KDM2A, and the ability of HP1 to localize to heterochromatin is reduced in KDM2A knockdowns, suggesting that KDM2A recruits these proteins [38]. In addition, KDM2A acts to repress transcription of pericentromeric satellite repeats, controlling centromeric integrity and genomic stability during mitosis [38, 41, 42]

There are contradictory reports as to the substrate for KDM2B. While some reports suggest that, similarly to KDM2A, KDM2B is mainly an H3K36me1/2 demethylase, other reports characterize it primarily as an H3K4me3 demethylase [36–38]. KDM2B has been shown to control transcriptional regulation of the

proliferation/senescence locus, p15/ink4b, through demethylation of H3K36me2 [37]. This finding provides a mechanism for a previous observation that the PRC2 polycomb complex regulates transcription of this locus, as KDM2B interacts with the EZH2 component of the PRC2 complex [43–45]. On the other hand, KDM2B has been shown to repress transcription of ribosomal genes and regulate cell morphology, chemokine expression, and apoptosis through demethylation of H3K4me3 [38, 39, 46]. Because the PHD domain of KDM2B is able to bind both H3K4me3 and H3K36me2, it is possible that KDM2B substrates are dictated mainly by the complexes associated with KDM2B during various cellular processes.

The KDM2 group is a good example of different very closely related enzymes acting on the same substrates in vitro or even in vivo, while exerting very different effects within the context of a cell. Subtle differences due to cellular localization, interacting cofactors, or temporal expression patterns are usually not accounted for by in vitro studies, which presumably is why conflicting results are obtained as to substrates and recognition of particular methylated histone residues.

KDM3 Group

The KDM3 group consists of three members, KDM3A, KDM3B and KDM3C, characterized by the presence of JMJC and FBOX domains and the ability to demethylate H3K9me1/me2. As H3K9 methylation is usually repressive, this group mediates transcriptional activation as opposed to repression.

KDM3A has a variety of roles, including transcriptional activation of metabolic, spermatogenesis-related and androgen receptor-target genes, as well as the control of reprogramming and the hypoxia response [47–50]. During spermatogenesis, KDM3A physically interacts with the activated androgen receptor (AR), and they co-localize to AR target gene promoters, where KDM3A removes the repressive H3K9me3 mark to activate transcription [47]. Similarly, in the hypoxia response, KDM3A interacts physically with the hypoxia response factor, HIF-1, which recruits KDM3A to the promoter of the glucose transporter gene, GLUT3, where it demethylates H3K9me3 and activates transcription [50]. KDM3A-mediated demethylation of the HOXA1 promoter can drive cell division, and, in ES cells, KDM3A appears to play a role in the reactivation of ES cell-specific gene expression [49, 51].

KDM3B is still poorly understood but has been shown to regulate cell proliferation through an interaction with the colorectal cancer-related metastatic protein, PRL-3 [52]. KDM3C expression is controlled by the ES cell-specific POU5F1 (OCT3/OCT4) transcription factor and, consequently, KDM3C has been implicated in the control of gene expression during ES cell differentiation and in pancreatic islets [53]. KDM3C interacts with the NSD3 histone methyltransferase complex, leading to a coordinated regulation of mouse testes development [54]. The exact mechanisms through which members of the KDM3 group exert their multiple functions have not been determined, but direct control of gene expression by KDM3A on genes regulating development has been recently reported [55].

KDM4 Group

The KDM4 group consists of five members, KDM4A/B/C/D and KDM4DL, each of which contains a JMJN domain in addition to a JMJC domain, two PHD domains, and two Tudor domains. KDM4A and KDM4C are capable of removing all three methylation states from H3K9 and H3K36, with higher activity on the trimethylated residues, while KDM4B and KDM4D are confined to H3K9me3 demethylation, and the substrates and functions for KDM4DL are unknown [56–58]. Both KDM4A and KDM4B can bind modified histones other than their substrates through their Tudor domains. While KDM4A can bind to methylated H4K20 and H3K4, KDM4B can bind only to methylated H4K20 [59–61]. Binding of KDM4A/KDM4B to H4K20 prevents P53BP1 recruitment to the same modified residue and thereby suppresses the DNA damage response. After the induction of DNA damage, KDM4A and KDM4B are degraded by the FBXO22-containing SCF E3 ubiquitin ligase, allowing P53BP1 to bind to the exposed methylated H4K20 and initiating the DNA damage response [59, 62].

Besides its role in the DNA damage response, the KDM4 group has additional cellular roles determined largely by its association partners. KDM4A acts both as a corepressor of E2F target genes through its interaction with HDAC1-3 and RB and as a coactivator of AR-target genes [63]. Overexpression of KDM4A leads to global changes in chromatin accessibility, accelerated cell cycle progression, and aberrant replication timing. These may be mediated through alterations in the localization of HP1γ, potentially explaining the linkage of KDM4 to multiple types of cancer [64].

KDM4B contributes to pericentromeric stability and chromosome segregation by demethylating H3K9me3 in pericentromeric chromatin [65]. It also directly targets the expression of cyclin-dependent kinase 6 (CDK6), which is essential for the G1/S transition, thereby contributing to the aberrant cell cycle progression phenotype of KDM4B mutants [66].

KDM4C, in concert with KDM1A, acts as a transcriptional activator for androgen receptor target genes and controls the self-renewal of ES cells by regulating the transcription of pluripotency-specific transcription factors [53, 67]. In mouse development it may regulate the expression of the proliferation-related transcription factors, MYC and KLF4 [68]. KDM4D has a number of disparate roles. It binds directly to p53 and activates p53 target genes, presumably through the removal of repressive H3K9 methylation marks on the promoter regions of these genes. The same study also found that KDM4D can act in an opposing manner, in a p53-independent pathway, to stimulate cell proliferation and survival, illustrating the complex balance of in vivo functions [69]. KDM4D also appears to be responsible for regulation of spermatogenesis via the activation of androgen-responsive genes and for demethylation of repressive H3K9 methylation marks surrounding the enhancers of tissue-specific genes; although this enhances expression, it is not sufficient for gene activation [70–72].

KDM5 Group

The KDM5 group members each contain five conserved domains, including both JMJN and JMJC domains, as well as ARID, two PHD, and Zinc finger domains. This family has four members (KDM5A through D) with KDM5A and KDM5B being located on the autosomes, whereas KDM5C and KDM5D are located on the X and Y chromosomes, respectively. All KDM5 family members have been shown to specifically recognize and demethylate H3K4me2/3 during the cell cycle and differentiation and, as such, are generally involved in transcriptional repression. KDM5A interacts with the SIN3B complex and acts in an RB-dependent manner to silence RB target genes enabling senescence. It can also repress genes regulated by the Notch pathway [73–75]. Moreover, KDM5A is closely connected with the PRC2 complex. The PRC2–KDM5A complex imposes transcriptional silencing on target genes by coordinated trimethylation of H3K27me3 and demethylation of H3K4me3 [76]. Recently, KDM5A has also been implicated in the DNA damage response where, upon ionizing radiation-induced double-strand break formation, KDM5A accumulates at the sites of DNA damage. Other silencing marks, such as H3K27me3, are incorporated during the DNA damage response, and so it is plausible that KDM5A involvement in DNA repair is mediated through a polycomb-related mechanism [77].

KDM5B is a transcriptional repressor with roles in neural differentiation, senescence, and cellular proliferation [73, 78–80]. Through an association with the transcription factors MYC and TFAP2C, KDM5B acts to downregulate the cell cycle gene CDKN1A, thereby promoting cell cycle progression [79]. Similarly, in senescence, KDM5B associates with RB, promoting repression of RB-dependent cell cycle genes [81]. In contrast with the described functions of the other KDM5 family members, KDM5B is able to function as a transcriptional activator during self-renewal. This seemingly contradictory finding stems from the observation that KDM5B is recruited by the histone acetyltransferase complex member MORF4L1 to the bodies of actively transcribed genes where demethylation of H3K4 inhibits cryptic initiation of transcription, a process that interferes with efficient RNA polymerase elongation [82].

KDM5C has a role in transcriptional repression of a subset of neuronal genes through its association with REST, a transcription factor required for silencing of neuronal genes in non-neuronal tissues, histone acetyltransferases, HDAC1-2, and the histone methyltransferase EHMT2 [83]. KDM5C also interacts with PCNA through a PCNA interaction motif in its sequence, and this interaction is necessary for KDM5C association with chromatin [84]. The polycomb-like protein RING6A associates with KDM5D and promotes the demethylase activity of KDM5D in vitro. In vivo, KDM5D appears to target RING6A to developmentally controlled genes where they act in concert to repress transcription [85]. KDM5D might also play a role in spermatogenesis, more specifically in the regulation of meiosis, as it forms a complex with MSH5 and appears to target MSH5 to condensed chromatin during meiotic prophase [86]. Therefore, it is obvious that, even though KDM5 enzymes all have identical substrates, they perform very different roles within

cellular networks and their roles are defined by both their expression patterns and the complexes with which they are associated.

KDM6 Group

The KDM6 group (KDM6A-C) is characterized by the presence of several tetratricopeptide repeats, in addition to the JMJC-domain, KDM6A and KDM6B are both able to demethylate H3K27me2 and me3 but, to date, no demethylase activity has been detected for KDM6C [87–89]. Because H3K27me2/3 are repressive marks established mostly at gene promoters and in coding regions, KDM6A and KDM6B are transcriptional coactivators involved in multiple processes, including cell cycle progression, differentiation, development, and the inflammatory response. KDM6A co-purifies with the H3K4 methyltransferase complexes that contain the mixed-lineage leukemia proteins, MLL1-3, resulting in H3K27me3 demethylation along with H3K4me3 methylation and leading to the activation of gene expression (e.g., the HOX gene cluster) [90–92]. Such coordinated activation of gene expression by complexes exhibiting different but reinforcing activities, MLL-KDM6A and PRC2-KDM5A complexes, illustrates how an assembly of protein complexes with functionally synergistic activities can be an efficient way to control chromatin structure and dynamics.

Similarly to KDM6A, KDM6B is an important activator of HOX genes and bivalent promoters (promoters in ES cells that are marked by both activating and repressive marks) [85, 87, 93, 94]. However, in spite of extensive homology with KDM6A, KDM6B does not appear to associate with MLL; instead, it associates with KDM7A, an H3K9me1/2, H3K27me1/2, and H4K20me1 histone demethylase, along with proteins regulating transcription elongation [95]. KDM6B also has another role in that it cooperates with KDM4B to control the differentiation potential of human bone marrow mesenchymal stem/stromal cells, again illustrating collaborative efforts within the KDM family [96].

As mentioned previously, KDM6C has no detectable demethylase activity, and so it is unclear what role, if any, it does play. However, protein–protein interaction data from the Greenblatt lab indicates that, similarly to KDM6A, KDM6C co-purifies with components of the MLL complexes (unpublished data). It is possible, therefore, that it does have demethylation activity that requires the presence of previously uncharacterized cofactors, or rather functions as a recruitment/stabilization factor rather than a bona fide histone demethylase.

KDM7 Group

The KDM7 group includes three members (KDM7A-C) and is characterized by the presence of PHD and JMJC domains. KDM7A is an H3K9me1/2 and H3K27me1/2 demethylase, while KDM7B demethylates both H3K9me1/2 and H4K20me1, and

KDM7C seems to use exclusively H3K9me1/2 as a substrate [36, 97–99]. KDM7A is a positive transcriptional regulator that localizes to the nucleolus, where it is involved in the transcription of rDNA [97, 100, 101]. It is also important for activation of neuronal-specific genes and brain development and directly controls the transcription of FGF4, an oncogenic growth factor [102, 103].

KDM7B, in addition to its ability to demethylate H3K9me1/2, was the first demethylase found to act on H4K20me1, a modification important for cell cycle progression, neural differentiation, and brain development [98, 99]. KDM7B interacts directly with the CTD of RNA polymerase II and serves as a general coactivator, present at many active genes [104]. In addition, KDM7B acts in concert with activating factors E2F1, HCFC1, and SETD1A to demethylate H4K20me1 onE2F1-target genes [99].

KDM7C is a transcriptional activator of HNFA in liver and rDNA genes. Interestingly, unmodified KDM7C is inactive and requires PKA-dependent phosphorylation to bind, demethylate, and form a complex with ARID5B. This activated KDM7C/ARID5B complex is then recruited to promoters, where it demethylates H3K9me2, leading to transcriptional activation [105]. All members of KDM7 family are able to bind H3K4me3 marks but do not use them as a substrate. Instead, binding to the trimethylated lysine 4 on histone H3 increases the enzymatic activity of KDM7 family members towards H3K9me2, illustrating the influence of other domains within the same protein on the catalytic activity of histone demethylases [99, 106].

JARID2

JARID2 possesses both ARID and zinc finger domains in addition to JMJN and JMJC domains. However, it does not exhibit any detectable demethylase activity, acting instead as a recruitment factor for other chromatin-modifying complexes. JARID2 is a component of polycomb repressive complex 2 (PRC2), where it is important for targeting PRC2 to its target genes and modulating its activity [18, 107]. JARID2 also interacts with the histone methyltransferase SETDB1 to control the levels of H3K9me2/3 at the NOTCH locus, thereby regulating NOTCH expression [108]. Through its recruiting ability, JARID2 has been implicated in transcriptional regulation and ES cell differentiation.

JMJC-Only Demethylases

There are other JUMONJI family members which possess a JMJC domain but contain no other recognizable domains and thus are not grouped into specific families (Fig. 3). JMJC-only members are, by far, more obscure and less studied, with few exceptions. KDM8, an H3K36me2 demethylase, is known to participate in cell cycle progression and circadian systems [109–111]. NO66 is a histone demethylase with specificity for H3K4me3 and H3K36me2. Together with SP7/OSX, it regulates osteoblast differentiation by demethylating H3K4 and H3K36 and inhibiting

SP7/OSX-mediated promoter activation. It may also function in replication and remodeling of heterochromatic regions [112, 113]. JMJD6 is the only JMJC-containing protein which has been demonstrated to be capable of demethylating arginine residues on histones H3 and H4 in vitro, but it is not clear whether this is its primary in vivo function as it also demethylates nonhistone substrates [17]. JMJD6 is required for organogenesis, hematopoietic differentiation, and regulation of cytokine responses [114–116]. For other JMJC-only group members, such as HSPBAP1, HIF1AN, HR, MINA53, JMJD4, JMJD7, JMJD8, and TYW5, no demethylase activities, histone substrates, or binding sites have been identified. Several of them have known cellular roles, although the mechanisms of their actions remain unclear. HSPBAP1 might have a role in the cellular stress response [72]. HIF1AN plays a role in the response to hypoxic conditions by hydroxylation of asparagines residues within HIF, suppressing its transcriptional activity [19, 117]. HR could act as transcription factor regulating cell growth, possibly through the hyperactivation of WNT signaling pathways [54, 118]. MINA53 was identified in a screen searching for MYC targets, is present in the nucleolus and, as is the case for MYC, is involved in cell proliferation [119]. TYW5 is a tRNA hydroxylase that acts in the biosynthesis of a hypermodified nucleoside, hydroxywybutosine, which is essential for correct phenylalanine codon translation [20]. JMJD7-PLA2G4B is a read-through protein encoding a calcium-dependent phospholipase, while nothing is known about JMJD4 or JMJD8 aside from the fact that they contain C-terminal JMJC domains [120].

Demethylation of Nonhistone Substrates

Even though most histone demethylases work specifically on histone substrates, some are able to remove methyl marks from nonhistone proteins, thereby modifying their activities. The tumor suppressor protein p53 is one of the most extensively studied methylated proteins, as it can be methylated on several residues to various extents. The degree of methylation and the methylated residue determine p53 activity and its roles during cell cycle progression, DNA repair, and apoptosis. Several histone methyltransferases, including KMT3C, KMT5A, and KMT7, methylate p53 on different residues and in response to different environmental stimuli. However, KDM1A is the only demethylase that has, to date, been shown to demethylate p53, preventing p53 interaction with P53BP1, and thereby inducing apoptosis [27, 121, 122].

KDM1A can also relieve methylation on E2F1 and DNMT1. E2F1 is a transcription factor with a role in cell cycle progression and apoptosis, and its target genes include several pro-apoptotic factors in DNA damage-induced apoptosis. Methylation of E2F1 by KTM7 results in transcriptional activation, while demethylation by KDM1A destabilizes E2F1, inhibiting DNA damage-induced cell death, and promoting DNA repair [123]. DNA methyltransferase 1 (DNMT1) is also

methylated by KMT7, with effects that are opposite to that on E2F1. In this case, demethylation by KDM1A, rather than methylation by KMT7, stabilizes DNMT1, ultimately regulating DNA methylation [124]. This connection between KDM1A and DNMT1 then provides an important link between DNA methylation and histone modification in the regulation of chromatin dynamics. Furthermore, it seems KDM1A and KMT7 have opposing roles in the regulation of the methylation status of nonhistone substrates.

Only a few members of the JUMONJI family of lysine demethylases have so far been shown to demethylate nonhistone substrates. KDM2A demethylates the p65 subunit of the NF-κB, a regulator of immune and inflammatory responses, reversing the mark imposed by KMT3B, and leading to the inhibition of NF-κB signaling [125]. Interestingly, KDM2A expression is itself driven by NF-κB signaling in a feedback loop. KDM4A-C can demethylate trimethylated lysine peptides in several nonhistone proteins in vitro, including WIZ, CDYL1, CBS, and EHMT1 [126]. All of these proteins are found in chromatin-related complexes regulating transcription, illustrating yet another example of a cross talk within chromatin modification machinery.

KDM7C demethylation of the nonhistone protein, ARID5B, is necessary for its function as a histone demethylase cofactor. KDM7C is activated upon phosphorylation by a protein kinase (PKA) and can then demethylate ARID5B. Only upon ARID5B demethylation can the ARID5B–KDM7C complex be targeted to H3K9ME2 at promoters of KDM7C target genes [105]. Even though JMJD6 is able to demethylate arginine residues on histones in vitro, histone demethylation does not seem to be a primary function of JMJD6 in vivo. Instead, it functions as a lysyl hydroxylase for the splicing factor U2AF65, thereby regulating its pre-mRNA splicing activity [127]. There are several other putative histone demethylases for which no substrates or possible functions have been identified but that can potentially act on nonhistone substrates (Table 1).

The Role of Chromatin in Human Disease

The traditional carcinogenesis model proceeded from initiation (exposure to a carcinogen leading to a change in nucleotide sequence) through promotion (additional exposure to a carcinogen leading to enhanced cell division) to progression (malignant tumor formation). However, there are many carcinogens that do not lead to DNA sequence changes but, instead, affect chromatin structure. This altered chromatin structure can result in changes in gene expression or chromosome instability, leading to an imbalance among apoptosis, proliferation, and differentiation. Even in the absence of specific carcinogens, the same effects can be induced by mutations or defects in many of the chromatin-modifying enzymes, including DNA repair proteins.

The best studied examples of cancer susceptibility genes are BRCA1/2, and mutations in these DNA repair genes are responsible for 2–10 % of all breast cancers and 5–10 % of all ovarian cancers worldwide. The inheritance of BRCA1/2 mutations increases the risk of breast cancer by 50–80 %, the risk of contralateral breast cancer by 60 %, and the risk of ovarian cancer by 15–25 % (reviewed in [128]). The incomplete penetrance of the cancer phenotype illustrates that not all individuals with a particular genetic mutation will develop breast cancer and, conversely, patients with the same tumor type may have very divergent genetic or epigenetic changes. Presumably, the specific combination of genetic mutations, genetic interactions between gene products, differences in epigenetic programming, and exposure to environmental factors can all influence cancer predisposition and manifestation.

So far, mutations in a number of different chromatin-related proteins have been linked to serious developmental disorders. Alpha-thalassemia/mental retardation syndrome results from loss of function mutations in the chromatin-remodeling enzyme, ATRX, and may be due to changes in DNA methylation patterns [129, 130]. ICF syndrome, a rare autosomal recessive disorder characterized by immunodeficiency, instability of pericentromeric heterochromatin, mental retardation, and developmental defects, is linked to mutations within DNMT3B, a DNA methyltransferase [131]. Similarly, MECP2 contains a methyl-CpG recognition domain, and loss of function mutations in this protein cause Rett syndrome, the most frequent cause of mental retardation in females [132]. Rubinstein–Taybi Syndrome (RSTS) is characterized by congenital malformation and mental retardation stemming largely from the mutations in the histone acetylase CBP [133]. EGF-stimulated phosphorylation of histone H3 on serine 10 by the serine–threonine PKA is a leading cause of Coffin–Lowry Syndrome [134]. In addition to these more prevalent and better studied disorders, many more have been linked to the aberrant expression of chromatin modifying enzymes, especially many types of cancer.

The link between chromatin modifications and human disease is strong, and listing all of the currently known connections here would be impossible. Enrichment profiles of all known or predicted chromatin modification enzymes with links to human disease demonstrate the wide-ranging effects chromatin can have on many aspects of gene regulation and chromosome stability (Fig. 4a). Here, neoplasms or abnormal tissue masses show highest enrichment, but there are other highly enriched categories. Figure 4b shows hierarchical view of the most enriched subsets and illustrates how broad categories can be partitioned into smaller categories or single disorders.

From this, it is obvious that often members of the same complex are implicated in the same disease allowing prediction of disease-related genes and consequently predictions of novel therapeutic targets. Thus, identification of histone demethylases associated with disease and their potential partners can be advantageous in the identification of novel disease genes and therapeutic targets.