JmjC-domain-containing proteins and histone demethylation

Robert J. Klose, Eric M. Kallin* and Yi Zhang*

Abstract | Histone methylation has important roles in regulating gene expression and forms part of the epigenetic memory system that regulates cell fate and identity. Enzymes that directly remove methyl marks from histones have recently been identified, revealing a new level of plasticity within this epigenetic modification system. Here we analyse the evolutionary relationship between Jumonji C (JmjC)-domain-containing proteins and discuss their cellular functions in relation to their potential enzymatic activities.

Howard Hughes Medical Institute, Department of Biochemistry and Biophysics, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina 27599-7295, USA.
**These authors contributed equally to this work.*
Correspondence to Y.Z.
e-mail:
yi_zhang@med.unc.edu
doi:10.1038/nrg1945

The basic unit of chromatin consists of 146 bp of DNA wrapped around a histone octamer, which is composed of two copies of each of the four core histones: H2A, H2B, H3 and H4. The covalent modification of core histones modulates genome function by contributing additional epigenetic information. One such modification is methylation, which occurs on arginine and lysine residues, and is involved in regulating a wide range of processes including gene activity, chromatin structure, dosage compensation and epigenetic memory (reviewed in REF. 1).

The identification and characterization of histone methyltransferases have provided a wealth of knowledge regarding the effects of distinct histone methylation marks on chromatin function. In general, lysine (K) methylation at H3K9, H3K27 and H4K20 is associated with regions of transcriptionally silenced chromatin, whereas methylation at H3K4, H3K36 and H3K79 is associated with transcriptionally active regions[1]. As a more detailed understanding of histone methylation emerges, the generality of these correlations becomes less clear, because H3K9 methylation has been found in transcriptionally active genes[2] and H3K36 methylation actually represses intragenic transcription initiation[3]. The functional consequences of the histone methylation system seem to rely on proteins that recognize defined methylation marks and elicit functional effects on the surrounding chromatin. This is exemplified by HP1 (heterochromatin protein 1), which recognizes H3K9 methylation and confers transcriptional silencing to surrounding chromatin[4,5]. Furthermore, the methylation state (mono-, di- or trimethylation) of a modified lysine residue can, in some instances, dictate which effector proteins efficiently recognize methylated chromatin. For example, the BPTF (bromodomain PHD-finger transcription factor) protein component of the NURF (nucleosome-remodelling factor) chromatin-remodelling complex uses its PHD (plant homeodomain) to specifically target the NURF complex to regions containing histone H3K4 modified in the trimethyl state[6]. The specificity of this NURF targeting module is important for the regulation of Hox gene expression during development, indicating that the state of histone methylation, in addition to the site of lysine modification, is important in determining the functional outcome of this epigenetic modification.

Until recently, it was unclear whether enzymes capable of antagonizing histone methylation existed. Three distinct classes of such enzymes have now been characterized (FIG. 1). PADI4 (Petidylarginine deiminase 4) was the first to be identified; it functions as a histone deiminase that converts methyl-arginine to citrulline as opposed to directly reversing arginine methylation[7,8]. Although PADI4 has a clear role in antagonizing methyl-arginine modifications, it cannot strictly be considered a histone demethylase as it produces citrulline instead of an unmodified arginine. LSD1 (Lysine specific demethylase 1) was the founding member of a second class of enzymes that directly reverse histone H3K4 or H3K9 modifications by an oxidative demethylation reaction in which flavin is a cofactor[9,10]. Full enzymatic activity of LSD1 requires its association with other proteins, such as the CoREST (restin corepressor) complex, indicating that regulatory subunits can have a role in modulating demethylase activity[11,12]. The third and largest class of demethylase enzymes contain a Jumonji C (JmjC) domain and catalyse lysine demethylation of histones through an oxidative reaction that requires iron Fe(II) and α-ketoglutarate (αKG) as cofactors[13]. Unlike LSD1, which can only remove mono- and dimethyl lysine modifications, the JmjC-domain-containing histone demethylases (JHDMs) can remove all three histone lysine-methylation states. So far, JHDMs have been shown to reverse H3K36 (JHDM1) (REF13), H3K9

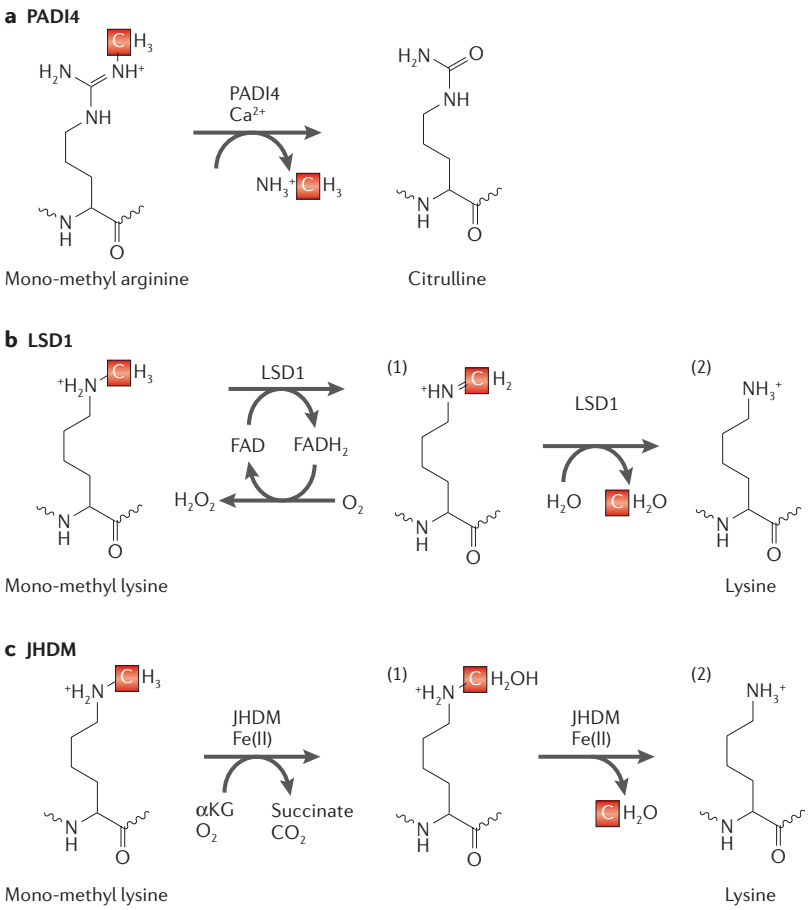


Figure 1 | **Chemical mechanism by which three distinct classes of enzymes antagonize histone methylation. a** | PADI4 (petidylarginine deiminase 4) is a Ca^{2+} dependent deiminase that antagonizes arginine methylation by demethylimination. Demethylimination of mono-methyl arginine is shown, with the methyl group shown in red. **b** | LSD1 (lysine specific demethylase 1) is proposed to mediate demethylation of mono and dimethylated lysine residues through an amine oxidation reaction that uses flavin (FAD) as a cofactor. Loss of the methyl group from monomethyl lysine occurs through an imine intermediate (1), which is hydrolysed to form formaldehyde by a non-enzymatic process (2). **c** | JHDM histone demethylases can demethylate mono-, di- and trimethylated lysine by an oxidative mechanism that requires Fe(II) and αKG as cofactors. Demethylation is thought to occur by direct hydroxylation of the methyl group (1), which results in an unstable hydroxymethyl product that is spontaneously released as formaldehyde (2).

(JHDM2A) (REF. 14) and both H3K9 and H3K36 (JHDM3 and JMJD2A–D) methylation[15–18]. Given that histone methylation has important roles in development and contributes to the cellular abnormalities observed in cancer, it seems likely that histone demethylases regulate a diverse array of cellular processes. Consistent with this notion, JHDM2A associates with the androgen receptor and is important for H3K9 demethylation during ligand-dependent activation of androgen-responsive genes[14]. The JHDM3 and JMJD2 H3K9 and H3K36 demethylases regulate gene expression[16], antagonize HP1 recruitment to chromatin[16–18] and are required for the proliferative capacity of specific cancer cell lines[17]. Further characterization of JHDMs should reveal more about their function and how they regulate the recognition of chromatin by effector proteins.

Cupin
A domain found within a superfamily of proteins that are characterized by their β-barrel tertiary structure (in Latin, *cupa* means small barrel). Proteins containing this domain include metal coordinating oxygenases and other proteins that lack enzymatic activity.

Here we present an analysis of JmjC-domain-containing proteins from yeast, worms, flies, mice and humans. By combining a bioinformatic analysis and a comprehensive literature survey, we provide new insights into JmjC-domain homology, protein domain architecture and enzymatic activity. Based on our analysis, we predict that many more JmjC-domain-containing proteins are likely to function as protein hydroxylases or histone demethylases. Furthermore, based on the available information about individual JmjC proteins, we speculate about the functional relationship to potential enzymatic activity within the context of individual bioinformatic groupings.

Function and structure of the JmjC domain

The JmjC domain was first defined based on the amino-acid similarities in the Jarid2 (Jumonji), Jarid1C (Smcx), and Jarid1A (RBP2) proteins[19–21]. Homology between the JmjC and cupin metalloenzyme domains[19,22] led to the identification of the JmjC-domain-containing factor inhibiting hypoxia (FIH) as an active protein oxygenase that can hydroxylate asparagine residues[23]. Given that a related group of oxygenases, the AlkB family of proteins, were previously shown to remove methylation from amine groups in modified DNA, it was predicted that chromatin-associated JmjC-domain-containing proteins might be involved in demethylation of modified arginine or lysine amine groups within histones[13,24]. This was confirmed by an unbiased activity-based biochemical purification, which identified a JmjC-domain-containing protein, JHDM1, as a H3K36-specific demethylase[13]. Based on the crystal structure of FIH and the recently solved structure of the catalytic domains of JHDM3A/JMJD2A, the JmjC domain has been shown to fold into eight β-sheets, thereby forming an enzymatically active pocket that coordinates Fe(II) and αKG[25–28] (FIG. 2a). Three amino-acid residues within the JmjC domain bind to the Fe(II) cofactor and two additional residues bind to αKG (FIG. 2b). Many JmjC-domain-containing proteins have conserved residues within the predicted cofactor-binding sites, indicating that these proteins could be active enzymes; others have amino-acid variations that are likely to abrogate enzymatic activity (FIG. 2b). To catalyse histone demethylation, the cofactor-bound JmjC domain is thought to produce a highly reactive oxoferryl species that hydroxylates the methylated substrate, allowing spontaneous loss of the methyl group as formaldehyde[29] (FIG. 1). Substrate specificity for JHDMs seems to rely on both the JmjC domain and additional domains found within each enzyme.

Phylogenetic analysis and domain alignment

Given the emerging importance of JmjC-domain-containing proteins in chromatin biology, we have explored this extensive family in the context of JmjC-domain evolution and domain architecture. Our aim is to provide a resource for further exploration of JmjC-domain function, and to aid in the identification of enzymatically active proteins. Through analysis of public protein-domain databases, we have extracted

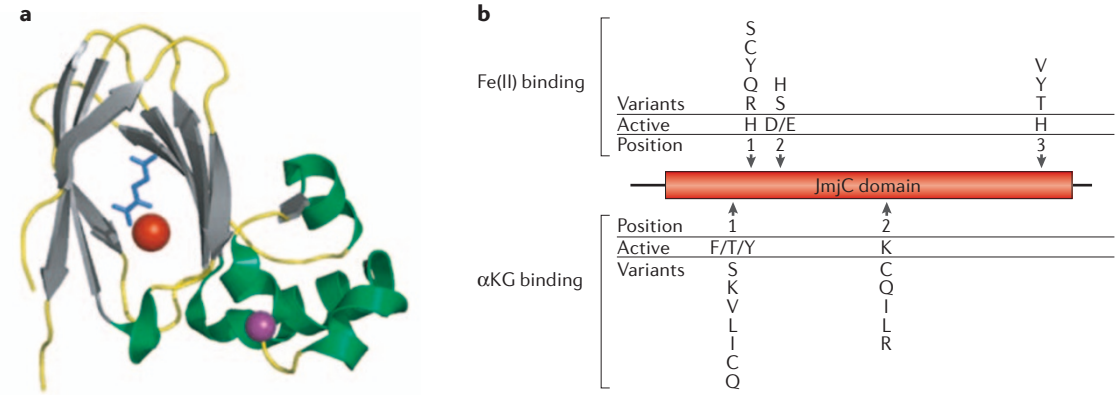


Figure 2 | **The JmjC domain contains residues required for Fe(II) and αKG binding. a** | A three-dimensional cartoon depicting the polypeptide backbone structure of the JmjC (Jumonji C) domain of JHMD3A/JMJD2A. The eight β-sheets of the cofactor-coordinating pocket are shown in grey, with the Fe(II) ion in red and αKG in blue. The α-helical region that associates with the zinc ion is shown in green and the zinc molecule in purple. **b** | A schematic representation of the JmjC domain showing the position of the Fe(II)-binding (top) and αKG-binding (bottom) residues. The amino-acid identity of residues within active hydroxylases and demethylases and amino-acid substitutions found in other JmjC-domain proteins are shown above and below the indicated cofactor-binding residue position.

a non-redundant set of 98 JmjC-domain-containing proteins from human (30), mouse (30), *Drosophila melanogaster* (13), *Caenorhabditis elegans* (13), fission yeast (7) and budding yeast (5) (BOX 1) (see Supplementary information S1 (table) for the complete list of proteins). A combination of multiple sequence alignment[30] and Bayesian inference phylogeny[31,32] was used to visualize the evolutionary relationship between the JmjC domains of the proteins within our database (FIG. 3). By merging information on the domain architecture of the full-length protein with JmjC-domain-based phylogeny, we have defined seven groups of evolutionarily conserved proteins, six of which have at least one additional protein domain besides the JmjC domain, and a seventh group that contains only a JmjC domain. To further analyse the JmjC domain within individual groups, we focused on amino-acid conservation within the predicted Fe(II) and αKG-binding sites to provide further insight into potential enzymatic activity (FIGS 4–9). In the following sections we discuss each of the JmjC-domain protein groups, commenting on potential enzymatic activity and, in some cases, on biological function, to provide a resource for the future investigation of JmjC-domain function.

The JHDM1 family

Members of the JHDM1 group are found in organisms from budding yeast to human. In addition to the JmjC domain, the human, mouse and *D. melanogaster* JHDM1 orthologues contain leucine-rich repeats (LRRs), an F-box domain and a CXXC zinc-finger domain (FIG. 4a). Using a novel activity-based histone demethylase assay, we recently identified the JHDM1 family of histone demethylases, and showed that the JmjC domain can specifically mediate Fe(II) and αKG-dependent histone demethylation[13]. Both human JHDM1 homologues (JHDM1A and JHDM1B) and their orthologue from budding yeast are H3K36 histone demethylases[13]. Little is known about the biological function of JHDM1 proteins, although the *C. elegans* orthologue can suppress spontaneous mutations[33] and the fission yeast orthologue functions to limit heterochromatic domains at the mating type locus[34].

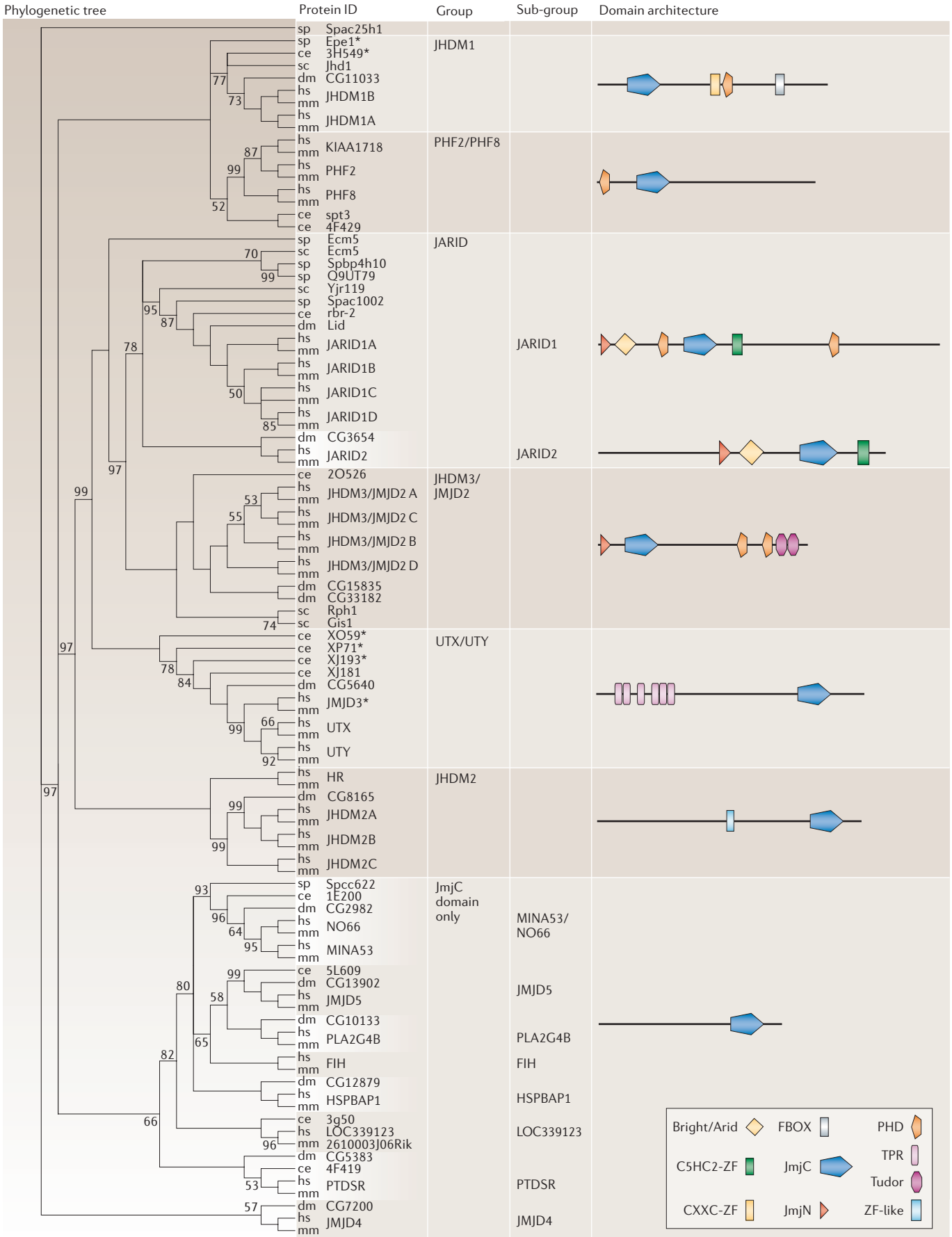
In addition to the JmjC domain, the JHDM1 group also contains two previously characterized functional domains, the F-box and CXXC zinc-finger domains. F-box proteins are known to associate with S-phase kinase-associated protein 1A (SKP1) to form the SKP1–cullin–F-box protein E3 ubiquitin ligase complex[35],

Oxoferryl
High-valent oxoiron (IV) intermediates that act as the oxidizing species in JmjC-catalysed demethylase reactions.

Bayesian inference of phylogeny
A statistical method based on a quantity called the posterior probability distribution, which is the probability of a phylogenetic tree conditioned on the observations (multiple sequence alignments). Phylogeny is calculated by determining the probability that the model (tree) is correct given the observed data (clustal alignment).

Box 1 | **Methods**

The SMART[97,98] and the PFAM domain databases[99] were used to obtain a list of all JmjC (Jumonji C)-domain-containing proteins from human, mouse, *Drosophila melanogaster*, *Caenorhabditis elegans*, fission yeast and budding yeast. This list was then curated by extracting the chromosome location of all entries. Any JmjC-domain-containing proteins that were linked to the same physical chromosome location were aligned using Multalin[100] to eliminate any redundancy in our curated JmjC-domain-containing data set. Within our curated JmjC-domain-containing protein database we have included 30 human proteins, 30 mouse proteins, 13 *D. melanogaster* proteins, 13 *C. elegans* proteins, 7 fission yeast proteins and 5 budding yeast proteins. A multiple sequence alignment of the JmjC domain was carried out using ClustalW[30,100] followed by phylogenetic-tree generation using a mixed amino-acid model and 2.2 million generations with sampling every thousandth generation using MrBayes[31,32]. The resulting tree was imported into Mega 3.1[101] for visualization and analysis. Based on the phylogenetic tree and the domain architecture of the full-length proteins we have classified the JmjC-domain-containing proteins into seven groups.



◀ Figure 3 | **Phylogenetic relationship of JmjC-domain-containing proteins from model organisms.** The phylogenetic relationship between the JmjC (Jumonji C)-domain-containing proteins from model organisms (Examples from *Schizosaccharomyces pombe* (sp); *Caenorhabditis elegans* (ce); *Saccharomyces cerevisiae* (sc); *Drosophila melanogaster* (dm); *Homo sapiens* (hs); *Mus musculus* (mm)) was determined by multiple sequence alignment and Bayesian inference analysis (BOX 1). For phylogenetic analysis, all posterior probabilities of clade partitions that were <100% are shown; the values represent the percentage of sampled trees used in the analysis that contained the consensus partition (left). By combining information from the phylogenetic analysis (left) and the domain architecture of the full-length protein (right), seven evolutionarily conserved groups of JmjC-domain-containing proteins were defined. The asterisk indicates proteins that were placed in a given group based on homology within the JmjC domain, but that lack some aspects of the domain architecture found in their related orthologues. In the JARID group, we have created two subgroups to emphasize the high level of homology between the JARID1 and JARID2 proteins, despite divergence in the spatial arrangement of the conserved domains. The JmjC-domain-only group has been divided into eight subgroups based on similarity within the JmjC domain, as determined by phylogenetic analysis alone. JmjN, Jumonji N domain; PHD, plant homeobox domain; TPR, tetracopeptide repeat domain; ZF-like, zinc-finger-like domain.

which suggests that JHDM1 might link histone demethylation to protein ubiquitylation. The CXXC zinc-finger domain is found in proteins that are involved in epigenetic regulation, and functions as a DNA-binding domain that recognizes non-methylated CpG DNA. This raises the possibility that the targeting of JHDM1-mediated histone demethylation is linked to DNA methylation status[36–38].

Proteins in the JHDM1 group have highly conserved residues within their predicted Fe(II) and αKG-binding sites (FIG. 4b). Mutation of these residues directly affects enzymatic activity both *in vitro* and *in vivo*[13]. This is exemplified in the fission yeast JHDM1 orthologue, Epe1, which has a naturally occurring histidine-to-tyrosine substitution in the third Fe(II)-binding residue that renders the enzyme inactive towards histone substrates[13]. Note that Epe1 still negatively regulates heterochromatin formation by interacting with Swi6 (the fission yeast HP1 orthologue) to counteract silencing. Surprisingly, this function relies on the JmjC domain, indicating that there is either an additional role for the JmjC domain of Epe1 that is independent of enzymatic activity[39], or an alternative enzymatic activity towards other non-histone substrates might exist. The other members of the JHDM1 group contain highly conserved cofactor binding residues, which strongly indicates that they function as H3K36 demethylases (FIG. 4).

The PHF2/PHF8 family

The PHF2/PHF8 group members are found in worms, mice and humans, and contain a PHD domain in addition to a JmjC domain (FIG. 5a). The subcellular localization of these proteins remains unknown, although most of them contain predicted nuclear localization signals (NLSs), indicating a role in nuclear function. Northern blot analysis indicates that PHF2 is ubiquitously expressed, and *in situ* hybridization has shown that the majority of *PHF2* gene expression is concentrated in the embryonic neural tube and root ganglia in mice[40]. Although the wild-type function of PHF8 remains to be determined, mutations in human *PHF8* cause inherited X-linked mental retardation (XLMR)[41,42]. Disease-causing mutations are predicted to result in truncations of PHF8, which remove a postulated NLS and presumably abrogate normal nuclear function[41,42]. The *C. elegans* PHF2/PHF8 orthologue 4F429 is enriched in the worm ovary, and RNAi-mediated knockdown causes low-penetrance embryonic lethality[43]. The functions of the other human PHF2/PHF8 homologue (KIAA1718) and the *C. elegans* orthologue spt-3 remain unknown.

Based on phylogenetic analysis, the JmjC domain of the PHF2/PHF8 group is closely related to that of the JHDM1 group (FIG. 3). With the exception of the mammalian PHF2 proteins, which have a histidine-to-tyrosine substitution within the third iron-binding residue of the JmjC domain, the Fe(II)-binding and αKG-binding residues in the PHF2/PHF8 group are identical to those found in JHMD1, making these proteins excellent candidates for histone demethylases (FIG. 5b). Mutations in genes that have roles in epigenetic regulation have previously been associated with XLMR[44]. The fact that truncation mutations within PHF8 result in XLMR indicates a potential link between histone methylation patterns and mental retardation.

The JARID1/JARID2 family

The JARID group of proteins contains two subgroups: JARID1 and JARID2 (FIGS 3,6). In higher eukaryotes, the JARID1 subgroup contains JmjN, AT-rich interactive, C5HC2-zinc-finger and PHD-finger domains (FIG. 6a). The JARID2 subgroup does not contain PHD domains, and its *Drosophila* homologue lacks the C5HC2-zinc-finger domain.

The JARID1 subgroup contains orthologues from yeast to humans. There are four distinct JARID1 members in mammals: JARID1A (RBP2), JARID1B (PLU1), JARID1C (SMCX) and JARID1D (SMCY). JARID1A was initially discovered as a retinoblastoma (RB)-binding protein in a yeast two-hybrid screen[45]. Subsequent functional analysis revealed a role for JARID1A in the activation of RB-mediated transcription and, paradoxically, as a factor that antagonizes RB function under certain conditions[46,47]. JARID1A also physically associates with several nuclear hormone receptors (NRs) to facilitate NR-mediated gene expression[48]. *JARID1B* transcripts are mainly found in the adult testis but are also transiently expressed during early development[49–51]. JARID1B interacts with the developmentally important transcription factors BF1 and PAX9 to potentiate transcriptional repression[52], and is frequently upregulated in breast cancers[49,51].

JARID1C and *JARID1D* lie on the X and Y chromosomes, respectively. No cellular role has been defined for these proteins, although JARID1D was first identified as a male-specific antigen that contributes to the sex-specific tissue-transplantation rejection response[53,54]. Disruption of the *D. melanogaster JARID1* orthologue, *Little imaginal disks* (*Lid*), results in defects that phenocopy mutations in trithorax proteins[55]. Given the involvement of human JARID1 proteins in transcriptional regulation, LID might modulate transcription as part of the trithorax system.

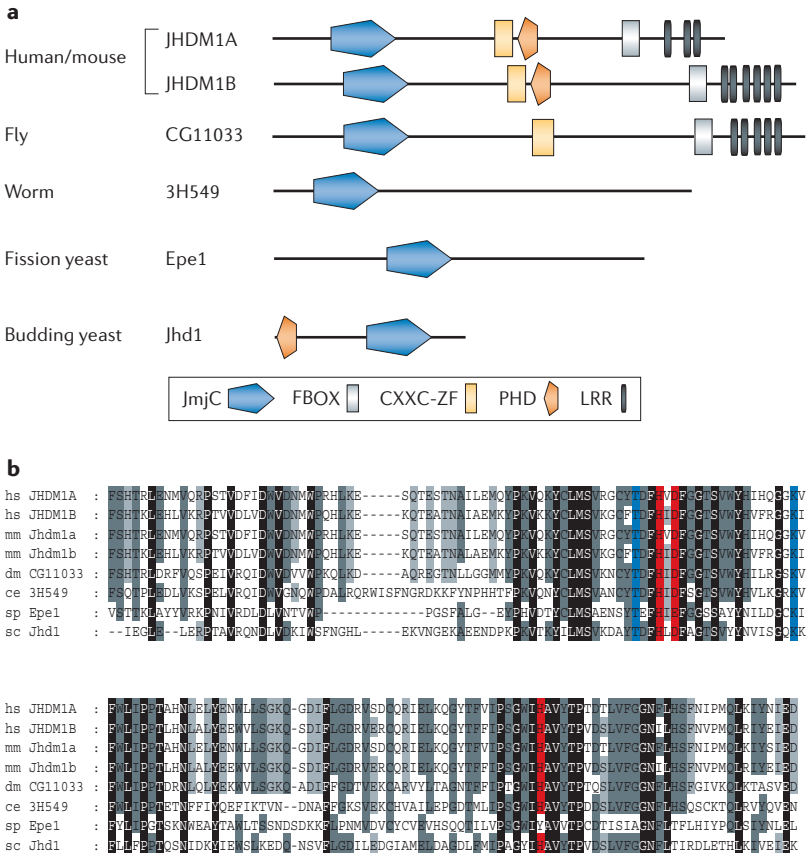


Figure 4 | **JHDM1 proteins are H3K36 demethylases. a** | The domain architecture of JHDM1 proteins is highly conserved in flies, mice and humans, but orthologues found in lower eukaryotes lack several of the domains found in higher eukaryotes. (JmjC, Jumonji C domain; PHD, plant homeobox domain; LRR, leucine rich repeats) **b** | Multiple sequence alignment of the JHDM1 group JmjC domains shows a high degree of homology within the predicted Fe(II)-binding (red) and αKG-binding sites (blue), suggesting that most JHDM1 orthologues are likely to be H3K36 demethylases. A substitution mutation within the third Fe(II)-binding residue of the *Schizosaccharomyces pombe* (sp) orthologue, Epe1, abrogates its H3K36 demethylase activity. Examples from *Homo sapiens* (hs), *Mus musculus* (mm), *Drosophila melanogaster* (dm), *Caenorhabditis elegans* (ce) and *Saccharomyces cerevisiae* (sc).

Members of the JARID2 subgroup are found in *D. melanogaster* and have orthologues in higher organisms. Mouse *Jarid2* was first identified in a gene-trap screen as an important factor in neural tube formation, and was named *Jumonji* (in reference to the Japanese character) because of the abnormal cruciform shape that the neural groove formed in the homozygous null mouse[20,56]. JARID2/JUMONJI has important roles in organogenesis; it functions as a transcriptional corepressor of cardiac genes, and controls cellular proliferation by interacting with transcription factors such as myocyte-specific enhancer factor 2, Nkx2.5, GATA4, and RB[57–60].

The function of the JmjC domain seems to be different within the two JARID subgroups, as the amino acids required for enzymatic function are intact in most members of the JARID1 subgroup but completely lacking in the JARID2 subgroup (FIG. 6b). The extensive conservation of the JmjC domain among JARID1 proteins, and their documented roles in transcriptional regulation, indicate that some of these proteins will constitute functional hydroxylases or demethylases (FIG. 6b).

The mammalian JARID1 homologues can function as transcriptional co-activators or corepressors, indicating that their substrate specificity could be uniquely tailored to their distinct functions in modulating transcription. Studies in *D. melanogaster* have shown that there is a delicate balance between the activation functions of the trithorax proteins and the repressive functions of the Polycomb (PcG) proteins in epigenetic memory and gene regulation[61]. It is tempting to speculate that LID, the sole JARID1 orthologue in *D. melanogaster*, has evolved as part of the trithorax system to keep its target loci free of repressive H3K27 methylation, which is mediated by PcG proteins[55]. Although both fission and budding yeast lack H3K27 methylation, they both have at least one JARID1 orthologue (Spac1002 and Yjr119c, respectively) with a JmjC domain, which contains cofactor binding residues that are compatible with enzymatic activity (FIG. 6b). Therefore, in agreement with the divergent roles of the human JARID1 group members in transcriptional regulation, additional JARID1 group members might have evolved in mammals to expand the diversity of JARID1 function through altered substrate specificity. Further analysis of the potential enzymatic activity of JARID1 proteins might identify novel histone demethylase enzymes with diverse functions in transcriptional regulation and epigenetic memory.

The JHDM3/JMJD2 family

The JHDM3/JMJD2 group of proteins has orthologues from yeast to humans. In higher eukaryotes, proteins of this family contain JmjN, PHD and Tudor domains in addition to the JmjC domain (FIG. 7a). There are four *JHDM3/JMJD2* genes (*JHDM3A/JMJD2A*, *JMJD2B*, *JMJD2C* and *JMJD2D*) in the human genome, each of which has orthologues in the mouse. It is thought that *JMJD2D* gave rise to two additional human genes, *JMJD2E* and *JMJD2F*, through local retrotransposition[62]. The *JMJD2D*, *JMJD2E* and *JMJD2F* genes encode shorter protein products that lack the C-terminal PHD and Tudor domains found in JHDM3A/JMJD2A, JMJD2B and JMJD2C. JHDM3A/JMJD2A is the only functionally-characterized member of the mammalian JHDM3/JMJD2 family, although amplification of *JMJD2C* (*GASC1*) has been implicated in oesophageal cancers[63]. JHDM3A/JMJD2A was originally identified as a transcriptional repressor associated with the NCoR corepressor complex[64], but has also been shown to physically interact with RB to repress E2F target genes[65]. More recently, it was suggested that the tandem Tudor domain of JHDM3A/JMJD2A potentially contains a chromatin-targeting module that directly binds methylated H3K4, H3K9 and H4K20 (REFS 66,67).

Several groups have recently shown that the mammalian JHDM3/JMJD2 proteins are functional histone demethylases that target H3K9 and H3K36, and require both the JmjN and JmjC domains for activity[15–18]. The dual substrate specificity of JHDM3/JMJD2 proteins indicates a previously unrealized functional link between H3K36 and H3K9 methylation, and suggests

Tudor domains
A repeated domain first identified in the *Drosophila melanogaster* Tudor protein, which has subsequently been identified in other proteins as a domain capable of mediating protein–nucleotide and protein–protein interactions. Recently, some Tudor domains have been shown to specifically associate with methylated lysine residues.

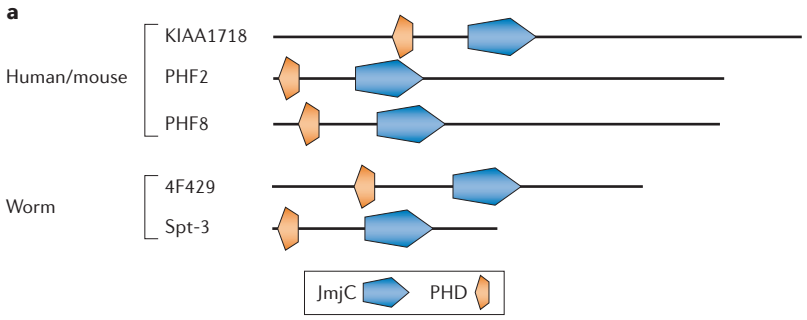


Figure 5 | **PHF2/PHF8 proteins are related to the JHDM1 histone demethylases.** **a** | PHF2/PHF8 orthologues are found in worms, mice and humans but absent from flies. Each member of this group has a JmjC (Jumonji C) domain and a closely associated PHD (plant homeobox) domain. **b** | With the exception of mouse and human PHF2, the PHF2/PHF8 group of proteins are conserved within the predicted Fe(II)-binding (red) and αKG-binding (blue) residues of the JmjC domain. The close similarity of the PHF2/PHF8 JmjC domain with that of JHDM1 proteins suggests that these proteins are excellent histone demethylase candidates. Examples from *Homo sapiens* (hs), *Mus musculus* (mm) and *Caenorhabditis elegans* (ce).

that JHDM3/JMJD2 proteins might function to remove H3K9 and H3K36 methylation from transcribed ORFs during the transition from the active state to the repressed state[2,68]. RNAi-mediated knockdown of the *C. elegans JHDM3/JMJD2* orthologue perturbed H3K9 and H3K36 methylation levels, which resulted in CEP-1/p53-dependent germ-cell apoptosis and caused defects in the progression of meiotic double-strand break repair[15]. It remains to be seen if this function of the worm JHDM3/JMJD2 homologue depends on its demethylase activity and whether this function is conserved in higher eukaryotes.

The JHDM3/JMJD2 group of proteins can demethylate trimethyl lysine, in contrast to all other characterized histone demethylases, which target di- or monomethyl lysine. The crystal structure of the N-terminal region of JHDM3A/JMJD2A (which contains the JmjN and JmjC domains) in association with Fe(II) and αKG provided the first detailed view of the core catalytic domain of an active histone demethylase (FIG. 2A)[28]. The structure of the JHMD3A/JMJD2A catalytic domain is similar to the FIH asparagine hydroxylase, with the exception that the potential position-1 residue that coordinates αKG is predicted to lie 13 residues C-terminal of the position-1 Fe(II)-binding residue, which indicates that there is some structural flexibility in the residues that are utilized for this interaction. The JHDM3A/JMJD2A structure also showed that a zinc molecule has a structural role in the formation of an active enzyme. Consistent with the requirement of the JmjN domain for enzymatic activity, the crystal structure shows that this domain folds into the JmjC domain and contributes to the overall structure of the protein. Because the existing crystal structure does not include the histone substrate, it is difficult to predict which feature confers the unique ability of this enzyme to remove the trimethyl modification state. Bioinformatic analysis of JHDM3/JMJD2 family members has also failed to provide any clues as to why JHDM3/JMJD2 proteins have this property (FIG. 7b).

In addition to the JHDM3/JMJD2 proteins in higher eukaryotes, budding yeast also has two JHMD3/JMJD2 orthologues, Rph1 and Gis1, which function as transcriptional repressors[69]. Analysis of the JmjC domains of these orthologues indicates that the position-three Fe(II)-binding residue is substituted in Gis1 but conserved in Rph1. Given that budding yeast chromatin lacks histone H3K9 methylation, Rph1 might retain H3K36 demethylase activity in this organism.

The UTX/UTY family

The ubiquitously transcribed tetratricopeptide repeat (TPR) UTX/UTY group of proteins have orthologues from worms to humans, and contain a tetratricopeptide repeat in addition to the JmjC domain (FIG. 8a). This group also includes the JMJD3 proteins (which have been identified in humans and mice) that lack the TPR domain. We include JMJD3 proteins in this group because of their extensive homology with UTX and UTY, both within and outside the JmjC domain. Three additional UTX/UTY/JMJD3-related proteins, which also lack the TPR domain, are found in *C. elegans*, which indicate these gene products might have a unique role in nematodes. The mammalian *UTY* gene resides on the Y chromosome and is expressed in most tissues in the male mouse[70]. The X-linked *UTX* gene escapes X inactivation in females and is ubiquitously expressed[71]. The function of the UTX/UTY proteins is largely unknown, except for the observation that UTY contributes to male-specific antigens, which can lead to the sex-specific tissue-transplantation rejection response[70]. The expression and function of the UTX/UTY/JMJD3 orthologues in lower eukaryotes remain uncharacterized.

With the exception of the mouse UTY protein and one of the *C. elegans* proteins, the JmjC domains of UTX/UTY orthologues in worms, flies and mammals contain conserved residues within the predicted cofactor binding sites that are compatible with enzymatic activity (FIG. 8b). The mammalian JMJD3 proteins have a lysine-to-arginine substitution in the second αKG-binding site but, because the charge properties of these two amino acids are similar, this substitution might not abrogate cofactor binding. Therefore, the UTX/UTY/JMJD3 family of proteins are probably enzymes and warrant further experimental characterization to determine whether they localize to the nucleus and contribute to chromatin metabolism.

NCoR corepressor complex
An HDAC-containing protein complex that has general roles in the transcriptional repression of hormone-regulated genes through interaction with unliganded nuclear hormone receptors.

Tetratricopeptide repeat
A structural motif responsible for mediating protein–protein interactions. The tetracopeptide repeat motif consists of tandem repeats of 34 amino-acid residues.

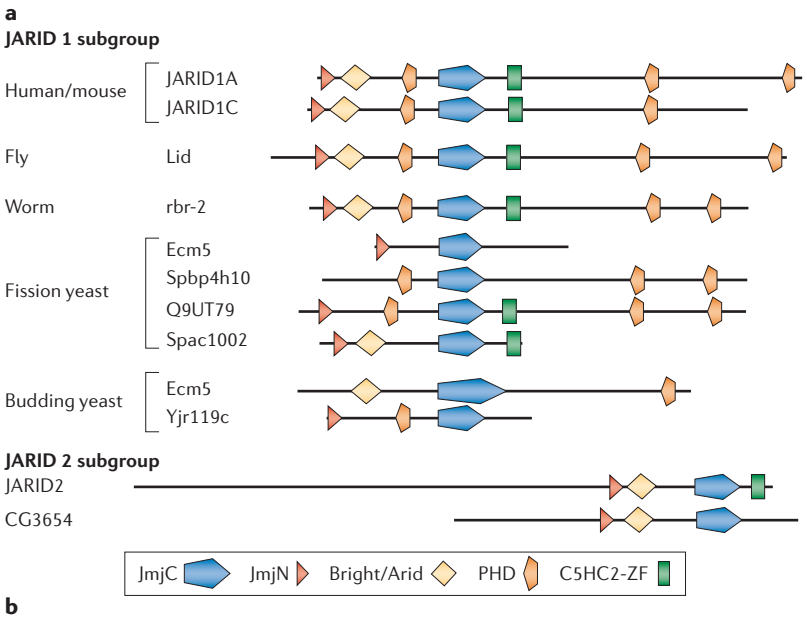


Figure 6 | **The JARID1/2 group contains potentially active enzymes. a** | The domain architecture of proteins within the JARID1/2 group is conserved; the JARID1 and JARID2 subgroups subdivision is based on spatial arrangement of these domains. (JmjC, Jumonji C domain; JmjN, Jumonji N domain; PHD, plant homeobox domain). **b** | Multiple sequence alignment demonstrates that the majority of the JARID1 proteins contain conserved residues within the Fe(II)-binding site (red) and αKG-binding site (blue), which are compatible with enzymatic activity. By contrast, the residues important for cofactor binding are substituted in the JARID2 subgroup and in some *Schizosaccharomyces pombe* (sp) proteins of the JARID1 subgroup (JmjC domains with substituted residues are denoted with an asterisk). Examples from *Homo sapiens* (hs), *Drosophila melanogaster* (dm), *Caenorhabditis elegans* (ce) and *Saccharomyces cerevisiae* (sc).

The JHDM2 family

The JHDM2 family of proteins have orthologues from flies to humans; they possess JmjC and modified zinc-finger domains[72] (FIG. 9a). The mammalian JHDM2 family is comprised of four human proteins: hairless (HR), JHDM2A, JHDM2B, and JHDM2C. *HR* was originally identified in mice, and its mutation results in universal congenital alopecia[73]. Mutation in its human homologue also causes hereditary alopecia[74]. *HR* is expressed in the skin and the brain[73,74], where it probably functions as a corepressor of the thyroid hormone receptor[75].

JHDM2A was originally identified in a testis-specific cDNA library, and *in situ* hybridization indicates that it is specifically expressed in meiotic and post-meiotic male germ cells[76]. Using the same activity-based assay that successfully identified JHDM1, we have recently demonstrated that the JHDM2A is a mono- and dimethyl H3K9-specific demethylase[14]. In response to hormone treatment, JHDM2A associates with the androgen receptor (AR) and contributes to AR-mediated gene activation, probably by keeping the promoter free of H3K9 methylation[14]. Human *JHDM2B* (*5qNCA*) is located in a region of chromosome 5 that is frequently deleted in myeloid leukaemia[72]. Ectopic expression of JHDM2B in cell lines carrying this deletion inhibits clonogenic growth, indicating that JHDM2B might have a role in cancer. Given that JHDM2B is an active H3K9 demethylase (K. Yamane and Y. Z., unpublished observations), it will be important to determine whether its enzymatic activity is involved in cell proliferation. JHDM2C (TRIP8) interacts with the ligand-binding domain of rat TRβ in a hormone-dependent manner[77], although the functional significance of this interaction has not been analysed. An interesting feature that is shared among JHDM2 proteins is their involvement in NR-mediated functions, indicating that removal of histone methylation might be an important part of the NR-based gene-regulation system.

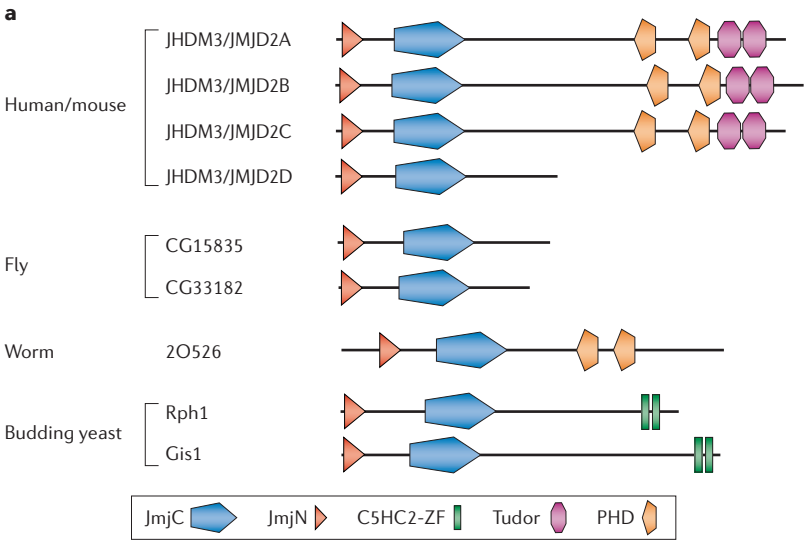
The JmjC domain of HR is the most divergent within the JHDM2 group, and all but one of the predicted cofactor-binding sites contain substituted residues, indicating that this protein is likely to be deficient in enzymatic activity (FIG. 9b). The remaining JHDM2 proteins have a highly conserved JmjC domain and probably constitute novel H3K9 histone demethylases (FIG. 9b). Surprisingly, fission yeast lacks orthologues of both the known JmjC-domain-containing H3K9 demethylases (JHDM2 and JHDM3/JMJD2) but contains H3K9 methylation. So, H3K9 methylation might not be dynamically regulated in *S. pombe*. Alternatively, one or more additional histone demethylases might be capable of removing H3K9 methylation in this organism.

Another feature of the JHDM2 and UTX/UTY JmjC domain is a 20–32 amino-acid sequence insertion between β-sheets 4 and 5, which is predicted to form a helical secondary structure (FIGS 8,9). Despite this insertion, JHDM2 is enzymatically active, indicating that there is a certain degree of structural flexibility associated with the formation of an enzymatically active JmjC domain.

The JmjC-domain-only family

The seventh group that we discuss contains several JmjC-domain-containing proteins that, apart from the JmjC domain, contain no other recognizable protein domains. This group forms its own branch that is based on homology within the JmjC domain, and includes all of the JmjC-domain proteins that are known to localize to the cytoplasm (FIH, PLA2G4B and HSPBAP1). We propose that these proteins, of which only the MINA53/NO66 group (see below) has orthologues in yeast, might have diverged in higher eukaryotes to carry out functions that are independent of histone demethylation.

FIH. FIH was the first JmjC-domain-containing protein that was shown to be enzymatically active. It functions as an asparagine protein hydroxylase for the hypoxia inducible factor α (HIFα) transcription factor[23]. Under conditions of low oxygen, the HIFα protein translocates from the



b

```
hs JHDM3A  : EKHVDEWNIGRLRIILDLVEKESGITIEGVNTPLLYFGMWKQSFAWTTEDMDLYSINYLHFGEPKSWYSVPPEHGKRLERLAKGF
hs JHDM3B  : DDDVAQWNIGSLRIILDMVERECGTIEGVNTPLLYFGMWKTTFAWTTEDMDLYSINYLHFGEPKSWYAIPPEHGKRLERLAIGF
hs JHDM3C  : DEGVDEWNIARLNIVLDVVEEECGISIEGVNTPLLYFGMWKTTFAWTTEDMDLYSINYLHFGEPKSWYAIPPEHGKRLERLAQGF
hs JHDM3D  : DENTKQWNLGHLGIIQDLLEKECGVVIEGVNTPLLYFGMWKTTFAWTTEDMDLYSINYLHFGEPKTWYVVPPEHGQRLERLAREL
mm Jhdm3a  : EQHVDEWNIGRLKIILDLVEKESGITIEGVNTPLLYFGMWKNSFAWTTEDMDLYSINYLHFGEPKSWYSVPPEHGKRLERLAKGF
mm Jhdm3b  : DDDVAQWNIGNLRIILDMVERECGTIEGVNTPLLYFGMWKTTFAWTTEDMDLYSINYLHFGEPKSWYAIPPEHGKRLERLAIGF
mm Jhdm3c  : DEGVDEWNIARLNIVLDVVEEECGISIEGVNTPLLYFGMWKTTFAWTTEDMDLYSINYLHFGEPKSWYAIPPEHGKRLERLAQGF
mm Jhdm3d  : GENTQHWNMGHLGSLLDVLKQDHDIVIEGVNTPLLYFGMWKTTFAWTTEDMDLYSINYLHFGQPKTWYAVPPEHGRRLERLAREL
dm CG15835 : DEDLDVWNIGRLDIILNLVNTDYNIIIDGVNTALLYFGMWKSSFAWTTEDMDLYSINYLHFGAPKTWYAIPPAYGRRLEKLANET
dm CG33182 : DTDQDSWNINRLGIILDYVNKDYNIQIDGVNTALLYFGMWKTTFAWTTEDMDLYSINYLHFGAPKTWYVVPPECQRKLEKVANQY
ce 2O526   : DAQVEEWNMNRLGIIL----EDTNYEIKGVNTVLLYFGMYKTTFAWTADMDLYSINFLHFGAPKTWFAISSSHADQFERFMSQQ
sc Rph1    : PEGLNVWNVAKLPNILD----HMETKVPGVNDSLLYAGLWKASFSWTLDQDLYSINYLHFGAPKQWYSIPQEDRFKFYKFMQEQ
sc Gis1    : PYDLTLWNLNNLPDSIN-----------SSNRRLLTGQSKQIIPWHLEQNKCSINYLHFGAPKQWYSIPSANTDQFLKILSKE

hs JHDM3A  : FPGS---AQSQEAFIRHKMTLISPLMLKKYGIPFDKVTQEAGEFMITFPYGYHAGFNLGFNCAESTNFATRRWIEYGKQAVLCSC
hs JHDM3B  : FPGS---SQGQDAFIRHKMTLISPIILKKYGIPFSRITQEAGEFMITFPYGYHAGFNLGFNCAESTNFATLRWIDYGKVATQCTC
hs JHDM3C  : FPSS---SQGQDAFIRHKMTLISPSVLKKYGIPFDKITQEAGEFMITFPYGYHAGFNLGFNCAESTNFATVRWIDYGKVAKLCTC
hs JHDM3D  : FPGS---SRGQGAFIRHKVALISPTVLKENGIPFNRITQEAGEFMVTFPYGYHAGFNLGFNCAEAINFATPRWIDYGKMASQCSC
mm Jhdm3a  : FPGS---AQSQEAFIRHKMTLISPLMLKKYGIPFDKVTQEAGEFMITFPYGYHAGFNLGFNCAESTNFATRRWIEYGKQAVLCSC
mm Jhdm3b  : FPGS---SQGQDAFIRHKMTLISPIILKKYGIPFSRITQEAGEFMITFPYGYHAGFNLGFNCAESTNFATLRWIDYGKVATQCTC
mm Jhdm3c  : FPSS---SQGQDAFIRHKMTLISPSVLKKYGIPFDKITQEAGEFMITFPYGYHAGFNLGFNCAESTNFATVRWIDYGKVAKLCTC
mm Jhdm3d  : FPGS---SQGQDAFIRHKVALISPTVLKENGIPFGRITQEAGEFMVTFPYGYHAGFNLGFNCAEAINFATPRWIDYGKVASQCSC
dm CG15835 : FSEN---YQEQNFYIRHKMTMISPKVLRQHNLPYNKITQEAGEFMITFPFGYHAGFNLGFNCAESTNFASKRWIEYGKRASICRC
dm CG33182 : FEAS---YKNQNFYIRHKMTLISPQILKQHDVPVSKITQEAGEFMITFPFGYHAGFNLGFNCAESTNFAMERWIEYGKRAVQCTC
ce 2O526   : FSYQNEYAPQQKAFIRHKTYLVTPRLLRQAGIPYATMVQRPNEFLITFPRGYHVGFNLGYNLAESTNFASQRWIDYGKDAVLCDC
sc Rph1    : FPEE---AKNQPEFIRHKMFLASPKLLQENGIRCNEIVHHEGEFMITYPYGYHAGFNYGYNLAESVLFALEEWLPIGKKLAGKCHC
sc Gis1    : PSSN---KENQPAFIRHQNIITSPDFLRKNNIKFNRVVGFQHEMLITFPMCMYSGFNYGYNFGEGIEGLLDQQAVVRKQPLKGGC
```

Figure 7 | **The JHDM3/JMJD2 proteins are H3K9/K36 demethylases. a** | JHDM3/JMJD2 orthologues are found from yeast to human and all group members contain a JmjC (Jumonji C) and JmjN (Jumonji N) domain. Members of this group contain additional C-terminal domains, which are probably involved in protein targeting (PHD, plant homeobox domain). **b** | With the exception of the budding yeast protein Gis1, all JHDM3/JMJD2 group members have conserved residues within the Fe(II)-binding site (red) and αKG-binding site (blue), indicating that uncharacterized members of this group could be functional histone demethylases. Examples from *Homo sapiens* (hs), *Mus musculus* (mm), *Drosophila melanogaster* (dm), *Caenorhabditis elegans* (ce) and *Saccharomyces cerevisiae* (sc).

cytoplasm to the nucleus where it activates the transcription of genes that promote adaptation to hypoxia. Under normal oxygen conditions, HIFα is hydroxylated by FIH in the cytoplasm on a single asparagine, which can inhibit binding of the p300/CBP histone acetyltransferase[23,25,78]. The structure of FIH in complex with Fe(II) and αKG has proved to be a useful template for assigning predicted cofactor-binding sites and secondary structure to other JmjC-domain-containing proteins[24–27] (Supplementary information S2 (figure)). The fact that FIH is a protein hydroxylase indicates that some JmjC-domain-containing proteins will function to modify proteins instead of catalysing demethylation. Because FIH is only found in mice and humans, it remains possible that the asparagine hydroxylase activity of FIH is an evolutionary adaptation that has arisen from the underlying enzymatic function of the JmjC domain in demethylation.

MINA53/NO66. The MINA53/NO66 group has orthologues from yeast to human. Myc-induced nuclear antigen 53 (MINA53) was originally identified in a screen for transcripts that are upregulated after Myc transgene expression in cultured cells[79]. It is also significantly upregulated in lung macrophages of coal miners who have been exposed to mineral dust[80]. MINA53 accumulates in the nucleus and is concentrated in the nucleolus[79,81]. Knockdown of *MINA53* using small interfering RNAs reduces proliferation of transformed rodent and human cell lines, indicating a potential role in cell proliferation[79,80]. A closely related protein, NO66, was identified as a component of *Xenopus laevis* nucleoli, where it interacts with components of pre-ribosomal complexes[82]. NO66 also localizes to some nuclear heterochromatic foci during the late stages of S-phase, indicating that it might have a role in heterochromatin formation.

The JmjC domains of the mammalian MINA53 proteins differ from the remainder of the MINA53/NO66 group, as they have a leucine instead of a threonine or phenylalanine in the first αKG-binding residue (Supplementary information S2 (figure)). It is unclear whether this substitution is compatible with enzymatic activity, but the remainder of the proteins in the MINA53/NO66 group have residues within the JmjC domain that are compatible with enzymatic activity (Supplementary information S2 (figure)). The nucleolar localization of MINA53/NO66 proteins indicates a possible role in modulation of rRNA cap methylation and/or arginine/lysine methylation of heterogeneous nuclear ribonucleoproteins. It remains to be seen whether these proteins have roles in ribosome maturation.

PLA2G4B. Human cytosolic phospholipase A2β (PLA2G4B) is a cytosolic protein. It contains an N-terminal JmjC domain which is linked to a calcium-sensing C2 domain and a C-terminal phospholipase domain. The PLA2G4B proteins vary in domain organisation among members. For example, the human and dog *PLA2G4B* orthologues contain all three domains, whereas orthologues from mice and flies have only the N-terminal portion of the protein, which contains the JmjC domain.

Little is known about human PLA2G4B, with the exception that it is ubiquitously expressed and that its phospholipase domain can hydrolyse phospholipids that contain arachidonic acid[83]. Human paralogues of PLA2G4B utilize this phospholipase activity as part of the normal biosynthesis of inflammatory lipids, and have been implicated in a wide variety of normal cellular processes[84]. The JmjC-domain residues that are predicted

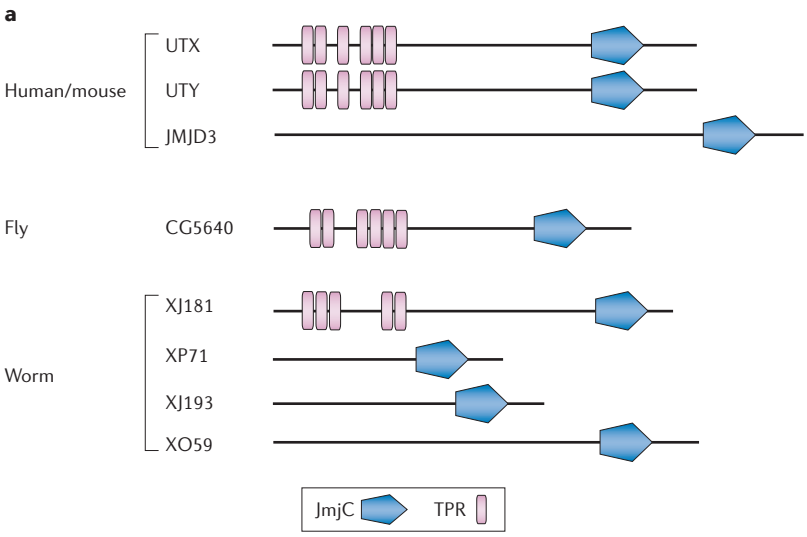


Figure 8 | **UTX/UTY proteins are poorly characterized proteins with conserved JmjC domains. a** | UTX/UTY orthologues are found from worms to humans and are characterized by a JmjC (Jumonji C) domain and TPR (tetratricopeptide) repeats. **b** | High levels of conservation are evident within the Fe(II)-binding site (red) and αKG-binding site (blue) of UTX/UTY proteins, indicating that these are potentially active enzymes. Yellow shading indicates the amino-acid insertion within the UTX/UTY JmjC domain. Examples from *Homo sapiens* (hs), *Mus musculus* (mm), *Drosophila melanogaster* (dm) and *Caenorhabditis elegans* (ce).

to be important for cofactor binding are conserved in PLA2G4B orthologues, indicating that some members of this family might have two independent enzymatic activities (Supplementary information S2 (figure)).

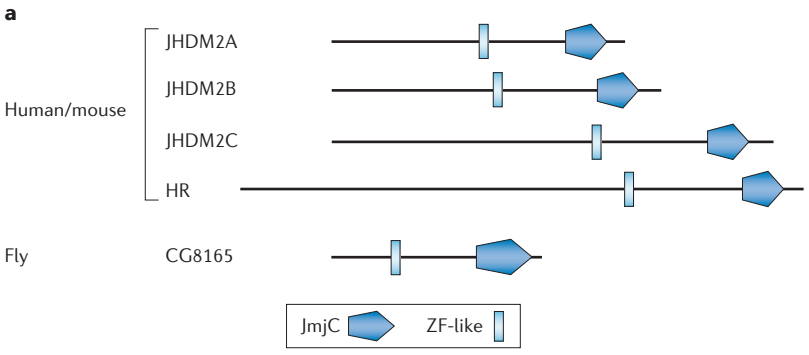
HSPBAP1. Heat shock 27kDa associated protein 1 (HSPBAP1; also known as PASS1) was identified in a yeast two-hybrid screen using heat shock protein 27kDa (HSP27) as bait. Orthologues of this protein exist from fly to human[85]. HSPBAP1 is expressed in a wide variety of tissues and localizes to the cytoplasm. Association of HSPBAP1 with HSP27 inhibits the capacity of HSP27 to protect cells against sublethal heat shock[85]. A recent study has implicated HSPBAP1 as a fusion partner of disrupted in renal carcinoma 3 (DIRC3) in familial renal cancer[86]. The residues that are predicted to be important for cofactor binding are conserved in HSPBAP1 orthologues, indicating that these proteins might constitute functional hydroxylases with enzymatic roles in the cytoplasm (Supplementary information S2 (figure)).

PTDSR. The mouse phoshatidylserine receptor (*PTDSR*) was originally identified in a screen for factors that bind phosphatidylserine during clearance of apoptotic cells[87]. *Ptdsr* is widely expressed in mouse tissues, and homozygous *Ptdsr*-null mice die perinatally with abnormalities in the development of the lungs, eyes, kidneys, intestine and liver[87–89]. Based on its affinity for phosphatidylserine, which is exposed on the outer plasma membrane of cells undergoing apoptosis, it was originally proposed that *Ptdsr* was an important mediator in the recognition and removal of apoptotic cells. Several subsequent studies have provided evidence supporting a role for PTDSR in cell engulfment. Recent analysis of the subcellular localization of PTDSR has raised doubts about its role in the clearance of apoptotic cells because it is predominantly found in the nucleus[90,91]. Furthermore, careful analysis of *Ptdsr*-null mice has failed to demonstrate an essential role in the clearance of apoptotic cells[89], which indicates that the biologically relevant role of PTDSR might be to modulate gene expression[90]. The predicted cofactor-binding sites in the JmjC domain of PTDSR are compatible with enzymatic activity, raising the possibility that this group of proteins might possess hydroxylase/demethylase activity (Supplementary information S2 (figure)).

JMJD5, JMJD4 and L0C339123. The functions of proteins of the JMJD5, JMJD4 and the L0C339123 groups remain unknown. Members of the JMJD4 and JMJD5 groups are the only JmjC-domain-containing proteins that have a serine residue in place of the normal threonine in the first αKG-binding site (Supplementary information S2 (figure)). Despite the similarity of these residues in molecular composition and charge, there seems to be a bias for inclusion of threonine as opposed to serine in this position in most JmjC-domain-containing proteins. This could be due to an inherent preference for threonine, as opposed to the less bulky serine residue, in forming an enzymatically competent JmjC-domain active site. The JmjC domain of the LOC339123 family contains substitutions in both the Fe(II)- and αKG-binding sites, indicating that these proteins are unlikely to be functional enzymes (Supplementary information S2 (figure)).

Concluding remarks

Most histone lysine methylation is mediated by a large family of methyltransferase enzymes that contain an enzymatic SET domain[1]. However, it was only in 2000 that the widespread importance of the SET domain was discovered[92]. Since then, there has been an exponential growth in our understanding of how histone methylation contributes to chromatin function through the regulation of gene activity, chromatin



b

```
hs JHDM2A  : MFSRFDDLMANIPLPEYTRRDGKLNLASRLFNYFVRPDLCPKMYNAYGLITPEDRKYGTNLHLDVSDAANVMVY
hs JHDM2B  : MFTRFEDLMENLPLPEYTKRDGRLNLASRLFSYFVRPDLCPKMYNAYGLITAEDRRVGTNLHLDVSDAVNVMVY
hs JHDM2C  : MFARYEDLLKSLPLPEYCNPEGKFNLASHLFGFFVRPDLCPRLCSAYGVVAAKQHDIGTNLHIEVSDVVNILVY
hs HR      : DTSRVENLAASLPLPEYCALHGKLNLASYLFPGLALRPLEPQLWAAYGVSPHRGH-LGTKNLCVEVADLVSILVH
mm Jhdm2a  : MFSRFDDLMANIPLPEYTRRDGKLNLASRLFNYFVRPDLCPKMYNAYGLITPEDRKYGTNLHLDVSDAANVMVY
mm Jhdm2b  : MFTRFEDLMENLPLPEYTKRDGRLNLASRLFSYFVRPDLCPKMYNAYGLITAEDRRVGTNLHLDVSDAVNVMVY
mm Jhdm2c  : MFTRYEDLFRCLPLPEYCNPEGKFNLASHLFGFFVRPDLCPRLCSAYGVAAAKQHDIGTNLHIEASDVVNVLVY
mm Hr      : DASRVQNLASSLPLPEYCAHQGKLNLASYLPLGLTLHPLEPQLWAAYGVNSHRGH-LGTKNLCVEVSDLISILVH
dm CG8165  : LFTRFADLMKGLPMPEYTLRTGNLNIASCLPKMFVPPDLCPKMYNAYGSALHPDK--GTKNLCLDISDAVNIMVY

hs JHDM2A  : VGIPK-GQCE-QEEEVLKTIQDGDSDELTIKRFIEGKEKPGAIWHIYAAKDTEKIREFLKKVSEEQGQENPADHD
hs JHDM2B  : VGIPI-GEGA-HDEEVLKTIDEGDADEVTKQRIHDGKEKPGAIWHIYAAKDAEKIRSLLRKVGEEQGQENPFDHD
hs JHDM2C  : VGIAK-GNGILSKAGILKKFEEEDLDDILRKRLKDSSEIPGAIWHIYAGKDVDKIREFLQKISKEQGLEVLPEHD
hs HR      : ADTPLPAWHR-AQKDFLSGLDGEG--------LWSPGSQVSTVWHVFRAQDAQRIRRFLQMVCPAGAGALEEGAP
mm Jhdm2a  : VGIPK-GQCE-QEEEVLRTIQDGDSDELTIKRFIEGKEKPGAIWHIYAAKDTEKIREFLKKVSEEQGQDNPADHD
mm Jhdm2b  : VGIPV-GEGA-HDEEVLKTIDEGDADEVTKQRIHDGKEKPGAIWHIYAAKDAEKIRELLRKVGEEQGQENPFDHD
mm Jhdm2c  : VGIAK-GNGVLSKAGILKKFEEEELDDVLRKILKDSSEIPGAIWHIYAGKDVDKIREFLQKISKEQGLEVLPEHD
mm Hr      : AEAQLPPWYR-AQKDFLSGLDGEG--------LWSPGSQTSTVWHVFRAQDAQRIRRFLQMVCPAGAGTLEEGAP
dm CG8165  : VGIPQDGDTRPQMAATQKAIEIGGCDYITRARCQSPDVLPGAIWHIFPARDADKIRDLLNRVTLEKGFRLEEDHD

hs JHDM2A  : PIHDQSWYLDRSLRKRLHQEYGVQGWAIVQFLGDVVFIPAGAPHQVHNLYSCIKVAEDFVSPEHVKHCFWLTQEF
hs JHDM2B  : PIHDQSWYLQTLRKRLYEEYGVQGWAIVQFLGDAVFIPAGAPHQVHNLYSCIKVAEDFVSPEHVKHCFRLTQEF
hs JHDM2C  : PIRDQSWYVNKKLRQRLLEEYGVRTCTLIQFLGDAIVLPAGALHQVQNFHSCIQVTEDFVSPEHLVESFHLTQEL
hs HR      : ----GSCYLDAGLRRRLREEWGVSCWTLLQAPGEAVLVPAGAPHQVQGLVSTVSVTQHFLSPETSALSAQLCHQG
mm Jhdm2a  : PIHDQSWYLDRSLRKRLYQEYGVQGWAIVQFLGDVVFIPAGAPHQVHNLYSCIKVAEDFVSPEHVKHCFWLTQEF
mm Jhdm2b  : PIHDQSWYLLQILRKRLFEEYGVQGWAIVQFLGDAVFIPAGAPHQVHNLYSCIKVAEDFVSPEHVKHCFRLTQEF
mm Jhdm2c  : PIRDQSWYVNRKLRQRLLEEYGVRACTLIQFLGDAIVLPAGTLHQVQNFHSCVQVTEDFVSPEHLVQSFHLTQEL
mm Hr      : ----GSCYLDAGLRRRLREEWGVSCWTLLQAPGEAVLVPAGAPHQVQGLVSTVSVTQHFLSPETSALSAQLYHQG
dm CG8165  : PIHDQNWYLLDKLRARLFKEYGVEGHPIVQCLGDAVFIPAGAPHQVQNLHNCIKVAEDFVSPENITHCYHLTHEF
```

Figure 9 | **JHDM2 proteins are H3K9 demethylases. a** | JHDM2 proteins are found from flies to humans and have a zinc-finger (ZF)-like domain in addition to a JmjC (Jumonji C) domain. **b** | With the exception of the human and mouse Hairless (HR) proteins, the Fe(II)-binding site (red) and αKG-binding site (blue) are conserved amongst JHDM2 orthologues, indicating that these proteins are probably functional H3K9 demethylases. Yellow shading indicates the amino-acid insertion within the JHDM2 JmjC domain. Examples from *Homo sapiens* (hs), *Mus musculus* (mm) and *Drosophila melanogaster* (dm).

structure, dosage compensation and epigenetic memory[1]. With the identification of enzymes that antagonize histone methylation[7–9,13], we are on the cusp of another rapid advancement in our understanding. Although it is clear that not all JmjC-domain-containing proteins are functional enzymes, our analysis suggests that many of these proteins satisfy cofactor-binding requirements and seem to be excellent candidates for histone or non-histone protein demethylases. The specific roles of individual protein demethylases in normal cellular function remain to be shown, but it is clear that histone methylation can be dynamically regulated in a manner that is analogous to acetylation or phosphorylation.

Here we have analysed 98 non-redundant JmjC-domain proteins from yeast, worms, flies, mice and humans. This analysis has allowed us to categorize this extensive family of proteins into seven groups based on the JmjC-domain homology and full-length protein domain architecture (FIG. 3). By defining JmjC-protein groups that include the relevant orthologues from commonly studied model organisms, we clarify the evolutionary relationship between members of this protein family and provide a resource that will benefit future functional analyses of JmjC-domain proteins. Furthermore, by highlighting proteins with potential enzymatic activity, and speculating on their possible substrates, we provide a focus for further enzymatic characterization. The JHDMs (JHDM1, JHDM2 and JHDM3/JMJD2) identified so far each contain multiple enzymatically active homologues in mice and humans, all with a similar substrate specificity. This indicates that, despite the large number of JmjC-domain proteins found in the human (30) and mouse (30) genomes, individual JmjC-protein groups might represent proteins that are differentially expressed, or have different chromatin-targeting mechanisms, but retain the same substrate specificity. It remains to be determined if this trend will continue as other JmjC-protein groups are characterized and novel enzymes are identified, but if so, this will certainly limit the potential for this class of enzymes to regulate the full range of histone methylation sites and individual modification states.

Enzymes that contain SET domains, much like JmjC-domain proteins, fall into defined protein families, some of which have redundant modification-site specificity[93]. The characterized SET-domain proteins that have activity towards H3K9 (Suv39H1/2, ESET, RIZ, G9A and GLP[1]) and H3K36 (NSD1 (REF. 1) and HYPB[94]) seem to use unique targeting mechanisms and result in the accumulation of different final modification states. Similar to the number and complexity of SET-domain enzymes that modify H3K9 and H3K36, the JmjC-domain-containing enzymes that counteract H3K9 (JHDM2A–C and JHDM3A–F) and H3K36 (JHMD1A–B and JHMD3A–F) methylation consist of multiple related proteins with unique specificity towards either the di- or trimethyl modification states. This observation suggests that the level of complexity within the histone methylation system might, in part, define the corresponding number and complexity of demethylase enzymes that antagonize these modifications. If this hypothesis holds true, an expanded modification/demodification system might have evolved in higher eukaryotes owing to the requirement for increased complexity in the mechanisms of targeting and regulating histone methylation at specific genomic loci. Of the characterized mammalian histone methyltransferase enzymes, only one is thought to be responsible for H3K27 methylation (EZH2), and one for H3K79 methylation (DOT1L). Given the reduced complexity in the methyltransferase enzymes that place these marks, it will be interesting to determine whether single enzymes are also responsible for H3K27 and H3K79 demethylation.

Following the identification of histone demethylase enzymes, the role of histone demethylation in normal cellular function is now being explored. So far, histone demethylases have been shown to counteract histone modifications that oppose the transcription state of a given gene, and to remove histone modifications during the transition from one transcriptional state to another[7–9,13,15–18]. With the availability of new information from genome-wide histone-modification profiling and studies of the dynamics of histone methylation patterns, avenues through which to explore how histone demethylases contribute to cellular function are becoming clear. Of particular interest is the recent observation that EZH2, a PcG-group H3K27 methyltransferase, functions in maintaining the undifferentiated state of embryonic stem (ES) cells[95,96]. When ES cells are induced to differentiate, a series of genes that contain H3K27 methylation, and are normally silenced, become activated concomitant with loss of H3K27 methylation[96]. It will be important to determine whether this reactivation relies on an active demethylation and, if so, to identify the responsible demethylase enzyme(s). Interestingly, JMJD3, a member of the UTX/UTY protein grouping, is upregulated during ES-cell differentiation, providing a possible candidate demethylase for H3K27 methylation[96]. Insight into how H3K27 methylation is regulated will influence our understanding of how the transition from repressed to active chromatin states is achieved, and provide opportunities to advance therapeutics related to stem-cell technology and cancer treatment. With the constantly expanding repertoire of site-specific histone demethylases, we look forward to determining how these unique enzymes might function in regulating transcription, epigenetic inheritance and cell fate.

1. Martin, C. & Zhang, Y. The diverse functions of histone lysine methylation. *Nature Rev. Mol. Cell Biol.* **6**, 838–849 (2005).
2. Vakoc, C. R., Mandat, S. A., Olenchock, B. A. & Blobel, G. A. Histone H3 lysine 9 methylation and HP1γ are associated with transcription elongation through mammalian chromatin. *Mol. Cell* **19**, 381–391 (2005).
3. Carrozza, M. J. *et al.* Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581–592 (2005).
4. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromodomain. *Nature* **410**, 120–124 (2001).
5. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001).
6. Wysocka, J. *et al.* A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 91–95 (2006).
7. Wang, Y. *et al.* Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* **306**, 279–283 (2004).
8. Cuthbert, G. L. *et al.* Histone deimination antagonizes arginine methylation. *Cell* **118**, 545–553 (2004).
References 7 and 8 identify PADI4 as the first enzyme capable of antagonizing histone arginine methylation.
9. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).
Identifies the first true histone demethylase enzyme, LSD1, which utilizes flavin as a cofactor to carry out an oxidative demethylation reaction.
10. Metzger, E. *et al.* LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437**, 436–439 (2005).
11. Lee, M. G., Wynder, C., Cooch, N. & Shiekhattar, R. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* **437**, 432–435 (2005).
12. Shi, Y. J. *et al.* Regulation of LSD1 histone demethylase activity by its associated factors. *Mol. Cell* **19**, 857–864 (2005).
13. Tsukada, Y. *et al.* Histone demethylation by a family of JmjC-domain-containing proteins. *Nature* **439**, 811–816 (2006).
The authors used a novel *in vitro* histone demethylase assay to biochemically purify and characterize the first JmjC-domain-containing histone demethylase enzyme.
14. Yamane, K. *et al*. JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* **125**, 483–495 (2006).
The authors biochemically purified a novel JmjC-domain-containing histone H3K9 demethylase and demonstrated its role in androgen receptor-mediated gene activation.
15. Whetstine, J. R. *et al.* Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **125**, 467–481 (2006).
16. Klose, R. *et al.* The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and 36. *Nature* **442**, 312–316 (2006).
17. Cloos, P. A. *et al.* The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* **442**, 307–311 (2006).
18. Fodor, B. D. *et al.* Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes Dev.* **20**, 1557–1562 (2006).
References 15 to 18 demonstrate that the JHDM3/JMJD2 histone demethylase enzymes are capable of removing the trimethyl modification state.
19. Clissold, P. M. & Ponting, C. P. JmjC: cupin metalloenzyme-like domains in jumonji, hairless and phospholipase A2β. *Trends Biochem. Sci.* **26**, 7–9 (2001).
The authors identify bioinformatic and structural similarities between metalloenzymes and JmjC-domain-containing proteins.
20. Takeuchi, T. *et al.* Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation. *Genes Dev.* **9**, 1211–1222 (1995).
The authors cloned the Jumonji protein in a gene-trap screen for factors involved in neural tube formation. The JmjC domain was later named on the basis of its presence in the Jumonji protein.
21. Balciunas, D. & Ronne, H. Evidence of domain swapping within the Jumonji family of transcription factors. *Trends Biochem. Sci.* **25**, 274–276 (2000).
22. Dunwell, J. M. & Gane, P. J. Microbial relatives of seed storage proteins: conservation of motifs in a functionally diverse superfamily of enzymes. *J. Mol. Evol.* **46**, 147–154 (1998).
23. Lando, D. *et al.* FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. *Genes Dev*. **16**, 1466–1471 (2002).
The authors identify the JmjC-domain-containing protein FIH as an asparaginyl hydroxylase enzyme and for the first time demonstrate enzymatic activity for a mammalian JmjC-domain protein.
24. Trewick, S. C., McLaughlin, P. J. & Allshire, R. C. Methylation: lost in hydroxylation? *EMBO Rep.* **6**, 315–320 (2005).
The authors propose that JmjC-domain-containing proteins might function as histone demethylases based on the function of the fission yeast JmjC-domain protein Epe1 in regulating silent chromatin structure.
25. Dann, C. E., Bruick, R. K. & Deisenhofer, J. Structure of factor-inhibiting hypoxia-inducible factor 1: an asparaginyl hydroxylase involved in the hypoxic response pathway. *Proc. Natl Acad. Sci. USA* **99**, 15351–15356 (2002).
26. Elkins, J. M. *et al.* Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF1α. *J. Biol. Chem.* **278**, 1802–1806 (2003).
27. Lee, C., Kim, S. J., Jeong, D. G., Lee, S. M. & Ryu, S. E. Structure of human FIH-1 reveals a unique active site pocket and interaction sites for HIF-1 and von Hippel–Lindau. *J. Biol. Chem.* **278**, 7558–7563 (2003).
28. Chen, Z. *et al.* Structural insights into histone demethylation by JMJD2 family members. *Cell* **125**, 691–702 (2006).
A report of the first crystal structure of the enzymatic domain of an active JmjC-domain-containing histone demethylase, JHDM3A/JMJD2A.
29. Clifton, I. J. *et al.* Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. *J. Inorg. Biochem.* **100**, 644–669 (2006).
30. Chenna, R. *et al.* Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31**, 3497–3500 (2003).
31. Huelsenbeck, J. P. & Ronquist, F. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).
32. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
33. Pothof, J. *et al.* Identification of genes that protect the *Caenorhabditis elegans* genome against mutations by genome-wide RNAi. *Genes Dev.* **17**, 443–448 (2003).
34. Ayoub, N. *et al.* A novel JmjC domain protein modulates heterochromatization in fission yeast. *Mol. Cell. Biol.* **23**, 4356–4370 (2003).
The authors show that Epe1 has important roles in modulating heterochromatin formation and provide a basis for speculation that the JmjC domain might have histone demethylase activity.
35. Bai, C. *et al.* SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* **86**, 263–274 (1996).
36. Ayton, P. M., Chen, E. H. & Cleary, M. L. Binding to nonmethylated CpG DNA is essential for target recognition, transactivation, and myeloid transformation by an MLL oncoprotein. *Mol. Cell. Biol.* **24**, 10470–10478 (2004).
37. Jorgensen, H. F., Ben-Porath, I. & Bird, A. P. MBD1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. *Mol. Cell. Biol.* **24**, 3387–3395 (2004).
38. Lee, J. H. & Skalnik, D. G. CpG-binding protein (CXXC finger protein 1) is a component of the mammalian SET1 histone H3–Lys4 methyltransferase complex, the analogue of the yeast SET1/COMPASS complex. *J. Biol. Chem.* **280**, 41725–41731 (2005).
39. Zofall, M. & Grewal, S. I. Swi6/HP1 recruits a JmjC-domain protein to facilitate transcription of heterochromatic repeats. *Mol. Cell* **22**, 681–692 (2006).
The authors demonstrate that Epe1 functions in heterochromatin to promote Pol II accessibility by counteracting repressive chromatin.
40. Hasenpusch-Theil, K. *et al.* PHF2, a novel PHD finger gene located on human chromosome 9q22. *Mamm. Genome* **10**, 294–298 (1999).

41. Laumonnier, F. *et al.* Mutations in PHF8 are associated with X linked mental retardation and cleft lip/cleft palate. *J. Med. Genet.* **42**, 780–786 (2005).
42. Siderius, L. E. *et al.* X linked mental retardation associated with cleft lip/palate maps to Xp11.3–q21.3. *Am. J. Med. Genet.* **85**, 216–220 (1999).
43. Fernandez, A. G. *et al.* New genes with roles in the *Caenorhabditis elegans* embryo revealed using RNAi of ovary-enriched ORFeome clones. *Genome Res.* **15**, 250–259 (2005).
44. Chelly, J. & Mandel, J. L. Monogenic causes of X linked mental retardation. *Nature Rev. Genet.* **2**, 669–680 (2001).
45. Defeo-Jones, D. *et al.* Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene product. *Nature* **352**, 251–254 (1991).
46. Kratzke, R. A. *et al.* Partial inactivation of the RB product in a family with incomplete penetrance of familial retinoblastoma and benign retinal tumors. *Oncogene* **9**, 1321–1326 (1994).
47. Benevolenskaya, E. V., Murray, H. L., Branton, P., Young, R. A. & Kaelin, W. G. Jr. Binding of pRB to the PHD protein RBP2 promotes cellular differentiation. *Mol. Cell* **18**, 623–635 (2005).
48. Chan, S. W. & Hong, W. Retinoblastoma-binding protein 2 (Rbp2) potentiates nuclear hormone receptor-mediated transcription. *J. Biol. Chem.* **276**, 28402–28412 (2001).
49. Lu, P. J. *et al.* A novel gene (PLU-1) containing highly conserved putative DNA/chromatin binding motifs is specifically upregulated in breast cancer. *J. Biol. Chem.* **274**, 15633–15645 (1999).
50. Madsen, B. *et al.* PLU-1, a transcriptional repressor and putative testis-cancer antigen, has a specific expression and localisation pattern during meiosis. *Chromosoma* **112**, 124–132 (2003).
51. Madsen, B. *et al.* Characterisation and developmental expression of mouse Plu-1, a homologue of a human nuclear protein (PLU-1) which is specifically upregulated in breast cancer. *Mech. Dev.* **119**, S239–S246 (2002).
52. Tan, K. *et al.* Human PLU-1 Has transcriptional repression properties and interacts with the developmental transcription factors BF-1 and PAX9. *J. Biol. Chem.* **278**, 20507–20513 (2003).
53. Wang, W. *et al.* Human H-Y: a male-specific histocompatibility antigen derived from the SMCY protein. *Science* **269**, 1588–1590 (1995).
54. Scott, D. M. *et al.* Identification of a mouse male-specific transplantation antigen, H-Y. *Nature* **376**, 695–698 (1995).
55. Gildea, J. J., Lopez, R. & Shearn, A. A screen for new trithorax group genes identified little imaginal discs, the *Drosophila melanogaster* homologue of human retinoblastoma binding protein 2. *Genetics* **156**, 645–663 (2000).
56. Jung, J., Mysliwiec, M. R. & Lee, Y. Roles of JUMONJI in mouse embryonic development. *Dev. Dyn.* **232**, 21–32 (2005).
57. Takeuchi, T., Kojima, M., Nakajima, K. & Kondo, S. *Jumonji* gene is essential for the neurulation and cardiac development of mouse embryos with a C3H/He background. *Mech. Dev.* **86**, 29–38 (1999).
58. Jung, J., Kim, T. G., Lyons, G. E., Kim, H. R. & Lee, Y. Jumonji regulates cardiomyocyte proliferation via interaction with retinoblastoma protein. *J. Biol. Chem.* **280**, 30916–30923 (2005).
59. Kim, T. G., Chen, J., Sadoshima, J. & Lee, Y. Jumonji represses atrial natriuretic factor gene expression by inhibiting transcriptional activities of cardiac transcription factors. *Mol. Cell. Biol.* **24**, 10151–10160 (2004).
60. Kim, T. G., Jung, J., Mysliwiec, M. R., Kang, S. & Lee, Y. Jumonji represses alpha-cardiac myosin heavy chain expression via inhibiting MEF2 activity. *Biochem. Biophys. Res. Commun.* **329**, 544–553 (2005).
61. Ringrose, L. & Paro, R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* **38**, 413–443 (2004).
62. Katoh, M. & Katoh, M. Identification and characterization of JMJD2 family genes *in silico*. *Int. J. Oncol.* **24**, 1623–1628 (2004).
63. Yang, Z. Q. *et al.* Identification of a novel gene, *GASC1*, within an amplicon at 9p23–24 frequently detected in esophageal cancer cell lines. *Cancer Res.* **60**, 4735–4739 (2000).
64. Zhang, D., Yoon, H. G. & Wong, J. JMJD2A is a novel *N*–CoR-interacting protein and is involved in repression of the human transcription factor achaete scute-like homologue 2 (ASCL2/Hash2). *Mol. Cell. Biol.* **25**, 6404–6414 (2005).
65. Gray, S. G. *et al.* Functional characterization of JMJD2A, a histone deacetylase- and retinoblastoma-binding protein. *J. Biol. Chem.* **280**, 28507–28518 (2005).
66. Kim, J. *et al.* Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.* **7**, 397–403 (2006).
67. Huang, Y., Fang, J., Bedford, M. T., Zhang, Y. & Xu, R. M. Recognition of hstone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* **312**, 748–751 (2006).
68. Bannister, A. J. *et al.* Spatial distribution of di- and trimethyl lysine 36 of histone H3 at active genes. *J. Biol. Chem.* **280**, 17732–17736 (2005).
69. Jang, Y. K., Wang, L. & Sancar, G. B. RPH1 and GIS1 are damage-responsive repressors of PHR1. *Mol. Cell. Biol.* **19**, 7630–7638 (1999).
70. Greenfield, A. *et al.* An H-YDb epitope is encoded by a novel mouse Y chromosome gene. *Nature Genet.* **14**, 474–478 (1996).
71. Greenfield, A. *et al.* The UTX gene escapes X inactivation in mice and humans. *Hum. Mol. Genet.* **7**, 737–742 (1998).
72. Hu, Z. *et al.* A novel nuclear protein, 5qNCA (LOC51780) is a candidate for the myeloid leukemia tumor suppressor gene on chromosome 5 band q31. *Oncogene* **20**, 6946–6954 (2001).
73. Cachon-Gonzalez, M. B. *et al.* Structure and expression of the hairless gene of mice. *Proc. Natl Acad. Sci.* USA **91**, 7717–7721 (1994).
74. Ahmad, W. *et al.* Alopecia universalis associated with a mutation in the human hairless gene. *Science* **279**, 720–724 (1998).
75. Potter, G. B. *et al.* The hairless gene mutated in congenital hair loss disorders encodes a novel nuclear receptor corepressor. *Genes Dev.* **15**, 2687–2701 (2001).
76. Hoog, C., Schalling, M., Grunder-Brundell, E. & Daneholt, B. Analysis of a murine male germ cell-specific transcript that encodes a putative zinc finger protein. *Mol. Reprod. Dev.* **30**, 173–181 (1991).
77. Lee, J. W., Choi, H. S., Gyuris, J., Brent, R. & Moore, D. D. Two classes of proteins dependent on either the presence or absence of thyroid hormone for interaction with the thyroid hormone receptor. *Mol. Endocrinol.* **9**, 243–254 (1995).
78. Metzen, E. *et al.* Intracellular localisation of human HIF-1α hydroxylases: implications for oxygen sensing. *J. Cell Sci.* **116**, 1319–1326 (2003).
79. Tsuneoka, M., Koda, Y., Soejima, M., Teye, K. & Kimura, H. A novel myc target gene, *mina53*, that is involved in cell proliferation. *J. Biol. Chem.* **277**, 35450–35459 (2002).
80. Zhang, Y. *et al.* The human mineral dust-induced gene, *MDIG*, is a cell growth regulating gene associated with lung cancer. *Oncogene* **24**, 4873–4882 (2005).
81. Eilbracht, J., Kneissel, S., Hofmann, A. & Schmidt-Zachmann, M. S. Protein NO52 — a constitutive nucleolar component sharing high sequence homologies to protein NO66. *Eur. J. Cell Biol.* **84**, 279–294 (2005).
82. Eilbracht, J. *et al.* NO66, a highly conserved dual location protein in the nucleolus and in a special type of synchronously replicating chromatin. *Mol. Biol. Cell* **15**, 1816–1832 (2004).
83. Pickard, R. T., Strifler, B. A., Kramer, R. M. & Sharp, J. D. Molecular cloning of two new human paralogs of 85 kDa cytosolic phospholipase A2. *J. Biol. Chem.* **274**, 8823–8831 (1999).
84. Hirabayashi, T., Murayama, T. & Shimizu, T. Regulatory mechanism and physiological role of cytosolic phospholipase A2. *Biol. Pharm. Bull.* **27**, 1168–1173 (2004).
85. Liu, C., Gilmont, R. R., Benndorf, R. & Welsh, M. J. Identification and characterization of a novel protein from Sertoli cells, PASS1, that associates with mammalian small stress protein hsp27. *J. Biol. Chem.* **275**, 18724–18731 (2000).
86. Bodmer, D., Schepens, M., Eleveld, M. J., Schoenmakers, E. F. & Geurts van Kessel, A. Disruption of a novel gene, *DIRC3*, and expression of DIRC3-HSPBAP1 fusion transcripts in a case of familial renal cell cancer and t(2;3)(q35;q21). *Genes Chromosomes Cancer* **38**, 107–116 (2003).
87. Fadok, V. A. *et al.* A receptor for phosphatidylserine-specific clearance of apoptotic cells. *Nature* **405**, 85–90 (2000).
88. Kunisaki, Y. *et al.* Defective fetal liver erythropoiesis and T lymphopoiesis in mice lacking the phosphatidylserine receptor. *Blood* **103**, 3362–3364 (2004).
89. Bose, J. *et al.* The phosphatidylserine receptor has essential functions during embryogenesis but not in apoptotic cell removal. *J. Biol.* **3**, 15 (2004).
90. Cui, P., Qin, B., Liu, N., Pan, G. & Pei, D. Nuclear localization of the phosphatidylserine receptor protein via multiple nuclear localization signals. *Exp. Cell Res.* **293**, 154–163 (2004).
91. Cikala, M. *et al.* The phosphatidylserine receptor from *Hydra* is a nuclear protein with potential Fe(II) dependent oxygenase activity. *BMC Cell Biol.* **5**, 26 (2004).
92. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599 (2000).
93. Schotta, G. *et al.* A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **18**, 1251–1262 (2004).
94. Sun, X. J. *et al.* Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J. Biol. Chem.* **280**, 35261–35271 (2005).
95. O'Carroll, D. *et al.* The polycomb-group gene *Ezh2* is required for early mouse development. *Mol. Cell. Biol.* **21**, 4330–4336 (2001).
96. Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006).
97. Letunic, I. *et al.* SMART 5: domains in the context of genomes and networks. *Nucleic Acids Res.* **34**, D257–D260 (2006).
98. Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl Acad. Sci. USA* **95**, 5857–5864 (1998).
99. Bateman, A. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **32**, D138–D141 (2004).
100. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881–10890 (1988).
101. Kumar, S., Tamura, K. & Nei, M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform.* **5**, 150–163 (2004).

Acknowledgements

R.J.K. is funded by the Canadian Institutes of Health Research. Work in the Zhang laboratory is funded by grants from the US National Institutes of Health and the Howard Hughes Medical Institute.

Competing interests statement

The authors declare no competing financial interests.

DATABASES

The following terms in this article are linked online to:

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
Lid
OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
X linked mental retardation (XLMR)
UniProtKB: http://ca.expasy.org/sprot
Epe1 | FIH | HP1 | HR | HSPBAP1 | JARID1A | JARID1B | JARID1C | JARID1D | JHDM1A | JHDM1B | JHDM2A | JHDM2B | JHDM2C | JHDM3 | JMJD3 | LSD1 | MINA53 | NO66 | PADI4 | PHF2 | PHF8 | PLA2GB4B | PTDSR

FURTHER INFORMATION

ClustalW: http://www.ebi.ac.uk/clustalw
Mega 3.1: http://www.megasoftware.net
MrBayes: http://mrbayes.csit.fsu.edu
Multalin: http://prodes.toulouse.inra.fr/multalin/multalin.html
PFAM: http://pfam.wustl.edu
SMART: http://smart.embl-heidelberg.de
The Zhang laboratory homepage:
http://www.med.unc.edu/~zhangyi/lab.htm

SUPPLEMENTARY INFORMATION

See online article: S1 (table) | S2 (figure)
Access to this links box is available online.