Journal of Integrative Plant Biology 2008, **50** (7): 886–896

Comparative Analysis of JmjC Domain-containing Proteins Reveals the Potential Histone Demethylases in *Arabidopsis* and Rice

Falong Lu[1,2†], Guanglin Li[1†], Xia Cui[1†], Chunyan Liu[1], Xiu-Jie Wang[1] and Xiaofeng Cao[1*]

([1]*State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, the Chinese Academy of Sciences*, Beijing 100101, China;
[2]*Graduate University, the Chinese Academy of Sciences*, Beijing 100039, China)

Abstract

Histone methylation homeostasis is achieved by controlling the balance between methylation and demethylation to maintain chromatin function and developmental regulation. In animals, a conserved Jumonji C (JmjC) domain was found in a large group of histone demethylases. However, it is still unclear whether plants also contain the JmjC domain-containing active histone demethylases. Here we performed genome-wide screen and phylogenetic analysis of JmjC domain-containing proteins in the dicot plant, *Arabidopsis*, and monocot plant rice, and found 21 and 20 JmjC domain-containing, respectively. We also examined the expression of JmjC domain-containing proteins and compared them to human JmjC counterparts for potential enzymatic activity. The spatial expression patterns of the *Arabidopsis* JmjC domain-containing genes revealed that they are all actively transcribed genes. These active plant JmjC domain-containing genes could possibly function in epigenetic regulation to antagonize the activity of the large number of putative SET domain-containing histone methyltransferase activity to dynamically regulate histone methylation homeostasis.

Key words: At JMJ; demethylase; demethylation; histone; Jumonji C; methylation; Os JMJ.

Lu F, Li G, Cui X, Liu C, Wang XJ, Cao X (2008). Comparative analysis of JmjC domain-containing proteins reveals the potential histone demethylases in *Arabidopsis* and rice. *J. Integr. Plant Biol.* **50**(7), 886–896.

Available online at www.jipb.net

In eukaryotic cells, the N-terminal tails of histones are subjected to various post-translational modifications which play vital roles in chromatin functions, such as transcriptional regulation, chromatin remodeling, DNA repair and cell cycle regulation (Strahl and Allis 2000; Jenuwein and Allis 2001; Kouzarides 2007). Histone methylation is one of the most complex modifications in which one, two or three methyl groups could be added to a lysine. Each of these three types of methylation has a different unique biological function, and methylation at different sites is also found to play distinct roles. In animals, histone H3K4, K36 and K79 methylation are found to predominantly associate with actively transcribing genes, whereas H3K9, H3K27 and H4K20 methylation are related to silent regions (Martin and Zhang 2005). In addition, arginine residues can also be mono-, asymmetric or symmetric dimethylated by type I or type II protein arginine methyltransferases (Zhang and Reinberg 2001).

Due to the vital roles of methylation in transcription and other cellular function regulations, the methylation status at local chromatins should be tightly controlled. In contrast to histone acetylation which is known to be dynamically regulated by histone acetyltransferases and deacetylases, histone methylation has long been regarded as an irreversible process until a breakthrough discovery which demonstrated that lysine demethylase 1 (KDM1), also known as lysine-specific demethylase1 (LSD1), confers the enzymatic activity to remove the methyl group from mono- and dimethylated but not trimethylated lysine 4 of histone H3 (H3K4) (Shi et al. 2004). Compared with the large number of SET domain-containing histone methyltransferases, there are much fewer known genes belonging to the same

Received 3 Apr. 2008 Accepted 14 Apr. 2008

Supported by the Hi-Tech Research and Development (863) Program of China (2006AA10A101), the State Key Basic Research and Development Plan of China (2005CB522400), the National Natural Science Foundation of China (30771209 to C. Liu and 30621001 to X. Cao) and the Chinese Academy of Sciences (CXTD-S2005-2) to X. Cao.

†These authors contributed equally to this work.

*Author for correspondence.

Tel: +86 10 6486 9203;

Fax: +86 10 6487 3428;

E-mail: <xfcao@genetics.ac.cn>.

doi: 10.1111/j.1744-7909.2008.00692.x

family as KDM1/LSD1, therefore there might exist other types of histone demethylases other than KDM1/LSD1 to keep the balance of methylation homeostasis. Shortly after the identification of KDM1/LSD1, a large family of Jumonji C (JmjC) domain-containing proteins have been discovered as histone demethylases (Tsukada et al. 2006; Yamane et al. 2006). Biochemical analysis revealed that both KDM1/LSD1 and JmjC domain-containing proteins directly reverse histone methylation by oxidative demethylation reaction, however, KDM1/LSD1 requires flavin whereas JmjC domain proteins require Fe(II) and α-ketoglutarate (αKG) as cofactors. Bioinformatic analysis uncovers 30 JmjC domain-containing proteins in human, which could be classified into seven groups (Klose et al. 2006a). Biochemical evidence from a number of studies indicates that KDMs confer substrate specificity, such as KDM2A/JHDM1A for H3K36me1/2, KDM5/JARID1 for H3K4me1/2/3, KDM4/JHDM3 for H3K9me2/3 and H3K36me2/3, KDM6 for H3K27me2/3, and KDM3/JHDM2 for H3K9me1/2 (Klose et al. 2006a; Allis et al. 2007). In addition, JMJD6 is shown to be a histone arginine demethylase that targets methylated histone H3R2 and H4R3 (Chang et al. 2007). These histone demethylases have been shown to be involved in transcriptional activation/repression, transcription elongation and genome integrity. The importance of lysine demethylation is illustrated by impaired KDM3A/JHDM2a, KDM4C/JHDM3C and KDM5C/JARID1C causing abnormal spermatogenesis, carcinogenesis and X-linked mental retardation, respectively (Cloos et al. 2006; Iwase et al. 2007; Okada et al. 2007).

Except for H3K79 and H4K20 methylation, both lysine and arginine methylation is relatively conserved between plants and animals (Johnson et al. 2004; Zhang et al. 2007). In *Arabidopsis* and rice, there are 39 and 38 SET domain-containing proteins present in each genome, some of which confer histone lysine methyltransferase activity (Baumbusch et al. 2001; Liang et al. 2003; Ding et al. 2007a, 2007b). In addition, four histone arginine methyltransferases and five putative histone arginine methyltransferases are found in *Arabidopsis* (Niu et al. 2007; Pei et al. 2007; Yan et al. 2007). In plants, histone methylation is important for chromatin stability, normal development and cellular memory. For example, SET domain-containing protein KYP/SUVH4 and SDG714, two histone H3K9 methyltransferases in *Arabidopsis* and rice, respectively, are essential for maintaining DNA methylation in the CNG context for chromatin stability (Jackson et al. 2002; Ding et al. 2007b). ATX1, another SET domain-containing histone H3K4 methyltransferase, is essential for floral organ development by activating homeotic genes (Alvarez-Venegas et al. 2003). EFS/SDG8 (Early Flowering in Short days/SET Domain Group 8), a histone lysine methyltransferase, is needed to prevent plant from early flowering by modifying the chromatin of *Flowering Locus C* (*FLC*), a key flowering time repressor in *Arabidopsis* (Kim et al. 2005; Zhao et al. 2005; Xu et al. 2007). AtPRMT10 and AtPRMT5, two histone arginine methyltransferases, are involved in *FLC*-dependent flowering time control (Niu et al. 2007; Pei et al. 2007; Wang et al. 2007; Schmitz et al. 2008).

Histone methylation is a conserved mechanism between animals and plants. Lines of evidence suggest that histone demethylases play important roles in plants as their human counterparts in homeostasis of histone methylation. Flowering locus D (FLD) and its homologs, homologs of human KDM1/LSD1, are known to repress *FLC* expression possibly through demethylation of H3K4 (He et al. 2003; Jiang et al. 2007). The emerging importance of JmjC proteins in *Arabidopsis* has been shown by the identification of developmental abnormalities associated with the JmjC loss-of-function mutants. For example, plants impaired of *Early Flowering 6* (*At ELF6*/*JMJ11*) and *Relative of Early Flowering 6* (*At REF6*/*JMJ12*) are early and late flowering, respectively (Noh et al. 2004); *Maternal Effect Embryo arrest 27* (*At MEE27*/*JMJ15*) is involved in gametocyte development (Pagnussat et al. 2005) and *Increase in Bonsai Methylation 1* (At *IBM1*/*JMJ25*), which is related to the KDM3/JHDM2 group, is essential to prevent the spreading of silencing from heterochromatin to active genes (Saze et al. 2008). In spite of the important roles of JmjC proteins in developmental regulation in *Arabidopsis*, no biochemical evidence has shown that these proteins are active histone demethylases. In addition, rice is one of the most important crops and is a model species for monocots. However, no JmjC proteins have been reported in rice and the importance of lysine or arginine demethylation remains elusive.

Herein, we carry out a systematic JmjC domain-containing proteins identification in *Arabidopsis* and rice by phylogenetic analysis. As the structural analysis of KDM4A/JHDM3A shows that the activity and specificity of JmjC proteins are determined by the JmjC domain, we predict whether these proteins are active histone demethylase by focusing on the conserved Fe(II) and αKG binding sites within the JmjC domain. Furthermore, we confirm that these JmjC domain-containing genes are actively transcribed by analyzing spatial expression of them in *Arabidopsis*. We propose that these *Arabidopsis* and rice proteins are conserved and important for histone methylation homeostasis and normal plant development.

Results

Domain architecture and evolution of all JmjC domain-containing proteins in plants

The JmjC domain-containing histone demethylases share a highly conserved JmjC domain and comprise a large gene family in plants. After intensive bioinformatics analysis, we obtained a non-redundant set of 71 JmjC domain-containing proteins, including 21 proteins from *Arabidopsis thaliana*, 20 from *Oryza sativa* and 30 from human. Based on the JmjC

domain sequences, we constructed a phylogenetic tree of human, *A. thaliana* and *O. sativa* JmjC domain-containing proteins (Figure 1). By merging information between the phylogenetic tree and the domain architecture of each protein, we classified the JmjC domain-containing proteins into eight groups: KDM6/JMJD3 group, KDM5/JARID1 group, KDM4/JHDM3 group, KDM3/JHDM2 group, KDM2/JHDM1 group, PHF group, JMJD6 group, and JmjC domain-only group. The JmjC domain-only group includes proteins with only JmjC domain but no other domains and not belonging to any other groups. It is worth noting that proteins belonging to the KDM6/JMJD3 group, KDM2/JHDM1 group and PHF group are missing in both *Arabidopsis* and rice.

To further analyze the JmjC domain-containing proteins in *Arabidopsis* and rice, we focused on domain character and amino acid conservation within the predicted Fe(II) and αKG binding sites so that further information for potential active demethylase activities can be provided. In the following sections we will discuss potential enzymatic activity, target specificity on histone methylation and biological function of each JmjC domain-containing group, in order to provide information for investigating JmjC domain functions both in *Arabidopsis* and rice.

KDM5/JARID1 group

Human KDM5/JARID1 group proteins are found to be active histone demethylases using H3K4me3, H3K4me2 and H3K4me1 as substrates (Christensen et al. 2007; Iwase et al. 2007; Klose et al. 2007; Lee et al. 2007; Xiang et al. 2007b). In higher plants, six *Arabidopsis* and three rice proteins were identified as members of this group by phylogenetic analysis. Of these nine members, two subgroups are divided according to their domain architecture. At JMJ17 and Os JMJ708 share the same domain architecture with their human counterparts, including the domains JmjN, ARID, JmjC, zf-C5HC2 and PHD. The other seven proteins show extensive similarity to human KDM5/JARID1 family proteins in their JmjC domain. Instead of zf-C5HC2 and PHD domains, six of these proteins have FYRN and FYRC domains at the C-terminal. It has been shown that the PHD domain in human KDM5/JARID1C preferentially binds to H3K9me3. Interestingly, FYRN and FYRC domains are normally found in trithorax and its homologs, a group of histone H3K4 methyltransferases (Finn et al. 2006). It is possible that FYRN and FYRC domains harbor chromatin-binding activity and target these enzymes to some specific chromatin regions other than the PHD targeting site, H3K9me3. Note that At JMJ19 is a relatively short protein that lacks the C-terminal part of FYRN and FYRC domains compared with other proteins.

Crystal structure of the KDM4A/JHDM3A catalytic core sequence shows that Fe(II) is chelated by three absolutely conserved residues (His188, Glu190 and His276) within the JmjC domain. Two additional residues (Thr185 and Lys206) are required for αKG binding (Chen et al. 2006; Klose et al. 2006a). Substitutions of the Glu that binds to Fe(II) to Asp, the Thr that binds to αKG to Tyr or Phe and the Lys that binds to αKG to Arg are compatible with histone demethylation activity (Klose et al. 2006a; Agger et al. 2007; De Santa et al. 2007; Hong et al. 2007; Lan et al. 2007; Xiang et al. 2007a). Successful prediction of demethylation activity of JmjC domain-containing proteins is achieved based on these features (Klose et al. 2006a). Therefore, we set out to predict whether *Arabidopsis* and rice JmjC domain-containing proteins have the demethylation activity according to the conservation of these Fe(II) and αKG binding amino acids. All the proteins but At JMJ19 and Os JMJ708 in KDM5/JARID1 group have conserved amino acids of His(H), Glu(E) and His(H) for Fe(II), and Phe(F) and Lys(K) for αKG binding within the cofactor binding sites (Figure 2), suggesting that they are likely to be active histone demethylases. Based on the highest similarity and domain structure, At JMJ17 demethylates histone H3K4 as its human counterparts. Further, we propose that Os JMJ703, At JMJ16, Os JMJ704, At JMJ14, At JMJ18 and At MEE27/JMJ15 are potential active histone H3K4 demethylases that target a different subset of targets. Further investigation of these proteins may extend our knowledge about roles of histone H3K4 methylation in plant development.

KDM4/JHDM3 group

The KDM4/JHDM3 group contains four members in human, three members in *Arabidopsis* and five members in rice. These proteins share the same features of JmjN and JmjC domains at their N-termini but have different C-terminal domains. Human KDM4A/JHDM3A, KDM4B/JHDM3B and KDM4C/JHDM3C have PHD and Tudor domains at the C-termini. Tudor domain of KDM4A/JHDM3A has been shown to bind H3K4me3, H3K9me3, H4K20me3 and H3K20me2 (Kim et al. 2006). Proteins of this group have two kinds of C-terminal domains in *Arabidopsis* and rice. One is C5HC2 type zinc finger (zf-C5HC2) in At JMJ13, Os JMJ706 and Os JMJ707. The other is four tandem repeats of C2H2 type zinc finger ($4\times$zf-C2H2) in At REF6/JMJ12, Os JMJ705, Os JMJ702, At ELF6/JMJ11 and Os JMJ701. Zf-C5HC2 is a putative DNA binding domain. The $4\times$zf-C2H2 domain in At REF6/JMJ12 and At ELF6/JMJ11 could be well modeled to $5\times$zf-C2H2 found in glioma-associated oncogene homolog 1 (GLI1) protein which has been demonstrated to have a specific DNA binding activity (Pavletich and Pabo 1993). Therefore, plant proteins in this group might bind directly to a specific DNA sequence, whereas human KDM4/JHDM3 proteins might bind to a specific chromatin region by interacting with some specific combination of modified histones. This indicates that these At JMJ and Os JMJ proteins may define a new type of transcriptional repressor or activator but not co-repressor or co-activator.

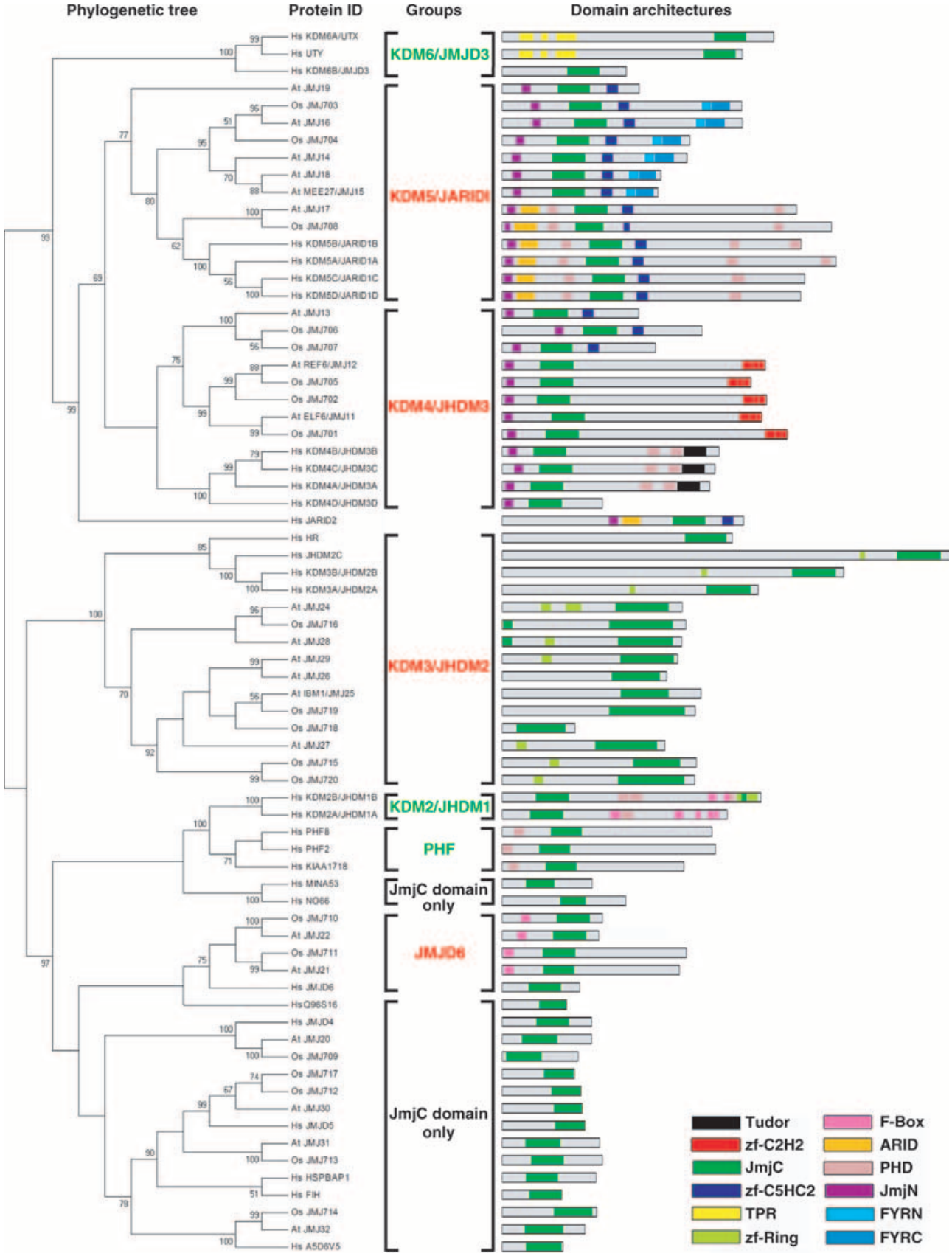

Figure 2. KDM5/JARID1 group proteins contain potential H3K4 demethylases in *Arabidopsis* and rice.

Cofactor binding sites within Jumonji C (JmjC) domain of KDM5/JARID1 group proteins from human, *Arabidopsis* and rice were aligned. The conserved residues compatible with the demethylation activity within the Fe(II) binding site are highlighted in red and those in the αKG binding site are indicated in blue. Proteins from *Homo sapiens*, *Arabidopsis thaliana* and *Oryza sativa* are indicated with Hs, At and Os prefix, respectively.

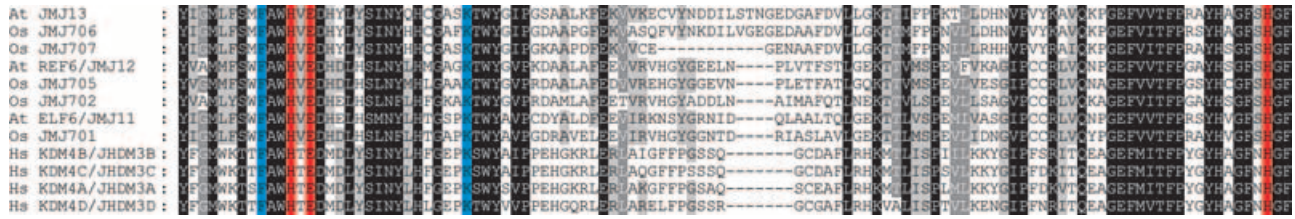

Figure 3. KDM4/JHDM3 group proteins are potential active histone demethylases in *Arabidopsis* and rice.

Cofactor binding sites within Jumonji C (JmjC) domain of KDM4/JHDM3 group proteins from human, *Arabidopsis* and rice were aligned. The conserved residues compatible with the demethylation activity within the Fe(II) binding site are highlighted in red and those in the αKG binding site are highlighted in blue. Proteins from *Homo sapiens*, *Arabidopsis thaliana* and *Oryza sativa* are indicated with Hs, At and Os prefix, respectively.

Human KDM4/JHDM3 family proteins are shown to actively demethylate histone H3K9me2/3 and H3K36me2/3 (Cloos et al. 2006; Klose et al. 2006b; Whetstine et al. 2006). All the *Arabidopsis* and rice proteins of this group have conserved Fe(II) and αKG binding amino acids within the cofactor binding site (Figure 3), suggesting that these proteins are likely to be active histone demethylases, and are likely to target tri- and dimethylated H3K9 and H3K36. Although there is a high sequence similarity between At ELF6/JMJ11 and At REF6/JMJ12, they were found to participate in different genetic pathways in flowering time regulation. At REF6/JMJ12 was shown to control flowering time through regulating acetylation status of *FLC* in the autonomous pathway, whereas At ELF6/JMJ11 functioned in the photoperiod pathway (Noh et al. 2004). In particular, in the *at ref6/jmj12* mutant, the late flowering phenotype was shown to be associated with increased histone

←

Figure 1. Phylogenetic tree of 71 Jumonji C (JmjC) domain-containing proteins based on JmjC domain sequences.

The neighbor-joining phylogenetic tree was constructed by using Poisson correction distance for amino acid sequences in the MEGA package. Percentage bootstrap values of 1 000 replications ($>$ 50%) are shown on internal branches. Proteins from human, *Arabidopsis thaliana* and *Oryza sativa* are indicated with Hs, At and Os prefix, respectively. Groups in red indicate that they are conserved among human and *Arabidopsis* and rice, groups in green indicate those are not conserved and group in black is JMJC domain-only proteins not classified in the above groups and not classified further. The domain architecture for each protein is represented by band chart. Different colors are used for different domains and are indicated at the bottom right of the figure. Protein database IDs for At JMJ and Os JMJ proteins are: At ELF6/JMJ11 (Q6BDA0), At REF6/JMJ12 (Q9STM3), At JMJ13 (Q9FJS0), At JMJ14 (Q8GUI6), At MEE27/JMJ15 (O64752), At JMJ16 (Q9FRS3), At JMJ17 (Q9SH34), At JMJ18 (Q8W4M0), At JMJ19 (Q8L7T6), At JMJ20 (Q67ZB6), At JMJ21 (Q9M9E8), At JMJ22 (Q67XX3, Q9FG15), At JMJ24 (Q27GN5), At IBM1/JMJ25 (Q9SSE9), At JMJ26 (O65384), At JMJ27 (Q8VYB9), At JMJ28 (Q8H1S7), At JMJ29 (O48794), At JMJ30 (Q8RWR1), At JMJ31 (IPI00518924), At JMJ32 (Q9LZU2), Os JMJ701 (A3AE68), Os JMJ702 (Q2QTX9), Os JMJ703 (Q53WJ1), Os JMJ704 (Q0DJC2), Os JMJ705 (Q5N712), Os JMJ706 (Q8W3G5), Os JMJ707 (A3AAD3), Os JMJ708 (A3BFN7), Os JMJ709 (Q5JKD1), Os JMJ710 (Q2R2A5), Os JMJ711 (Q75LR4), Os JMJ712 (A3C049), Os JMJ713 (Q5ZC07), Os JMJ714 (A3C030), Os JMJ715 (Q6AUV7), Os JMJ716 (Q0DRY1), Os JMJ717 (Q6YVS8), Os JMJ718 (Q6H405), Os JMJ719 (Q0E4N0) and Os JMJ720 (Q6K7P0).

acetylation level in the *FLC* locus. It is possible that the change of histone acetylation level in the *FLC* locus is caused by deregulation of histone methylation. The roles that At REF6/JMJ12 play in histone demethylation are waiting to be elucidated. Further, this may give us new insight into chromatin regulation of the *FLC* locus that determines flowering time.

KDM3/JHDM2 group

The KDM3/JHDM2 group contains four members in human, six members in *Arabidopsis* and five members in rice, which have a JmjC domain at the C-terminal. Some of the members have a putative Ring type zinc finger (zf-Ring) domain ahead of it. The zf-Ring domain is shown to be essential for the demethylation activity of KDM3A/JHDM2A (Yamane et al. 2006). KDM3A/JHDM2A can demethylate H3K9me2 and H3K9me1 but not H3K9me3. When the JmjC domain sequence of these groups were aligned together, we found that At JMJ29, At JMJ26, At IBM1/JMJ25, Os JMJ719, Os JMJ718 and At JMJ27 have the conserved Fe(II) and αKG binding amino acids within the cofactor binding site (Figure 4), but At JMJ24, Os JMJ716, At JMJ28, Os JMJ715 and Os JMJ720 do not. It is likely that At JMJ29, At JMJ26, At IBM1/JMJ25, Os JMJ719, Os JMJ718 and At JMJ27 are active histone demethylases targeting H3K9me2 and H3K9me1. A recent report showed that mutation in At IBM1/JMJ25 resulted in DNA methylation spreading from a transposon, LINE, to nearby coding gene, *Bonsai*, and showed pleiotropic developmental phenotypes (Saze et al. 2008). Moreover, the pleiotropic phenotypes can be suppressed by mutation in a histone H3K9 histone methyltransferase KYP/SUVH4. Taken together, these clues strongly suggest that At IBM1/JMJ25 is a histone H3K9 demethylase. However, direct biochemical evidence supporting this hypothesis is still to be determined. It is worth noting that mutation in At IBM1/JMJ25 does not affect DNA methylation at a global level. This also suggests that other members within this group should participate in controlling H3K9 methylation and DNA methylation.

JMJD6 group

The JMJD6 group contains one protein in human, two in *Arabidopsis* and two in rice. Human JMJD6 has no other obvious protein domain besides the JmjC domain and is shown to be histone arginine demethylase that can demethylate H3R2me2 and H4R3me2 (Chang et al. 2007). *Arabidopsis* and rice homologs of JMJD6 have an additional F-Box domain in the N-terminal. The F-Box domain defines some ubiquitin E3 ligase activity, suggesting a link between protein demethylation and ubiquitination. The F-Box domain is also found in human JHDM1A and JHDM1B. However, the function of links between these two distinct enzymatic activities remains to be defined.

The JmjC domain of *Arabidopsis* and rice proteins of this group differs from that of human JMJD6 because the Thr residue of the first αKG binding site is substituted with Ala or Ser (arrowhead in Figure 5). Whether such substitution affects their potential αKG binding activity is still unknown. However, in all JmjC domain proteins analyzed so far, there is a bias to Thr (Klose et al. 2006a).

JmjC domain-only group

There are eight JmjC domain-only proteins in human, four in *Arabidopsis* and five in rice without other recognizable protein domains except for JmjC and JmjN domains. All *Arabidopsis* and rice JmjC domain-only proteins except At JMJ31 and Os JMJ713 have conserved Fe(II) and αKG binding amino acids within the cofactor binding site (Figure 6), indicating that these proteins might be active demethylases. However, the substrate preference is unknown. Further investigation on these enzymes might reveal that these proteins have distinct enzymatic activity.

Spatial expression of *Arabidopsis* JmjC domain-containing proteins

To determine whether these genes are active in *Arabidopsis*, we used reverse transcription polymerase chain reaction (RT-PCR) to test their spatial expression in different tissues, including cultured cell, rosette leaf, stem, inflorescence, silique and root. Because at least one primer of each primer pair used in the RT-PCR spans two exons, only spliced cDNA, but not genomic DNA, can be amplified. The results showed that all these 21 genes were expressed under normal growth conditions, indicating that all these genes are active expressing genes (Figure 7). Most of these genes are ubiquitously expressed with a slightly higher expression in rosette leaves and inflorescences. *At MEE27*/*JMJ15* is mainly expressed in siliques and *At JMJ16* is mainly expressed in inflorescences, indicating that they may confer specialized function at reproductive developmental stage.

Discussion

JmjC domain proteins represent a large family of histone demethylases in both animals and plants. In the present study, we found 21 JmjC domain-containing proteins in *Arabidopsis* and 20 in rice. We analyzed the domain structure of these 41 proteins and compared them with human counterparts. Also, by analyzing the potential Fe(II) and αKG binding site, we predict that 15 *Arabidopsis* and 13 rice JmjC domain-containing proteins are very likely to confer histone demethylase activity. In addition, we demonstrated that the *Arabidopsis* JmjC domain-containing genes are active genes by analyzing their spatial expression patterns.

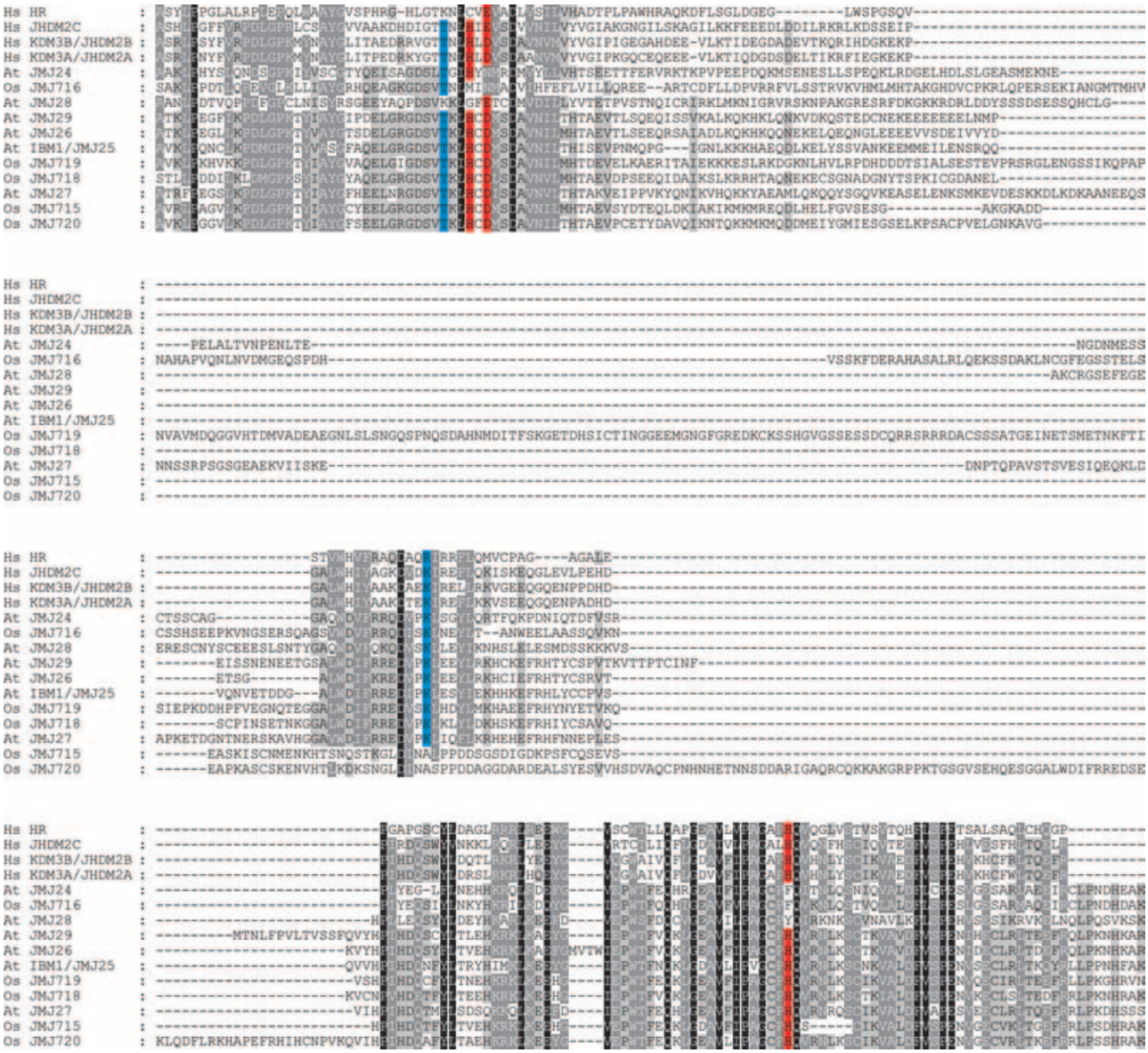

Figure 4. KDM3/JHDM2 group proteins contain potential H3K9 demethylases in *Arabidopsis* and rice.

Cofactor binding sites within the Jumonji C (JmjC) domain of KDM3/JHDM2 group proteins from human, *Arabidopsis* and rice were aligned. The conserved residues compatible with the demethylation activity within the Fe(II) binding site are highlighted in red and those in the αKG binding site are indicated in blue. Proteins from *Homo sapiens*, *Arabidopsis thaliana* and *Oryza sativa* are indicated with Hs, At and Os prefix, respectively.

However, plants lack some groups of proteins found in human including the KDM6/JMJD3, KDM2/JHDM1 and PHF groups. KDM2/JHDM1 is the first identified histone H3K36me2 and H3K36me1 demethylase (Tsukada et al. 2006). Given that histone H3K36 methylation is found to be highly abundant and plays an important role in flowering time regulation (Johnson et al. 2004; Zhao et al. 2005; Zhang et al. 2007), we believe that the H3K36 methylation level should be tightly controlled by making a good balance between methylation and demethylation. The KDM4/JHDM3 group proteins might account for part of the demethylation of methylated H3K36. However, KDM4/JHDM3 group proteins in human predominantly demethylate H3K36me3 and weakly demethylate H3K36me2. One possibility is that *Arabidopsis* and rice KDM4/JHDM3 family proteins are also capable of demethylating H3K36me1. The other possibility is that some other JmjC domain proteins demethylate H3K36me1 in *Arabidopsis* and rice. The molecular function of the PHF group is still unknown. We noticed that they showed great similarity

Figure 5. *Arabidopsis* and rice JMJD6 group proteins contain amino acid substitution in the first αKG binding site.

Cofactor binding sites within the Jumonji C (JmjC) domain of JMJD6 group proteins from human, *Arabidopsis* and rice were aligned. The conserved residues compatible with the demethylation activity within the Fe(II) binding site are highlighted in red and those in the αKG binding site are indicated in blue. Proteins from *Homo sapiens*, *Arabidopsis thaliana* and *Oryza sativa* are indicated with Hs, At and Os prefix, respectively. The first conserved αKG binding residue is substituted with Ala or Ser in *Arabidopsis* and rice proteins within this group and is indicated by arrowhead.

Figure 6. *Arabidopsis* and rice Jumonji C (JmjC) domain-only proteins contain some potential active histone demethylases.

Cofactor binding sites within the JmjC domain of JmjC domain-only group proteins from human, *Arabidopsis* and rice were aligned. The conserved residues compatible with the demethylation activity within the Fe(II) binding site are highlighted in red and those in the αKG binding site are indicated in blue. Proteins from *Homo sapiens*, *Arabidopsis thaliana* and *Oryza sativa* are indicated with Hs, At and Os prefix, respectively.

to KDM2/JHDM1 group protein in their JmjC domain, probably because they have similar enzymatic activity but target different chromatin contents. The KDM6/JMJD3 group contains three members in human, among which Hs KDM6A/UTX and Hs KDM6B/JMJD3 have been shown to demethylate H3K27me3 and H3K27me2 (Swigut and Wysocka 2007). Histone H3K27 methylation is highly conserved between animals and plants and is required for the silencing of homeobox transcription factors that are important for proper patterning during development. H3K27 methylation level is developmentally regulated. For example, *AGAMOUS* (*AG*), a floral homeotic gene required for stamen and carpel identity, is silenced in vegetative tissue by histone H3K27 methylation and is activated in the center of floral primordia prior to the initiation of stamen and carpel primordia (Schubert et al. 2006; Ito et al. 2007). Because H3K27 demethylation occurs during cell differentiation, there should be some JmjC domain-containing proteins responsible for active H3K27 demethylase activity in plants.

Histone methylation, which is catalyzed by a large group of SET domain proteins or protein arginine methyltransferases, is known to play a great role in plant development (Ng et al. 2007). Roles of JmjC domain-containing proteins in plant development are just emerging. Until now, it has been known that At ELF6/JMJ11 and At REF6/JMJ12 are required for proper flowering timing (Noh et al. 2004), that At MEE27/JMJ15 is involved in gametocyte development (Pagnussat et al. 2005) and that mutation in At IBM1/JMJ25 causes pleiotropic phenotypes (Saze et al. 2008). However, how these proteins control histone methylation, and how histone methylation regulated by these proteins controls development, is still unknown. Further study of the roles these proteins play in development and the biochemical mechanism underlining these, in conjugation with studying histone methyltransferases in development, shall further extend our knowledge about how histone methylation programs plant development.

Materials and Methods

Sequences of JmjC domain-containing proteins

The amino acid sequences of all JmjC domain-containing proteins from human, *Arabidopsis thaliana* and *Oryza sativa*, were downloaded from UniProt (http://www.ebi.ac.uk/uniprot/), InterPro (http://www.ebi.ac.uk/interpro), SMART (http://smart.embl-heidelberg.de) and TIGR (http://www.tigr.org) databases. Unreported JmjC domain-containing proteins were identified by comparing known JmjC domain-containing proteins with the Refseq

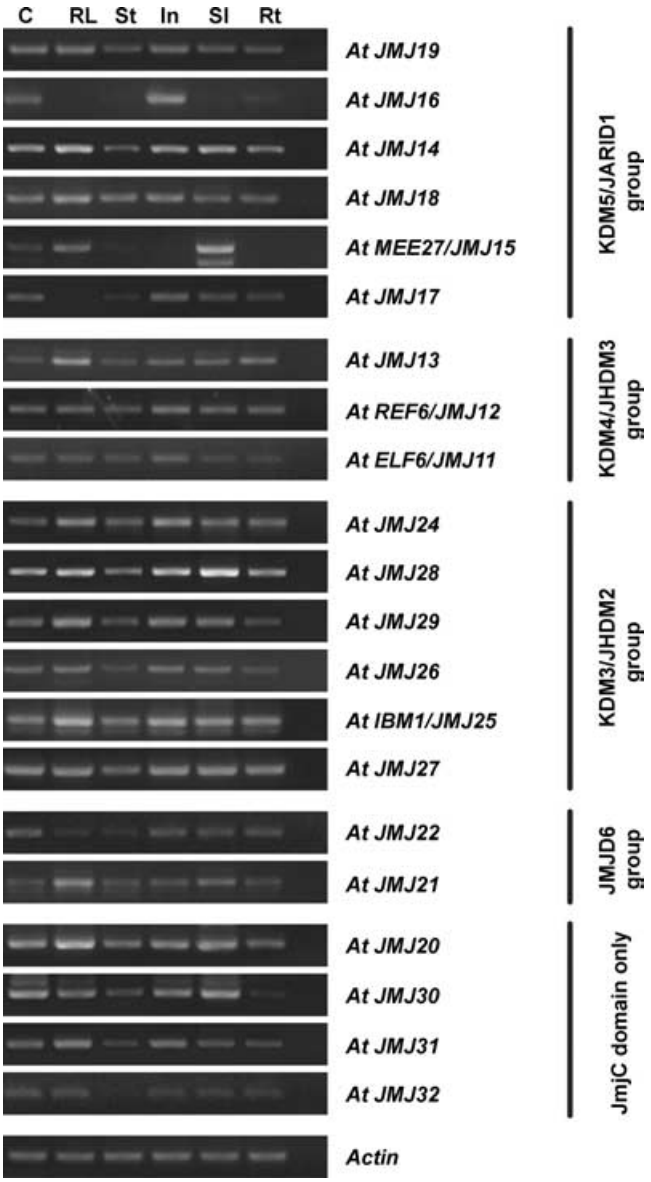

Figure 7. Spatial expression pattern of Jumonji C (JmjC) domain-containing genes in *Arabidopsis*.

Genes are arranged as the groups discussed above and this is also indicated on the right. *Actin* is used as an internal control that indicates equal amount cDNA is used. C, callus; In, inflorescence; RL, rosette leaf; Rt, root; SI, silique; St, stem.

protein sequences of the above species using the BLASTP program. After removing redundant and variant sequences, we included 30 human proteins, 21 *A. thaliana* proteins, and 20 *O. sativa* proteins in our JmjC domain-containing protein database. The denominations of JmjC domain-containing proteins used in this study were according to the criteria in ChromDB (http://chromdb.org/).

Protein domain analysis

All the domains in the 71 JmjC domain-containing proteins were analyzed using the hidden Markov model (HMM) algorithm in SMART and Pfam (http://pfam.wustl.edu). After removing the redundant and/or low score (e>1.0E-10) domains, we obtained a full domain architecture of all the JmjC domain-containing proteins by using Perl scripts.

Sequence alignment and phylogenetic tree construction

A multiple sequence alignment of all the JmjC domains was performed using ClustalW (Thompson et al. 1994) with default parameters, then the phylogenetic tree of all the JmjC domain-containing proteins was constructed and visualized by software MEGA (ver.3.1) using the neighbor-joining method (Kumar et al. 2004). The aligned sequences were visualized using GeneDoc software.

Expression of JmjC domain-containing proteins

Arabidopsis thaliana wild-type Columbia (Col) plants were grown at 23 °C under long day conditions (light:dark cycle, 16:8 h). Rosette leaf, stem, inflorescence and silique tissues were collected from 6 week-old plants grown in soil. Root was collected from 13-d-old plants planted in MS plate containing 3% sucrose. Total RNAs were extracted using Trizol reagent and were reverse transcribed with MMLV reverse transcriptase according to the manufacture's instruction. Semi-quantitative PCR was performed by amplifying these samples for appropriate cycles ranging 24–32 before reaching the stationary phase. Primers used for RT-PCR are as follows: AT3G48430 (*At REF6*/*JMJ12*) for cx1812 TGTGCCATCCAGAGTACCCCAGGA and cx1813 GGGATGCACCTGATGTGACATCCA; AT5G04240 (*At ELF6*/*JMJ11*) for cx1814 CCTTTGTTTCAAGAGTGCCACGATCA and cx1815 TTTCCTGAGCATCTGTCTCTCCTTGTCTC; AT1G09060 (*At JMJ24*) for cx1822 TATCCAGACTGATTTTGTGTCACGCCC and cx1823 GCTTGGTTGCCTCGTCTAAGTTGTGAGA; AT1G63490 (*At JMJ17*) for cx1842 GCAGCACTTAAAAGAGGGACAAACACTTGA and cx1843 GGCTCTGTTGCTCTGGCGGGA; AT5G06550 (*At JMJ22*) for cx1844 TTGCAGAGAAAGTTCCGGTTTTGGA and cx1845 GAAATTCTAACACATTCAACAAATTACTCCTGCTA; AT4G21430 (*At JMJ28*) for cx1892 AGCCTGATTCTGTGAAGAAGTTAGGTTTTGA and cx1893 TCCATGGTTCAACGTCAAATTCTTCCTTA; AT4G20400 (*At JMJ14*) for cx1894 CTGGCGAAGAAAGTGGATGGTGTTTTA and cx1895 TGGATCAAGAGCCTCAACCACCTTAATTA; AT3G20810 (*At JMJ30*) for cx1896 TGTTCTGATTGGAAGCAAGAGCTTGTGA and cx1897 CGATATTGTCTAGATCAACCTGACTAGAGTTGCA; AT5G63080 (*At JMJ20*) for cx1898 GGTGGGAAAGGATCTTGGACACCTCTA and cx1899 CATTTAAATTCATTCCTGTGTTGGCAGCA; AT5G19840 (*At JMJ31*) for cx1900 TGATAGACAGAGAAATGAGCCTGTTAGTCTCGA and cx1901 CGAAAGAACTGGGTCCTATCTTCTT-

CACCA; AT4G00990 (*At JMJ27*) for cx1902 CGGGCTCAGGAGAAGCTGAGAAAGTAATTA and cx1903 CCACCTTTATGCAAGACTGTCTATTCCTCA; AT1G11950 (*At JMJ26*) for cx1904 GTAACACTAAGTGAGGAGCAAAGGTCTGCA and cx1905 CCTTTGTGCACGACTTGAGATTCCGA; AT1G78280 (*At JMJ21*) for cx1906 ATTTAACCATAGGAGATCCCATCTGCGA and cx1907 CCAAACAGTTTGCTCAACTTGTTCCCA; AT3G07610 (*At IBM1*/*JMJ25*) for cx1908 GTAGCGCTTGACTTCGTCTCACCTGA and cx1909 GCTTCGGAAGACTTCTCTTCACTTGGA; AT2G34880 (*At MEE27*/*JMJ15*) for cx1910 AGATATGTGAGGATGCATCCATCATGGA and cx1911 CCTCAGTGGCCTGAACAATAAACGGA; AT1G62310 (*At JMJ29*) for cx1912 TCAGACGCGGTGAATATTCTAACGCA and cx1913 GACTTTTGTGCACGACTTTAGATTCCGA; AT1G30810 (*At JMJ18*) for cx1914 GAACCTTGGTGTATCGGTCGAGCCTA and cx1915 TTCAATGGCCTGAACTATAGAGGGTGAGA; AT2G38950 (*At JMJ19*) for cx1916 CCAAAGAAGAGTCATAATCCTGTGATGATGA and cx1917 CAGTTTGATCCGTGACTTGATTATGTTGGA; AT5G46910 (*At JMJ13*) for cx1918 CGAAAGAAGACGGATATACACCGTATTCCA and cx1919 TCAGTCTCTTAAGTTCCTGATGATGGTCAGA; AT1G08620 (*At JMJ16*) for cx1920 ACAGTCCATTGTTCATGGTTTACTTGGAGA and cx1921 AAACCTCTTGAAGTGAGCTTAGTTCCTCCA; and AT3G45880 (*At JMJ32*) for cx1922 CCGAGATACTGATGCGTTCAAGCTTGA and cx1923 AACTTGCTTCCCTCTGCATCACTGGA.

References

Agger K, Cloos PA, Christensen J, Pasini D, Rose S, Rappsilber J et al. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in *HOX* gene regulation and development. *Nature* **449**, 731–734.

Allis CD, Berger SL, Cote J, Dent S, Jenuwien T, Kouzarides T et al. (2007). New nomenclature for chromatin-modifying enzymes. *Cell* **131**, 633–636.

Alvarez-Venegas R, Pien S, Sadder M, Witmer X, Grossniklaus U, Avramova Z (2003). ATX-1, an *Arabidopsis* homolog of trithorax, activates flower homeotic genes. *Curr. Biol.* **13**, 627–637.

Baumbusch LO, Thorstensen T, Krauss V, Fischer A, Naumann K, Assalkhou R et al. (2001). The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Res.* **29**, 4319–4333.

Chang B, Chen Y, Zhao Y, Bruick RK (2007). JMJD6 is a histone arginine demethylase. *Science* **318**, 444–447.

Chen Z, Zang J, Whetstine J, Hong X, Davrazou F, Kutateladze TG et al. (2006). Structural insights into histone demethylation by JMJD2 family members. *Cell* **125**, 691–702.

Christensen J, Agger K, Cloos PA, Pasini D, Rose S, Sennels L et al. (2007). RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. *Cell* **128**, 1063–1076.

Cloos PAC, Christensen J, Agger K, Maiolica A, Rappsilber J, Antal T et al. (2006). The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* **442**, 307–311.

De Santa F, Totaro MG, Prosperini E, Notarbartolo S, Testa G, Natoli G (2007). The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* **130**, 1083–1094.

Ding B, Zhu Y, Gao J, Yu Y, Cao K, Shen WH et al. (2007a). Molecular characterization of three rice SET-domain proteins. *Plant Sci.* **172**, 1072–1078.

Ding Y, Wang X, Su L, Zhai J, Cao S, Zhang D et al. (2007b). SDG714, a histone H3K9 methyltransferase, is involved in Tos17 DNA methylation and transposition in rice. *Plant Cell* **19**, 9–22.

Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T et al. (2006). Pfam: clans, web tools and services. *Nucleic Acids Res.* **34**, D247–251.

He Y, Michaels SD, Amasino RM (2003). Regulation of flowering time by histone acetylation in *Arabidopsis*. *Science* **302**, 1751–1754.

Hong S, Cho YW, Yu LR, Yu H, Veenstra TD, Ge K (2007). Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc. Natl. Acad. Sci. USA* **104**, 18439–18444.

Ito T, Ng KH, Lim TS, Yu H, Meyerowitz EM (2007). The homeotic protein AGAMOUS controls late stamen development by regulating a jasmonate biosynthetic gene in *Arabidopsis*. *Plant Cell* **19**, 3516–3529.

Iwase S, Lan F, Bayliss P, de la Torre-Ubieta L, Huarte M, Qi HH et al. (2007). The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell* **128**, 1077–1088.

Jackson JP, Lindroth AM, Cao X, Jacobsen SE (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**, 556–560.

Jenuwein T, Allis CD (2001). Translating the histone code. *Science* **293**, 1074–1080.

Jiang D, Yang W, He Y, Amasino RM (2007). *Arabidopsis* relatives of the human lysine-specific Demethylase1 repress the expression of FWA and FLOWERING LOCUS C and thus promote the floral transition. *Plant Cell* **19**, 2975–2987.

Johnson L, Mollah S, Garcia BA, Muratore TL, Shabanowitz J, Hunt DF et al. (2004). Mass spectrometry analysis of *Arabidopsis* histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Res.* **32**, 6511–6518.

Kim J, Daniel J, Espejo A, Lake A, Krishna M, Xia L et al. (2006). Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.* **7**, 397–403.

Kim SY, He Y, Jacob Y, Noh YS, Michaels S, Amasino R (2005). Establishment of the vernalization-responsive, winter-annual habit in *Arabidopsis* requires a putative histone H3 methyl transferase. *Plant Cell* **17**, 3301–3310.

Klose RJ, Kallin EM, Zhang Y (2006a). JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* **7**, 715–727.

Klose RJ, Yamane K, Bae Y, Zhang D, Erdjument-Bromage H, Tempst P et al. (2006b). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* **442**, 312–316.

Klose RJ, Yan Q, Tothova Z, Yamane K, Erdjument-Bromage H, Tempst P et al. (2007). The retinoblastoma binding protein RBP2 is an H3K4 demethylase. *Cell* **128**, 889–900.

Kouzarides T (2007). Chromatin modifications and their function. *Cell* **128**, 693–705.

Kumar S, Tamura K, Nei M (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinform.* **5**, 150–163.

Lan F, Bayliss PE, Rinn JL, Whetstine JR, Wang JK, Chen S et al. (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**, 689–694.

Lee MG, Norman J, Shilatifard A, Shiekhattar R (2007). Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. *Cell* **128**, 877–887.

Liang YK, Wang Y, Zhang Y, Li SG, Lu XC, Li H et al. (2003). *OsSET1*, a novel SET-domain-containing gene from rice. *J. Exp. Bot.* **54**, 1995–1996.

Martin C, Zhang Y (2005). The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **6**, 838–849.

Ng DW, Wang T, Chandrasekharan MB, Aramayo R, Kertbundit S, Hall TC (2007). Plant SET domain-containing proteins: structure, function and regulation. *Biochim. Biophys. Acta* **1769**, 316–329.

Niu L, Lu F, Pei Y, Liu C, Cao X (2007). Regulation of flowering time by the protein arginine methyltransferase AtPRMT10. *EMBO Rep*. **8**, 1190–1195.

Noh B, Lee SH, Kim HJ, Yi G, Shin EA, Lee M et al. (2004). Divergent roles of a pair of homologous jumonji/zinc-finger-class transcription factor proteins in the regulation of *Arabidopsis* flowering time. *Plant Cell* **16**, 2601–2613.

Okada Y, Scott G, Ray MK, Mishina Y, Zhang Y (2007). Histone demethylase JHDM2A is critical for Tnp1 and Prm1 transcription and spermatogenesis. *Nature* **450**, 119–123.

Pagnussat GC, Yu HJ, Ngo QA, Rajani S, Mayalagu S, Johnson CS et al. (2005). Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. *Development* **132**, 603–614.

Pavletich NP, Pabo CO (1993). Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. *Science* **261**, 1701–1707.

Pei Y, Niu L, Lu F, Liu C, Zhai J, Kong X et al. (2007). Mutations in the Type II protein arginine methyltransferase AtPRMT5 result in pleiotropic developmental defects in *Arabidopsis*. *Plant Physiol*. **144**, 1913–1923.

Saze H, Shiraishi A, Miura A, Kakutani T (2008). Control of genic DNA methylation by a jmjC domain-containing protein in *Arabidopsis thaliana*. *Science* **319**, 462–465.

Schmitz RJ, Sung S, Amasino RM (2008). Histone arginine methylation is required for vernalization-induced epigenetic silencing of FLC in winter-annual *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **105**, 411–416.

Schubert D, Primavesi L, Bishopp A, Roberts G, Doonan J, Jenuwein T et al. (2006). Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J.* **25**, 4638–4649.

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA et al. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953.

Strahl BD, Allis CD (2000). The language of covalent histone modifications. *Nature* **403**, 41–45.

Swigut T, Wysocka J (2007). H3K27 demethylases, at long last. *Cell* **131**, 29–32.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.

Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P et al. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439**, 811–816.

Wang X, Zhang Y, Ma Q, Zhang Z, Xue Y, Bao S et al. (2007). SKB1-mediated symmetric dimethylation of histone H4R3 controls flowering time in *Arabidopsis*. *EMBO J.* **26**, 1934–1941.

Whetstine JR, Nottke A, Lan F, Huarte M, Smolikov S, Chen Z et al. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **125**, 467–481.

Xiang Y, Zhu Z, Han G, Lin H, Xu L, Chen CD (2007a). JMJD3 is a histone H3K27 demethylase. *Cell Res.* **17**, 850–857.

Xiang Y, Zhu Z, Han G, Ye X, Xu B, Peng Z et al. (2007b). JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer. *Proc. Natl. Acad. Sci. USA* **104**, 19226–19231.

Xu L, Zhao Z, Dong A, Soubigou-Taconnat L, Renou JP, Steinmetz A et al. (2007). Di- and tri- but not mono-methylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in *Arabidopsis thaliana*. *Mol. Cell. Biol.* doi:10.1128/MCB.01607-07

Yamane K, Toumazou C, Tsukada Y, Erdjument-Bromage H, Tempst P, Wong J et al. (2006). JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* **125**, 483–495.

Yan D, Zhang Y, Niu L, Yuan Y, Cao X (2007). Identification and characterization of two closely related histone H4 arginine 3 methyltransferases in *Arabidopsis thaliana*. *Biochem. J.* **408**, 113–121.

Zhang K, Sridhar VV, Zhu J, Kapoor A, Zhu JK (2007). Distinctive core histone post-translational modification patterns in *Arabidopsis thaliana*. *PLoS ONE* **2**, e1210.

Zhang Y, Reinberg D (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* **15**, 2343–2360.

Zhao Z, Yu Y, Meyer D, **Wu C, Shen WH** (2005). Prevention of early flowering by expression of FLOWERING LOCUS C requires methylation of histone H3 K36. *Nat. Cell Biol.* **7**, 1256–1260.

(Handling editor: Chun-Ming Liu)