SCIENCE CHINA
Life Sciences

THEMATIC ISSUE: Plant non-coding RNAs and epigenetics
• REVIEW •

February 2018 Vol.61 No.2: 170–177
https://doi.org/10.1007/s11427-017-9163-4

Structure and mechanism of plant histone mark readers

Rui Liu[1†], Xueqin Li[1,2†], Wei Chen[1] & Jiamu Du[1*]

[1]*Shanghai Center for Plant Stress Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 201602, China;*
[2]*University of Chinese Academy of Sciences, Beijing 100049, China*

Received June 9, 2017; accepted August 11, 2017; published online October 9, 2017

In eukaryotes, epigenetic-based mechanisms are involved in almost all the important biological processes. Amongst different epigenetic regulation pathways, the dynamic covalent modifications on histones are the most extensively investigated and characterized types. The covalent modifications on histone can be "read" by specific protein domains and then subsequently trigger downstream signaling events. Plants generally possess epigenetic regulation systems similar to animals and fungi, but also exhibit some plant-specific features. Similar to animals and fungi, plants require distinct protein domains to specifically "read" modified histones in both modification-specific and sequence-specific manners. In this review, we will focus on recent progress of the structural studies on the recognition of the epigenetic marks on histones by plant reader proteins, and further summarize the general and exceptional features of plant histone mark readers.

epigenetics, histone mark, histone modifications, structure, plant

Citation: Liu, R., Li, X., Chen, W., and Du, J. (2018). Structure and mechanism of plant histone mark readers. Sci China Life Sci 61, 170–177. https://doi.org/10.1007/s11427-017-9163-4

INTRODUCTION

Epigenetic regulation is defined as a stably heritable phenotype that occurs without alterations in the DNA sequence (Berger et al., 2009). In eukaryotes, the epigenetic regulation functions in most biological processes, especially at the chromatin level. As a result of the extensive studies during the last two decades, the biochemical basis underlying the epigenetic regulation has been largely revealed as the changing of local conformation or properties of chromatin, including the dynamic covalent post-translational modifications on histones, DNA methylation, non-coding RNA-mediated processes, and the dynamic higher ordered structure of chromatin. The histone modifications could create new binding sites or eliminate existing binding sites for downstream effectors by adding or removing of certain chemical groups on histone as marks, which have been extensively studied both structurally and functionally in animal systems.

Histone modification is one of most studied and well-characterized epigenetic regulation pathway (Kouzarides, 2007), which is of a dynamic and diversified nature. Several different modification types have been identified and studied, such as lysine methylation, arginine methylation, lysine acetylation, serine/threonine phosphorylation, as well as the recently identified new marks such as lysine crotonylation (Kouzarides, 2007; Li et al., 2017). Different modifications can be deposited at different sites on the histone tails in both sequence-specific and modification-specific manners, generating the so called "histone code" to guide various downstream signaling pathways (Jenuwein and Allis, 2001; Kouzarides, 2007). Usually, a certain histone code may indicate a defined event that translates to a specific functional output, such as that histone 3 lysine 4 tri-methylation

†Contributed equally to this work
*Corresponding author (email: jmdu@sibs.ac.cn)

(H3K4me3) and H3K27me3 stand for gene activation and gene repression, respectively (Jenuwein and Allis, 2001; Kouzarides, 2007). Besides of the traditional histone modification, histone proteins have lots of variants which can combine with histone modification to lead to more diversified functional outputs (Li and Fang, 2015). The functional readout of the histone code requires specific reader modules to recognize the histone modification and transfer the signal to the downstream effector. A diverse set of histone mark reader modules have been identified in animals, such as plant homeodomain (PHD) and chromodomain for methyllysine recognition, and bromodomain and double PHD for acetyllysine recognition (Patel and Wang, 2013). In plants, most identified histone mark readers are similar to those identified in animals. However, the recognition mechanisms also have plant-specific features. In this review, we will mainly focus on plant epigenetic mark readers which possess reported 3-dimentional structures and can lead to mechanistic insights. Through structural analysis, we will discuss the general features of plant epigenetic readers and compare them with their animal counterparts, as well as outline their plant-specific properties. We list some of the structures of known plant histone mark readers, especially at the reader-histone peptide level, in Table 1, with the analysis of the structures and their functional consequences outlined in the following sections.

RECOGNITION OF METHYLATED H3K4 BY ASHH2

The CW-type zinc finger consists of a small domain of 50–60 residues that is conserved in both animals and plants, but is found only in a small number of proteins (Perry and Zhao, 2003). Most CW domain-containing proteins are involved in epigenetic regulation. It is quite interesting that although both animals and plants possess CW domains-containing proteins, none of these animal and plant CW domain-containing proteins are orthologs, suggesting a distinct evolutionary pathway (Hoppmann et al., 2011). In plants, an H3K36me3 methyltransferase *Arabidopsis* ASH1 HOMOLOG 2 (ASHH2) possesses a CW domain ahead of its histone methyltransferase domain, suggesting a potential role in the regulation of ASHH2 methyltransferase function. The *in vitro* pull-down and surface plasmon resonance assays of *Arabidopsis* ASHH2 CW domain showed that it can bind to methylated H3K4 mark with a preference for H3K4me1, while exhibiting a lower binding affinity for H3K4me2/3 (Hoppmann et al., 2011). The solution nuclear magnetic resonance (NMR)-based structure of the ASHH2 CW domain indicated the formation of an aromatic-lined methyllysine-binding pocket formed by a pair of aromatic amino acids, similar to other canonical methyllysine reader pockets (Figure 1A) (Hoppmann et al., 2011; Patel, 2016). Nevertheless, in the absence of a structure of the CW-H3K4me1 complex, it is not possible to understand the structural basis for the observed preference for H3K4me1, because there are only a very limited number of proteins that show a preference for the lower methylation state. By contrast, the CW domain-containing protein mouse Microrchidia 3 (Morc3) possesses a similar CW domain with an aromatic cage formed by two aromatic residues, too, which resembles the ASHH2 CW domain (Andrews et al., 2016; Li et al., 2016b; Liu et al., 2016a). However, mouse Morc3 shows a clear preference towards the higher methylation mark H3K4me3 (Andrews et al., 2016; Li et al., 2016b; Liu et al., 2016a). Therefore, it will be interesting to carry out further structural studies on ASHH2 CW-H3K4me1 recognition in attempts to reveal the structural basis for the differential preference between H3K4me1 and H3K4me3 marks by similar aromatic cage-containing CW domains from different proteins.

Table 1 A list of the plant histone mark readers modules discussed in this paper

Protein	Domain	Histone peptide	PDB code	Reference
ASHH2	CW	No peptides	2L7P	(Hoppmann et al., 2011)
MRG2	Chromo	H3(1–11)K4me3	4PL6	(Bu et al., 2014)
	Chromo	H3(31–41)K36me3	4PLI	
MRG701	Chromo	No peptide	5IN1	(Liu et al., 2016b)
ZMET2	Chromo	H3(1–15)K9me2	4FT2	(Du et al., 2012)
	BAH	H3(1–32)K9me2	4FT4	
SHH1	SAWADEE	H3(1–15)K9me3	4IUR	(Law et al., 2013)
	SAWADEE	H3(1–15)K9me2	4IUT	
	SAWADEE	H3(1–15)K4me1K9me1	4IUV	
	SAWADEE	H3(1–15)K9me1	4IUU	
	SAWADEE	No peptide	4IUP	
ORC1b	BAH/PHD	H3(1–15)	5HH7	(Li et al., 2016a)
ATXR5	PHD	H3(1–10)	5VAB	(Bergamin et al., 2017)

RECOGNITION OF H3K4ME3 AND H3K36ME3 BY PLANT MRG PROTEINS

The Morf related gene proteins (MRG) are conserved from yeast to animals and plants with a similar domain architecture consisting of an N-terminal chromodomain and a C-terminal featured MRG domain (Bu et al., 2014; Larschan et al., 2007; Sun et al., 2008; Xu et al., 2008; Zhang et al., 2006). The chromodomain of MRG proteins have been shown to bind to H3K36me3 mark with weak but measureable affinity in animals and fungi, but they can also bind to H3K4me3 in plants (Bu et al., 2014; Sun et al., 2008; Xu et al., 2008; Zhang et al., 2006). In animals or fungi, the H3K36me3 mark is concentrated in the 3′-end of the transcribed gene and serves as the transcriptional activation and alternative splicing signals (Kolasinska-Zwierz et al., 2009). By contrast, the H3K36me3 mark is localized in the 5′-end of the coding gene and overlapped with the H3K4me3 mark in plants (Roudier et al., 2011). In *Arabidopsis*, the two MRG proteins MRG1 and MRG2 can use their chromodomains to recognize H3K36me3 and H3K4me3 to further regulate flowering time (Bu et al., 2014). The structures of MRG2 chromodomain in complexes with H3K36me3 and H3K4me3 peptides highlight the tight accommodation of the trimethyllysine by an aromatic cage formed by four aromatic residues (Figure 1B)

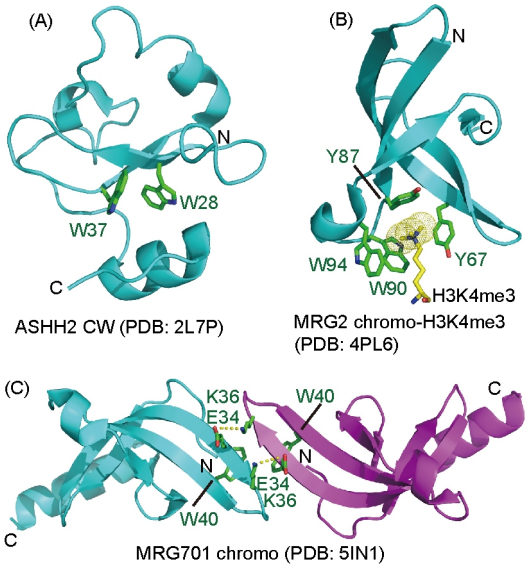

Figure 1 Structures of ASHH2 CW domain and plant MRG chromodomains. (A), The solution NMR structure of ASHH2 CW domain is colored in cyan. The two aromatic residues, which potentially contribute to methyllysine recognition, are highlighted in green stick representation. (B), The crystal structure of *Arabidopsis* MRG2 chromodomain in complex with H3K4me3 peptide. Only the trimethyllysine of the peptide can be observed in the complex. The peptide and the aromatic cage residues are highlighted in stick representation. (C), The crystal structure of rice MRG701 chromodomain. The two chromodomains, which are colored in cyan and magenta, respectively, form a homodimer with salt bridge, hydrogen bonding and hydrophobic interactions. The hydrogen bonds are highlighted in dashed lines.

(Bu et al., 2014). However, only the methyllysine, but none the flanking sequence residues, can be observed in the structure of the complex (Figure 1B), which is probably due to the weak binding affinity (0.7–0.8 mmol L^{-1}) (Bu et al., 2014). A later study showed that despite the low binding affinity, the chromodomain of plant MRG proteins can form multimers in a zig-zag orientation to align more proteins on the H3K36me3 enriched chromatin region and hence conceivably enhance the overall binding affinity (Figure 1C) (Liu et al., 2016b). This mechanism is unique in green plant but not exists in animal or fungi MRG proteins, suggesting a plant specific epigenetic regulation pattern (Liu et al., 2016b).

RECOGNITION OF METHYLATED H3K9 BY PLANT DNA METHYLATION-RELATED PROTEINS

In plants, the abundant non-CG DNA methylation, including CHG (H denotes A, T, or C) and CHH methylation, is highly-associated with H3K9me2/1 marks (Du et al., 2015). In the heterochromatin region, the CHG and CHH methylation are faithfully maintained by two chromomethylases (CMTs), named CMT3 and CMT2, respectively. These two proteins adopt similar domain architectures consisting of an N-terminal bromo-adjacent homology (BAH) domain, a C-terminal DNA methyltransferase domain, and a chromodomain unexpectedly positioned inside the primary sequence of the DNA methyltransferase domain (Du, 2016). BAH and chromo domains are primarily found in epigenetic-related proteins and serve as histone mark reader modules (Blus et al., 2011; Yang and Xu, 2013), suggesting a possible role in targeting the DNA methyltransferase domain. Structural studies on ZMET2, the equivalent of CMT3 in maize, revealed that both BAH and chromo domains can recognize the H3K9me2 mark by utilizing their conserved aromatic cages (Figure 2A and B) (Du et al., 2012). The BAH and chromo domains each possesses three aromatic residues constituting the aromatic cage to accommodate the methyllysine using a similar recognition mechanism, as observed in other classic methyllysine readers (Figure 2A and B) (Du et al., 2012; Patel, 2016). The simultaneous recognition of the H3K9me2 mark by the two domains is essential for the *in vivo* function of CMT3, given that mutations disrupting the aromatic cage of either BAH or chromo domain, impair the *in vivo* CHG methylation pattern (Du et al., 2012). Therefore, the dual recognition of H3K9me2 by the two domains of CMT3 definitively targets the DNA methyltransferase to H3K9me2-enriched chromatin loci and further reinforce the pattern of H3K9me2-directed CHG DNA methylation (Du et al., 2012). In addition to CMT3, the CHH methyltransferase CMT2 can also recognize H3K9me2 marks by its BAH and chromo domains (Stroud et al., 2014). However, their binding preferences are different in that CMT3 shows almost

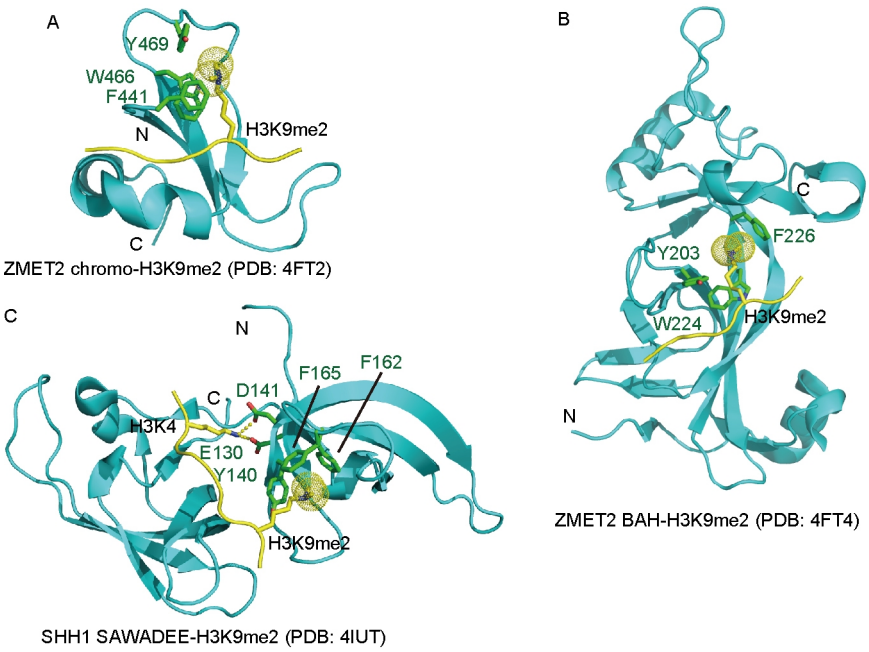

Figure 2 The structures of plant DNA methylation-related histone readers in complexes with H3K9me2 peptides. A, The crystal structure of ZMET2 in complex with H3K9me2 with the peptide bound to the chromodomain. The chromodomain is shown in a cyan ribbon and the peptide in yellow. The aromatic cage residues and H3K9me2 are highlighted in stick representation. B, The crystal structure of ZMET2 in complex with H3K9me2 with the peptide bound to the BAH domain. C, The crystal structure of SHH1 SAWADEE domain in complex with H3K9me2 peptide. The recognition includes the accommodation of the H3K9me2 by an aromatic cage and the recognition of unmethylated H3K4 by hydrogen bonding and salt bridge interactions, which are highlighted by dashed lines.

equal binding affinities to H3K9me1/2/3, but CMT2 has a significant preference for H3K9me2 over H3K9me1/3 (Stroud et al., 2014). Structural studies on CMT2 may shed light on the molecular mechanism underlying different binding preference between CMT3 and CMT2.

Besides heterochromatin, non-CG DNA methylation is also enriched in some small patches of euchromatic regions in plants (Du et al., 2015). In these regions, non-CG methylation is established and maintained by a plant specific RNA-directed DNA methylation (RdDM) pathway (Du et al., 2015). This pathway requires two steps: the upstream siRNA biogenesis by the plant-specific RNA polymerase IV (Pol IV), and the downstream non-coding RNA biogenesis by plant-specific Pol V (Matzke et al., 2015; Matzke and Mosher, 2014). The siRNA and non-coding RNA can base-pair and further direct the *de novo* DNA methyltransferase DOMAINS REARRANGED METHYLASE 2 (DRM2) to facilitate DNA methylation in certain chromatin regions (Du, 2016; Matzke et al., 2015; Matzke and Mosher, 2014). Indeed, the RdDM-controlled DNA methylation loci are determined by the specific targeting of Pol IV and Pol V. Although the elucidation of the overall mechanism for Pol IV and Pol V targeting is still a mystery, some progress has shown that some epigenetic factors, especially H3K9me and methylated DNA, play key roles in this process (Johnson et al., 2014; Law et al., 2013).

A Pol IV binding protein, SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1, also known as DNA-BINDING TRANSCRIPTION FACTOR 1, DTF1) was identified independently by two groups through immuno-precipitation followed by mass spectra (IP-MS) and a genetic screen, respectively (Law et al., 2011; Liu et al., 2011). SHH1 possesses a featured C-terminal SAWADEE domain which is a plant-specific domain with unique conserved cysteine and histidine residues that can potentially coordinate a Zn^{2+} ion (Mukherjee et al., 2009). The subsequent biochemical and structural studies uncovered that SHH1 SAWADEE domain functions as an H3K9me-binding domain (Law et al., 2013; Zhang et al., 2013). Overall, the SAWADEE domain adopts a tandem tudor-like conformation with a Zn^{2+} ion coordinated within the second tudor domain (Figure 2C) (Law et al., 2013). It is still not known whether the zinc-binding motif possesses a specific function or just simply plays a structural role. The H3K9me2 peptide is captured within the interface of the two tudor subdomains (Figure 2C), similar to the Ubiquitin-like PHD and RING finger domains 1 (UHRF1) tandem tudor-H3K9me3 complex (Law et al., 2013; Nady et al., 2011). The classic aromatic cage of the SAWADEE domain can accommodate the H3K9me1/2/3 equally well and a negatively charged pocket can specifically recognize the unmodified H3K4 or the lower H3K4me1 methylation state (Figure 2C) (Law et al., 2013). This binding is essential to target the RdDM to the gene silencing H3K9me histone mark and unmodified H3K4 to facilitate the coupling of DNA methylation to H3K9 methylation (Du et al., 2015). It turns out that although 44% Pol IV controlled loci are

targeted by SHH1, no additional histone binding protein has been identified to date to target the remaining Pol IV loci or Pol V to a specific histone mark. Considering the close connection between RdDM and histone modification, especially H3K9me, further studies are necessary for the discovery of more histone mark readers that target the RdDM machinery to certain histone marks.

RECOGNITION OF UNMODIFIED H3 BY DNA REPLICATION-RELATED PROTEINS ORC1B AND ATXR6

Compared with histone modifications that create new binding sites for reader proteins, the unmodified state of histone tails can prevent the binding of modified histone mark-specific readers and can be recognized by unmodified histone-specific readers, representing another type of readout.

DNA replication represents one of the most important biological events that ensures the faithful inheritance of genomic information to offspring (Aladjem, 2007). The initiation of DNA replication requires assembling of the origin recognition complex (ORC) at certain chromatin loci. The ORC subunit 1 (ORC1) is distinct from other ORC subunits in that it possesses an N-terminal BAH domain that is conserved from yeast, to plants and animals (Duncker et al., 2009). Yeast can employ the Orc1p BAH domain to mediate interaction with the silence information regulator 1 protein (Sir1p) (Hou et al., 2005; Hsu et al., 2005), while in metazoan the ORC1 BAH domain can recognize the DNA replication mark H4K20me2 that links DNA replication to a histone mark (Kuo et al., 2012). In plants, an exceptional feature of ORC1 is that a PHD finger is embedded inside the primary sequence of the BAH domain, making it questionable which type of histone mark guides DNA replication in plants. In genome-wide studies, it has been shown that plant DNA replication associates with H3K4me3/2, H2A.Z and H4K5Ac marks (Costas et al., 2011), and a pull-down assay indicated that *Arabidopsis* ORC1b may be associated with the H3K4me3 mark (de la Paz Sanchez and Gutierrez, 2009). However, a more quantitative isothermal titration calorimetry (ITC) measurement indicated that the plant ORC1 BAH-PHD cassette recognizes unmodified H3(1–15) peptide (Li et al., 2016a). The crystal structure of *Arabidopsis* ORC1b BAH-PHD in complex with an H3(1–15) peptide revealed a novel unmodified H3 binding mode (Figure 3A) (Li et al., 2016a). The two domains have a quite unusual arrangement in that the PHD is embedded inside the BAH domain in the primary sequence but structurally it protrudes out from the BAH domain to form an independent domain (Figure 3A) (Li et al., 2016a). The unmodified H3 peptide is specifically clamped between the BAH and PHD domains, with extensive interactions with the two domains from opposite sides (Figure 3A) (Li et al., 2016a). The recognition strictly depends on the unmodified state of H3A1, H3R2, H3T3 and H3K4, especially the specific recognition of H3R2 and H3K4 by an extensive hydrogen bond network (Figure 3A) (Li et al., 2016a). The PHD finger dominates in the recognition of H3A1, H3K4, and H3R8, while the BAH domain dominates in H3T3 recognition (Li et al., 2016a). H3R2 and H3Q5 interact with both the BAH and PHD domains, resulting in the mixed combinatorial readout of a series of histone residues by the two domains coordinately (Li et al., 2016a). This multivalent recognition mode is unique in plants and has never been observed in animals or fungi, indicating a possible plant-specific combinatorial readout and a potential link between plant DNA replication and unmodified H3.

In the heterochromatin region of plants, DNA replication requires a specific histone mark H3K27me1, which is deposited by two plant specific histone methyltransferases *ARABIDOPSIS* TRITHORAX-RELATED PROTEIN 5 (ATXR5) and ATXR6 (Jacob et al., 2009; Jacob et al., 2010). The two proteins have an N-terminal PHD finger that recognize

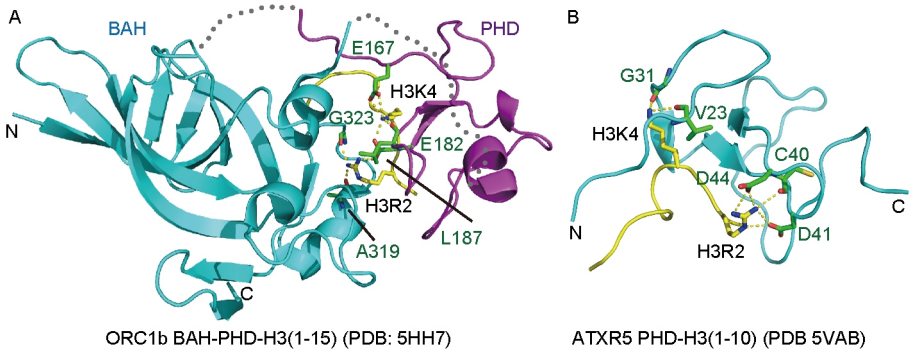

Figure 3 Structures of plant DNA replication-related histone readers ORC1b and ATXR5 PHD finger in complexes with unmodified H3 peptides. A, Crystal structure of *Arabidopsis* ORC1b BAH-PHD cassette in complex with an unmodified H3(1–15) peptide with the BAH and PHD domains colored in cyan and magenta, respectively. The recognition features specific hydrogen bonding and salt bridge interactions with unmodified H3R2 and H3K4, with the hydrogen bonds highlighted by dashed lines. B, Crystal structure of ATXR5 PHD finger in complex with unmodified H3(1–10) peptide. H3K4 and H3R2 are specifically recognized by extensive hydrogen bonding interactions.

unmodified H3K4 and a C-terminal Su(var)3-9, enhancer of zeste, trithorax (SET) domain that specifically methylate H3K27 on H3.1 but not H3.3 (Jacob et al., 2010; Jacob et al., 2014). The structure of ATXR5 PHD finger in complex with an unmodified H3(1–10) peptide reveals that ATXR5 PHD finger uses a canonical mode to bind to the unmodified H3R2 and H3K4 (Figure 3B) (Bergamin et al., 2017; Li and Li, 2012). An extensive hydrogen bond network is formed to achieve the specific recognition to unmodified H3R2 and H3K4 (Figure 3B) (Bergamin et al., 2017). The recognition of H3K4 by the PHD finger can enhance the binding between ATXR5 to the substrate nucleosome (Bergamin et al., 2017). This would prevent the enzyme from functioning at the active euchromatic region, which is marked by the H3K4me3 mark. Interestingly, the PHD finger of ATXR5 can specifically interact with the SET domain in the presence of cofactor substrate *S*-adenosyl-*L*-methionine (SAM), providing an addition layer for the regulation of the ATXR5 catalytic function by the PHD finger, though the structure is still unavailable (Bergamin et al., 2017).

It is interesting that although there is no report on the functional relationship between plant DNA replication and the unmodified H3R2 and H3K4, the two available structures of the plant DNA replication-related proteins, ORC1b and ATXR5, have reader modules for recognition of unmodified H3R2 and H3K4. The two proteins both have distinct domain architectures other than their animal counterparts, revealing a plausible plant-specific regulatory mechanism of DNA replication by unmodified H3, although the possibility requires to be further elucidated by additional functional investigation.

CONCLUSIONS AND PERSPECTIVES

In general, most of the currently reported plant epigenetic mark reader modules are similar to those found in animals, such as the common PHD finger, chromodomain and BAH domain, as well as other reader modules (Liu and Min, 2016). It is probably due to the conserved function of these domains during evolution that very limited plant-specific epigenetic mark reader modules have been identified. One reported plant-specific reader is the SAWADEE domain, which is conserved in plants but does not exist in animals. The structural studies demonstrate that SAWADEE is a type of variant version of tandem tudor with a featured zinc-binding motif within the second tudor (Law et al., 2013). Therefore, the histone mark reader modules in plants are most likely to be restricted to the same types as have already been identified in animals or alternately just animal epigenetic reader derivations, with limited variation. The histone-binding substrate of the same reader domain may vary in plants and animals. For instance, ORC1 BAH domain was reported to recognize H4K20me2 in animals (Kuo et al., 2012), in contrast to the recognition of non-histone protein in yeast (Hou et al., 2005; Hsu et al., 2005). On the other hand, it turns out to recognize unmodified H3 together with a fused PHD finger in plants (Li et al., 2016a), indicating the gain of a different function in plants during evolution. All the current available structures of plant histone mark readers are involved in methyllysine or unmodified lysine readout. The recognition of these marks employ the classical aromatic cage to accommodate the methyllysine, or a hydrogen bond network to specifically bind unmodified lysine, indicating a common recognition mechanism as in animals (Patel, 2016). Structurally, all these recognition mechanisms are within the general context that has been established by the extensive studies in animal histone mark readers. At the whole protein level, a particular feature of the plant histone mark reading proteins is that several of the reader domains graft into the primary sequence of another domain but structurally extend out to form independent structural and functional domains, such as the methylatransferase domain-embedded chromodomain of plant CMTs and the BAH domain-embedded PHD finger of plant ORC1b. This is different from the canonical manner of tandem-linked domains required to connect two domains. A plausible reason might be that the extensive and active transposons in plant genomes make the grafting events happen more frequently in plants.

In the past decades, extensive plant epigenetic studies have accumulated a plenty of genetic and genomic data. Further investigation of more biochemical and structural studies to the plant field may contribute to the discovery of more plant-specific epigenetic mark readers and to reveal their readout mechanisms. It would be interesting to investigate whether there are some absolute plant-specific novel epigenetic mark reader modules, or whether there is novel type of histone mark substrate for known reader modules in plants, and whether there are any plant-derived novel epigenetic mark recognition mechanisms. Indeed, plants provide a very powerful research system to investigate the *in vivo* function of histone mark readers. There are available mutant libraries of seeds and it is quite convenient to carry out genetic and genomic studies with the plant mutants. This offers an opportunity to study the function of a certain gene and the crosstalk between different genes at the individual level. The structural and biochemical studies at the molecular level and the genetics and genomic studies at the individual level are complemented and the combination of them would shape a comprehensive understanding of the function of a certain histone reader. Therefore, the plant system provides a different way to investigate the epigenetic regulation mechanism and is a strong tool for the structural and functional analyses of the histone mark readers both *in vivo* and *in vitro*.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the National Key R&D Program of China (2016YFA0503200), National Natural Science Foundation of China (31622032), and the Chinese Academy of Sciences to Jiamu Du.*

Aladjem, M.I. (2007). Replication in context: dynamic regulation of DNA replication patterns in metazoans. Nat Rev Genet 8, 588–600.

Andrews, F.H., Tong, Q., Sullivan, K.D., Cornett, E.M., Zhang, Y., Ali, M., Ahn, J.W., Pandey, A., Guo, A.H., Strahl, B.D., Costello, J.C., Espinosa, J.M., Rothbart, S.B., and Kutateladze, T.G. (2016). Multivalent chromatin engagement and inter-domain crosstalk regulate MORC3 ATPase. Cell Rep 16, 3195–3207.

Bergamin, E., Sarvan, S., Malette, J., Eram, M.S., Yeung, S., Mongeon, V., Joshi, M., Brunzelle, J.S., Michaels, S.D., Blais, A., Vedadi, M., and Couture, J.F. (2017). Molecular basis for the methylation specificity of ATXR5 for histone H3. Nucleic Acids Res 45, 6375–6387.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes Dev 23, 781–783.

Blus, B.J., Wiggins, K., and Khorasanizadeh, S. (2011). Epigenetic virtues of chromodomains. Crit Rev Biochem Mol Biol 46, 507–526.

Bu, Z., Yu, Y., Li, Z., Liu, Y., Jiang, W., Huang, Y., and Dong, A.W. (2014). Regulation of *Arabidopsis* flowering by the histone mark readers MRG1/2 via interaction with CONSTANS to modulate FT expression. PLoS Genet 10, e1004617.

Costas, C., de la Paz Sanchez, M., Stroud, H., Yu, Y., Oliveros, J.C., Feng, S., Benguria, A., López-Vidriero, I., Zhang, X., Solano, R., Jacobsen, S.E., and Gutierrez, C. (2011). Genome-wide mapping of *Arabidopsis thaliana* origins of DNA replication and their associated epigenetic marks. Nat Struct Mol Biol 18, 395–400.

de la Paz Sanchez, M., and Gutierrez, C. (2009). *Arabidopsis* ORC1 is a PHD-containing H3K4me3 effector that regulates transcription. Proc Natl Acad Sci USA 106, 2065–2070.

Du, J. (2016). Structure and mechanism of plant DNA methyltransferases. Adv Exp Med Biol 945, 173–192.

Du, J., Johnson, L.M., Jacobsen, S.E., and Patel, D.J. (2015). DNA methylation pathways and their crosstalk with histone methylation. Nat Rev Mol Cell Biol 16, 519–532.

Du, J., Zhong, X., Bernatavichute, Y.V., Stroud, H., Feng, S., Caro, E., Vashisht, A.A., Terragni, J., Chin, H.G., Tu, A., Hetzel, J., Wohlschlegel, J.A., Pradhan, S., Patel, D.J., and Jacobsen, S.E. (2012). Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell 151, 167–180.

Duncker, B.P., Chesnokov, I.N., and McConkey, B.J. (2009). The origin recognition complex protein family. Genome Biol 10, 214.

Hoppmann, V., Thorstensen, T., Kristiansen, P.E., Veiseth, S.V., Rahman, M.A., Finne, K., Aalen, R.B., and Aasland, R. (2011). The CW domain, a new histone recognition module in chromatin proteins. EMBO J 30, 1939–1952.

Hou, Z., Bernstein, D.A., Fox, C.A., and Keck, J.L. (2005). Structural basis of the Sir1-origin recognition complex interaction in transcriptional silencing. Proc Natl Acad Sci USA 102, 8489–8494.

Hsu, H.C., Stillman, B., and Xu, R.M. (2005). Structural basis for origin recognition complex 1 protein-silence information regulator 1 protein interaction in epigenetic silencing. Proc Natl Acad Sci USA 102, 8519–8524.

Jacob, Y., Bergamin, E., Donoghue, M.T.A., Mongeon, V., LeBlanc, C., Voigt, P., Underwood, C.J., Brunzelle, J.S., Michaels, S.D., Reinberg, D., Couture, J.F., and Martienssen, R.A. (2014). Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. Science 343, 1249–1253.

Jacob, Y., Feng, S., LeBlanc, C.A., Bernatavichute, Y.V., Stroud, H., Cokus, S., Johnson, L.M., Pellegrini, M., Jacobsen, S.E., and Michaels, S.D. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. Nat Struct Mol Biol 16, 763–768.

Jacob, Y., Stroud, H., Leblanc, C., Feng, S., Zhuo, L., Caro, E., Hassel, C., Gutierrez, C., Michaels, S.D., and Jacobsen, S.E. (2010). Regulation of heterochromatic DNA replication by histone H3 lysine 27 methyltransferases. Nature 466, 987–991.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science 293, 1074–1080.

Johnson, L.M., Du, J., Hale, C.J., Bischof, S., Feng, S., Chodavarapu, R.K., Zhong, X., Marson, G., Pellegrini, M., Segal, D.J., Patel, D.J., and Jacobsen, S.E. (2014). SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. Nature 507, 124–128.

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. Nat Genet 41, 376–381.

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

Kuo, A.J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J.K., Patel, D.J., and Gozani, O. (2012). The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. Nature 484, 115–119.

Larschan, E., Alekseyenko, A.A., Gortchakov, A.A., Peng, S., Li, B., Yang, P., Workman, J.L., Park, P.J., and Kuroda, M.I. (2007). MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. Mol Cell 28, 121–133.

Law, J.A., Du, J., Hale, C.J., Feng, S., Krajewski, K., Palanca, A.M.S., Strahl, B.D., Patel, D.J., and Jacobsen, S.E. (2013). Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. Nature 498, 385–389.

Law, J.A., Vashisht, A.A., Wohlschlegel, J.A., and Jacobsen, S.E. (2011). SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. PLoS Genet 7, e1002195.

Li, M., and Fang, Y.D. (2015). Histone variants: the artists of eukaryotic chromatin. Sci China Life Sci 58, 232–239.

Li, S., Yang, Z., Du, X., Liu, R., Wilkinson, A.W., Gozani, O., Jacobsen, S.E., Patel, D.J., and Du, J. (2016a). Structural basis for the unique multivalent readout of unmodified H3 tail by *Arabidopsis* ORC1b BAH-PHD cassette. Structure 24, 486–494.

Li, S., Yen, L., Pastor, W.A., Johnston, J.B., Du, J., Shew, C.J., Liu, W., Ho, J., Stender, B., Clark, A.T., Burlingame, A.L., Daxinger, L., Patel, D.J., and Jacobsen, S.E. (2016b). Mouse MORC3 is a GHKL ATPase that localizes to H3K4me3 marked chromatin. Proc Natl Acad Sci USA 113, E5108–E5116.

Li, Y., and Li, H. (2012). Many keys to push: diversifying the 'readership' of plant homeodomain fingers. Acta Biochim Biophys Sin 44, 28–39.

Li, Y., Zhao, D., Chen, Z., and Li, H. (2017). YEATS domain: Linking histone crotonylation to gene regulation. Transcription 8, 9–14.

Liu, J., Bai, G., Zhang, C., Chen, W., Zhou, J., Zhang, S., Chen, Q., Deng, X., He, X.J., and Zhu, J.K. (2011). An atypical component of RNA-directed DNA methylation machinery has both DNA methylation-dependent and -independent roles in locus-specific transcriptional gene silencing. Cell Res 21, 1691–1700.

Liu, Y., and Min, J. (2016). Structure and function of histone methylation-binding proteins in plants. Biochem J 473, 1663–1680.

Liu, Y., Tempel, W., Zhang, Q., Liang, X., Loppnau, P., Qin, S., and Min, J. (2016a). Family-wide characterization of histone binding abilities of human CW domain-containing proteins. J Biol Chem 291, 9000–9013.

Liu, Y., Wu, H., Yu, Y., and Huang, Y. (2016b). Structural studies on MRG701 chromodomain reveal a novel dimerization interface of MRG proteins in green plants. Protein Cell 7, 792–803.

Matzke, M.A., Kanno, T., and Matzke, A.J.M. (2015). RNA-directed DNA methylation: the evolution of a complex epigenetic pathway in flowering plants. Annu Rev Plant Biol 66, 243–267.

Matzke, M.A., and Mosher, R.A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet 15, 394–408.

Mukherjee, K., Brocchieri, L., and Bürglin, T.R. (2009). A comprehensive classification and evolutionary analysis of plant homeobox genes. Mol Biol Evol 26, 2775–2794.

Nady, N., Lemak, A., Walker, J.R., Avvakumov, G.V., Kareta, M.S., Achour, M., Xue, S., Duan, S., Allali-Hassani, A., Zuo, X., Wang, Y.X., Bronner, C., Chédin, F., Arrowsmith, C.H., and Dhe-Paganon, S. (2011). Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. J Biol Chem 286, 24300–24311.

Patel, D.J. (2016). A structural perspective on readout of epigenetic histone and DNA methylation marks. Cold Spring Harb Perspect Biol 8, a018754.

Patel, D.J., and Wang, Z. (2013). Readout of epigenetic modifications. Annu Rev Biochem 82, 81–118.

Perry, J., and Zhao, Y. (2003). The CW domain, a structural module shared amongst vertebrates, vertebrate-infecting parasites and higher plants. Trends Biochem Sci 28, 576–580.

Roudier, F., Ahmed, I., Bérard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., Bouyer, D., Caillieux, E., Duvernois-Berthet, E., Al-Shikhley, L., Giraut, L., Després, B., Drevensek, S., Barneche, F., Dèrozier, S., Brunaud, V., Aubourg, S., Schnittger, A., Bowler, C., Martin-Magniette, M.L., Robin, S., Caboche, M., and Colot, V. (2011). Integrative epigenomic mapping defines four main chromatin states in *Arabidopsis*. EMBO J 30, 1928–1938.

Stroud, H., Do, T., Du, J., Zhong, X., Feng, S., Johnson, L., Patel, D.J., and Jacobsen, S.E. (2014). Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. Nat Struct Mol Biol 21, 64–72.

Sun, B., Hong, J., Zhang, P., Dong, X., Shen, X., Lin, D., and Ding, J. (2008). Molecular basis of the interaction of *Saccharomyces cerevisiae* Eaf3 chromo domain with methylated H3K36. J Biol Chem 283, 36504–36512.

Xu, C., Cui, G., Botuyan, M.V., and Mer, G. (2008). Structural basis for the recognition of methylated histone H3K36 by the Eaf3 subunit of histone deacetylase complex Rpd3S. Structure 16, 1740–1750.

Yang, N., and Xu, R.M. (2013). Structure and function of the BAH domain in chromatin biology. Crit Rev Biochem Mol Biol 48, 211–221.

Zhang, H., Ma, Z.Y., Zeng, L., Tanaka, K., Zhang, C.J., Ma, J., Bai, G., Wang, P., Zhang, S.W., Liu, Z.W., Cai, T., Tang, K., Liu, R., Shi, X., He, X.J., and Zhu, J.K. (2013). DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. Proc Natl Acad Sci USA 110, 8290–8295.

Zhang, P., Du, J., Sun, B., Dong, X., Xu, G., Zhou, J., Huang, Q., Liu, Q., Hao, Q., and Ding, J. (2006). Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. Nucleic Acids Res 34, 6621–6628.