

***Arabidopsis thaliana*: a powerful model organism to explore histone modifications and their upstream regulations**

Yang Yu, Sihan Wang, Ziqin Wang, Renwei Gao & Joohyun Lee

To cite this article: Yang Yu, Sihan Wang, Ziqin Wang, Renwei Gao & Joohyun Lee (2023) *Arabidopsis thaliana*: a powerful model organism to explore histone modifications and their upstream regulations, Epigenetics, 18:1, 2211362, DOI: [10.1080/15592294.2023.2211362](https://doi.org/10.1080/15592294.2023.2211362)

To link to this article: <https://doi.org/10.1080/15592294.2023.2211362>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 17 May 2023.



[Submit your article to this journal](#)



Article views: 3569



[View related articles](#)



[View Crossmark data](#)



Citing articles: 9 [View citing articles](#)

REVIEW



Arabidopsis thaliana: a powerful model organism to explore histone modifications and their upstream regulations

Yang Yu*, Sihan Wang*, Ziqin Wang, Renwei Gao, and Joohyun Lee

Division of Natural and Applied Sciences, Duke Kunshan University, Kunshan, Jiangsu, China

ABSTRACT

Histones are subjected to extensive covalent modifications that affect inter-nucleosomal interactions as well as alter chromatin structure and DNA accessibility. Through switching the corresponding histone modifications, the level of transcription and diverse downstream biological processes can be regulated. Although animal systems are widely used in studying histone modifications, the signalling processes that occur outside the nucleus prior to histone modifications have not been well understood due to the limitations including non viable mutants, partial lethality, and infertility of survivors. Here, we review the benefits of using *Arabidopsis thaliana* as the model organism to study histone modifications and their upstream regulations. Similarities among histones and key histone modifiers such as the Polycomb group (PcG) and Trithorax group (TrxG) in *Drosophila*, Human, and *Arabidopsis* are examined. Furthermore, prolonged cold-induced vernalization system has been well-studied and revealed the relationship between the controllable environment input (duration of vernalization), its chromatin modifications of *FLOWERING LOCUS C (FLC)*, following gene expression, and the corresponding phenotypes. Such evidence suggests that research on *Arabidopsis* can bring insights into incomplete signalling pathways outside of the histone box, which can be achieved through viable reverse genetic screenings based on the phenotypes instead of direct monitoring of histone modifications among individual mutants. The potential upstream regulators in *Arabidopsis* can provide cues or directions for animal research based on the similarities between them.

ARTICLE HISTORY

Received 24 November 2022
Revised 7 April 2023
Accepted 28 April 2023

KEYWORDS

Arabidopsis thaliana; model organisms; histone modifications; epigenetics

Introduction


Eukaryotic nuclear DNA is packed into chromatin, a condensed structure whose basic unit is the nucleosomes [1]. The core of nucleosomes is composed of a histone octamer containing two copies of histone H2A, H2B, H3, and H4 proteins, and an approximately 146-base-pair DNA segment wrapped around it [2,3]. Other than providing structural support, histones protrude their highly basic side chains from the nucleosome, receiving extensive covalent modifications on the tails, affecting inter-nucleosomal interactions and thus altering the chromatin structure and DNA accessibility [4–6]. Various histone modifications can be recognized by effector modules, which then transduce signals for further chromatin structural or functional changes [7]. The status of histone

modification can switch by ‘writers’ and ‘erasers,’ underlying the dynamics of the chromatin structure and the process of transcriptional initiation and elongation [5,8]. The restructuring and remodelling of chromatin result in controlling the level of transcription through the corresponding histone modifications [9,10]. These modifications transduce the inheritance of gene expression patterns without altering the underlying DNA sequences, but by adapting chromatin structures [11,12].

Histone modifications have attracted significant attention in biomedical research due to their involvement in cancer onset, neurological diseases, and embryonic development [13–15]. Because of the limitations of human research, studying the basic science of histone

CONTACT Joohyun Lee  joohyun.lee@duke.edu  Division of Natural and Applied Sciences, Duke Kunshan University, Jiangsu, China

*These authors contributed equally to this manuscript.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/15592294.2023.2211362>

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

modifications in other organisms is an essential requirement. The ideal model organism for studying histone modifications should possess orthologous genes and the usual characteristics that make a good model system, including well-studied genomic features, the feasibility to maintain and reproduce in the lab settings with a short generation time, and the capacity to generate mutants [16]. Additionally, the model organism should enable researchers to interconnect multiple layers of genetic alterations with phenotypic changes to determine the role of these genes in fundamental biochemical processes [17–19]. Among animal models, *Drosophila* and mice are the most chosen model organisms for histone modification studies because of their well-studied genetic frameworks and the conserved molecular pathways that are relevant to human diseases [20,21]. However, there are challenges that impede a broader inspection of the histone modifications using these animal models. First, the knock-out of a key histone regulator frequently leads to embryonic deaths and/or infertility in the animal models, which limits the spectrum of the study [22–24]. As an alternative, gene knock-down is performed as one of the major approaches in animal models to examine the function of a gene through targeted reduction of its expression level [25–28]. The available knock-out mutants listed in major databases of *Drosophila* and mice cannot cover all epigenetic pathways. As an example, in the Mouse Genome Informatics (<http://www.informatics.jax.org/>) database, 189 epigenetics-related genes can be found. By matching the gene symbols with records in Mutant Mouse Resource & Research Centers (MMRRC), a national mutant mice repository, only 35 knock-out strains are found to be available [29]. The mutant availability of *Arabidopsis* enables comprehensive reverse genetic screenings of the upstream regulators in the signalling pathway using subsequent phenotypic changes as the indicator, which reflects the level of environmental inputs and corresponding switching of histone modifications [30]. The available mutants allow a high-throughput screening of phenotypes to study the upstream signalling, without direct monitoring of histone

modification in each mutant line required at the initial genetic screening. Second, transgenerational inheritance of environment-induced histone modifications is still not consistently observed in animal models [31,32]. The complex and changeable information behind histone modifications may not be able to accurately flow from parents to offspring, leaving potential obstacles to the research across multiple generations. Also, studies on the topic including cancer, ageing, and drug development have addressed the significance of the upstream signalling pathways that convert the environmental stimuli to epigenetic changes [33–35]. Due to the limitation of mutant screening in animal models, the exploration of upstream regulators from the environment involved in signalling histone modification from the outside of the nucleus has been rarely accomplished, and the profile of upstream signalling pathways of histone modifications across development remained uncovered [36].

Various studies have shown that the mechanisms of histone modifications and their molecular components are highly conserved evolutionarily across higher eukaryotes including plants, suggesting that plants research has the potential to unravel epigenetic pathways that are yet not completely understood and complement research conducted by animal models [37–40]. According to The Arabidopsis Information Resource (TAIR), a comprehensive *Arabidopsis thaliana* information database, 131 genes are annotated with epigenetics-related functions and all of their corresponding knock-out mutants are available, which implies that they are not infertile or lethal [41]. Moreover, there are over 260,000 individual mutant lines of *Arabidopsis thaliana* generated via *Agrobacterium*-mediated T-DNA mutagenesis (SALK, SAIL, and GABI-KAT lines), covering almost all *Arabidopsis thaliana* genes [42–45]. These cost-effective and easy-to-handle random insertional mutants can be easily accessed in multiple *Arabidopsis thaliana* stock centres worldwide, such as Arabidopsis Biological Resource Center (ABRC) (<https://abrc.osu.edu/>), Nottingham Arabidopsis Stock Centre (NASC) (<https://arabidopsis.info>) and GABI-Kat

(<https://www.gabi-kat.de>). Additionally, plants are immobile and so are more plastic in post-embryonic development to react to changing environmental conditions [46], which then minimizes behavioural artefacts. Consequently, a robust inheritance of histone modifications is observed in *Arabidopsis*, making it a good candidate to investigate cues of epigenetic memory [38,47].

The research using plant models has contributed to understanding the process that upstream environment inputs alter histone modifications and gene expression patterns [48]. For example, the well-studied vernalization system, a temperature-sensing process by which exposure to the prolonged cold during winter, leads to an epigenetic switch that permits flowering in the spring, provides a quantitatively controllable experimental system for histone modifications. In such a case, environmental factors regulate the establishment and maintenance of epigenetic modifications, which eventually affect gene expressions and phenotypic changes [49,50]. *Arabidopsis thaliana* enables the study of cold-induced environmental epigenetics of *FLOWERING LOCUS C* (*FLC*) for the following reasons: (1) competence to flower via vernalization is determined by chromatin modifications of *FLC* expression, (2) cold exposure can be quantitatively controlled, and (3) mutants of chromatin control in vernalization process can be easily identified by monitoring flowering time. Polycomb Repressive Complex 2 (PRC2) and Trithorax group (TrxG) complexes are found in both plants and animals and confer the balancing of *FLC* expression via quantitative depositions of H3K27me3 and H3K4me3 on *FLC* chromatin. During cold exposure, *FLC* expression is gradually reduced via increased Histone 3 lysine 27 trimethylation (H3K27me3) deposition on *FLC* chromatin by PRC2, and the decrease of Histone 3 lysine 4 trimethylation (H3K4me3) deposition was also monitored [30,51–53]. After vernalization, *FLC* silencing is stably maintained even when the plant grows in warm conditions via PRC1. Because *FLC* expression in *Arabidopsis* is the only determinant to promote flowering in spring, the flowering time is correlated with the duration of cold treatment, deposition of

H3K27me3 on *FLC* chromatin, and the level of *FLC* expression. This example demonstrates the potential of using *Arabidopsis* as a controllable system to explore mechanisms and components transducing signals from the environmental stimuli to the responses inside the nucleus.

Similarity of histone proteins between animals and plants

Histone proteins are highly conserved in eukaryotes, with structural constraint likely resulting from necessary assembly into the histone octamers and functional constraint from compact binding on DNA [54–58]. Among the histone family members, histones H3 and H4 are the most evolutionarily conserved and they bind to the DNA terminal segments that enter or leave the nucleosome, playing a more outstanding role in chromatin formation regulation [55,59–61]. Based on the evolutionary analysis of histone H3 protein, a high level of sequential conservation is maintained among eukaryotes, and ancient genome duplication events led to functional histone variants like H3.1 and H3.3 [62–65]. H3.3 differs from H3.1 at only three or four amino acids across species [66–69]. The expression of H3.1 is mainly detected during the S phase, which is coupled with DNA synthesis during replication [70–73]. In contrast, H3.3 is expressed throughout the whole cell cycle, serving as the replacement variant for the DNA-synthesis-independent deposition pathway [74,75]. Histone H4 proteins are also highly conserved among higher eukaryotes, which can be explained that the H4 genes have experienced strong purifying selection and birth-and-death evolution [76]. In addition, despite the independent emergence of many histone variants across animals and plants, their amino acid sequences show a limited number of differences at approximately the same loci, implying a selective pressure in evolutionary history [77].

The sequence alignment of the N-terminal tails in *Arabidopsis thaliana* histone subunits H3.1, H3.3, and H4 with those of other species indicates a high level of conservation across organisms (Table 1 and Figure S1). As expected, *Arabidopsis thaliana* H3.1, H3.3, or H4 shows a high level of identity and similarity with the human H3.1 (98%

Table 1. The identity and similarity of N-terminus tail in histone subunits H3.1, H3.3, and H4 of *Arabidopsis thaliana* across organisms.

Species	Histone 3.1		Histone 3.3		Histone 4	
	Identity	Similarity	Identity	Similarity	Identity	Similarity
<i>Caenorhabditis elegans</i>	N/A	N/A	100%	100%	100%	100%
<i>Canis lupus familiaris</i>	98%	100%	98%	100%	100%	100%
<i>Drosophila melanogaster</i>	N/A	N/A	98%	100%	98%	100%
<i>Homo sapiens</i>	98%	100%	98%	100%	100%	100%
<i>Macaca mulatta</i>	98%	100%	98%	100%	100%	100%
<i>Mus musculus</i>	98%	100%	98%	100%	100%	100%
<i>Rattus norvegicus</i>	98%	100%	98%	100%	100%	100%

identity and 100% similarity), H3.3 (98% identity and 100% similarity) or H4 (100% identity and 100% similarity). However, it is important to note that sequence similarities may not always represent the same function as shown by previous research which compares *Arabidopsis UVR8* receptor sequence with RCC1 protein family in animals [78].

The basic C-terminal domain (CTD) of histone proteins is critical for chromatin binding and the N-terminal tail of histones modulates nucleosome structure and function [79,80]. Covalent modifications on the N-terminal tail of histone, including methylation and acetylation, are found in mammals and plants [5,81,82]. Most organisms share the modifications in the same amino acids, including H3K36, H3K4, and H3K27, resulting in similar transcriptional controls [83,84]. Lys-to-Met (K-to-M) substitutions of a key amino acid of histone H3, as an example, showed dominant-negative inhibition in all humans, *Drosophila*, and *Arabidopsis* [77,85–88]. Studies using human embryonic kidney–293T (HEK293T) cells suggested that exogenously expressed histone H3 Lys-36 to Met (H3K36M) mutant protein resulted in a global reduction of endogenous H3K36 methylation [85,86,89,90]. In *Drosophila*, overexpression of H3K36M in the eye decreases the H3K36me₂, suggesting its dominant-negative effect on lysine methyltransferases [91]. In *Arabidopsis*, overexpressed H3K36M also induces a global H3K36 hypomethylation in the same dominant-negative manner [85,92].

Another example is the histone H3.3, which is required for reprogramming events during development across species and is associated with genes that are actively transcribing in both animals and plants [93]. Histone H3.3 is encoded by two genes in *Drosophila* (*H3.3A* and *H3.3B*) and human

(*H3f3a* and *H3f3b*), and the knock-out mutants of one of either genes show decreased viability, partial lethality or infertility of the survivors [69,94,95]. Knock-down studies in mice also showed that the complete depletion of H3.3 caused early embryonic lethality, suggesting the significance of H3.3 in animal development [96]. In *Arabidopsis*, *HTR4*, *HTR5*, and *HTR8* encode H3.3. Although the triple knockout is lethal, *HTR4-HTR5*-double-knockout-*HTR8*-RNAi lines could survive [93]. Research on this mutant found a reduced linker H1 deposition and thus increased gene body DNA methylation [93]. The viable H3.3 mutant indicates that *Arabidopsis* can contribute to understanding the underlying mechanisms of histone proteins by its various knock-out mutants. Such examples also support that histone research with plants will complement animal models in answering unresolved questions that can be difficult to address with animal models.

Similarity of histone modifier between animals and plants: PcG and polycomb complexes

Other than homologies in histone proteins and modification types, the shared or similar modifying mechanisms and conserved modifiers between *Arabidopsis* and animal models are also of great importance. Polycomb group (PcG) proteins were originally identified in *Drosophila* as the repressors of homoeotic (*Hox*) genes during embryonic development [97]. PcG homologs were then identified in various species including multicellular plants, animals, and some unicellular organisms, which suggests that the emergence of PcG proteins may trace back to early in eukaryotic evolution [98,99]. Although

sequence similarities are not always correlated to the functional similarity between species [100,101], inter-species protein homology searching is a tool to recognize cues for function prediction [102–104]. In higher eukaryotes, PcG proteins constitute a global gene silencing system that is essential for maintaining cell identity and regulating growth and development [105,106]. PcG proteins are assembled in several distinct transcriptional-repressive complexes, among which the two major groups are Polycomb repressive complex 1 (PRC1) and 2 (PCR2) [107]. PRC2 catalyses mono-, di- and trimethylations of lysine 27 of histone H3 (H3K27me1/2/3)

and PRC1 functions to maintain the H3K27me3 established by PRC2 as well as introduce H2A lysine 119 ubiquitination (H2AK119ub1) [108–111].

PRC2 establishes H3K27 methylation in various organisms through the cooperation of the highly conserved core complex [112,113]. The *Drosophila* PRC2 core comprises four subunits: Enhancer of zeste (E(Z)), Extra sex comb (ESC/ESCL), Suppressor of zeste 12 (SU(Z)12), and NURF55 [114] (Figure 1). The former three subunits confer the histone methyltransferase activity and RBBP4/7 functions of PRC2 to bind to unmodified nucleosomes, which establish and

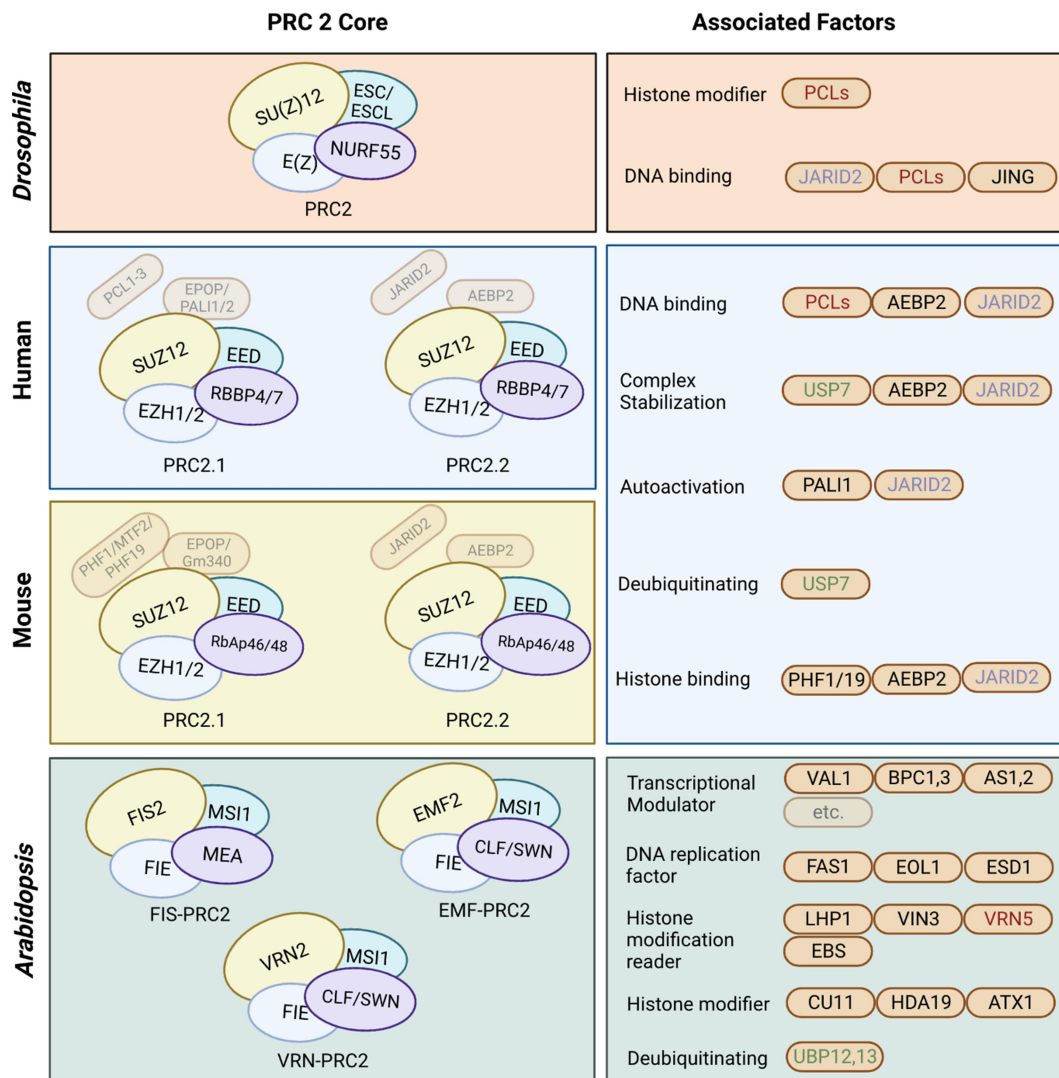


Figure 1. Comparison of PRC2 core components and their associated factors among *Drosophila*, human, mouse, and Arabidopsis. The interaction of PRC2 core components is represented among *Drosophila*, human, and Arabidopsis. The core components used the same colour between the organisms indicate corresponding orthologous proteins. The associated factors are classified by their functions, and their colour-coding indicates orthologous proteins between organisms.

maintain chromatin structures [115–117]. Correspondingly, functional and structural homologs of PRC2 core components in humans are E(Z) (EZH1/2), ESC/ESCL (EED), SU(Z)12 (SUZ12), and NURF55 (RBBP4/7) [118]. In mouse, the functional and structural homologs of PRC2 core components are E(Z) (EZH1/2), ESC/ESCL (EED), SU(Z)12 (SUZ12), and NURF55 (RbAp46/48) [119]. In Arabidopsis, homologous genes encoding the PRC2 core complex are also found, which include E(Z) (MEA, CLF, and SWN), ESC/ESCL (FIE), SU(Z)12 (FIS2, EMF2, and VRN2), and NURF55 (MSI1) [99], and these components are all orthologs with *Drosophila*, mice, and human, which are represented in the same colour of Figure 1. PRC2 mediates transcriptional repression via globally monitored H3K27me3 in *Drosophila* [114,120,121], mammals [122], and plants including Arabidopsis [123]. PRC2-mediated histone methylation requires targeted recruitment of PRC2 factors to the specific region of the chromatin [124,125]. Although distinct models are proposed to explain recruitment processes at a certain level, cross-species similarity is found in all key steps of each model. In *Drosophila*, PRC2 interacts with DNA-binding proteins to recognize and catalyse H3K27 methylation at Polycomb Response Elements (PREs) loci and DNA fragments with transcriptional factor (TF) binding site, with the latter contributes to the PRC2 recruitment [108,126]. Loci like PREs are also found in mammals and Arabidopsis with fixed-binding transcription factor families [127–131]. However, a PRE-KO experiment in *Drosophila* demonstrates that PREs may not be a necessary condition for H3K27me3 modification [132] and functional PREs are not globally identified in the genomes of mammals yet [133,134]. In mammals, PRC2 tends to bind unmethylated CG-rich CpG islands at the promoters of target genes [126,135,136]. Besides, many long noncoding RNAs (lncRNAs), including HOTAIR, Kcnq1ot1, and Braveheart, could recruit PRC2 through specific RNA-protein binding in mammals [137]. In Arabidopsis, lncRNAs such as COLDAIR [138], COOLAIR [139], and COLDWRAP [140] also participate in the recruitment of PRC2 to *FLC* chromatin. Besides, R-loop,

a DNA-RNA hybrid structure, is also suggested to promote PRC2 recruitment in *Drosophila*, mammalian, and Arabidopsis through similar mechanisms [141–143].

PRC2 core complex can gain functional specificity by associating with partner proteins that can regulate PRC2 methyltransferase activity and/or its recruitment to the specific sites of the genome [126]. This gaining function via the associated factors mechanism was also conserved across species (Figure 1). In *Drosophila*, Polycomb-like (Pcl)'s binding to PRC2 is essential for high levels of H3K27me3 [144]. Pcl, Jarid2, and Jing also contribute to the recruitment of PRC2 to PRE [111]. In mammals, the N-terminal portion of SUZ12 interacts with various additional associated factors, leading to at least two different PRC2 assemblies, PRC2.1 (core assembly plus PCL1–3, EPOP, and PALI1/2) and PRC2.2 (core assembly plus AEBP2 and JARID2) with distinct biochemical properties [145]. Similarly, three PRC2 complexes, FIS-PRC2, EMF-PRC2, and VRN-PRC2, were found in Arabidopsis to control various developmental processes [146,147]. In human, associated factors including PCLs, AEBP2, and JARID2 play a role in DNA binding and complex recruitment; AEBP2 and JARID2 in complex stabilization; JARID2 and PALI1 in allosteric autoactivation; and PHF1/19, JARID2 and AEBP2 in binding to modified histones [83]. Similarly, in Arabidopsis, a complex protein–protein interaction network is formed through interactions among PRC2 core components and associated factors that share both functional and structural homology with counterparts found in other species. Those factors can be classified as transcriptional activators and repressors (VAL1, BPC1, BPC3, AS1, AS2, etc.), DNA replication factors (FAS1, ESD7, EOL1), histone modification readers (LHP1, VIN3, VRN5, SHL, EBS), histone modifiers (CU11, ATX1, HDA19), and ubiquitin C-terminal hydrolases [83,148]. Arabidopsis VIN3 and its homolog VRN5 are required in vernalization-induced histone modification and have multiple conserved domains with human PCL proteins that regulate the PRC2 recruitment and enzymatic activity. Arabidopsis shares multiple Jarid2 homologs (JM14, JM16, JM17) with animals [149]. Arabidopsis LHP1, a transcription

modulator in PRC2 is a homolog of human HP1 that is considered exclusive in PRC1 that participates in transcriptional regulation. UBP12 and UBP13 that remove mono-ubiquitination of histone H2A in Arabidopsis show sequence similarity to human ubiquitin-specific protease USP7 [150–156].

PRC1 is less evolutionarily conserved between animals and plants because of the divergence events early in animal evolution [157]. The core subunits of canonical PRC1 (cPRC1) found in *Drosophila* are the E3 (Sce/dRING and Psc/Su(z) 2) monoubiquitin ligase module and associated Pc, Ph, and SCM [150]. Ph and Pc are essential for higher-order chromatin folding of *Hox* clusters and Pc specifically binds to H3K27me3 [158,159]. SCM is a functional link and a key mediator between PRC1 and PRC2 in transcriptional silencing [160]. Most of the components of cPRC are missing in plants, but in Arabidopsis, three homologs of Psc (AtBMI1a, AtBMIb, and AtBMI1c) and two homologs of dRING (AtRING1a and AtRING1b) have been identified [146], which are a PRC1-like complex that incorporates some PRC1 homologs, but also plant-specific factors [147]. The previous model describes that the cPRC1 complex recognizes methylation of the histone by PRC2 and in turn maintains the trimethylation mediated by PRC2 and catalyses suppressive mono-ubiquitination. The same H3K27me3-governing function at major PcG targets, such as *FLC*, is executed by Arabidopsis LHP1, an H3K27me3-binding factor [161]. Evidence also indicates that the activity of PRC1 may be independent of PRC2 [162].

Similarity of histone modifier and chromatin remodeler between animals and plants: TrxG and trithorax complexes

Trithorax group (TrxG) proteins function to establish, maintain, and transmit the gene expression patterns mainly through methylating lysine 4 of histone H3 (H3K4) and remodelling nucleosomes [163,164]. TrxG is first identified in *Drosophila* as the transcription activator of *Hox* gene, which has the opposite to transcription-suppressive PcG proteins [165–167]. Similar to PcG, homologs of the TrxG were found in broader

species, across yeast, animals, and plants [168–171], and they can be classified into three groups based on their functions [172]. The first group is the SET domain-containing factors that are involved in transcriptional activation, and more specifically, related to di- and tri-methylation of H3K4 [133,173,174]. Such SET-containing proteins are found in both animals and plants, such as SET1 and TRX in *Drosophila*, MLLs in humans, SET1A/B and MLLs in mice, and ATXs and SDGs in Arabidopsis [163,175,176] (Figure 2). Another group is the ATP-dependent chromatin-remodelling factors, such as BRM in *Drosophila*, BRG1 and hBRM in humans, and BRM and SYD in Arabidopsis [163,177,178]. Proteins from the third group can combine with certain DNA sequences, and it consists of some unclassified proteins from the former two groups and some histone modifiers and chromatin remodellers [163]. Collectively, TrxG proteins are assembled into multiprotein complexes to exert their functions as a whole complex [163]. A major group of TrxG complex includes COMPASS and COMPASS-like, which catalyses the trimethylation of the histone H3 lysine 4 (H3K4me3). Another group, represented by SWI/SNF, mediates ATP-dependent chromatin remodelling and nucleosome structure altering [172,179].

COMPASS complexes, also known as SET1A/B, mediate H3K4me3 as an activation mark at non-transcribing regions of the gene, including 5' regions and promoters in animals and plants [180–182]. COMPASS has a core of four (in Arabidopsis) or five (in human, mice, and *Drosophila*) component modules (Figure 2), which is necessary for the complex assembly of COMPASS in both mammals and plants [183]. The COMPASS core is highly conserved in almost all higher eukaryotes [184]. In human, SET domain-containing protein SET-2/SET1 is a conserved methyltransferase [184,185], Ash2 is required for H3K4me3, DPY30 is required for H3K4me2 and H3K4me3, and RBBP5 and WDR5 are required for the assembly of COMPASS complex. The *Drosophila*'s homologs are SET1/2 (dSet), WDR5 (Wds), Ash2 (Ash2), RbBP5 (Rbbp5), and Dpy30 (Dpy30), respectively. In mice, they are SET1/2 (SET1A/B), WDR5 (WDR5), Ash2 (ASH2L), RbBP5 (RBBP5), and

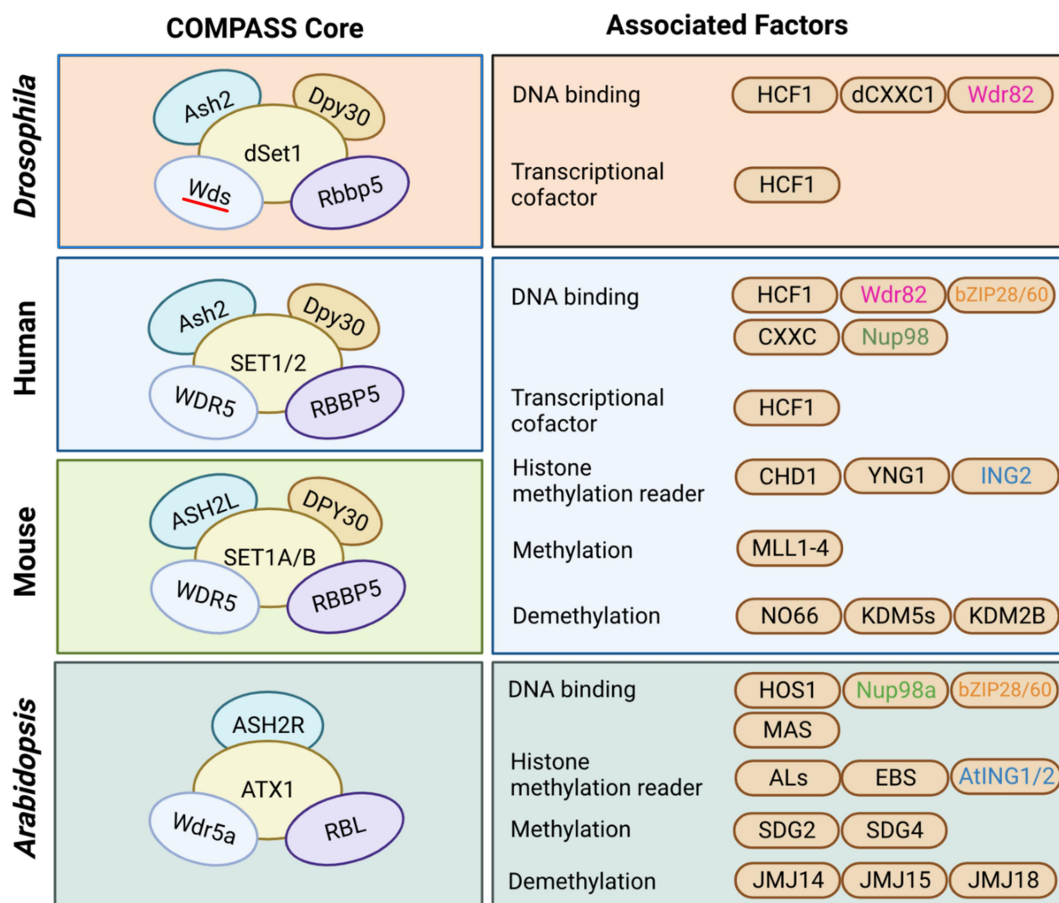


Figure 2. Comparison of COMPASS core components and their associated factors among *Drosophila*, mouse, human, and Arabidopsis. The interaction of COMPASS core components is represented among *Drosophila*, mouse, human, and Arabidopsis. The core components used the same colour between the organisms indicate corresponding orthologous proteins. The associated factors are classified by their functions, and their colour-coding indicates orthologous proteins between organisms.

Dpy30 (DPY30) [186]. Arabidopsis retains four out of five homologs of the human COMPASS complex core components [182]. The homolog of human SET1/2, ASH2, and RbBP5 in Arabidopsis are ATX1, ASH2R, and RBL. The Arabidopsis homologs for human WDR5 are WDR5a and WDR5b. WDR5a binds to histone H3 tails and can form the core complex with ASH2R and RBL together [182], while WDR5b does not interact with RBL to form the structural platform of the COMPASS complex.

COMPASS also acquires certain functions by interacting with associate factors (Figure 2). In both animals and plants, several associated factors of the COMPASS complex are responsible for the writing, reading, and erasing of the histone methylation [187]. In mammals, MLL1-4 performs as the writer of histone methylation [187-190] with SDG2 and SDG4 acting as the

writers in Arabidopsis [188,189]. The readers in Arabidopsis are ALs, AtING1,2, EBS, etc., whereas in mammals they are CHD1, YNG1, ING2, etc. [187,191,192]. NO66, KDM2B, and KDM5A/B/C/D are the erasers in animals, while JMJ14, JMJ15, and JMJ18 are the demethylases in Arabidopsis [187,193-196]. Some associated factors mediate the recruitment of the COMPASS component. For example, bZIP28 and bZIP60 can interact with WDR5 and Ash2 respectively in Arabidopsis, and with SWD3 and BRE2 in mammals, to direct the COMPASS complex to the promoters of unfolded protein response (UPR) genes [197,198]. In Arabidopsis, MAS can recruit COMPASS to MAF4 by interacting with WDR5a [199]. In mammals, Nup98, is a subset of nucleoporins that can recruit COMPASS complex to promoters to regulate H3K4me3 [200]. In humans, HCF1, CxxC,

Nup98, Wdr82, bZIP28, and bZIP60 facilitate DNA binding, whose functional homologs in *Drosophila* are HCF1, dCxxC1, and Wdr82 while in Arabidopsis are HOS1, MAS bZIP28 and bZIP60.

SWI/SNF functions to mediate ATP-dependent chromatin remodelling, one of the basic mechanisms required for the compaction of the nucleosome [201–203]. SWI/SNF complexes were originally discovered in *Saccharomyces cerevisiae*, and similar complexes were characterized later in *Drosophila* [204,205]. Conserved among eukaryotes, SWI/SNF complexes utilize the energy from ATP hydrolysis to slide the DNA around the nucleosome, reposition the nucleosome and allow transcriptional apparatus to be recruited to the DNA [172]. The catalytic core subunit of the complex shows cross-species conservation in mammals and Arabidopsis. BRM/BRG1 in mammals is homologous to SWI/SNF in yeast [206]. In human, one catalytic subunit, BRG1 or BRM, associates with multiple non-catalytic subunits, forming three complexes: canonical BAF, Polybromo-associated BAF, and noncanonical BAF [207]. BAFs are important in multiple cellular processes, and their dysfunctions are the risk factors for cancer [203,208]. Although there is no plant endogenous SWI/SNF complex, two homologs to SWI/SNF catalytic subunits, BRM and SYD, are found in Arabidopsis and they participate in the development and phytohormone responses [209].

Initially discovered in *Drosophila*, PcG proteins collectively act as repressors of *Hox* genes, while the gene activation is mediated by TrxG proteins [165]. Later work has broadened the spectrum of regulatory elements that can recruit PcG and TrxG factors, which are named PcG and TrxG response elements (PREs and TREs) [210–212]. Trx associates with CREB-binding protein (CBP), mediates H3K27 acetylation, and prevents H3K27me3 at Polycomb target genes [213]. TrxG protein Ash1 antagonizes PcG-mediated silencing via H3K36 dimethylation at *HOX* genes [214,215]. On contrary, Pc inhibits the H3K27 acetylation by blocking the acetyltransferase activity of CBP and consequently favours H3K27 methylation [216]. With the opposite regulatory functions, the TrxG and PcG complexes compete dynamically, maintaining the

balance between genetic repression and activation in various organisms. In Arabidopsis, PcG protein CLF acts as a transcriptional repressor of the *AGAMOUS* (AG) gene through H3K27me3 while TrxG protein ATX1 acts to keep high-level AG transcription, suggesting the antagonist function of ATX1 [217].

Why Arabidopsis: open questions from environmental signal to histone box

More recent studies on the epigenetic components of cellular memory and plasticity have revealed that control of chromatin status is not limited to the modifications performed by the multiple complexes mentioned in the previous section. The chromatin structure is also shaped by the histone sliding and the incorporation of histone variants (e.g., H2A.Z) mediated by four subfamilies of ATP-dependent chromatin-remodelling enzymes: imitation switch (ISWI), chromodomain helicase DNA-binding (CHD), SWI/SNF, and INO80 [218,219]. Moreover, the histone modifications and DNA methylation are found to be interdependent processes, which together determine chromatin status [220,221]. Although histone modifications are highly conserved in the evolution of eukaryotes, DNA methylation has unique mechanisms observed in mammals and plants [222,223]. Mammalian genomes show particularly high CG methylation levels (70%–80% in general) [222]. CG methylation is established through DNA methyltransferases 3 alpha (DNMT3A) and DNMT3B and is maintained by DNMT1 upon replication [224]. CG methylation is implicated in the repressed gene expression in the promoter region but is associated with actively transcribed gene bodies [224]. In Arabidopsis, DNA methylation occurs in three cytosine sequence contexts: CG, CHG, and CHH (H=A, T or C). CG methylation is maintained by METHYLTRANSFERASE 1 (MET1), an orthologue of the mammalian DNMT1 [225]. CHG methylation and asymmetric CHH methylation are site-specifically methylated by CHROMOMETHYLASE 2 (CMT2) and METHYLTRANSFERASE 2 (DRM2) [226]. Gene body methylation is most frequent in constitutively and moderately expressed genes but not in genes with variable expression [227,228]. These

findings indicate that there are multiple species-specific mechanisms evolved that intertwined together to control the chromatin status within the nucleus. Meanwhile, outside the nucleus, the signalling pathway connecting environmental stimuli and the intracellular histone modifications has received more attention in recent years, becoming a significant factor in understanding the field of biology, disease, and evolution [229–231]. However, how environmental signals are transferred from the plasma membrane to the nucleus has been still uncovered. Although there have been efforts to identify the upstream regulators outside the ‘histone box’ using large-scale data analyses, grossly perturbed epigenomes in each mutant make it hard to identify if the change of histone modifications compared to the wild type is direct effects [232].

Identification of the upstream regulatory components using mutant can be challenging in some animal models due to lethality and/or sterility. For example, loss-of-function mutations on genes encoding Polycomb proteins frequently cause lethality in mammalian development [233,234]. Homozygous mutations on *EZH2*, *EED*, and *SUZ12*, the core components of the PRC2 complex, lead to early embryonic lethality in mice, and mutants in H3K4 demethylase or methyltransferase produce initially healthy animals which give birth to an increasing percentage of sterile offspring after each generation [232,235,236]. Furthermore, controlling environmental input to establish a quantitative system in animal models and the corresponding phenotype caused by the same input would be challenging.

Considering these aspects, *Arabidopsis thaliana* is a powerful model to identify upstream regulators of histone modifications. Moreover, the vernalization system in *Arabidopsis* brings advantages to reveal the relationship between the controllable environment input (duration of vernalization) correlate phenotypic change, chromatin modifications, and following gene expression. *FLC* locus is the key to conferring the cold-induced chromatin modification via the balance of the chromatin modification between PRC2 and TrxG, resulting in corresponding flowering time phenotypes [122,237,238]. Before cold exposure, the FRIGIDA (FRI) complex activates *FLC* together

with the deposition of H3K4me3 mediated by TrxG, leading to the consequent suppression of flowering and delayed flowering time [239–241] (Figure 3, left panel). During cold exposure, the expression of a cold-specific PRC2 component, *VIN3*, is gradually induced, and *FLC* expression is gradually reduced via increased H3K27me3 deposition on *FLC* chromatin. After vernalization, regardless of *FLC* activation by FRI, PRC2-mediated *FLC* silencing is stably maintained even when the plant grows in warm conditions because of the function of PRC1 [242] (Figure 3, middle panel). The *vin3* mutant is vernalization insensitive, resulting in delayed flowering after vernalization [243]. In addition, three long noncoding RNAs (lncRNAs) generated from *FLC*, *COOLAIR* [138], *COLD AIR* [244], and *COLDWRAP* [140] are up-regulated during vernalization. *VIN3* and three non-coding RNAs, *COLD AIR*, *COOLAIR*, and *COLDWRAP* have essential roles in the PRC2-mediated silencing of *FLC*. Interestingly, the plant homeodomain (PHD) finger domain of *VIN3* is found to be critical for PRC2 recruitment in human cells [155,245]. Human polycomb-like (PCL) proteins, hPCL1, hPCL2, and hPCL3, contain plant homeodomain (PHD) finger domains, and participate in PRC2 recruitment to the target locus for the increase of H3K27me3 [245–248]. A human H3K27me3-methyltransferase, *EZH2*, binds to PHD finger domain of hPCL1 (PHF1) and methylates its target genes [155,245]. Therefore, *VIN3* in the PRC2 complex is highly likely to have a homologous function to PCLs in human.

In a previous genetic screening, multiple upstream regulators outside the histone box in *Arabidopsis* were identified [249]. *VIN3* is not expressed without cold exposure, thus the authors performed mutagenesis to isolate mutants that induce the expression of *VIN3* without vernalization using *VIN3-GUS* transgenic reporter lines treated ethyl methane sulphonate for random mutations [249]. *VIN3* is a PRC2 component that suppresses *FLC* expression with cold exposure, the early flowering mutants with less expression of *FLC* are strong candidates for negative regulators of *VIN3* without vernalization. If the candidates negatively regulate *VIN3*, the *vin3/candidate* double mutant

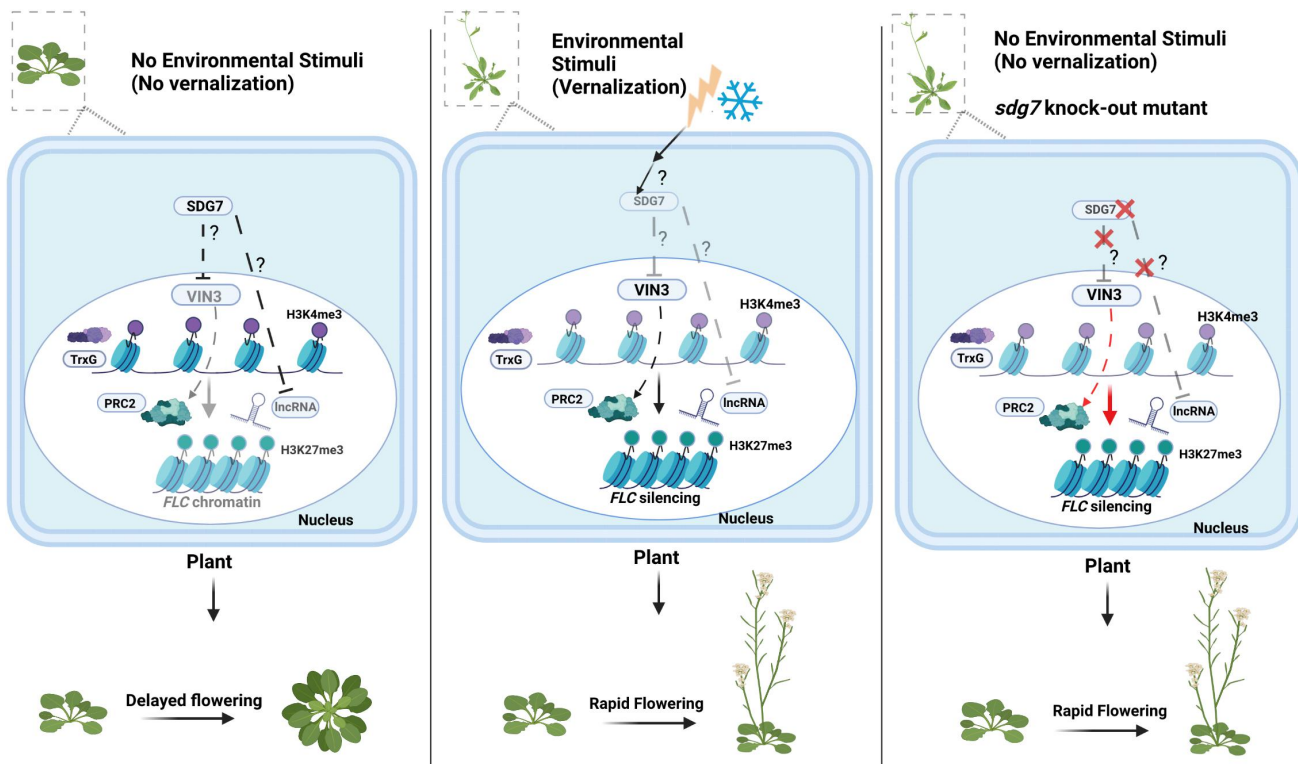


Figure 3. Vernalization pathway without or with prolonged cold exposure as an environmental stimulus in *Arabidopsis*. Quantitative correlations between the duration of cold exposure and the histone modifications results in changes of *FLC* chromatin status and its expression. Because the flowering time in vernalization pathway depends on *FLC* expression, the flowering time is switched based on the quantitative environmental input (left and middle panel). The *sdg7* mutant is unable to suppress the expression of the PRC2 component VIN3, as well as the lncRNAs COOLAIR and COLDAIR. This results in a decrease in *FLC* expression via an increase in H3K27me3, leading to rapid flowering without cold exposure (right panel).

should show the *vin3* mutant phenotype (delayed flowering after vernalization). Therefore, a mutant of the upstream negative regulator of VIN3 should exhibit: (1) rapid induction of the VIN3-*GUS* reporter and thus native VIN3, (2) rapid flowering without cold exposure depending on the chromatin status of *FLC*, and (3) vernalization-insensitive flowering with the loss of VIN3 in a double mutant. This experiment identified a mutant that meets the three criteria, the causative locus encodes a methyltransferase, SET DOMAIN GROUP7 (SDG7), and the mutant has a missense mutation in a highly conserved proline residue (P199L) in SET domain-containing proteins (*sdg7-1*) [249]. Furthermore, lncRNAs, COOLAIR, and COLDAIR were also induced in *sdg7* mutants without cold exposure. Increased deposition of H3K27me3 on *FLC* chromatin, resulting in decreased *FLC* expression, suggesting that SDG7 is an upstream regulator of VIN3 (Figure 3, right

panel). Interestingly, SDG7 does not have a nuclear localization signal, and is not targeted at the nucleus [249,250]. Lys methylation activity of SDG7 was detected *in vitro* with an aquaporin substrate targeted to the plasma membrane [250], but no histone methyltransferase activity was detected *in vitro*. These data suggest that SDG7 is involved in the vernalization via Lys methylation of extranuclear proteins as a signalling process to control PRC2-mediated *FLC* silencing. Although Lys methylation of extranuclear proteins is widespread, its biological significance is unclear [251–253]. *Arabidopsis* SDG7 is an example of a potential upstream regulator of the PRC2 outside of the histone box (Figure 4, left panel), suggesting that *Arabidopsis* plant can be a model to identify upstream regulators from the environment to the nucleus using the controllable vernalization system. In plants, the quantitatively controlled input level tightly associates with the chromatin modification

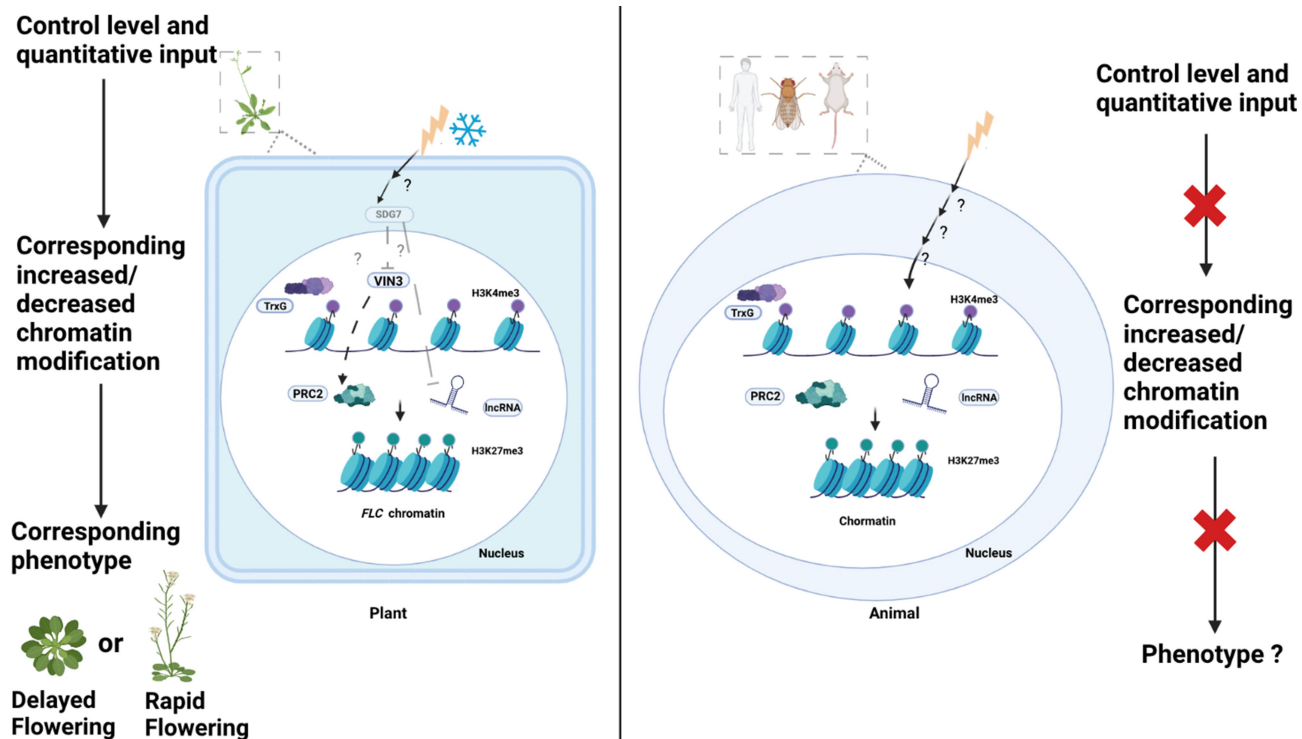


Figure 4. Comparison of regulations of the histone modification pathway between plants and animals in three aspects: environmental input, corresponding histone modifications, and resulting phenotypic changes.

change that is revealed by the corresponding phenotypes, whereas in animals, the lack of such association hinders such phenotypic screening (Figure 4, right panel).

Discussion

The involvements of epigenetic regulation, especially histone modifications in the response to environmental changes, have been revealed, but there are missing pieces that demonstrate how the extracellular signals are transduced into the nucleus. *Arabidopsis thaliana* can serve as exceptional experimental subjects to study the components of epigenetic signalling induced by environmental cues. It will also be a powerful model that opens up more opportunities for the understanding of histone modifications, especially to understand how upstream regulators located outside of the nucleus work to transduce environmental signals. Because of the high similarity of histone proteins, chromatin modification modules, the key protein complexes and associated components in animals and plants, the upstream signalling and its components could be conserved at the

molecular level. In addition, findings in animal models have already provided numerous key inspirations guiding the studies in plant, i.e., the discovery of Polycomb and Trithorax proteins which are first identified in *Drosophila* [123]. The findings in *Arabidopsis thaliana*, vice versa, have great potential to provide cues for animal studies. *Arabidopsis* has made a significant contribution to the field of small RNA and related epigenetic mechanisms, such as the finding of the Argonaute gene and RNA-directed DNA methylation [254]. The chromatin remodelling factor DECREASE IN DNA METHYLATION 1 (DDM1) was first discovered in *Arabidopsis* mutant screenings and later mammalian ortholog Lsh was found with the same effect in centromeric methylation [254–256]. These efforts suggest that using *Arabidopsis* as a model organism for epigenetic studies can also make a contribution to research related to human health. Epigenetic switches of mammalian stem cells have been extensively studied. However, stem cell research using plants can be beneficial as an extension of mammalian research. The animal and plant kingdoms have diverged since 16 billion years ago, and

the only conserved protein that specifically participated in the stem cell function among these two kingdoms is the retinoblastoma-related (RBR) protein [257,258]. Interestingly, their epigenetic modifying patterns sustaining the pluripotency of stem cells are dramatically parallel [259]. These patterns include the alteration of nucleosome conditions, histone acetylation or deacetylation, H3K27 or H3K4 methylation, chromatin compaction, and chromatin assembly or disassembly, where corresponding proteins and regulators are functioning following conserved general patterns in the epigenetic regulations of stem cells [259]. Investigations into stem cell reprogramming properties in plants and animals have also revealed the high conservation in their core regulatory logic and general machinery [257,258]. In addition, plant stem cells are more dependent on cell-to-cell communication, which is consistent with their need for receiving signals from the environment [258]. Combining the fact that differentiation and dedifferentiation naturally occur throughout their life span in plant stem cells, it can be a beneficial experimental subject to investigate the basic mechanisms of pluripotency restoration of cells [257,260].

The ease of cultivation provides another reason for applying plants as the model organisms for epigenetic research such as the short generation time of fewer than 2 months, the smaller size that makes it easier to handle larger scale samples in growth facilities, and the potential to maintain epigenomes between generations on account of the self-pollinations [261]. Furthermore, the small pan-genome size (~135Mb) with a significantly low content of interspersed repetitive DNA sequences may minimize redundancy issues at a genetic level [262–264]. Easier seed storage is also beneficial compared to the maintenance of mammalian offsprings, which require extensive care to be kept alive and healthy under genetic mutation or experimental treatments.

Acknowledgments

This study was supported by the Startup fund at Duke Kunshan University, 2022 Summer Research Scholarship

fund, 2021 Interdisciplinary Seed Grant, Syneer and Wang-Cai Seed Grant 2022, and Wang-Cai Biochemistry Lab Grant

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The work was supported by the Duke Kunshan University [Startup].

Author contributions

Y. Y., S. W., Z. W., R. G., and J. L. contributed to the manuscript writing.

Data availability statement

histone sequences used in this study are available at the UniProt (<https://www.uniprot.org/>) [265] and NCBI RefSeq (<https://www.ncbi.nlm.nih.gov/refseq/>) [266]. Accession numbers are listed in Supplement Table 1.

References

- [1] Audia JE, Campbell RM. Histone modifications and cancer. *Cold Spring Harb Perspect Biol.* 2016;8(4):a019521.
- [2] Martire S, Banaszynski LA. The roles of histone variants in fine-tuning chromatin organization and function. *Nat Rev Mol Cell Biol.* 2020;21(9):522–541.
- [3] Van Holde KE, Sahasrabudhe CG, Shaw BR. A model for particulate structure in chromatin. *Nucleic Acids Res.* 1974;1(11):1579–1586.
- [4] Olins DE, Olins AL. Chromatin history: our view from the bridge. *Nat Rev Mol Cell Biol.* 2003;4(10):809–814.
- [5] Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res.* 2011;21(3):381–395.
- [6] Koster MJ, Snel B, Timmers HT. Genesis of chromatin and transcription dynamics in the origin of species. *Cell.* 2015;161(4):724–736.
- [7] Yun M, Wu J, Workman JL, et al. Readers of histone modifications. *Cell Res.* 2011;21(4):564–578.
- [8] Gillette TG, Hill JA. Readers, writers, and erasers: chromatin as the whiteboard of heart disease. *Circ Res.* 2015;116(7):1245–1253.
- [9] Handy DE, Castro R, Loscalzo J. Epigenetic modifications: basic mechanisms and role in cardiovascular disease. *Circulation.* 2011;123(19):2145–2156.
- [10] Portela A, Esteller M. Epigenetic modifications and human disease. *Nature Biotechnol.* 2010;28(10):1057–1068.

- [11] Simmons D. Epigenetic influence and disease. *Nat Educ.* 2008;1(1):6.
- [12] Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. *Nat Rev Genet.* 2016;17(8):487–500.
- [13] Zhao Z, Shilatifard A. Epigenetic modifications of histones in cancer. *Genome Bio.* 2019;20(1):245.
- [14] Alaskhar Alhamwe B, Khalaila R, Wolf J, et al. Histone modifications and their role in epigenetics of atopy and allergic diseases. *Allergy Asthma Clin Immunol.* 2018;14(1):39. DOI:10.1186/s13223-018-0259-4
- [15] Kuehner JN, Bruggeman EC, Wen Z, et al. Epigenetic regulations in neuropsychiatric disorders. *Front Genet.* 2019;10:268.
- [16] Ankeny RA, Leonelli S. What's so special about model organisms? *Stud Hist Philos Sci A.* 2011;42(2):313–323.
- [17] Davis RH. The age of model organisms. *Nat Rev Genet.* 2004;5(1):69–76.
- [18] Bonini NM, Berger SL. The sustained impact of model organisms—in genetics and epigenetics. *Genetics.* 2017;205(1):1–4.
- [19] Anderson KV, Ingham PW. The transformation of the model organism: a decade of developmental genetics. *Nature Genet.* 2003;33(S3):285–293.
- [20] Halim MA, Tan FHP, Azlan A, et al. Ageing, *drosophila melanogaster* and epigenetics. *Malays J Med Sci.* 2020;27(3):7–19.
- [21] Blewitt M, Whitelaw E. The use of mouse models to study epigenetics. *Cold Spring Harb Perspect Biol.* 2013;5(11):a017939.
- [22] Herz HM, Madden LD, Chen Z, et al. The H3K27me3 demethylase dUTX is a suppressor of notch- and Rb-dependent tumors in *Drosophila*. *Mol Cell Biol.* 2010;30(10):2485–2497.
- [23] Copur Ö, Müller J. The histone H3-K27 demethylase Utx regulates HOX gene expression in *Drosophila* in a temporally restricted manner. Vol. 140. Cambridge, England: Development; 2013. pp. 3478–3485.
- [24] Shalaby NA, Sayed R, Zhang Q, et al. Systematic discovery of genetic modulation by Jumonji histone demethylases in *Drosophila*. *Sci Rep.* 2017;7(1):5240.
- [25] Dudley KJ, Revell K, Whitby P, et al. Genome-Wide Analysis in a Murine Dnmt1 Knockdown Model Identifies Epigenetically Silenced Genes in Primary Human Pituitary Tumors. In: *Molecular cancer research.* 2008;6(10):1567–1574.
- [26] Cooper A, Butto T, Hammer N, et al. Inhibition of histone deacetylation rescues phenotype in a mouse model of birk-barel intellectual disability syndrome. *Nat Commun.* 2020;11(1):480. DOI:10.1038/s41467-019-13918-4
- [27] Anreiter I, Kramer JM, Sokolowski MB Epigenetic mechanisms modulate differences in *Drosophila* foraging behavior. *Proc Natl Acad Sci USA.* 2017;114:12518–12523.
- [28] Ordway AJ, Teeters GM, Weasner BM, et al. A multi-gene knockdown approach reveals a new role for Pax6 in controlling organ number in *Drosophila*. Cambridge, England: Development; 2021. p. 148.
- [29] Bult CJ, Blake JA, Smith CL, et al. Mouse genome database (MGD) 2019. *Nucleic Acids Res.* 2019;47(D1):D801–d6.
- [30] Bastow R, Mylne JS, Lister C, et al. Vernalization requires epigenetic silencing of *FLC* by histone methylation. *Nature.* 2004;427(6970):164–167.
- [31] Heard E, Martienssen Robert A. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell.* 2014;157(1):95–109.
- [32] Xavier MJ, Roman SD, Aitken RJ, et al. Transgenerational inheritance: how impacts to the epigenetic and genetic information of parents affect offspring health. *Human Reproduction Update.* 2019;25(5):519–541.
- [33] Liu F, Wang L, Perna F, et al. Beyond transcription factors: how oncogenic signalling reshapes the epigenetic landscape. *Nat Rev Cancer.* 2016;16(6):359–372.
- [34] Zannas AS, Arloth J, Carrillo-Roa T, et al. Lifetime stress accelerates epigenetic aging in an urban, African American cohort: relevance of glucocorticoid signaling. *Genome Bio.* 2015;16(1):266. DOI:10.1186/s13059-015-0828-5
- [35] Tough DF, Tak PP, Tarakhovsky A, et al. Epigenetic drug discovery: breaking through the immune barrier. *Nat Rev Drug Discov.* 2016;15(12):835–853.
- [36] Bagot RC, Labonté B, Peña CJ, et al. Epigenetic signaling in psychiatric disorders: stress and depression. *Dialogues Clin Neurosci.* 2014;16(3):281–295.
- [37] Long HK, Sims D, Heger A, et al. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. *Elife.* 2013;2:e00348.
- [38] Radford EJ. Exploring the extent and scope of epigenetic inheritance. *Nat Rev Endocrinol.* 2018;14(6):345–355.
- [39] Fuchs J, Demidov D, Houben A, et al. Chromosomal histone modification patterns – from conservation to diversity. *Trends Plant Sci.* 2006;11(4):199–208.
- [40] Feng S, Jacobsen SE. Epigenetic modifications in plants: an evolutionary perspective. *Curr Opin Plant Biol.* 2011;14(2):179–186.
- [41] Rhee SY, Beavis W, Berardini TZ, et al. The *Arabidopsis* information resource (TAIR): a model organism database providing a centralized, curated gateway to *Arabidopsis* biology, research materials and community. *Nucleic Acids Res.* 2003;31:224–228.
- [42] O'Malley RC, Barragan CC, Ecker JR. A user's guide to the *Arabidopsis* T-DNA insertion mutant collections. *Methods in molecular biology.* In: Jose M. Alonso, Anna N. Stepanova, editors. *Plant Functional Genomics.* Vol. 1284. Clifton, NJ: Springer 2015pp. 323–342
- [43] Alonso JM, Stepanova AN, Leisse TJ, et al. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*.

- Science. 2003;301(5633):653–657. DOI:10.1126/science.1086391
- [44] Woody ST, Austin-Phillips S, Amasino RM, et al. The *WiscDslox* T-DNA collection: an Arabidopsis community resource generated by using an improved high-throughput T-DNA sequencing pipeline. *J Plant Res.* 2007;120(1):157–165.
- [45] Sessions A, Burke E, Presting G, et al. A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell.* 2002;14(12):2985–2994. DOI:10.1105/tpc.004630
- [46] Pikaard CS, Mittelsten Scheid O. Epigenetic regulation in plants. *Cold Spring Harb Perspect Biol.* 2014;6(12):a019315.
- [47] Hauser MT, Aufsatz W, Jonak C, et al. Transgenerational epigenetic inheritance in plants. *Biochim Biophys Acta.* 2011;1809(8):459–468.
- [48] Pfluger J, Wagner D. Histone modifications and dynamic regulation of genome accessibility in plants. *Curr Opin Plant Biol.* 2007;10(6):645–652.
- [49] Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet.* 2012;13(2):97–109.
- [50] Skinner MK, Guerrero-Bosagna C. Environmental signals and transgenerational epigenetics. *Epigenomics.* 2009;1(1):111–117.
- [51] Sheldon CC, Rouse DT, Finnegan EJ, et al. The molecular basis of vernalization: the central role of *FLOWERING LOCUS C*. *Proc Natl Acad Sci USA.* 2000;97:3753–3758.
- [52] De Lucia F, Crevillen P, Jones AME, et al. A PHD-Polycomb Repressive Complex 2 triggers the epigenetic silencing of *FLC* during vernalization. *Proc Natl Acad Sci USA.* 2008;105:16831–16836.
- [53] Yang H, Howard M, Dean C. Antagonistic roles for H3K36me3 and H3K27me3 in the cold-induced epigenetic switch at *Arabidopsis FLC*. *Curr Biol.* 2014;24(15):1793–1797.
- [54] Mariño-Ramírez L, Jordan IK, Landsman D. Multiple independent evolutionary solutions to core histone gene regulation. *Genome Bio.* 2006;7(12):R122.
- [55] Mariño-Ramírez L, Kann MG, Shoemaker BA, et al. Histone structure and nucleosome stability. *Expert Rev Proteomics.* 2005;2(5):719–729.
- [56] Bönisch C, Hake SB. Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res.* 2012;40(21):10719–10741.
- [57] Stevens Kathryn M, Swadling Jacob B, Hoher A, et al. Histone variants in archaea and the evolution of combinatorial chromatin complexity. *Proc Natl Acad Sci USA.* 2020;117:33384–33395.
- [58] Joseph SR, Pálffy M, Hilbert L, et al. Competition between histone and transcription factor binding regulates the onset of transcription in zebrafish embryos. *Elife.* 2017;6:e23326.
- [59] Hyland EM, Cosgrove MS, Molina H, et al. Insights into the role of histone H3 and histone H4 core modifiable residues in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 2005;25(22):10060–10070.
- [60] Talbert PB, Henikoff S. Histone variants at a glance. *J Cell Sci.* 2021;134(6):134.
- [61] Long M, Sun X, Shi W, et al. A novel histone H4 variant H4G regulates rDNA transcription in breast cancer. *Nucleic Acids Res.* 2019;47(16):8399–8409. DOI:10.1093/nar/gkz547
- [62] Waterborg JH. Evolution of histone H3: emergence of variants and conservation of post-translational modification sites 1 this article is part of special issue entitled asilomar chromatin and has undergone the journal's usual peer review process. *Biochem Cell Biol.* 2012;90(1):79–95.
- [63] Szenker E, Ray-Gallet D, Almouzni G. The double face of the histone variant H3.3. *Cell Res.* 2011;21(3):421–434.
- [64] Qanmber G, Ali F, Lu L, et al. Identification of histone H3 (HH3) genes in *Gossypium hirsutum* revealed diverse expression during ovule development and stress responses. *Genes (Basel).* 2019;10(5):355.
- [65] Muhire BM, Booker MA, Tolstorukov MY. Non-neutral evolution of H3.3-encoding genes occurs without alterations in protein sequence. *Sci Rep.* 2019;9(1):8472.
- [66] Stroud H, Otero S, Desvoyes B, et al. Genome-wide analysis of histone H3.1 and H3.3 variants in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA.* 2012;109:5370–5375.
- [67] Ahmad K, Henikoff S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Molecular Cell.* 2002;9(6):1191–1200.
- [68] Ramachandran S, Henikoff S. Replicating nucleosomes. *Sci Adv.* 2015;1(7):e1500587.
- [69] Schwartz BE, Ahmad K. Transcriptional activation triggers deposition and removal of the histone variant H3.3. *Genes & deve.* 2005;19(7):804–814.
- [70] Tagami H, Ray-Gallet D, Almouzni G, et al. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell.* 2004;116(1):51–61.
- [71] Jacob Y, Bergamin E, Donoghue MTA, et al. Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science.* 2014;343(6176):1249–1253. DOI:10.1126/science.1248357
- [72] Almouzni G, Cedar H. Maintenance of epigenetic information. *Csh Perspect Biol.* 2016;8(5):a019372.
- [73] Mendiratta S, Gatto A, Almouzni G. Histone supply: multitiered regulation ensures chromatin dynamics throughout the cell cycle. *J Cell Bio.* 2019;218(1):39–54.
- [74] Shi L, Wen H, Shi X. The histone variant H3.3 in transcriptional regulation and human disease. *J Mol Biol.* 2017;429(13):1934–1945.
- [75] Clément C, Orsi GA, Gatto A, et al. High-resolution visualization of H3 variants during replication reveals

- their controlled recycling. *Nat Commun.* **2018**;9(1):3181. DOI:10.1038/s41467-018-05697-1
- [76] Piontkivska H, Rooney AP, Nei M. Purifying selection and birth-and-death evolution in the histone H4 gene family. *Mol Biol Evol.* **2002**;19(5):689–697.
- [77] Probst AV, Desvoyes B, Gutierrez C. Similar yet critically different: the distribution, dynamics and function of histone variants. *J Exp Bot.* **2020**;71(17):5191–5204.
- [78] Kliebenstein DJ, Lim JE, Landry LG, et al. Arabidopsis UVR8 regulates ultraviolet-B signal transduction and tolerance and contains sequence similarity to human regulator of chromatin condensation 1. *Plant Physiol.* **2002**;130(1):234–243.
- [79] Luque A, Collepardo-Guevara R, Grigoryev S, et al. Dynamic condensation of linker histone C-terminal domain regulates chromatin structure. *Nucleic Acids Res.* **2014**;42(12):7553–7560.
- [80] Morales V, Richard-Foy H. Role of histone N-terminal tails and their acetylation in nucleosome dynamics. *Mol Cell Biol.* **2000**;20(19):7230–7237.
- [81] Izzo A, Schneider R. Chatting histone modifications in mammals. *Brief Funct Genomics.* **2010**;9(5–6):429–443.
- [82] Miller JL, Grant PA. The role of DNA methylation and histone modifications in transcriptional regulation in humans. *Subcell Biochem.* **2013**;61:289–317.
- [83] Bieluszewski T, Xiao J, Yang Y, et al. PRC2 activity, recruitment, and silencing: a comparative perspective. *Trends Plant Sci.* **2021**;26(11):1186–1198.
- [84] Cui J, Zhang Z, Shao Y, et al. Genome-wide identification, evolutionary, and expression analyses of histone H3 variants in plants. *Bio Med Res Int.* **2015**;2015:341598.
- [85] Sanders D, Qian S, Fieweger R, et al. Histone lysine-to-methionine mutations reduce histone methylation and cause developmental pleiotropy. *Plant Physiol.* **2017**;173(4):2243–2252.
- [86] Sankaran SM, Gozani O. Characterization of H3.3K36M as a tool to study H3K36 methylation in cancer cells. *Epigenetics.* **2017**;12(11):917–922.
- [87] Lowe BR, Maxham LA, Hamey JJ, et al. Histone H3 mutations: an updated view of their role in chromatin deregulation and cancer. *Cancers (Basel).* **2019**;11(5):11.
- [88] Lam UTF, Tan BKX, Poh JJX, et al. Structural and functional specificity of H3K36 methylation. *Epigenetics & Chromatin.* **2022**;15(1):17.
- [89] Kernohan KD, Grynspan D, Ramphal R, et al. H3.1 K36M mutation in a congenital-onset soft tissue neoplasm. *Pediatric Blood & Cancer.* **2017**;64(12):64. DOI:10.1002/pbc.26633
- [90] Lewis PW, Müller MM, Koletsky MS, et al. Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science.* **2013**;340(6134):857–861.
- [91] Chaouch A, Berlandi J, Chen CCL, et al. Histone H3.3 K27M and K36M mutations de-repress transposable elements through perturbation of antagonistic chromatin marks. *Mol Cell.* **2021**;81(23):4876–90.e7. DOI:10.1016/j.molcel.2021.10.008
- [92] Zaghi M, Broccoli V, Sessa A. H3K36 methylation in neural development and associated diseases. *Front Genet.* **2020**;10:10.
- [93] Wollmann H, Stroud H, Yelagandula R, et al. The histone H3 variant H3.3 regulates gene body DNA methylation in *Arabidopsis thaliana*. *Genome Bio.* **2017**;18(1):94. DOI:10.1186/s13059-017-1221-3
- [94] Bush KM, Yuen BT, Barrilleaux BL, et al. Endogenous mammalian histone H3.3 exhibits chromatin-related functions during development. *Epigenetics & Chromatin.* **2013**;6(1):7.
- [95] Sakai A, Schwartz BE, Goldstein S, et al. Transcriptional and developmental functions of the H3.3 histone variant in *drosophila*. *Curr Biol.* **2009**;19(21):1816–1820.
- [96] Jang C-W, Shibata Y, Starmer J, et al. Histone H3.3 maintains genome integrity during mammalian development. *Genes & deve.* **2015**;29(13):1377–1392.
- [97] Lewis EB. A gene complex controlling segmentation in *Drosophila*. *Nature.* **1978**;276(5688):565–570.
- [98] Whitcomb SJ, Basu A, Allis CD, et al. Polycomb Group proteins: an evolutionary perspective. *Trends Genet.* **2007**;23(10):494–502.
- [99] Shaver S, Casas-Mollano JA, Cerny RL, et al. Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga *Chlamydomonas*. *Epigenetics.* **2010**;5(4):301–312.
- [100] Gerlt JA, Babbitt PC. Can sequence determine function? *Genome Bio.* **2000**;1(5):reviews0005.1.
- [101] Tian W, Skolnick J. How well is enzyme function conserved as a function of pairwise sequence identity? *J Mol Biol.* **2003**;333(4):863–882.
- [102] Chua HN, Sung W-K, Wong L. An efficient strategy for extensive integration of diverse biological data for protein function prediction. *Bioinformatics.* **2007**;23(24):3364–3373.
- [103] Cock JM, McCormick S. A large family of genes that share homology with CLAVATA3. *Plant Physiol.* **2001**;126(3):939–942.
- [104] Saudek V, Fraternali F. Cystinosin, MPDU1, sWEETs and KDELR belong to a well-defined protein family with putative function of cargo receptors involved in vesicle trafficking. *PLoS ONE.* **2012**;7(2):e30876.
- [105] Ringrose L, Paro R. Polycomb/Trithorax response elements and epigenetic memory of cell identity. Vol. 134. England: Development (Cambridge; **2007**. pp. 223–232.
- [106] Simon JA, Kingston RE. Mechanisms of Polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol.* **2009**;10(10):697–708.
- [107] Levine Stuart S, Weiss A, Erdjument-Bromage H, et al. The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. *Mol Cell Biol.* **2002**;22(17):6070–6078.

- [108] Laugesen A, Højfeldt JW, Helin K. Molecular mechanisms directing PRC2 recruitment and H3K27 methylation. *Mol Cell*. **2019**;74(1):8–18.
- [109] Uckelmann M, Davidovich C. Not just a writer: pRC2 as a chromatin reader. *Biochem Soc Trans*. **2021**;49(3):1159–1170.
- [110] Lavarone E, Barbieri CM, Pasini D. Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. *Nat Commun*. **2019**;10(1):1679.
- [111] Fischer S, Weber LM, Liefke R. Evolutionary adaptation of the Polycomb repressive complex 2. *Epigenetics & Chromatin*. **2022**;15(1):7.
- [112] Wang L, Joshi P, Miller EL, et al. A role for mono-methylation of histone H3-K27 in Gene Activity in *Drosophila*. *Genetics*. **2018**;208(3):1023–1036.
- [113] Wiles ET, Selker EU. H3K27 methylation: a promiscuous repressive chromatin mark. *Curr Opin Genet Dev*. **2017**;43:31–37.
- [114] Müller J, Hart CM, Francis NJ, et al. Histone methyltransferase activity of a *drosophila* polycomb group repressor complex. *Cell*. **2002**;111(2):197–208.
- [115] Chammas P, Mocavini I, Di Croce L. Engaging chromatin: pRC2 structure meets function. *Br J Cancer*. **2020**;122(3):315–328.
- [116] Murzina NV, Pei X-Y, Zhang W, et al. Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. *Structure*. **2008**;16(7):1077–1085. DOI:10.1016/j.str.2008.05.006
- [117] Nekrasov M, Wild B, Müller J. Nucleosome binding and histone methyltransferase activity of *Drosophila* PRC2. *EMBO Rep*. **2005**;6(4):348–353.
- [118] Ciferri C, Lander GC, Maiolica A, et al. Molecular architecture of human polycomb repressive complex 2. *Elife*. **2012**;1:e00005.
- [119] Vijayanathan M, Trejo-Arellano MG, Mozgová I. Polycomb repressive complex 2 in eukaryotes—an evolutionary perspective. *Epigenomes*. **2022**;6(1):6.
- [120] Cao R, Wang L, Wang H, et al. Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science*. **2002**;298(5595):1039–1043.
- [121] Czermin B, Melfi R, McCabe D, et al. *Drosophila* enhancer of zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal polycomb sites. *Cell*. **2002**;111(2):185–196.
- [122] Schwartz YB, Pirrotta V. Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet*. **2007**;8(1):9–22.
- [123] Pien S, Grossniklaus U. Polycomb group and trithorax group proteins in *Arabidopsis*. *Biochimica Et Biophysica Acta (BBA) - Gene Struct and Expression*. **2007**;1769(5–6):375–382.
- [124] Yang Y, Li G. Post-translational modifications of PRC2: signals directing its activity. *Epigenetics & Chromatin*. **2020**;13(1):47.
- [125] Guo Y, Zhao S, Wang GG. Polycomb Gene Silencing Mechanisms: pRC2 Chromatin Targeting, H3K27me3 ‘ReadoutAnd Phase Separation-Based Compaction. *Trends in Genet: TIG*. **2021**;37(6):547–565.
- [126] Aranda S, Mas G, Di Croce L. Regulation of gene transcription by Polycomb proteins. *Sci Adv*. **2015**;1(11):e1500737.
- [127] Xiao J, Jin R, Yu X, et al. *Cis* and *trans* determinants of epigenetic silencing by Polycomb repressive complex 2 in *Arabidopsis*. *Nature Genet*. **2017**;49(10):1546–1552. DOI:10.1038/ng.3937
- [128] Sing A, Pannell D, Karaïskakis A, et al. A vertebrate polycomb response element governs segmentation of the posterior hindbrain. *Cell*. **2009**;138(5):885–897.
- [129] Woo CJ, Kharchenko PV, Daheron L, et al. A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell*. **2010**;140(1):99–110.
- [130] Lodha M, Marco CF, Timmermans MC. The ASYMMETRIC LEAVES complex maintains repression of KNOX homeobox genes via direct recruitment of Polycomb-repressive complex2. *Genes & deve*. **2013**;27(6):596–601.
- [131] Berger N, Dubreucq B, Roudier F, et al. Transcriptional regulation of *Arabidopsis* *LEAFY COTYLEDON2* involves *RLE*, a *cis*-element that regulates trimethylation of histone H3 at lysine-27. *Plant Cell*. **2011**;23(11):4065–4078.
- [132] De S, Mitra A, Cheng Y, et al. Formation of a polycomb-domain in the absence of strong polycomb response elements. *PLoS Genet*. **2016**;12(7):e1006200.
- [133] Schuettengruber B, Chourrout D, Vervoort M, et al. Genome regulation by polycomb and trithorax proteins. *Cell*. **2007**;128(4):735–745.
- [134] Parreno V, Martinez A-M, Cavalli G. Mechanisms of Polycomb group protein function in cancer. *Cell Res*. **2022**;32(3):231–253.
- [135] Blackledge NP, Klose RJ. The molecular principles of gene regulation by Polycomb repressive complexes. *Nat Rev Mol Cell Biol*. **2021**;22(12):815–833.
- [136] Blackledge NP, Rose NR, Klose RJ. Targeting Polycomb systems to regulate gene expression: modifications to a complex story. *Nat Rev Mol Cell Biol*. **2015**;16(11):643–649.
- [137] Davidovich C, Cech TR. The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2. *RNA*. **2015**;21(12):2007–2022.
- [138] Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*. **2011**;331(6013):76–79.
- [139] Tian Y, Zheng H, Zhang F, et al. PRC2 recruitment and H3K27me3 deposition at *FLC* require FCA binding of *COOLAIR*. *Sci Adv*. **2019**;5(4):eaau7246.
- [140] Kim D-H, Sung S. Vernalization-triggered intragenic chromatin loop formation by long noncoding RNAs. *Dev Cell*. **2017**;40(3):302–12.e4.
- [141] Alecki C, Chiwara V, Sanz LA, et al. RNA-DNA strand exchange by the *drosophilapolycomb* complex PRC2. *Nat Commun*. **2020**;11(1):1781.

- [142] Skourti-Stathaki K, Triglia ET, Warburton M, et al. R-loops enhance polycomb repression at a subset of developmental regulator genes. *Molecular Cell*. 2019;73(5):930–45. e4.
- [143] Ariel F, Lucero L, Christ A, et al. R-Loop mediated trans action of the APOLO long noncoding RNA. *Molecular Cell*. 2020;77(5):1055–65.e4. DOI:10.1016/j.molcel.2019.12.015
- [144] Nekrasov M, Klymenko T, Fraterman S, et al. Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *Embo J*. 2007;26(18):4078–4088.
- [145] Healy E, Mucha M, Glancy E, et al. PRC2.1 and PRC2.2 synergize to coordinate H3K27 trimethylation. *Molecular Cell*. 2019;76(3):437–52.e6. DOI:10.1016/j.molcel.2019.08.012
- [146] Kim D-H, Sung S. Polycomb-mediated gene silencing in *Arabidopsis thaliana*. *Mol Cells*. 2014;37(12):841–850.
- [147] Grossniklaus U, Paro R. Transcriptional silencing by polycomb-group proteins. *Csh Perspect Biol*. 2014;6(11):6.
- [148] Godwin J, Farrona S. The importance of networking: plant polycomb repressive complex 2 and its interactors. *Epigenomes*. 2022;6(1):6.
- [149] Zheng B, Chen X. Dynamics of histone H3 lysine 27 trimethylation in plant development. *Curr Opin Plant Biol*. 2011;14(2):123–129.
- [150] Baile F, Gómez-Zambrano Á, Calonje M. Roles of Polycomb complexes in regulating gene expression and chromatin structure in plants. *Plant Commun*. 2022;3(1):100267.
- [151] Feng J, Lu J. LHP1 could act as an activator and a repressor of transcription in plants. *Front Plant Sci*. 2017;8. DOI:10.3389/fpls.2017.02041
- [152] Schoelz JM, Riddle NC. Functions of HP1 proteins in transcriptional regulation. *Epigenetics & Chromatin*. 2022;15(1):14.
- [153] Costa S, Dean C. Storing memories: the distinct phases of Polycomb-mediated silencing of *Arabidopsis FLC*. *Biochem Soc Trans*. 2019;47(4):1187–1196.
- [154] Walker E, Chang WY, Hunkapiller J, et al. Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation. *Cell Stem Cell*. 2010;6(2):153–166.
- [155] Sarma K, Margueron R, Ivanov A, et al. Ezh2 Requires PHF1 to Efficiently Catalyze H3 Lysine 27 Trimethylation in vivo. *Mol Cell Biol*. 2008;28(8):2718–2731.
- [156] Kraleman LEM, Liu S, Trejo-Arellano MS, et al. Removal of H2Aub1 by ubiquitin-specific proteases 12 and 13 is required for stable Polycomb-mediated gene repression in *Arabidopsis*. *Genome Biol*. 2020;21(1):144.
- [157] Gahan JM, Rentzsch F, Schnitzler CE The genetic basis for PRC1 complex diversity emerged early in animal evolution. *Proc Natl Acad Sci USA*. 2020;117:22880–22889.
- [158] Fischle W, Wang Y, Jacobs SA, et al. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes & deve*. 2003;17(15):1870–1881.
- [159] Cheutin T, Cavalli G. Loss of PRC1 induces higher-order opening of Hox loci independently of transcription during *Drosophila* embryogenesis. *Nat Commun*. 2018;9(1):3898.
- [160] Kang H, McElroy KA, Jung YL, et al. Sex comb on midleg (Scm) is a functional link between PcG-repressive complexes in *Drosophila*. *Genes & deve*. 2015;29(11):1136–1150.
- [161] Zhang X, Germann S, Blus BJ, et al. The *Arabidopsis* LHP1 protein colocalizes with histone H3 Lys27 trimethylation. *Nat Struct Mol Biol*. 2007;14(9):869–871.
- [162] Kahn TG, Dorafshan E, Schultheis D, et al. Interdependence of PRC1 and PRC2 for recruitment to polycomb response elements. *Nucleic Acids Res*. 2016;44:10132–10149.
- [163] Schuettengruber B, Martinez A-M, Iovino N, et al. Trithorax group proteins: switching genes on and keeping them active. *Nat Rev Mol Cell Biol*. 2011;12(12):799–814.
- [164] Kingston RE, Tamkun JW. Transcriptional regulation by trithorax-group proteins. *Cold Spring Harbor Perspect Biol*. 2014;6(10):a019349.
- [165] Kassis JA, Kennison JA, Tamkun JW. Polycomb and Trithorax Group Genes in *Drosophila*. *Genetics*. 2017;206(4):1699–1725.
- [166] Kennison JA. The polycomb and trithorax group proteins of *drosophila*: trans-regulators of homeotic gene function. *Annu Rev Genet*. 1995;29(1):289–304.
- [167] Kennison JA. Introduction to Trx-G and Pc-G genes. *Methods Enzymol*. 2004;377:61–70.
- [168] Alvarez-Venegas R, Avramova Z. Two *Arabidopsis* homologs of the animal *trithorax* genes: a new structural domain is a signature feature of the *trithorax* gene family. *Gene*. 2001;271(2):215–221.
- [169] Alvarez-Venegas R, Pien S, Sadler M, et al. ATX-1, an *Arabidopsis* homolog of trithorax, activates flower homeotic genes. *Curr Biol*. 2003;13(8):627–637.
- [170] Hanson RD, Hess JL, Yu BD, et al. Mammalian Trithorax and Polycomb-group homologues are antagonistic regulators of homeotic development. *Proc Natl Acad Sci USA*. 1999; 96:14372–14377.
- [171] Nislow C, Ray E, Pillus L. *SET1*, a yeast member of the *trithorax* family, functions in transcriptional silencing and diverse cellular processes. *Mol Biol Cell*. 1997;8(12):2421–2436.
- [172] Tang L, Nogales E, Ciferri C. Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Prog Biophys Mol Biol*. 2010;102(2–3):122–128.
- [173] Dillon SC, Zhang X, Trievel RC, et al. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol*. 2005;6(8):227.

- [174] Herz H-M, Garruss A, Shilatifard A. SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends Biochem Sci.* **2013**;38(12):621–639.
- [175] Gu Y, Nakamura T, Alder H, et al. The t (4; 11) chromosome translocation of human acute leukemias fuses the *ALL-1* gene, related to *Drosophila trithorax*, to the *AF-4* gene. *Cell.* **1992**;71(4):701–708.
- [176] Bantignies F, Goodman RH, Smolik SM. Functional interaction between the coactivator *Drosophila* CREB-binding protein and ASH1, a member of the trithorax group of chromatin modifiers. *Mol Cell Biol.* **2000**;20(24):9317–9330.
- [177] Eissenberg JC, Shilatifard A. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Dev Biology.* **2010**;339(2):240–249.
- [178] Pu L, Sung ZR. PcG and trxG in plants – friends or foes. *Trends Genet.* **2015**;31(5):252–262.
- [179] Peterson CL, Tamkun JW. The SWI-SNF complex: a chromatin remodeling machine? *Trends Biochem Sci.* **1995**;20(4):143–146.
- [180] Zhang X, Bernatavichute YV, Cokus S, et al. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* **2009**;10(6):R62.
- [181] Oh S, Park S, van Nocker S. Genic and global functions for Paf1C in chromatin modification and gene expression in *Arabidopsis*. *PLoS Genet.* **2008**;4(8):e1000077.
- [182] Jiang D, Kong NC, Gu X, et al. *Arabidopsis* COMPASS-like complexes mediate histone H3 lysine-4 trimethylation to control floral transition and plant development. *PLoS Genet.* **2011**;7(3):e1001330.
- [183] Wang Y, Ding Z, Liu X, et al. Architecture and subunit arrangement of the complete *Saccharomyces cerevisiae* COMPASS complex. *Sci Rep.* **2018**;8(1):17405.
- [184] Worden EJ, Zhang X, Wolberger C. Structural basis for COMPASS recognition of an H2B-ubiquitinated nucleosome. *Elife.* **2020**;9:e53199.
- [185] Dou Y, Milne TA, Ruthenburg AJ, et al. Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol.* **2006**;13(8):713–719.
- [186] Cenik BK, Shilatifard A. COMPASS and SWI/SNF complexes in development and disease. *Nat Rev Genet.* **2021**;22(1):38–58.
- [187] Xiao J, Lee U-S, Wagner D. Tug of war: adding and removing histone lysine methylation in *Arabidopsis*. *Curr Opin Plant Biol.* **2016**;34:41–53.
- [188] Cartagena JA, Matsunaga S, Seki M, et al. The *Arabidopsis* SDG4 contributes to the regulation of pollen tube growth by methylation of histone H3 lysines 4 and 36 in mature pollen. *Dev Biology.* **2008**;315(2):355–368.
- [189] Berr A, McCallum EJ, Ménard R, et al. *Arabidopsis* SET DOMAIN GROUP2 is required for H3K4 trimethylation and is crucial for both sporophyte and gametophyte development. *Plant Cell.* **2010**;22(10):3232–3248.
- [190] Shilatifard A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem.* **2012**;81(1):65–95.
- [191] Molitor AM, Bu Z, Yu Y, et al. *Arabidopsis* AL PHD-PRC1 complexes promote seed germination through H3K4me3-to-H3K27me3 chromatin state switch in repression of seed developmental genes. *PLoS Genet.* **2014**;10(1):e1004091.
- [192] Lee WY, Lee D, Chung WI, et al. *Arabidopsis* ING and Alfin1-like protein families localize to the nucleus and bind to H3K4me3/2 via plant homeodomain fingers. *Plant J.* **2009**;58(3):511–524.
- [193] Lu F, Cui X, Zhang S, et al. JM14 is an H3K4 demethylase regulating flowering time in *Arabidopsis*. *Cell Res.* **2010**;20(3):387–390.
- [194] Yang H, Mo H, Fan D, et al. Overexpression of a histone H3K4 demethylase, JM15, accelerates flowering time in *Arabidopsis*. *Plant Cell Rep.* **2012**;31(7):1297–1308.
- [195] Yang H, Han Z, Cao Y, et al. A companion cell-dominant and developmentally regulated H3K4 demethylase controls flowering time in *Arabidopsis* via the repression of *FLC* expression. *PLoS Genet.* **2012**;8(4):e1002664.
- [196] Kooistra SM, Helin K. Molecular mechanisms and potential functions of histone demethylases. *Nat Rev Mol Cell Biol.* **2012**;13(5):297–311.
- [197] Nawkar GM, Kang CH, Maibam P, et al. HY5, a positive regulator of light signaling, negatively controls the unfolded protein response in *Arabidopsis*. *Proc Natl Acad Sci USA.* **2017**;114:2084–2089.
- [198] Pastor-Cantizano N, Ko DK, Angelos E, et al. Functional diversification of ER stress responses in *Arabidopsis*. *Trends Biochem Sci.* **2020**;45(2):123–136.
- [199] Zhao X, Li J, Lian B, et al. Global identification of *Arabidopsis* lncRNAs reveals the regulation of MAF4 by a natural antisense RNA. *Nat Commun.* **2018**;9(1):5056.
- [200] Franks TM, McCloskey A, Shokhirev MN, et al. Nup98 recruits the Wdr82–Set1A/COMPASS complex to promoters to regulate H3K4 trimethylation in hematopoietic progenitor cells. *Genes & deve.* **2017**;31(22):2222–2234.
- [201] Wu JI, Lessard J, Crabtree GR. Understanding the words of chromatin regulation. *Cell.* **2009**;136(2):200–206.
- [202] Kadoch C, Crabtree GR. Mammalian SWI/SNF chromatin remodeling complexes and cancer: mechanistic insights gained from human genomics. *Sci Adv.* **2015**;1(5):e1500447.
- [203] Masliah-Planchon J, Bièche I, Guinebretière J-M, et al. SWI/SNF chromatin remodeling and human malignancies. *Annu Rev Pathol.* **2015**;10(1):145–171.
- [204] Dingwall AK, Beek SJ, McCallum CM, et al. The *Drosophila* snr1 and brm proteins are related to

- yeast SWI/SNF proteins and are components of a large protein complex. *Mol Biol Cell*. 1995;6(7):777–791.
- [205] Stern M, Jensen R, Herskowitz I. Five SWI genes are required for expression of the HO gene in yeast. *J Mol Biol*. 1984;178(4):853–868.
- [206] Wang W, Côté J, Xue Y, et al. Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *Embo J*. 1996;15(19):5370–5382. DOI:10.1002/j.1460-2075.1996.tb00921.x
- [207] Centore RC, Sandoval GJ, Soares LMM, et al. Mammalian SWI/SNF Chromatin Remodeling Complexes: emerging Mechanisms and Therapeutic Strategies. *Trends Genet*. 2020;36(12):936–950.
- [208] Mittal P, Roberts CWM. The SWI/SNF complex in cancer — biology, biomarkers and therapy. *Nat Rev Clin Oncol*. 2020;17(7):435–448.
- [209] Reyes JC. The many faces of plant SWI/SNF complex. *Mol Plant*. 2014;7(3):454–458.
- [210] Tillib S, Petruk S, Sedkov Y, et al. Trithorax- and Polycomb-group response elements within an Ultrabithorax transcription maintenance unit consist of closely situated but separable sequences. *Mol Cell Biol*. 1999;19(7):5189–5202.
- [211] Müller J, Kassiss JA. Polycomb response elements and targeting of Polycomb group proteins in *Drosophila*. *Curr Opin Genet Dev*. 2006;16(5):476–484.
- [212] Kassiss JA, Brown JL. Polycomb group response elements in *Drosophila* and vertebrates. *Adv Genet*. 2013;81:83–118.
- [213] Tie F, Banerjee R, Stratton CA, et al. CBP-mediated acetylation of histone H3 lysine 27 antagonizes *drosophila* polycomb silencing. *Development*. 2009;136(18):3131–3141.
- [214] Yuan W, Xu M, Huang C, et al. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem*. 2011;286(10):7983–7989.
- [215] Schmähling S, Meiler A, Lee Y, et al. Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC. Vol. 145. England: Development (Cambridge; 2018.
- [216] Tie F, Banerjee R, Fu C, et al. Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain. *Proc Natl Acad Sci USA*. 2016;113:E744–53.
- [217] Alvarez-Venegas R. Regulation by polycomb and trithorax group proteins in Arabidopsis. *The Arabidopsis Book*. Vol. 8. Washington, DC: PubMed Central; 2010. p. e0128.
- [218] Clapier CR, Iwasa J, Cairns BR, et al. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol*. 2017;18(7):407–422.
- [219] Giaimo BD, Ferrante F, Herchenröther A, et al. The histone variant H2A.Z in gene regulation. *Epigenetics & Chromatin*. 2019;12(1):37.
- [220] Li E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet*. 2002;3(9):662–673.
- [221] Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet*. 2009;10(5):295–304.
- [222] Li E, Zhang Y. DNA methylation in mammals. *Cold Spring Harb Perspect Biol*. 2014;6(5):a019133.
- [223] X-J H, Chen T, Zhu J-K. Regulation and function of DNA methylation in plants and animals. *Cell Res*. 2011;21(3):442–465.
- [224] Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*. 2019;20(10):590–607.
- [225] He L, Huang H, Bradai M, et al. DNA methylation-free Arabidopsis reveals crucial roles of DNA methylation in regulating gene expression and development. *Nat Commun*. 2022;13(1):1335.
- [226] Stroud H, Do T, Du J, et al. Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nat Struct Mol Biol*. 2014;21(1):64–72.
- [227] Zilberman D, Gehring M, Tran RK, et al. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genet*. 2007;39(1):61–69.
- [228] Bewick AJ, Schmitz RJ. Gene body DNA methylation in plants. *Curr Opin Plant Biol*. 2017;36:103–110.
- [229] Cavalli G, Heard E. Advances in epigenetics link genetics to the environment and disease. *Nature*. 2019;571(7766):489–499.
- [230] Molina-Serrano D, Kyriakou D, Kirmizis A. Histone modifications as an intersection between diet and longevity. *Front Genet*. 2019;10:192.
- [231] Rugowska A, Starosta A, Konieczny P. Epigenetic modifications in muscle regeneration and progression of Duchenne muscular dystrophy. *Clin Epigenetics*. 2021;13(1):13.
- [232] Casas E, Vavouri T. Mechanisms of epigenetic inheritance of variable traits through the germline. *Reproduction (Cambridge, England. Reproduction*. 2020;159(6):R251–r63.
- [233] Deevy O, Bracken AP. PRC2 functions in development and congenital disorders. *Development*. 2019;146(19). DOI:10.1242/dev.181354
- [234] Kerppola TK. Polycomb group complexes – many combinations, many functions. *Trends Cell Biol*. 2009;19(12):692–704.
- [235] Katz DJ, Edwards TM, Reinke V, et al. A *C. elegans* LSD1 demethylase contributes to germline immortality by reprogramming epigenetic memory. *Cell*. 2009;137(2):308–320.
- [236] Xiao Y, Bedet C, Robert VJP, et al. *Caenorhabditis elegans* chromatin-associated proteins SET-2 and ASH-2 are differentially required for histone H3 Lys 4 methylation in embryos and adult germ cells. *PNAS*. 2011;108(20):8305–8310.
- [237] Mendenhall EM, Bernstein BE. Chromatin state maps: new technologies, new insights. *Curr Opin Genet Dev*. 2008;18(2):109–115.

- [238] Schuettengruber B, Cavalli G. Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development*. 2009;136(21):3531–3542.
- [239] Lee I, Amasino RM. Effect of vernalization, photoperiod, and light quality on the flowering phenotype of *Arabidopsis* plants containing the *FRIGIDA* gene. *Plant Physiol*. 1995;108(1):157–162.
- [240] Michaels SD, Amasino RM. *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*. 1999;11(5):949–956.
- [241] Lee J, Amasino RM. Two FLX family members are non-redundantly required to establish the vernalization requirement in *Arabidopsis*. *Nat Commun*. 2013;4(1):2186.
- [242] Kim DH, Sung S. Genetic and epigenetic mechanisms underlying vernalization. *Arabidopsis Book*. 2014;12:e0171.
- [243] Sung S, Amasino RM. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature*. 2004;427(6970):159–164.
- [244] Swiezewski S, Liu F, Magusin A, et al. Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature*. 2009;462(7274):799–802.
- [245] Cao R, Wang H, He J, et al. Role of hPHF1 in H3K27 methylation and Hox gene silencing. *Mol Cell Biol*. 2008;28(5):1862–1872.
- [246] Inouye C, Remondelli P, Karin M, et al. Isolation of a cDNA encoding a metal response element binding protein using a novel expression cloning procedure: the one hybrid system. *DNA Cell Biol*. 1994;13(7):731–742.
- [247] Coulson M, Robert S, Eyre HJ, et al. The identification and localization of a human gene with sequence similarity to Polycomblike of *Drosophila melanogaster*. *Genomics*. 1998;48(3):381–383.
- [248] Wang S, Robertson GP, Zhu J. A novel human homologue of *Drosophila* polycomblike gene is up-regulated in multiple cancers. *Gene*. 2004;343(1):69–78.
- [249] Lee J, Yun JY, Zhao W, et al. A methyltransferase required for proper timing of the vernalization response in *Arabidopsis*. *Proc Natl Acad Sci USA*. 2015;112:2269–2274.
- [250] Sahr T, Adam T, Fizames C, et al. O-carboxyl- and N-methyltransferases active on plant aquaporins. *Plant & Cell Physiol*. 2010;51(12):2092–2104.
- [251] Cao X-J, Arnaudo AM, Garcia BA. Large-scale global identification of protein lysine methylation in vivo. *Epigenetics*. 2013;8(5):477–485.
- [252] Clarke SG. Protein methylation at the surface and buried deep: thinking outside the histone box. *Trends Biochem Sci*. 2013;38(5):243–252.
- [253] Huang J, Berger SL. The emerging field of dynamic lysine methylation of non-histone proteins. *Curr Opin Genet Dev*. 2008;18(2):152–158.
- [254] Jones AM, Chory J, Dangel JL, et al. The impact of *Arabidopsis* on human health: diversifying our portfolio. *Cell*. 2008;133(6):939–943.
- [255] Chan S-L, Henderson IR, Jacobsen SE. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat Rev Genet*. 2005;6(5):351–360.
- [256] Yan Q, Cho E, Lockett S, et al. Association of Lsh, a regulator of DNA methylation, with pericentromeric heterochromatin is dependent on intact heterochromatin. *Mol Cell Biol*. 2003;23(23):8416–8428.
- [257] Heidstra R, Sabatini S. Plant and animal stem cells: similar yet different. *Nat Rev Mol Cell Biol*. 2014;15(5):301–312.
- [258] Olariu V, Nilsson J, Jönsson H, et al. Different reprogramming propensities in plants and mammals: are small variations in the core network wirings responsible? *PLoS ONE*. 2017;12(4):e0175251.
- [259] Sang Y, Wu M-F, Wagner D. The stem cell—Chromatin connection. *Semin Cell Dev Biol*. 2009;20(9):1143–1148.
- [260] Laux T. The stem cell concept in plants: a matter of debate. *Cell*. 2003;113(3):281–283.
- [261] Koornneef M, Meinke D. The development of *Arabidopsis* as a model plant. *Plant J*. 2010;61(6):909–921.
- [262] Jiao W-B, Schneeberger K. Chromosome-level assemblies of multiple *Arabidopsis* genomes reveal hotspots of rearrangements with altered evolutionary dynamics. *Nat Commun*. 2020;11(1):989.
- [263] Pruitt RE, Meyerowitz EM. Characterization of the genome of *Arabidopsis thaliana*. *J Mol Biol*. 1986;187(2):169–183.
- [264] Goodman HM, Ecker JR, Dean C. The genome of *Arabidopsis thaliana*. *Proc Natl Acad Sci USA*. 1995;92:10831–10835.
- [265] Bateman A, Martin M-J, Orchard S, et al. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021;49(D1):D480–9. doi:10.1093/nar/gkaa1100.
- [266] O’Leary NA, Wright MW, Brister JR, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res*. 2016;44(D1):D733–45. DOI:10.1093/nar/gkv1189