

Review

Zn²⁺-Dependent Histone Deacetylases in Plants: Structure and EvolutionInmaculada Yruela,^{1,2,4,*} Carlos Moreno-Yruela,^{3,4,@} and Christian A. Olsen^{3,@}

Zn²⁺-dependent histone deacetylases are widely distributed in archaea, bacteria, and eukaryotes. Through deacetylation of histones and other biomolecules, these enzymes regulate mammalian gene expression, microtubule stability, and polyamine metabolism. In plants, they play essential roles in development and stress response, but little is known about their biochemistry. We provide here a holistic revision of plant histone deacetylase (HDA) phylogeny and translate recent lessons from other organisms. HDA evolution correlates with a gain of structural ductility/disorder, as observed for other proteins. We also highlight two recently identified Brassicaceae-specific HDAs, as well as unprecedented key mutations that would affect the catalytic activity of individual HDAs. This revised phylogeny will contextualize future studies and illuminate research on plant development and adaptation.

Diversification of Histone Deacetylases

Histone deacetylases (see [Glossary](#), named HDACs in mammals and HDAs in plants) are widely distributed in all kingdoms of life. These enzymes regulate gene expression among other biological processes, play a crucial role in development and cell-cycle progression, and their deregulation is associated with human diseases such as cancer and neurodegeneration [1–3].

In eukaryotes, nuclear HDACs deacetylate the histone proteins that scaffold DNA packing into nucleosomes, rendering them key epigenetic regulators [4–6]. Histone deacetylation leads to stronger histone–DNA interactions, stimulates chromatin condensation, and thus triggers transcriptional repression [7]. By contrast, HDACs delete the N^ε-acetyl-lysine recognition motif for bromodomain-containing transcriptional corepressors and activators [8]. Overall, given the opposing effects of histone acetyltransferases (HATs), there is a strict balance between acetylation and deacetylation of histones that constitutes a pivotal regulatory mechanism of gene expression [8]. Importantly, HDACs regulate the function of biomolecules other than histones, both in the nucleus and in the cytoplasm. For example, HDAC3 deacetylates phosphoglycerate kinase (PGK) [9], HDAC6 regulates microtubule dynamics via deacetylation of α -tubulin [10,11], and HDAC10 takes part in the metabolism of polyamines by deacetylating N^ε-acetylspermidine [12].

In plants, reversible histone acetylation affects genome stability, transcriptional regulation, development, and stress responses [13,14]. In addition, HDAs regulate the cell cycle and programmed cell death, in part via interaction with key transcription factors [15,16], and they have also been shown to deacetylate proteins other than histones [17].

There are three families of lysine deacetylases in eukaryotes, based on three conserved deacetylase domains that rely on different cofactors [18]. HDAs refer here to the Rpd3/Hda1 (reduced potassium-dependence 3/histone deacetylase 1) family of Zn²⁺-dependent hydrolases, which share similarity to yeast Rpd3, further grouped into classes I, II, and IV [19,20]. **Phylogenetic analyses** suggest that these three HDAC/HDA classes are constituted by ancient enzymes that evolve from bacterial

Highlights

Histone deacetylases are ubiquitous and highly conserved enzymes across all kingdoms of life.

Plant HDAs regulate gene expression and control development and stress response, but little is known about their structure and biochemistry.

Structural and biochemical studies have revealed diverse functions in mammals beyond deacetylation of histones and chromatin remodeling.

A revised and holistic phylogenetic analysis highlights plant HDAs that retain structural characteristics from their bacterial origin and HDAs that have diverged to fulfill potential family-specific functions.

Individual HDAs harbor active-site mutations that have not been identified in mammals or bacteria.

¹Estación Experimental de Aula Dei, Consejo Superior de Investigaciones Científicas (CSIC), Avda. Montañana 1005, 50059 Zaragoza, Spain

²Group of Biochemistry, Biophysics, and Computational Biology (GBsC), Institute for Biocomputation and Physics of Complex Systems (BIFI) and Universidad de Zaragoza (UNIZAR) Joint Unit to CSIC, Zaragoza, Spain

³Center for Biopharmaceuticals and Department of Drug Design and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Universitetsparken 2, 2100 Copenhagen, Denmark

⁴These authors contributed equally.

*Correspondence:

i.yruela@csic.es (I. Yruela).

@Twitter: @InmaYruela (I. Yruela),

@CarlosMYruela (C. Moreno-Yruela), and @ChristianAOlsen (C.A. Olsen)

HDAC-related proteins [19]. The sir2 (silent information regulator 2) family of NAD⁺-dependent enzymes (the sirtuins, class III HDACs) and the plant-specific HD2 family (or HD-tuins) possess catalytic domains unrelated to those of the Rpd3/Hda1 family [21,22] and are not the focus of this review.

The diversification of Zn²⁺-dependent deacetylases in all kingdoms of life indicates a high degree of evolutionary plasticity, and suggests that each of the HDA classes has a non-redundant and essential role. Plant Rpd3 homologs were first identified in maize in 1998 [23], and have been since characterized in arabidopsis (*Arabidopsis thaliana*), *Zea mays* (maize), *Oryza sativa* (rice), and other model and crop plant species [24–30]. Various HDA phylogenetic analyses have been published, including a limited number of species or focused on individual genomes (Box 1) [25,27,29–31]. In this review we provide a holistic evolutionary revision of the plant HDA family, including 60 species representative of 42 families (Figure S1 in the supplemental

Box 1. History of the Phylogeny of Plant HDAs

Plant Zn²⁺-dependent HDAs are divided into classes I, II, and IV, and they can be further grouped according to similarity to arabidopsis sequences [25,32]. Pandey *et al.* identified two clusters within class I containing the *AtHDA19* (also known as *HDA1*), and *AtHDA6* and *AtHDA7* genes, respectively (Figure 1A). Each cluster included dicot and monocot genes, suggesting a genome duplication origin before the divergence of the two lineages [32]. In addition, in the phylogeny of Pandey *et al.* class II HDAs include the *AtHDA5*, *AtHDA15*, and *AtHDA18* genes, and class IV contains *AtHDA2*, which has similarity to bacterial *Acu* and cyanobacterial glutamine synthetase [32]. Note that class IV was earlier designated as class III, but was then renamed class IV to distinguish it from the sir2 family. In this first analysis, arabidopsis orthologs, namely *AtHDA9*, *AtHDA10*, and *AtHDA17*, were included as class I HDAs, and *AtHDA10* and *AtHDA17* were suggested to be derived from *AtHDA9* by sequence rearrangements (Figure 1A) [32]. Later, a closer look revealed that these genes encode truncated proteins (158 and 142 residues, respectively) that lack essential HDA structural elements [25]. In addition, the *AtHDA8* and *AtHDA14* genes were unclassified, although they appeared to be related to class II [32]. Alinsug *et al.* later proposed a revised class II that included *AtHDA5*, *AtHDA8*, *AtHDA14*, *AtHDA15*, and *AtHDA18* (Figure 1B) [25]. In other studies, such as those conducted by Fu *et al.* in *Oryza sativa*, it has been suggested that the Rpd3/Hda1 HDA family could be divided into four alternative classes [117]. This review allows the elaboration of a revised HDA phylogenetic tree (Figure 1C) and the identification of key HDA variants with potential alternative functions. We conclude that *AtHDA7* and *AtHDA18* do not constitute canonical subclasses of HDAs, and are instead members of subclasses HDA6 of class I and HDA5 of class II HDAs, respectively (hereafter HDA6-2 and HDA5-2). They are associated with Brassicaceae lineage-specific genes that derived from a recent divergence event. In addition, the HDA9 subclass of class I is split into two subgroups, HDA9a and HDA9b, as a result of evolutionary divergence in the monocot lineage.

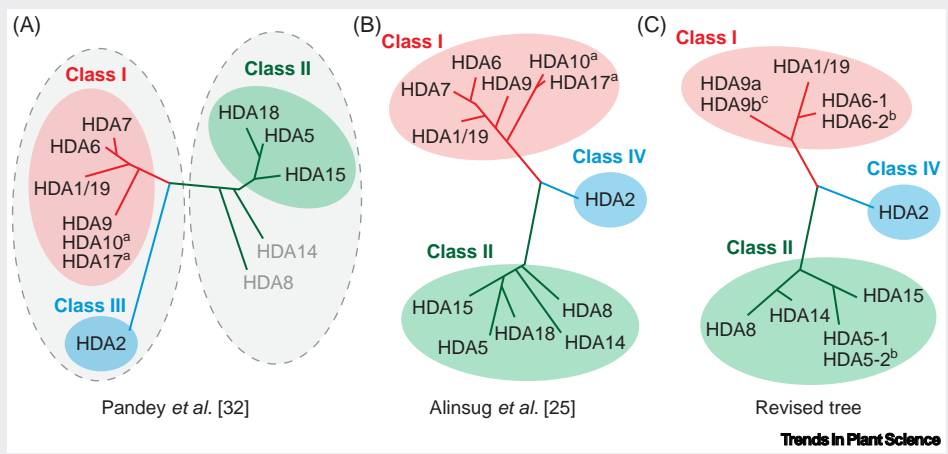


Figure 1. Diversification and Phylogeny Revision of Plant HDA Classes. Schemes of the phylogenetic cladograms proposed for plant Zn²⁺-dependent HDAs by (A) [32], (B) [25], and (C) this review. ^aTruncated sequences that lack the HDA catalytic pocket. ^bBrassicaceae-specific HDAs: HDA6-2 was previously named HDA7 and HDA5-2 was termed HDA18. ^cMonocot-specific subclass. Cladograms of panels (A) and (B) are adapted, with permission, from [32] and [25].

Glossary

- Angiosperms:** flowering plants, vascular plants that are distinguished by the production of seeds within an enclosed ovary. These constitute the most abundant and diverse group of land plants.
- Areaceae, Bromeliaceae, Musaceae, Zosteraceae:** families of monocotyledonous plants.
- Brassicaceae:** a family of dicots that includes economically important herbaceous plants. Their diversification is dated to ~100 Mya.
- Bryophytes:** non-vascular land plants that include liverworts, hornworts, and mosses. They belong to an old lineage dated to ~500 Mya.
- Chlorophytes:** 'green algae', a highly paraphyletic group of green plants that mostly live in aquatic environments. Their origin is dated to ~1200 Mya.
- Cladogram:** 'evolutionary tree', a diagram that shows predicted phylogenetic relationships among organisms, genes, or proteins. Cladograms use lines to distinguish between paralogs and orthologs during the course of evolution.
- Cyanobacteria:** blue-green bacteria, a phylum of photosynthetic prokaryotes. They were involved in the origin of primary plastids and chloroplasts through endosymbiotic events, and their origin is dated to 2700–1400 Mya.
- Dicots:** dicotyledons, a paraphyletic group of angiosperms that are distinguished by seeds that have two embryonic leaves or cotyledons. The divergence between monocots and dicots occurred at 145–230 Mya.
- Divergence:** also known as divergent evolution, the accumulation of differences between closely related populations within a species leading to speciation. This term can also be applied in molecular evolution when proteins derive from homologous genes, either orthologs or paralogs.
- Ductile region:** an intrinsically disordered region (IDR) of a protein that does not fold into a well-defined 3D structure. Ductile regions can be either entirely disordered or partially disordered, spanning only a few contiguous disordered residues (<10 aa) or containing long segments (≥30 aa) of contiguously disordered residues. Algorithms have been developed that can predict the likelihood that an amino acid in a protein sequence is structured or disordered [42–44].

information online). This allows the identification of key HDA variants with potential alternative functions. Furthermore, we discuss structural differences in the context of the recent literature on mammalian HDACs.

Phylogeny and Evolution of Plant HDAs

The extensive genomic information currently available on families of plants from **chlorophytes to angiosperms** enables precise revision of plant HDA phylogeny. Importantly, we verified the sequence composition of each protein so as to discard incomplete proteins and identify potentially inactive enzymes that lack essential amino acid residues.

Class I HDAs

Two major distinct clades are distinguished in **dicots** (Figure 1A) and **monocots** (Figure 1B): clade 1 with subclass HDA9, and clade 2 with subclasses HDA19 (also named HDA1, hereafter HDA1/19) and HDA6. Subclasses HDA1/19 and HDA6 share a common ancestor and have evolved independently of subclass HDA9. As expected, **orthologs** in **Brassicaceae** are clustered (bootstrap >85) in the three subclasses according to their phylogeny (Figure 1A).

Earlier studies in *Arabidopsis* reported a distinct ortholog, named HDA7 (At5g35600), which could belong to a fourth subclass of class I [25,32]. *AthaHDA7* and subsequent orthologs from the **Brassicaceae** species *Crub* (pink shepherd's purse), *Bnap* (colza), *Bole* (broccoli, cabbage), and *Brap* (napa cabbage) are found within subclass HDA6 and are positioned close to *AthaHDA6* (At5g63110), which indicates that they are **paralogs** (Figure 1A, light-blue circles). Furthermore, HDA7-type orthologs are not present in other related phylogenetic families, indicating that HDA7 is encoded by **Brassicaceae** lineage-specific genes that are associated with a recent **divergence** event. In fact, **Brassica** genomes have undergone whole-genome duplication (WGD) and lineage-specific whole-genome triplication (WGT) events followed by diploidization [33–35]. For clarity in the annotation, we recommend that the paralogs of *AthaHDA6* and *AthaHDA7* (and their respective orthologs) are renamed HDA6-1 and HDA6-2, respectively (Figure 1A).

Chlorophyte homologs are only present in clade 1 within subclass HDA9 (bootstrap >98, Figure 1), indicating that subclass HDA9 has a more ancient evolutionary origin than subclasses HDA1/19 and HDA6, and that these originated before the divergence of the monocot and dicot lineages. In addition, two distant **paraphyletic subgroups** (bootstrap >90) appear in the monocot subclass HDA9 (Figure 1B, named HDA9a and HDA9b). Interestingly, homologs from chlorophytes and the ancient freshwater alga *Chara braunii* (Cbra, ca. 450 Mya) [36,37] cluster close to the HDA9a subgroup (bootstrap >95) in dicot and monocot trees. This suggests that subgroup HDA9b diverged later as a result of a WGD event in the monocot lineage. Note that HDA9b orthologs in the monocot lineage are often annotated as HDA6 (i.e., *Acom*, XP_020098843.1; *Atau*, XP_020188071.1; *Doli*, OEL25058; *Pdac*, XP_008791160.1; *Sbic*, XP_002444249.1), which can lead to confusion because HDA6 constitutes a different subclass of class I. In the dicot phylogeny (Figure 1A), HDA9 of *Qsub* clustered together with green algae homologs, which could indicate an ancestral origin for this protein, but it could also reflect errors in the reported sequence.

Class II HDAs

Class II HDAs comprise two major clades: clade 1, with subclasses HDA8 and HDA14; and clade 2, with subclasses HDA5 and HDA15 (Figure 2). Subclasses HDA8 and HDA14 are only present in plants but still exhibit high sequence similarity to the conserved catalytic domains of HDA5 and HDA15, which are common to all eukaryotic organisms [26]. Thus, subclasses HDA8 and HDA14 likely diverged separately from HDA5 and HDA15.

Histone deacetylases (HDACs/HDAs):

Zn²⁺-dependent hydrolase enzymes in archaea, bacteria, and eukaryotes that cleave post-translational acyl modifications of the ε-amino groups of lysine residues in the proteome.

Lycophyte: a paraphyletic clade of vascular plants that constitutes an old lineage of land plants whose origin is dated to 400–425 Mya.

Monocots: monocotyledons, principally grasses, a paraphyletic group of angiosperms distinguished by seeds that contain a single embryonic leaf or cotyledon. They include most economically important crops. The divergence between monocots and dicots occurred at 145–230 Mya.

Monophyletic group: a group of organisms with a common ancestor, which can be an individual, a species, or a population.

Orthologs: copies of genes or proteins that result from speciation during the evolution of species.

Papaveraceae: the ancestral family of dicots.

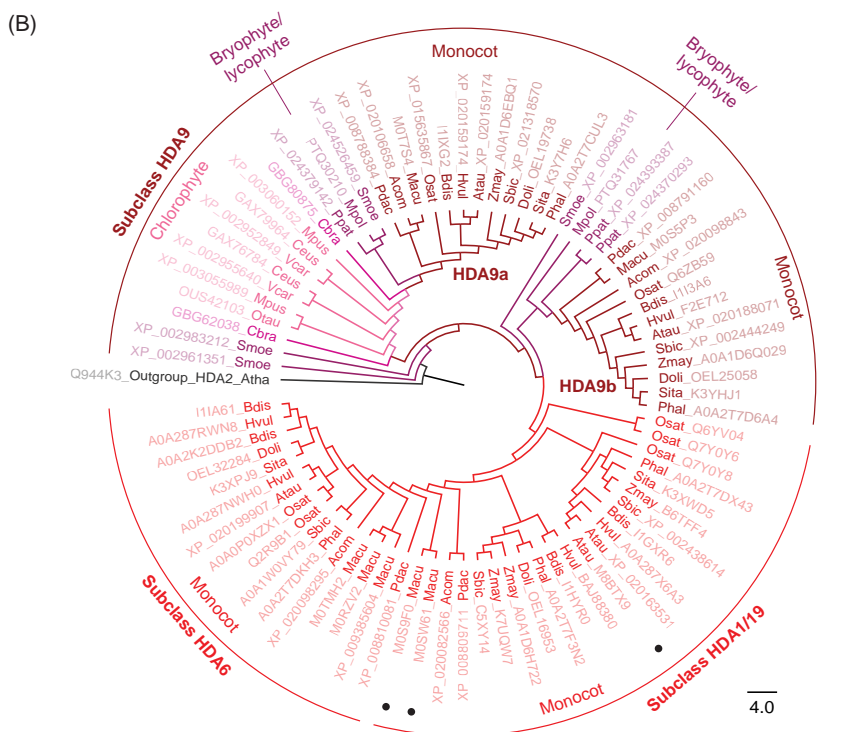
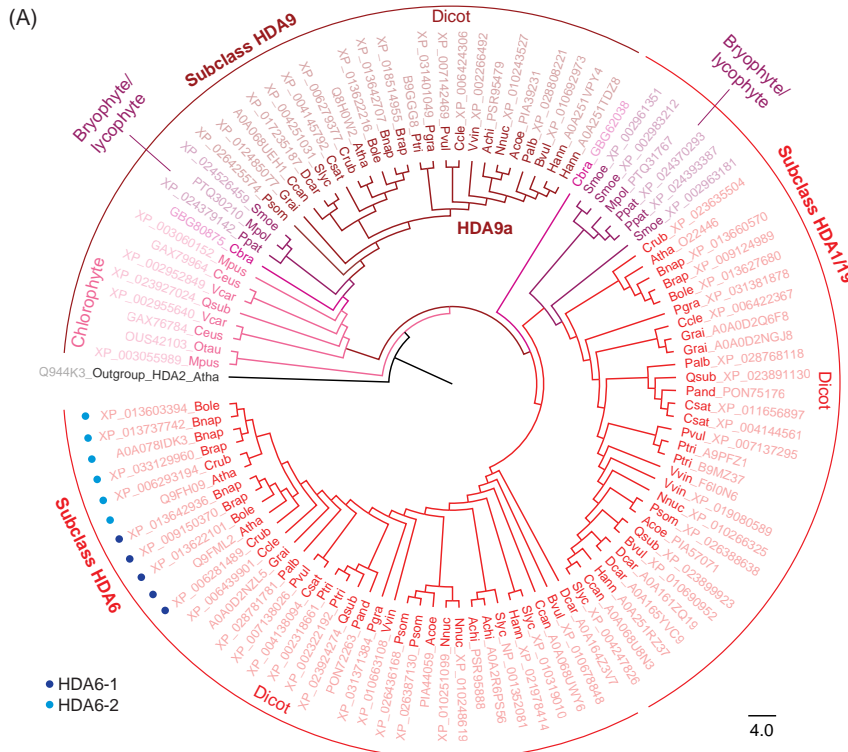
Paralogs: copies of genes or proteins that have descended side by side during the history of an organism.

Paraphyletic group: a group of organisms that includes an ancestor but not all of its descendants.

Phylogenetic analyses: computational methods that infer the evolutionary history and relationships among organism or species, or between genes, proteins, or organs.

Poaceae: also known as Gramineae, a large family of monocots that include most grass plants. Their flowers are distinguished by spikelets grouped into panicles or spikes. They underwent a whole-genome duplication (WGD) event at ~60 Mya.

Protein ductility: also known as protein disorder, flexibility in the 3D structure of a protein.



By analogy to AthaHDA7, previous reports revealed a distinct ortholog within class II named HDA18 (At5g61070) that was assigned to a different subclass [25,32]. The AthaHDA18 protein (Figure 2A, light-blue circles) is found here in the same **monophyletic group** (bootstrap >95) as AthaHDA5 (At5g61060), indicating that AthaHDA18 and AthaHDA5 are paralogs. In addition, orthologs of AthaHDA18 are identified in other Brassicaceae species such as Crub (pink shepherd's purse), Bnap (colza), Bole (cabbage) and Brap (napa cabbage), which suggests that HDA18 is also encoded by a Brassicaceae lineage-specific gene that is associated with a recent divergence event. For clarity in the annotation, we recommend that AthaHDA5 and AthaHDA18 and their respective orthologs are renamed HDA5-1 and HDA5-2, respectively (Figure 2A). The existence of these paralogs of HDA6 (class I) and HDA5 (class II) could help in understanding the divergence and adaptation mechanisms in Brassica plants.

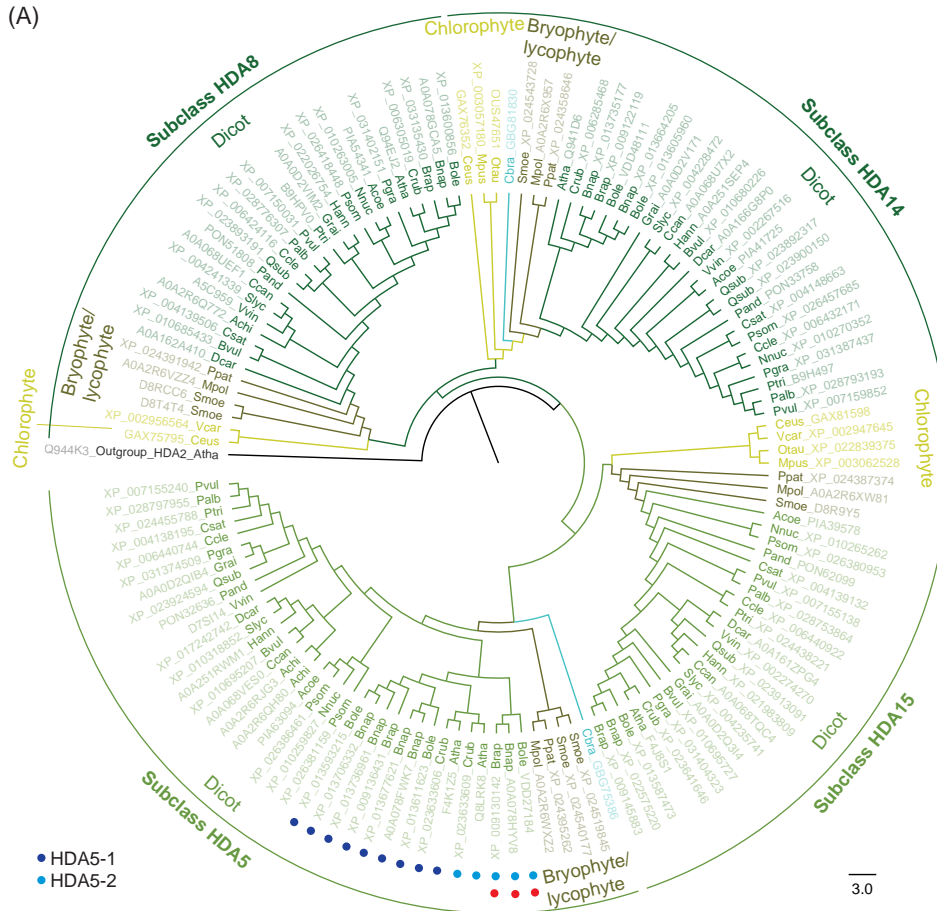
Class II Brassicaceae and **Poaceae** HDAs form separated monophyletic subgroups (bootstrap >98) in the HDA5, HDA8 clades, and in the HDA14 clade of Brassicaceae and the HDA15 clade of Poaceae, following the angiosperm phylogeny. An exception to this is OsatHDA5, which is grouped with Arecaceae, Bromeliaceae, and Musacea members (bootstrap >85). Furthermore, the dicot tree suggests that subclasses HDA8, HDA14, and HDA15 had an earlier ancestor in this lineage, and that they share a distant evolutionary relationship compared to subclass HDA5 (Figure 2A).

Lastly, some HDA5 orthologs from the distant **Papaveraceae** family (with respect to Brassicaceae) are annotated as HDA18-like (i.e., Psom, XP_026381159), and some HDA14 orthologs in monocot lineage are annotated as HDA10 (i.e., Zmay, XP_008673398), which can lead to confusion. The terms HDA10 and HDA17 have been previously associated with truncated members of class I HDAs in Arabidopsis [25,32]. However, the truncation has eliminated a large portion of the HDA catalytic domain, including key catalytic residues, indicating that these proteins have a different structure and function [25]. Therefore, we consider it more accurate to not include HDA10 and HDA17 in the plant HDA family tree (Box 1).

Class IV HDAs

Class IV HDAs are distributed into two major distinct clades: clade 1 with chlorophyte, **bryophyte**, **lycophyte**, and monocot sequences, and clade 2 with dicot sequences (Figure 3A). This clustering pattern is consistent with the diversification of the two lineages during evolution. In clade 1,

Figure 1. Phylogeny of Plant Class I Histone Deacetylases (HDAs). Phylogenetic cladograms including 23 families, 45 species, and 180 sequences of (A) dicot and (B) monocot class I HDAs. (A) Sequences belong to the dicot families: Actinidiaceae (*Actinidia chinensis*, Achi), Amarantaceae (*Beta vulgaris*, Bvul), Apiaceae (*Daucus carota*, Dcar), Asteraceae (*Helianthus annuus*, Hann), Brassicaceae (*Arabidopsis thaliana*, Atha; *Brassica napus*, Bnap; *Brassica oleracea*, Bole, *Brassica rapa*, Brap; *Capsella rubella*, Crub), Cannabaceae (*Parasponia andersonii*, Pand), Characeae (ancestral plant *Chara braunii*, Cbra), Cucurbitaceae (*Cucumis sativus*, Csat), Fabaceae (*Phaseolus vulgaris*, Pvul; *Prosopis alba*, Palb), Fagaceae (*Quercus suber*, Qsub), Lythraceae (*Punica granatum*, Pgra), Malvaceae (*Gossypium raimondii*, Grai), Nelumbonaceae (ancestral eudicot *Nelumbo nucifera*, Nnuc), Papaveraceae (ancestral eudicot *Papaver somniferum*, Psom), Ranunculaceae (ancestral eudicot *Aquilegia coerulea*, Acoe), Rubiaceae (*Coffea canephora*, Ccan), Rutaceae (*Citrus clementina*, Ccle), Salicaceae (*Populus trichocarpa*, Ptri), Solanaceae (*Solanum lycopersicum*, Slyc), Vitaceae (*Vitis vinifera*, Vvin). (B) Sequences belong to the monocot families: Arecaceae (*Phoenix dactylifera*, Pdac), Bromeliaceae (*Ananas comosus*, Acom), Musaceae (*Musa acuminata*, Macu), and Poaceae (*Aegilops tauschii*, Atau; *Brachypodium distachyon*, Bdis; *Dichanthelium oligosanthes*, Doli; *Hordeum vulgare*, Hvul; *Oryza sativa*, Osat; *Panicum hallii*, Phal; *Sorghum bicolor*, Sbic; *Setaria italica*, Sita; *Zea mays*, Zmay). Chlorophyte (*Chlamydomonas eustigma*, Ceus; *Micromonas pusilla*, Mpus; *Ostreococcus tauri*, Otai; *Volvox carteri*, Vcar), bryophyte (*Marchantia polymorpha*, Mpol; *Physcomitrella patens*, Ppat), and lycophyte (*Selaginella moellendorffii*, Smoe) species are also included. AthaHDA2 was selected as outgroup to highlight the clear evolutionary separation between class I and class IV HDAs. Sequence alignment and phylogenetic analysis were performed as described in [118]. The likelihood aLRT (approximate likelihood-ratio test) statistical test and a bootstrap value of 1000 were used.



monocot Poaceae sequences are separated into two clusters (bootstrap >90), and **Arecaceae**, **Bromeliaceae**, **Musaceae**, and **Zosteraceae** form distinct clusters close to bryophytes and chlorophytes, suggesting a common ancestor. Clade 2 follows the clustering pattern of dicot phylogeny, including separated clusters for the ancestral dicots Acoe, Nnuc, and Psom (bootstrap >90).

Evolution and Ductile Regions in the HDA Family

It is well documented that **protein ductility** increases in concert with organismal complexity and multicellularity [38,39]. This means that the proportion of flexible or **ductile regions** (often named intrinsically disordered regions, IDRs) in protein sequences generally increases in eukaryotes relative to bacterial organisms. It is predicted that prokaryotic proteomes contain 12–30% residues within long IDRs, whereas in eukaryotes this prediction is 33–50% [40–42]. This gain of IDRs in eukaryotes is hypothesized to be associated with differences in the genetic machinery [43]. Importantly, individual eukaryotic proteins with an endosymbiotic origin may maintain bacterial characteristics in their sequences [43].

Class I and II HDACs include characteristic long IDRs at the N-terminus and/or C-terminus [20], which are often indicative of protein interaction networks [44]. In class I, longer IDRs are predicted in the C-termini of plant subclasses HDA1/19 (150–160 amino acids, aa), HDA6 (80–100 aa), and the monocot subclass HDA9b (80–100 aa) compared to subclass HDA9a (40–50 aa). Accordingly, HDA9a sequences contain ~17% disordered residues, a percentage much lower than those predicted in HDA1/19, HDA6, and HDA9b (30–35%, Figure 3B), consistent with their ancient origin. Moreover, HDA9a sequences present a high similarity (up to 60%) with bacterial proteins (i.e., GeneBank ID: HHG08807.1, from metagenome analysis in Yellowstone National Park). In Brassicaceae, HDA6-2 orthologs lack the long flexible C-terminal tail that is characteristic of subclass HDA6, explaining the reduction in disordered residue fraction from ~33% to ~22%. In class II, subclass HDA15 presents a long IDR in the N-terminus (ca. 130 aa) that is not predicted in the other subclasses, thus increasing disordered residues to ~33% (Figure 3B) and suggesting an earlier eukaryotic origin. Similarly, some Brassicaceae HDA5-2 orthologs present up to 50% disordered residues, suggesting that they have a more recent origin than HDA5-1 and a clear increase in plasticity. Note, Brassica species underwent recent WGD and hexaploidization events (23–47 Mya), and this affected the evolution of many agriculturally important species [45,46].

Most **cyanobacterial** nuclear genomes contain two histone deacetylases that are homologs of classes II and IV, whereas no HDAs are encoded in the chloroplast genome. Cyanobacterial HDAs present high similarity (47–66%) to plant HDA14 and HDA2, respectively (Figure S2),

Figure 2. Phylogeny of Plant Class II Histone Deacetylases (HDAs). Phylogenetic cladograms including a total of 27 families, 44 species, and 104 sequences of (A) dicot and (B) monocot class II HDAs. (A) Sequences belong to the dicot families: Actinidiaceae (*A. chinensis*, Ach), Amarantaceae (*B. vulgaris*, Bvul), Apiaceae (*D. carota*, Dcar), Asteraceae (*H. annuus*, Hann), Brassicaceae (*A. thaliana*, Atha; *B. napus*, Bnap; *B. oleraceae*, Bole; *B. rapa*, Brap; *C. rubella* Crub), Cannabaceae (*P. andersonii*, Pand), Characeae (ancestral plant *C. braunii*, Cbra), Cucurbitaceae (*C. sativus*, Csat), Fabaceae (*P. vulgaris*, Pvul; *P. alba*, Palb), Fagaceae (*Q. suber*, Qsub), Lythraceae (*P. granatum*, Pgra), Malvaceae (*G. raimondii*, Grai), Nelumbonaceae (ancestral eudicot *N. nucifera*, Nnuc), Papaveraceae (ancestral eudicot *P. somniferum*, Psom), Ramunculaceae (ancestral eudicot, Acoe), Rubiaceae (*C. canephora*, Ccan), Rutaceae (*C. clementina*, Ccle), Salicaceae (*P. trichocarpa*, Ptri), Solanaceae (*S. lycopersicum*, Slyc), Vitaceae (*V. vinifera*, Vvin). Chlorophyte (*C. eustigma*, Ceus; *M. pusilla*, Mpus; *O. tauri*, Otau; *V. carteri*, Vcar), bryophyte (*M. polymorpha*, Mpol; *P. patens*, Ppat), and lycophyte (*S. moellendorffii*, Smoe) species are also included. (B) Sequences belong to the monocot families: Arecaceae (*P. dactylifera*, Pdac), Bromeliaceae (*A. coerulea*, Acom), Musaceae (*M. acuminata*, Macu), and Poaceae (*A. tauschii*, Atau; *B. distachyon*, Bdis; *D. oligosanthos*, Doli; *H. vulgare*, Hvul; *O. sativa*, Osat; *P. hallii*, Phal; *S. bicolor*, Sbic; *S. italica*, Sita; *Z. mays*, Zmay). Chlorophytes, bryophytes, and lycophytes were not included for simplicity. AthaHDA2 was again used as outgroup to highlight the separation from class IV. The sequence alignment and phylogenetic analysis were performed as described in Figure 1.

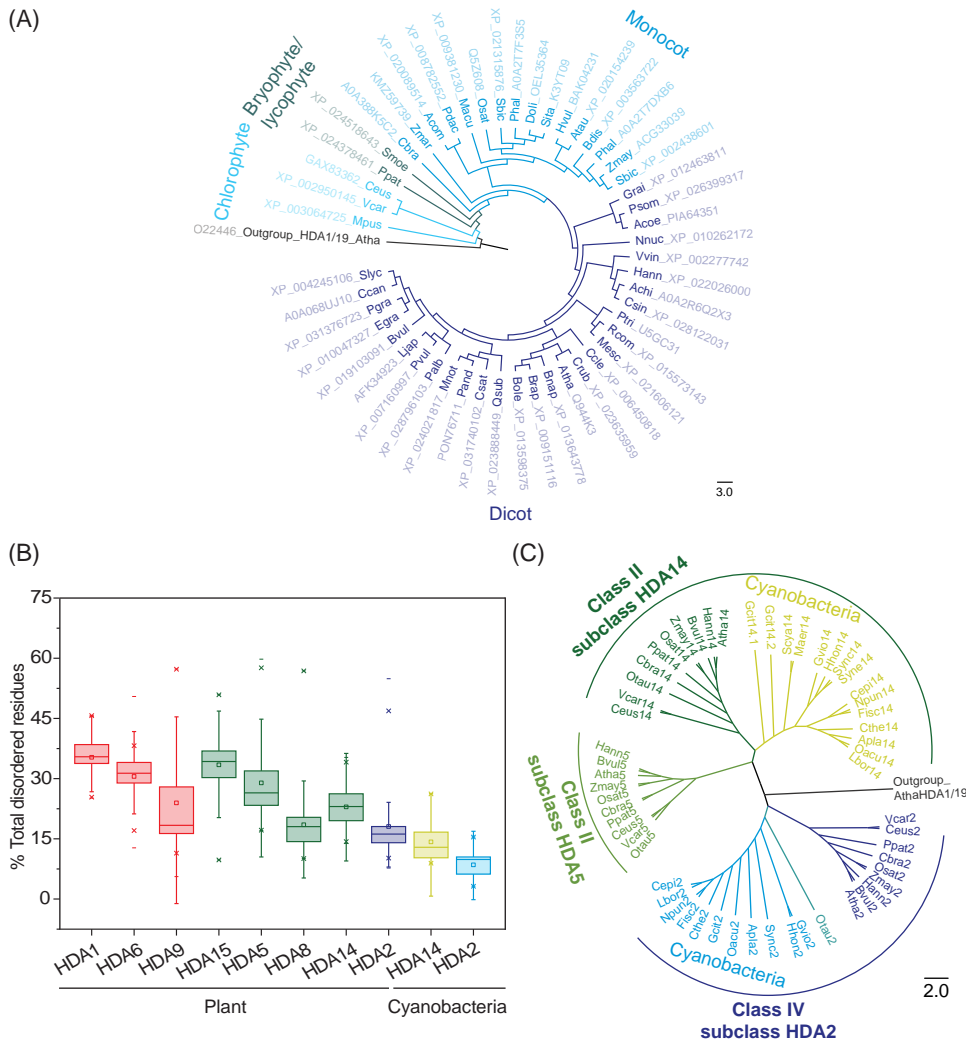


Figure 3. Phylogeny of Plant Class IV Histone Deacetylases (HDAs) and Ductility/Disorder in HDA Evolution. (A) Phylogenetic cladogram of plant class IV HDAs including 31 families, 51 species, and 51 sequences. Sequences belong to chlorophyte (Ceus, Mpus, Vcar), bryophyte (Ppat), lycophyte (Smoe), and monocot and dicot families: Actinidiaceae (Achi), Amarantaceae (Bvul), Asteraceae (Pdca), Bromeliaceae (Acum), Brassicaceae (Atha, Bnap, Bole, Brap, Crub), Cannabaceae (Pand), Characeae (ancestral plant Cbra), Cucurbitaceae (Csat), Euphorbiaceae (*Manihot esculenta*, Mesc; *R. communis*, Rcom), Fabaceae (*Lotus japonicus*, Ljap, Pvul, Palb), Fagaceae (Qsub), Lythraceae (Pgrr), Malvaceae (Grai), Moraceae (*Morus notabilis*, Mnot), Musaceae (Macu), Myrtaceae (*Eucalyptus grandis*, Egra), Nelumbonaceae (ancestral eudicot Nnuc), Papaveraceae (ancestral eudicot Psom), Poaceae (Atau, Bdis, Doli, Hvul, Osat, Phal, Sbic, Sita, Zmay), Ramunculaceae (ancestral eudicot Acoe), Rubiaceae (Ccan), Rutaceae (Cole), Salicaceae (Ptri), Solanaceae (Slyc), Theaceae (Csin), Vitaceae (Vvin), and Zosteraceae (*Zostera marina*, Zmar). Outgroup: AthaHDA1/19. Analyses were performed as described in Figure 1. (B) Box-plot distribution of the percentage of disordered residues in HDA proteins from cyanobacteria and plants, based on predictions from PONDRL VSL2b algorithm [119,120]. (C) Phylogenetic tree of HDAs of subclasses HDA2, HDA14, and HDA5 from cyanobacteria and plants. Sequences of cyanobacteria belong to: *Crinallium epipsammum* PCC 9333 (Cepi14, Cepi2); *Chroococcidiopsis thermalis* PCC 7203 (Cthe14, Cthe2); *Fischerella* sp. NIES-4106 (Fisc14, Fisc2); *Gloeothece citrififormis* PCC 7424 (Gcit14.1, Gcit14.2, Gcit2); *Gloebacter violaceus* PCC 7421 (Gvio14, Gvio2); *Halomicronema hongdechloris* C2206 (Hhon14, Hhon2); *Leptolyngbya boryana* dg5 (Lbor14, Lbor2); *Microcystis aeruginosa* NIES-843 (Maer14); *Nostoc punctiforme* PCC 73102 (Npun14, Npun2); *Oscillatoria acuminata* PCC 6304 (Oacu14, Oacu2); *Stanieria cyanosphaera* PCC 7437 (Scya14); *Synechococcus elongatus* PCC 11801 (Sync14, Sync2); and *Synechocystis* sp. PCC 6803 (Syne14).

which contain 14–20% of total disordered residues and are therefore closer to the prokaryotic proteome average (Figure 3B). This indicates that the HDA2 and HDA14 plant proteins conserve a sequence fingerprint related to their bacterial origin. Previous reports suggested that the relationship between eukaryote HDAC/HDAs and bacterial homologs is uncertain, and that multiple horizontal transfer events might have taken place [19]. Interestingly, the **cladogram** in Figure 3C shows that plant HDA2 and HDA14 are closely related to cyanobacterial histone deacetylase proteins (bootstrap >78), which could suggest that plant HDA genes were incorporated via endosymbiosis and subsequent transfer from the bacterial plasmid to the nuclear genome (plastid origin). Not surprisingly, 800–2000 genes in the arabidopsis genome are believed to have a cyanobacterial origin [47,48]. Moreover, HDA14 has been localized in the chloroplast in arabidopsis, where it deacetylates photosynthetic proteins [17].

The Zn²⁺-Dependent Active Site in Plant HDAs

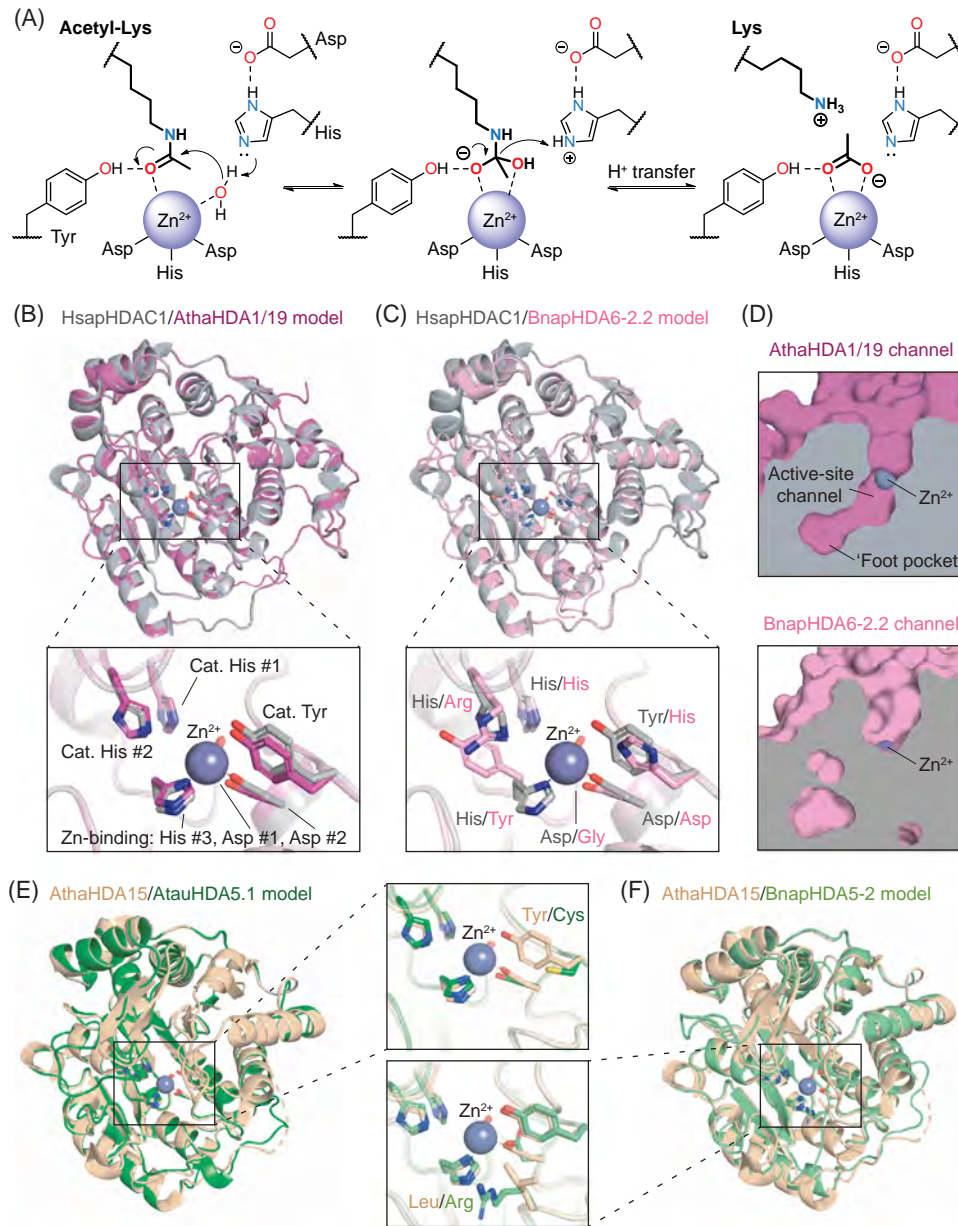
The sequence organization of HDACs/HDAs has been extensively reviewed [20,49–51]. In plants, bioinformatics tools have enabled prediction of HDA sequence architectures, as displayed in the Pfam and Interpro databases [25,29]. Class I, II, and IV HDAs harbor the conserved histone deacetylase domain (313–330 aa), which covers 60–75% of the entire sequence, and often a nuclear localization signal (NLS) at the N-terminus. An additional C-terminal NLS motif can be observed in class I HDA1/19 and class II HDA15. Conversely, the class II enzymes HDA8, HDA15, the single Brassicaceae HDA5-2 (before HDA18), and the class IV member HDA2 all contain nuclear export signal (NES) motifs [25]. HDA15 contains a zinc-finger RanBP2 domain located close to the N-terminus, which is also found in mammalian HDAC6 and has been linked to ubiquitin binding and targeting for lysosomal degradation [52,53].

The HDAC/HDA catalytic domain has been characterized by analysis of several crystal structures [51,54,55] which helped to identify residues that are essential for Zn²⁺ cofactor binding and amide-bond hydrolysis (Figure 4A) [54–57]. In brief, the Zn²⁺ ion is bound to contiguous Asp and His residues, and to an additional Asp residue in a separate loop. Once the N^ε-acetyl-lysine substrate is bound in the active site, a water molecule attacks the amide and forms a tetrahedral intermediate that is stabilized by a Tyr residue (Figure 4A). This step is also catalyzed by two His-Asp dyads, of which one appears to serve as a proton-exchange catalyst whereas the other helps to stabilize the transition state [57]. The cycle concludes with formation of the Lys and acetate products, assisted by the protonated His-Asp dyad (Figure 4A). Most HDA protein sequences analyzed in this review conserve these essential residues. However, some exceptions in each class are worth discussing.

Class I HDAs

Class I plant HDAs are, on average, ~470 residues in length, and they share up to 58–68% sequence identity with human class I HDACs 1–3 and 40–50% with HDAC8. Plant sequences conserve (i) the catalytic residues His148, His149, and Tyr311 (referred to AthaHDA1/19, At4g38130, Uniprot O22446), that are equivalents to His140, His141, and Tyr303 in human HDAC1 (Uniprot Q13547) [56–58], in which the catalytic Tyr is preceded by a conserved poly (Gly) motif; and (ii) the Zn²⁺-binding residues Asp184, His186, and Asp272 that are equivalents to Asp176, His178, and Asp264 [58]. Surprisingly, the HDA1/19 orthologs Atau M8BTX9, Macu M0SW61, and Macu M0S9F0 are exceptions to these rules because they have lost the catalytic residues equivalent to His148 and His149, indicating complete loss of deacetylase activity (Figure 1B, black circles).

Of particular note are mutations that are present in Brassicaceae HDA6-2 orthologs (before HDA7, Figure 1A, light-blue circles), such as substitution of the catalytic Tyr residue by His in



Trends In Plant Science

Figure 4. 3D Structural Alignment and Key Active-Site Mutations In Histone Deacetylases. (A) Mechanism of catalysis of Zn^{2+} -dependent histone deacetylases, showing residues involved in Zn^{2+} binding, transition state stabilization, and proton transfer [57]. The last step involves two proton transfers, one proton from the His-Asp dyad and a second proton from the acetate moiety. A single His-Asp dyad is shown for simplicity. (B) Superposition of AthaHDA1/19 (O22446, e-value 4.2×10^{-88}) with the X-ray crystal structure of *Homo sapiens* (Hsap) HDAC1 (PDB 4BKX, [58]). Amino acid positions are, for HsapHDAC1: His140 (#1), His141 (#2), Asp176 (#1), His178 (#3), Asp264 (#2), and Tyr303; and for AthaHDA1/19: His148 (#1), His149 (#2), Asp184 (#1), His186 (#2), Asp272 (#2), and Tyr311. (C) Superposition of BnapHDA6-2.2 (XP_013737742, e-value 7.3×10^{-80}) with the X-ray crystal structure of HsapHDAC1 (PDB 4BKX [58]). BnapHDA6-2.2 amino acid positions are: His150 (His/His), Arg151, Gly186, Tyr188 (His/Tyr), Asp272, and His311 (Tyr/His). Abbreviation: Cat., catalytic residue. (D) Section view of the active-site channel in AthaHDA1/19 and BnapHDA6-2.2. The conserved position of the Zn^{2+} ion is shown for BnapHDA6-2.2, although the lack of coordinating residues makes its presence unlikely. (E) Superposition of AtauHDA5.1 (XP_020157471, e-value 1.7×10^{-52}) with the X-ray crystal structure

(Figure legend continued at the bottom of the next page.)

Bnap (HDA6-2.1: A0A078IDK3 and HDA6-2.2: XP_013737742), Bole (XP_013603394), and Brap (XP_033129960) (Figure S3). This modification resembles the active-site mutation that is characteristic of class IIa mammalian HDACs, and which abolishes deacetylase activity [59,60]. However, in these four HDA6-2 proteins it is accompanied by further mutations in the equivalents to the catalytic His149 (H-to-R), the Zn²⁺-binding residues Asp184 (D-to-G) and His186 (H-to-Y, only in Bole XP_013603394 and in Bnap XP_013737742), and a Gly residue in the poly(Gly) motif (G-to-Q, Figure S3). Although there is no experimental data on the catalytic activity of HDA6-2 proteins, it appears unlikely that these four proteins act as deacetylases. The HDA6-2 orthologs of *Atha* (Q9FH09) and *Crub* (XP_006293194) also contain mutations in the Zn²⁺-binding residues equivalent to Asp184 (D-to-G) and His186 (H-to-P). Therefore, we expect that the HDA6-2 proteins are not catalytically active histone deacetylase enzymes and that they might have adopted a different role in Brassicaceae biology. Multiple histone deacetylases exert their function in multiprotein complexes that assemble through protein–protein interactions [61,62]. In this regard, it is also remarkable that Brassicaceae HDA6-2 proteins lack the long flexible C-terminal tail that is characteristic of subclass HDA6, and this would likely reduce their ability to interact with protein partners. These features could have resulted from high selection pressure during plant speciation, affording a novel function. Future studies will assess whether HDA6-2 proteins fulfill different biological roles than HDA6-1 and earlier HDA6 members.

X-ray crystal structures of class I HDACs have been recently determined at high resolution from diverse bacteria and mammals, most of them fixed by cocrystallization through inhibitor binding [51,61]. These structural studies have been useful in understanding the configuration of the HDAC active site and the mechanism of catalysis [54]. No class I crystal structures have so far been reported in plants, but these structures could help to advance the plant HDA field through 3D structural alignment, similarly to how the bacterial HDLP structure launched research on mammalian HDACs [63].

Superposition of class I *Atha*HDA1/19 (O22446, Figure 4B) and of the non-canonical BnapHDA6-2.2 (XP_013737742, Figure 4C) with human HDAC1 (PDB 4BKX, 67% and 49% identity, respectively) [58] reveals high correlation in the structure of the catalytic domain, which indicates that homology models could be built via thorough refinement. In fact, the predicted catalytic and Zn²⁺-binding residues in *Atha*HDA1/19 are at the exact positions of those in human HDAC1 (Figure 4B) [56,58,61]. The Zn²⁺ cofactor appears to be buried at the bottom of the characteristic 11 Å hydrophobic channel that accommodates the N^ε-acetyl-lysine substrate [56,57]. This channel is extended into a ‘foot pocket’ that is only present in class I HDACs in mammals [51,64], but this could be because it was present in the crystal structure that was employed for alignment. In the case of BnapHDA6-2.2, the mutations in catalytic and Zn²⁺-binding residues completely reshape the active-site channel and would not accommodate the Zn²⁺ cofactor (Figure 4D). Thus, the 3D alignment supports a different biochemical activity and biological role for the Brassicaceae-specific HDA6-2 proteins.

Class II HDAs

Class II HDAs are, on average, ~500 residues in length and share up to 30–45% sequence identity with their human homologs (class IIa: HDACs 4, 5, 7, 9; class IIb: HDACs 6 and 10). All plant sequences have one catalytic domain, in contrast to mammalian class IIb HDACs, of

of *Atha*HDA15 (PDB 6J6T). Amino acid positions are, for *Atha*HDA15: His276 (#1), His277 (#2), Asp313 (#1), His315 (#3), Asp404 (#2), and Tyr444; and for *Ata*HDA5.1: His101 (#1), His102 (#2), Asp138 (#1), His140 (#2), Asp227 (#2), and Cys267. (F) Superposition of BnapHDA5-2 (A0A078HYV8, e-value 1.9×10^{-47}) with the X-ray crystal structure of *Atha*HDA15 (PDB 6J6T). BnapHDA5-2 amino acid positions are: His140 (#1), His141 (#2), Asp163 (#1), His165 (#2), Asp228 (#2), Tyr268, and the ‘gatekeeper’ Arg235 (Leu411 in *Atha*HDA15). Superpositions were obtained from HHpred (<https://toolkit.tuebingen.mpg.de/>) [121,122].

which HDAC6 has two catalytic domains and HDAC10 has one complete and one partial domain [51]. In plants, class II HDAs conserve (i) the catalytic residues His276, His277, and Tyr444 (referred to AthaHDA15, At3g18520, Uniprot F4J8S1) that are equivalent to His610, His611, and Tyr782 in the catalytic domain 2 (CD2) of human HDAC6 (Uniprot Q9UBN7) [65], and (ii) the Zn²⁺-binding residues Asp313, His315, and Asp404 that are equivalent to Asp649, His651, and Asp742. Surprisingly, the HDA5 ortholog of grass Atau (XP_020157471) contains a Cys residue in place of the catalytic Tyr, a difference that has not been reported to date in any organism (Figure 2B, red circle).

The catalytic site of most class II HDAs, similarly to class I and class IV, includes a Leu residue at the rim of the active site. This position has been found to be substituted by Lys in catalytic domain 1 (CD1) of mammalian HDAC6 [65], and by Glu in HDAC10 [12]. In these two cases the alternative amino acid serves as a 'gatekeeper' that provides substrate selectivity towards acetylated C-terminal Lys residues and polyamines, respectively [12,65]. In plants, Leu411 (referred to AthaHDA15) is changed to Gln in subclass HDA8, and to Arg in variants of Brassicaceae subclass HDA5 of Bnap (A0A078HYV8), Bole (VDD27184), and Brap (XP_009130142) (Figure 2A, red circles), but no Lys or Glu variants are found.

The sole 3D structure reported of a plant HDA is of the class II AthaHDA15 (PDB 6J6T, 2.36 Å resolution) [55]. This X-ray crystal structure shows tetrameric oligomerization for the histone deacetylase domain, and each monomer contains a typical ~10 Å channel with the catalytic Zn²⁺ ion at the bottom (Figure 4E). In addition, the N-terminal zinc-finger domain (aa 1–146) assists and stabilizes the dimerization of the histone deacetylase domain *in vitro* and *in vivo*, similarly to mammalian HDACs 1 and 2 [61,66], and is essential for its enzymatic activity [55]. As expected, the Zn²⁺ cofactor is bound to Asp313, His315, and Asp404, by analogy to HDAC6_CD2 (PDB 5EFN) [65], in which the predicted catalytic Tyr side chain is pointed towards it. Accordingly, Asp313, His315, and also His276 and His277, have been shown to be essential for the enzymatic activity of AthaHDA15 [55,67]. Superposition of the HDA5 ortholog of the grass Atau (XP_020157471, 60% identity) with the reported AthaHDA15 structure reveals that the active site Cys residue (Cys267) is in a position and orientation equivalent to that of the missing catalytic Tyr (Figure 4E). Because the corresponding Tyr is involved in stabilization of the tetrahedral intermediate formed during acetamide substrate hydrolysis [57], it seems unlikely that AtauHDA5 would be an active deacetylase, unlike close orthologs from arabidopsis and other species [68]. Superposition of BnapHDA5-2 (A0A078HYV8, 59% identity) with the structure of AthaHDA15 shows that the conserved Leu at the entrance of the active site is mutated to Arg (Arg234), and is positioned where it could act as 'gatekeeper' (Figure 4F). This could indicate that Brassicaceae HDA5-2 variants have a substrate specificity more similar to that of mammalian HDAC6_CD1, which harbors a Lys residue at the same position [65].

Class IV HDAs

Plant class IV HDAs are homologs of mammalian HDAC11 [19,25], are on average ~360 residues in length, and share up to 60% sequence identity with the human homolog. These sequences conserve (i) the catalytic residues His200, His201, and Tyr361 (referred to AthaHDA11, At5g26040) that are equivalent to His141, His142, and Tyr304 in human HDAC11 (Uniprot Q96DB2), and (ii) the Zn²⁺-binding residues Asp238, His240, and Asp318, the equivalents of Asp181, His183, and Asp261. Remarkable exceptions are the monocot Doli (OEL35364) and the non-legume symbiotic Pand (PON76711) orthologs that harbor mutations of the catalytic Tyr residues (Tyr267Asp and Tyr323Arg, respectively). These mutations in the active site have not been reported previously, and would likely cause loss of deacetylase activity for these enzymes or indicate an alternative catalytic mechanism. Interestingly, Pand diverged from legumes >100 million years ago and is capable of fixing N₂ via symbiosis [69].

Biological Function of HDAs in Plants

Plant HDAs were first extracted in 1988 [70], and their inhibition was key for studying plant histone acetylation in the 1990s [71–73]. Studies since then have focused on their mRNA expression, sub-cellular localization, and the elucidation of their function via genetic and transcriptional manipulation [10,25,74–76]. By contrast, the biochemical evaluation of their catalytic activity and substrate preferences has not received as much attention until recent years [55,77,78]. Because the biological significance of HDAs in plants has been reviewed extensively [13,62,79–81], we only include a brief summary here (Table 1).

The class I enzymes HDA1/19, HDA6, and HDA9, which have attracted most attention, are localized in the nucleus and are heavily associated with chromatin remodeling [62,80,81]. In particular, HDA6 deacetylates histones at transposable elements and helps to maintain DNA methylation at those sites, which results in gene silencing [82–84]. This function is regulated by Ser phosphorylation at the C-terminus of HDA6, which promotes its interaction with histone methyltransferases that act

Table 1. Subcellular Localization and Representative Functions of Plant HDAs

Protein	Localization	Function	Refs
Class I			
HDA1/19	Nucleus	Circadian rhythm regulation Flowering time Immune response Phytohormone signaling Plant growth Seed maturation and germination Stress response	[97] [88,89,94] [87,109] [88,92,95] [44,88] [92,93,110] [22,76,88]
HDA6-1	Nucleus	Chromatin silencing Circadian rhythm regulation Flowering time Phytohormone signaling Plant growth Stress response	[78,82–84] [85,86,90,97] [88,89,94] [88,95,111] [44,88] [88]
HDA6-2 (HDA7)	Nucleus	Plant growth Seed maturation and germination	[112] [112]
HDA9	Nucleus	Flowering time Immune response Phytohormone signaling Stress response Thermomorphogenesis	[6,99] [102] [101,103] [98,100,101,104] [103]
Class II			
HDA5-1	Nucleus Cytoplasm	Flowering time	[68]
HDA5-2 (HDA18)	Nucleus Cytoplasm	Root development	[74]
HDA8	Nucleus	Plant growth	[113]
HDA14	Chloroplast Mitochondria	Microtubule dynamics Photosynthesis regulation	[10] [17]
HDA15	Nucleus Cytoplasm	Photomorphogenesis Photosynthesis regulation Seed germination Stress response	[67] [106] [75] [76]
Class IV			
HDA2	Nucleus	Phytohormone signaling Stress response	[114] [114]

on deacetylated histones [78]. HDA6 also interacts with the histone demethylases LDL1 and LDL2 to repress the transcription of circadian rhythm-related genes [85,86]. In addition, HDA1/19 and HDA6 interact directly with a large number of transcription factors and chromatin modifiers [44,87–95], in clear analogy to the mammalian network of epigenetic deacetylase complexes [58,61,96]. In arabidopsis and other plants, HDA1/19 and HDA6 bind to the histone deacetylation complex 1 (HDC1) protein [88], which is the centerpiece for multiple protein–protein interactions with histone- and RNA-binding proteins [44,94], and to Sin3-like transcriptional regulators [94,95]. HDC1-containing complexes regulate phytohormone signaling including abscisic acid (ABA) [88,95], fruit cell proliferation [95], flowering time [94], and organ size [88]. HDA1/19 and HDA6 also form complexes with the corepressor TOPLESS (TPL) and with pseudo-response regulators (PRRs) involved in the circadian cycle [90,97]. Interestingly, HDA1/19 forms the brassinosteroid signaling-related factor (BES1)–TPL–HDA1/19 complex that also regulates ABA signaling [92].

HDA9, although less extensively characterized, has also been proposed to form complexes with a repressive epigenetic function [98–101]. HDA9 interacts with the SANT domain-containing powerdress (PWR) and with the highly expressed osmotically responsive gene 15 (HOS15) proteins in the multifunctional HDA9–PWR–HOS15 complex [6,81,98,102]. HDA9 and PWR can also act as transcriptional activators in the context of thermomorphogenesis where they act as inducers of the synthesis of the phytohormone auxin [103]. Other epigenetic protein partners connected to HDA9 are the transcription factors VAL1 and VAL2, that are related to flowering time [99], and the ABA-insensitive 4 (ABI4) and WRKY53 factors that regulate stress responses [100,101,104]. Remarkably, the literature on HDA9 has increased greatly over recent years (and even during the preparation of this review), which highlights the rapid development of the field (Table 1) [81].

Class II HDAs are less well understood, but they appear to fulfill highly diverse functions across several cell compartments (Table 1) [26]. HDA14 was first identified to deacetylate α -tubulin in the cytoplasm [10], similarly to HDAC6 in mammalian cells [105]. Alternatively, HDA14 has been found in mitochondria and chloroplasts [17,26] where it can regulate proteins involved in photosynthesis via deacetylation [17]. HDA15 and HDA5 are found both in the cytoplasm [26] and the nucleus, where they overlap in function with class I HDAs and modulate genes involved in the temperature stress response [76] and flowering [68]. Similarly to class I HDAs, HDA15 forms complexes with transcription factors in arabidopsis. For instance, HDA15 interacts with the phytochrome-interacting factors 1 (PIF1) and 3 (PIF3) to repress seed germination and chlorophyll biosynthesis-related genes, respectively [75,106]. HDA15 is also recruited by elongated hypocotyl 5 (HY5) to regulate light-responsive genes and photomorphogenesis [67]. Interestingly, HDA15 can shuttle between the nucleus and the cytoplasm in response to light [26]. HDA5, on the other hand, relies on phosphorylation and interaction with 14-3-3 proteins for cytoplasmic localization [26], similarly to mammalian class IIa HDACs [107,108]. HDA15 has also been reported recently to be phosphorylated at the C-terminus but, in this case, phosphorylation alters its activity and subnuclear localization [55]. Both the class II HDA8 and the class IV HDA2, that are localized in the nucleus, remain largely uncharacterized (Table 1).

Concluding Remarks and Future Perspectives

Studies over the past decade have revealed that plant HDAs play crucial roles in diverse biological processes [62]. HDAs take part in epigenetic regulatory complexes and modulate development, flowering, the circadian clock, and the response to different stresses [13,62,80,81]. However, their mechanisms of action remain largely unknown, in part owing to the lack of structural and biochemical studies. In this regard, the recent development of functional assays [12,96,115] and crystallization protocols [12,58,65,116] for mammalian HDACs could be translated to plant

Outstanding Questions

Have Brassicaceae HDAs evolved independently of those of other dicots? If so, are they similar to canonical HDAs or are they involved in processes that are not observed in other species?

What evolutionary and adaptive advantages have HDA9b isozymes provided in the Poaceae family of monocots? To what extent has the increase of ductile regions in HDA9b influenced their functionality?

Which additional structural properties and functionalities led to the gain of ductile regions during plant HDA evolution?

How are plant HDAs structured in space? Are their catalytic domains similar to the reported 3D structures of human HDACs?

Is the HDAC catalytic mechanism conserved in plant HDAs? What N^{ϵ} -acetyllysine substrates are hydrolyzed by each of the individual isoenzymes?

What are the biochemical and functional consequences of the mutations that are found in key catalytic residues of class I and class II HDAs?

What are the functions of HDA2 and, overall, of class IV HDAs? Are they involved in plant epigenetics and development, or do they regulate alternative processes?

HDA and provide important insight. In this review, we provide a foundation for these advances by presenting the overall HDA phylogeny and highlighting key structural variants that warrant further investigation (see Outstanding Questions). It is our hope that the revised and updated phylogeny will also set a standard for future annotation and classification of plant HDAs. In addition, evolutionary analyses reveal, first, that Brassica species express two additional and functionally distinct HDAs, and second, that the intriguing HDA2 and HDA14 proteins retain structural characteristics from their bacterial origin. The collective findings presented in this review will underpin the elucidation of the activity and role of each individual plant HDA and, ultimately, facilitate the application of plant epigenetics to solve future challenges in ecology and crop production.

Acknowledgments

This work was supported by the Government of Aragón-European Regional Development Fund (FEDER; grant E35_17R to I. Y.).

Supplemental Information

Supplemental information associated with this article can be found online at <https://doi.org/10.1016/j.tplants.2020.12.011>.

References

- Choudhary, C. *et al.* (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834–840
- Haberland, M. *et al.* (2009) The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat. Rev. Genet.* 10, 32–42
- Falkenberg, K.J. and Johnstone, R.W. (2014) Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat. Rev. Drug Discov.* 13, 673–691
- Goldberg, A.D. *et al.* (2007) Epigenetics: a landscape takes shape. *Cell* 128, 635–638
- Meier, K. and Brehm, A. (2014) Chromatin regulation: how complex does it get? *Epigenetics* 9, 1485–1495
- Park, H.J. *et al.* (2019) HOS15 interacts with the histone deacetylase HDA9 and the Evening complex to epigenetically regulate the floral activator GIGANTEA. *Plant Cell* 31, 37–51
- Hebbes, T.R. *et al.* (1988) A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* 7, 1395–1402
- Arrowsmith, C.H. *et al.* (2012) Epigenetic protein families: a new frontier for drug discovery. *Nat. Rev. Drug Discov.* 11, 384–400
- Wang, S. *et al.* (2015) Insulin and mTOR pathway regulate HDAC3-mediated deacetylation and activation of PGK1. *PLoS Biol.* 13, e1002243
- Tran, H.T. *et al.* (2012) *Arabidopsis thaliana* histone deacetylase 14 (HDA14) is an α -tubulin deacetylase that associates with PP2A and enriches in the microtubule fraction with the putative histone acetyltransferase ELP3. *Plant J.* 71, 263–272
- Ustinova, K. *et al.* (2020) The disordered N-terminus of HDAC6 is a microtubule-binding domain critical for efficient tubulin deacetylation. *J. Biol. Chem.* 295, 2614–2628
- Hai, Y. *et al.* (2017) Histone deacetylase 10 structure and molecular function as a polyamine deacetylase. *Nat. Commun.* 8, 15368
- Liu, X. *et al.* (2014) Transcriptional repression by histone deacetylases in plants. *Mol. Plant* 7, 764–772
- Shen, Y. *et al.* (2015) Histone acetylation enzymes coordinate metabolism and gene expression. *Trends Plant Sci.* 20, 614–621
- Gallinari, P. *et al.* (2007) HDACs, histone deacetylation and gene transcription: from molecular biology to cancer therapeutics. *Cell Res.* 17, 195–211
- Verdin, E. and Ott, M. (2015) 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat. Rev. Mol. Cell Biol.* 16, 258–264
- Hartl, M. *et al.* (2017) Lysine acetylome profiling uncovers novel histone deacetylase substrate proteins in *Arabidopsis*. *Mol. Syst. Biol.* 13, 949
- Park, S.-Y. and Kim, J.-S. (2020) A short guide to histone deacetylases including recent progress on class II enzymes. *Exp. Mol. Med.* 52, 204–212
- Gregoret, I.V. *et al.* (2004) Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. *J. Mol. Biol.* 338, 17–31
- Yang, X.J. and Seto, E. (2008) The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nat. Rev. Mol. Cell Biol.* 9, 206–218
- Haigis, M.C. and Guarente, L.P. (2006) Mammalian sirtuins – emerging roles in physiology, aging, and calorie restriction. *Genes Dev.* 20, 2913–2912
- Ueda, M. *et al.* (2019) Transcriptome analysis of the hierarchical response of histone deacetylase proteins that respond in an antagonistic manner to salinity stress. *Front. Plant Sci.* 10, 1323
- Rossi, V. *et al.* (1998) Identification and characterisation of an RPD3 homologue from maize (*Zea mays* L.) that is able to complement an rpd3 null mutant of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 258, 288–296
- Hollender, C. and Liu, Z. (2008) Histone deacetylase genes in *Arabidopsis* development. *J. Integr. Plant Biol.* 50, 875–885
- Alinsug, M.V. *et al.* (2009) Phylogenetic analysis, subcellular localization, and expression patterns of RPD3/HDA1 family histone deacetylases in plants. *BMC Plant Biol.* 9, 37
- Alinsug, M.V. *et al.* (2012) Subcellular localization of class II HDAs in *Arabidopsis thaliana*: nucleocytoplasmic shuttling of HDA15 is driven by light. *PLoS One* 7, e30846
- Chu, J. and Chen, Z. (2018) Molecular identification of histone acetyltransferases and deacetylases in lower plant *Marchantia polymorpha*. *Plant Physiol. Biochem.* 132, 612–622
- Yang, C. *et al.* (2018) Characterization and subcellular localization of histone deacetylases and their roles in response to abiotic stresses in soybean. *BMC Plant Biol.* 18, 226
- Imran, M. *et al.* (2020) Histone deacetylase (HDAC) gene family in allotetraploid cotton and its diploid progenitors: in silico identification, molecular characterization, and gene expression analysis under multiple abiotic stresses, DNA damage and phytohormone treatments. *Int. J. Mol. Sci.* 21, 321
- Yuan, L. *et al.* (2020) Genome-wide identification of the HDAC family proteins and functional characterization of CsHD2C, a HD2-type histone deacetylase gene in tea plant (*Camellia sinensis* L. O. Kuntze). *Plant Physiol. Biochem.* 155, 898–913
- Cigliano, R.A. *et al.* (2013) Genome-wide analysis of histone modifiers in tomato: gaining an insight into their developmental roles. *BMC Genomics* 14, 57
- Pandey, R. *et al.* (2002) Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana*

- suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Res.* 30, 5036–5055
33. Guo, Y.L. *et al.* (2009) Recent speciation of *Capsella rubella* from *Capsella grandiflora*, associated with loss of self-incompatibility and an extreme bottleneck. *Proc. Natl. Acad. Sci. U. S. A.* 106, 5246–5251
 34. Donoghue, M.T. *et al.* (2011) Evolutionary origins of Brassicaceae specific genes in *Arabidopsis thaliana*. *BMC Evol. Biol.* 11, 47
 35. Liu, S. *et al.* (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5, 3930
 36. Delwiche, C.F. and Cooper, E.D. (2015) The evolutionary origin of a terrestrial flora. *Curr. Biol.* 25, R899–R910
 37. Nishiyama, T. *et al.* (2018) The *Chara* genome: secondary complexity and Implications for plant terrestrialization. *Cell* 174, 448–464
 38. Yruela, I. *et al.* (2017) Evidence for a strong correlation between transcription factor protein disorder and organismic complexity. *Genome Biol. Evol.* 9, 1248–1265
 39. Niklas, K.J. *et al.* (2018) The evolutionary origins of cell type diversification and the role of intrinsically disordered proteins. *J. Exp. Bot.* 69, 1437–1446
 40. Schadt, E. *et al.* (2011) The relationship between proteome size, structural disorder and organism complexity. *Genome Biol.* 12, R120
 41. Xue, B. *et al.* (2012) Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Struct. Dyn.* 30, 137–149
 42. Oldfield, C.J. and Dunker, A.K. (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* 83, 553–584
 43. Yruela, I. and Contreras-Moreira, B. (2012) Protein disorder in plants: a view from the chloroplast. *BMC Plant Biol.* 12, 165
 44. Perrella, G. *et al.* (2016) The histone deacetylase complex 1 protein of *Arabidopsis* has the capacity to interact with multiple proteins including histone 3-binding proteins and histone 1 variants. *Plant Physiol.* 171, 62–70
 45. Kagale, S. *et al.* (2014) Polyploid evolution of the Brassicaceae during the Cenozoic era. *Plant Cell* 26, 2777–2791
 46. Parkin, I.A. *et al.* (2014) Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biol.* 15, R77
 47. Rujan, T. and Martin, W. (2001) How many genes in *Arabidopsis* come from cyanobacteria? An estimate from 386 protein phylogenies. *Trends Genet.* 17, 113–120
 48. Stegemann, S. *et al.* (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8828–8833
 49. Marmorstein, R. (2001) Structure of histone deacetylases: insights into substrate recognition and catalysis. *Structure* 9, 1127–1133
 50. Yang, X.-J. and Seto, E. (2003) Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression. *Curr. Opin. Genet. Dev.* 13, 143–153
 51. Maolanon, A.R. *et al.* (2016) Innovative strategies for selective inhibition of histone deacetylases. *Cell Chem. Biol.* 23, 759–768
 52. Hook, S.S. *et al.* (2002) Histone deacetylase 6 binds polyubiquitin through its zinc finger (PAZ domain) and copurifies with deubiquitinating enzymes. *Proc. Natl. Acad. Sci. U. S. A.* 99, 13425–13430
 53. Ferreira de Freitas, R. *et al.* (2018) Identification and structure–activity relationship of HDAC6 zinc-finger ubiquitin binding domain inhibitors. *J. Med. Chem.* 61, 4517–4527
 54. Porter, N.J. and Christianson, D.W. (2019) Structure, mechanism, and inhibition of the zinc-dependent histone deacetylases. *Curr. Opin. Struct. Biol.* 59, 9–18
 55. Chen, C.Y. *et al.* (2020) Structure of *Arabidopsis* HISTONE DEACETYLASE 15. *Plant Physiol.* 184, 1585–1600
 56. Weerasinghe, S.V.W. *et al.* (2008) Residues in the 11 Å channel of histone deacetylase 1 promote catalytic activity: implications for designing isoform-selective histone deacetylase inhibitors. *J. Med. Chem.* 51, 5542–5551
 57. Gantt, S.M. *et al.* (2016) General base-general acid catalysis in human histone deacetylase 8. *Biochemistry* 55, 820–832
 58. Millard, C.J. *et al.* (2013) Class I HDACs share a common mechanism of regulation by inositol phosphates. *Mol. Cell* 51, 57–67
 59. Lahm, A. *et al.* (2007) Unraveling the hidden catalytic activity of vertebrate class IIa histone deacetylases. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17335–17340
 60. Bradner, J.E. *et al.* (2010) Chemical phylogenetics of histone deacetylases. *Nat. Chem. Biol.* 6, 238–243
 61. Millard, C.J. *et al.* (2017) Targeting class I histone deacetylases in a ‘complex’ environment. *Trends Pharmacol. Sci.* 38, 363–377
 62. Chen, X. *et al.* (2020) Functions and mechanisms of plant histone deacetylases. *Sci. China Life Sci.* 63, 206–216
 63. Finnin, M.S. *et al.* (1999) Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature* 401, 188–193
 64. Bressi, J.C. *et al.* (2010) Exploration of the HDAC2 foot pocket: synthesis and SAR of substituted N-(2-aminophenyl) benzamides. *Bioorg. Med. Chem. Lett.* 20, 3142–3145
 65. Hai, Y. and Christianson, D.W. (2016) Histone deacetylase 6 structure and molecular basis of catalysis and inhibition. *Nat. Chem. Biol.* 12, 741–747
 66. Taplick, J. *et al.* (2001) Homo-oligomerisation and nuclear localisation of mouse histone deacetylase 1. *J. Mol. Biol.* 308, 27–38
 67. Zhao, L. *et al.* (2019) HY5 interacts with the histone deacetylase HDA15 to repress hypocotyl cell elongation in photomorphogenesis. *Plant Physiol.* 180, 1450–1466
 68. Luo, M. *et al.* (2015) Regulation of flowering time by the histone deacetylase HDA5 in *Arabidopsis*. *Plant J.* 82, 925–936
 69. Dupin, S.E. *et al.* (2019) The non-legume *Parasponia andersonii* mediates the fitness of nitrogen-fixing rhizobial symbionts under high nitrogen conditions. *Front. Plant Sci.* 10, 1779
 70. Sendra, R. *et al.* (1988) Characterization of pea histone deacetylases. *Plant Mol. Biol.* 11, 857–866
 71. Waterborg, J.H. *et al.* (1990) Dynamic histone acetylation in alfalfa cells. Butyrate interference with acetate labeling. *Biochim. Biophys. Acta* 1049, 324–330
 72. Georgieva, E.I. *et al.* (1991) Histone acetylation in *Zea mays*. II. Biological significance of post-translational histone acetylation during embryo germination. *J. Biol. Chem.* 266, 18751–18760
 73. Brosch, G. *et al.* (1995) Inhibition of maize histone deacetylases by HC toxin, the host-selective toxin of *Cochliobolus carbonum*. *Plant Cell* 7, 1941–1950
 74. Liu, C. *et al.* (2013) HDA18 affects cell fate in *Arabidopsis* root epidermis via histone acetylation at four kinase genes. *Plant Cell* 25, 257–269
 75. Gu, D. *et al.* (2017) Identification of HDA15-PIF1 as a key repression module directing the transcriptional network of seed germination in the dark. *Nucleic Acids Res.* 45, 7137–7150
 76. Shen, Y. *et al.* (2019) *Arabidopsis* histone deacetylase HDA15 directly represses plant response to elevated ambient temperature. *Plant J.* 100, 991–1006
 77. Fong, P.M. *et al.* (2006) *Arabidopsis thaliana* histone deacetylase 1 (AtHD1) is localized in euchromatic regions and demonstrates histone deacetylase activity in vitro. *Cell Res.* 16, 479–488
 78. Yu, C.W. *et al.* (2017) HISTONE DEACETYLASE 6 acts in concert with histone methyltransferases SUVH4, SUVH5, and SUVH6 to regulate transposon silencing. *Plant Cell* 29, 1970–1983
 79. Ma, X. *et al.* (2013) Histone deacetylases and their functions in plants. *Plant Cell Rep.* 32, 465–478
 80. Luo, M. *et al.* (2017) Plant responses to abiotic stress regulated by histone deacetylases. *Front. Plant Sci.* 8, 2147
 81. de Rooij, P.G.H. *et al.* (2020) The diverse and unanticipated roles of histone deacetylase 9 in coordinating plant development and environmental acclimation. *J. Exp. Bot.* 71, 6211–6225
 82. Probst, A.V. *et al.* (2004) *Arabidopsis* histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats. *Plant Cell* 16, 1021–1034
 83. To, T.K. *et al.* (2011) *Arabidopsis* HDA6 regulates locus-directed heterochromatin silencing in cooperation with MET1. *PLoS Genet* 7, e1002055

84. Liu, X. *et al.* (2012) HDA6 directly interacts with DNA methyltransferase MET1 and maintains transposable element silencing in *Arabidopsis*. *Plant Physiol.* 158, 119–129
85. Hung, F.Y. *et al.* (2018) The *Arabidopsis* LDL1/2–HDA6 histone modification complex is functionally associated with CCA1/LHY in regulation of circadian clock genes. *Nucleic Acids Res.* 46, 10669–10681
86. Hung, F.Y. *et al.* (2019) The LDL1/2–HDA6 histone modification complex interacts with TOC1 and regulates the core circadian clock components in *Arabidopsis*. *Front. Plant Sci.* 10, 233
87. Kim, K.C. *et al.* (2008) *Arabidopsis* WRKY38 and WRKY62 transcription factors interact with histone deacetylase 19 in basal defense. *Plant Cell* 20, 2357–2371
88. Perrella, G. *et al.* (2013) Histone deacetylase complex 1 expression level titrates plant growth and abscisic acid sensitivity in *Arabidopsis*. *Plant Cell* 25, 3491–3505
89. Gu, X. *et al.* (2013) Photoperiodic regulation of flowering time through periodic histone deacetylation of the florigen gene FT. *PLoS Biol.* 11, e1001649
90. Wang, L. *et al.* (2013) Transcriptional corepressor TOPLESS complexes with pseudoresponse regulator proteins and histone deacetylases to regulate circadian transcription. *Proc. Natl. Acad. Sci. U. S. A.* 110, 761–766
91. Gao, M.J. *et al.* (2004) An auxin-responsive SCARECROW-like transcriptional activator interacts with histone deacetylase. *Plant Mol. Biol.* 55, 417–431
92. Ryu, H. *et al.* (2014) Control of early seedling development by BES1/TPL/HDA19-mediated epigenetic regulation of ABI3. *Nat. Commun.* 5, 4138
93. Gao, M.J. *et al.* (2015) SCARECROW-LIKE 15 interacts with HISTONE DEACETYLASE 19 and is essential for repressing the seed maturation programme. *Nat. Commun.* 6, 7243
94. Ning, Y.-Q. *et al.* (2019) The HDA19 histone deacetylase complex is involved in the regulation of flowering time in a photoperiod-dependent manner. *Plant J.* 98, 448–464
95. Zhang, Z. *et al.* (2020) Genome-wide target mapping shows histone deacetylase complex1 regulates cell proliferation in cucumber fruit. *Plant Physiol.* 182, 167–184
96. Song, Y. *et al.* (2020) Mechanism of crosstalk between the LSD1 demethylase and HDAC1 deacetylase in the CoREST complex. *Cell Rep.* 30, 2699–2711
97. Lee, H.G. *et al.* (2019) The *Arabidopsis* Sin3–HDAC complex facilitates temporal histone deacetylation at the CCA1 and PRR9 loci for robust circadian oscillation. *Front. Plant Sci.* 10, 171
98. Mayer, K.S. *et al.* (2019) HDA9–PWR–HOS15 is a core histone deacetylase complex regulating transcription and development. *Plant Physiol.* 180, 342–355
99. Zeng, X. *et al.* (2020) HISTONE DEACETYLASE 9 functions with Polycomb silencing to repress FLOWERING LOCUS C expression. *Plant Physiol.* 182, 555–565
100. Zheng, Y. *et al.* (2020) Histone deacetylase hda9 and wrky53 transcription factor are mutual antagonists in regulation of plant stress response. *Mol. Plant* 13, 598–611
101. Baek, D. *et al.* (2020) Histone deacetylase HDA9 With ABI4 contributes to abscisic acid homeostasis in drought stress response. *Front. Plant Sci.* 11, 143
102. Yang, L. *et al.* (2020) HOS15 and HDA9 negatively regulate immunity through histone deacetylation of intracellular immune receptor NLR genes in *Arabidopsis*. *New Phytol.* 226, 507–522
103. van der Woude, L.C. *et al.* (2019) HISTONE DEACETYLASE 9 stimulates auxin-dependent thermomorphogenesis in *Arabidopsis thaliana* by mediating H2A.Z depletion. *Proc. Natl. Acad. Sci. U. S. A.* 116, 25343–25354
104. Zheng, Y. *et al.* (2016) Histone deacetylase HDA9 negatively regulates salt and drought stress responsiveness in *Arabidopsis*. *J. Exp. Bot.* 67, 1703–1713
105. Zhang, Y. *et al.* (2003) HDAC-6 interacts with and deacetylates tubulin and microtubules in vivo. *EMBO J.* 22, 1168–1179
106. Liu, X. *et al.* (2013) PHYTOCHROME INTERACTING FACTOR 3 associates with the histone deacetylase HDA15 in repression of chlorophyll biosynthesis and photosynthesis in etiolated *Arabidopsis* seedlings. *Plant Cell* 25, 1258–1273
107. Grozinger, C.M. and Schreiber, S.L. (2000) Regulation of histone deacetylase 4 and 5 and transcriptional activity by 14-3-3-dependent cellular localization. *Proc. Natl. Acad. Sci. U. S. A.* 97, 7835–7840
108. Nishino, T.G. *et al.* (2008) 14-3-3 regulates the nuclear import of class IIa histone deacetylases. *Biochem. Biophys. Res. Commun.* 377, 852–856
109. Niu *et al.* (2019) SIZ1-mediated SUMOylation of TPR1 suppresses plant immunity in *Arabidopsis*. *Mol. Plant* 12, 215–228
110. Long, J.A. *et al.* (2006) TOPLESS regulates apical embryonic fate in *Arabidopsis*. *Science* 312, 1520–1523
111. Hao, Y. *et al.* (2016) Histone deacetylase HDA6 enhances brassinosteroid signaling by inhibiting the BIN2 kinase. *Proc. Natl. Acad. Sci. U. S. A.* 113, 10418–10423
112. Cigliano, R.A. *et al.* (2013) Histone deacetylase ATHDA7 is required for female gametophyte and embryo development in *Arabidopsis*. *Plant Physiol.* 163, 431–440
113. Zheng, B. *et al.* (2014) An ARID domain-containing protein within nuclear bodies is required for sperm cell formation in *Arabidopsis thaliana*. *PLoS Genet.* 10, e1004421
114. Pischke, M.S. *et al.* (2006) A transcriptome-based characterization of habituation in plant tissue culture. *Plant Physiol.* 140, 1255–1278
115. Moreno-Yruela, C. *et al.* (2018) Histone deacetylase 11 is an epsilon-N-mristoyllysine hydrolase. *Cell Chem. Biol.* 25, 849–856
116. Watson, P.J. *et al.* (2012) Structure of HDAC3 bound to co-repressor and inositol tetrakisphosphate. *Nature* 481, 335–340
117. Fu, W. *et al.* (2007) Sequence and expression analysis of histone deacetylases in rice. *Biochem. Biophys. Res. Commun.* 356, 843–850
118. Becana, M. *et al.* (2020) Plant hemoglobins: a journey from unicellular green algae to vascular plants. *New Phytol.* 227, 1618–1635
119. Peng, K. *et al.* (2006) Length-dependent prediction of protein intrinsic disorder. *BMC Bioinforma.* 7, 208
120. Peng, Z.L. and Kurgan, L. (2012) Comprehensive comparative assessment of in-silico predictors of disordered regions. *Curr. Protein Pept. Sci.* 13, 6–18
121. Sanchez, R. and Sali, A. (1997) Evaluation of comparative protein structure modeling by MODELLER-3. *Proteins* 50–58
122. Zimmermann, L. *et al.* (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* 430, 2237–2243