

RESEARCH

Open Access



Genome-wide identification and analysis of GH1-containing H1 histones among poplar species

Ping Li¹, Jing Wang¹, Qimin Zhang¹, Anmin Yu¹, Rui Sun¹ and Aizhong Liu^{1*}

Abstract

Histone H1s are basic nuclear proteins, which played key role in the binding of DNA and nucleosome, eventually the stability of eukaryotic chromatin. In most species, H1s possess an evolutionarily conserved nucleosome-DNA binding globular domain (GH1), which is conserved between species, especially in mammals. However, there is limited information on the phylogeny, structure and function of H1s in poplar. In the present research, 21 GH1-containing proteins found in *Populus trichocarpa* were classified into three subgroups (H1s, Myb (SANK) GH1 and AT-hook GH1) based on their domains. The *Populus* H1 proteins contained lysine-rich N-, C-terminal tails and a conserved GH1 domain, particularly the characteristic amino acids in the helix and strand structures of the five H1 subtypes. The phylogenetic and structure diversity analysis of GH1 proteins across different *Populus* species and model plants revealed three conserved subgroups with characteristic amino acids. The variation in the number of members across the five subtypes was consistent with the evolutionary relationships among *Populus* species. The conserved characteristic amino acids among same *Populus* subtype can be served as markers for subtype identification. Furthermore, the abundance analysis of H1s in *Populus* indicated their unique functions in young tissues and stages, which may be related to DNA methylation. The consistent expression pattern of H1 across *Populus* species was in accordance with collinearity pairs. Present analyses provided valuable information on the diversity and evolution of H1s in *Populus*, advocating further research of H1s in plants.

Keywords Histone H1, GH1 domain, *Populus*, DNA methylation, Evolution

Introduction

Eukaryotic DNA packed with histones in nucleosomes, constitutes the basic chromatin organization that shapes the 3-dimensional structure of the genome and gene expression [1]. Unlike the core histones (H2A, H2B, H3 and H4), histone H1 links the nucleosomes to DNA and protects the linkage of linker DNA and nucleosomes [2, 3]. Linker histones (H1s) bind to nucleosomes via

electrostatic interactions, is essential for the stabilization of eukaryotic chromatin structure, and folding and compaction dynamics of nucleosome [4, 5]. Histone H1, conserved in higher eukaryotes, contains a tripartite structure including a conserved central globular domain (GD), a short N-terminal domain (NTD) and a long C-terminal domain (CTD) [6]. Unlike core histones, H1 is often highly variable in terms of the number of variants (or subtypes), and sequence divergence among different eukaryotes, or cell types [7]. For instance, 12 H1 variants with different expression patterns have been reported in different tissue/ cell types in human and mice [1, 8]. The distribution density of H1s and the binding or unbinding status of H1s and DNA, in particular, broadly affect

*Correspondence:

Aizhong Liu
liuazhong@mail.kib.ac.cn

¹ Key Laboratory for Forest Resource Conservation and Utilization in the Southwest Mountains of China (Ministry of Education), College of Forestry, Southwest Forestry University, Kunming, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

the opening/ closing dynamics of chromatin accessible regions and regulate gene expression [9, 10].

In plants, the subtypes and functions of different H1s have been studied. Three *Arabidopsis* H1s harbored a conserved GD, similar to those in animals, but displayed low sequence similarity among the subtypes (H1-1, H1-2 and H1-3) [11]. Five linker histones identified in castor bean presented conserved structure (GD) but different sequences [12]. Plant H1s can also regulate gene expression and plant growth through epigenetic mechanisms. In *Arabidopsis*, H1 regulates transcriptional silencing of genes and transposable elements (TEs) by targeting methylated DNA sequences, altering nucleosome organization, and modifying the methylation levels of TEs [13, 14]. H1s can assist H2A.W in promoting chromatin compaction, maintaining the density of heterochromatin, and preventing DNA methylation in *Arabidopsis* [15], while its deficiency affects seed dormancy, flowering time, lateral root formation, stomata development, and callus formation, mediated by influence heterochromatin structure, histone acetylation, and methylation [16]. Overexpression of *Arabidopsis* H1 in tobacco induced chromatin structural changes, affecting genes associated with developmental processes [17].

H1 histones also play important roles in plants under stress conditions. The function of different H1 subtypes are various. The interaction between Pin1 and histone H1 can regulate the residence time of phosphorylation to affect the stabilization of chromatin, which plays an important role in pathogen response [18]. Loss of H1s can activate pericentromeric highly DNA-methylated TEs under heat stress to influence plant heat stress resistance [19]. H1.3 (a subclass of minor H1 variants), is required for DNA methylation and stomatal functioning under normal or water-deficient conditions [4]. In addition to growth regulation and stress response, H1s are involved in the plant reproductive development. The histone H1 depletion during *Arabidopsis* male gametogenesis leads to activation of TEs by relaxing heterochromatin through DEMETER-directed DNA demethylation [20]. Histone H1 can influence gene imprinting in *Arabidopsis* via DNA methylation changes in gene promoter regions (MEA, FWA, FIS2) in the endosperm [21]. The suppression of H1 expression in *Arabidopsis* can influence DNA methylation patterns and heritable development [22].

There are numerous evidences on the importance of histone H1, which varied among subtype the subtypes, however, limited information is available on features of linker proteins between different species. Genome-wide identification and characterization analysis was an effective method to begin exploring the subtypes and functions of linker proteins, which are structurally conserved, but differ in types, numbers, and functions [23]. Poplar,

a model tree species, possess significant economic and ecological values, including fiber, timber, biofuel, bioremediation, and animal feed [24]. In this study, we identified 21 GH1 domain proteins in *P. trichocarpa*, which were grouped into three subtypes, similar to *Arabidopsis thaliana* homologs. The distribution of GH1 members across different *Populus* species (39 in *Populus euphratica*, 44 in *Populus alba*, 39 in *Populus tomentosa*, 23 in *Populus deltoids*) was consistent with plant evolution. The H1 showed higher abundance of hydrophilic amino acids (Lys and His) in GD, especially between helices and strands and variability among H1 members, which ensured their function in DNA binding. All poplar species contained five H1 subtypes with characteristic recognition sites, however these may be associated with distinct functions. The expression analysis revealed role of H1s in plant primary growth, which may be associated with cell division proliferation and DNA methylation. Overall this study provides several insights into different aspects of poplar H1s, advocating further investigations into the functions of H1s in plants.

Results

Genome-wide identification and characterization of histone H1 proteins in *P. trichocarpa*

Genome-wide similarity search of *P. trichocarpa* using *Arabidopsis* H1 proteins identified 21 *P. trichocarpa* proteins containing the conserved GH1 domains (Table 1). In addition to the GH1 domain, AT-hook and Myb domains were also present and used for sub-categorization of some GH1 proteins. The *P. trichocarpa* 21 GH1 proteins were classified into three subgroups, i) H1 subgroup (9 members) containing conserved GH1 domain, flanked by unstructured NTD and CTD tail, typical structure of H1, except protein Potri.005G069700 with long CTD, ii). Myb (SANK) GH1 subgroup (6 members) containing an additional Myb (SANK) domain in the NTD and a Coiled Coil domain in the CTD, and iii). AT-hook GH1 subgroup (6 members) containing multiple (5–8) AT-hook motifs in the CTD along with the GH1 domains towards NTD (Fig. 1, Table 1).

To explore the DNA binding function of GH1 domain in linker histone, the structure features of identified GH1 proteins were analyzed. All the GH1 proteins contained GH1 domain ranged from 70 to 72 amino acids in length (Table 1). The sequence analysis of *P. trichocarpa* GH1 proteins and the representative protein 3D structure revealed the presence of three Helices (Helix I, II and III) and β -strand, which were typical characteristic of the GH1 domain in linker histones (Figs. 1, and 2). At the junctions between helices and strands, many hydrophilic amino acids were detected. The conserved His5 (Helix I) and His38 (Helix II) (as screen in Potri.002G043100)

Table 1 Protein characteristic of GH1 proteins in *P. trichocarpa*

Protein ID	Subtypes	Protein Length (aa)				Lysine content				Theoretical pI				Number of positively charged residues (Asp + Glu)				Subcellular location
		total	NTD	GH1	CTD	Whole	NTD	CTD	Whole	NTD	CTD	Whole	NTD	CTD	Whole	NTD	CTD	
Potri.002G043100	H1	286	54	71	161	25.90%	18.50%	32.20%	10.61	9.82	11.46	80	13	52	nucl-			
Potri.005G219800	H1	285	54	71	160	25.60%	16.70%	30.70%	10.75	9.9	11.69	80	12	52	nucl-			
Potri.013G042700	H1	203	12	72	119	27.60%	36.40%	32.00%	11.62	13.3	12.54	64	3	43	nucl-			
Potri.008G162300	H1	181	33	72	76	25.40%	18.80%	32.90%	10.58	10.6	10.6	48	5	25	nucl-			
Potri.010G076800	H1	197	34	72	91	26.40%	21.20%	33.00%	10.81	10.7	11.12	55	6	31	nucl-			
Potri.002G199900	H1	195	18	72	105	22.60%	22.20%	26.40%	10.5	11.26	10.85	53	4	32	nucl-			
Potri.005G116600	H1	202	49	72	81	22.30%	20.40%	28.70%	10.35	9.4	11.48	53	13	25	nucl-			
Potri.007G014200	H1	179	49	72	58	20.10%	14.30%	29.70%	10.22	8.07	11.4	44	11	19	nucl-			
Potri.005G069700	H1-nonspecific	1028	43	72	913	8.40%	9.50%	8.40%	4.73	5.33	4.61	32	4	20	nucl-			
Potri.018G066300	GH1-Myb (SANK)	279	114	70	95	9.70%	10.70%	11.30%	9.08	9.26	4.71	43	18	11	nucl-			
Potri.006G147400	GH1-Myb (SANK)	275	116	70	89	9.50%	9.60%	12.70%	8.33	7.93	4.82	40	16	11	nucl-			
Potri.009G087000	GH1-Myb (SANK)	296	116	70	110	9.10%	10.40%	8.30%	9.5	9.56	4.94	48	19	15	nucl-			
Potri.001G292700	GH1-Myb (SANK)	296	114	70	112	8.80%	10.70%	5.00%	9.59	9.19	5.61	49	18	17	nucl-			
Potri.014G004900	GH1-Myb (SANK)	292	115	70	107	12.00%	11.50%	13.20%	9.82	10.11	7.82	51	22	17	nucl-			
Potri.007G005000	GH1-Myb (SANK)	335	115	70	150	11.90%	12.40%	12.70%	9.95	10.16	9.62	59	23	26	nucl-			
Potri.015G040200	GH1-HMGA1(AT hook)	482	44	71	367	7.50%	0.00%	8.30%	9.69	5.14	10.43	16	0	6	nucl-			
Potri.012G048500	GH1-HMGA1(AT hook)	473	44	71	358	7.60%	3.90%	8.50%	9.4	4.13	9.99	16	0	6	nucl-			
Potri.004G087500	GH1-HMGA1(AT hook)	189	8	71	110	8.50%	12.50%	9.50%	10.9	3.79	12.26	31	0	24	nucl-			
Potri.017G129400	GH1-HMGA1(AT hook)	190	8	71	111	8.90%	12.50%	9.40%	10.56	3.79	12.31	33	0	25	nucl-			
Potri.008G132400	GH1-HMGA1(AT hook)	207	32	71	104	8.20%	0.00%	10.00%	11.23	3.67	12.78	31	0	24	nucl-			
Potri.010G109600	GH1-HMGA1(AT hook)	215	40	71	104	7.90%	0.00%	10.10%	11.09	3.67	12.75	30	0	23	nucl-			

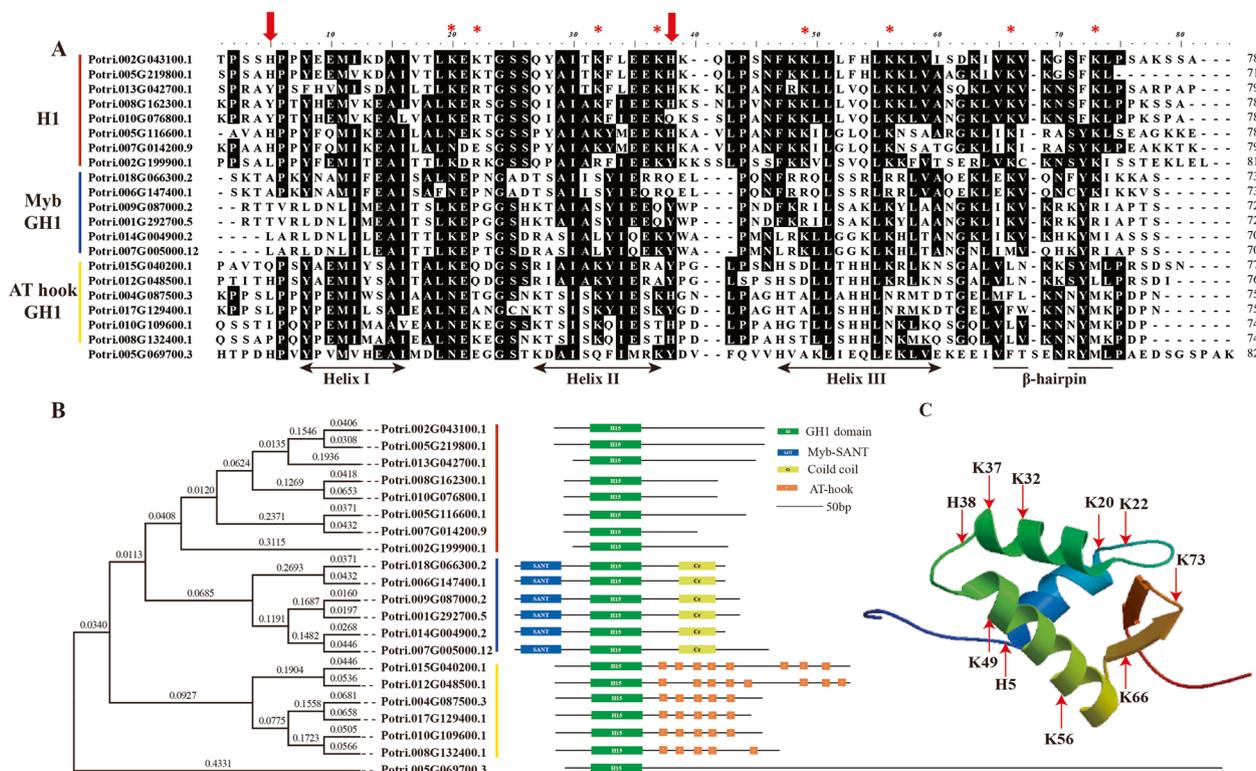


Fig. 1 Protein sequence alignment, phylogenetic tree, and domain structure display of GH1 domains in all GH1 proteins of *P. trichocarpa*. **A** Alignment of amino acid sequence of GH1 domains. The sites with asterisks or arrows above are the putative conserved basic binding sites of H1 proteins. **B** Neighbor-joining phylogenetic tree and domain architecture of GH1 domain proteins in *P. trichocarpa*. Phylogenetic tree was built with protein sequences of GH1 domains of *P. trichocarpa* GH1 proteins using MEGA 5.05 (bootstrap values of 1,000 replicate). **C** 3D model and conserved binding residues (in red) in GH1 domain of Potri.002G043100 using I-TASSER. Potri.002G043100 model was built using 4qlc.1.K template and 3.50 Å X-ray method with 0.39 sequence similarity, 0.26 coverage, 60–130 range

were located near the duplex DNA, consistent with the recognized helix and strand structure in the GH1 model (Fig. 1A, FigS1). Additionally, conserved basic residues such as Lys20 and Lys22 (β-strand), Lys32 (Helix II), and Lys37 (3' end of Helix II), Lys49 and Lys56 (Helix III), Lys66 and Lys73 (β-hairpin) were identified in the H1 of *P. trichocarpa* (Fig. 1A, FigS1). The 3D model of the GH1 domain showed that all conserved residues were located on the surface (FigS1C), facilitating DNA binding of H1. The GH1 domains of other two GH1 subgroup proteins contained fewer conserved residues, even with no conserved residues (Fig. 1A).

Except for the long GH1 protein Potri.005G069700, sequence analysis identified only 8 H1 proteins containing lysine-rich NTD and CTD (mostly over 30%), which may facilitate their DNA-binding function. The lower lysine content in Myb (SANK) GH1 and AT-hook GH1 may be insufficient for the DNA-binding function of the GH1 domain (Table 1).

A neighbor-joining (NJ) phylogenetic tree was constructed by aligning the protein sequences of 21 GH1 protein domains from *P. trichocarpa* (Fig. 1). Except

for the long GH1 protein Potri.005G069700, the other GH1 proteins were placed into three subgroups, i) eight H1 typical GH1 domain proteins, ii) six Myb (SANK) GH1 proteins, iii) six AT-hook GH1 proteins (Fig. 1B). Based on the evidence above, we considered the 8 GH1 domain contained to be the true linker histone H1s in *P. trichocarpa*. The distinct separation in phylogeny and structure of these GH1 domain proteins indicated the differentiation of H1, Myb (SANK) GH1 and AT-hook GH1, aiding in the identification of linker histones.

Variability of histone H1 proteins among different *Populus* species

To investigate the extent of divergence of histone H1 proteins in *Populus*, the GH1 proteins were also identified in *P. euphratica*, *P. alba*, *P. tomentosa* and *P. deltoids* (Table S1). The lengths of GH1 domains ranged from over 50 to less than 72 amino acids in *Populus* species (Table S1). Full-length sequences of 166 *Populus* GH1 proteins (39 in *P. euphratica*, 44 in *P. alba*, 39 in *P. tomentosa*, and 23 in *P. deltoids*) and three *Arabidopsis* H1s were utilized to construct a maximum likelihood

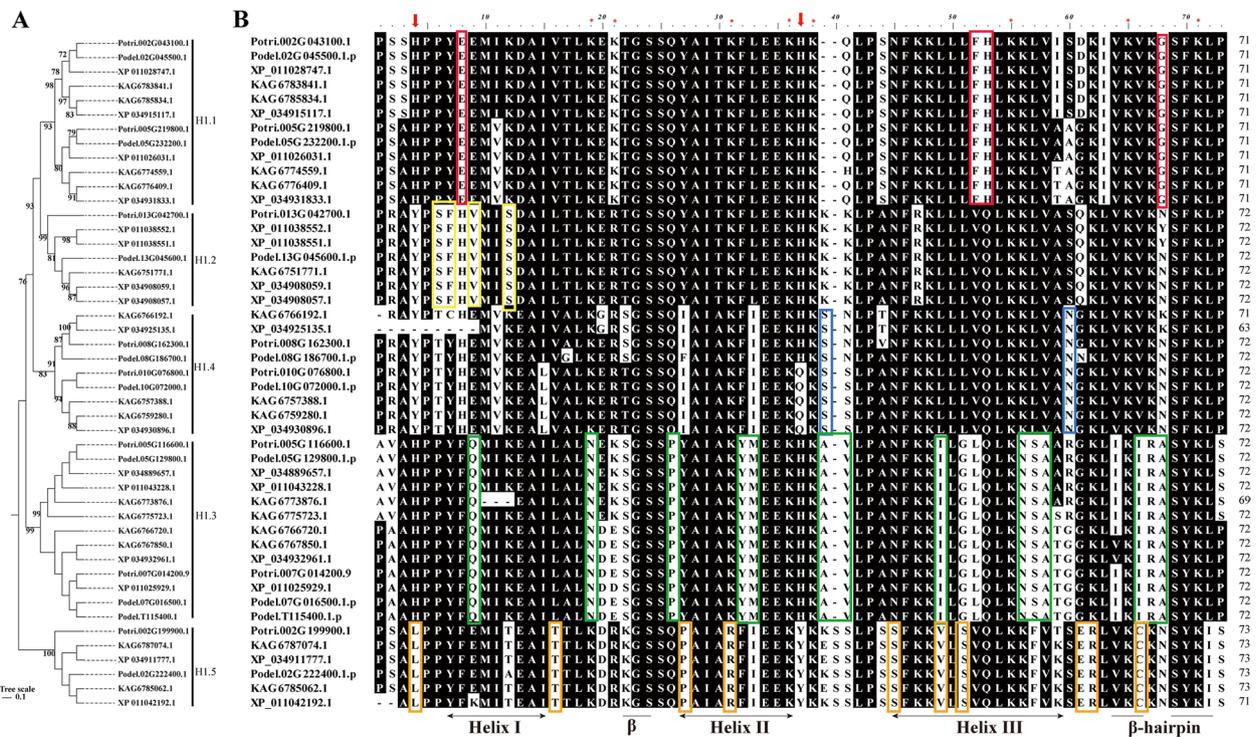


Fig. 2 Sequence and phylogeny analysis of H1 in *Populus*. **A** Maximum likelihood phylogenetic tree based on protein sequences of the GH1 domain of *Populus* H1 proteins. **B** Amino acid sequence alignment of GH1 domains of *Populus* H1. The sites in the red boxes are conserved binding site residues in H1.1, the sites in the yellow boxes are conserved binding site residues in H1.2, the sites in the blue boxes are conserved binding site residues in H1.4, the sites in the green boxes are conserved binding site residues in H1.3, the sites in the orange boxes are conserved binding site residues in H1.5

(ML) phylogenetic tree, and classified into three main subgroups: H1s, Myb (SANK) GH1, and AT hook GH1, except abnormal H1 (FigS2). Similar subgroups were also identified in model plants (FigS3, Table S2). The distribution of GH1 subgroups varied among *Populus* species. Key differences include that *P. alba* had a relatively large number of AT—hook GH1s (25), while *P. euphratica* had a notably high count of Myb (SANK) GH1s (22). *P. tomentosa* had the most H1s (14) among the species mentioned. This diverse distribution of GH1 proteins across *Populus* species may be related to their

evolutionary history (FigS4), potentially indicating functional differences among these species.

Sequence and phylogenetic analysis revealed that all *Populus* species possess abnormal H1 proteins, characterized by a long C-terminal tail and low lysine content. The number of abnormal H1 variants varied among *Populus* species, with most species retaining one variant, except for *P. alba* (3 variants) and *P. tomentosa* (2 variants) (Table 2). The abnormal H1 proteins formed distinct phylogenetic branches compared to typical H1 proteins, highlighting their divergence (FigS2).

Table 2 Different GH1 subgroups of different *Populus* species

Species	Genome size (Mb)	Myb (SANK) GH1	AT-hook GH1	H1					abnormal H1
				H1.1	H1.2	H1.3	H1.4	H1.5	
<i>Populus trichocarpa</i>	434.29	6	6	2	1	2	2	1	1
<i>Populus alba</i>	416.99	25	7	2	2	2	2	1	3
<i>Populus tomentosa</i>	739.8	11	12	4	1	4	3	2	2
<i>Populus deltoids</i>	428.64	7	6	2	1	3	2	1	1
<i>Populus euphratica</i>	496.03	22	9	2	2	2	0	1	1

Sequence analysis of *Populus* H1 identified the conserved GH1 domain, which included three characteristic helices and a β -strand structure (Fig. 2). The GH1 domain sequences from five *Populus* species classified them into five distinct subtypes (Fig. 2). The *Populus* H1 proteins were categorized into five variants, designated as H1.1, H1.2, H1.3, H1.4, and H1.5, with the exception of *P. euphratica*, which lacked H1.4 subtypes (Fig. 2, Table 2). H1.1, conserved in *A. thaliana*, was present in all *Populus* species. *P. alba* and *P. euphratica* had two H1.2, more than other species. H1.3, involved in plant abiotic stress responses, was unevenly distributed. H1.5 was rare in most *Populus* species, except *P. tomentosa* with the largest genome.

Through sequence analysis, we identified conserved hydrophilic residues in the *Populus* GH1 domains (Fig. 2). The 10 putative DNA binding residues in the histone linker H1 identified in *P. trichocarpa* (Fig. 1A) were also conserved across H1.1 subtypes in different *Populus* species (Fig. 2). The number of these conserved H1.1 proteins varied among different *Populus* species: *P. trichocarpa*, *P. alba*, *P. euphratica*, and *P. deltoids* each had two H1.1 protein, while *P. tomentosa* had four. In addition to the conserved hydrophilic sites, characteristic sites of H1.1, such as Glu8, Phe52, His53, and Gly68, were also identified (Fig. 2). The second group H1 proteins in *Populus* had conserved Lys-/Arg- residues in helix II, helix III and the β -hairpin structure, but His5 and Lys22 were replaced by Tyr5 and Arg22, respectively (Fig. 1, Fig. 2). The number of H1.2 proteins was low: *P. trichocarpa*, *P. deltoids* and *P. tomentosa* each had one, while *P. alba* and *P. euphratica* each had two. These H1.2 proteins were identified by their characteristic sequences Ser5, Phe6, Val8, and Ser11 (Fig. 2). The fourth group of H1 proteins were conserved, with His5, Lys22, and His32 replaced by Tyr4, Arg21, and Gln37, respectively. All H1s in this group, except for those with missing sequences, had a mutation at the His32 site (Fig. 2). *P. trichocarpa*, *P. deltoids*, and *P. alba* each had two H1.4 proteins, while *P. tomentosa* had three (Table 2). They also had characteristic conserved residues at Ser39 and Asn60 (Fig. 2). The third group of H1 proteins had a conserved substitute of Lys38 for Asn39 (Fig. 2). All H1.3 proteins had lost sequences at the N-terminals and had a conserved substituted of Lys22 for Glu21. The number of H1.3 proteins varied: *P. euphratica*, *P. trichocarpa* and *P. alba* conserved each had two; *P. deltoids* had three, *P. tomentosa* had four. The characteristic sequences of H1.3 included Gln9, Asn19, Pro26, Tyr32, Met33, Ala39, Val40, Ile49, Asn56, Ser57, Ala58, Ile66, Arg67 and Ala68, which were highlighted in green boxes (Table 2). The fifth group of H1 proteins had substitutions of Leu4, Arg21, Arg31 and Tyr37 for His5, Lys22, Lys32 and His38, respectively

(Fig. 2). *P. euphratica*, *P. trichocarpa*, *P. deltoids* and *P. alba* each had one H1.5, while *P. tomentosa* had two (Table 2). The characteristic sequences of H1.5 included Leu4, Thr16, Pro27, Arg31, Ser44, Val49, Ser51, Glu61, Arg62, Cys66 (Fig. 2). The same H1 subtype in different *Populus* species may indicate conservation within species. The variation in the number of different H1 subtypes among different *Populus* species may suggest evolutionary divergence (FigS4). The characteristic sequences in the coding regions of different H1 subtypes could serve as DNA markers for *Populus* (FigS5).

Collinearity analysis of *P. trichocarpa* H1 proteins

To detect the origin relationships of H1 from different *Populus* species, we employed collinearity analysis involving the *P. trichocarpa* genome and across different poplars. The intraspecific collinearity analysis of *P. trichocarpa* H1 predominantly distributed among members of the same H1 subtypes. However, some collinearity pairs were also observed between H1.1, H1.2, H1.4 members (FigS6). The k_a/k_s values between *P. trichocarpa* H1 and their homologs from other *Populus* species were all less than 1, which indicated that these H1 homologous were undergoing purifying selection (negative selection) (Table S3). Based on the intergenetic collinearity analysis, H1 from five *Populus* species represented enriched collinearity relationships with *P. trichocarpa* H1 (Fig. 3). The collinearity analysis revealed a distinct pattern among H1 subtypes. Collinearity pairs were predominantly and uniformly enriched within the same H1 subtypes, indicating a high degree of genomic conservation at the subtype-level. This suggests that genes within the same H1 subtype may have evolved in a more coordinated manner. Moreover, we also observed the presence of collinearity pairs distributed across different H1 members, specifically among H1.1, H1.2, and H1.4. As detailed in Table S3, these inter-subtype collinearity relationships imply complex evolutionary connections between these particular H1 variants.

Tissue-specific expressional analyses of *P. trichocarpa* GH1 proteins

We retrieved publicly available transcriptomic data to explore the expression profiles of histone *H1* genes at different development stages and in various tissues in *P. trichocarpa*. As shown in Fig. 4 and Table S4, the three subtypes of GH1 proteins exhibited distinct expression patterns, with H1s demonstrating higher expression levels. H1.1, H1.2 and one H1.4 variant (*Potri.010G076800*) exhibited high expression levels across all the examined tissues. H1.3 displayed high expression levels in young plant tissues, particularly in dormant buds and young leaves. One H1.4 variant (*Potri.008G162300*) and H1.5

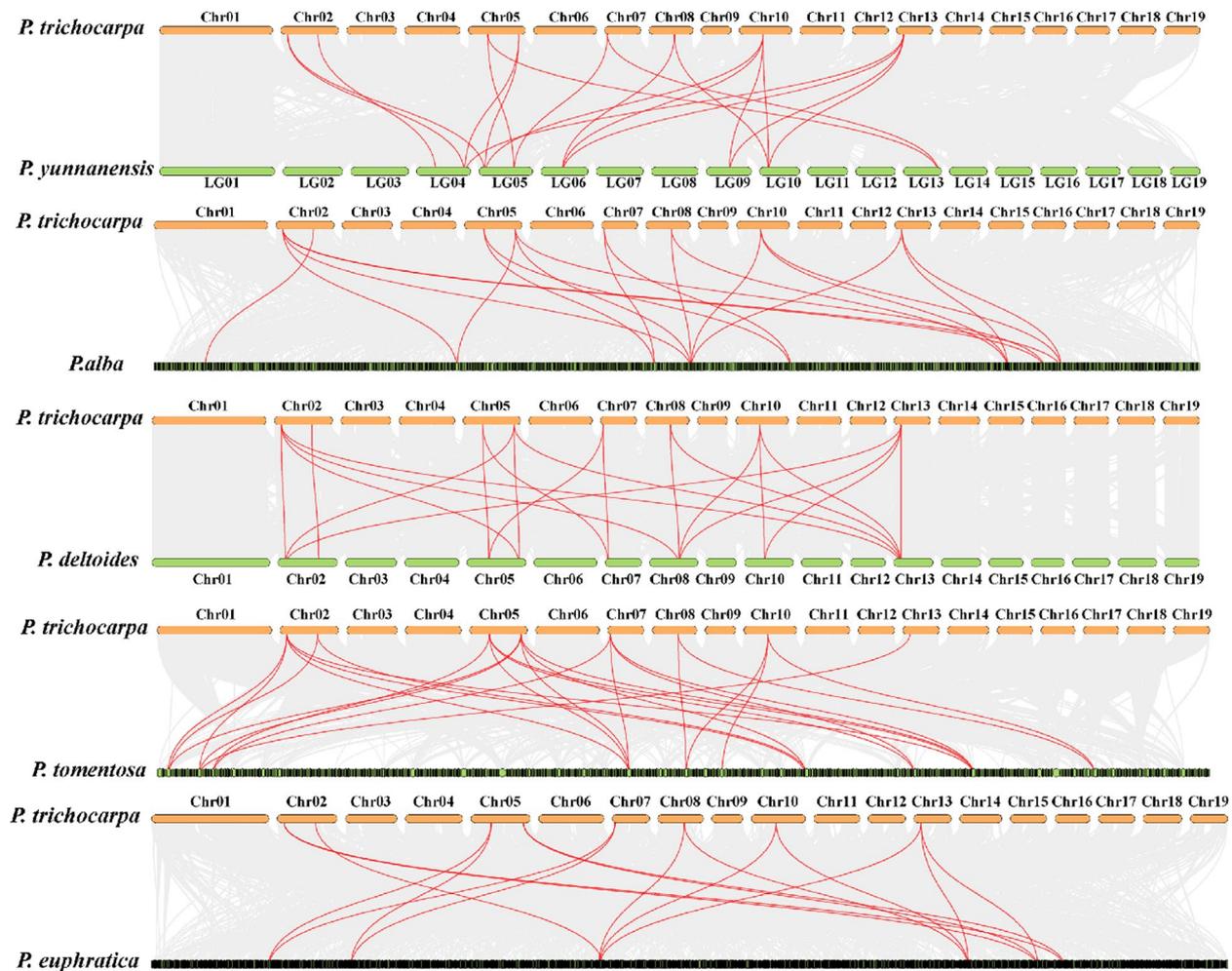


Fig. 3 Collinearity analysis of poplar H1. Orange bars represent chromosomes of *P. trichocarpa*, green and black represent chromosomes or scaffolds of other poplars. Red lines connect collinear H1 genes, and gray background lines represent all collinear gene pairs between genomes. Species names are marked on the left side of the chromosomes

exhibited consistently low expression levels throughout the developmental stages, except in young leaves and root tips. The atypical H1 variant (*Potri.005G069700*) also exhibited low expression level across all tissues. The differential expression of histone genes was consistent with the subgroup members, especially H1.1, H1.2 and H1.3. The high enrichment in young tissues suggests that H1 plays a crucial role during periods of rapid growth. The main proteins related to DNA methylation also presented similar expression pattern during different tissues (FigS7).

Expressional analyses of H1 proteins in different *Populus* species

From the expression data of drought stress and recovery treatment across different tissues of *P. alba*, we observed distinct expression patterns of H1 proteins

(Fig. 5A). In general, histone H1 expression patterns showed distinct characteristics across different conditions and species. In drought-recovery and normal tissues, certain H1 variants maintained stable low expression, such as one H1.2 (*XP_034908057.1*) and H1.5. Conversely, H1.1, H1.3, and one H1.4 (*XP_034930896.1*) exhibited high expression levels. Additionally, one H1.4 (*XP_034925135.1*) and one H1.2 (*XP_034908059.1*) accumulated under specific conditions, suggesting specialized functions.

When considering species-specific H1 expressions, *P. deltoides* presented unique patterns. H1.1, H1.2, and two H1.3 variants were highly induced under biotic stress and in developmental tissues, particularly in active buds, indicating their crucial roles in early growth processes. Meanwhile, H1.5, one H1.3, one H1.4, and abnormal H1 showed low expression across various

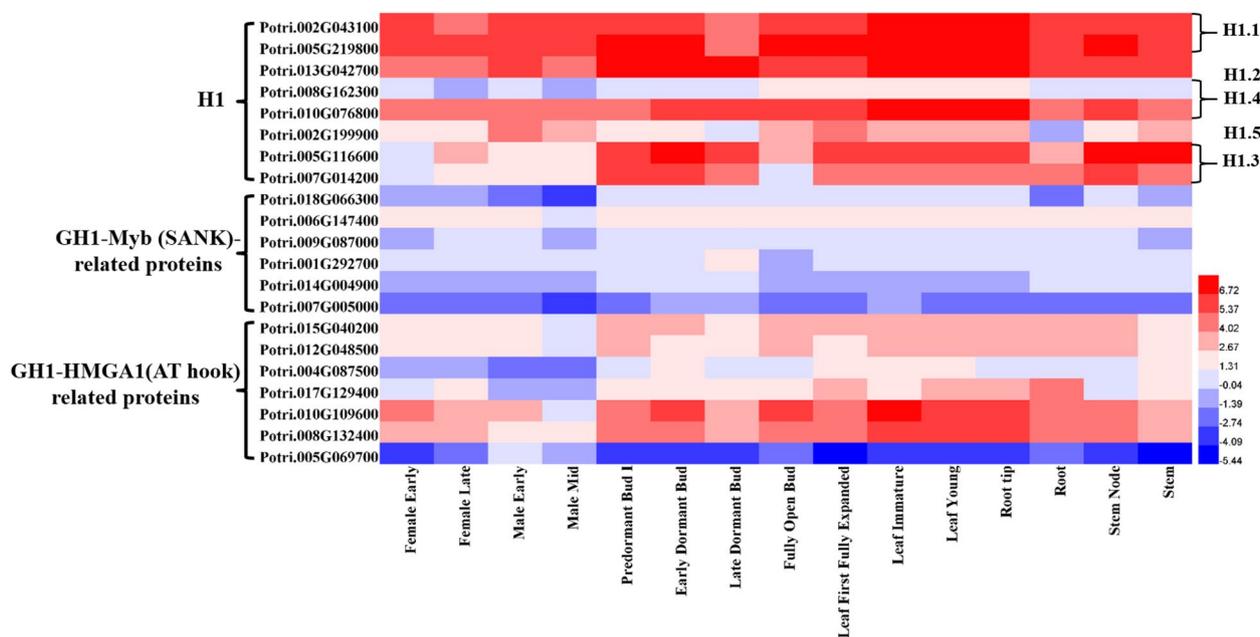


Fig. 4 Heatmap of expression profiles of GH1 proteins in *P. trichocarpa*. Expression data is based on \log_2 (FPKM) values from Phytozome 13. Female early: spring female plant; female late: winter female plant; male early: spring male plant; male mild: winter male plant; predormant bud 1: bud in winter; early dormant bud: bud in late winter; late dormant bud: bud in spring; fully open bud: bud in late spring; leaf first fully expanded: first opened leaves; leaf immature: normal growth leaves; leaf young: bottom leaves; root tip and root; stem node and stem

treatments. One H1.4 (*Podel.10G072000.1.p*) has low expression during the stress stage but accumulated in the early growth phase (Fig. 5B, Table S4).

In *P. euphratica*, one H1.1 (*XP_011028747.1*) and one H1.3 (*XP_011043228.1*) were induced by all treatments, highlighting their significance in both stress responses and development. *XP_011025929.1* (H1.4) was highly induced in most cases, except in seeds. On the contrary, *XP_011038552.1* (H1.2), *XP_011042192.1* (H1.5), and abnormal H1 (*XP_011031098.1*) were inhibited. *XP_011026031.1* (H1.1) and *XP_011038551.1* (H1.2) had low expression levels, except in growth-stage tissues. *XP_011025929.1* (H1.3) accumulated during both stress and growth, and these expression changes implied differential gene functions (Fig. 5C, Table S4).

Expressional verification of poplar H1 proteins

To investigate whether H1 proteins respond to different stress and growth stage, we collected tissues from different growth stages and stress treatment samples of *P. yunnanensis* (which is evolutionary closely related to *P. trichocarpa*) [25], and analyzed the expression changes of eight H1 genes using qRT-PCR. As shown in Fig. 6A, six out of the eight H1 genes exhibited significant expression changes. In contrast to *EF1*(control), *Poyun10755* and *Poyun30287* showed no detectable expression during any of the experimental growth stages as represented of their homologs (*Potri.002G199900*, *Potri.007G014200*) in *P.*

trichocarpa (Table S4, Table S5). Four H1 genes, including *Poyun12377*(H1.1), *Poyun13213* (H1.1), *Poyun23177* (H1.4) and *Poyun24440* (H1.2) showed high expression levels during the pre-germinated bud stage. *Poyun14213* (H1.3) displayed differential expression across growth stages, with higher expression levels observed in young tissues. The expression of *Poyun16041* (H1.4) was significantly higher during young stages, except in old leaves. All H1 genes were significantly induced by Abscisic Acid (ABA), except for *Poyun10755* (H1.5, Fig. 6B). Under salt and cold stress conditions, the expression of most H1 genes was significantly inhibited. With the exception of *Poyun14213* (H1.3), the expression of most H1 genes was inhibited under drought stress. These results suggested that most H1 genes may play distinct roles in young growth stages, tissues, and stress responses at the transcriptional level.

Discussion

Histone protein, H1 play key role in packing of DNA into nucleosomes and maintenance of the chromatin structure [4, 5]. Consistent with earlier analyses of linker histones in *Arabidopsis* and castor bean [12, 26], the subgroups of *Populus* GH1 proteins can be classified as H1s, Myb (SANK) GH1, and AT-hook GH1 (Fig. 1, Table S1). The subtype composition may be influenced by post-translational modifications (PTMs), which modulates chromatin structure, and affects transcriptional status

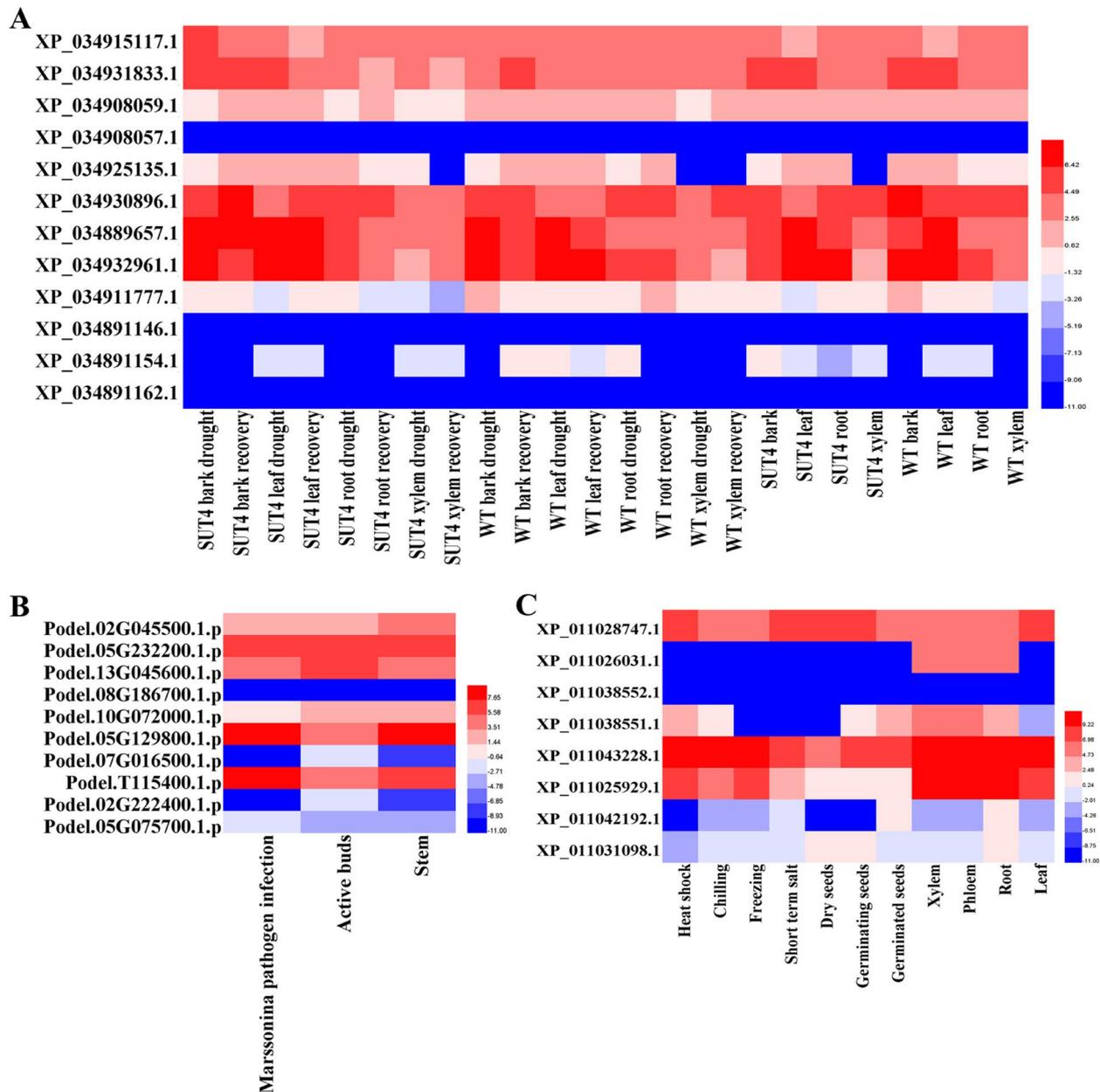


Fig. 5 Heatmap of expression profiles of H1s in *P. alba* (A), *P. deltoids* (B) and *P. euphratica* (C). Expression data was obtained from RNA-seq analysis using raw data from the NCBI database with accession numbers reported in reference articles. Log₂(FPKM) values were used for the heatmap

of genes during normal and disease conditions in mammals and plants [8, 27, 28]. The phylogenetic tree of *P. trichocarpa* GH1 proteins and other plants showed subgroups specific to H1s, Myb (SANK) GH1, and AT-hook GH1 related proteins (Fig. 1, Table S1, FigS1). Furthermore, the classification of GH1 protein subtypes also suggested conserved nature of H1 proteins across different plant genera (FigS2, FigS3, Table S2).

The analyses of protein sequences and conserved motifs revealed that all GH1 proteins and H1s harbored GH1 domain, which was relatively conserved during plant evolution (FigS3). The H1 proteins showed same domain architecture as in other eukaryotes [29]. In addition to the conserved GH1 domains, *Populus* H1s showed diversity in N-terminal and low complexity domains C-terminals. The N-terminal and C-terminal regions of *Populus* H1s play crucial roles in the protein's interaction

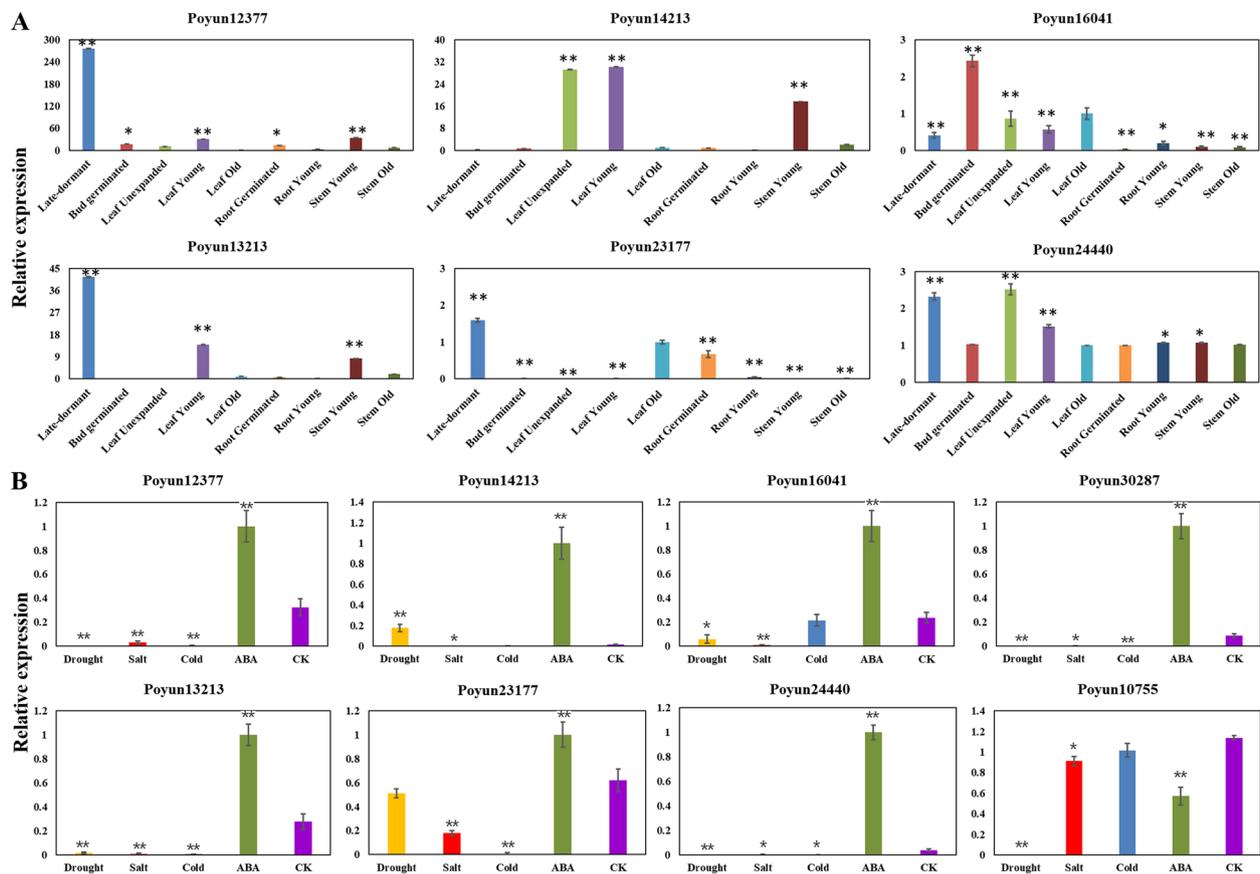


Fig. 6 Relative expression pattern of H1 in *P. yunnanensis*. **A** Relative expression pattern of H1 on different growth stages in *P. yunnanensis*. **B** Relative expression pattern of H1 under different stress and treatment in *P. yunnanensis*. EF1 (*Poyun37990*) was used as an internal control. Data are means of three biological replicates, and error bars represent \pm SE from three independent experiments, each performed with 2–3 leaves from three separate plants. Asterisks denote significant differences determined by the LSD test. Specifically, * indicates $p < 0.05$ (significant difference), and ** indicates $p < 0.01$ (highly significant difference)

with DNA. The observed diversity in these regions is likely to be associated with the kinetic properties of DNA binding, which may lead to variations in how quickly and tightly the H1 proteins bind to DNA [29].

The long GH1 protein Potri.005G069700, with low lysine content, a long C-terminal tail, and aberrant DNA binding subunits (Table 1, Fig. 2), may belong to a special category of H1, which may be evolutionarily important across all *Populus* species. These aberrant H1 proteins were generally maintained as a single copy among the *Populus* species, except *P. alba* (3 aberrant H1s) and *P. tomentosa* (2 aberrant H1s) (FigS2, Table S1). The significance of aberrant H1 proteins required further investigation.

Compared to the GH1 domain proteins found in *Arabidopsis*, a significantly higher number of 21 GH1 proteins were identified in *P. trichocarpa* [22, 26], which may be related to certain evolutionary events (Table 1). Presence of similar number of GH1 proteins in *P. trichocarpa* (21

and *P. deltoids* (23) reflected their close evolutionary relationship (Table 2, FigS4). These findings supported the prevailing view that H1 domain proteins have evolved throughout plant evolution [25, 26]. Moreover, the number of H1 proteins also showed copy number variations during the course of evolution, as *P. tomentosa* harbored the maximum number of H1 proteins, which is consistent with its evolutionary position and triploid background (Table 2, FigS4) [25, 30]. *P. trichocarpa*, *P. deltoids* and *Palba* shared a similar evolutionary relationship and possessed a similar number of H1 proteins, particularly H1.1, H1.4, and H1.5 (Table 2, FigS4) [25]. *P. euphratica* had lost H1.4, possibly due to its distinct evolutionary trajectory within the *Populus* genus. As shown in Table 2 and Figure S4, *P. euphratica* has diverged more significantly from other *Populus* species compared to the rest of the group. Another contributing factor to this genomic variation might be the relatively low conservation level of H1 proteins. Previous research has indicated that H1

proteins exhibit a lower degree of sequence conservation across different organisms [31]. H1 variants are important for organizing of higher-order chromatin structures, which suggesting the functional differentiation according to the differential number of H1 proteins in poplar [32]. The differential distribution of GH1 proteins among model plants was consistent with their diverse evolutionary histories (Table S2) [26].

Based on the sequence analysis, the H1 proteins of *Populus* can be classified into five subtypes H1.1, H1.2, H1.3, H1.4 and H1.5. H1.1 and H1.3, which shared same branches with *A. thaliana*, were named based on the classification and nomenclature in *A. thaliana* (FigS2) [26]. Although, the number of H1 subtypes varied among different *Populus* species, the structure and hydrophilic amino acids were conserved. At the junctions between helices and strands, hydrophilic amino acids play a crucial role in DNA binding [33]. All H1.1 subtypes had conserved hydrophilic amino acids (Lys-residues) and recognition sites (His-) (Fig. 2). The H1.2, H1.3, H1.4 and H1.5 subtypes also conserved contained hydrophilic amino acids. Compared to H1.1 subtype, residues within the same subtype members were substituted with the same hydrophilic amino acids (Fig. 2). The conserved characteristic residues have been identified among same subtype members of *Populus* species, such as four conserved amino acids in H1.1 and H1.2 subtypes, two hydrophilic amino acids in H1.4 subtypes, more than 10 conserved amino acid sites in H1.3 and H1.5 subtypes (Fig. 2). Higher content positively charged residues of H1s were required during the DNA binding of H1 [32]. The characteristic amino acids involved in same H1 subtypes can be used as DNA markers in *Populus* (Fig. 2, FigS5). Gene duplication events are always associated with evolution in plant, collinearity analysis of H1 revealed the origins and duplication events, particularly among members of same H1 subtype members in *P. trichocarpa*. The enriched collinearity relationship members were also highly expression during different tissues and stress treatment (Fig. 3, FigS6, Table S4).

Assessing relapse after chemotherapy through the quantitative analysis of histone H1 variants suggesting that protein abundance may infer their function [34]. In this study, H1 proteins of different *Populus* species exhibited extremely variable expression across various tissues and development stages (Fig. 4–6, Table S4). Most H1 proteins (H1.1, H1.2 and H1.3) maintained high expression levels across different tissues, particularly in young plant tissues, which were similarity to those observed in cucumbers [35]. The atypical H1 maintained low expression levels across all tissues and during stress treatments. Under biotic stress, H1.3 played a lesser role compared to its role during plant growth and abiotic stress, as

observed in *Arabidopsis* [4]. The low expression level of atypical H1 and H1.4 across different *Populus* species and treatments may be related to the number of evolutionary subtypes [25]. These results suggested the various functions of different H1 proteins during different tissue and stress treatment. The activated H1 members reported on other plants were also active across *Populus* genomes, revealing their evolutionary similarities (Fig. 6) [36].

The differential expression level of different H1 members may be involved in modulating differential binding affinity or related to differential cellular regulation [29, 37]. The PTM sites of H1s (located in GH domain and tails) for phosphorylation, acetylation, methylation, ubiquitination, and ADP-ribosylation, regulate the function of linker histones [38]. The activity of major DNA methylation enzymes, such as cytosine-specific methyl transferase (CMT2, CMT3) and DNA cytosine-5-methyltransferase 3-related (DRM2) were induced in the young tissues and development stages (FigS7) [39]. The nucleus location of histones and their expression in the centromere of cucumber imply their regulation function [35]. The abundance of DNA cytosine-5 methyl transferase (DNMT), and lysine-specific histone demethylase (LSD1) were significantly increased during young tissues (predormant bud, young leaf, root tip etc.). These results suggested the role of H1s in rapid growth stages, which may be related to DNA methylation (FigS7) [40]. Beyond the immediate connection with DNA methylation in rapid growth, histone variant modifications have broader implications. Histone variant sequence modifications and their expression levels are not only essential for maintaining chromosomal integrity but also have a profound impact on damaged chromatin dynamics [41]. In the context of *Populus* evolution, the variations in the number and expression of H1s tell a complex story. They are not merely numerical or expression-level changes; instead, they serve as a molecular record of the species' evolutionary journey. This work not only enhances our knowledge of H1s in plants, but also paves the way for future research on epigenetic regulation, with potential applications in plant breeding and genetic engineering.

Materials and methods

Identification of GH1 proteins

Genome data of *P. trichocarpa* (black cottonwood), *P. euphratica* (Euphrates poplar), *P. alba* (white poplar), *P. tomentosa* (Chinese white poplar) and *P. deltoides* (eudicots) were obtained from the Genome data of NCBI ([https:// www.ncbi.nlm.nih.gov/genome/?term=Populus](https://www.ncbi.nlm.nih.gov/genome/?term=Populus)). The genome data of *P. yunnanensis* were obtained from the National Genomics Data Center (NGDC, <https://ngdc.cnc.ac.cn>) with accession number PRJCA010101 [25]. Histone H1 proteins from *Arabidopsis* were obtained from

the *Arabidopsis* Information Resource (TAIR, <https://www.arabidopsis.org>) using the known genes [22, 26, 42].

Sequence similarity searches for *P. trichocarpa* were performed using the BLASTP function of the standalone BLAST+ tool (NCBI-blast-2.7.1+) with *Arabidopsis* H1 proteins (At1g06760, At2g30620, and At2g18050) [43]. *P. trichocarpa* H1 proteins containing conserved domain sequences were used as queries to identify potential *Populus* proteins in *P. euphratica*, *P. alba*, *P. tomentosa* and *P. deltoids* with a maximum E-value of 1e-5 (Table S1). The lysine content of histone H1 and terminal domains was calculated using ProtParam (<http://us.expasy.org/tools/protparam.html>). Subcellular location predictions were obtained using WoLF PSORT [44].

Phylogenetic analysis of GH1 proteins and *Populus* species

The phylogenetic tree of GH1 domain proteins in *P. trichocarpa* was constructed using their protein sequences of GH1 domains in MEGA 5.05 (<http://www.megasoftware.net/history.php>) with neighbor-joining (NJ) method. The phylogenetic tree of H1 proteins in various *Populus* species was constructed in MEGA 5.05 using the ML method with the protein sequences of their GH1 domains [45, 46]. The phylogenetic tree of H1 proteins from represent plants was constructed in MEGA 5.05 using the ML method with their whole length protein sequences. The bootstrap values reported for each branch represented the percentage of 1,000 replicate trees that included that branch. The rooted species tree of *Populus* was obtained by the OrthoFinder method with species orthologs, which revealed the evolutionary relationship of *Populus* species [47]. The collinearity analysis of H1 coding genes was obtained by TBtools [48].

Protein structure and conserved motif analysis of GH1

The functional domains of *Populus* H1 sequences were analyzed using SMART (<http://smart.embl-heidelberg.de/>). The lysine content of *P. trichocarpa* H1 sequences was calculated using ExPaSy (<https://www.expasy.org/>) [49]. The protein model was constructed using I-TASSER [50]. The most similarity template was used to build 3D models, such as 4qlc.1.K template of Potri.002G043100.

Expression patterns of H1 genes in different *Populus* species

P. trichocarpa transcriptome data were obtained from Phytozome 13 (<https://phytozome-next.jgi.doe.gov/>). The data was utilized to analyze the expression profiles of *P. trichocarpa* H1 genes across various developmental stages (spring, winter female and male plants) and tissues (bud (predormant, early dormant, late dormant and fully open), leaf (first fully, immature and young), root (tip and whole), stem (node and whole)).

The transcriptome FastQ data of various *Populus* species were download with SRA-Explorer using their accession number reported in the articles (<https://sra-explorer.info>) [51–61]. To compare the expression of H1 genes across different *Populus* species under various stress, we compared the transcriptome data to their background genomes. Transcript reconstruction was performed using hisat2 software [62] following alignment with Samtools. The merging of assembled read partitions was evaluated using StringTie [63]. The expression levels of each gene were quantified and normalized using Fragments Per Kilobase of transcript per Million mapped reads (FPKM). The expression heatmaps were constructed based on the Log₂ values of the FPKM.

qRT-PCR assays

The *P. yunnanensis* materials were collected from Kunming (E102° 74N25° 17), and were well grown in the culture room of Southwest Forestry University, Kunming, under natural conditions. Fresh and healthy dormant buds (late dormant, germinated bud), leaf (unexpanded, young, old), stems (young, old), roots (germinated, young) were collected. For stress treatment: 200 mM NaCl was added for the salt treatment; plants were subjected to 24 h 4°C for the cold treatment; 1 weeks no-water treatment was executed for the drought treatment; 100 mM ABA was added for the ABA treatment, following the previously reported treatment methods [64]. Total RNAs were extracted from different tissues of *P. yunnanensis* using the RNAPrep Pure Plant Plus Kit (Cat. DP441, Tiangen, Beijing, China), following the manufacturer's instructions. 1 µg RNA was treated with DNaseI and reverse-transcribed with oligo (dT) and the PrimeScriptTMRT reagent Kit (Takara, Japan). The relative expression levels of individual genes were measured using gene-specific primers by real-time quantitative PCR (qRT-PCR) analysis, carried out in a 20 µL reaction mix containing 1 µL of diluted cDNA template and SYBR Premix Ex TaqII (Takara, Japan) on a Bio-Rad CFX96. The elongation factor 1-alpha (EF1) gene (*Poyun37990*, homolog of *Potri.009G018600*) served as the internal control [65]. The internal control and data analysis were conducted according to previously reported methods [64]. The primer sequences used for qRT-PCR are listed in Table S6.

Conclusions

The study aimed to understand the structure, evolution, and function of GH1 proteins and H1s in *Populus*. Phylogenetic analysis classified GH1 members into three variants, and five H1 subtypes were identified with characteristic

amino acids for subgroup distinction. The characteristic amino acids of *Populus* H1 subtypes can serve as markers for subgroup distinction, enhancing our understanding of H1s' evolutionary relationships and structural conservation. Expression analysis showed functional differentiation of H1s in primary growth and DNA methylation regulation, indicating specialized roles in *Populus* development. DNA methylation was found crucial for H1s' function in tissue development and stress responses. Future work should focus on leveraging the identified characteristic markers for more accurate H1 subtype identification in *Populus*. Additionally, experiments on these subtypes will further clarify their functions in primary growth and DNA methylation. These efforts will not only deepen our understanding of H1s in *Populus* but also lay a foundation for broader research on H1s across the plant kingdom.

Abbreviations

GH1	Nucleosome-DNA binding globular domain
H1	Linker histone
GD	Central globular domain
NTD	N-terminal domain
CTD	C-terminal domain
TE	Transposable elements
ML	Maximum likelihood
ABA	Abscisic Acid
PTM	Post-translational modification

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11456-6>.

Supplementary Material 1. Fig S1 Secondary structure of GH1 domain of *P. trichocarpa* H1 proteins.

Supplementary Material 2. Fig S2 Phylogenetic analysis of *Populus* GH1 proteins.

Supplementary Material 3. Fig S3 Phylogenetic tree of GH1 proteins during model plant species

Supplementary Material 4. Fig S4 Phylogenetic tree of *Populus* species with OrthoFinder.

Supplementary Material 5. Fig S5 Sequence analysis of *Populus* H1 coding sequences.

Supplementary Material 6. Fig S6 Collinearity analysis of H1 coding genes of *P. trichocarpa*

Supplementary Material 7. Fig S7 Expression pattern of DNA methylation transferase

Supplementary Material 8. Table S1 Sequence similarity searching during different *Populus* species.

Supplementary Material 9. Table S2 Sequence similarity searching during different plant

Supplementary Material 10. Table S3 Collinearity pairs among *Populus* H1s and their ka/ks values.

Supplementary Material 11. Table S4 Expression data of *Populus* H1 for heatmap.

Supplementary Material 12. Table S5 The list of *P. trichocarpa* H1 homologs in *P. yunnanensis*.

Supplementary Material 13. Table S6 The primer sequences used for qRT-PCR

Acknowledgements

No conflict of interest is declared.

Authors' contributions

P.L. wrote the main manuscript text, J.W. and Q.Z. contributed to Fig. 5, A.Y. prepared the methods of data analyze, R.S. checked the plant material, A.L. did the writing framework. All authors reviewed and agreed to the published version of the manuscript.

Funding

This study was supported by Yunnan Fundamental Research Projects (202301AT070216, 202201AU070072), the National Natural Science Foundation of China (32460375), and the Project of Yunnan Provincial Department of Education Science Research (2024Y577).

Data availability

Genome sequences of *Populus* were obtained from the Genome data of NCBI (<https://www.ncbi.nlm.nih.gov/genome/?term=Populus>). The genome data of *P. yunnanensis* obtained from the National Genomics Data Center (NGDC, <https://ngdc.cncb.ac.cn>) with accession number PRJCA010101. Histone H1 proteins from *Arabidopsis* were obtained from the *Arabidopsis* Information Resource (TAIR, <https://www.arabidopsis.org>). *P. trichocarpa* transcriptome data were obtained from Phytozome 13 (<https://phytozome-next.jgi.doe.gov/>). The raw data of *Populus* transcriptome were download with SRA-Explorer (<https://sra-explorer.info>) using their accession numbers (SRR064169, SRR064170, SRR12020515, SRR12020518, SRR12020521).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 12 January 2025 Accepted: 6 March 2025

Published online: 24 March 2025

References

1. Brockers K, Schneider R. Histone H1, the forgotten histone. *Epigenomics*. 2019;11(4):363–6.
2. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ: Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 1997;389(6648):251–60.
3. Allan J, Cowling GJ, Harborne N, Cattini P, Craigie R, Gould H. Regulation of the higher-order structure of chromatin by histones H1 and H5. *J Cell Biol*. 1981;90(2):279–88.
4. Rutowicz K, Puzio M, Halibart-Puzio J, Lirski M, Kotliński M, Kroteń MA, Knizewski L, Lange B, Muszewska A, Śniegowska-Świerk K, et al. A Specialized Histone H1 Variant Is Required for Adaptive Responses to Complex Abiotic Stress and Related DNA Methylation in *Arabidopsis*. *Plant Physiol*. 2015;169(3):2080–101.
5. Fyodorov DV, Zhou BR, Skoultschi AI, Bai Y. Emerging roles of linker histones in regulating chromatin structure and function. *Nat Rev Mol Cell Biol*. 2018;19(3):192–206.
6. Roque A, Ponte I, Suau P. Post-translational modifications of the intrinsically disordered terminal domains of histone H1: effects on secondary structure and chromatin dynamics. *Chromosoma*. 2017;126(1):83–91.
7. Mayor R, Izquierdo-Bouldstridge A, Millán-Ariño L, Bustillos A, Sampaio C, Luque N, Jordan A. Genome distribution of replication-independent histone H1 variants shows H1.0 associated with nucleolar domains and H1X associated with RNA polymerase II-enriched regions. *J Biol Chem*. 2015;290(12):7474–91.

8. Andrés M, García-Gomis D, Ponte I, Suau P, Roque A. Histone H1 post-translational modifications: Update and future perspectives. *Int J Mol Sci.* 2020;21(16):5941.
9. Parseghian MH, Newcomb RL, Winokur ST, Hamkalo BA. The distribution of somatic H1 subtypes is non-random on active vs. inactive chromatin: distribution in human fetal fibroblasts. *Chromosome Res.* 2000;8(5):405–24.
10. Parseghian MH, Newcomb RL, Hamkalo BA. Distribution of somatic H1 subtypes is non-random on active vs. inactive chromatin II: distribution in human adult fibroblasts. *J Cell Biochem.* 2001;83(4):643–59.
11. Gantt JS, Lenvik TR. *Arabidopsis thaliana* H1 histones. Analysis of two members of a small gene family. *Eur J Biochem.* 1991;202(3):1029–39.
12. Guo J, Li P, Yu A, Chapman MA, Liu A. Genome-wide characterization and evolutionary analysis of linker histones in castor bean (*Ricinus communis*). *Front Plant Sci.* 2022;13:1014418.
13. Choi J, Lyons DB, Kim MY, Moore JD, Zilberman D. DNA Methylation and Histone H1 jointly repress transposable elements and aberrant intragenic transcripts. *Mol Cell.* 2020;77(2):310–323.e317.
14. Shintomi K, Iwabuchi M, Saeki H, Ura K, Kishimoto T, Ohsumi K. Nucleosome assembly protein-1 is a linker histone chaperone in *Xenopus* eggs. *Proc Natl Acad Sci U S A.* 2005;102(23):8210–5.
15. Bourguet P, Picard CL, Yelagandula R, Pélissier T, Lorković ZJ, Feng S, Pouch-Pélissier MN, Schmücker A, Jacobsen SE, Berger F, et al. The histone variant H2A.W and linker histone H1 co-regulate heterochromatin accessibility and DNA methylation. *Nat Commun.* 2021;12(1):2683.
16. Rutowicz K, Lirski M, Mermaz B, Teano G, Schubert J, Mestiri I, Kroteń MA, Fabrice TN, Fritz S, Grob S, et al. Linker histones are fine-scale chromatin architects modulating developmental decisions in *Arabidopsis*. *Genome Biol.* 2019;20(1):157.
17. Prymakowska-Bosak M, Przewłoka MR, Iwkiewicz J, Egierszordorf S, Kuraś M, Chaubet N, Gigot C, Spiker S, Jerzmanowski A. Histone H1 overexpressed to high level in tobacco affects certain developmental programs but has limited effect on basal cellular functions. *Proc Natl Acad Sci USA.* 1996;93(19):10250–5.
18. Jinasena D, Simmons R, Gyamfi H, Fitzkee NC. Molecular Mechanism of the Pin1-Histone H1 Interaction. *Biochemistry.* 2019;58(6):788–98.
19. Liu S, de Jonge J, Trejo-Arellano MS, Santos-González J, Köhler C, Hennig L. Role of H1 and DNA methylation in selective regulation of transposable elements during heat stress. *New Phytol.* 2021;229(4):2238–50.
20. He S, Vickers M, Zhang J, Feng X. Natural depletion of histone H1 in sex cells causes DNA demethylation, heterochromatin decondensation and transposon activation. *eLife* 2019;8(8):e42530.
21. Rea M, Zheng W, Chen M, Braud C, Bhangu D, Rognan TN, Xiao W. Histone H1 affects gene imprinting and DNA methylation in *Arabidopsis*. *Plant J.* 2012;71(5):776–86.
22. Wierzbicki AT, Jerzmanowski A. Suppression of histone *H1* genes in *Arabidopsis* results in heritable developmental defects and stochastic changes in DNA methylation. *Genetics.* 2005;169(2):997–1008.
23. Jiang D, Berger F. Histone variants in plant transcriptional regulation. *Biochim Biophys Acta.* 2017;1860(1):123–30.
24. Stettler RF, Bradshaw HD, Science JF. Biology of *Populus* and its implications for management and conservation. 1997, 43(3):457–457.
25. Shi T, Zhang X, Hou Y, Jia C, Dan X, Zhang Y, Jiang Y, Lai Q, Feng J, Feng J, et al. The super-pangenome of *Populus* unveils genomic facets for its adaptation and diversification in widespread forest trees. *Mol Plant.* 2024;17(5):725–46.
26. Kotliński M, Knizewski L, Muszewska A, Rutowicz K, Lirski M, Schmidt A, Baroux C, Ginalski K, Jerzmanowski A. Phylogeny-Based systematization of *arabidopsis* proteins with histone H1 globular domain. *Plant Physiol.* 2017;174(1):27–34.
27. Charbonnel C, Rymarenko O, Da Ines O, Benyahya F, White CI, Butter F, Amiard S. The Linker Histone GH1-HMGA1 is involved in telomere stability and DNA damage repair. *Plant Physiol.* 2018;177(1):311–27.
28. Kotliński M, Rutowicz K, Knizewski Ł, Palusiński A, Olędzki J, Fogtman A, Rubel T, Koblowska M, Dadlez M, Ginalski K, et al. Histone H1 variants in *Arabidopsis* are subject to numerous post-translational modifications, both conserved and previously unknown in histones, suggesting complex functions of H1 in plants. *PLoS One.* 2016;11(1):e0147908.
29. Flanagan TW, Brown DT. Molecular dynamics of histone H1. *Biochem Biophys Acta.* 2016;1859(3):468–75.
30. Ren Y, Zhang J, Wang G, Liu X, Li L, Wang J, Yang M. The relationship between insect resistance and tree age of transgenic triploid *Populus tomentosa* Plants. *Front Plant Sci.* 2018;9:53.
31. Bayona-Feliu A, Casas-Lamesa A, Carbonell A, Climent-Cantó P, Tatarski M, Pérez-Montero S, Azorín F, Bernués J. Histone H1: Lessons from Drosophila. *Biochem Biophys Acta.* 2016;1859(3):526–32.
32. Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet RM. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature.* 1993;362(6417):219–23.
33. Wolffe AP. Histone H1. *Int J Biochem Cell Biol.* 1997;29(12):1463–6.
34. Noberini R, Morales Torres C, Savoia EO, Brandini S, Jodice MG, Bertalot G, Bonizzi G, Capra M, Diaferia G, Scaffidi P et al. Label-Free Mass Spectrometry-Based Quantification of Linker Histone H1 Variants in Clinical Samples. *Int J Mol Sci.* 2020;21(19):7330.
35. Wang Y, Li Y, Zhou F, Zhang L, Gong J, Cheng C, Chen J, Lou Q. Genome-wide characterization, phylogenetic and expression analysis of Histone gene family in cucumber (*Cucumis sativus* L.). *Int J Biol Macromolecules.* 2023;230:123401.
36. Cannon SB, Mitra A, Baumgarten A, Young ND, May G. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* 2004;4:10.
37. Raghuram N, Carrero G, Stasevich TJ, McNally JG, Thng J, Hendzel MJ. Core Histone Hyperacetylation Impacts Cooperative Behavior and High-Affinity Binding of Histone H1 to Chromatin. *Biochemistry.* 2010;49(21):4420–31.
38. Izzo A, Schneider R. The role of linker histone H1 modifications in the regulation of gene expression and chromatin dynamics. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms.* 2016;1859(3):486–95.
39. Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ, Jacobsen SE. Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nat Struct Mol Biol.* 2014;21(1):64–72.
40. Yaari R, Katz A, Domb K, Harris KD, Zemach A, Ohad N. RdDM-independent de novo and heterochromatin DNA methylation by plant CMT and DNMT3 orthologs. *Nat Commun.* 2019;10(1):1613.
41. Ferrand J, Rondinelli B, Polo SE. Histone variants: guardians of genome integrity. *Cells.* 2020;9(11):2424.
42. He S, Yu Y, Wang L, Zhang J, Bai Z, Li G, Li P, Feng X. Linker histone H1 drives heterochromatin condensation via phase separation in *Arabidopsis*. *The Plant cell.* 2024; 36(5):1829–43.
43. Hu G, Kurgan L. Sequence similarity searching. *Curr Protoc Protein Sci.* 2019;95(1):e71.
44. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 2007;35(Web Server issue):W585–587.
45. Ao D, Li S, Jiang S, Luo J, Chen N, Meurens F, Zhu J. Inter-relation analysis of signaling adaptors of porcine innate immune pathways. *Mol Immunol.* 2020;121:20–7.
46. Ayaz A, Saqib S, Huang H, Zaman W, Lü S, Zhao H. Genome-wide comparative analysis of long-chain acyl-CoA synthetases (LACSs) gene family: A focus on identification, evolution and expression profiling related to lipid synthesis. *Plant physiology and biochemistry : PPB.* 2021;161:1–11.
47. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 2019;20(1):238.
48. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: An integrative toolkit developed for interactive analyses of big biological Data. *Mol Plant.* 2020;13(8):1194–202.
49. Tang T, Yu A, Li P, Yang H, Liu G, Liu L. Sequence analysis of the Hsp70 family in moss and evaluation of their functions in abiotic stress responses. *Sci Rep.* 2016;6:33650.
50. Zheng W, Zhang C, Li Y, Pearce R, Bell EW, Zhang Y. Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell reports methods.* 2021;1(3):100014.
51. Zhang J, Feng J, Lu J, Yang Y, Zhang X, Wan D, Liu J. Transcriptome differences between two sister desert poplar species under salt stress. *BMC Genomics.* 2014;15(1):337.
52. Xue LJ, Frost CJ, Tsai CJ, Harding SA. Drought response transcriptomes are altered in poplar with reduced tonoplast sucrose transporter expression. *Sci Rep.* 2016;6:33655.
53. Berta M, Giovannelli A, Sebastiani F, Camussi A, Racchi ML. Transcriptome changes in the cambial region of poplar (*Populus alba* L.) in response to water deficit. *Plant Biol.* 2010;12(2):341–54.

54. Wang J, Tian Y, Li J, Yang K, Xing S, Han X, Xu D, Wang Y. Transcriptome sequencing of active buds from *Populus deltoides* CL. and *Populus × zhaiguanheibaiyang* reveals phytohormones involved in branching. *Genomics*. 2019;111(4):700–9.
55. Ning K, Ding C, Huang Q, Zhang W, Yang C, Liang D, Fan R, Su X. Transcriptome profiling revealed diverse gene expression patterns in poplar (*Populus × euramericana*) under different planting densities. *PLoS ONE*. 2019;14(5):e0217066.
56. Chao Q, Gao ZF, Zhang D, Zhao BG, Dong FQ, Fu CX, Liu LJ, Wang BC. The developmental dynamics of the *Populus* stem transcriptome. *Plant Biotechnol J*. 2019;17(1):206–19.
57. Chen J, Yin W, Xia X. Transcriptome Profiles of *Populus euphratica* upon Heat Shock stress. *Curr Genomics*. 2014;15(5):326–40.
58. Chen J, Tian Q, Pang T, Jiang L, Wu R, Xia X, Yin W. Deep-sequencing transcriptome analysis of low temperature perception in a desert tree, *Populus euphratica*. *BMC Genomics*. 2014;15(1):326.
59. Yu L, Ma J, Niu Z, Bai X, Lei W, Shao X, Chen N, Zhou F, Wan D. Tissue-specific transcriptome analysis reveals multiple responses to salt stress in *populus euphratica* seedlings. *Genes*. 2017;8(12):372.
60. Zhang C, Luo W, Li Y, Zhang X, Bai X, Niu Z, Zhang X, Li Z, Wan D. Transcriptomic Analysis of Seed Germination Under Salt Stress in Two Desert Sister Species (*Populus euphratica* and *P. pruinosa*). *Front Genet*. 2019;10:231.
61. Tang S, Liang H, Yan D, Zhao Y, Han X, Carlson JE, Xia X, Yin W. *Populus euphratica*: the transcriptomic response to drought stress. *Plant Mol Biol*. 2013;83(6):539–57.
62. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37(8):907–15.
63. Kovaka S, Zimin AV, Pertea GM, Razaghi R, Salzberg SL, Pertea M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol*. 2019;20(1):278.
64. Li P, Wang J, Jiang D, Yu A, Sun R, Liu A. Function and characteristic analysis of candidate PEAR Proteins in *Populus yunnanensis*. *Int J Mol Sci*. 2023;24(17):13101.
65. Regier N, Frey B. Experimental comparison of relative RT-qPCR quantification approaches for gene expression studies in poplar. *BMC Mol Biol*. 2010;11:57.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.