BMC Plant Biology

RESEARCH

Open Access

Comprehensive characterization of poplar *HSP20* gene family: genome-wide identification, stress-induced expression profiling, and protein interaction verifications

Lincui Shi[1†], Yude Kang[2†], Ling Ding[1], Liejia Xu[1], Xiaojiao Liu[1], Anmin Yu[1], Aizhong Liu[1*] and Ping Li[1*]

Abstract

Background Heat shock proteins (HSP20s) are crucial components in plant stress responses, acting as molecular chaperones to safeguard cellular integrity and prevent abnormal protein aggregation. While extensive research has been conducted on HSP20s in various plant species, limited information is available regarding the HSP20 protein family in poplar (*Populus yunnanensis*), a species of significant ecological and economic importance native to southwestern China.

Results To elucidate the distribution, structural features, and functional characteristics of HSP20 proteins in *P. yunnanensis*, a combination of bioinformatics tools and experimental validation was utilized. A total of 53 PyHSP20s were identified within the *P. yunnanensis* genome and classified into 12 subfamilies: CI, CII, CIII, CIV, CV, CVI, CVII, MI, MII, ER, CP, and Px containing 24, 1, 1, 1, 2, 2, 14, 3, 1, 1, 2, and 1 HSP20 proteins, respectively. Classification was based on subcellular localization and phylogenetic relationships, revealing subfamilies with varying exon–intron structures and conserved motifs. The 3D structures analysis showed significant differentiation, with the CI subfamily PyHSP20s exhibiting 8 β-sheets, compared to 7 β-sheets in other subfamilies. Additionally, the N-terminal arms displayed heterogeneity in length and sequence. The 53 PyHSP20s were unevenly distributed across 15 chromosomes, with tandem segmental duplications explaining the expansion of subfamilies, particularly CI, CV, CVI, and CVII. The analysis of cis-elements associated with stress response and hormone regulation underscored the critical role of PyHSP20 in stress adaptation. Expression profiling via database analysis and qRT-PCR confirmed the responsiveness of PyHSP20s to multiple stressors, including salt, mannitol, drought, heat, and abscisic acid (ABA). Furthermore, Yeast Two-Hybrid (Y2H) assays demonstrated potential regulatory interactions between PyHSP20s and other functional proteins involved in stress responses.

Conclusions These findings provide a comprehensive understanding of the classification, structural differentiation, and functional roles of PyHSP20s in *P. yunnanensis*, thereby establishing a foundation for future functional investigations into this protein family.

Keywords HSP20, Poplar, Stress response, Protein interaction

[†]Lincui Shi and Yude Kang contributed equally to this work.

*Correspondence:
Aizhong Liu
liuaizhong@mail.kib.ac.cn
Ping Li
liping2020@swfu.edu.cn
Full list of author information is available at the end of the article

© 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/.

Background

Heat shock proteins (HSPs) are ubiquitous, ATP-independent chaperones that play critical roles, particularly under fluctuating conditions [1]. These proteins are essential for cellular protection and homeostasis, facilitating protein folding, preventing aggregation, and promoting degradation under stress conditions [2]. HSPs are classified into different families based on their molecular weight, including HSP40, HSP60, HSP70, HSP90, and small HSPs [3].

Small heat shock proteins (sHSPs) are a group of proteins with monomeric masses ranging from 16–40 kDa and a conserved C-terminal α-crystallin domain (ACD) consisting of approximately 90 amino acids [4]. sHSPs are key components of the cellular protein system and participate in various biological processes, including cell growth, regulation, and stress adaptation [5]. Most sHSPs have molecular weights in the range of 15–22 kDa and are also known as HSP20 [6]. The primary amino acid sequences of HSP20 proteins include a variable N-terminal domain, a highly conserved C-terminal region (known as the ACD domain), and a C-terminal region [7]. In angiosperms, eleven subfamilies of HSP20 proteins can be classified based on their amino acid sequence identity and subcellular targeting signals [6]. CP (chloroplast), MTII (mitochondria), ER (endoplasmic reticulum), PX (peroxisomes), CIII (nucleus), and MTI/CP (mitochondrial and chloroplast double locational proteins) are six classes of HSP20 proteins with specific targeting signals. In addition to these six classes, five subfamilies that lack targeted organelle signals but are responsible for targeting the cytosol are also classified (CI, CII, CIV, CV, and CVI) [8]. HSP20 proteins can prevent the aggregation of misfolded proteins in vitro due to their ATP-independent chaperone activity [9]. The polydisperse state of HSP20 proteins reveals various oligomeric states during cell growth and regulation [10]. The dynamic structure and chaperone binding ability of HSP20 proteins enable them to bind and sequester target proteins promptly after the production of abnormal proteins induced by stress and defense responses [11]. HSP20 proteins maintain membrane integrity through their amphitropic and non-transmembrane signals by associating with membranes and capturing unfolding proteins to prevent their irreversible aggregation, which is particularly important under stress conditions [12].

During the heat response of *Agaves tequilana*, the enrichment of HSP20 proteins enhances the thermoregulatory capacity of leaves and protects the shoot apical meristem (SAM) and folded leaves from high temperatures [13]. The increased expression of HSP20 during heat pretreatment plays a probable role in enhancing chilling tolerance of banana fruit [14]. Heterologously expressed HSP20s, isolated from thermotolerant bacteria such as *Escherichia coli*, exhibit enhanced tolerance to heat, cold, acid, alkaline, and hyperosmotic stresses [15]. In another study, HSP20 proteins have been shown to enhance tolerance to heat, salt, osmotic and drought stress, not only in prokaryotic and eukaryotic microorganisms, but also in higher plants [16].

The number of HSP20s varies among different plants. The subfamilies also display variations across various plant species. Potato contains 48 HSP20s, categorized into 12 distinct subfamilies [17]. Forty-eight VvHSP20s in grape are organized into 11 subfamilies based on phylogenetic analysis and subcellular localization [18]. Thirty-two HSP20 proteins from *Coix lacryma-jobi* L. are organized into 11 subfamilies [19], while forty-four PpHSP20s from peach are categorized into 10 distinct subfamilies [20]. Thirty CsHSP20s from cucumber are grouped into 11 subfamilies [21], and thirty-eight HSP20 proteins from *Sorbus pohuashanensis* are categorized into 10 distinct subfamilies [22]. Forty-one HSP20 proteins from *Cynodon transvaalensis* (African bermudagrass) are organized into 12 subfamilies [23].

Populus yunnanensis, a poplar species native to southwestern China, is highly valuable for forestry production and environmental conservation due to its rapid growth and environmental adaptability [24, 25]. In this study, bioinformatics methods were employed to identify *HSP20* genes in *P. yunnanensis*. Through sequence alignment and functional domain verification, a total of 53 *HSP20* genes were identified in *P. yunnanensis*. By constructing phylogenetic trees of HSP20 proteins from *P. yunnanensis* and *Arabidopsis thaliana*, combined with subcellular localization prediction and sequence-structure analysis, the 53 HSP20 proteins from *P. yunnanensis* were classified into 12 subfamilies: CP, MTII, ER, PX, CIII, MTI, CP, CI, CII, CIV, CV, and CVI, which are also conserved among the other five *Populus* species (*Populus euphratica*, *Populus deltoides*, *Populus trichocarpa*, *Populus tomentosa*, *Populus alba*). Sequence and protein structure analysis revealed that the HSP20 proteins in *P. yunnanensis* are conserved within subfamilies. The 53 HSP20 proteins were dispersed across the 15 chromosomes of *P. yunnanensis*. Collinearity analysis revealed 8 pairs of initial HSP20 proteins in *P. yunnanensis*. *Cis*-element analysis revealed that the promoter regions of most *HSP20* genes harbor binding sites for response elements related to drought, salt, and ABA (Abscisic acid), suggesting their potential involvement in plant stress response mechanisms. This conclusion was further supported by gene expression analysis. Given that HSP20 proteins exhibit binding and chaperone properties, we confirmed their interactions with ABA response proteins, other heat shock proteins, and kinases via yeast two-hybrid (Y2H)

assays. This study lays the foundation for further investigations into the function of HSP20 and the utilization of stress resistance gene resources for genetic improvement in *P. yunnanensis*.

Results

Identification and sequence analysis of *P. yunnanensis* HSP20s

A total of 57 candidate HSP20 proteins were initially identified in *P. yunnanensis* via local BLASTP searches against 19 HSP20 proteins from *A. thaliana* (Table S1). Subsequently, using these initial 57 candidates, an additional 77 candidate PyHSP20 proteins were identified via local BLASTP searches in the *P. yunnanensis* genome. Following domain detection using CDD (pfam00011), SMART, and molecular weight screening (Table 1) [4, 6], 53 PyHSP20 proteins were confirmed. The physicochemical characteristics of these proteins varied significantly. The protein lengths ranged from 137 amino acids (Poyun21268) to 331 amino acids (Poyun17694), with molecular weights ranging from 15.7 kDa (Poyun02516) to 36.7 kDa (Poyun17694). The theoretical isoelectric points (pI) ranged from 4.62 (Poyun34740) to 9.23 (Poyun24554), with 12 proteins classified as basic. The GRAVY scores, which indicate hydropathy, ranged from −0.94 to −0.059, suggesting strong hydrophilicity across all PyHSP20 proteins. Aliphatic indices ranged from 60.61 to 98.39, reflecting varying degrees of protein stability. The characteristic ACD domain (pfam00011) in PyHSP20 proteins spanned from 73 to 107 amino acids. Subcellular localization predictions indicated that 38 PyHSP20s were likely located in the cytoplasm, 7 in the nucleus, 2 in the chloroplast, 4 in the mitochondrion, 1 in the endoplasmic reticulum, and 1 in the peroxisome, suggesting that the cytoplasm is the primary localization site for HSP20s in *P. yunnanensis* (Table S2).

Protein and gene structure analysis

To investigate the structural features of PyHSP20 proteins, we conducted an analysis of motifs, protein domains, and gene structures (Fig. 1). Using MEME software, we identified 10 distinct conserved motifs within the PyHSP20 proteins. Motifs 1, 3, and 4 were the most frequently occurring, present in 34, 49, and 37 proteins, respectively. Additionally, motif 2 was found in 30 proteins, motif 5 in 18, motif 6 in 29, motif 7 in 21, motif 8 in 13, motif 9 in 20, and motif 10 in 17 proteins.

Conserved domain analysis revealed that nearly all PyHSP20 proteins possessed the ACD domain, with exceptions being Poyun01954, and Poyun18465, which belonged to the IbpA superfamily, and Poyun36443, which belonged to the HSP20 superfamily [26]. Additionally, some PyHSP20 proteins exhibited extra conserved domains, such as those in the Iso_dh superfamily (Poyun38293), PTZ00121 superfamily (Poyun17694), and PRK10263 and PRK12323 superfamily (Poyun08023), in addition to the key heat-shock protein domains.

Gene structure analysis showed that 24 *PyHSP20* genes consisted of a single exon, 26 *PyHSP20* genes had two exons, 2 *PyHSP20* genes had three exons, and the remaining *PyHSP20* genes had six exons.

Phylogenetic analysis and classification of PyHSP20s

To better understand the phylogenetic relationships and classification of PyHSP20s, an unrooted phylogenetic tree was constructed using the protein sequences of 53 PyHSP20s and 19 AtHSP20s (Fig. 2). Based on the phylogenetic relationships and subcellular locations of HSP20 proteins of *P. yunnanensis* and *A. thaliana* (Table 1), the 53 PyHSP20s were classified into 12 subfamilies: 24 CI (22 localized in the cytosol and 2 in the nucleus), 1 ER (endoplasmic reticulum), 2 CVI (cytosol), 1 Px (peroxisome), 1 CIV (cytosol), 2 CV (nucleus), 3 MI (mitochondria), 2 CP (plastid), 1 MII (mitochondria), 14 CVII (11 in the cytosol and 3 in the nucleus), 1 CII (cytosol), and 1 CIII (cytosol). Additionally, we analyzed the phylogenetic relationships of HSP20 proteins among different poplar species. All five poplar species analyzed contained 12 subfamilies of HSP20s (Figure S1).

The structures of PyHSP20s were conserved among evolutionarily related proteins and even extended beyond the boundaries of subfamilies. Three protein models (Poyun18465, Poyun06896, Poyun08063) were constructed using PyHSP20 proteins, all of which contained characteristic β folds. Two of these models (Poyun18465 and Poyun06896) additionally included α-helices (Fig. 3). The protein structure model of Poyun18465 was constructed based on the A0A251JRI8.1.A template (small heat shock protein domain-containing protein) with 78.57% sequence identity. A model quality estimate (MQE) of 0.78 GMQE was generated using ProMod3 version 3.4.0. Five PyHSP20s from the MII, Px, ER, CII, and CIV subfamilies exhibited structural similarities to Poyun18465, possessing two conserved α helixes (α1-α2), seven β folds (β1-β7), and highly conserved sequence motifs between β2-β3 (Pro-Gly), β5-β6 (Leu-Pro) and β6-β7 (Gly) (Fig. 3A). The protein structure model of Poyun06896 was constructed based on the A0A836Y6J7.1.A template with 85.33% sequence identity and a MQE of 0.69 GMQE was obtained using ProMod3 3.4.0. Eight PyHSP20s from the MI, CP, CVI and CVII subfamilies displayed structural similarity to Poyun06896, including one conserved α-helix and seven β-fold (β1-β7), with conserved sequence motifs between β2-β3 (Pro and Val), β5-β6 (Pro) and β6-β7 (Gly-Val-Leu) (Fig. 3C). The protein structure model of Poyun08063

Table 1 Characteristics of PyHSP20s in *P. yunnanensis*

ID	Number of amino acids	Molecular weight (Da)	Theoretical pI	Aliphatic index	Grand average of hydropathicity (GRAVY)	pfam00011 Domain	subcellular localization
Poyun01954	243	27,398.19	8.28	82.63	−0.647	142–243	mitochondrion
Poyun02516	147	15,707.25	8.76	88.78	−0.059	35–140	cytoplasm
Poyun06896	150	17,050.18	4.68	71.33	−0.314	61–146	cytoplasm
Poyun15799	237	26,786.38	7.9	65.02	−0.741	130–237	chloroplast
Poyun17694	331	36,790.86	8.64	63.75	−0.761	23–99	cytoplasm
Poyun24210	222	23,702.7	5.7	69.73	−0.319	123–215	cytoplasm
Poyun24897	192	21,790.05	6	98.39	−0.414	71–153	endoplasmic reticulum
Poyun31249	198	22,418.05	5.38	74.29	−0.721	32–112	cytoplasm
Poyun36443	238	26,204.72	6.92	77.73	−0.55	140–238	chloroplast
Poyun38019	229	25,228.92	6.53	77.95	−0.516	125–223	mitochondrion
Poyun02471	144	16,201.64	9.01	82.57	−0.466	40–143	cytoplasm
Poyun02640	158	18,231.67	5.98	68.42	−0.639	54–157	cytoplasm
Poyun03516	144	16,360.72	5.34	83.06	−0.605	31–134	cytoplasm
Poyun03517	144	16,360.72	5.34	83.06	−0.605	31–134	cytoplasm
Poyun03518	144	16,360.72	5.34	83.06	−0.605	31–134	cytoplasm
Poyun05362	157	17,560.19	6.16	74.52	−0.489	51–138	cytoplasm
Poyun06689	140	15,887	5.99	80	−0.656	36–139	cytoplasm
Poyun08063	161	18,314.7	5.98	70.81	−0.631	55–160	cytoplasm
Poyun08064	158	18,275.81	6.77	69.68	−0.685	54–157	cytoplasm
Poyun08065	154	17,750.21	6.19	75.19	−0.626	50–153	cytoplasm
Poyun08066	158	18,292.64	5.22	72.72	−0.647	54–157	cytoplasm
Poyun08067	145	16,710.09	6.2	78.62	−0.792	38–141	cytoplasm
Poyun08071	160	18,266.47	5.08	66.94	−0.652	56–159	cytoplasm
Poyun17195	162	18,605.03	6.35	70.37	−0.725	58–161	cytoplasm
Poyun18465	156	17,566	6.43	79.81	−0.537	48–154	cytoplasm
Poyun21268	137	15,747.83	5.24	86.86	−0.411	34–112	cytoplasm
Poyun22158	162	18,322.72	6.2	70.37	−0.68	58–161	cytoplasm
Poyun22160	162	18,322.77	6.77	70.37	−0.682	58–161	cytoplasm
Poyun22161	162	18,251.68	6.2	70.37	−0.656	58–161	cytoplasm
Poyun24552	169	19,233.14	6.9	85.38	−0.367	22–98	cytoplasm
Poyun24553	168	19,162.02	6.96	84.7	−0.421	22–107	cytoplasm
Poyun27967	172	19,271.14	6.52	85	−0.276	22–108	cytoplasm
Poyun28484	152	17,534.83	5.71	69.14	−0.705	48–151	cytoplasm
Poyun28485	152	17,465.72	5.49	69.14	−0.681	48–151	cytoplasm
Poyun34740	150	16,943.05	4.62	72.73	−0.325	61–146	cytoplasm
Poyun38186	156	18,040.38	6.19	64.23	−0.796	51–155	cytoplasm
Poyun38292	155	17,820.24	6.19	73.48	−0.683	50–154	cytoplasm
Poyun39254	158	18,011.36	6.19	72.72	−0.64	54–156	cytoplasm
Poyun39298	247	27,359.89	6.11	80.04	−0.71	37–112	cytoplasm
Poyun24554	201	23,037.49	9.23	78.01	−0.531	19–105	cytoplasm
Poyun07304	180	20,439.98	4.67	77.94	−0.495	81–169	cytoplasm
Poyun18517	225	25,633.88	5.99	65.47	−0.696	140–220	cytoplasm
Poyun18518	214	23,572.79	9.18	76.96	−0.484	119–203	mitochondrion
Poyun18860	212	23,786.75	5.75	77.69	−0.64	116–189	mitochondrion
Poyun04237	196	22,378.79	5.01	60.61	−0.656	87–183	nucleus
Poyun08023	323	35,986.8	9.19	67.62	−0.94	23–110	nucleus
Poyun20302	202	23,098.88	5.13	70.45	−0.554	87–194	nucleus
Poyun26228	266	29,702.92	9.15	75.49	−0.491	168–265	nucleus
Poyun38293	189	21,863.09	9.16	74.66	−0.814	70–174	nucleus
Poyun39253	227	26,521.11	9.01	67.31	−0.856	123–226	nucleus
Poyun39300	208	23,722.28	9.16	79.13	−0.602	35–107	nucleus
Poyun16995	155	17,693.15	6.25	74.13	−0.693	30–135	cytoplasm
Poyun33147	141	15,943.08	6.98	76.67	−0.613	31–138	peroxisome

The subcellular localization of PyHSP20s just list the most likely localization from WOLF PSORT (https://wolf-psort.hgc.jp/)

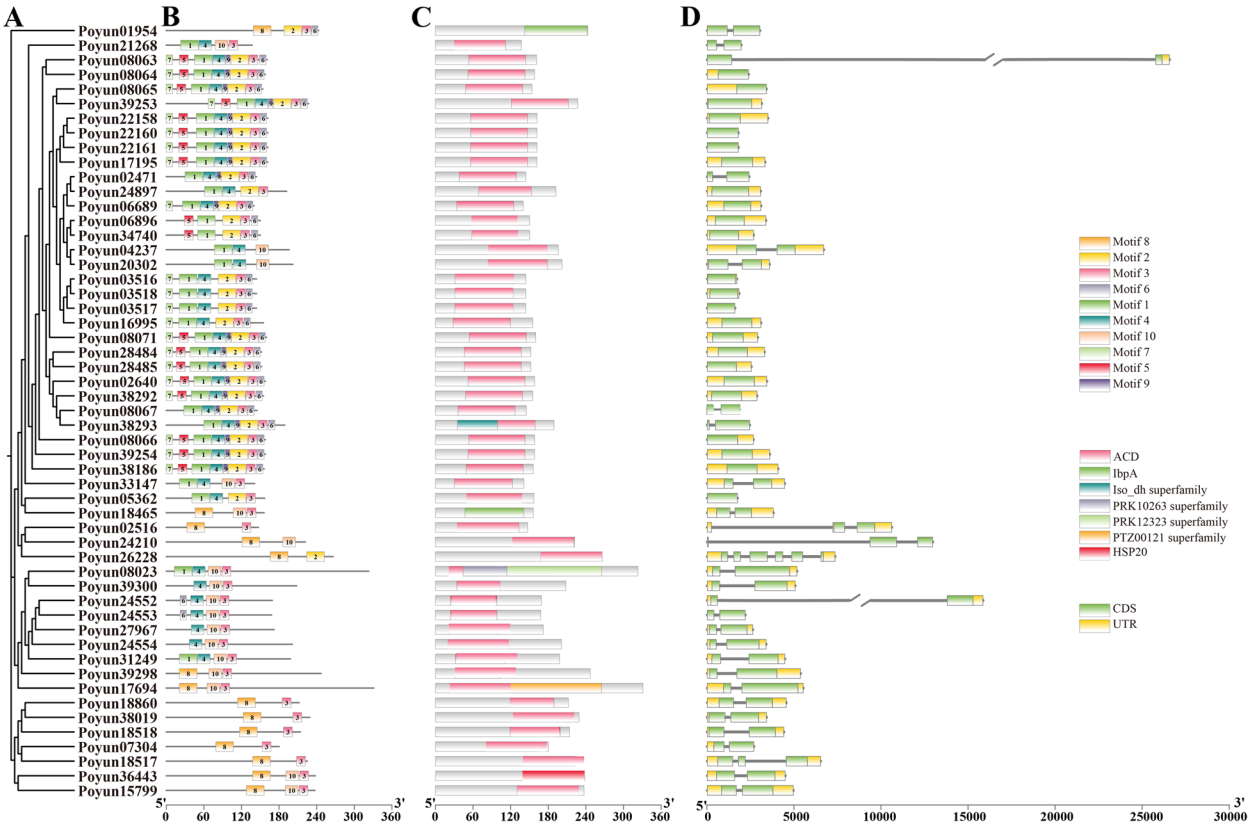

Fig. 1 Phylogenetic analysis, conserved motifs, domains, and gene structure of PyHSP20s. **A** Phylogenetic tree constructed using 53 PyHSP20s protein sequences with the maximum likelihood method. **B** Conserved motifs of PyHSP20 proteins predicted using Multiple Em for Motif Elicitation (MEME) with a 10 threshold value. C. Schematic diagram of the main domains of PyHSP20 proteins. D. Gene structure of the PyHSP20 coding genes

was constructed based on the A0A059AGU1.1.A template (sHSP domain-containing protein) with 80.25% sequence identity, and a MQE of 0.75 GMQE was generated using ProMod3 3.4.0. All 24 CI PyHSP20s exhibited structural similarity to Poyun08063, containing six conserved η-helices and eight β folds (β1-β8), with highly conserved sequence motifs between β3-β4 (Phe, Leu-Pro and Val) and β6-β7 (Phe, Arg, Leu-Pro, Lys, and Lys-Ala) (Fig. 3B).

Chromosome location and gene duplication of *PyHSP20* genes

Based on the chromosome locations of HSP20s in *P. yunnanensis*, 53 *PyHSP20* genes were found to be unevenly distributed across 15 chromosomes (Fig. 4A). Chromosomes 1 and 19 had the highest number of *PyHSP20* genes (Fig. 4A). Chromosome 3 with 7 genes, exhibited concentrated distribution sites. Chromosome 10 had 5 *PyHSP20* genes. Chromosomes 2, 6 and 7 each had 4 *PyHSP20* genes. Chromosomes 9 and 12 each contained 3 *PyHSP20* genes, and chromosome 8 had 2 *PyHSP20* genes. Additionally, chromosomes 11, 14, 15, 16 and 18 each carried only one *PyHSP20* gene. Based on these findings, we hypothesized that tandem duplications and segment duplications might account for the expansion of *PyHSP20* genes. To further investigate the evolutionary relationships between different *PyHSP20* genes, a collinearity analysis was performed. The eight pairs of *PyHSP20* genes exhibited similarities in their sequences and shared origins (Fig. 4B, Table S3). The collinearity pairs were exclusively identified within the same subfamily groups (CI, CV, CVI, and CVII), suggesting that the expansion of PyHSP20 genes resulted from tandem replication within the same subfamilies.

The predicted *cis*-elements in the promoter regions of *PyHSP20* genes

Cis-elements were predicted to act as binding sites for regulatory proteins in the promoter regions of genes, thereby facilitating the regulation of gene expression. In the promoter regions of *PyHSP20* genes, numerous *cis*-elements associated with stress response and hormone responsiveness were identified, with light-responsive elements being the most abundant (Fig. 5, Table S4).

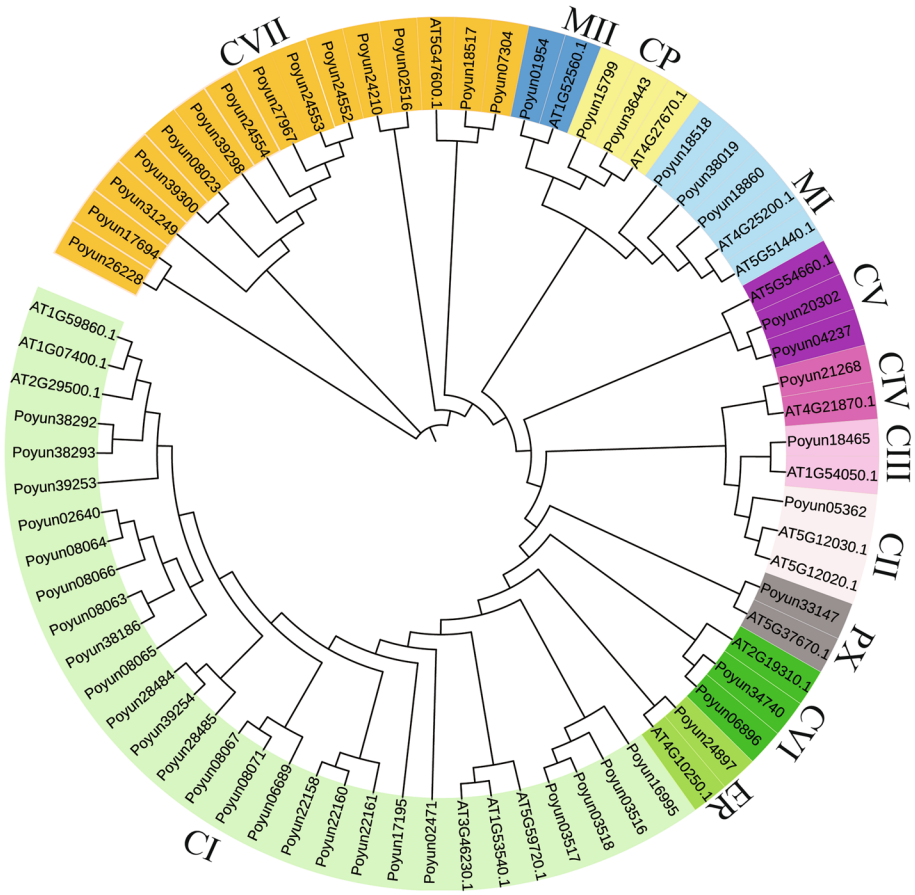

Fig. 2 Phylogenetic tree of HSP20s of *P. yunnanensis* and *A. thaliana*. Phylogenetic tree constructed using protein sequences of HSP20s from *P. yunnanensis* and *A. thaliana*. The twelve subfamilies are marked by different colors

Notably, all *PyHSP20*s contained light-responsive elements in their promoter regions, a characteristic feature of plant functional genes. Anaerobic induction-related elements were the second most abundant *cis*-elements in the *PyHSP20* promoters, present in 49 *PyHSP20* promoters, excluding *Poyun08063*, *Poyun28484*, *Poyun39253*, and *Poyun18517*. This suggested the involvement of PyHSP20s in abiotic stress responses. ABA-responsive elements, which were crucial for stress response and gene regulation, were universally distributed in the promoters of 41 *PyHSP20* members across all subfamilies. MeJA (Methyl Jasmonate) -responsive elements were identified in 43 *PyHSP20* promoters, excluding members of the CVI, ER, and CP subfamilies. MYB binding elements were identified in 35 *PyHSP20* promoters, excluding members of the CIII and ER subfamilies. Additionally, elements related to plant growth and hormone regulation, as well as stress-responsive elements (e.g., zein metabolism, salicylic acid, gibberellin, auxin, defense and low-temperature responsiveness), were identified in the promoters of more than twenty *PyHSP20s*. The widespread distribution of these detected elements, associated with regulatory hormones and stress responsiveness in the *PyHSP20* promoters highlighted their functional associations.

Relative expression levels of *PyHSP20s*

To predict the functions of different HSP20 subfamily members, we examined the expression levels of *PyHSP20* homologous in *Populus trichocarpa*, the closest evolutionary relative to *P. yunnanensis* [28], across various tissues and hormone treatments using expression data from Phytozome. Within each HSP20 subfamily, we identified several representative members whose expression was significantly induced by hormones, particularly stress-related hormones such as abscisic acid (ABA), methyl jasmonate (MeJA), and Salicylic acid (SA). Notably, members of the CI, CII, and MII subfamilies exhibited significant induction (Fig. 6A). Additionally, expression patterns of different

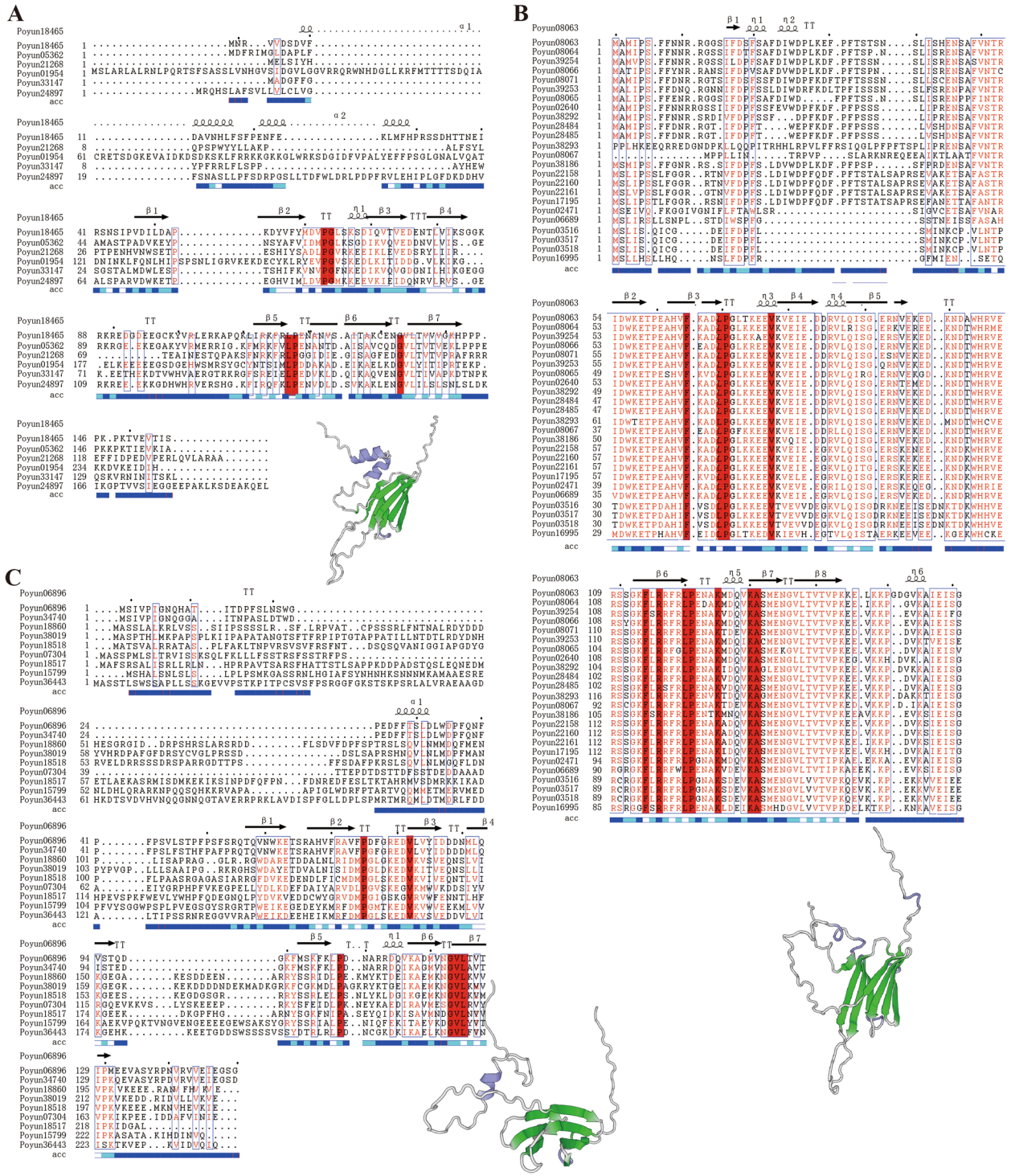

Fig. 3 Sequence and structure analysis of PyHSP20s based on their 3D models. Multiple sequence alignment of PyHSP20s were analyzed using ESPript 3.0 based on different protein models built with the SWISS-MODEL Workspace of Expasy. **A** Sequence alignment of five PyHSP20 proteins based on protein model of Poyun18465 built with A0A251JRI8.1.A template (SHSP domain-containing protein). **B** Sequence alignment of 24 CI PyHSP20 proteins based on protein model of Poyun08063 built with A0A251JRI8.1.A template (SHSP domain-containing protein). **C** Sequence alignment of 8 PyHSP20 proteins based on protein model of Poyun06896 built with A0A836Y6J7.1.A template. Highly conserved residues are highlighted in red and blue boxes, whereas variable residues are highlighted in black. Conserved amino acids: P (Pro), G(Gly), L (Leu), V (Val), F (Phe), R (Arg), K (Lys), A (Ala). The structure markers on the top and bottom of the sequence alignment correspond to the 3D models built on SWISS-MODEL [27]. A helices were highlighted in purple, and β folds were highlighted in green

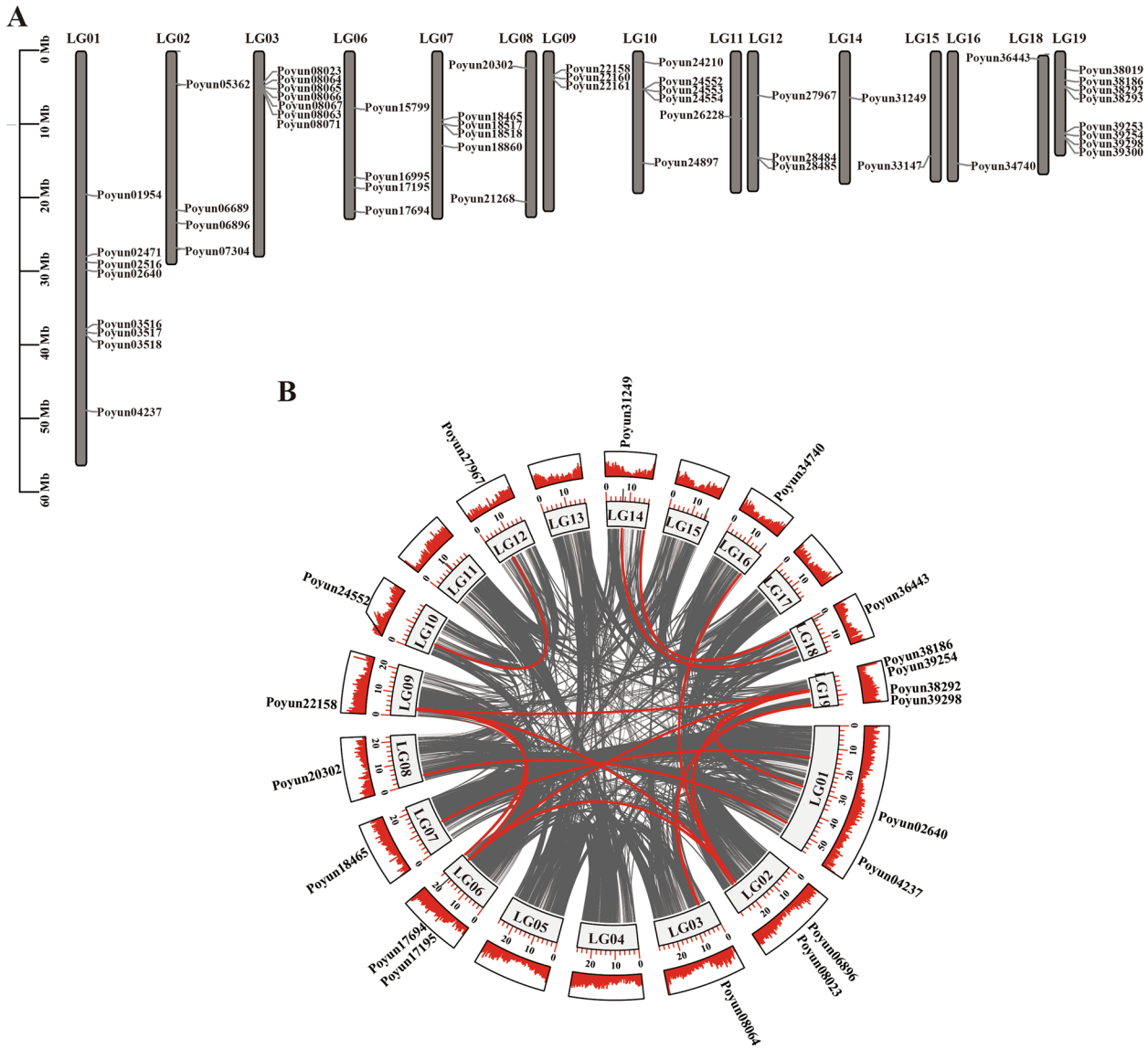

Fig. 4 Chromosome location and Collinearity analysis of PyHSP20. **A** Chromosome location of PyHSP20 genes. **B** Collinearity pairs of *PyHSP20* genes were connected with red lines. LG represented chromosome

PyHSP20 subfamilies were observed in our RNA-seq data of *P. yunnanensis* under NaCl treatment (Fig. 6B). These variations in expression further indicated functional differentiation among the HSP20s.

To investigate the functional differentiation of HSP20s in *P. yunnanensis* during stress responses, we measured the relative expression of representative *PyHSP20*s under various stress conditions using qRT-PCR. All detected *PyHSP20*s exhibited stress responses under stress conditions, with relative expression levels significantly different from those of the untreated controls. This finding was consistent with the number of *cis*-elements related to the stress response detected (Fig. 7). Specifically, *Poyun17195*, *Poyun28484*, *Poyun38186*, *Poyun38292*, *Poyun16995*, and *Poyun22158*, belonging to the CI subfamily of HSP20s, were significantly induced under salt, heat, drought, mannitol and ABA treatments. Notably, except for *Poyun16995*, which was evolutionarily closest to the ER subfamily and was primarily induced by heat stress, all other tested CI *PyHSP20* members were most strongly induced by salt stress. *Poyun15799* (CP subfamily), *Poyun18860* (MI subfamily), and *Poyun34740* (CVI subfamily) exhibited similar changes in expression under stress conditions, with the greatest increase in expression under heat stress and relatively high expression levels

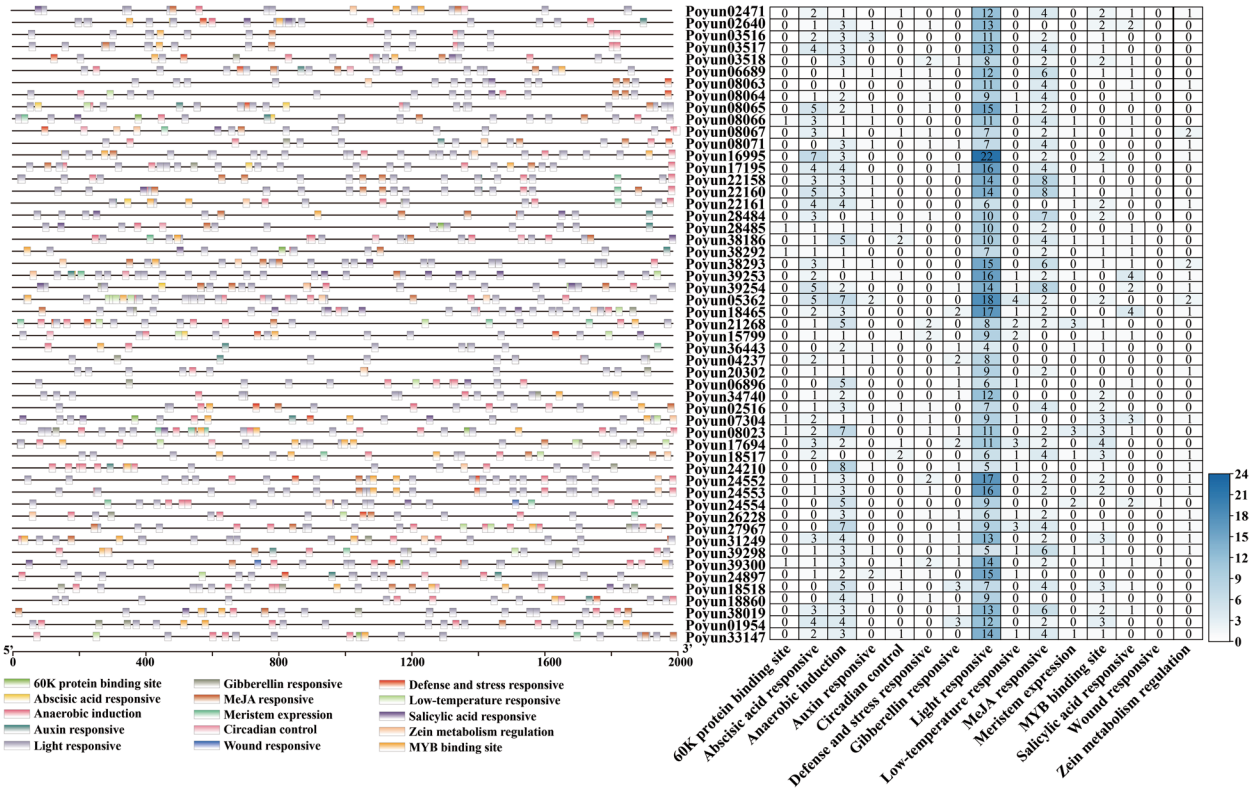

Fig. 5 *Cis*-elements predicted within *PyHSP20s* promoters. Different colored boxes on the left represented different *cis*-elements in the promoters of PyHSP20s within 2000 bp sequences before their transcription start sites. The right color-coded matrix summarized the number of *cis*-elements in the promoters of *PyHSP20s*

under salt, drought, ABA, drought and mannitol treatments. *Poyun05362*, a representative member of the CII *PyHSP20* subfamily, exhibited the highest expression under salt stress, followed by mannitol and heat stress. However, its expression level was lower under drought and ABA stress compared to untreated conditions, despite the presence of numerous ABA-related *cis*-elements on its promoter. These observations indicated significant differences in expression trends among subfamily members in response to various stress conditions, highlighting the functional differentiation of PyHSP20s and their regulation by promoter *cis*-elements.

The proteins predicted to interact with PyHSP20s underwent a Y2H assay

HSP20 exhibited polymerization activity, enabling it to bind with other proteins. The interaction network of the PyHSP20s predicted via STRING revealed that most PyHSP20s can interact with other heat-shock proteins. Some PyHSP20s were predicted to potentially interact with functional proteins and enzymes, including shikimate kinase and chloroplastic isoform proteins (Fig. 8A; Table S6). Interaction relationships between different HSP20 proteins were predicted (Fig. 8B).

To further investigate the function of PyHSP20s, we conducted Y2H assays to examine the interactions between PyHSP20s and their predicted interacting proteins. During the examination of PyHSP20s' interactions, Poyun18860 (MI) exhibited the highest number of interactions and was found to bind with CHLOROPLAST UNUSUAL POSITIONING 1 (CHUP1, Poyun15301), Shikimate kinase-like protein (Poyun32474), and PyHSP20s (Poyun05362 (CII), Poyun01954 (MII)). Shikimate kinase-like protein (SKL2) has been reported to function as a drought and salinity stress response factor, enhancing rice stress resistance by binding with the ABA stress ripening (ASR) protein and the ABA signaling pathway [29]. CHLOROPLAST UNUSUAL POSITIONING 1 (CHUP1) functions as a plant-specific actin polymerization factor, regulating the movement of chloroplasts [30]. The protein binding relationships of PyHSP20s with CHUP1 and SKL provided valuable insights into the function of PyHSP20s under stress. The interactions of

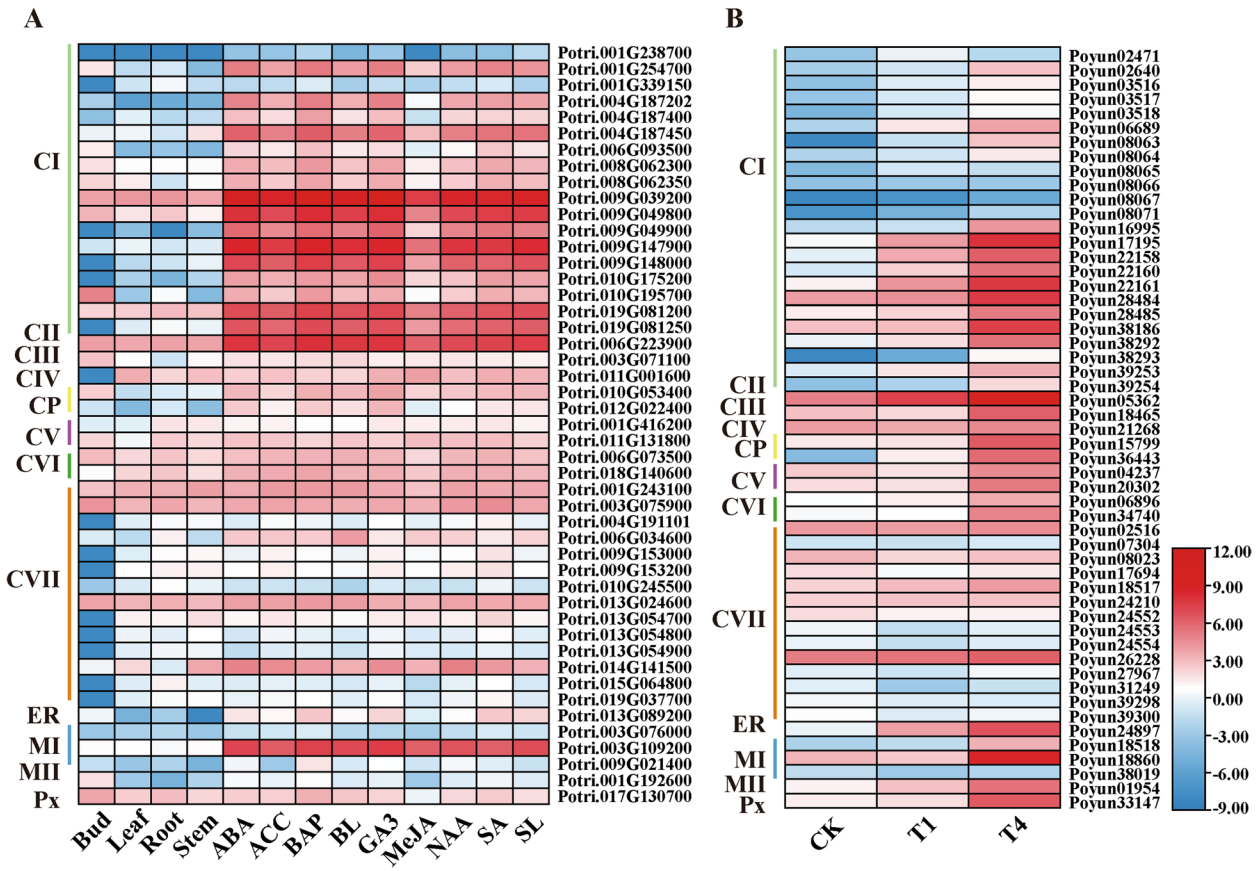

Fig. 6 The expression profiles of PyHSP20s and their homologous in *P. trichocarpa* across different tissues and under stress conditions. **A** Expression profiles of PyHSP20 homologous in *P. trichocarpa* across different tissues and hormone treatments. Expression data for various tissues (root, stem, bud, leaf) and hormone-treated leaves (ABA (Abscisic acid), SA (Salicylic acid), MeJA (Methyl Jasmonate), ACC (1-Aminocyclopropane-1-carboxylic acid), BAP (6-Benzylaminopurine), BL (Brassinolide), GA (Gibberellic Acid), NAA (Naphthaleneacetic acid), SL (Strigolactone)) were retrieved from Phytozome 13. **B** Expression profiles of *PyHSP20*s under salt stress conditions in *P. yunnanensis*. RNA-seq data: T1, two days of 25 mM NaCl treatment; T4, the fourth cycle of NaCl treatment (each cycle consisting of two days of NaCl treatment followed by three days of recovery, as described in the methods); CK, untreated control. All heatmaps were generated using TBtools with log2-transformed FPKM values (Table S5)

PyHSP20s with different activation and binding partners revealed their functional flexibility.

Discussion

Small heat shock proteins (sHSPs) are ubiquitous chaperone proteins that play crucial roles in maintaining ATP-independent protein stability by preventing protein misfolding and aggregation [1]. In addition to heat stress, plant sHSPs also play roles in tolerating various types of stress [31]. With advances in sequencing technology, numerous sHSPs have been identified across various plant species, such as *Arabidopsis*, potato, peach, and so on [17, 20, 32]. However, limited information is available regarding HSP20 proteins in *Populus* species, particularly *P. yunnanensis*, an economically important tree species native to China [24]. In this study, we identified a total of 53 HSP20 proteins in *P. yunnanensis*, which were classified into twelve subfamilies based on their phylogenetic relationships and subcellular localization. In *P. yunnanensis*, both protein and gene structures were conserved among HSP20 proteins within the same subfamily. The eight pairs of PyHSP20s identified through collinearity analysis offered insights into gene expansion and duplication events. The expression profiles of the PyHSP20s also confirmed the functional conservation of subfamily members under stress conditions, reflecting their cooperative ability with functional proteins.

The number (53) of *P. yunnanensis* HSP20 proteins was more similar to those of perennial plants than to those of annual herbs (Table 1) [18, 20–23, 32, 33]. The number of HSP20 proteins in poplar species was comparable (Figure S1). The diversity of the physicochemical properties among PyHSP20 proteins indicated functional differentiation, particularly in subcellular localization (Table 1) [34]. Protein and gene structures varied among PyHSP20 members. Conserved short

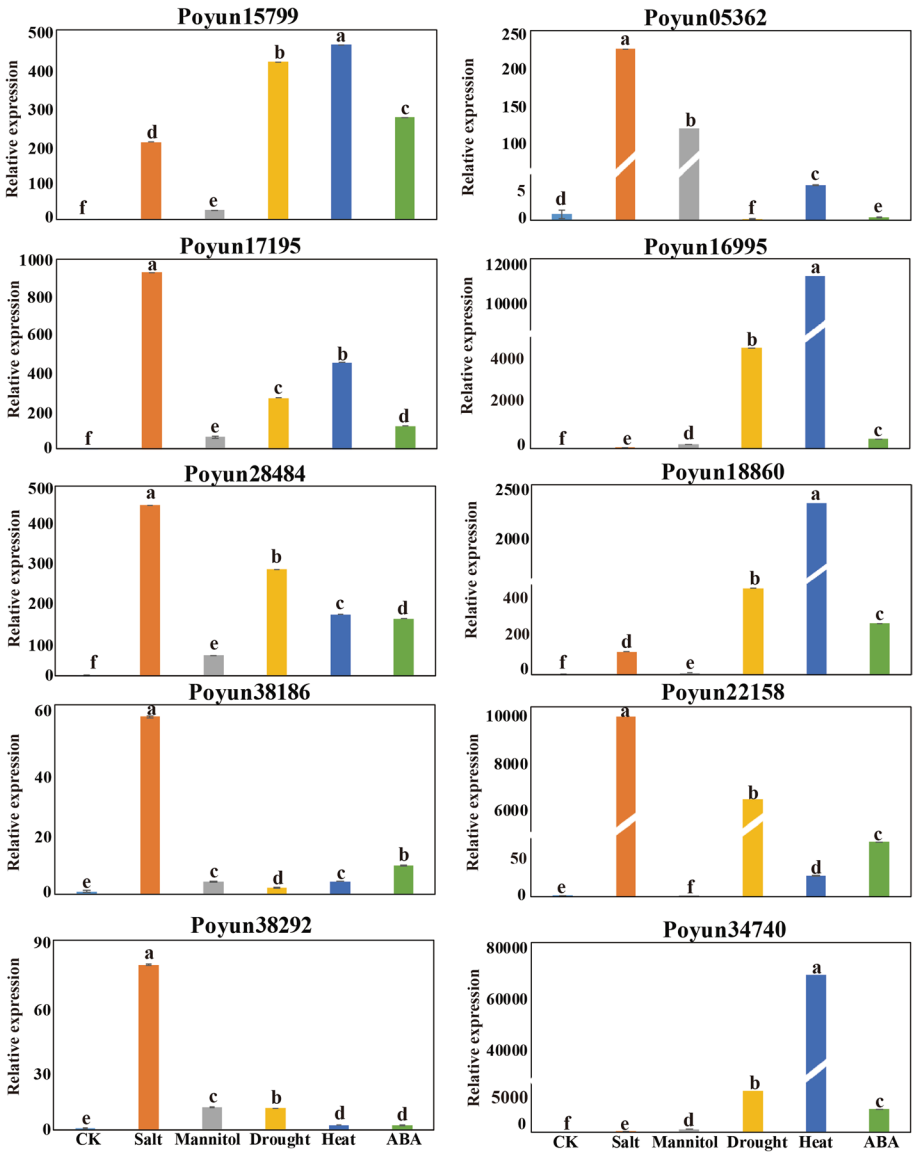

Fig. 7 Relative expression of *PyHsp20*s under stress. Relative expression levels of *PyHSP20*s under various stress conditions were depicted. Different colored boxes represented different stress and hormone treatment: blue for un-treated control; red for salt treatment; grey for D-mannitol treatment; yellow for drought treatment; purple for heat treatment; green for ABA treatment. Error bars indicated the standard error of three replicates. Significant differences ($p < 0.05$) were denoted by different letters above the bars

motifs were identified and arranged in different orders at various positions within PyHSP20 sequences [35]. The varying numbers of exons in *PyHSP20*s correlated with motif distribution and phylogenetic relationships (Fig. 1). Most *PyHSP20*s had fewer than three exons, except for three *PyHSP20*s (Poyun26228, Poyun18517, Poyun24210) belonging to the same phylogenetic branch. Conversely, gene lengths of *PyHSP20*s varied due to the presence of long introns, which didn't affect protein sequence length or composition but may influence gene function, regulation and evolution [36, 37]. IbpA proteins, a type of HSP20 involved in heat stress resistance, are primarily found in bacteria [38]. The presence of IbpA domains in some PyHSP20s (Poyun01954, Poyun18465) may provide insights into their evolution and origin. The identification of additional domains in Poyun01954 and Poyun18465 could clarify the specificity of their structure and function [39].

The classification of HSP20 proteins, based on their phylogenetic relationships and subcellular localization, is a widely adopted approach [6]. A total of 53 PyHSP20s

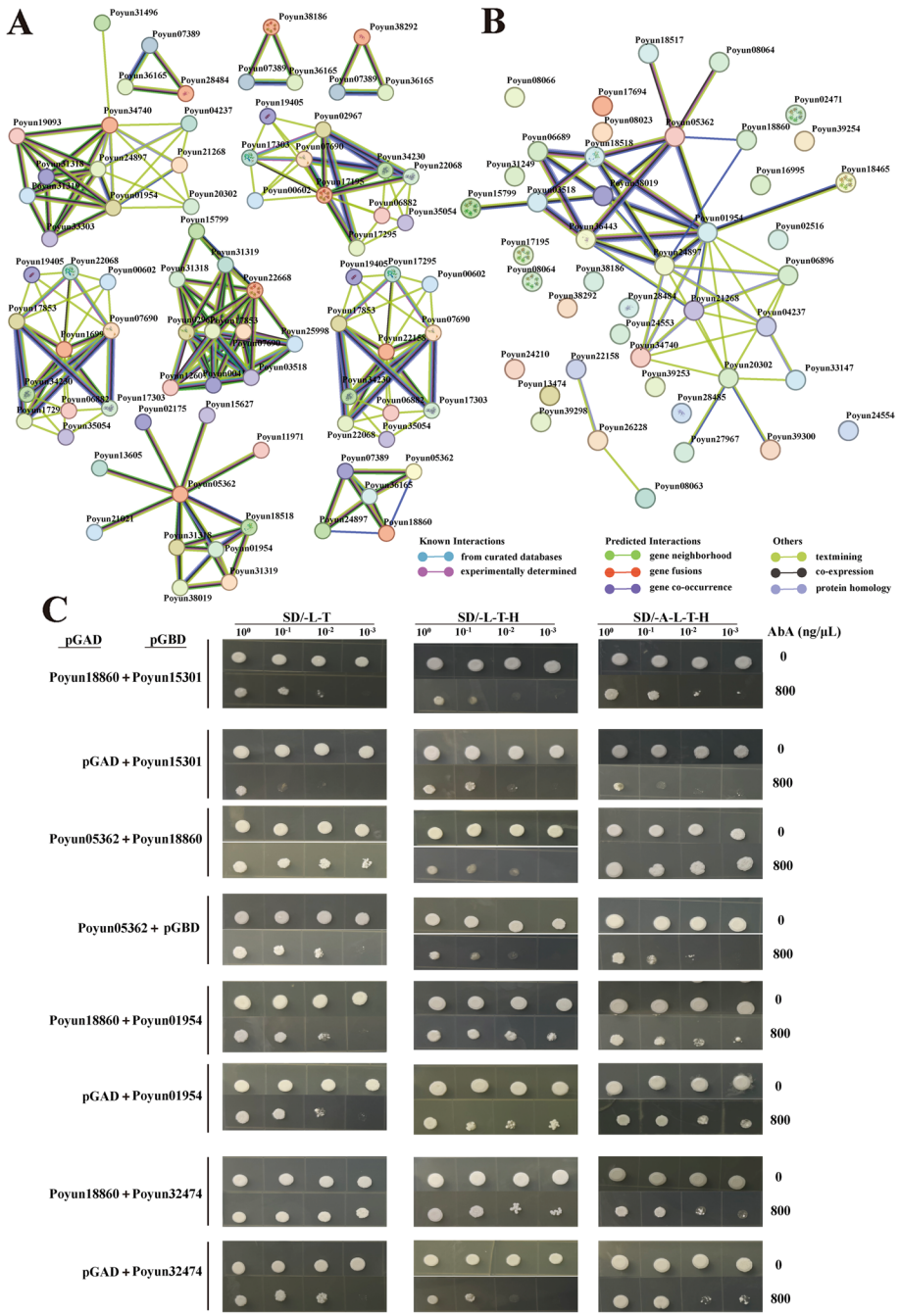

Fig. 8 Interacting proteins and interaction relationship of PyHSP20s predicted in STRING and Y2H assay. **A** The interacted proteins of PyHSP20s were presented and predicted in STRING. **B** The interaction relationships of PyHSP20s were speculated based on their sequences and structures in STRING. Different colored lines represented various interaction types of proteins, with details marked under the Figures. **C** Protein–protein interactions detected using the Yeast Two-Hybrid (Y2H) assay. The coding sequences of test proteins were fused to the activation domain vector (pGADT7) and the GAL4 DNA binding domain vector (pGBKT7), respectively. The negative control involved removing one of the coding sequences of interacting proteins. SD/-L–T, SD/-H–L-T, and SD/-A-H–L-T represented SD-Leu-Trp, SD-His-Leu-Trp, and SD-Ade-His-Leu-Trp culture medium, respectively. Yeast cultures with varying concentrations (10^0, 10^{-1}, 10^{-2}, 10^{-3}) were used for testing. AbA (Aurothiomalate A, 800 ng/μL) was added to facilitate selection

were classified into twelve subfamilies (Fig. 2), consistent with the classifications observed in *P. euphratica*, *P. deltoides*, *P. trichocarpa*, *P. tomentosa*, *P. alba* (Figure S1) and *Arabidopsis* [32]. However, the classification differed from that of *Prunus persica*, which comprises 13 subfamilies, and peach, which contains 10 subfamilies [20, 40]. This subfamily classification may offer valuable insights into the functional differentiation of HSP20 proteins [32]. The cytoplasmic PyHSP20s formed the largest cluster, facilitating chaperone binding and the capture of unfolded proteins [11, 12]. The CI subfamily of PyHSP20 represented the largest HSP20 subfamily, consistent with findings in potato, peach, *Arabidopsis*, *Coix* and grape [17–20, 32], underscoring the evolutionary conservation of HSP20 proteins across species. The α-crystallin domain (ACD), a conserved hallmark of sHSP20s, comprises seven or eight antiparallel β strands that are essential for maintaining their binding function [8]. Although all PyHSP20 proteins possessed an ACD domain (Fig. 3), we generated three distinct 3D protein models for PyHSP20s using complete protein sequences and diverse sHSP templates from various species (Fig. 3). These models revealed significant differences in both sequence and structure among the PyHSP20 proteins [27]. The CI subfamily, the largest among PyHSP20s, contains eight β-strands, exceeding the number found in other subfamilies, which may explain its functional specificity. The conserved amino acids located within the β-strand regions further contribute to the binding affinity of HSP20 [41]. Beyond the core ACD domain, PyHSP20s feature an N-terminal arm with variable lengths and sequence diversity, potentially influencing their function and structure [42]. Compared to other PyHSP20 subfamilies, the CI subfamily exhibited shorter N-terminal and C-terminal arms, indicative of functional divergence [43].

Both the longest chromosome, LG01, and the shortest chromosome, LG19, contained the largest number of *PyHSP20* genes, indicating an uneven distribution of *PyHSP20* (Fig. 4). Unlike those in peach, potato and grape, *PyHSP20* genes were not located in the terminal regions of chromosomes [17, 18, 20]. Gene duplication events are the primary drivers of gene family expansion and genome evolution [44]. Following collinearity analysis, 16 *PyHSP20* genes were identified, all belonging to the CI, CV, CVI, and CVII subfamilies (Table S3). These PyHSP20s were localized in the cytoplasm, possessed similar molecular weights, and shared similar amino acid sequences, elucidating the origins of HSP20 members within the same subfamily. The expansion of the *HSP20* gene family in *P. yunnanensis* may differ from that in peach, where a collinear relationship exists between the CI and ER subfamilies [20, 40]. Transcriptional regulation plays a crucial role in controlling gene expression levels [45]. The *cis*-elements detected in the promoters of *HSP20* genes revealed the transcriptional regulatory functions of transcription factors and hormone signals, including ABA, MeJA, and MYB (Fig. 5, Table S4). ABA is a plant hormone extensively involved in plant stress responses, and had been found to increase abiotic stress tolerance in creeping bentgrass through the function of AsHSP17 [46]. The exogenous application of MeJA enhances stress tolerance in perennial ryegrass by regulating the expression of various pathway genes [47]. The increased expression of PyHSP20 homologous in *P. trichocarpa* and PyHSP20s under hormone treatment and stress may be consistent with the regulatory effects of their promoters (Fig. 6). The enhanced expression of representative *PyHSP20*s, as verified by qRT-PCR, further confirmed the function of HSP20 under abiotic stress (Fig. 7). In addition to transcription activation, the formation of protein complexes with interacting proteins, which act as molecular chaperones, is a key functional mechanism by which HSP20 prevents protein misfolding and irreversible aggregation [7]. SKL2, a shikimate kinase-like protein, has been identified as a factor responding to drought and salinity stress. It enhances rice's resistance to these stresses by interacting with the ABA stress ripening (ASR) protein and modulating ABA signaling responses [29]. CHUP1 (CHLOROPLAST UNUSUAL POSITIONING 1) is a plant-specific actin polymerization factor that regulates chloroplast mobility [30]. The interaction between PyHSP20s from the CII, MI and MII subfamilies, along with SKL2, and CHUP1, as confirmed by Y2H assays, provided insights into the functional mechanisms and stress response pathways of PyHSP20s (Fig. 8, Table S6). These findings suggested that the interactions between PyHSP20s and functional proteins may enhance the stress response of PyHSP20s and improve plant stress tolerance.

Conclusions

In summary, the 53 PyHSP20s identified in this study were categorized into 12 subfamilies: CI, CII, CIII, CIV, CV, CVI, CVII, MI, MII, ER, CP, and Px. The variations in member numbers, gene structures, and protein structures were correlated with their protein subcellular locations, consistent with observations in five other poplar species. The β-fold structures, characteristic of the PyHSP20 ACD domains, exhibited differences in β-fold numbers, lengths, and sequence-variable N-terminal arms across different PyHSP20 subfamilies. Analysis of gene locations and duplications revealed that tandem segmental duplications significantly contributed to the expansion of *PyHSP20* genes, particularly those located in the cytoplasm. The *cis*-elements related to the stress response, detected in the promoters,

provided insights into the stress response mechanisms of PyHSP20s, as verified through database searches and qRT-PCR assays. The interactions among PyHSP20s, ABA response proteins, and other HSP proteins also elucidated the mechanisms behind the induced expression of PyHSP20s. Overall, this study provided a solid foundation for further functional research on *PyHSP20* family genes in poplar.

Materials and methods

Identification of HSP20 proteins in *P. yunnanensis*

P. yunnanensis HSP20 proteins were initially identified through homolog analysis utilizing protein sequences from 19 *A. thaliana* HSP20 proteins via the Basic Local Alignment Search Tool (BLAST, version: blast-2.14.1+, downloaded on Oct 9, 2023), with an E-value threshold set at 1e-5 [6]. The genome data used in this study were obtained from the National Genomics Data Center (NGDC, https://ngdc.cncb.ac.cn) with the accession number PRJCA008692 [48]. The candidate *P. yunnanensis* HSP20 proteins were validated via the Conserved Domain Database (CDD; https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, accessed on Apr 8, 2024) and SMART (http://smart.embl.de/, Mar 23, 2024), with the Pfam domain pfam00011 as the validation criterion [49, 50]. To ensure the comprehensive identification of HSP20 in *P. yunnanensis*, the initially identified HSP20 proteins were used as queries for further homology analysis within the *P. yunnanensis* genome [51].

Physical, chemical, and phylogenetic analysis of PyHSP20s

The physical and chemical properties of all *P. yunnanensis* candidate HSP20 proteins were evaluated using the ProtParam tool on Expasy (https://web.expasy.org/protparam/, Apr 8, 2024). The subcellular locations of the 53 PyHSP20s were predicted using WOLF PSORT (https://wolf-psort.hgc.jp/; accessed on Apr 10, 2024). To examine the phylogenetic relationships among the 53 PyHSP20s, a phylogenetic tree was constructed based on the protein sequences of PyHSP20s and 19 *A. thaliana* HSP20 proteins using MEGA11 with maximum likelihood (ML) method and 1000 bootstrap replicates [52]. Sequence alignment was conducted using BioEdit (version 7.0.9.0) for the complete protein sequences of the 53 PyHSP20s.

Analysis of the motifs, domains, gene structure and *cis*-elements of PyHSP20s

The conserved motifs were identified using Multiple Em for Motif Elicitation (MEME, https://meme-suite.org/meme/tools/meme, version 5.5.5, Apr 15, 2024). The conserved domains of PyHSP20s were identified using the Conserved Domain Database (CDD, accessed on Apr 15, 2024). The motifs, domains and gene structures were visualized using the BioSequence Structure Illustrator package in TBtools (version 2.083) [53]. The conserved *cis*-elements of PyHSP20s were identified using PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on Apr 8, 2024) with 2 kb sequences upstream of the start codon and visualized using the BioSequence Structure Illustrator package in TBtools.

Chromosome location and collinearity analysis of *PyHSP20*s

The chromosomal locations of *PyHSP20*s in *P. yunnanensis* were mapped using the Show Genes on the Chromosome package in TBtools. Collinearity analysis of PyHSP20s was conducted using the Comparative Genomics package in TBtools.

Plant material treatment, qRT-PCR and RNA-seq

Poplar cuttings were cultivated in pots containing a 3:1:1 mixture of humus, quartz sand, and perlite in the greenhouse at Southwest Forestry University (E102.74°, N25.17°) under controlled environmental conditions, with temperatures maintained at 25 °C during the light period and 18 °C during the dark period. Natural light cycles were employed to simulate photoperiodic conditions. Two-month-old poplar plants were subjected to various abiotic stress treatments, including ABA (50 μmol/mL ABA for 1 day), heat (45℃ for 1 day), drought (withholding water for 2 days), mannitol (25% D-mannitol for 1 day), and salt (150 mM NaCl for 1 day). These treatments were conducted following the protocols established in our previous study [54]. For each treatment, ten healthy plants were used. After the stress treatments, leaves from both treated and control plants were immediately flash-frozen in liquid nitrogen and stored at −80 °C for subsequent analysis. In this study, two-month-old plants were subjected to a cycling treatment involving incremental salt (NaCl) concentrations over multiple cycles, each followed by a 3-day recovery period. The treatment regimen consisted of four cycles with increasing NaCl concentrations: the first cycle used 25 mM NaCl for 2 days, the second cycle used 50 mM NaCl for 2 days, and the third and fourth cycles both used 75 mM NaCl for 2 days. After each salt treatment, plants were allowed a 3-day recovery period. For sampling, T1 represented samples collected after the first cycle of 25 mM NaCl treatment, while T4 included samples from the end of the fourth cycle, which included exposures to 25 mM, 50 mM, and two rounds of 75 mM NaCl, each followed

by recovery periods. For RNA-seq analysis, leaves were collected from five plants and sequenced on the Illumina platform by OEbiotech Corporation in Shanghai, P.R. China.

Total RNA was extracted from the treated plant leaves using the RNAprep Pure Plant Plus Kit (Cat. DP441; Tiangen, Beijing, China). 1 μg of total RNA was reverse transcribed using the EasyScript® All-in-One First-Strand cDNA Synthesis SuperMix for qPCR Reagent Kit (Transgene, Beijing, China). The relative expression levels of the *HSP20* genes were measured using qRT-PCR with TransStart® Green qPCR SuperMix (Transgene, Beijing, China) and Bio-Rad CFX96 system. The internal control and data analysis were conducted according to previously reported methods [54]. The primers used for qRT-PCR are listed in Table S7.

Protein interaction relationship prediction and yeast two-hybrid assays

The potential interacting proteins of PyHSP20 were predicted using STRING (version 12.0, https://cn.string-db.org/cgi/input?sessionId=bMVUUvWVnIVl&input_page_show_search=on) based on protein sequences derived from the *Populus trichocarpa* genome. Yeast two-hybrid assays were employed to verify the interactions between HSP20 proteins and their potential interacting proteins. The coding sequences for verification were separately cloned and inserted into pGADT7 (AD, Clontech, USA) and pGBKT7 (BD, Clontech, USA) using Phanta Max Super-Fidelity DNA Polymerase (Vazyme p505, Nanjing, China) and the ClonExpress II One Step Cloning Kit (Vazyme c112, Nanjing, China). Yeast two-hybrid assays were performed following the instructions provided in the Matchmaker™ GAL4 Two-Hybrid System 3 & Libraries User Manual (PT3247-1, Clontech, USA) and previous methods [54]. The primers utilized for gene cloning are listed in Table S8.

Abbreviations

HSP	Heat shock protein
ACD	α-Crystallin domain
Y2H	Yeast Two-Hybrid assay
ABA	Abscisic Acid
RNA	Ribonucleic acid
pI	Isoelectric points
GRAVY	Grand average of hydropathicity
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
SA	Salicylic acid
GA	Gibberellin
MeJA	Methyl Jasmonate

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06264-9.

Supplementary Material 1: Figure S1. Phylogenetic tree of all poplar HSP20 proteins.

Supplementary Material 2: Table S1. 57 candidate HSP20 proteins identified in *P. yunnanensis* using local BLASTP with 19 *A. thaliana* HSP20s.

Supplementary Material 3: Table S2. The subcellular localization message of PyHSP20s.

Supplementary Material 4: Table S3. Data for collinearity analysis pairs.

Supplementary Material 5: Table S4. *Cis*-elements of *PyHSP20* promoters.

Supplementary Material 6: Table S5. Expression data for *PyHSP20*s and their homologous in *P. trichocarpa*.

Supplementary Material 7: Table S6. Interaction relationships of PyHSP20s and other proteins.

Supplementary Material 8: Table S7. Primers used in qRT-PCR.

Supplementary Material 9: Table S8. Primers used for vector construction.

Acknowledgements

We would like to express our sincere gratitude to all the colleagues in our laboratory for providing discussion and technical assistance. We would like to thank our editors and reviewers for their constructive comments that significantly improved the manuscript.

Author's contributions

C.L.S. and D.Y.K. worked on the experiments and provided the initial resulting images. L.D. and J.L.X. worked on the experiments. J.X.L. and M.A.Y. participated in the data analysis. Z.A.L. provided funding and framework suggestion. P.L. designed this research and wrote the manuscript. All the authors read and approved the final manuscript.

Funding

This work was financially supported by the Yunnan Fundamental Research Projects (202301AT070216, 202201AU070072), and the National Natural Science Foundation of China (32460375).

Data availability

The datasets supporting the conclusions of this article are included within the article and its additional files. The poplar sequences in this article were downloaded from BIG (https://ngdc.cncb.ac.cn/) with accession number PRJCA010101. The Arabidopsis HSP20 sequences can be found in TAIR (https://www.arabidopsis.org/). The expression profiling data of *Populus trichocarpa* can be found in Phytozome (https://phytozome-next.jgi.doe.gov/). The RNA-seq raw sequence data of *P. yunnanensis* have been deposited on the National Centre for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) repository with accession number PRJNA1222559 (https://www.ncbi.nlm.nih.gov/sra/PRJNA1222559).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details
[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. [2]Yunnan Provincial Key Laboratory for Conservation and Utilization of in-Forest Resource, Southwest Forestry University, Kunming, Yunnan, China.

Received: 2 December 2024 Accepted: 14 February 2025
Published online: 25 February 2025

References

1. Fu X. Chaperone function and mechanism of small heat-shock proteins. Acta Biochim Biophys Sin. 2014;46(5):347–56.
2. Voellmy R, Boellmann F. Chaperone regulation of the heat shock protein response. Adv Exp Med Biol. 2007;594:89–99.
3. Yildiz MT, Tutar L, Giritlioğlu NI, Bayram B, Tutar Y. MicroRNAs and heat shock proteins in breast cancer biology. Methods in molecular biology (Clifton, NJ). 2022;2257:293–310.
4. Van Montfort R, Slingsby C, Vierling E. Structure and function of the small heat shock protein/alpha-crystallin family of molecular chaperones. Adv Protein Chem. 2001;59:105–56.
5. Carra S, Alberti S, Arrigo PA, Benesch JL, Benjamin IJ, Boelens W, Bartelt-Kirbach B, Brundel B, Buchner J, Bukau B, et al. The growing world of small heat shock proteins: from structure to functions. Cell Stress Chaperones. 2017;22(4):601–11.
6. Waters ER, Vierling E. Plant small heat shock proteins - evolutionary and functional diversity. New Phytol. 2020;227(1):24–37.
7. Waters ER. The evolution, function, structure, and expression of the plant sHSPs. J Exp Bot. 2013;64(2):391–403.
8. Basha E, O'Neill H, Vierling E. Small heat shock proteins and α-crystallins: dynamic proteins with flexible functions. Trends Biochem Sci. 2012;37(3):106–17.
9. Mogk A, Bukau B. Role of sHsps in organizing cytosolic protein aggregation and disaggregation. Cell Stress Chaperones. 2017;22(4):493–502.
10. Baldwin AJ, Lioe H, Robinson CV, Kay LE, Benesch JL. αB-crystallin polydispersity is a consequence of unbiased quaternary dynamics. J Mol Biol. 2011;413(2):297–309.
11. Hilton GR, Lioe H, Stengel F, Baldwin AJ, Benesch JL. Small heat-shock proteins: paramedics of the cell. Top Curr Chem. 2013;328:69–98.
12. Nakamoto H, Vígh L. The small heat shock proteins and their clients. Cellular and molecular life sciences : CMLS. 2007;64(3):294–306.
13. Luján R, Lledías F, Martínez LM, Barreto R, Cassab GI, Nieto-Sotelo J. Small heat-shock proteins and leaf cooling capacity account for the unusual heat tolerance of the central spike leaves in Agave tequilana var. Weber. Plant Cell Environ. 2009;32(12):1791–803.
14. He LH, Chen JY, Kuang JF, Lu WJ. Expression of three sHSP genes involved in heat pretreatment-induced chilling tolerance in banana fruit. J Sci Food Agric. 2012;92(9):1924–30.
15. Sato Y, Okano K, Honda K. Effects of small heat shock proteins from thermotolerant bacteria on the stress resistance of Escherichia coli to temperature, pH, and hyperosmolarity. Extremophiles : life under extreme conditions. 2024;28(1):12.
16. Jiang C, Xu J, Zhang H, Zhang X, Shi J, Li M, Ming F. A cytosolic class I small heat shock protein, RcHSP17.8, of Rosa chinensis confers resistance to a variety of stresses to Escherichia coli, yeast and Arabidopsis thaliana. Plant Cell Environ. 2009;32(8):1046–59.
17. Li Y, Li Z, Dong L, Tang M, Zhang P, Zhang C, Cao Z, Zhu Q, Chen Y, Wang H, et al. Histone H1 acetylation at lysine 85 regulates chromatin condensation and genome stability upon DNA damage. Nucleic Acids Res. 2018;46(15):7716–30.
18. Ji XR, Yu YH, Ni PY, Zhang GH, Guo DL. Genome-wide identification of small heat-shock protein (HSP20) gene family in grape and expression profile during berry development. BMC Plant Biol. 2019;19(1):433.
19. Hua Y, Liu Q, Zhai Y, Zhao L, Zhu J, Zhang X, Jia Q, Liang Z, Wang D. Genome-wide analysis of the HSP20 gene family and its response to heat and drought stress in Coix (Coix lacryma-jobi L.). BMC genomics. 2023;24(1):478.
20. Zhang C, Zhang Y, Su Z, Shen Z, Song H, Cai Z, Xu J, Guo L, Zhang Y, Guo S, et al. Integrated analysis of HSP20 genes in the developing flesh of peach: identification, expression profiling, and subcellular localization. BMC Plant Biol. 2023;23(1):663.
21. Huang J, Hai Z, Wang R, Yu Y, Chen X, Liang W, Wang H. Genome-wide analysis of HSP20 gene family and expression patterns under heat stress in cucumber (Cucumis sativus L.). Front Plant Sci. 2022;13:968418.
22. Qi X, Di Z, Li Y, Zhang Z, Guo M, Tong B, Lu Y, Zhang Y, Zheng J. Genome-wide identification and expression profiling of heat shock protein 20 gene family in Sorbus pohuashanensis (Hance) Hedl under abiotic stress. Genes. 2022;13(12):2241.
23. Cui F, Taier G, Wang X, Wang K. Genome-wide analysis of the HSP20 gene family and expression patterns of HSP20 genes in response to abiotic stresses in *Cynodon transvaalensis*. Front Genet. 2021;12:732812.
24. Chengzhong H, Pengyan C, Xiutao Z, Anan D, Dexin W, Peiyao X. A survey of research progress on gene resources of Populus yunnanensis. J Southwest Forestry Univ. 2010;35(1):12–7.
25. Wang Z, Zhang W, Bai G, Lu Q, Li X, Zhou Y, Yang C, Xiao Y, Lang M. Highly resilient and fatigue-resistant poly(4-methyl-ε-caprolactone) porous scaffold fabricated via thiol-yne photo-crosslinking/salt-templating for soft tissue regeneration. Bioactive materials. 2023;28:311–25.
26. Laskowska E, Wawrzynów A, Taylor A. IbpA and IbpB, the new heat-shock proteins, bind to endogenous *Escherichia coli* proteins aggregated intracellularly by heat shock. Biochimie. 1996;78(2):117–22.
27. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296-w303.
28. 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.
29. Jiang Y, Peng X, Zhang Q, Liu Y, Li A, Cheng B, Wu J. Regulation of drought and salt tolerance by OsSKL2 and OsASR1 in rice. Rice (New York, NY). 2022;15(1):46.
30. Kong SG, Yamazaki Y, Shimada A, Kijima ST, Hirose K, Katoh K, Ahn J, Song HG, Han JW, Higa T, et al. CHLOROPLAST UNUSUAL POSITIONING 1 is a plant-specific actin polymerization factor regulating chloroplast movement. Plant Cell. 2024;36(4):1159–81.
31. Sun W, Van Montagu M, Verbruggen N. Small heat shock proteins and stress tolerance in plants. Biochem Biophys Acta. 2002;1577(1):1–9.
32. Siddique M, Gernhard S, von Koskull-Döring P, Vierling E, Scharf KD. The plant sHSP superfamily: five new members in *Arabidopsis thaliana* with unexpected properties. Cell Stress Chaperones. 2008;13(2):183–97.
33. Yang M, Zhang Y, Zhang H, Wang H, Wei T, Che S, Zhang L, Hu B, Long H, Song W, et al. Identification of MsHsp20 gene family in Malus sieversii and functional characterization of MsHsp16.9 in heat tolerance. Front Plant Sci. 2017;8: 1761.
34. Golan Y, Berman B, Assaraf YG. Heterodimerization, altered subcellular localization, and function of multiple zinc transporters in viable cells using bimolecular fluorescence complementation. J Biol Chem. 2015;290(14):9050–63.
35. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. Nucleic Acids Res. 2015;43(W1):W39-49.
36. Poverennaya IV, Roytberg MA. Spliceosomal introns: features, functions, and Evolution. Biochemistry Biokhimiia. 2020;85(7):725–34.
37. Yan J, Ma Z, Xu X, Guo AY. Evolution, functional divergence and conserved exon-intron structure of bHLH/PAS gene family. Molecular genetics and genomics : MGG. 2014;289(1):25–36.
38. Kuczynska-Wisnik D, Kçdzierska S, Matuszewska E, Lund P, Taylor A, Lipinska B, Laskowska E. The Escherichia coli small heat-shock proteins IbpA and IbpB prevent the aggregation of endogenous proteins denatured in vivo during extreme heat shock. Microbiology (Reading, England). 2002;148(Pt 6):1757–65.
39. Rosier Olimpio Pereira I, Saes Parra Abdalla D. Soy isoflavones reduce heat shock proteins in experimental atherosclerosis. Eur J Nutr. 2006;45(3):178–86.
40. Lian X, Wang Q, Li T, Gao H, Li H, Zheng X, Wang X, Zhang H, Cheng J, Wang W, et al. Phylogenetic and transcriptional analyses of the *HSP20* gene family in peach revealed that PpHSP20-32 is involved in plant height and heat tolerance. Int J Mol Sci. 2022;23(18):10849.

41. Nomoto A, Nishinami S, Shiraki K. Affinity of aromatic amino acid side chains in amino acid solvents. Biophys Chem. 2022;287:106831.
42. Kriehuber T, Rattei T, Weinmaier T, Bepperling A, Haslbeck M, Buchner J. Independent evolution of the core domain and its flanking sequences in small heat shock proteins. FASEB J. 2010;24(10):3633–42.
43. Poulain P, Gelly JC, Flatters D. Detection and architecture of small heat shock protein monomers. PLoS ONE. 2010;5(4): e9990.
44. Kong H, Landherr LL, Frohlich MW, Leebens-Mack J, Ma H, dePamphilis CW. Patterns of gene duplication in the plant *SKP1* gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth. The Plant journal : for cell and molecular biology. 2007;50(5):873–85.
45. Cui Y, Cao Q, Li Y, He M, Liu X. Advances in cis-element- and natural variation-mediated transcriptional regulation and applications in gene editing of major crops. J Exp Bot. 2023;74(18):5441–57.
46. Sun X, Sun C, Li Z, Hu Q, Han L, Luo H. AsHSP17, a creeping bentgrass small heat shock protein modulates plant photosynthesis and ABA-dependent and independent signalling to attenuate plant response to abiotic stress. Plant, Cell Environ. 2016;39(6):1320–37.
47. Nie G, Zhou J, Jiang Y, He J, Wang Y, Liao Z, Appiah C, Li D, Feng G, Huang L, et al. Transcriptome characterization of candidate genes for heat tolerance in perennial ryegrass after exogenous methyl Jasmonate application. BMC Plant Biol. 2022;22(1):68.
48. Sang Y, Long Z, Dan X, Feng J, Shi T, Jia C, Zhang X, Lai Q, Yang G, Zhang H, et al. Genomic insights into local adaptation and future climate-induced vulnerability of a keystone forest tree in East Asia. Nat Commun. 2022;13(1):6541.
49. Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, et al. The conserved domain database in 2023. Nucleic Acids Res. 2023;51(D1):D384-d388.
50. Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. Nucleic Acids Res. 2021;49(D1):D458-d460.
51. Liu D, Zheng K, Wang Y, Zhang Y, Lao R, Qin Z, Li T, Zhao Z. Harnessing an arbuscular mycorrhizal fungus to improve the adaptability of a facultative metallophytic poplar (*Populus yunnanensis*) to cadmium stress: physiological and molecular responses. J Hazard Mater. 2022;424(Pt B):127430.
52. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38(7):3022–7.
53. Chen C, Wu Y, Li J, Wang X, Zeng Z, Xu J, Liu Y, Feng J, Chen H, He Y, et al. TBtools-II: a "one for all, all for one" bioinformatics platform for biological big-data mining. Mol Plant. 2023;16(11):1733–42.
54. 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.

Publisher's Note

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