Genome-Wide Analysis of C2H2 Zinc-Finger Family Transcription Factors and Their Responses to Abiotic Stresses in Poplar (*Populus trichocarpa*)

Quangang Liu[1], Zhanchao Wang[1], Xuemei Xu[2], Haizhen Zhang[1], Chenghao Li[1]*

1 State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, Heilongjiang, People's Republic of China, **2** Library of Northeast Forestry University, Harbin, Heilongjiang, People's Republic of China

* chli0@163.com

Abstract

Background

C2H2 zinc-finger (C2H2-ZF) proteins are a large gene family in plants that participate in various aspects of normal plant growth and development, as well as in biotic and abiotic stress responses. To date, no overall analysis incorporating evolutionary history and expression profiling of the C2H2-ZF gene family in model tree species poplar (*Populus trichocarpa*) has been reported.

Principal Findings

Here, we identified 109 full-length C2H2-ZF genes in *P. trichocarpa*, and classified them into four groups, based on phylogenetic analysis. The 109 C2H2-ZF genes were distributed unequally on 19 *P. trichocarpa* linkage groups (LGs), with 39 segmental duplication events, indicating that segmental duplication has been important in the expansion of the C2H2-ZF gene family. Promoter *cis*-element analysis indicated that most of the C2H2-ZF genes contain phytohormone or abiotic stress-related *cis*-elements. The expression patterns of C2H2-ZF genes, based on heatmap analysis, suggested that C2H2-ZF genes are involved in tissue and organ development, especially root and floral development. Expression analysis based on quantitative real-time reverse transcription polymerase chain reaction indicated that C2H2-ZF genes are significantly involved in drought, heat and salt response, possibly via different mechanisms.

Conclusions

This study provides a thorough overview of the *P. trichocarpa* C2H2-ZF gene family and presents a new perspective on the evolution of this gene family. In particular, some C2H2-ZF genes may be involved in environmental stress tolerance regulation. *PtrZFP2*, *19* and *95* showed high expression levels in leaves and/or roots under environmental stresses.

Additionally, this study provided a solid foundation for studying the biological roles of C2H2-ZF genes in *Populus* growth and development. These results form the basis for further investigation of the roles of these candidate genes and for future genetic engineering and gene functional studies in *Populus*.

Introduction

Zinc-finger proteins (ZFPs) are the largest family of transcription regulators in plants [1]. The term "zinc finger" (ZF) refers to a protein domain, in which a zinc ion is surrounded by cysteine (Cys) residues and/or the histidine (His) residues to stabilize its three-dimensional structure, which comprises a two-stranded antiparallel beta sheet and a helix. Based on the number and location of the Cys and His residues, the ZFPs are divided into different types: C2H2, C2HC, C2HC5, C2C2, CCCH, C3HC4, C4, C4HC3, C6, and C8 [2,3]. The first plant-specific ZFP was identified in *Petunia* and interacts with the promoter region of the 5-enolpyruvylshikimate-3-phosphate synthase gene (EPSPS) [4]. Subsequently, further zinc-finger transcription factors (TFs) have been identified in other plants and their contributions to important biological processes during vegetative growth, reproductive development and stress responses have been noted [5–8].

The C2H2 ZFPs, also called the TFIIIA-type ZFPs, can be represented as X2-Cys-X(2–4)-Cys-X12-His-X(3–5)-His, where X represents any amino acid and where the precise spacing varies between the two cysteines and between the two histidines [9]. Among the ZFP types, C2H2-ZFPs are one of the most widespread transcription factor families in eukaryotes [10]. *In silico* analysis has shown that C2H2-ZF genes represent ~3% of all genes in mammals, ~2.3% in Diptera and ~0.8% in *Saccharomyces cerevisiae* [11,12]. In plants, the C2H2 zinc-finger gene family is large. There are 176, 189 and 124 members in *Arabidopsis*, rice and foxtail millet, respectively [13–15]. Compared with other eukaryote C2H2-ZFPs, there are two main structural features in most of the plant C2H2 ZFPs. First, in plant multiple-fingered C2H2-ZFPs, the zinc-finger domains are separated by long spacers that vary in length and sequence, whereas the C2H2-ZFPs of yeast and animals are mostly clustered and separated by a short spacer (six to eight amino acids) [16,17]. Second, most plant C2H2-ZFPs include an invariant sequence, QALGGH, located inside the zinc-finger helices, whereas yeast and animal C2H2-ZFPs do not have this motif [1].

C2H2-ZFPs differ widely in their structures and functions. Many ZFPs function as TFs, and although these proteins are thought to mainly interact with DNA or chromatin, some are also thought to bind to RNA and proteins [18–20]. Therefore C2H2-ZFPs not only participate in transcriptional regulation, but also act in RNA metabolism and in other biological activities. They are involved in a wide range of processes, including plant growth and development, and in biotic and abiotic stress responses in plants. A number of C2H2-ZFPs are involved in trichome initiation, seed germination, floral organogenesis, primary microRNA biogenesis and heat stress in *Arabidopsis* [21–25]; cold, drought, oxidative and salt stress in rice [26,27]; cold and drought stress in soybean [28]; pathogen defense in *Capsicum annuum* [29]; and osmotic, cold and mechanical stress in poplar [30]. These studies showed that C2H2-ZFPs are likely to be associated with multiple physiological processes and stress responses.

Poplar (*Populus trichocarpa*) is a valuable forest resource, making it an important commercial and ecological species [31]. *P. trichocarpa* is also a model plant, whose whole genome sequence is available [32]. Compared with the extensive studies of C2H2-ZF genes in many

other plant species, little research has been conducted in *Populus* so far. Therefore, the genome-wide identification and expression analysis of the C2H2-ZF gene family is important. Here, we report a systematic study of this gene family. We identified 109 C2H2-ZF genes and analyzed their phylogenetic relationships, chromosomal locations, gene structures, conserved protein motifs and promoter *cis*-elements. The expression profiles of this gene family in different tissues and organs of *Populus*, and under biotic and abiotic stress conditions, were analyzed using data from heatmap and quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR) analyses. Our results provided a subset of potential candidate genes that may be used for future genetic engineering and gene functional studies in *Populus*.

Materials and Methods

Ethics Statement

No specific permissions were required for the described field studies. Our university, the Northeast Forestry University in Harbin, is not privately-owned. The field studies did not involve endangered or protected species.

Identification and characteristics of the C2H2-ZF gene family

The genomic library, cDNA library and protein database of *P. trichocarpa* were downloaded from the Phytozome v10.0 (http://phytozome.jgi.doe.gov/pz/portal.html) and NCBI (http://www.ncbi.nlm.nih.gov/) databases. The hidden Markov model (HMM) profile of the C2H2-ZF gene family (protein family ID: PF00096) was downloaded from the Protein family database (Pfam 27.0, http://pfam.xfam.org/) [33]. All of the located sequences were further analyzed manually to confirm the presence of a C2H2-ZF domain (SM000355) using the SMART (http://smart.embl-heidelberg.de/) database [34]. Finally, the obtained genes were compared with the C2H2-ZF gene family in PlnTFDB v3.0 (http://plntfdb.bio.uni-potsdam.de/v3.0/) [35]. Closely related genes from *Arabidopsis* were obtained from the *Arabidopsis* Information Resource (TAIR, http://www.arabidopsis.org/index.jsp). WoLF PSORT (http://wolfpsort.org/) [36] was used to predict the subcellular localization of C2H2-ZFPs. The ExPasy site (http://web.expasy.org/protparam/) [37] was used to calculate the molecular weight and isoelectric point (pI) of the deduced polypeptides.

Phylogenetic analysis

Multiple sequence alignment of the full-length protein sequences was performed using Clustal X (version 1.83) [38] and adjusted manually using BioEdit 7.1 software [39]. Phylogenetic analyses using the neighbor-joining method in MEGA 5.0 [40] and a bootstrap test carried out with 1000 iterations were performed to construct unrooted phylogenetic trees.

Exon/intron structure analysis and identification of conserved motifs

The Gene structure display server (GSDS 2.0, http://gsds.cbi.pku.edu.cn/index.php) [41] was used to generate the exon/intron organization. The Multiple Expectation Maximization for Motif Elucidation (MEME) system (Version 4.9.1, http://meme.nbcr.net/meme/) was used to identify conserved motifs for each C2H2-ZF gene [42]. Structural motif annotation was performed using the Pfam and SMART tools.

Chromosomal location

Genes were mapped on chromosomes by identifying their chromosomal locations provided in the PopGenIE v3 database (http://www.popgenie.org/) [43]. Physical map was constructed

using Adobe Illustrator CS5 (Adobe Systems Incorporated). Blocks of the same color represent the homologous chromosome fragments. Genes were considered to have undergone segmental duplication if they were located on duplicated chromosomal segments [44]. Meanwhile, to search for potential duplicated *Populus* C2H2-ZF genes, MCScanX software (http://chibba. pgml.uga.edu/mcscan2/) was used [45], apart from them falling into duplicated blocks. Genes separated by five or fewer gene loci within a physical distance of 100 kb were considered tandem duplicates. A schematic view of the reorganization of homologous chromosomes segments was based on the most recent account of whole-genome duplication in *P. trichocarpa* [32]. Synonymous (Ks) and nonsynonymous substitution (Ka) rates were calculated according to previous study [46].

Promoter *cis*-element analysis

Promoter sequences (2 kb upstream of the translation start site) of all C2H2-ZF genes were obtained from the Phytozome v10.0 database. PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [47] was used to analyze the sequence of the C2H2-ZF gene promoters, and to predict and locate their *cis*-elements.

Gene Ontology (GO) Annotation

The functional grouping of C2H2-ZFP sequences and the analysis of annotation data were executed using Blast2GO v3.0 [48]. Blast2GO annotation associates genes or transcripts with GO terms using hierarchical vocabularies. Genes are described in terms related to three categories of GO classification: biological processes, molecular functions and cellular components.

exNorthern and exHeatMap analysis

The expression profile of each gene was obtained by evaluating its expressed sequence tag (EST) representation among 17 cDNA libraries derived from different tissues and/or developmental stages using the exNorthern tool of the PopGenIE v2 database (http://www.popgenie.org/). The 17 libraries were derived from several taxa of *Populus* [49]. The exHeatMap tool at PopGenIE v2 was used to visualize the heatmap of C2H2-ZF genes under different stress conditions. The heatmap data can be directly downloaded using the accession numbers of genes via the exHeatmap tool in the PopGenIE v2 database.

Plant materials and stress treatments

Clonally propagated *P. trichocarpa* (genotype Nisqually-1) were cultured in half-strength Murashige and Skoog medium under long-day conditions (16 h light/8 h dark) at 23–25°C. Stress treatment conditions and sampling timings were conducted following the previous method with minor modification [50,51]. These clones were exposed to 150 mM mannitol, 200 mM sodium chloride (NaCl) and 42°C for drought, salinity and heat stress treatments, respectively. Young leaves and roots from three different plants were collected at different time points (0, 3, 6, 12 and 24 h) after treatment. Three biological replicates were performed for each stress treatment. Each experiment was repeated at least three times with independent sample preparation to obtain reproducible results. All samples were immediately frozen in liquid nitrogen and stored at −80°C until analysis.

RNA isolation and qRT-PCR verification

The cetyltrimethylammonium bromide method, with minor modifications [52], was used to extract total RNA from roots and leaves. The first-strand cDNA synthesis was carried out with

the ReverTra Ace qPCR RT Master Mix with a gDNA Remover Kit (TOYOBO, Osaka, Japan), in accordance with the manufacturer's instructions. Primer Premier 5 was used to design primers with melting temperatures of 55–60°C, primer lengths of 18–25 bp, and amplicon lengths of 101–221 bp. Details of the primers are shown in S1 Table. SYBR Premix Ex Taq II (TaKaRa, Dalian, China) was used to perform qRT-PCR in accordance with the manufacturer's instructions. Reactions were prepared in a total volume of 20 μL, containing 10 μL of 2×SYBR Premix, 2 μL of cDNA template, 6 μL of ddH$_2$O, and 1 μL of each primer to make a final concentration of 10 μM. The *P. trichocarpa* actin gene (GenBank ID: XM_002298674) [53] was used as a reference gene. The PCR conditions and relative gene expression calculations were as previously described [54].

Results

Identification of C2H2-ZF genes in *Populus*

The HMM profile of the Pfam C2H2-ZF domain (protein family ID: PF00096) was used as the query to identify C2H2-ZF genes in the *Populus trichocarpa* genome. One hundred and twenty-one candidate C2H2-ZF genes were identified. All the C2H2-ZF candidates were analyzed manually using the SMART (SM000355) database to verify the presence of the C2H2-ZF domain. Finally, 109 C2H2-ZF genes were identified. Our result was roughly in agreement with PlnTFDB, where 110 members of the C2H2-ZF gene family were deposited for *Populus*. This number is less than the number present in the *Arabidopsis*, rice and foxtail millet genomes (176, 189 and 124, respectively) [13–15]. The C2H2-ZFPs have previously been named as AT-ZFP in *Arabidopsis thaliana*, hence we have named *Populus* proteins as PtrZFP (C2H2-ZFPs of *Populus trichocarpa*) [13].

The identified *P. trichocarpa* C2H2-ZF genes had molecular masses ranging from 17626.35 to 191195.46 Da. The encoded proteins varied from 161 to 1685 amino acids (aa), with an average of 404 aa. The pI values of the predicted proteins were varied; for example, PtrZFP59 had a pI of 4.75, whereas that of PtrZFP47 was 9.91. WoLF PSORT was used to predict the location of the predicted proteins in the plant cell. One hundred and three C2H2-ZFPs were predicted as nuclear proteins, four as chloroplast proteins and two as cytoplasmic proteins. SMART database was used to comfirm the number of C2H2 motifs. Based on the number of C2H2 motifs, the C2H2-ZF genes could be classified into two groups: a motif-rich group (PtrZFP16, 18, 67 and 70, each with 4–6 motifs) and a motif-poor group (for all the other proteins, the number of motifs varied from 1 to 3). The details of other characteristics of the nucleic acid and protein sequences are provided in S2 Table.

Of the identified 109 C2H2-ZF genes, 28 genes showed more than one gene model, which could be attributed to alternative splicing, in the Phytozome v10.0 database. These were *PtrZFP10, 11, 12, 13, 14, 16, 18, 26, 27, 31, 36, 42, 54, 57, 58, 59, 61, 72, 73, 76, 82, 86, 95, 96, 103, 105, 107* and *109. PtrZFP31, 76* and *82* have the largest number (six) of transcripts. *PtrZFP16, 26, 27, 54, 86* and *105* encode 3 to 5 transcripts. Other genes only encode two cDNAs. In the case of *PtrZFP82*, out of the six cDNAs, Potri.005G164100.6 does not have a ZF motif because of alternative splicing. Out of the five cDNAs available for *PtrZFP54*, Potri.006G116500.2 and Potri.006G116500.5 do not encode a ZF motif. *PtrZFP105* also encodes five alternatively spliced cDNAs, but Potri.001G160700.3 and Potri.001G160700.5 do not have any ZF motifs. In the case of *PtrZFP86*, Potri.006G227600.1, Potri.006G227600.2 and Potri.006G227600.3 encode two, three and three fingers, respectively. Of the three transcripts representing *PtrZFP26*, Potri.012G143200.3 does not have any ZF motifs (S3 Table).

Phylogenetic analysis, gene structure and conserved motifs of C2H2-ZF genes in *Populus*

To examine the phylogenetic relationships among the C2H2-ZF domain proteins in *Populus*, an unrooted phylogenetic tree was constructed from alignments of the full-length C2H2-ZFP sequences (Fig 1A). The 109 C2H2-ZF genes were classified into four groups (I, II, III and IV) containing 38, 28, 17, and 26 members, respectively. Almost 8000 pairs of paralogous genes are present in the *Populus* genome [55]. Based on the phylogenetic analysis, we identified 40 sister pairs, all of which had strong bootstrap support (>96%) (S4 Table).

Gene structural diversity and conserved motif divergence were a possible mechanism for the evolution of multigene families [44]. To gain further insights into the structural diversity of *Populus* C2H2-ZF genes, we analyzed the exon/intron organization in the full-length cDNAs with their corresponding genomic DNA sequences of individual C2H2-ZF genes in *Populus* (Fig 1B). Most closely related C2H2-ZF members within the same subgroups shared similar gene structures in terms of either intron numbers or exon lengths. For example, the C2H2-ZF group I and II genes had zero to three introns with exception of *PtrZFP82*. By contrast, the gene structure appeared to be more variable in subgroups III and IV, which had the largest number of exon/intron structure variants, with striking distinctions.

Putative protein motifs were predicted using MEME to further reveal the diversification of C2H2-ZF genes in *Populus* (Fig 1C). Fifteen distinct motifs were identified (S5 Table). Motifs 1, 3, 6, 14 and 15 represented different types of C2H2-ZF. These different types were found in all 109 *Populus* C2H2-ZFPs. Besides the C2H2-ZF domain, a known functional domain (Motif 9) was identified in the C2H2-ZFPs. Sixteen members contained the hexapeptide motif 'DLELRL' (Motif 9) in the C-terminal region. The hexapeptide motif acts as a repression domain and has been studied in *Arabidopsis* [56]. Furthermore, nine unidentified conserved motifs were observed. Only eight proteins (PtrZFP1, 4, 5, 6, 19, 22, 69 and 98) contained motif 11.

Different types of C2H2-ZFs have been identified in rice, *Arabidopsis* and petunia [10,13, 14]. The first type includes a plant-specific conserved domain 'QALGGH', which is classified as the Q-type. The 62 identified Q-type C2H2-ZFPs in *Populus* were fewer than those present in the foxtail millet (97) [15], more than in bread wheat (47) [57] and similar to those in *Arabidopsis* (64) [13] and rice (65) [14]. Certain modifications of the Q-type ZF were designated as the M-type. The C-type does not include any conserved motif in the ZF region compared with the Q-type and M-type. Some of the ZFs contained highly conserved motifs in the C-type ZF and the flanking regions, and were named Z-type [14]. The details of the C2H2-ZF types are shown in S6 Table. *Populus* C2H2-ZFPs contained 1 to 6 ZF domains. Seventy-six proteins possess a single C2H2-ZF domain, which included three types of C2H2-ZF (Q-, Z- and C-types). Among them, 52 proteins contained a Q-type ZF, followed by 19 proteins with the C-type ZF and five proteins with the Z-type ZF. A significantly reduced number (23) contained two ZFs, among which 1 protein contained two C-type ZFs and five proteins contained two Q-type ZFs. For the others, two proteins contained a Q- and M-type ZF combination and 15 contained a C- and Z-type combination. There were six proteins were three-fingered proteins, of which two proteins contained all C-type ZFs. Only one protein had four ZFs, which was an M- and C-type combination. Two five-fingered proteins comprised C- and Z-type ZFs, but the number of C- and Z-type was different. The single 6-fingered protein contained all C-type ZFs (S7 Table).
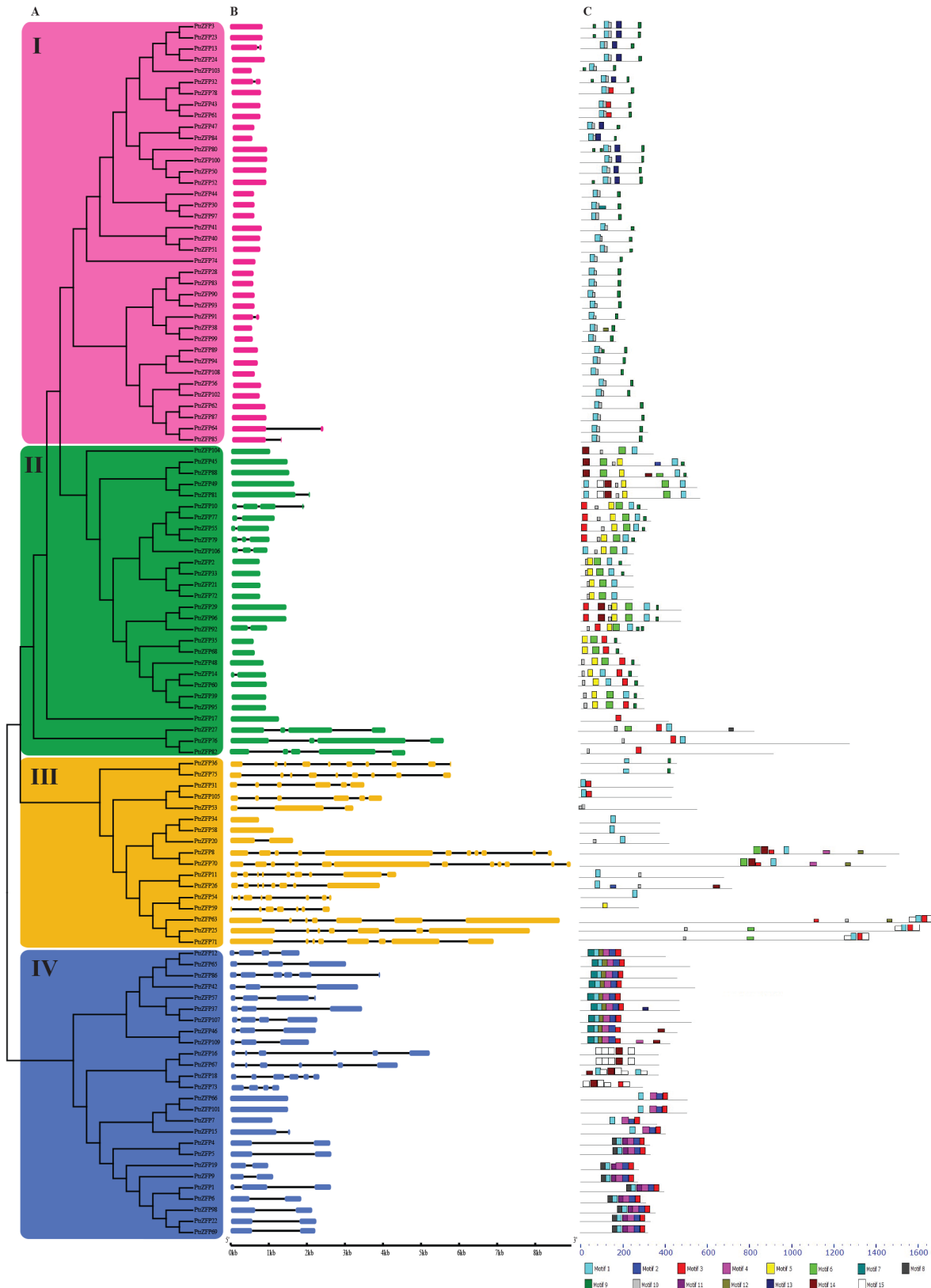
Fig 1. Phylogenetic relationships, gene structure and motif compositions of *Populus* C2HC2-ZF genes. A. Multiple alignment of 109 full-length amino acid sequences of *Populus* C2HC2-ZF genes, executed by ClustalX 1.83. The phylogenetic tree was constructed using MEGA 5.0 and the neighbor-joining

method. Support values from a bootstrap analysis with 1,000 replicates are specified at each node. The four major phylogenetic subgroups, designated as I to IV, are marked with different colored backgrounds. **B.** Exon/intron structures of *Populus* C2HC2-ZF genes. Exon/intron structures were obtained from the Gene Structure Display Server. Exons and introns of each subgroup are represented by particular colored boxes and black lines, respectively. **C.** Schematic representation of the conserved motifs identified by MEME. Each colored box represents a motif and black lines represent non-conserved sequences.

Chromosomal locations and duplications of *Populus* C2H2-ZF genes

In silico mapping of the gene loci showed that the 109 *Populus* C2H2-ZF genes were distributed across all 19 linkage groups (LGs). All of the *Populus* C2H2-ZF genes were distributed across the LGs. However, the distribution of *Populus* C2H2-ZF genes across the LGs was uneven. LG I has the largest number (12) of genes, followed by 11 on LG X. In contrast, only one or two genes were found on each of LG VII, LG XI, LG XIII, LG XV and LG XIX ([Fig 2](#)).

Previous studies revealed that the *Populus* genome has undergone at least three rounds of genome-wide duplication, followed by multiple segmental duplications, tandem duplications and transposition events, such as retrotransposition and replicative transposition [32]. In particular, the segmental duplication associated with the salicoid duplication event that occurred 65 million years ago contributed significantly to the expansion of many multigene families [54,58,59]. To determine the possible relationship between the C2H2-ZF genes and potential segmental duplications, we mapped the *Populus* C2H2-ZF genes to the duplicated blocks established in previous studies [32]. The distribution of C2H2-ZF genes relative to the corresponding duplicate blocks is shown in [Fig 2](#). About 79% (86 of 109) of *Populus* C2H2-ZF genes were preferentially retained duplicates located in both duplicated regions. Eleven duplicated blocks
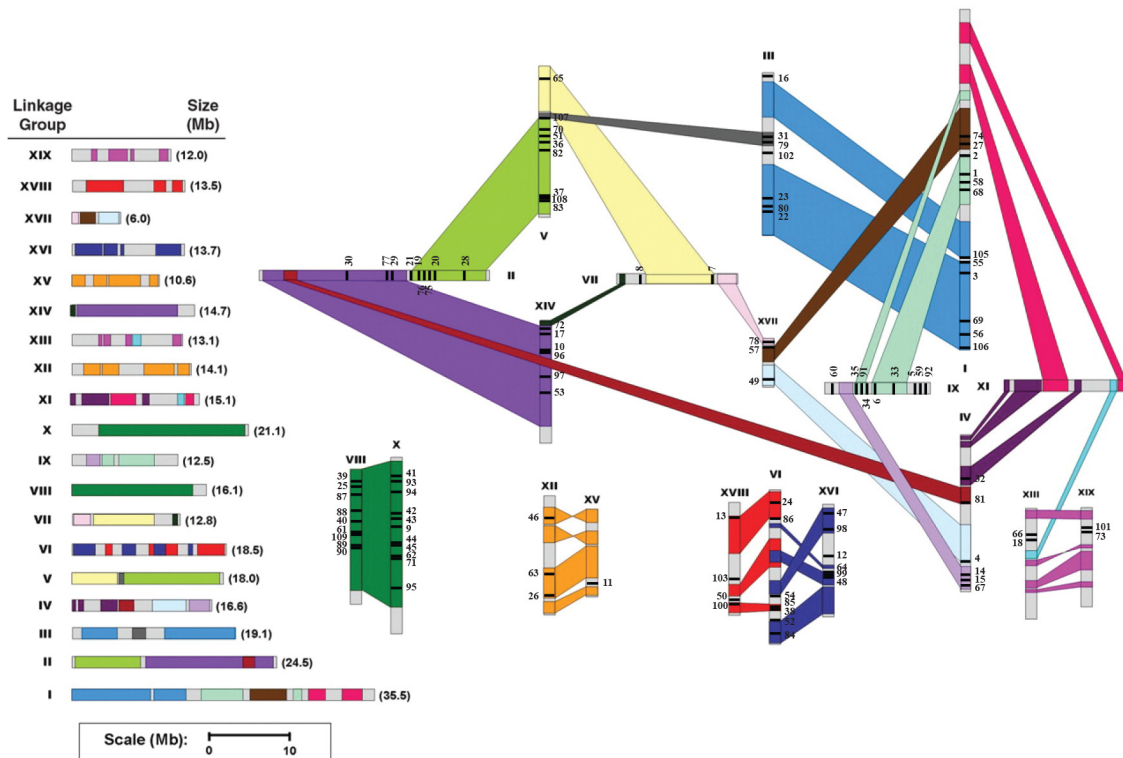

Fig 2. Chromosomal locations of *Populus* C2HC2-ZF genes. One hundred and nine C2HC2-ZF genes were mapped to 19 linkage groups (LG). A schematic view of chromosome reorganization caused by recent whole-genome duplication in *Populus* is shown (adapted from [32]). Segmental duplicated homologous blocks are indicated by the same color. The scale represents mega bases (Mb). The LG numbers are indicated above each bar.

only contained C2H2-ZF genes in one of the blocks and lacked duplicates in the corresponding block. By contrast, 23 C2H2-ZF genes were located outside any duplicated blocks, suggesting that dynamic changes may have occurred after segmental duplication, leading to the loss of some genes.

Comparative analysis of the C2H2-ZF genes in *Populus* and rice

To investigate the evolutionary relationship of the C2H2-ZF gene families, the full-length amino acid sequences of the 109 proteins from *P. trichocarpa* and 189 from rice (*Oryza sativa*) [14] were used to construct a phylogenetic tree. In rice, three plant-specific clusters including distinct types of C2H2-ZF have been identified compared with yeast and *Arabidopsis* [14]. In this study, we also observed the three clusters, named group a, b and c, comprising 16, 5 and 38 proteins, respectively (Fig 3). With few exceptions, group a proteins had a single C2H2-ZF domain, which contained Z2- to Z5-type ZFs; group b members possessed two consecutive Q-
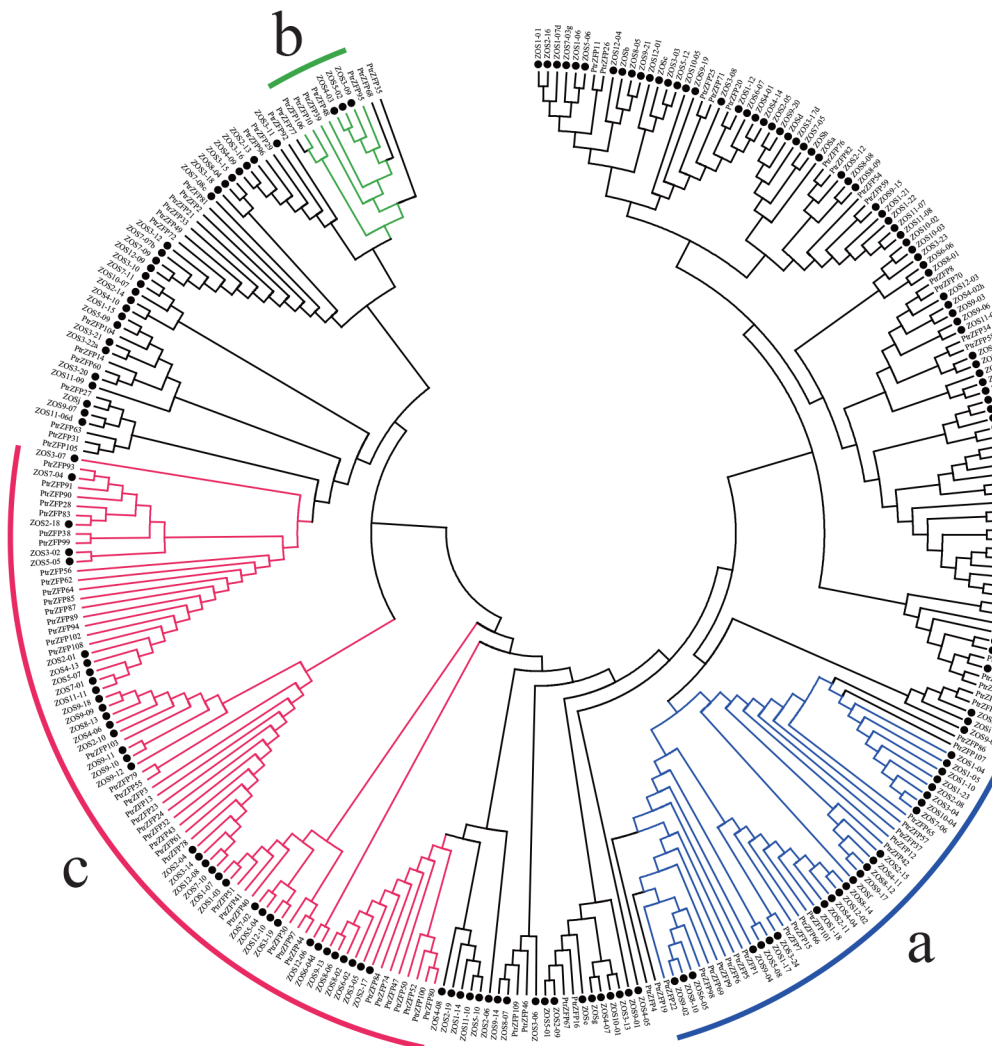

Fig 3. Phylogenetic tree of C2H2-ZF proteins from *Populus* and rice. One hundred and nine C2H2-ZF proteins in *Populus* and 189 in rice were aligned using Clustal X 1.83 and the phylogenetic tree was constructed using MEGA 5.0 by the neighbor-joining method. The Bootstrap value was 1,000 replicates. The rice proteins start with Z and are marked with a solid dot to distinguish them from the *Populus* C2HC2-ZF proteins. The three plant-specific clusters were designated as a, b and c, and indicated in a specific color.

doi:10.1371/journal.pone.0134753.g003

type ZFs, whereas the single Q-type ZF proteins were clustered in group c. The C2H2-ZFPs in groups a, b and c may be plant-specific. The other protein sequences, which mainly contained C- and Z1-type ZFs, appeared in plants and other species. According to the phylogenetic analysis of *Populus*, group a, b and c were also classified to group IV, II and I, respectively (Figs 1A and 3). It is noteworthy that some group a proteins from *Populus* not only had a Z2- to Z5-type ZF, but also contained a C-type ZF. Additionally, ZOS1-15/ZOS5-09/PtrZFP104, ZOS3-20/PtrZFP14/PtrZFP60 and ZOS9-07/ZOS11-06d/ZOSj/PtrZFP27 formed three small, unique clusters in the phylogenetic analysis. PtrZFP55 and 79, which contained two types of ZF (Q- and M1-type) were also clustered in group c (Fig 3).

Promoter *cis*-element analysis

Cis-elements play key roles in the transcriptional regulation of genes that control abiotic stress responses, such as drought and heat stress [60]. Meanwhile, phytohormones, such as salicylic acid (SA), jasmonic acid (JA), ethylene (ET), and abscisic acid (ABA), are essential for plants' ability to adapt to abiotic stresses, by inducing the interaction between TFs and corresponding *cis*-elements [61,62]. To identify putative *cis*-acting regulatory DNA elements in the C2H2-ZF genes, their promoter sequences (2 kb upstream of the translation start site) were obtained from the Phytozome v10.0 database, and the *cis*-elements of these 109 promoters were examined using the PlantCARE database. The C2H2-ZF gene family promoters harbored multiple *cis*-elements related to phytohormone and environmental stress signal responsiveness, such as MBS (MYB binding site, involved in drought-inducibility), HSE (heat stress-responsive element), ABRE (abscisic acid-responsive element), W-Box (WRKY binding site, involved in abiotic stress responsiveness), ERE (ethylene-responsive element) and TCA-element (salicylic acid-responsive element) (S8 Table). The majority of C2H2-ZF genes contained *cis*-elements related to phytohormone and environmental stress signal responsiveness. *PtrZFP4* and *92* have nine *cis*-elements (ABRE, CGTCA-motif, ERE, G-Box, HSE, MBS, TCA-element, TGACG-motif and W-Box), whereas *PtrZFP17*, *40* and *94* have only two *cis*-elements (S9 Table).

Gene Ontology (GO) Annotation

The biological processes, molecular functions and cellular components of *Populus* C2H2-ZF genes were investigated based on the putative assignment of Gene Ontology (GO) terms using Blast2GO v3.0 (Fig 4, S10 Table). The results showed that the 109 C2H2-ZF genes putatively participated in diverse biological processes. Of the nine terms of biological processes defined by Blast2Go, most *Populus* C2H2-ZFPs were predicted to function in the metabolic process (~29%) and the cellular process (~27%), followed by cellular component organization or biogenesis (~9%) and the developmental process (~9%). Molecular function prediction indicated that all 109 C2H2-ZFPs were annotated as small molecule or/and ion binding (~82%), which is in accordance with the molecular role of C2H2-ZFP in DNA and metal ion binding. In addition, some C2H2-ZFPs were involved in transcription factor activity (~12%) and catalytic activity (~6%). In addition, cellular component prediction showed that 20 *Populus* C2H2-ZFPs were localized in the cell part (~48%) and organelle (~48%), respectively. Only two C2H2-ZFPs were membrane-enclosed lumen-localized (~4%) (Fig 4, S10 Table).

Expression profile of *Populus* C2H2-ZF genes in different tissues or organs

Publicly available expressed sequence tags (ESTs) are considered useful tools to study gene expression profiles using the Digital Northern tool [63]. We first carried out a preliminary analysis of C2H2-ZF gene expression across different tissues and organs by counting the
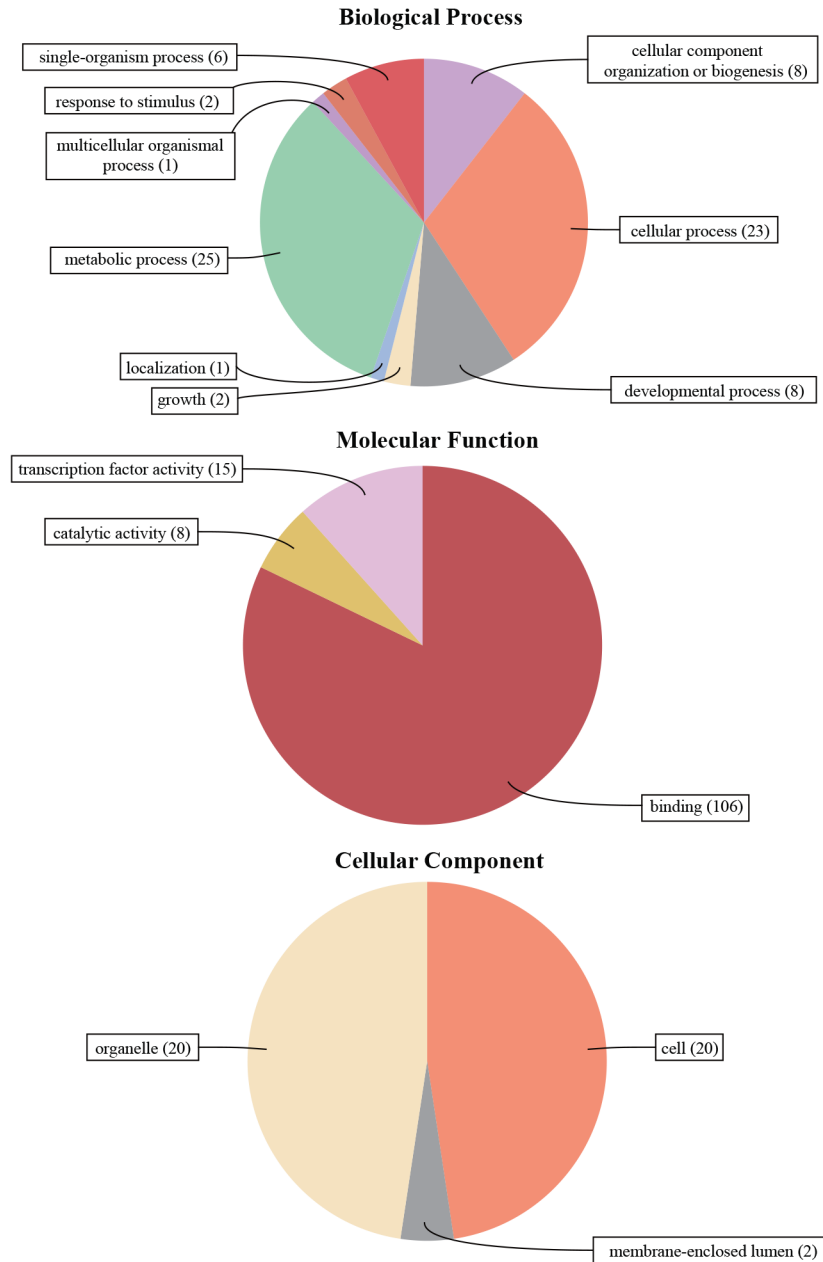
Fig 4. Gene Ontology (GO) results for *Populus* C2HC2-ZF proteins. GO analysis of 109 C2HC2-ZFP sequences predicted for their involvement in biological processes, molecular functions and cellular components. For the results presented as detailed bar diagrams, see S10 Table.

doi:10.1371/journal.pone.0134753.g004

frequencies of ESTs in different *Populus* cDNA libraries. A complete search of the digital expression profiles in PopGenIE yielded 33 *Populus* C2H2-ZF genes in the cDNA libraries. As illustrated in Fig 5, the C2H2-ZF gene family had a varied expression pattern across various developmental stages of tissues and organ. Among the 33 genes, one showed the highest transcript accumulation in young leaves, one in the shoot meristem, four in senescing leaves, two in bark, eight in roots, four in flower buds, four in apical shoots, four in dormant buds, three in
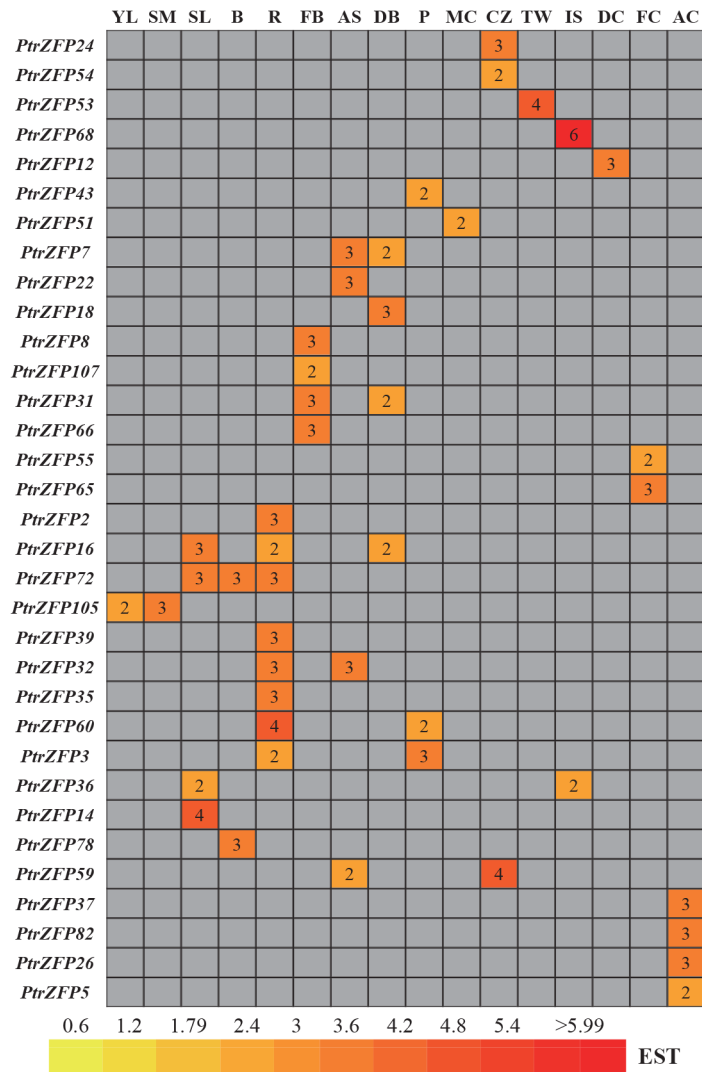
Fig 5. *In silico* EST analysis of *Populus* C2HC2-ZF genes. EST frequency for each gene was calculated by evaluating its EST representation among 16 cDNA libraries accessible using the exNorthern tool at the PopGenIE v2 database. The number in the box represents the frequencies of EST counts in different libraries. YL: young leaves, SM: shoot meristem, SL: senescing leaves, B: bark, R: roots, FB: flower buds, AS: apical shoot, DB: dormant buds, P: petioles, MC: male catkins, CZ: cambial zone, TW: tension wood, IS: imbibed seeds, DC: dormant cambium, FC: female catkins, AC: active cambium.

doi:10.1371/journal.pone.0134753.g005

petioles, one in male catkins, three in the cambial zone, one in tension wood, two in imbibed seeds, one in the dormant cambium, two in female catkins and four in the active cambium. A putative paralogous gene (*PtrZFP54/59*) had the highest transcript accumulation in the cambial zone. However, the expression patterns of a few gene pairs, including *PtrZFP14/60*, *PtrZFP31/105*, *PtrZFP35/68, and PtrZFP37/107*, were significantly different, although they were paralogous genes (Figs [1] and [5]). These results were similar to a previous study of the SPL gene family in *Populus* [64]. Moreover, *PtrZFP53* was highly expressed in tension wood. However, some genes (*PtrZFP8, 31, 51, 55, 65, 66,* and *107*) were mainly expressed during flower development.

Expression profiles of *Populus* C2H2-ZF genes under various stresses

Drought is the main environmental stress encountered by most land plants during their life span. To gain further insights into the potential roles of *Populus* C2H2-ZF genes in drought tolerance, we analyzed the expression profiles of *Populus* C2H2-ZF genes in response to drought stress using the publicly available heatmap data. The heatmap data can be directly downloaded using the accession numbers of genes via the exHeatmap tool in the PopGenIE v2 database (http://www.popgenie.org/). According to the present data set, 44 genes were upregulated in leaves and 20 genes were upregulated in roots under drought stress (Fig 6A). Thirteen genes were upregulated in leaves and roots. Thus, the site of activity of the gene was also different under drought stress. This result was consistent with previous analyses, which revealed that the C2H2-ZF gene family contributed to variation in the drought response [26,28]. In other biotic and abiotic stresses, beetle damage induced the upregulation of 29 C2H2-ZF genes and mechanical wounding induced the upregulation of 22 genes (Fig 6B). Notably, *PtrZFP2, 21, 28, 33, 35, 68*, and *72* were upregulated by drought, beetle and mechanical damage (Fig 6A and 6B).

Validation of C2H2-ZF gene expression under drought, heat and salt stress by qRT-PCR

To verify the expression profiles of *Populus* C2H2-ZF genes obtained by the heatmap analysis, qRT-PCR was performed for 51 selected genes (44 genes upregulated in leaves and 20 genes upregulated in roots, including 13 genes upregulated both in leaves and roots) under drought stress. The genes up- or downregulated by more than 1.5-fold were considered significantly differentially expressed. The results were broadly consistent with the heatmap data: 41 genes were induced, one gene (*PtrZFP71*) was suppressed and two genes (*PtrZFP23* and *40*) showed no change in leaves under drought stress (Fig 7). In roots, 16 genes were induced, one gene (*PtrZFP88*) was suppressed and three gene (*PtrZFP64, 95* and *96*) were not influenced by drought (Fig 8). Notably, *PtrZFP2, 11, 29, 33, 50, 68, 85* and *95* were upregulated significantly at all time points in leaves, whereas *PtrZFP33* and *72* were upregulated significantly at all time points in roots (Figs 7 and 8). In addition, *PtrZFP2, 15, 21, 28, 29, 33, 35, 50, 68* and *72* not only exhibited a high expression level in leaves, but were also upregulated in roots. The results were consistent with those observed in heatmap analysis (Figs 6A, 7 and 8); however, the levels of upregulation were slightly different.

Subsequently, the expression profiles of the 51 selected genes were further analyzed under heat and salt stress. Under heat stress treatment, 34 genes were induced, two genes (*PtrZFP29* and *33*) were suppressed and eight genes (*PtrZFP9, 26, 37, 50, 72, 81, 101* and *103*) were unchanged in leaves (Fig 7). Seventeen genes were induced, two genes (*PtrZFP21* and *72*) were suppressed and one gene (*PtrZFP68*) was not influenced in roots after heat treatment (Fig 8). In leaves, 10 genes (*PtrZFP2, 3, 15, 23, 35, 40, 51, 57, 62* and *71*) were upregulated significantly at all time points, whereas six genes (*PtrZFP29, 39, 49, 50, 88* and *96*) were upregulated significantly at all time points in roots. *PtrZFP2, 15, 28, 35, 64, 85* and *88* showed high expression levels in leaves and roots. *PtrZFP19* was strongly upregulated (>40-fold) at 24 h after heat treatment (Figs 7 and 8). Under salt stress, 30 genes were induced, seven genes (*PtrZFP9, 15, 44, 51, 71, 88* and *101*) were suppressed and seven genes (*PtrZFP7, 23, 38, 40, 85, 87* and *102*) were not affected in leaves (Fig 7). Five genes were induced, seven genes (*PtrZFP15, 28, 29, 35, 42, 68* and *85*) were suppressed and eight genes (*PtrZFP10, 33, 49, 60, 64, 72, 85* and *96*) were unchanged in roots after NaCl treatment (Fig 8). Among these genes, *PtrZFP2, 5, 21, 26, 29, 57, 68, 72* and *91* were upregulated significantly at all time points in leaves; no genes were upregulated significantly at all time points in roots. Notably, *PtrZFP2, 21* and *50* displayed significantly high expression levels in

Fig 6. Differential expression of *Populus* C2HC2-ZF genes under different stress conditions. The heatmap was visualized using the exHeatmap tool in the PopGenIE v2 database. Blue and red indicate low and high levels of transcript abundances, respectively. **A.** Heatmap showing 92 *Populus* C2HC2-ZF genes in leaves and roots under drought stress. **B.** Heatmap showing 92 *Populus* C2HC2-ZF genes under beetle and mechanical stress.
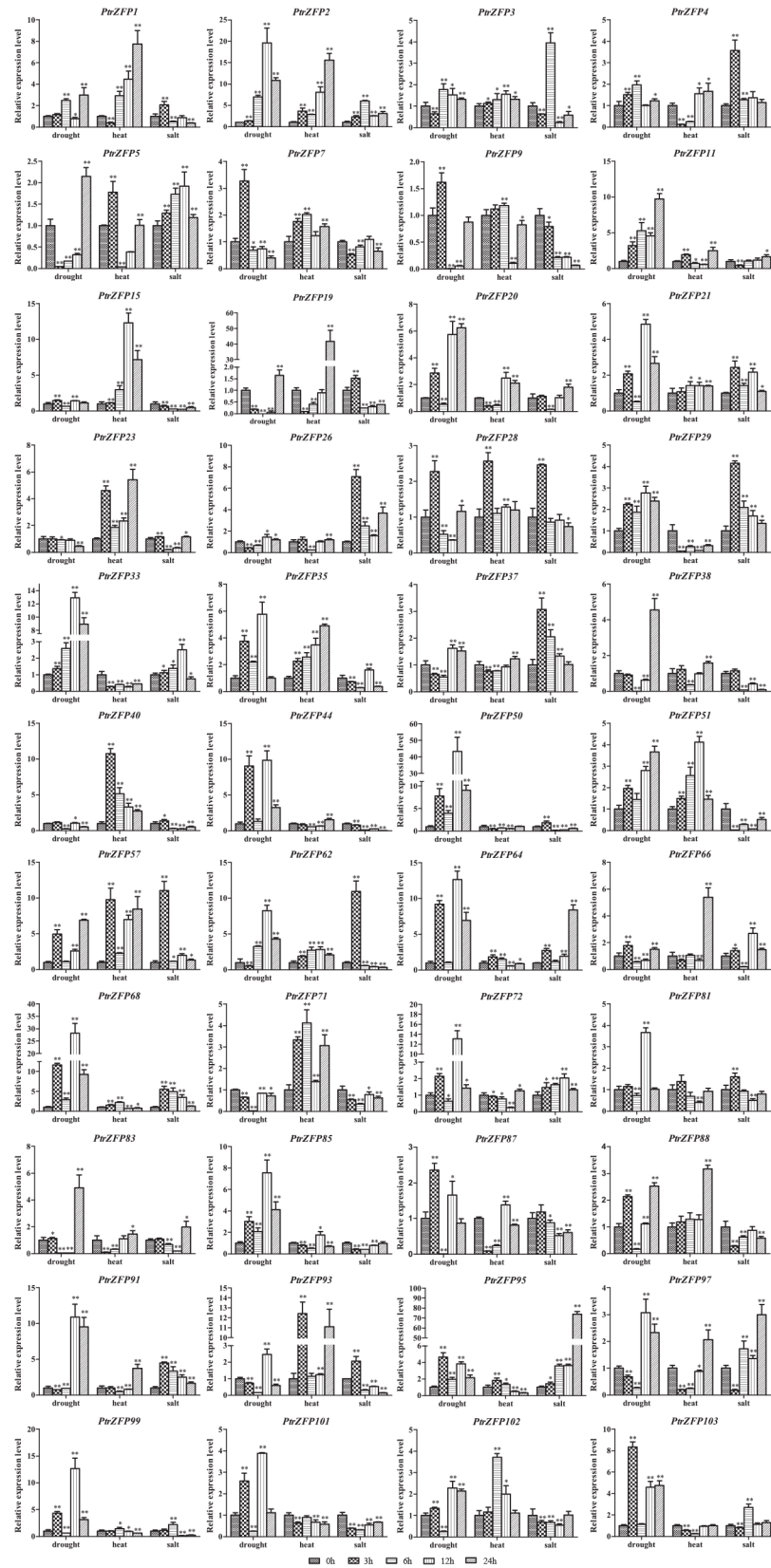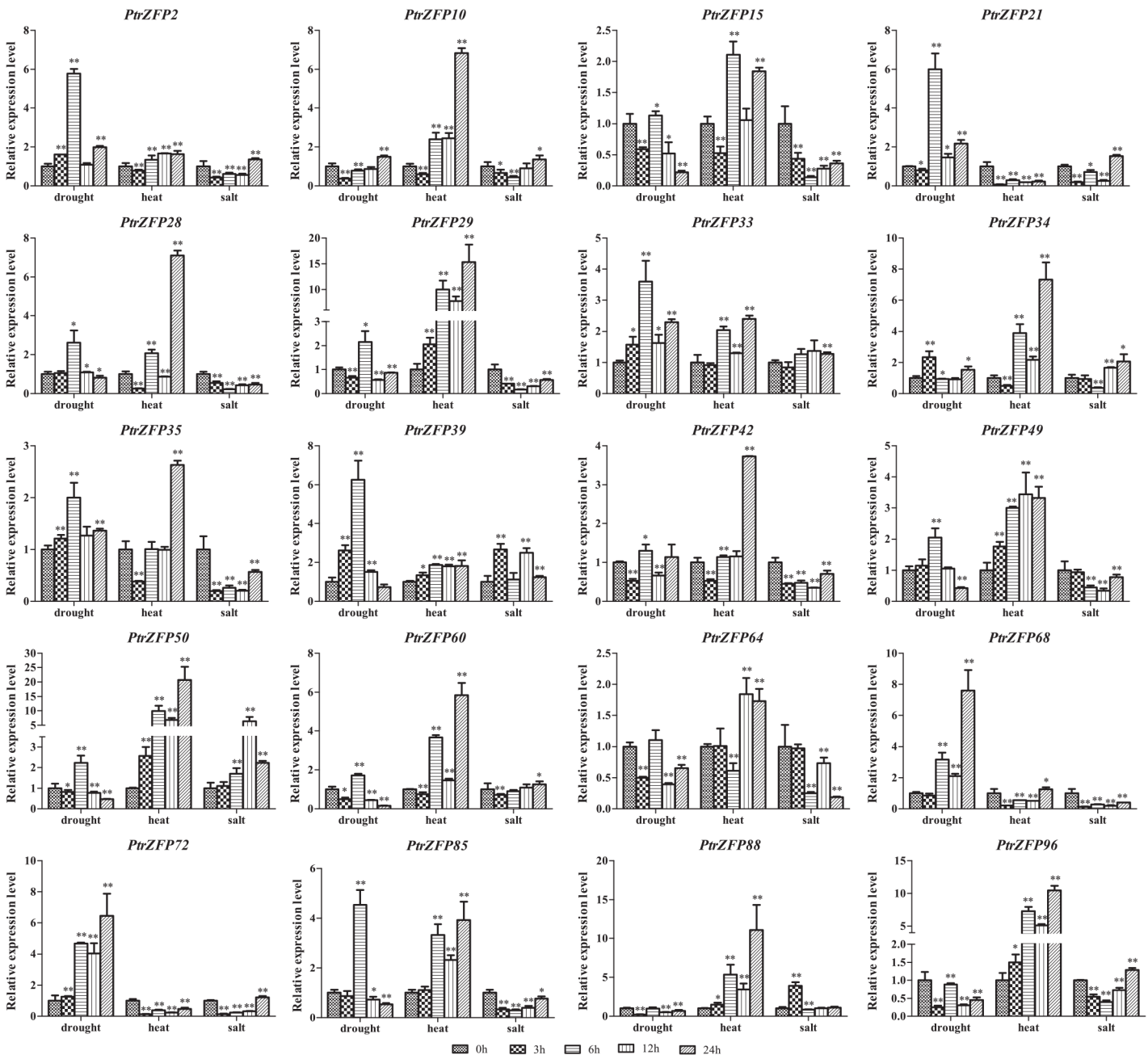
doi:10.1371/journal.pone.0134753.g006

Fig 7. Expression analysis of 44 selected C2HC2-ZF genes in leaves under drought, heat and salt stresses using qRT-PCR. The relative mRNA abundance of 44 selected C2HC2-ZF genes was normalized with respect to the reference gene (*Actin1*). The x-axis represents time points after stress treatments. Error bars represent the standard deviations from three biological replicates. Asterisks indicate stress treatment groups that showed a significant difference in transcript abundance compared with the control group (* $P < 0.05$, ** $P < 0.01$).

Fig 8. Expression analysis of 20 selected C2HC2-ZF genes in roots under drought, heat and salt stresses using qRT-PCR. The relative mRNA abundance of 20 selected C2HC2-ZF genes was normalized with respect to the reference gene (*Actin1*). The x-axis represents time points after stress treatments. Error bars represent the standard deviations from three biological replicates. Asterisks indicate stress treatment groups that showed a significant difference in transcript abundance compared with the control group (* $P < 0.05$, ** $P < 0.01$).

leaves and roots. *PtrZFP95* was greatly upregulated (>70-fold) at 24 h after NaCl treatment (Figs 7 and 8).

In summary, our results showed that these candidate C2H2-ZF genes responded positively to more than one kind of stress. For example, among the 41 drought-inducible genes in leaves, 32 were upregulated by heat and 28 by salt. Meanwhile, among the 16 drought-inducible genes in roots, 15 were upregulated by heat and four by salt. *PtrZFP2* demonstrated a significantly positive response to all three stresses at some time points after treatment, whether in leaves or roots. It is also noteworthy that all upregulated gene promoter sequences contained *cis*-elements related to phytohormone and abiotic stresses, such as ABRE, G-box, HSE, MBS, TCA-element and W-Box. However, we also observed that some genes were downregulated under various stresses: *PtrZFP71* was suppressed in leaves by drought and salt, whereas *PtrZFP29* was suppressed in leaves by heat and suppressed in roots by salt.

Discussion

Preliminary analysis of the C2H2-ZF gene family has been performed in the model plants *Arabidopsis* and rice [13,14]. However, this family has not previously been studied in *P. trichocarpa*, a model forest tree. In this report, we identified 109 full-length C2H2-ZF genes in the *P. trichocarpa* genome, each of which contains at least one C2H2-ZF motif. The lengths of these sequences varied significantly, implying a high degree of complexity among the C2H2-ZF genes. About 94% (103 of 109) of the C2H2-ZF genes were predicted as nuclear proteins, whereas the other six genes were predicted to be located in the chloroplast or cytoplasm. This suggested that these genes had special functions compared with the other members in this family. Furthermore, based on the WoLF PSORT analyses, we were able to roughly determine their gene localizations; however, experimental verification is required for more accurate localization. Based on the phylogenetic analysis, we identified 40 paralogous pairs among the 109 *Populus* C2H2-ZF genes. Among the 40 gene pairs, 39 belong to the same group and identical subcellular localizations were predicted for their proteins (Fig 1A, S2 and S4 Tables). However, PtrZFP35 and PtrZFP68 were predicted to be localized in the chloroplast and nucleus, respectively, even though they were both placed in group II. These results indicated that the same phylogenetic grouping based on sequence similarity did not necessarily correspond to the same subcellular localization. Therefore, homologous genes may show differences in gene function and signal transduction. These results were similar to those of a previous analysis of the C3HC4-type RING finger gene family in *Populus* [65].

Exon–intron increase or decrease can be caused by integrations and realignments of gene fragments. Therefore, gene structural variation plays a major role in the evolution of gene families [66]. The current study provided an example of such diversification in the form of a C2H2-ZF gene (*PtrZFP34*) with only one exon, whereas other genes in the same phylogenetic group (group III) have seven or eight exons. Moreover, *PtrZFP70* has 12 exons, whereas *PtrZFP20*, which is similar to *PtrZFP70*, only has two exons. Previous studies showed that many plant-specific Q-type C2H2-ZFPs play important roles in diverse environmental stress responses, as well as in various plant developmental and physiological processes [14,57]. In this study, about 57% (62 of 109) of C2H2-ZFPs from *Populus* had Q-type ZF domains, which was a larger number compared with other experimental models such as *Arabidopsis* (36%) and rice (34%). This result suggested that Q-type C2H2-ZFPs are more important for woody plants. However, only eight proteins (PtrZFP1, 4, 5, 6, 19, 22, 69 and 98) contained motif 11, which suggested that these proteins may have special functions. The similar gene structures and conserved motifs of C2H2-ZF genes and proteins in the same subfamilies may provide additional

support for the phylogenetic analysis. Conversely, the differences in gene structure and motif composition among different subfamilies indicated that they might be functionally divergent.

Gene duplication events, including tandem and segmental duplication, play an important role in genomic expansions and realignments [67,68]. Gene duplication has been reported for many plant transcription factor gene families, such as the NAC, CCCH and HD-ZIP families [44,69,70]. To verify this, the mechanisms involved in the expansion of C2H2-ZF members in *Populus* were examined. Among the 40 sister pairs, only one pair (*PtrZFP37/107*) of the C2H2-ZF genes appeared to have undergone tandem duplication, based on their more than 96% similarity at the amino acid level. By contrast, 39 segmental duplication events were identified, suggesting the existence of low-tandem and high-segmental repetitions in the C2H2-ZF gene family. The result was similar to that observed for the WRKY duplications in *P. trichocarpa* [71].

In this study, we compared the members of the C2H2-ZF gene family in *Populus* with those in rice and found that three groups were plant-specific clusters including distinct types of C2H2-ZF. Compared with the rice, some proteins in the *Populus* group a had added a C-type ZF domain, which suggested that woody plants might have undergone a series of changes during the evolutionary process. In addition, three small unique clusters (ZOS1-15/ZOS5-09/PtrZFP104, ZOS3-20/PtrZFP14/PtrZFP60 and ZOS9-07/ZOS11-06d/ZOSj/PtrZFP27) may have undergone independent evolutionary trajectories from the other clusters. Two 2-fingered proteins (PtrZFP55 and 79) were assigned to group c, which suggested that they might share a common origin with other single Q-type finger members.

It is accepted that *cis*-elements play key roles in the transcriptional regulation of genes controlling various abiotic stress and phytohormone responses. Plant hormones play central roles in the ability of plants to adapt to changing environments. Many abiotic stress-related and phytohormone-related *cis*-elements, including MBS, HSE, ABRE, W-Box, ERE and TCA-elements, have been identified [72–74]. All of these *cis*-elements were observed in the present study (S4 Table). Each gene of this family contained at least two *cis*-elements related to phytohormone or abiotic stress signal responsiveness. *PtrZFP17*, *40* and *94* have only two *cis*-elements, which suggested that this gene might not be associated with abiotic stress. By contrast, *PtrZFP4* and *92* have nine *cis*-elements, which suggested that these genes might have important functions under different abiotic stresses. These results were consistent with expression profiles of *Populus* C2H2-ZF genes under various stresses. *PtrZFP17* and *94* showed no change under drought and mechanical damage, whereas *PtrZFP4* was upregulated in leaves at some time points after drought, heat and salt treatment.

Evidence suggested that C2H2-ZF transcription factors are involved widely in the integration and development of many organs and tissues, such as seed maturation [75], floral development [76], secondary metabolism and cell wall structure [77]. Using *in silico* analysis, *PtrZFP105*, identified in the present study, was observed to be preferentially expressed in young leaves, suggesting a role in the regulation of adaxial leaf fate. *PtrZFP2*, *3*, *16*, *32*, *35*, *39*, *60* and *72* showed high expression levels in roots, indicating that these genes may affect root development. The putative paralogous genes (*PtrZFP54/59*) showed high transcript accumulation in the cambial zone, suggesting that they are involved in the same regulation network of biological processes. *PtrZFP8*, *31*, *51*, *55*, *65*, *66* and *107* were mainly expressed in flower buds, male catkins and female catkins, indicating that these genes might be closely related to flower differentiation and floral organ formation. Notably, *RID1* (LOC_Os10g28330), a *PtrZFP65* orthologous gene, acts as a master switch from vegetative to floral development in rice [76]. This suggested that *PtrZFP65* might have functions similar to *RID1* during floral development in *P. trichocarpa*. Interestingly, *PtrZFP107* was preferentially expressed in *Populus* flower buds, whereas *ENY* (AT5G66730), an *Arabidopsis* ortholog of *PtrZFP107*, is involved in seed

maturation [75]. This result suggested that their functions might vary in different plant species. In addition, *PtrZFP53* was abundantly expressed in tension wood. *Populus*, as an important tree species for a large variety of wood-based products, produces abundant wood (secondary cell wall) compared with herbaceous plants. Tension wood is mainly characterized by abnormal fibers that are poorly lignified and have an additional thick layer in the secondary cell wall [78]. Therefore, tension wood is involved in the formation of wood. Considering the high expression level of *PtrZFP53* in tension wood, we hypothesized that the C2H2-ZF gene family might contribute to wood formation. This conclusion is consistent with previous research on cotton fiber quality [77]. *Populus* C2H2-ZF genes may also be involved in other biological processes, such as shoot and cambium development, hinted at by their abundant expression in the shoot meristem and active cambium. Gene ontology annotation also supported this hypothesis. Determination of the functions and mechanisms of action of the identified C2H2-ZF members requires further experiments.

In the life cycle of a tree, growth and productivity are frequently threatened by environmental stresses, such as beetle damage, drought, heat and high salinity. These stresses may cause fatal damage to trees [79–81]. However, many stress-related genes are induced to help plants adapt to these biotic and abiotic stresses. In this study, *PtrZFP2*, *21*, *28*, *33*, *35*, *68* and *72* were upregulated in drought, beetle and mechanical damage, suggesting that these genes play essential roles under multiple stresses. In plants, leaves and roots are the most important organs for resisting abiotic stress. Plant leaves provide an adaptive mechanism for plants undergoing abiotic stress by increasing stomatal closure, decreasing transpiration rate and reducing the leaf area [82,83]. Roots can sense soil changes in the surrounding environment under abiotic stress conditions, and send a series of signals to the leaves and shoots to reduce root damage and maintain plant growth in spite of water shortage [84]. Based on *in silico* analysis and our qRT-PCR analysis, 32 genes were upregulated in leaves in the comparison between drought and heat, 28 genes were upregulated in leaves in the comparison between drought and salt, 15 genes were upregulated in roots in the comparison between drought and heat, and four genes were upregulated in roots in the comparison between drought and salt. *PtrZFP2* acts as a positive regulator under all three stresses in leaves and roots. Notably, *ZFP252* (GenBank: AAO46041.1), a *PtrZFP2* orthologous gene, can enhance drought and salt tolerance in rice [85]. Similar results were found in potato [86]. Our study found that *PtrZFP2* was also involved in response to heat stress. These results indicated that C2H2-ZF genes might be involved in the substantial common regulatory systems or cross-talk triggered by different stresses. Many abiotic stresses ultimately result in dehydration and osmotic imbalance of plant cells; therefore, it is not surprising that there is a large overlap of genes induced by drought, heat and salt stresses [81,87]. These genes may have shared roles in two or more stresses. Interestingly, some drought-upregulated genes (*PtrZFP29* and *33*) were downregulated by heat stress, which indicated that two sets of C2H2-ZF genes are involved in response to drought and heat stress, respectively. The huge expression differences suggest that these genes carry out different physiological and biochemical functions to adapt to complicated challenges. Therefore, it would be interesting to undertake further functional studies of these C2H2-ZF genes to establish the interactions among the particular pathways that are activated during the drought, heat and salt stress responses.

Conclusions

In the present study, we performed a comprehensive analysis of phylogenetic relationships, chromosomal locations, gene structures, conserved motifs, *cis*-elements and expression profiles of the C2H2-ZF gene family in *Populus*. One hundred and nine full-length C2H2-ZF genes in

the *Populus* genome were identified, which were clustered phylogenetically into four distinct subfamilies. The genes were non-randomly distributed across 19 LGs, and segmental duplications had contributed significantly to the expansion of the *Populus* C2H2-ZF gene family. *Cis*-elements in the C2H2-ZF genes provided clues to their functions and expression in specific tissues or organs, as well as under different biotic and abiotic stresses. We identified a subset of *Populus* C2H2-ZF genes with putative functional roles in drought, heat and salt responses. The information obtained in the current study could help to select appropriate candidate genes for further functional characterization to unravel their divergent roles.

Supporting Information

S1 Table. Primers for qRT-PCR of 51 selected C2HC2-ZF genes were designed using Primer Premier 5 (F represents a forward primer, R represents a reverse primer).
(XLSX)

S2 Table. List of 109 C2HC2-ZF genes identified in *P. trichocarpa* and their sequence characteristics (aa, amino acids; Da, Dalton).
(XLSX)

S3 Table. Alternative splicing of *Populus* C2HC2-ZF genes.
(XLSX)

S4 Table. Divergence between C2HC2-ZF genes pairs in *Populus*. Gene pairs were identified at the terminal nodes (> 96% identical) of the phylogenetic tree shown in Fig 1A. Synonymous (Ks) and nonsynonymous substitution (Ka) rates are presented for each pair. Gene pairs created by tandem duplication or segmental duplication are indicated in the table.
(XLSX)

S5 Table. Motif sequences of C2HC2-ZF genes identified in *P. trichocarpa*.
(XLSX)

S6 Table. Types of C2H2-ZF motifs.
(XLSX)

S7 Table. Numbers and types of *Populus* C2H2-ZFPs.
(XLSX)

S8 Table. Phytohormone and abiotic stress related *cis*-elements.
(XLSX)

S9 Table. Abiotic stress and phytohormone response elements in *Populus* C2H2-ZFP promoters.
(XLSX)

S10 Table. Details of the Gene Ontology annotation of *Populus* C2HC2-ZFP sequences.
(XLSX)

S1 Text. Information for 109 C2HC2-ZF genes. A complete list of the 109 C2HC2-ZF gene sequences of *P. trichocarpa* identified in this study. Genomic DNA and amino acid sequences were downloaded from Phytozome v10.0 (http://phytozome.jgi.doe.gov/pz/portal.html) and NCBI (http://www.ncbi.nlm.nih.gov/) databases.
(TXT)

Author Contributions

Conceived and designed the experiments: QL CL. Performed the experiments: QL ZW. Analyzed the data: QL HZ XX CL. Contributed reagents/materials/analysis tools: QL ZW CL. Wrote the paper: QL CL.

References

1. Takatsuji H. Zinc-finger proteins: the classical zinc finger emerges in contemporary plant science. Plant Mol Biol. 1999; 39: 1073–1078. PMID: 10380795

2. Moore M, Ullman C. Recent developments in the engineering of zinc finger proteins. Brief Funct Genomic Proteomic. 2003; 1: 342–355. PMID: 15239882

3. Berg JM, Shi Y. The galvanization of biology: a growing appreciation for the roles of zinc. Science. 1996; 271: 1081–1085. PMID: 8599083

4. Takatsuji H, Mori M, Benfey PN, Ren L, Chua NH. Characterization of a zinc finger DNA-binding protein expressed specifically in *Petunia* petals and seedlings. Embo J. 1992; 11: 241–249. PMID: 1740109

5. Hu W, Ma H. Characterization of a novel putative zinc finger gene. The Plant Journal. 2006; 45: 399–422. PMID: 16412086

6. Huang F, Chi Y, Meng Q, Gai J, Yu D. GmZFP1 encoding a single zinc finger protein is expressed with enhancement in reproductive organs and late seed development in soybean (*Glycine max*). Mol Biol Rep. 2006; 33: 279–285. PMID: 17077988

7. Mukhopadhyay A, Vij S, Tyagi AK. Overexpression of a zinc-finger protein gene from rice confers tolerance to cold, dehydration, and salt stress in transgenic tobacco. Proc Natl Acad Sci USA. 2004; 101: 6309–6314. PMID: 15079051

8. Colasanti J, Yuan Z, Sundaresan V. The indeterminate gene encodes a zinc finger protein and regulates a leaf-generated signal required for the transition to flowering in maize. Cell. 1998; 93: 593–603. PMID: 9604934

9. Pabo CO, Peisach E, Garant RA. Design and selection of novel Cys2His2 zinc finger proteins. Annu Rev Biochem. 2001; 70: 313–340. PMID: 11395410

10. Kubo K, Sakamoto A, Kobayashi A, Rybka Z, Kanno Y, Nakagawa H, et al. Cys2/His2 zinc-finger protein family of petunia: evolution and general mechanism of target-sequence recognition. Nucleic Acids Res. 1998; 26: 608–615. PMID: 9421523

11. Chung HR, Schafer U, Jackle H, Bohm S. Genomic expansion and clustering of ZAD-containing C2H2 zinc-finger genes in *Drosophila*. Embo Rep. 2002; 3: 1158–1162. PMID: 12446571

12. Bohm S, Frishman D, Mewes HW. Variations of the C2H2 zinc finger motif in the yeast genome and classification of yeast zinc finger proteins. Nucleic Acids Res. 1997; 25: 2464–2469. PMID: 9171100

13. Englbrecht CC, Schoof H, Böhm S. Conservation, diversification and expansion of C2H2 zinc finger proteins in the *Arabidopsis thaliana* genome. BMC Genomics. 2004; 5: 39. PMID: 15236668

14. Agarwal P, Arora R, Ray S, Singh AK, Singh VP, Takatsuji H, et al. Genome-wide identification of C2H2 zinc-finger gene family in rice and their phylogeny and expression analysis. Plant Mol Biol. 2007; 65: 467–485. PMID: 17610133

15. Muthamilarasan M, Bonthala VS, Mishra AK, Khandelwal R, Khan Y, Roy R, et al. C2H2 type of zinc finger transcription factors in foxtail millet define response to abiotic stresses. Functional & Integrative Genomics. 2014; 14: 531–543.

16. Sakamoto H. *Arabidopsis* Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. Plant Physiol. 2004; 136: 2734–2746. PMID: 15333755

17. Klug A, Schwabe JW. Protein motifs 5. Zinc fingers. Faseb J. 1995; 9: 597–604. PMID: 7768350

18. Ciftci-Yilmaz S, Mittler R. The zinc finger network of plants. Cell Mol Life Sci. 2008; 65: 1150–1160. doi: 10.1007/s00018-007-7473-4 PMID: 18193167

19. Gamsjaeger R, Liew C, Loughlin F, Crossley M, Mackay J. Sticky fingers: zinc-fingers as protein-recognition motifs. Trends Biochem Sci. 2007; 32: 63–70. PMID: 17210253

20. Lee BM, Xu J, Clarkson BK, Martinez-Yamout MA, Dyson HJ, Case DA, et al. Induced fit and "Lock and Key" recognition of 5S RNA by zinc fingers of transcription factor IIIA. J Mol Biol. 2006; 357: 275–291. PMID: 16405997

21. Dinneny JR, Weigel D, Yanofsky MF. *NUBBIN* and *JAGGED* define stamen and carpel shape in *Arabidopsis*. Development. 2006; 133: 2285.

22. Joseph MP, Papdi C, Kozma-Bognar L, Nagy I, Lopez-Carbonell M, Rigo G, et al. The *Arabidopsis* zinc finger protein3 interferes with abscisic acid and light signaling in seed germination and plant development. Plant Physiol. 2014; 165: 1203–1220. PMID: 24808098

23. Yang L, Liu Z, Lu F, Dong A, Huang H. SERRATE is a novel nuclear regulator in primary microRNA processing in *Arabidopsis*. The Plant Journal. 2006; 47: 841–850. PMID: 16889646

24. Zhou Z, Sun L, Zhao Y, An L, Yan A, Meng X, et al. Zinc finger protein 6 (ZFP6) regulates trichome initiation by integrating gibberellin and cytokinin signaling in *Arabidopsis thaliana*. New Phytol. 2013; 198: 699–708. doi: 10.1111/nph.12211 PMID: 23506479

25. Li S, Zhou X, Chen L, Huang W, Yu D. Functional characterization of *Arabidopsis thaliana* WRKY39 in heat stress. Mol Cells. 2010; 29: 475–483. doi: 10.1007/s10059-010-0059-2 PMID: 20396965

26. Huang J, Sun S, Xu D, Yang X, Bao Y, Wang Z, et al. Increased tolerance of rice to cold, drought and oxidative stresses mediated by the overexpression of a gene that encodes the zinc finger protein ZFP245. Biochem Bioph Res Co. 2009; 389: 556–561.

27. Sun SJ, Guo SQ, Yang X, Bao YM, Tang HJ, Sun H, et al. Functional analysis of a novel Cys2/His2-type zinc finger protein involved in salt tolerance in rice. J Exp Bot. 2010; 61: 2807–2818. doi: 10.1093/jxb/erq120 PMID: 20460361

28. Luo X, Bai X, Zhu D, Li Y, Ji W, Cai H, et al. GsZFP1, a new Cys2/His2-type zinc-finger protein, is a positive regulator of plant tolerance to cold and drought stress. Planta. 2012; 235: 1141–1155. doi: 10.1007/s00425-011-1563-0 PMID: 22160567

29. Kim SH, Hong JK, Lee SC, Sohn KH, Jung HW, Hwang BK. *CAZFP1*, Cys2/His2-type zinc-finger transcription factor gene functions as a pathogen-induced early-defense gene in *Capsicum annuum*. Plant Mol Biol. 2004; 55: 883–904. PMID: 15604723

30. Gourcilleau D, Lenne C, Armenise C, Moulia B, Julien JL, Bronner G, et al. Phylogenetic study of plant Q-type C2H2 zinc finger proteins and expression analysis of poplar genes in response to osmotic, cold and mechanical stresses. Dna Res. 2011; 18: 77–92. doi: 10.1093/dnares/dsr001 PMID: 21367962

31. Taylor G. *Populus*: *Arabidopsis* for forestry. Do we need a model tree? Ann Bot-London. 2002; 90: 681–689.

32. Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science. 2006; 313: 1596–1604. PMID: 16973872

33. Finn RD. Pfam: clans, web tools and services. Nucleic Acids Res. 2006; 34: D247–D251. PMID: 16381856

34. Letunic I. SMART 4.0: towards genomic data integration. Nucleic Acids Res. 2004; 32: 142D–144D.

35. Riaño-Pachón D, Ruzicic S, Dreyer I, Mueller-Roeber B. PlnTFDB: an integrative plant transcription factor database. BMC Bioinformatics. 2007; 8: 42. PMID: 17286856

36. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, et al. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 2007; 35: W585–W587. PMID: 17517783

37. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In: John MW (ed) The proteomics protocols handbook. New York: Humana Press; 2005:571–607.

38. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 1997; 25: 4876–4882. PMID: 9396791

39. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In Nucleic acids symposium series. 1999; 41: 95–98.

40. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011; 28: 2731–2739. doi: 10.1093/molbev/msr121 PMID: 21546353

41. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2014: btu817.

42. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009; 37: W202–W208. doi: 10.1093/nar/gkp335 PMID: 19458158

43. Sjödin A, Street NR, Sandberg G, Gustafsson P, Jansson S. The *Populus* Genome Integrative Explorer (PopGenIE): a new resource for exploring the *Populus* genome. New Phytol. 2009; 182: 1013–1025. doi: 10.1111/j.1469-8137.2009.02807.x PMID: 19383103

44. Hu R, Qi G, Kong Y, Kong D, Gao Q, Zhou G. Comprehensive analysis of NAC domain transcription factor gene family in *Populus trichocarpa*. BMC Plant Biol. 2010; 10: 145. doi: 10.1186/1471-2229-10-145 PMID: 20630103

45. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012; 40: e49. doi: 10.1093/nar/gkr1293 PMID: 22217600

46. Chen X, Chen Z, Zhao H, Zhao Y, Cheng B, Xiang Y. Genome-wide analysis of soybean HD-Zip gene family and expression profiling under salinity and drought treatments. PLoS One. 2014; 9: e87156. doi: 10.1371/journal.pone.0087156 PMID: 24498296

47. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. Nucleic Acids Res. 2002; 30: 325–327. PMID: 11752327

48. Conesa A, Gotz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics. 2008; 2008: 619832. doi: 10.1155/2008/619832 PMID: 18483572

49. Street N, Sjödin A, Bylesjö M, Gustafsson P, Trygg J, Jansson S. A cross-species transcriptomics approach to identify genes involved in leaf development. BMC Genomics. 2008; 9: 589. doi: 10.1186/1471-2164-9-589 PMID: 19061504

50. Song Y, Ci D, Tian M, Zhang D. Comparison of the physiological effects and transcriptome responses of *Populus simonii* under different abiotic stresses. Plant Mol Biol. 2014; 86: 139–156. doi: 10.1007/s11103-014-0218-5 PMID: 25002226

51. Bae EK, Lee H, Lee JS, Noh EW. Drought, salt and wounding stress induce the expression of the plasma membrane intrinsic protein 1 gene in poplar (*Populus alba* x *P. tremula* var. *glandulosa*). Gene. 2011; 483: 43–48. doi: 10.1016/j.gene.2011.05.015 PMID: 21640804

52. Chang S, Puryear J, Cairney J. A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Rep. 1993; 11: 113–116.

53. Chu Y, Huang Q, Zhang B, Ding C, Su X. Expression and molecular evolution of two DREB1 genes in black poplar (*Populus nigra*). PLoS One. 2014; 9: e98334. doi: 10.1371/journal.pone.0098334 PMID: 24887081

54. Wang Z, Zhang H, Yang J, Chen Y, Xu X, Mao X, et al. Phylogenetic, expression, and bioinformatic analysis of the ABC1 gene family in *Populus trichocarpa*. The Scientific World Journal. 2013; 2013: 1–11.

55. Guo YH, Yu YP, Wang D, Wu CA, Yang GD, Huang JG, et al. GhZFP1, a novel CCCH-type zinc finger protein from cotton, enhances salt stress tolerance and fungal disease resistance in transgenic tobacco by interacting with GZIRD21A and GZIPR5. New Phytol. 2009; 183: 62–75. doi: 10.1111/j.1469-8137.2009.02838.x PMID: 19402879

56. Hiratsu K, Mitsuda N, Matsui K, Ohme-Takagi M. Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in *Arabidopsis*. Biochem Bioph Res Co. 2004; 321: 172–178.

57. Kam J, Gresshoff PM, Shorter R, Xue GP. The Q-type C2H2 zinc finger subfamily of transcription factors in *Triticum aestivum* is predominantly expressed in roots and enriched with members containing an EAR repressor motif and responsive to drought stress. Plant Mol Biol. 2008; 67: 305–322. doi: 10.1007/s11103-008-9319-3 PMID: 18347915

58. Barakat A, Bagniewska-Zadworna A, Choi A, Plakkat U, DiLoreto DS, Yellanki P, et al. The cinnamyl alcohol dehydrogenase gene family in *Populus*: phylogeny, organization, and expression. BMC Plant Biol. 2009; 9: 26. doi: 10.1186/1471-2229-9-26 PMID: 19267902

59. Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM. Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. Plant Physiol. 2009; 149: 981–993. doi: 10.1104/pp.108.132795 PMID: 19091872

60. Nakashima K, Yamaguchi-Shinozaki K, Shinozaki K. The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. Front Plant Sci. 2014; 5: 170. doi: 10.3389/fpls.2014.00170 PMID: 24904597

61. Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, et al. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. Curr Opin Plant Biol. 2006; 9: 436–442. PMID: 16759898

62. Santner A, Estelle M. Recent advances and emerging trends in plant hormone signalling. Nature. 2009; 459: 1071–1078. doi: 10.1038/nature08122 PMID: 19553990

63. Ohlrogge J, Benning C. Unraveling plant metabolism by EST analysis. Curr Opin Plant Biol. 2000; 3: 224–228. PMID: 10837264

64. Li C, Lu S. Molecular characterization of the SPL gene family in *Populus trichocarpa*. BMC Plant Biol. 2014; 14: 131. doi: 10.1186/1471-2229-14-131 PMID: 24884654

65. Liu Q, Yang J, Wang Z, Xu X, Mao X, Li D, et al. Genome-wide classification, identification and expression profile of the C3HC4-type RING finger gene family in poplar (*Populus trichocarpa*). Plant Mol Biol Rep. 2015 Mar 22. doi: 10.1007/s11105-015-0870-1

66. Xu G, Guo C, Shan H, Kong H. Divergence of duplicate genes in exon-intron structure. Proc Natl Acad Sci USA. 2012; 109: 1187–1192. doi: 10.1073/pnas.1109047109 PMID: 22232673

67. Vision TJ, Brown DG, Tanksley SD. The origins of genomic duplications in *Arabidopsis*. Science. 2000; 290: 2114–2117.

68. Kumar R, Tyagi AK, Sharma AK. Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. Mol Genet Genomics. 2011; 285: 245–260. doi: 10.1007/s00438-011-0602-7 PMID: 21290147

69. Chai G, Hu R, Zhang D, Qi G, Zuo R, Cao Y, et al. Comprehensive analysis of CCCH zinc finger family in poplar (*Populus trichocarpa*). BMC Genomics. 2012; 13: 253. doi: 10.1186/1471-2164-13-253 PMID: 22708723

70. Hu R, Chi X, Chai G, Kong Y, He G, Wang X, et al. Genome-wide identification, evolutionary expansion, and expression profile of homeodomain-leucine zipper gene family in poplar (*Populus trichocarpa*). PLoS One. 2012; 7: e31149. doi: 10.1371/journal.pone.0031149 PMID: 22359569

71. He H, Dong Q, Shao Y, Jiang H, Zhu S, Cheng B, et al. Genome-wide survey and characterization of the WRKY gene family in *Populus trichocarpa*. Plant Cell Rep. 2012; 31: 1199–1217. doi: 10.1007/s00299-012-1241-0 PMID: 22371255

72. Narusaka Y, Nakashima K, Shinwari ZK, Sakuma Y, Furihata T, Abe H, et al. Interaction between two *cis*-acting elements, ABRE and DRE, in ABA-dependent expression of *Arabidopsis rd29A* gene in response to dehydration and high-salinity stresses. Plant J. 2003; 34: 137–148. PMID: 12694590

73. Diaz-Martin J, Almoguera C, Prieto-Dapena P, Espinosa JM, Jordano J. Functional interaction between two transcription factors involved in the developmental regulation of a small heat stress protein gene promoter. Plant Physiol. 2005; 139: 1483–1494. PMID: 16244139

74. Wei Q, Hu P, Kuai BK. Ectopic expression of an *Ammopiptanthus mongolicus* H$^+$-pyrophosphatase gene enhances drought and salt tolerance in *Arabidopsis*. Plant Cell, Tissue and Organ Culture (PCTOC). 2012; 110: 359–369.

75. Feurtado JA, Huang D, Wicki-Stordeur L, Hemstock LE, Potentier MS, Tsang EW, et al. The *Arabidopsis* C2H2 zinc finger INDETERMINATE DOMAIN1/ENHYDROUS promotes the transition to germination by regulating light and hormonal signaling during seed maturation. Plant Cell. 2011; 23: 1772–1794. doi: 10.1105/tpc.111.085134 PMID: 21571950

76. Wu C, You C, Li C, Long T, Chen G, Byrne ME, et al. *RID1*, encoding a Cys2/His2-type zinc finger transcription factor, acts as a master switch from vegetative to floral development in rice. Proc Natl Acad Sci USA. 2008; 105: 12915–12920. doi: 10.1073/pnas.0806019105 PMID: 18725639

77. Al-Ghazi Y, Bourot S, Arioli T, Dennis ES, Llewellyn DJ. Transcript profiling during fiber development identifies pathways in secondary metabolism and cell wall structure that may contribute to cotton fiber quality. Plant Cell Physiol. 2009; 50: 1364–1381. doi: 10.1093/pcp/pcp084 PMID: 19520671

78. Lafarguette F, Leplé J, Déjardin A, Laurans F, Costa G, Lesage-Descauses M, et al. Poplar genes encoding fasciclin-like arabinogalactan proteins are highly expressed in tension wood. New Phytol. 2004; 164: 107–121.

79. Syvertsen JP, Garcia-Sanchez F. Multiple abiotic stresses occurring with salinity stress in citrus. Environ Exp Bot. 2014; 103: 128–137.

80. Jenkins MJ, Runyon JB, Fettig CJ, Page WG, Bentz BJ. Interactions among the mountain pine beetle, fires, and fuels. Forest Sci. 2014; 60: 489–501.

81. Bi H, Parekh J, Li Y, Murphy S, Lei Y. Adverse influences of drought and temperature extremes on survival of potential tree species for commercial environmental forestry in the dryland areas on the western slopes of New South Wales, Australia. Agr Forest Meteorol. 2014; 197: 188–205.

82. Caldeira CF, Bosio M, Parent B, Jeanguenin L, Chaumont F, Tardieu F. A hydraulic model is compatible with rapid changes in leaf elongation under fluctuating evaporative demand and soil water status. Plant Physiol. 2014; 164: 1718–1730. doi: 10.1104/pp.113.228379 PMID: 24420931

83. Huang XY, Chao DY, Gao JP, Zhu MZ, Shi M, Lin HX. A previously unknown zinc finger protein, DST, regulates drought and salt tolerance in rice via stomatal aperture control. Genes Dev. 2009; 23: 1805–1817. doi: 10.1101/gad.1812409 PMID: 19651988

84. Lynch J. Root architecture and plant productivity. Plant Physiol. 1995; 109: 7–13. PMID: 12228579

85. Xu DQ, Huang J, Guo SQ, Yang X, Bao YM, Tang HJ, et al. Overexpression of a TFIIIA-type zinc finger protein gene *ZFP252* enhances drought and salt tolerance in rice (*Oryza sativa* L.). FEBS letters. 2008; 582(7): 1037–1043. doi: 10.1016/j.febslet.2008.02.052 PMID: 18325341

86. Tian ZD, Zhang Y, Liu J, Xie CH. Novel potato C2H2-type zinc finger protein gene, *StZFP1*, which responds to biotic and abiotic stress, plays a role in salt tolerance. Plant biology. 2010; 12(5): 689–697. doi: 10.1111/j.1438-8677.2009.00276.x PMID: 20701691

87. Ludwig AA, Saitoh H, Felix G, Freymark G, Miersch O, Wasternack C, et al. Ethylene-mediated cross-talk between calcium-dependent protein kinase and MAPK signaling controls stress responses in plants. Proc Natl Acad Sci USA. 2005; 102: 10736–10741. PMID: 16027369