Analysis of the *Prefoldin* Gene Family in 14 Plant Species

Jun Cao*

Institute of Life Science, Jiangsu University, Zhenjiang, China

Prefoldin is a hexameric molecular chaperone complex present in all eukaryotes and archaea. The evolution of this gene family in plants is unknown. Here, I identified 140 *prefoldin* genes in 14 plant species. These prefoldin proteins were divided into nine groups through phylogenetic analysis. Highly conserved gene organization and motif distribution exist in each prefoldin group, implying their functional conservation. I also observed the segmental duplication of maize *prefoldin* gene family. Moreover, a few functional divergence sites were identified within each group pairs. Functional network analyses identified 78 co-expressed genes, and most of them were involved in carrying, binding and kinase activity. Divergent expression profiles of the maize *prefoldin* genes were further investigated in different tissues and development periods and under auxin and some abiotic stresses. I also found a few *cis*-elements responding to abiotic stress and phytohormone in the upstream sequences of the maize *prefoldin* genes. The results provided a foundation for exploring the characterization of the *prefoldin* genes in plants and will offer insights for additional functional studies.

**Keywords:** prefoldin, phylogeny, evolution, functional divergence, expression

**INTRODUCTION**

Molecular chaperones can stabilize, interact with, or assist other proteins to acquire functionally active conformations, without being present in their final structures (Hartl, 1996). Chaperonins play a role in protein assembly, folding, trafficking and degradation, and are crucial for cellular development (Hartl et al., 2011). Group I chaperonins are found in eubacteria and endosymbiotic organelles, whereas group II ones are present in archaea and eukaryotes. They share a common structure with different functions (Yébenes et al., 2011).

Prefoldin is a group II chaperonin. Archaea prefoldin possesses two subunits (*prefoldin* α and β) and polymerizes to an α2β4 hexamer (Leroux et al., 1999). Eukaryotic prefoldin consists of six subunits; two α subunits (PFD3 and PFD5) and four β subunits (PFD1, PFD2, PFD4, and PFD6) (Siegert et al., 2000). The structure confers selective substrate specificity for target proteins. They form a jellyfish-like heterohexameric complex to deliver newly synthesized unfolded proteins to cytosolic chaperonins containing TCP-1 (CCT) for protein folding. In addition, prefoldin also protects unfolded proteins during CCT capturing and releasing proteins (Siegert et al., 2000; Möckli et al., 2007).

Previous studies have indicated that prefoldin plays a central role in cellular development. Through binding to tubulin and actin, yeast prefoldin facilitated productive folding and binding inside the chaperonin cavity. Deleting of single or multiple subunits of prefoldin in yeast usually...
results in disruption of cytoskeletal structures (Vainberg et al., 1998). Similarly, studies on prefoldin function have revealed that down regulation of prefoldin can decreased the level of endogenous α-tubulin by 95%, and that of actin by 30% (Lundin et al., 2008). Deletion of prefoldin genes results in cytoskeletal defects, slow growth and cold sensitivity in yeast (Geissler et al., 1998; Siegers et al., 1999). In Caenorhabditis elegans, reduction of functional prefoldins caused embryonic lethality, and silencing of prefoldin subunits 1, 2, 3, and 6 exhibited reduced microtubule growth (Lundin et al., 2008). Mutation of prefoldin 5 or prefoldin 1 leads to a variety of neurodegenerative effects through a reduction of microfilaments and microtubules, such as mucus clearance defects, hydrocephaly, and loss of nerve bundles, in mice (Cao et al., 2008; Lee et al., 2011). In Arabidopsis, lesions in prefoldin 6 resulted in impaired microtubule organization and dynamics, which was associated with reduced plant size, defects in cell division, and so on (Gu et al., 2008). In addition, prefoldin 3 and prefoldin 5 mutants also displayed important changes in microtubule organization and developmental patterns. Furthermore, both mutants were sensitive to salt, implying important role of cytoskeleton in plant tolerance to salt stress (Rodríguez-Milla and Salinas, 2009). DELLA proteins directly interact with the prefoldin complex in a gibberellin (GA)-dependent manner. When GA is present, the complex is localized to the cytoplasm. When GA is absent, it stays in the nucleus, and then affects microtubule orientation (Locascio et al., 2013). A recent study has shown that prefoldin can bind chromatin and plays an important role in gene transcription (Millán-Zambrano et al., 2013). Bud27, an ATP-independent prefoldin-like molecular chaperone, can also regulate the gene expression transcribed by the RNA pol II in yeast (Miron-Garcia et al., 2014).

Six prefoldin members [PFD1 (At2g07340), PFD2 (At5g22480), PFD3 (At5g49510), PFD4 (At1g08780), PFD5 (At5g23290), and PFD6 (At1g29990)] have been identified in the Arabidopsis genome (Hill and Hemmingsen, 2001). Phylogenetic analysis showed that these six genes were divided into different evolutionary branches, suggesting the functional divergence among them (Hill and Hemmingsen, 2001). Moreover, only several of them have been functionally identified (Gu et al., 2008; Rodriguez-Milla and Salinas, 2009). Although prefoldin can form jellyfish-shaped hexameric complexes consisting of two α-type and four β-type subunits (Siegrist et al., 2000; Martin-Benito et al., 2002), smaller individual motifs and more divergent patterns are not recognized among each prefoldin proteins. The recent availability of genome sequences of some models plant species provides an opportunity to study the evolution of prefoldin gene family. In this study, I have identified the prefoldin gene family from 14 plant species, and each species comprises 6–24 genes. Considering their significant developmental and physiological role, it is of considerable interest to us to study the evolution of plant prefoldin genes. Here, I performed integrated analyses to unravel the evolutionary mechanisms of the plant prefoldin protein family especially for maize. It will provide a useful basis for further functional studies of these gene family.

**MATERIALS AND METHODS**

**Prefoldin Sequence Retrieval and Identification in 14 Plant Species**

To identify potential prefoldin gene in 14 completely sequenced plant genomes, I first used six Arabidopsis prefoldin sequences previous identified (Hill and Hemmingsen, 2001) as queries to perform BLAST searches against the phytozome database\(^1\) with \(-1\) expect (E) threshold. In addition, a keyword “prefoldin” was also used to perform searching in this study. The Pfam database (Punta et al., 2012) was used to confirm the encoded prefoldin based on the presence or not of the prefoldin subunit domain. Next, the ProtParam tool\(^2\) and the CELLO v2.5 server\(^3\) (Yu et al., 2004) were used to determine the physicochemical parameters and subcellular localization of the prefoldin proteins, respectively.

**Phylogenetic Analyses of the Prefoldin Gene Family**

To further explore the evolutionary relationship of plant prefoldins, multiple sequence alignments and phylogenetic analysis of the prefoldin proteins were performed with MUSCLE 3.52 (Edgar, 2004) and MEGA v5 (Tamura et al., 2011), respectively. And neighbor-joining (NJ) method was used to perform phylogenetic analyses of the prefoldin proteins with bootstrap analysis of 1,000 replicates (Cao, 2012). Furthermore, maximum likelihood and PhyML methods were also used to construct additional trees for validating the result from NJ tree.

**Chromosomal Location, Genomic Duplication and Inference of Duplication Time**

I used the annotation data of the prefoldin genes on MaizeSequence\(^4\) for chromosomal location. Paralogous regions of the genomes were first predicted with SyMAP v3.4 (Soderlund et al., 2011). If two genes coming from the paraloguous regions were also located at the terminal evolutionary branch in the phylogenetic tree, they were thought to be derived from common ancestral duplication. To calculate the duplication time of the prefoldin paralogs, pairwise alignment of the maize prefoldin gene pairs was performed using an embedded program ClustalW (codons) in MEGA v5 (Tamura et al., 2011). Next, the \(K_s\) and \(K_t\) values of paralogous genes were estimated using K-Estimator 6.0 program (Comeron, 1999). The \(K_s\) value was used to calculate the duplication date (\(T = K_s/2\kappa\)), assuming clock-like rates (\(\lambda\)) of 6.5 × 10⁻⁹ for maize (Gaut et al., 1996).

**Exon-Intron Structure and Conserved Motifs Analysis**

The organizations of prefoldin genes were analyzed in these plant lineages by comparing their coding and genomic sequence

\(^1\)http://www.phytozome.net
\(^2\)http://web.expasy.org/protparam
\(^3\)http://cello.life.nctu.edu.tw
\(^4\)http://www.maizesequence.org
information in the Phytozome\(^3\) and NCBI databases. In addition, MEME program\(^4\) (Bailey et al., 2006) was used to identify finer motifs in the candidate plant prefoldin protein sequences. Parameter of maximum number of motifs is 15.

Functional Divergence and Gene Co-expression Analyses

In the process of protein evolution, some residues are highly conserved, while others are highly variable. To further investigate the divergence between different groups of prefoldin proteins, DIVERGE (version 2.0) (Gu, 1999, 2001) was used to analyze the type-I functional divergence. The functional divergence between two groups was measured as the coefficient of divergence. When the coefficient equals 0; it means that the evolutionary rate of the duplicate genes at each site is entirely consistent. Vice versa, when the coefficient is greater than 0, the evolutionary rate of the duplicate genes at some critical amino-acid residues is different. The software will predict these sites responding for the functional genes at some critical amino-acid residues is different. The coefficient equals 0; it means that the evolutionary rate of the duplicate genes at each site is entirely consistent. Vice versa, when the coefficient is greater than 0, the evolutionary rate of the duplicate genes at some critical amino-acid residues is different. The software will predict these sites responding for the functional divergence.

To further analysis the relation between maize prefoldin and other genes, I also used the Co-expression Browser (COB)\(^7\) (Schaefer et al., 2014) to explore their networks. The domestication maize genotype was selected for co-expression analysis. To define the transcription start site (TSS) of each prefoldin gene in maize, I first collected their expressed sequence tag information. The TSS positions of prefoldin genes were used for the expression analysis of the prefoldin genes. Expression data were normalized and viewed in the Genesis (v 1.7.6) program (Sturn et al., 2002).

Microarray-Based Expression Analysis

Maize microarray data (GSE27004) (Sekhon et al., 2011) were used for the expression analysis of the prefoldin genes. Expression data were normalized and viewed in the Genesis (v 1.7.6) program (Sturn et al., 2002).

Plant Treatment, RNA Isolation, Quantitative Real-Time PCR (QRT-PCR), and Promoter Sequence Analysis

Endosperm can provide the necessary nutrition for early germination and growth when maize is grown in water. In this study, 10-day-old maize seedlings after germination were used to test the expression profiles of prefoldin genes under IAA, low temperature, drought, and salt stresses. For low temperature treatment, the maize seedlings were placed at 4°C environment for 3 h. And the seedlings were dried for 3 h between folds of tissue paper at 23 ± 1°C for drought treatment (Xia et al., 2012). The maize seedlings were put into 10 µM IAA and 150 mM NaCl solutions for 24 h for auxin and salt stress treatments, respectively. Control (CK) seedlings were normally grown at 23 ± 1°C with a photoperiod of 14 h light and 10 h dark. Three biological replicates were performed for qRT-PCR analysis. Total RNA was extracted with the TRIzol total RNA extraction kit (Sangon). RNase free DNase-I was used to remove genomic DNA. Next, I used M-MLV (TakaRa) to perform reverse transcription, followed by quantitative assays of each diluted cDNA using an ABI 7500 sequence detection system. All 13 maize prefoldin genes were selected for qRT-PCR analysis, and their primers are listed in Supplementary Table S2. Actin 1 (GRMZM2G126010) gene in maize was used as the endogenous control. And 2\(^−ΔΔCT\) method (Livak and Schmittgen, 2001) was used to calculate their relative expression level. t-Test was used to perform a significant analysis.

To define the transcription start site (TSS) of each prefoldin gene in maize, I first collected their expressed sequence tag information. The TSS positions of prefoldin genes were used

| Lineage | Organism          | Genome size (Mb)* | Number of predicted genes* | Number of prefoldin genes |
|---------|-------------------|-------------------|---------------------------|--------------------------|
| Algae   | Chlamydomonas reinhardtii | 120.41            | 14488                     | 8                        |
| Moss    | Physcomitrella patens     | 477.95            | 35936                     | 9                        |
| Lycophytes | Selaginella moellendorffii | 212.5            | 34782                     | 7                        |
| Dicots  | Solanum lycopersicum     | 781.51            | 27466                     | 8                        |
|         | Vitis vinifera          | 486.26            | 28268                     | 9                        |
|         | Arabidopsis thaliana    | 119.67            | 33583                     | 9                        |
|         | Populus trichocarpa     | 485.67            | 42577                     | 13                       |
|         | Cucumis sativus         | 244.82            | 21320                     | 7                        |
|         | Medicago truncatula     | 314.48            | 45000                     | 6                        |
|         | Glycine max             | 973.49            | 50202                     | 24                       |
| Monocots| Oryza sativa           | 382.78            | 30534                     | 10                       |
|         | Brachypodium distachyon | 272.06            | 26250                     | 8                        |
|         | Zea mays                | 2065.7            | 39454                     | 13                       |
|         | Sorghum bicolor         | 739.15            | 33081                     | 9                        |
| Total   |                    |                   |                           | 140                      |

*The data come from www.ncbi.nlm.nih.gov/genome/
FIGURE 1 | Phylogenetic relationship, gene structure and motif composition of the prefoldin genes in plants. The phylogenetic tree is constructed and classified into nine major groups (I–IX). The insertion positions of 0, 1, and 2 phase introns are marked with red, green, and blue lines, respectively. Different motifs of the prefoldin proteins are displayed by different colored boxes.
as references to determine their upstream promoter sequences. In this study, 1,000-bp upstream promoter sequences were acquired for further analyses. In addition, some abiotic stress- and phytohormone-responsive elements were identified in the promoter regions of the maize prefoldin genes using PLACE\(^8\) (Higo et al., 1999).

**RESULTS AND DISCUSSION**

**Identification of Prefoldin Genes in 14 Plant Species**

To identify the prefoldin genes in plants, I first used the amino-acid sequences of *Arabidopsis* prefoldin (Hill and Hemmingsen, 2001) to perform BLAST searches in the phytozome database\(^9\). In addition, a keyword “prefoldin” search was also performed. All the putative prefoldin protein sequences were subjected to the Pfam analysis (Punta et al., 2012) to verify the reliability of the results based on the presence or not of the prefoldin subunit domain. As a result, I identified 140 *prefoldin* genes from 14 plants. The number of *prefoldin* genes ranged from 6 to 24 in each species (Table 1). There are 24 *prefoldin* genes existing in the soybean genome, while the members in other species range from about six in alfalfa, nine in *Arabidopsis*, sorghum, grape and *Physcomitrella patens*, 10 in rice, and 13 in maize and poplar. Previous study has identified 6 *prefoldin* genes in *Arabidopsis* (Hill and Hemmingsen, 2001). In this study, other three *prefoldins* (At1g03760, At1g26660, and At1g49245) were also found in this species. There are about 33,583, 39,454, and 42,577 genes in the *Arabidopsis*, maize and poplar genomes, respectively, which are 9.9, 29.2, and 39.4% larger than that of rice (30,534), respectively, implying a disproportion between the numbers of *prefoldin* genes and the sizes of predicted genomes. The *prefoldin* genes in plants encode highly hydrophilous polypeptides (from −0.003 to −0.988 in grand average hydrophobicities) with about from 16 to 612 amino acids and predicted pIs (from 4.21 to 10.2) (Supplementary Table S1). CELLO v2.5 server\(^10\) (Yu et al., 2004) was used to further predict the localization of the prefoldin proteins. The results showed that most of the candidate prefoldins were probably localized to the cytoplasm, suggesting that the prefoldin proteins participate in the cytoplasmic folding of tubulin and actin monomers (Zhao et al., 2005; Gu et al., 2008). Several plant prefoldins have also been localized in the nucleus (Supplementary Table S1), suggesting their functional relevance in DNA repair or integration, transcription and gene regulation (Millán-Zambrano and Chávez, 2014; Mirón-García et al., 2014).

**Phylogenetic Relationships, Gene Organization, and Motif Analysis**

To examine the phylogenetic relationships and evolutionary history of prefoldins in plants, I first constructed an unrooted tree and classified this gene family into nine groups (Group I-IX) (Figure 1) based on the observed topological structure of tree and sequence similarity. In addition, other methods, such as, maximum likelihood, PhyML methods, were also used to reconstruct the phylogenetic trees of prefoldin family, and very similar results were got as well as the tree topology of NJ method. Here, I employed the NJ tree for further analysis. This group classification was also supported by other evidences, such as intron-exon organization, and motif composition, which will be described below. Phylogenetic relationships of some eukaryotic prefoldin subunits have been illustrated (Hill and Hemmingsen, 2001). In which, six different clusters were generated with six prefoldin subunit sequences through phylogenetic analysis. In this study, I also found that all members of Groups II and IV had high homology with eukaryotic PFD5 and PFD3, respectively, which form α subunits of the prefoldin complex. While the
FIGURE 3 | Location of the prefoldin genes and segmental duplication regions on maize chromosomes. The SyMAP database (Soderlund et al., 2011) was used to determine the segmental duplication regions.

members of Group V, Group VII, Group VIII, and Group IX represented four β subunits of PFD2, PFG4, PFD1 and PFD6, respectively. In addition to the six prefoldin subunit sequences, I also found and used three other prefoldin genes (At1g03760, At1g26660, and At1g49245) in Arabidopsis for BLAST searching. Because less conservation exists in prefoldin sequences, these sequences from 14 plant species were divided into nine groups (Figure 1). Each group contained 6–20 members. All terrestrial plants evolved from the primitive chlorophycean species in Viridiplantae (Misumi et al., 2008). Therefore, the origin of this gene may date back to the primitive algae. My search for prefoldin genes in Chlamydomonas reinhardtii yielded about eight members. This number remained relatively constant with the complexity of the genomes in plant evolution (Figure 2). The

TABLE 2 | Inference of duplication time of prefoldin paralogous pairs in maize.

| Paralogous pairs | $K_s$  | $K_s$  | $K_s/K_s$ | Data (million years ago) |
|------------------|--------|--------|-----------|----------------------------|
| AC203985.4_FGT005/GRMZM2G135354 | 0.01906 | 0.42986 | 0.04434 | 39.08 |
| GRMZM2G010065/GRMZM2G014676 | 0.02581 | 0 | – | 0 |
| GRMZM2G099909/GRMZM2G023347 | 0.00793 | 0.19994 | 0.03966 | 18.18 |
| GRMZM2G010065/GRMZM2G014676 | 0.02172 | 0.23698 | 0.09089 | 21.73 |
| GRMZM2G014676/GRMZM2G0119740 | 0.09121 | 0.19455 | 0.46883 | 17.69 |
| GRMZM2G010065/GRMZM2G0119740 | 1.26364 | 0.89611 | 1.41014 | 81.46 |
most obvious expansion of prefoldin gene family occurred in soybean, containing about 17% (24 paralogs) of the 140 identified prefoldins, which may be due to recent genome duplication occurring about 13 MYA (Schmutz et al., 2010).

In general, intron gain or loss generates structural complexity, which is a key evolutionary mechanism of most gene families (Cao, 2012; Chen and Cao, 2014; Cao et al., 2015). To further insight into the evolutionary relationships of the plant prefoldin genes, I investigated the exon-intron structure of each member and compared with each other. The position and distribution of introns of each prefoldin gene are shown in Figure 1. Each prefoldin gene contains 0–13 introns (Figure 1). Fifteen out of 20 members in Group V possessed a minimum of zero intron each, whereas GRMZM2G392932 possessed a maximum of thirteen introns. Similar exon-intron structure occurred in most prefoldin genes of same group, implying conserved evolutionary relationship. For instance, most prefoldin gene in Group VII has two introns with same insertion position and phase distribution, whereas most members within Group IV contain five introns. The structural diversity in different groups suggests multiple origins of gene ancestry.

Next, to obtain more insights into the diversity of motif compositions in the 140 prefoldin proteins, I used the MEME motif search tool (Bailey et al., 2006) to search for the conserved motifs in these proteins, and found 15 conserved motifs (Figure 1). As described above, nine groups were classified among these prefoldin genes based on the phylogenetic analyses. Noticeably, I also found common motif compositions in most members of each group, implying functional similarities between them (Figure 1). Several groups (such as Groups I, IV, and V) possess five motifs, while Group VI only has one motif. Except for members in Group VI, motif 3 was shared by most other prefoldin proteins. I also found some additional distinct motifs in specific groups, such as, motif 2 restricted in Group I, II, and III; motifs 14 and 15 in Group I; motif 9 in Group III; and motif 13, 7, 6, and 11 in Group IV, V, VIII, and IX, respectively. Previous study has indicated that low conservation exists between prefoldin sequences (Hill and Hemmingsen, 2001). My motif composition analyses of the prefoldin proteins also confirm it. The diversity of the prefoldin sequences may increase the complexity of the binding substrate, thus greatly expand the functional scope of the prefoldin complex. Further investigation may be required to determine if any of these distinct motifs across different groups also have a unique functional role.

### Chromosomal Position and Duplication of Prefoldin Genes in Maize

Gene duplication, including segmental duplication, tandem duplication and retroposition, plays an important role in the evolution of organism (Chen Y et al., 2014; Cao and Li, 2015). To investigate the duplication mechanisms of prefoldin genes, I first examined their physical locations on chromosomes and found that the prefoldin genes were scattered in the maize genome (Figure 3). Further, all prefoldin genes were located in the duplicated segments of chromosomes in maize. Three of 6 pairs (GRMZM2G023347/GRMZM2G070061, GRMZM2G099909/GRMZM2G102580, and GRMZM2G010065/GRMZM2G392932) were retained in the maize duplication event (Figure 3). Next, I also used Ks to estimate the evolutionary dates of duplicated prefoldin genes (Table 3). The results indicated that duplication events of maize 3/6 pairs occurred within the past 17.69–21.73 million years, consistent with the time of subsequent genome duplication events in maize (Gaut et al., 1996). In addition, the earliest segmental duplication event was also observed around 81.46 MYA in the prefoldins of maize (GRMZM2G014676/GRMZM2G119740) (Kellogg, 2001). Therefore, expansion of prefoldin genes in maize might have

### Table 3 | Functional divergence estimated in prefoldin paralogs.

| Groups comparison | \(\theta^1\) | SE\(^2\) | LRT\(^3\) | \(N(0.5)^4\) | \(N(0.8)^4\) |
|------------------|----------|----------|--------|----------|----------|
| I/II             | 0.826543 | 0.239298 | 17.32376 | 15 | 15 |
| I/III            | 0.803522 | 0.190311 | 12.43868 | 13 | 4 |
| I/V              | 0.646644 | 0.285272 | 8.070123 | 15 | 7 |
| I/IV             | 0.027228 | 0.201279 | 1.794084 | 1 | 1 |
| V/VI             | 1.105756 | 0.246338 | 13.44324 | 15 | 14 |
| V/VI             | 1.689677 | 0.300763 | 31.56156 | 15 | 15 |
| V/II             | 1.053435 | 0.198181 | 25.42036 | 15 | 15 |
| V/IX             | 0.014082 | 0.230382 | 1.732052 | 1 | 1 |
| V/III            | 0.765629 | 0.156564 | 6.121142 | 6 | 3 |
| III/IV           | 0.960633 | 0.189864 | 18.39889 | 15 | 8 |
| III/V            | 0.323706 | 0.276028 | 3.636844 | 6 | 2 |
| II/VI            | 0.407484 | 0.300177 | 0.682597 | 1 | 0 |
| II/IV            | 0.977057 | 0.268244 | 13.87541 | 15 | 15 |
| II/II            | 0.811230 | 0.216721 | 13.90022 | 15 | 6 |
| II/IX            | −0.278640 | 0.410256 | 0.078849 | 0 | 0 |
| III/IV           | 0.514812 | 0.177764 | 6.660248 | 4 | 1 |
| III/VI           | 0.507173 | 0.345447 | 0.811551 | 1 | 0 |
| III/VI           | 0.796963 | 0.304516 | 8.12796 | 15 | 13 |
| V/III            | 0.187283 | 0.378168 | 1.031081 | 3 | 0 |
| V/IV             | 0.996719 | 0.203875 | 19.93697 | 15 | 14 |
| IV/III           | 0.199838 | 0.371056 | 0.945486 | 1 | 0 |
| IV/II            | 0.483363 | 0.269882 | 1.506023 | 1 | 1 |
| IV/VI            | 0.563231 | 0.320949 | 6.915311 | 15 | 12 |
| VI/III           | 0.636617 | 0.225039 | 13.53771 | 15 | 11 |
| VI/VI            | 0.254421 | 0.133029 | 4.686944 | 2 | 1 |
| VI/IV            | 0.389670 | −0.851500 | 0.209424 | 0 | 0 |
| VI/VI            | 1.083898 | 0.393359 | 7.594005 | 15 | 15 |
| VII/VI           | 1.951755 | 0.280780 | 12.16217 | 15 | 15 |
| VII/IX           | 0.844957 | 0.220248 | 10.71504 | 14 | 4 |
| V/IX             | −0.619080 | −0.380780 | 2.643254 | 0 | 0 |
| V/VI             | 0.515722 | 0.419894 | 1.903136 | 13 | 2 |
| VI/II            | 0.932783 | 0.312437 | 10.24838 | 15 | 15 |
| II/III           | 0.359807 | 1.001866 | 0.075892 | 0 | 0 |
| VIII/VI          | 1.021701 | 0.258359 | 12.72446 | 15 | 15 |
| VIII/IX          | 1.647830 | 0.628639 | 6.871029 | 15 | 15 |
| VIII/IX          | −0.048070 | 0.311906 | 0.876434 | 1 | 0 |

1\(\theta\): the coefficient of functional divergence. 2\(\text{SE}\): standard error. 3\(\text{LRT}\): likelihood ratio test. 4\(N(0.5)\) and \(N(0.8)\): the numbers of divergent residues when the cut-off value is 0.5 and 0.8, respectively.
occurred due to the large-scale segmental duplication events in evolution. Segmental duplication contributes to the prefoldin family gene expansion in maize.

**Functional Divergence Analyses of the Prefoldin Proteins**

To further study whether amino-acid substitutions cause adaptive functional diversification, the program DIVERGE (Gu, 1999, 2001) was used to estimate type-I functional divergence between different prefoldin groups. Thirty-six pairs of paralogous members were compared and estimated the evolutionary rate of each amino-acid sites. The results indicated that the coefficient of functional divergence ($\theta$) values between 29 group pairs was less than 1 (Table 3), indicating significant site-specific alteration in selective constraints of most members of the prefoldin group pairs. Further, I also predicted a few critical residues associated with the functional divergence. For example, about 15 critical sites were predicted in Group I/II, I/VII, I/VIII, V/VI, V/VII, VII/VIII, and VII/IX pairs, while no site was predicted in the Group II/IX, IV/IX, and V/IX pairs. An example of detailed distribution of the functional divergence sites of Group VI/VIII pairs was shown in Figure 4. Further analyses indicated that one divergence site is located in the first residue. Other fourteen predicted sites are located in the N-terminal of $\alpha$-helice 2 ($\alpha$2). As we know, prefoldin can form a heterohexameric jellyfish like structure with coiled-coil tentacles (Siegert et al., 2000). These tentacles are involved in substrate binding. In this study, I found that some amino-acid sites on $\alpha$-helice 2 (tentacle) present different sequence composition, implying an increase in substrate specificity for target proteins. Higher theta values ($\theta$) exit in Group I/VI, I/VII, I/VIII, V/VI, V/VII, VII/VIII, and VII/IX pairs (Table 3), suggesting a higher evolutionary rate between them. Thus, due to the different evolutionary rates, the prefoldin genes diverge functionally from each other. The amino-acid mutations promoted the functional evolution and divergence of prefoldin genes, as an adaptation of the species to the changing environment (Cao et al., 2011).

**Functional Network Analysis of the Maize Prefoldins**

Genes involved in the same biological process are often coordinately expressed, and thus their co-expression information will be as a factor to understand the biological process (Eisen et al., 1998). To further explore which genes are possibly interacted by maize prefoldin, I performed a coexpression analyses with the COB$^{11}$ (Schaefer et al., 2014). I found that 7 of 13 maize prefoldins were co-expressed in this network, resulting in 347 interactions exhibited by 78 genes (Figure 5). Among the 347 interactors identified, 23 and 17 genes were coexpressed with GRMZM2G014676 and GRMZM2G135691, respectively. GRMZM1G014676 is an orthologous gene of Arabidopsis PFD5 (At5g23290), while GRMZM2G135691 is an orthologous gene of Arabidopsis PFD1 (At2g07340) (Figure 1). Moreover, PFD5 plays an essential role in Arabidopsis tolerance to salt stress (Rodríguez-Milla and Salinas, 2009). They may also have similar biology functions. The network analysis of

---

$^{11}$http://csbio.cs.umn.edu/cob
FIGURE 5 | Prefoldin genes interaction network. Seven prefoldin genes are mapped to the co-expression database and reveal a total of 78 unique genes that showed 347 interactions in maize.

FIGURE 6 | Expression profiles of the maize prefoldin genes in 20 different developmental stages.
maize prefoldins has a reference value for the functional research of prefoldin genes. Some of these interactors include carrier proteins, such as equilibrative nucleoside transporter (ENT), mitochondrial phosphate carrier (MPC), acyl carrier protein (ACP), and so on. ENT (GRMZM2G002391) mediated nucleosides transport across plasma membrane influenced Arabidopsis growth and pollen germination (Bernard et al., 2011). MPC (GRMZM2G118208) is a mitochondrial solution carrier protein, which delivers phosphate across the inner mitochondrial membrane and function in plant development and resistance (Haferkamp and Schmitz-Esser, 2012). As a small acidic protein with conserved Asp-Ser-Leu (DSL) motif, ACP (GRMZM2G149580) functions in root nodule symbiosis in soybean (Wang et al., 2014). I also found that another interactor, an auxin binding protein (ABP1), is required for binding to auxin at low concentration and involved in the embryogenesis, auxin signaling and cell division (Chen X et al., 2014; Xu et al., 2014). In addition, an Arabidopsis cyclin-dependent kinase (CDK) G2 regulated the salinity stress response and was associated with the control of flowering time (Ma et al., 2015). Coexpression analyses reflect the correlation of the expression profiles of different genes, and are suggestive in tracing the genes in the same biological process or pathway. My results indicated that some proteins possessed carrying, binding, and kinase function were usually coexpressed with maize prefoldins (Figure 5). The prefoldin genes may be involved in some related molecular processes by interacting with these interactors. Thus, whether these interactors could function with prefoldin need further experimental verification. The approaches and results open a new way to explore the molecular mechanism of prefoldin dynamics. Also, this interaction can also help to study the key regulatory steps in these processes or pathways. It will be advantageous to screen candidate genes for further functional researches.

Expression Patterns and Promoter Sequence Analysis of the Maize Prefoldin Genes

The gene expression patterns may provide important clues to their functions (Cao and Shi, 2012). Here, I first tested spatial and temporal specific expression patterns of maize prefoldin genes in 20 tissues using microarray data. I detected 10 probes of the maize prefoldin transcripts, and other three transcripts (GRMZM2G049390, GRMZM2G535911, and GRMZM2G014676) with no detectable probes were not
further analyzed in this study. The results showed that the expression abundance of the prefoldin genes detected presented great variability in maize (Figure 6). Some prefoldins (AC203985.4_FGT005, GRMZM2G135354, and GRMZM2G392932) expressed exclusively in developing embryo and endosperm may improve seed yield and quality. Notably, due to the great expression in maize silk, GRMZM2G135691 may play a key role in silk development.

Next, I also investigated the responses of maize prefoldin genes under drought, low temperature and salt stresses using 10-day-old seedlings. Quantitative RT-PCR was used to analyze the expression profiles of 13 prefoldin genes in maize under various stresses (Figure 7). The expression of GRMZM2G135354 and GRMZM2G049390 was significantly upregulated during drought, whereas other prefoldin genes were downregulated. The expression of over 84 percent of maize prefoldins was significantly repressed under salt, whereas GRMZM2G135354 and GRMZM2G049390 were strongly induced, suggesting that the prefoldin genes might present different responses to salinity or drought stress. Similar results also occurred with low-temperature. Under low temperature stress, most prefoldin genes were significantly downregulated (Figure 7). For instance, the expression level of GRMZM2G135691 decreased about 50 times. However, the expression of some prefoldin genes (GRMZM2G010065, and GRMZM2G049390) was induced. Previous study has shown that osmotic stress caused by drought and salinity can change microtubule orientation, and then regulates primary root elongation in Arabidopsis (Liu et al., 2014). Moreover, low temperature also leads to the depolymerization of microtubule (Bartolo and Carter, 1991). My expression profiles of maize prefoldin genes under these stresses indicated that most prefoldin members were significantly repressed. Therefore, it can be inferred that under the drought, salinity, or low temperature stress, the decreased expression of prefoldin genes may inhibit or change the development of microtubules. Next, the auxin response of these maize prefoldin genes was also investigated. Among the 13 prefoldin genes detected, three members (GRMZM2G392932, GRMZM2G014676, and GRMZM2G049390) were upregulated during IAA treatment, suggesting important roles in regulating IAA response. Nevertheless, the expression of GRMZM2G135354 and GRMZM2G355911 was unchanged after IAA treatment, whereas others were steadily decreased during the treatment (Figure 7).

To further explore the potential regulatory mechanism of maize prefoldin genes in phytohormone and abiotic stress stimuli, I used the PLACE web server12 (Higo et al., 1999) to identify their putative cis-elements. Here, 1,000-bp promoter regions of the maize prefoldin genes were analyzed. A few phytohormones and abiotic stress regulatory elements were found. They included auxin response factor (ARF)-binding site (S000270); tissue-specific expression and auxin-inducible (S000273); auxin-inducible (S000370); IAA/SA-inducible (S000024); SA-inducible (S000390); drought-inducible (S000414 and S000174);

12http://www.dna.affrc.go.jp/PLACE/
cold/drought-inducible (S000153); and salt-inducible (S000453) response regulatory elements. I also found that all maize prefoldin genes contained multiple regulatory elements in their promoter regions (Figure 8), suggesting that these phytohormones and abiotic stresses regulated the expression of maize prefoldin genes. Comparing the distribution of nine cis-elements in their promoter sequences, I found variation between all sister pairs of maize prefoldin genes, implying divergent expression profiles in the duplicated genes. It may be the reason of subfunctionalization or neofunctionalization (Prince and Pickett, 2002).

**SUMMARY**

I performed a comparative analysis of the plant prefoldin gene family in this study, and found that the prefoldin family could be divided into nine groups by phylogenetic analyses. Gene organization and motif compositions of the prefoldin members were highly conserved in each group, implying their functional conservation. Segmental duplication contributed to maize prefoldin family gene expansion. A few critical residues associated with the functional divergence were predicted. Functional network analyses identified some co-expressed genes, which usually have binding, carrying and kinase activity. Furthermore, the differential expression profiles of the maize prefoldin genes suggested functional divergence during development and IAA and some abiotic stress treatments. Some cis-elements responsive to abiotic stress and phytohormone were also found in the upstream sequence of the maize prefoldin genes. These data will provide useful insights for further functional investigation of this gene family in the future.

**AUTHOR CONTRIBUTIONS**

JC designed and carried out the experiments and wrote the manuscript. The author read and approved the manuscript.

**ACKNOWLEDGMENTS**

This project is supported by grants from the National Science Foundation of China (No. 31100923), the National Science Foundation of Jiangsu Province (BK2011467), the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), and Jiangsu University “Youth Backbone Teacher Training Project” (2012–2016).

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00317

---

**REFERENCES**

Bailey, T. L., Williams, N., Misleh, C., and Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34, W369–W373. doi: 10.1093/nar/gkl198

Bartolo, M. E., and Carter, J. V. (1991). Microtubules in mesophyll cells of nonacclimated and cold-acclimated spinach: visualization and responses to freezing, low temperature, and dehydration. *Plant Physiol.* 97, 175–181. doi: 10.1104/pp.97.1.175

Bernard, C., Traub, M., Kunz, H. H., Hach, S., Trentmann, O., and Möhlmann, T. (2011). Equilibrative nucleoside transporter 1 (ENT1) is critical for pollen germination and vegetative growth in *Arabidopsis*. *J. Exp. Bot.* 62, 4627–4637. doi: 10.1093/jxb/err183

Cao, J. (2012). The pectin lyases in *Arabidopsis thaliana*: evolution, selection and expression profiles. *PLoS ONE* 7:e46944. doi: 10.1371/journal.pone.0046944

Cao, J., Huang, J., Yang, Y., and Hu, X. (2011). Analyses of the oligopeptide transporter gene family in poplar and grape. *BMC Genomics* 12:465. doi: 10.1186/1471-2164-12-465

Cao, J., and Li, X. (2015). Identification and phylogenetic analysis of late embryogenesis abundant proteins family in tomato (*Solanum lycopersicum*). *Planta* 241, 757–772. doi: 10.1007/s00425-014-2215-y

Cao, J., Li, X., Lv, Y., and Ding, L. (2015). Comparative analysis of the phytoctyanin gene family in 10 plant species: a focus on *Zea mays*. *Front. Plant Sci.* 6:515. doi: 10.3389/fpls.2015.00515

Cao, J., and Shi, F. (2012). Evolution of the RALF gene family in plants: gene duplication and selection patterns. *Evol. Bioinform. Online* 8, 271–292. doi: 10.4137/EBO.S9652

Cao, S., Carlsson, G., Osipowich, A. B., Llanes, J., Lin, Q., Hoek, K. L., et al. (2008). Subunit 1 of the prefoldin chaperone complex is required for lymphocyte development and function. *J. Immunol.* 181, 476–484. doi: 10.4049/jimmunol.181.1.476

---

Chen, X., Grandont, L., Li, H., Hauschild, R., Paque, S., Abuzeineh, A., et al. (2014). Inhibition of cell expansion by rapid ABP1-mediated auxin effect on microtubules. *Nature* 516, 90–93. doi: 10.1038/nature13889

Chen, Y., and Cao, J. (2014). Comparative genomic analysis of the Sm gene family in rice and maize. *Gene* 539, 238–249. doi: 10.1016/j.gene.2014.02.006

Chen, Y., Hao, X., and Cao, J. (2014). Small auxin upregulated RNA (SAUR) gene family in maize: identification, evolution, and its phylogenetic comparison with *Arabidopsis*, rice, and sorghum. *J. Integr. Plant Biol.* 56, 133–150. doi: 10.1111/jipb.12127

Cameron, J. M. (1999). K-Emulator: calculation of the number of nucleotide substitutions per site and the confidence intervals. *Bioinformatics* 15, 763–764. doi: 10.1093/bioinformatics/15.9.763

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14863–14868. doi: 10.1073/pnas.95.25.14863

Gaut, B. S., Morton, B. R., McCaig, B. C., and Clegg, M. T. (1996). Substitution rate comparisons between grasses and palms: synonymous rate differences at the Adh parallel rate differences at the plastid gene rbcL. *Proc. Natl. Acad. Sci. U.S.A.* 93, 10274–10279. doi: 10.1073/pnas.93.19.10274

Geissler, S., Siegers, K., and Schiebel, E. (1998). A novel protein complex promoting formation of functional alpha- and gamma-tubulin. *EMBO J.* 17, 952–966. doi: 10.1093/emboj/17.4.952

Gu, X. (1999). Statistical methods for testing functional divergence after gene duplication. *Mol. Biol. Evol.* 16, 1664–1674. doi: 10.1093/oxfordjournals.molbev.a026080

Gu, X. (2001). Maximum-likelihood approach for gene family evolution under functional divergence. *Mol. Biol.*
Evil. 18, 453–464. doi: 10.1093/oxfordjournals.molbev.a03824
Gu, Y., Deng, Z., Paredes, A. R., DeBolt, S., and Wang, Z. Y. (2008). Prefoldin 6 is required for normal microtubule dynamics and organization in Arabidopsis. Proc. Natl. Acad. Sci. U.S.A. 105, 18064–18069. doi: 10.1073/pnas.0808018105
Haferkamp, L. and Schmitz-Esser, S. (2012). The plant mitochondrialcarrier family: functional and evolutionary aspects. Front. Plant Sci. 3:2. doi: 10.3389/fpls.2012.00002
Hartl, F. U. (1996). Molecular chaperones in cellular protein folding. Nature 381, 571–579. doi: 10.1038/381571a0
Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and protein degradation. Nature 475, 324–332. doi: 10.1038/nature10137
Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res. 27, 297–300. doi: 10.1093/nar/27.1.297
Hill, J. E., and Hemmingsen, S. M. (2001). Arabidopsis thaliana CYCLIN-DEPENDENT KINASE G2 regulates salinity stress response and salt tolerance. Mol. Plant. 2, 526–534. doi: 10.1093/mpspf016
Schaefer, R. J., Brikshne, R., Springer, N. M., and Myers, C. L. (2014). Discovering functional modules across diverse maize transcriptomes using COB, the Co-expression Browser. PLoS ONE 9:e99193. doi: 10.1371/journal.pone.0099193
Schmutz, J., Cannon, B. S., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. Nature 463, 178–183. doi: 10.1038/nature08670
Sekhon, R. S., Lin, H., Childs, K. L., Hansey, C. N., Buell, C. R., de Leon, N., et al. (2011). Genome-wide atlas of transcription during maize development. Plant J. 66, 553–563. doi: 10.1111/j.1365-313X.2011.04527.x
Siegers, K., Waldmann, T., Leroux, M. R., Grein, K., Shcchevchenko, A., Schiebel, E., et al. (1999). Compartmentation of protein folding in vivo: sequestration of non-native polypeptide by the chaperonin-GimC system. EMBO J. 18, 75–84. doi: 10.1093/emboj/cbo18.1.75
Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F. U., and Moarefi, I. (2000). Structure of the molecular chaperone prefoldin: unique interaction of multiple coiled coil tertiaries with unfolded proteins. Cell 103, 621–632. doi: 10.1016/S0092-8674(00)01653-3
Soderlund, C., Bomhoff, M., and Nelson, W. M. (2011). SyMAP v3.4: a turnkey synteny system with application to plant genomes. Nucleic Acids Res. 39:e68. doi: 10.1093/nar/gkr123
Sturm, A., Quackenbush, J., and Trajanowski, Z. (2002). Genesis: cluster analysis of microarray data. Bioinformatics 18, 207–208. doi: 10.1093/bioinformatics/18.1.207
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739. doi: 10.1093/molbev/msr121
Vanier, I. E., Lewis, S. A., Rommelaere, H., Ampe, C., Vandenkerckhove, J., Klein, H. L., et al. (1998). Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. Cell 93, 863–873. doi: 10.1016/S0092-8674(00)81446-4
Wang, J., Tóth, K., Tanaka, K., Nguyen, C. T., Yan, Z., Brechenmacher, L., et al. (2014). A soybean acyl carrier protein, GmACP, is important for root nodule symbiosis. Mol. Plant Microbe Interact. 27, 415–423. doi: 10.1093/mp/27.4.415
Xia, K., Wang, R., Ou, X., Fang, Z., Tian, C., Duan, J., et al. (2012). OsTIR1 and OsAFB2 downregulation leads to more tillers, early flowering and less tolerance to salt and drought in rice. PLoS ONE 7:e30039. doi: 10.1371/journal.pone.0030039
Xu, T., Dai, N., Chen, J., Nagawa, S., Cao, M., Li, H., et al. (2014). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based
on n-peptide compositions. *Protein Sci.* 13, 1402–1406. doi: 10.1110/ps.03479604
Zhao, H., Wang, Q., Zhang, H., Liu, Q., Du, X., Richter, M., et al. (2005). UXT is a novel centrosomal protein essential for cell viability. *Mol. Biol. Cell* 16, 5857–5865. doi: 10.1091/mbc.E05-08-0705

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AB and the handling Editor declared a shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 Cao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.