

Article

## Systematic Analysis of the Maize PHD-Finger Gene Family Reveals a Subfamily Involved in Abiotic Stress Response

Qianqian Wang †, Jinyang Liu †, Yu Wang, Yang Zhao, Haiyang Jiang \* and Beiji Cheng \*

Key Laboratory of Crop Biology of Anhui Province, Anhui Agricultural University, No. 130, Changjiang West Road, Hefei 230036, China; E-Mails: qqwangahau@163.com (Q.W.); liujinyang315@163.com (J.L.); yuwangahau@163.com (Y.W.); zhaoyang2013@ahau.edu.cn (Y.Z.)

† These authors contributed equally to this work.

\* Authors to whom correspondence should be addressed;

E-Mails: hyjiang@ahau.edu.cn (H.J.); beijiucheng@ahau.edu.cn (B.C.);

Tel./Fax: +86-551-6578-6021 (H.J. & B.C.).

Academic Editor: Jianhua Zhu

Received: 20 July 2015 / Accepted: 16 September 2015 / Published: 30 September 2015

---

**Abstract:** Plant homeodomain (PHD)-finger proteins were found universally in eukaryotes and known as key players in regulating transcription and chromatin structure. Many PHD-finger proteins have been well studied on structure and function in animals. Whereas, only a few of plant PHD-finger factors had been characterized, and majority of PHD-finger proteins were functionally unclear. In this study, a complete comprehensive analysis of maize PHD family is presented. Sixty-seven PHD-finger genes in maize were identified and further divided into ten groups according to phylogenetic analysis that was supported by motif and intron/exon analysis. These genes were unevenly distributed on ten chromosomes and contained 12 segmental duplication events, suggesting that segmental duplications were the major contributors in expansion of the maize PHD family. The paralogous genes mainly experienced purifying selection with restrictive functional divergence after the duplication events on the basis of the  $K_a/K_s$  ratio. Gene digital expression analysis showed that the PHD family had a wide expression profile in maize development. In addition, 15 potential stress response genes were detected by promoter *cis*-element and expression analysis. Two proteins ZmPHD14 and ZmPHD19 were located in the nucleus. These results provided a solid base for future functional genome study of

the PHD-finger family in maize and afforded important clues for characterizing and cloning potentially important candidates in response to abiotic stresses.

**Keywords:** PHD family; genome-wide analysis; duplication; maize; abiotic stress

---

## 1. Introduction

The zinc-finger protein is a kind of protein with a “finger” domain, and the zinc-finger domains are rich in cysteine or histidine. Conservative cysteine residues and histidine residues can stabilize the normal space structure by combining zinc ions. According to the arrangement characteristics of conservative cysteine residues and histidine residues, Zinc-finger proteins are divided into several types including Really Interesting New Genes (RING), LIM (Lin11, Isl-1 and Mec-3), plant homeodomain (PHD), *etc.* [1]. The plant homeodomain (PHD) fingers are widely distributed and conserved throughout eukaryotes and most proteins are located in the nucleus, which has been found from yeast to man [2]. The PHD proteins contain one or more PHD-finger domains those are composed of about 60 amino acids with the structural characteristic of Cys4-His-Cys3, which is similar to RING and LIM domains [2,3]. The number of bases between the cysteine and histidine are relatively conserved, as is that between the cysteine residues, and the second amino acid residue before the last pair of cysteines is typically an aromatic amino acid such as tryptophan [4]. These core amino acids residues combine with two zinc ions and are separated by three loop regions [5]. Recent studies show that the PHD-finger domain can specifically bind to histone H3 trimethylated at lysine [6].

Since the first PHD-finger protein HAT3.1 was found in *Arabidopsis* [5], more and more PHD-finger proteins were identified as involved in a number of physiological and biochemical process in fungus, animals or plants, including the regulation of chromatin structure and transcription, *etc.* [7]. Through combination with nucleosomal histones, some PHD-finger proteins regulate the chromatin state [8]. For example, in *Arabidopsis*, the PHD-finger protein DUET was required for organization and progression of chromosomes during male meiosis [9]; VIM1, an *Arabidopsis* PHD-finger protein could associate with histone to participate in the regulation of the chromatin state [10]; MMD1, encoding a PHD domain-containing protein, is expressed preferentially in male meiocytes, which are involved in chromatin remodeling and transcriptional events required for successful progression during meiosis [11]. The latest research shows that PHD proteins play important roles in methylation maintenance [12,13]. The mammals PHD protein UHRF1 is crucial in the maintenance of DNA methylation [14] and ATX1 and ATX2 with the activity of histone methyltransferase regulated the transcription of certain genes [15]. The PHD-finger protein is reported to mediate the protein–protein interactions [16], which are already known to be involved in various human diseases [17], for example, UHRF1, a PHD protein plays an important role in breast carcinogenesis by forming a complex with other proteins [18]. In addition, some PHD-fingers were considered to take part in the regulation of ubiquitination [19], and were detected as binding targets of rare phosphatidylinositol phosphates (*PtdInsPs*) which play crucial roles in DNA damage signaling [20]. Furthermore, a number of PHD proteins play significant roles in plant response to abiotic stresses; six soybean PHD proteins were

identified as regulators of abscisic acid (ABA) signaling and conferred salt tolerance in transgenic *Arabidopsis* [21].

Although the number of PHD-finger homologs identified is increasing in different species, most putative PHD family members remain unrecognized and uncharacterized in each organism, and few comprehensive analyses of the PHD family in evolution have been conducted. It would be of great significance to get a better understanding of the multiple PHD family members, because of their potentially important functions in many central biological processes. Maize is an important cereal crop and has become a model plant for the study of genetics, evolution and other basic biological research [22]. The availability of the maize genome sequences has provided an excellent opportunity for whole-genome annotation, classification and comparative genomics research [23,24]. In this paper, we performed a comprehensive analysis for PHD family in maize genome, including the phylogenetic relationship, gene structure, chromosomal localization and gene duplication analysis. Seventeen maize PHD-finger factors that might participate in the process of various abiotic stresses were identified, and their expression patterns were detected in response to ABA, PEG and NaCl treatment. All these results provide a better understanding of the maize PHD-finger family on evolutionary history and functional mechanisms.

## 2. Results

### 2.1. Identification of PHD Proteins in Maize

To identify the *PHD* genes in the maize genome, the consensus protein sequences of the PHD-finger Hidden Markov Model (HMM) profile were employed as a query to search against the maize genome database with the BLASTP program. A total of 95 candidate PHD-finger protein sequences were identified in maize. Based on sequences similarity analysis, 28 redundant PHD-finger sequences were discarded. In order to confirm the reliability of putative PHD-finger members in maize, the amino acid sequences of the 67 proteins were searched for the presence of PHD domains with Pfam and SMART programs. The results showed that all of the 67 non-redundant PHD-finger proteins in maize contained the PHD conserved domain. Subsequently, the 67 PHD genes were named *ZmPHD1* to *ZmPHD67* according to their locations in chromosomes (Table 1). The length of all maize PHD proteins was between 72 and 2379 amino acids with an average of 756 amino acids. The detailed information of *ZmPHD* genes is listed in Table 1, including chromosome location, protein length (aa) and gene identifier.

**Table 1.** The detailed information of maize PHD family members.

Gene Name	Gene Identifier	Protein Size (aa)	5' End	3' End	Chromosome
<i>ZmPHD1</i>	GRMZM5G862565	1579	100,335,533	100,348,389	I
<i>ZmPHD2</i>	GRMZM5G866423	287	209,536,012	209,538,518	I
<i>ZmPHD3</i>	GRMZM5G813111	252	276,110,228	276,115,825	I
<i>ZmPHD4</i>	AC225147.4	250	292,875,307	292,879,433	I
<i>ZmPHD5</i>	GRMZM2G403562	1900	144,629,199	144,637,545	II
<i>ZmPHD6</i>	GRMZM2G013936	751	155,881,115	15,590,458	II
<i>ZmPHD7</i>	GRMZM2G047316	241	158,280,766	158,284,014	II

Table 1. Cont.

Gene Name	Gene Identifier	Protein Size (aa)	5' End	3' End	Chromosome
<i>ZmPHD8</i>	GRMZM2G409224	1566	198,098,543	198,115,319	II
<i>ZmPHD9</i>	GRMZM2G181158	345	201,127,832	201,131,690	II
<i>ZmPHD10</i>	GRMZM2G158194	556	9,042,243	9,044,511	III
<i>ZmPHD11</i>	GRMZM2G085266	971	33,200,782	32,090,831	III
<i>ZmPHD12</i>	GRMZM2G069886	1214	114,434,585	114,444,762	III
<i>ZmPHD13</i>	GRMZM2G081350	818	133,742,782	133,749,539	III
<i>ZmPHD14</i>	GRMZM2G153087	257	135,907,092	135,912,402	III
<i>ZmPHD15</i>	GRMZM2G080917	257	170,106,279	170,110,892	III
<i>ZmPHD16</i>	GRMZM2G025703	173	171,031,966	171,035,433	III
<i>ZmPHD17</i>	GRMZM2G314546	1577	197,392,618	197,405,889	III
<i>ZmPHD18</i>	GRMZM2G068331	697	10,228,455	10,233,101	IV
<i>ZmPHD19</i>	GRMZM2G008259	172	11,389,157	11,413,875	IV
<i>ZmPHD20</i>	GRMZM2G059266	558	40,354,679	40,358,758	IV
<i>ZmPHD21</i>	GRMZM5G871463	72	79,000,726	79,005,980	IV
<i>ZmPHD22</i>	GRMZM2G107807	255	120,625,503	120,629,020	IV
<i>ZmPHD23</i>	GRMZM2G067019	418	136,070,311	136,101,261	IV
<i>ZmPHD24</i>	GRMZM2G385338	1812	234,187,863	234,200,009	IV
<i>ZmPHD25</i>	GRMZM2G391413	1322	239,318,755	239,331,611	IV
<i>ZmPHD26</i>	GRMZM2G110952	1166	2,576,103	2,593,650	V
<i>ZmPHD27</i>	GRMZM2G087482	1376	7,443,884	7,459,423	V
<i>ZmPHD28</i>	GRMZM2G466292	527	65,425,225	65,434,544	V
<i>ZmPHD29</i>	GRMZM2G466270	1290	65,442,955	65,456,381	V
<i>ZmPHD30</i>	GRMZM2G063864	255	174,409,725	174,413,665	V
<i>ZmPHD31</i>	GRMZM2G045544	727	203,646,189	203,650,857	V
<i>ZmPHD32</i>	GRMZM2G368206	1465	210,502,142	210,518,679	V
<i>ZmPHD33</i>	GRMZM2G039895	555	1,537,396	1,541,397	VI
<i>ZmPHD34</i>	GRMZM2G134214	849	54,775,354	54,785,293	VI
<i>ZmPHD35</i>	GRMZM2G103230	1147	58,263,970	58,269,763	VI
<i>ZmPHD36</i>	GRMZM2G473258	321	94,516,049	94,518,519	VI
<i>ZmPHD37</i>	GRMZM2G168249	808	107,094,897	107,103,424	VI
<i>ZmPHD38</i>	GRMZM2G330024	1108	109,391,739	109,405,354	VI
<i>ZmPHD39</i>	GRMZM2G148810	253	124,705,685	124,710,238	VI
<i>ZmPHD40</i>	GRMZM2G016817	256	147,643,515	147,650,880	VI
<i>ZmPHD41</i>	GRMZM2G412492	999	8,826,124	8,831,882	VII
<i>ZmPHD42</i>	GRMZM2G097726	219	13,348,616	13,356,443	VII
<i>ZmPHD43</i>	GRMZM2G013794	976	72,242,570	72,287,055	VII
<i>ZmPHD44</i>	GRMZM2G091265	212	105,907,436	105,911,988	VII
<i>ZmPHD45</i>	GRMZM2G316191	2379	150,159,272	150,178,742	VII
<i>ZmPHD46</i>	GRMZM2G128176	1712	171,564,760	171,575,198	VII
<i>ZmPHD47</i>	GRMZM2G434715	771	174,765,898	174,774,731	VII

Table 1. Cont.

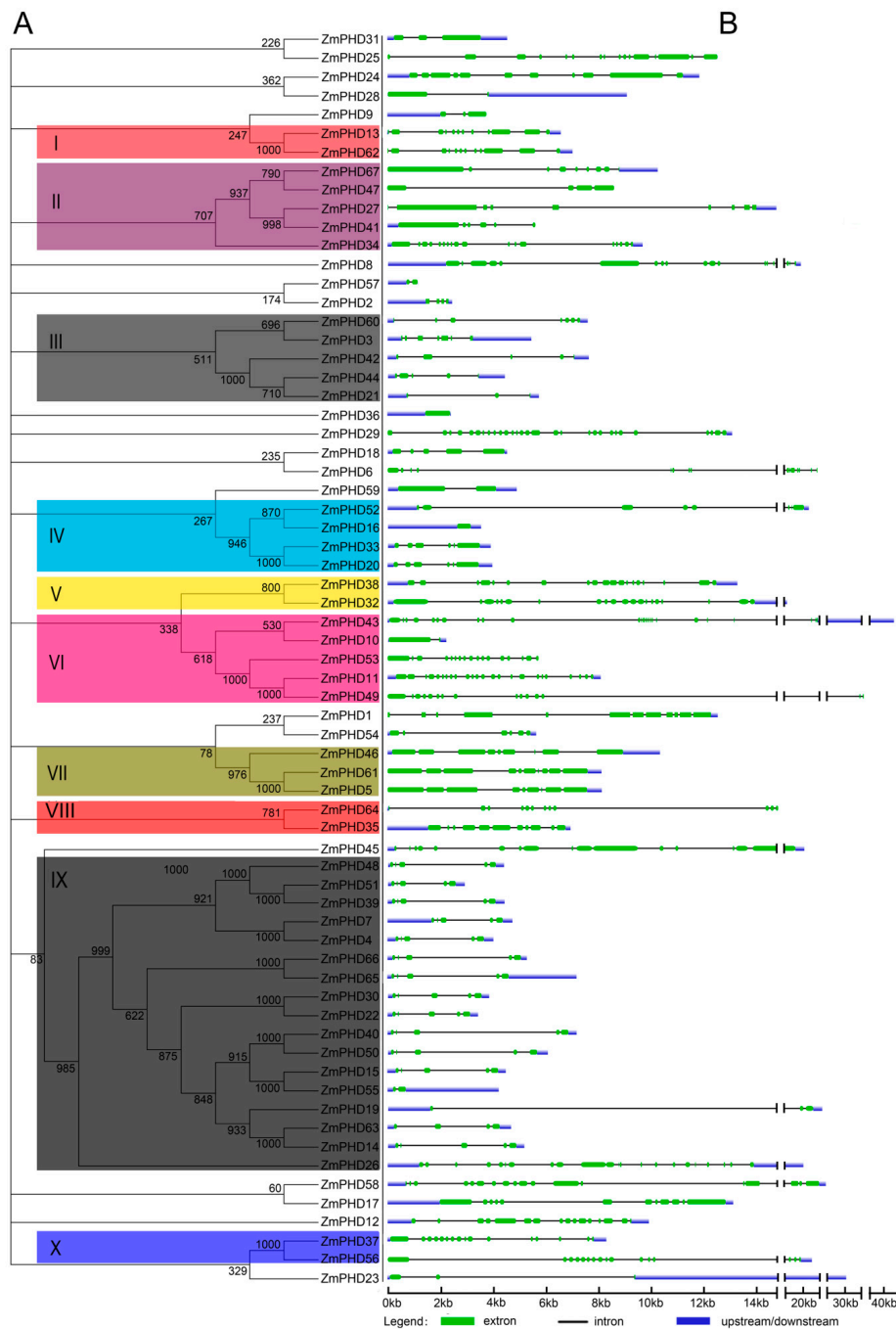
Gene Name	Gene Identifier	Protein Size (aa)	5' End	3' End	Chromosome
<i>ZmPHD48</i>	GRMZM5G893976	241	15,274,043	15,278,560	VIII
<i>ZmPHD49</i>	GRMZM2G372928	759	25,354,740	25,391,185	VIII
<i>ZmPHD50</i>	GRMZM2G158918	256	103,454,290	103,460,620	VIII
<i>ZmPHD51</i>	GRMZM2G017142	253	125,536,101	125,539,093	VIII
<i>ZmPHD52</i>	GRMZM2G335720	836	137,142,057	137,164,518	VIII
<i>ZmPHD53</i>	GRMZM2G170412	870	147,718,464	147,724,348	VIII
<i>ZmPHD54</i>	GRMZM2G149587	381	156,906,971	156,912,759	VIII
<i>ZmPHD55</i>	GRMZM2G172001	255	164,744,187	164,748,517	VIII
<i>ZmPHD56</i>	GRMZM2G455243	852	12,345,652	12,367,654	IX
<i>ZmPHD57</i>	GRMZM2G156129	108	29,157,055	29,158,213	IX
<i>ZmPHD58</i>	GRMZM2G472428	1990	30,050,060	30,074,947	IX
<i>ZmPHD59</i>	GRMZM5G889372	868	136,475,636	136,480,652	IX
<i>ZmPHD60</i>	GRMZM2G178072	249	152,186,947	152,194,716	IX
<i>ZmPHD61</i>	GRMZM2G314661	1869	4,871,817	4,880,152	X
<i>ZmPHD62</i>	GRMZM2G365888	819	27,266,656	27,273,862	X
<i>ZmPHD63</i>	GRMZM2G115424	245	31,307,377	31,312,172	X
<i>ZmPHD64</i>	GRMZM2G404426	421	65,737,118	65,752,360	X
<i>ZmPHD65</i>	GRMZM2G156088	249	119,469,068	119,473,971	X
<i>ZmPHD66</i>	GRMZM2G050495	248	119,702,189	119,707,621	X
<i>ZmPHD67</i>	GRMZM2G038050	1339	149,547,787	149,558,325	X

## 2.2. Phylogenetic and Structural Analysis

To gain further insights into the evolutionary relationships of maize *PHD* genes, we constructed an unrooted phylogenetic tree with 67 maize PHD-finger proteins sequences (Figure 1A), and investigated the exon/intron structures of individual *ZmPHD* genes by comparing their cDNA sequences with corresponding genomic DNA sequences (Figure 1B). To facilitate the research and analysis, the *ZmPHD* family was divided into 10 groups (group I to group X) according to the bootstrap values (>500) in the Neighbor-Joining (NJ) tree, while some *ZmPHD* members were beyond the 10 groups because of low bootstrap values (<500) of the NJ tree, which was also shown in previous study [24]. Group IX contained 17 members and was the largest clade of all the groups, which represented the 25.37% of the total *ZmPHD* proteins. Whereas, groups I, V, VIII and X only had two members.

Observation of the exon/intron structure in maize *PHD* genes (Figure 1B) revealed the number of introns varied from 1 to 28, with the exception of *ZmPHD16* and *ZmPHD36* having no introns, and *ZmPHD29* having the most introns (28) among the *ZmPHD* family. In addition, most maize *PHD* genes clustered in the same group shared highly similar exon/intron distribution patterns, including the exon length and intron numbers. For example, two genes locating in group X both had 15 introns, and the members belonging to group VII contained eight introns. Likewise, most genes in group IX possessed four introns except *ZmPHD19*, -26, -55, -63, and -65, which had 2, 19, 1, 3 and 5 introns, respectively. By comparison, the genes in groups II and III showed great diversity in exon length as

well as the intron numbers (Figure 1B). In summary, the exon/intron structures of maize *PHD* genes were basically in line with the phylogenetic relationship.



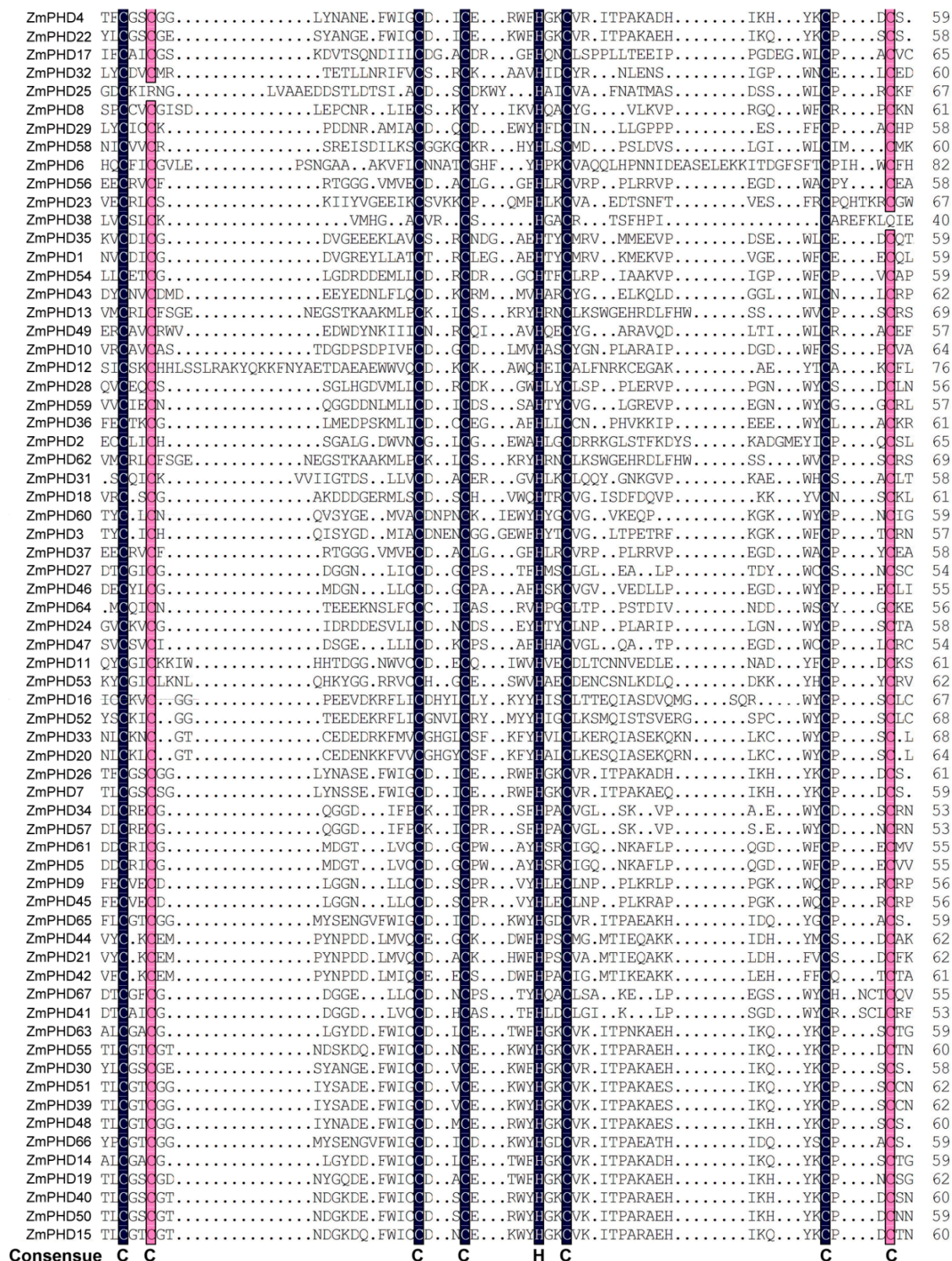
**Figure 1.** Phylogenetic relationships and gene structure of maize *PHD* genes. (A) An unrooted tree is generated with the MEGA4.0 software using the full-length amino acid sequences of the 67 maize *PHD* proteins by the Neighbor-Joining (NJ) method, with 1000 bootstrap replicates. The tree shows ten major phylogenetic groups (group I to X) indicated with different colored backgrounds; and (B) Exon/intron organization of maize *PHD* genes. Green boxes indicate exons and grey lines represent introns, and the untranslated regions (UTRs) are indicated by blue boxes. The sizes of exons and introns can be estimated using the scale at the bottom.

### 2.3. The Sequence Analysis of Maize PHD-Finger Domain and Motifs

To have a further understanding of the similarity between the maize PHD-finger domains, we aligned 67 maize PHD-finger domain sequences (Figure 2). The length of the 67 sequences varied from 40 to 60 amino acids, and the highlighted figures in black and red areas were the Zn ion binding sites with seven cysteines and one histidine residue. The PHD-finger domain was conservative because of the Zn ion binding sites contributing to the structural stability of domain. Based on the result of sequence alignment, we found the maize PHD-finger domain consensus sequence was C-X(1-2)-C-X(8-19)-C-X(2-4)-C-X(4-6)-H-X2-C-X(11-26)-C-X(2-3)-C, which was basically consistent with the previous findings [25].

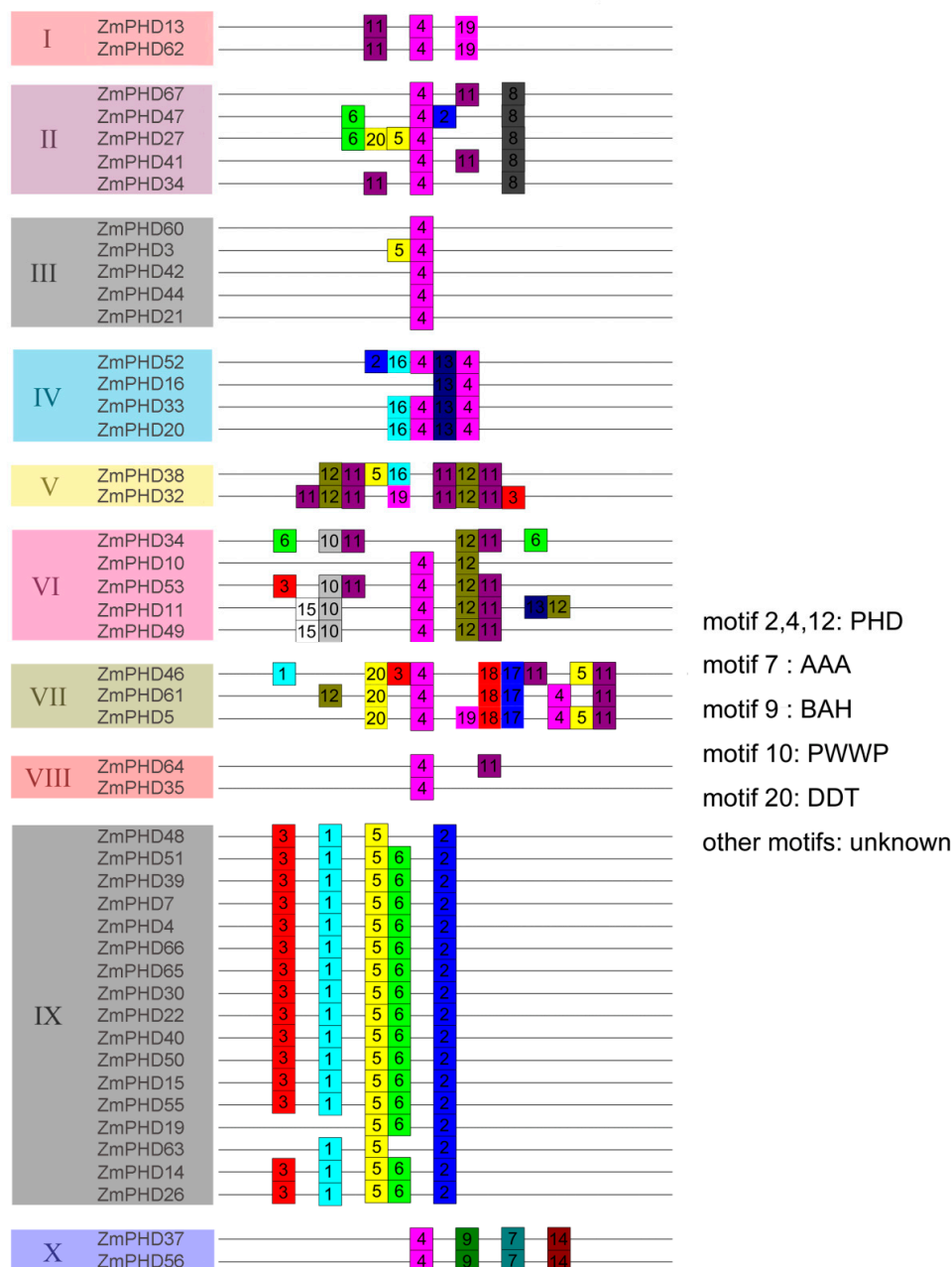
To gain more insight into the diversity of motif compositions among ZmPHDs, maize PHD proteins from group I to X were subjected to the MEME tool. Twenty conserved motifs that designated as motif 1 to motif 20 in ZmPHD proteins were identified (Figure 3). The details of each motif are shown in Table S1. In addition, the SMART and Pfam were employed to annotate the 20 identified putative motifs (Figure 3). Furthermore, to make the MEME results concise and clear, two or more adjacent motifs, representing the same domain, were merged and displayed as one domain district. Of the 20 motifs, half remained as function unknown in SMART and Pfam databases. Motifs 2, 4 and 12 stood for the conserved PHD domain and were found one or more times among all the ten group members. Most maize PHD proteins within the same groups shared highly similar motif compositions and distribution, which implied the ZmPHD members within the same groups might share similar functions. However, great difference were also observed between different groups. For example, proteins in group I possessed motifs 4, 9 and 11, while group IX members contained motifs 1, 2, 3, 5 and 6 (Figure 3). Besides, some motifs that were exclusively observed in a particular group, suggested that these motifs might contribute to specific functions of that group. For instance, motif 9 represents the Bromo-adjacent Homology domain (BAH) domain that existed only in group X members, and shown to play an important role in DNA methylation, replication and transcriptional regulation [26]; motif 10 is a PWWP domain, named after a conserved Pro-Trp-Trp-Pro motif, only found in group VI members except the ZmPHD10, which is methyl-lysine recognition motif that is involved in a series of cellular processes by binding to histone-4 methylated at lysine-20, such as maintaining genome stability and regulating cell-cycle progression [27,28]; moreover, motif 20 representing the DNA binding homeobox and Different Transcription factors (DDT) domain was merely present in the group VII and ZmPHD27 and motif 7 representing the ATPases Associated with diverse cellular Activities (AAA) domain was only belong to group VII as well as motif 9 (Figure 3). The motif distribution among ZmPHD proteins further supported the closely evolutionary relationships among ZmPHD proteins as well as the reliability of the phylogenetic analysis.





**Figure 2.** Multiple sequence alignment of PHD-finger domain of 67 maize PHD proteins. The shading of the alignment presents identical residues in black and similar residues in red, and the high conserved amino are marked at the bottom. The PHD-finger domain of each ZmPHD members is corresponding to the first PHD motif in the upstream of each PHD protein in Figure 3, respectively.



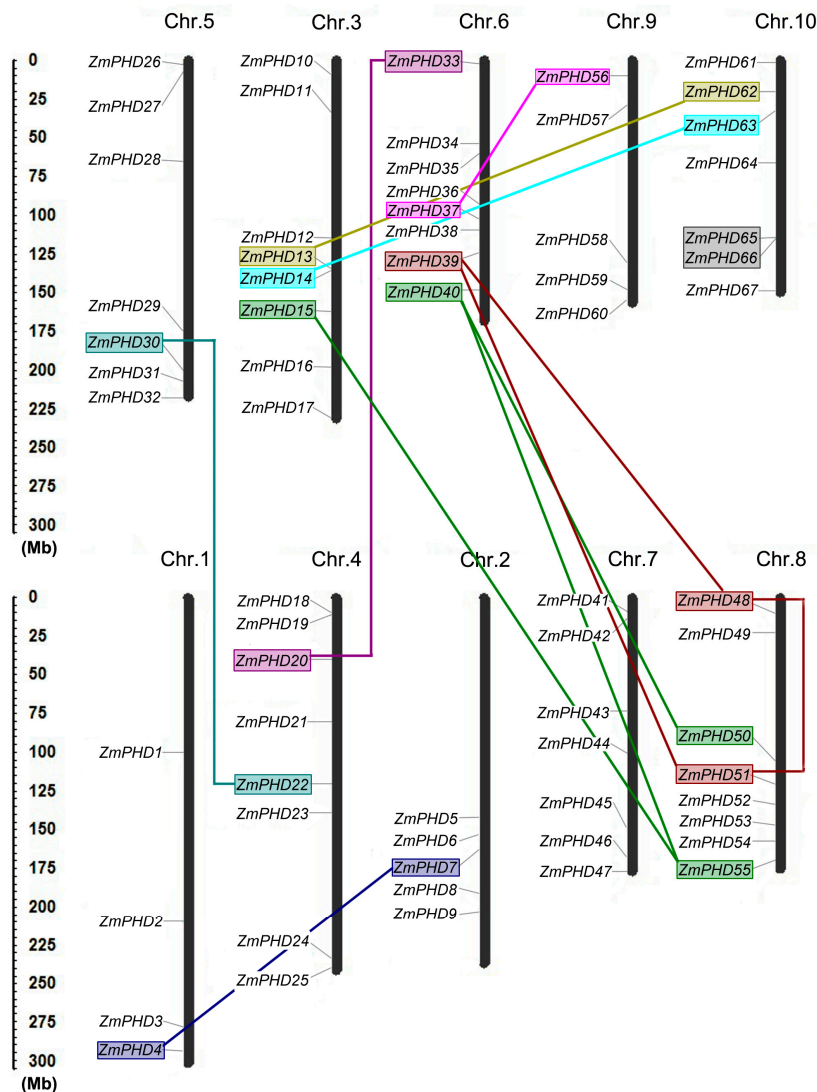


**Figure 3.** Distribution of conserved motifs in maize PHD members. All motifs were identified by MEME using the complete amino acid sequences of ZmPHD proteins. Different motifs are indicated by different colors boxes numbered 1–20, and the length of each box in the proteins does not represent the actual motif size. The annotation of each motif is listed on the right. PHD: plant homeodomain; AAA: ATPases Associated with diverse cellular Activities; BAH: Bromo-adjacent Homology domain; PWWP: Pro-Trp-Trp-Pro motif; DDT: DNA binding homeobox and Different Transcription factors.

2.4. Chromosomal Locations and Duplications of ZmPHDs

Sixty-seven maize *PHD* genes were placed on 10 maize chromosomes according to their positions in the maize genome (Figure 4), which showed that maize *PHD* genes were unevenly distributed on the 10 maize chromosomes. Chromosome 3, 6 and 8 contained highest number of *ZmPHD* genes (8);

only four members were assigned on chromosome 1. The precise location of all *PHD* genes on the ten maize chromosomes is listed in detail in Table 1. Distribution of these *PHD* genes on individual chromosome was also irregular; for example; genes on chromosome 2 were distributed on the lower end of the arm; while genes were basically equally distributed on chromosome 7 (Figure 4).



**Figure 4.** Chromosomal locations of maize *PHD* genes. The left scale represents the megabases (Mb). Chromosome numbers are shown at the top of each vertical black bar. The rough location of each maize *PHD* genes is marked with the grey line that link the gene names with the chromosome bars. The segmental duplication gene pairs are marked with color boxes and joined by corresponding color lines, and the grey box indicates the tandem duplication gene pairs.

Gene duplication events, including tandem and segmental duplications, are of great importance to expand the number of gene family members [29]. Potential duplication events were investigated with the purpose of elucidating the expanded mechanism of the maize *PHD* gene family, which had occurred during the process of evolution. Based on the phylogenetic and comparative analysis of the *ZmPHD* genes, 12 gene pairs (*ZmPHD13/62*, *ZmPHD20/33*, *ZmPHD51/39*, *ZmPHD48/39*,

*ZmPHD48/51*, *ZmPHD4/7*, *ZmPHD22/30*, *ZmPHD40/50*, *ZmPHD15/55*, *ZmPHD63/14*, *ZmPHD37/56*, *ZmPHD40/55*) were identified to be involved in the segmental duplication events. Among the 12 segmental duplication pairs, the high frequency of segmental duplication appeared between chromosomes 6 and 8, which contained four segmental duplications. It is noteworthy that only one gene pair *ZmPHD65/66* was involved in tandem duplication. These results suggest that segmental duplication events are the main gene duplication events in the maize PHD-finger family, and might play major roles in the amplification of the maize PHD family, which is consistent with previous research [24].

To explore the selection in duplication and divergence of PHDs in maize, the non-synonymous (*Ka*), synonymous (*Ks*) and *Ka/Ks* were calculated for each pairs of duplicated *ZmPHD* genes. The *Ka* and *Ks* were used to examine the course of divergence after duplication, and the *Ka/Ks* ratio was applied to measure selective pressure of duplicated gene pairs. Generally, the value of *Ka/Ks* > 1 indicated positive selection that accelerated the evolution; the *Ka/Ks* ratio = 1 signified neutral selection; while *Ka/Ks* ratio < 1 stood for negative selection or purifying selection. The value of *Ka/Ks* of 13 duplicated gene pairs in maize varied from 0.084 to 0.764 with an average of 0.286.; in addition, the majority of them were less than 0.3, which indicated that the duplicated *ZmPHD* genes were under strong negative selection during evolution (Table 2). Based on the estimations for *Ks*, the time of 13 duplicated evens were calculated with a substitution rate of  $6.5 \times 10^{-9}$  substitutions per site per year [30], which were ranged from 7.38 to 74.53 million years (Mya) (Table 2).

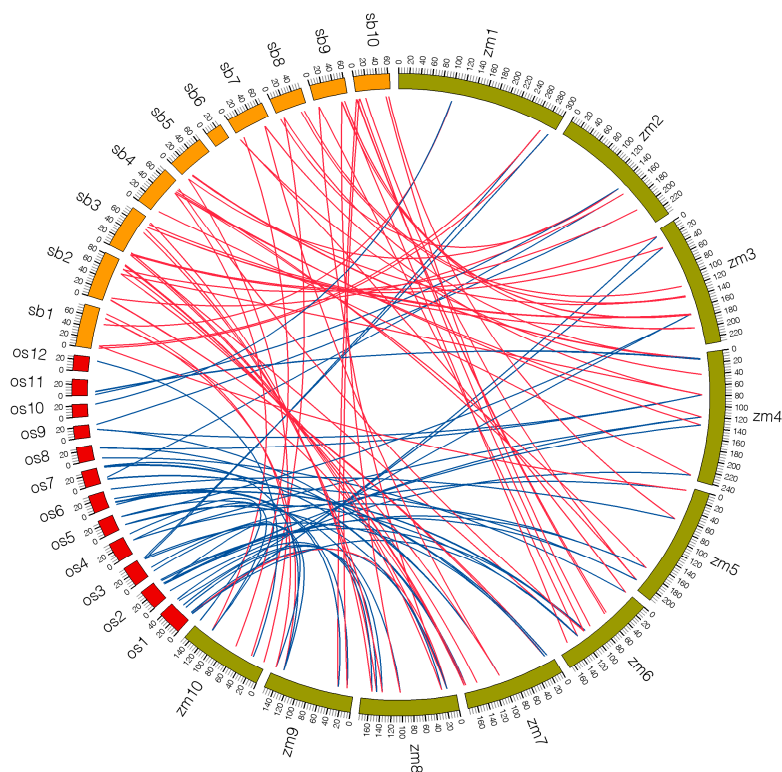
**Table 2.** *Ka/Ks* analysis and divergence time estimated for maize duplicated PHD paralogs.

Paralogous Pairs	<i>Ks</i>	<i>Ka</i>	<i>Ka/Ks</i>	Duplication Date (MY)	Duplicate Type
<i>ZmPHD4/7</i>	0.969	0.175	0.181	74.53	segmental
<i>ZmPHD13/62</i>	0.148	0.031	0.209	16.07	segmental
<i>ZmPHD14/63</i>	0.176	0.026	0.147	13.53	segmental
<i>ZmPHD15/55</i>	0.686	0.436	0.635	52.76	segmental
<i>ZmPHD20/33</i>	0.23	0.078	0.339	17.69	segmental
<i>ZmPHD22/30</i>	0.151	0.019	0.125	11.61	segmental
<i>ZmPHD37/56</i>	0.188	0.025	0.132	14.46	segmental
<i>ZmPHD40/50</i>	0.131	0.02	0.152	10.07	segmental
<i>ZmPHD40/55</i>	0.7	0.535	0.764	53.84	segmental
<i>ZmPHD48/39</i>	0.657	0.112	0.17	50.53	segmental
<i>ZmPHD48/51</i>	0.627	0.109	0.173	48.23	segmental
<i>ZmPHD51/39</i>	0.236	0.02	0.084	18.15	segmental
<i>ZmPHD65/66</i>	0.096	0.059	0.614	7.38	tandem

### 2.5. Microsynteny Analysis among Maize, Sorghum and Rice

To elucidate the evolutionary relationship of PHD families among maize, sorghum and rice, comparative analysis was performed to identify orthologous *PHD* genes. Seventy-one orthologous gene pairs were found between maize and sorghum. Conversely, only 54 orthologous gene pairs were detected between maize and rice (Figure 5, Tables S2 and S3). The number of orthologous genes between maize and sorghum was far greater than that between maize and rice, which might be due to the closer relationship between maize and sorghum [31]. In addition, a portion of collinear

gene pairs between maize and sorghum were not identified between maize and rice, such as *ZmPHD35/Sb10g001106.1*, *ZmPHD63/Sb08g006530.1* and *ZmPHD40/Sb09g020610.1* and so on, which implied these orthologous gene pairs appear behind the divergence of the progenitors of maize and rice (Figure 5, Tables S2 and S3). However, two orthologous gene pairs (*ZmPHD57/LOC\_Os06g01170.1*, *ZmPHD62/LOC\_Os12g34330.1*) were not found between maize and sorghum that might due to the genes loss in the evolution of sorghum. Furthermore, we also found that two or more PHD genes from sorghum and rice were orthologous of the same maize *PHD* gene; these genes are probably paralogous gene pairs and play important roles in the amplification of *PHD* gene family in the process of evolution. For example, *Sb04g023220.1* and *Sb06g017810.1* are orthologous genes to *ZmPHD30*, as well as *LOC\_Os02g48800.1* and *LOC\_Os06g20410.1* to *ZmPHD31* (Figure 5, Tables S2 and S3).

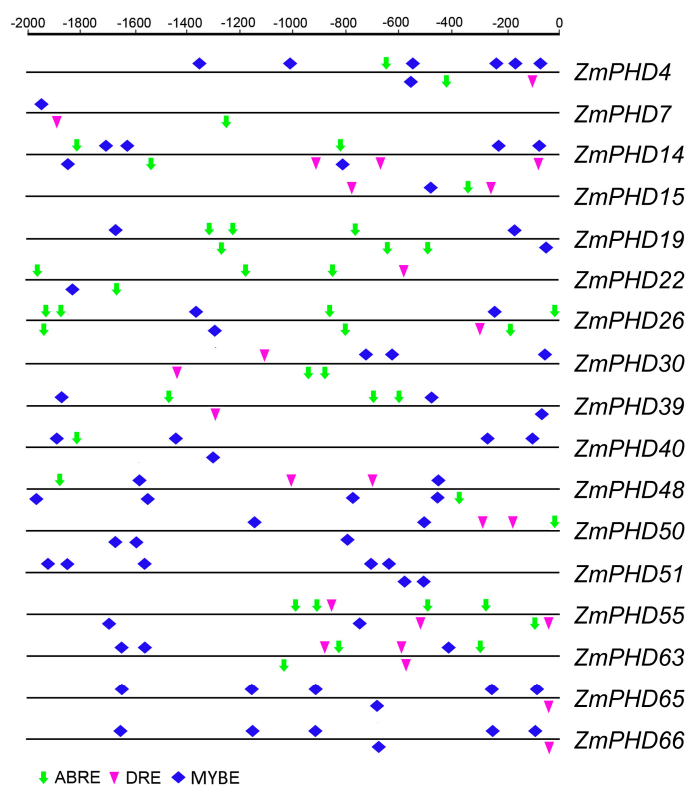


**Figure 5.** Microsynteny of PHD regions across maize, sorghum and rice. The maize, sorghum and rice chromosomes are shown in different color boxes and labeled zm, sb and os, respectively. Numbers along each chromosome box indicate sequence lengths in megabases. Red lines represent the syntenic relationships between maize and sorghum PHD regions, and blue lines represent the syntenic relationships between maize and rice regions.

## 2.6. Analysis of the Promoters of Potential Abiotic Stress-Responsive PHD Genes

It was reported that six soybean *GmPHD*-type transcription regulators enhanced stress tolerance in transgenic *Arabidopsis* [21]. We analyzed the conserved motifs for these 6 *GmPHD* genes by MEME, which showed highly similarity with group IX PHD genes (Figure S1). The phylogenetic analysis also showed that they shared closely evolutionary relationship (Figure S1). As proteins functions were closely related to their structures, we conjectured the 17 *ZmPHD* genes might also play a role under adversity stress in maize. This observation prompted us to investigate possible stress-responsive

*cis*-elements in the promoter regions of the 17 maize *PHD* genes by searching against the PLACE database. Three types of *cis*-elements, including the ABRE (ABA responsive element), DRE (dehydration-responsive element) and MYB-binding site *cis*-element (MYBE) were detected in current study [32]. The results showed that each of the 17 *PHD* genes contained one or more types of *cis*-elements of ABRE, DRE or in their 2000 bp promoter sequences (Figure 6), which were key elements for stress responsiveness. Although the phylogenetic analysis showed close relationships among these 17 *ZmPHD* genes (Figure S1), we found surprising differences in the numbers of the three *cis*-elements in their promoter regions (Table S4). For example, the promoter regions of *ZmPHD14* and *ZmPHD63* contained multiple putative ABREs, DREs and MYBEs. In contrast, only MYBEs were detected in the promoter of *ZmPHD51*. Moreover, we also found that the *cis*-elements were not conserved in the promoter regions of the five segmental duplicated gene pairs (*ZmPHD22/30*, *ZmPHD40/50*, *ZmPHD14/63*, *ZmPHD40/55*, *ZmPHD48/51*) when compared to the tandem duplicated one (*ZmPHD65/66*). This observation indicated that the five segmental duplication pairs might possess different regulatory feature. In addition, each of these six paralogs contained at least one ABRE or MYBE element, which was the central *cis*-elements signal transduction in response to stress conditions (Figure 6, Table S4). Thus, we concluded that these duplicated genes might share similar regulatory pathway in some respects.



**Figure 6.** *Cis*-elements in the promoter regions of maize putative stress-responsive *PHD* genes. The stress-responsive *cis*-elements distributed on the sense strand and reverse strand are shown above and below the grey lines, respectively. ABRE, MYB and DRE core sequences are indicated by green thick drop-down arrows, pink triangles and blue squares, respectively. ABRE: ABA responsive element; MYB: MYB-binding site *cis*-element; DRE: dehydration-responsive element.





employed to investigate the expression patterns of the group IX genes using three-week-old maize leaves under ABA, NaCl, and PEG treatment, respectively. The expression levels of 17 *PHD* genes are shown in Figure 8 except for *ZmPHD65*, -66, which were not expressed during the whole treatment stage under three stress treatments (Figure 8). Of the 15 *PHD*-finger genes, 14 genes were obviously up-regulated in response to ABA stress, besides, *ZmPHD15*, which was up-regulated at 3 h after ABA treatment and down-regulated thereafter. Notably, five out of the 14 up-regulated genes showed great amplitude of variation of expression level (>3-fold), including *ZmPHD7*, -14, -19, -30 and -39, while the remaining nine genes *ZmPHD4*, -22, -26, 40, -48, -50, -51, -55 and -63 showed slight expression changes (<2-fold) (Figure 8A). Under NaCl treatment, the expression levels of *ZmPHD15*, -30, -39, -48, -50, -51 and -63 were mildly regulated (<2-fold), but the other seven genes were highly induced (Figure 8B). Among the 15 *PHD*-finger genes *ZmPHD14* and -22 were dramatically down-regulated across all time points. The *ZmPHD4*, -7, -19, -26, and -55 were strongly up-regulated at the early stage after NaCl treatment, whereas, their expression revealed relatively decreases thereafter, but still higher than the control. All the group IX genes under PEG treatment, especially *ZmPHD22*, -30, and -39 were strongly up-regulated by more than 8- to 30-fold compared to the control (Figure 8C). It is interesting that a large proportion of the up-regulated genes were down-regulated at 1h after PEG treatment, and then reached a peak expression level at 6 h, while the *ZmPHD4*, -15, -19 and -26 reached the peak expression level at 3 h.

By comparing with the relative expression of *PHD*-finger genes in group IX under ABA, NaCl and PEG treatment, we found *ZmPHD14*, -19 showed strong induction or repression under all three stress treatments, while the rest genes exhibited highly regulated just in one or two treatment, for *ZmPHD15*, -48, -51, and -63 showed highly regulated by PEG stress, but exhibited slightly expression change under other two treatments, *ZmPHD22*, and -55 were strongly regulated just in PEG and NaCl treatment (Figure 8). Of nine pairs of segmental duplicated genes in the group IX, most exhibited similar expression profiles under three stresses treatment, such as *ZmPHD14/63*, *ZmPHD40/50* following PEG treatment, *ZmPHD39/51* following NaCl treatment and *ZmPHD22/30*, *ZmPHD14/63* following ABA treatment. However, the expression profiling of some segmental duplicated genes was different, for *ZmPHD15/55* under PEG treatment and *ZmPHD4/7* under ABA treatment (Figure 8).

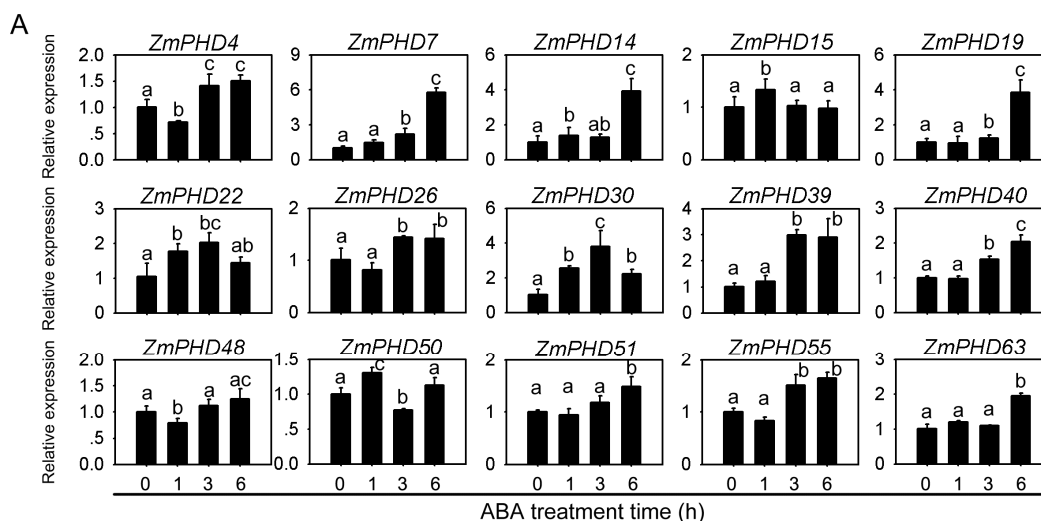
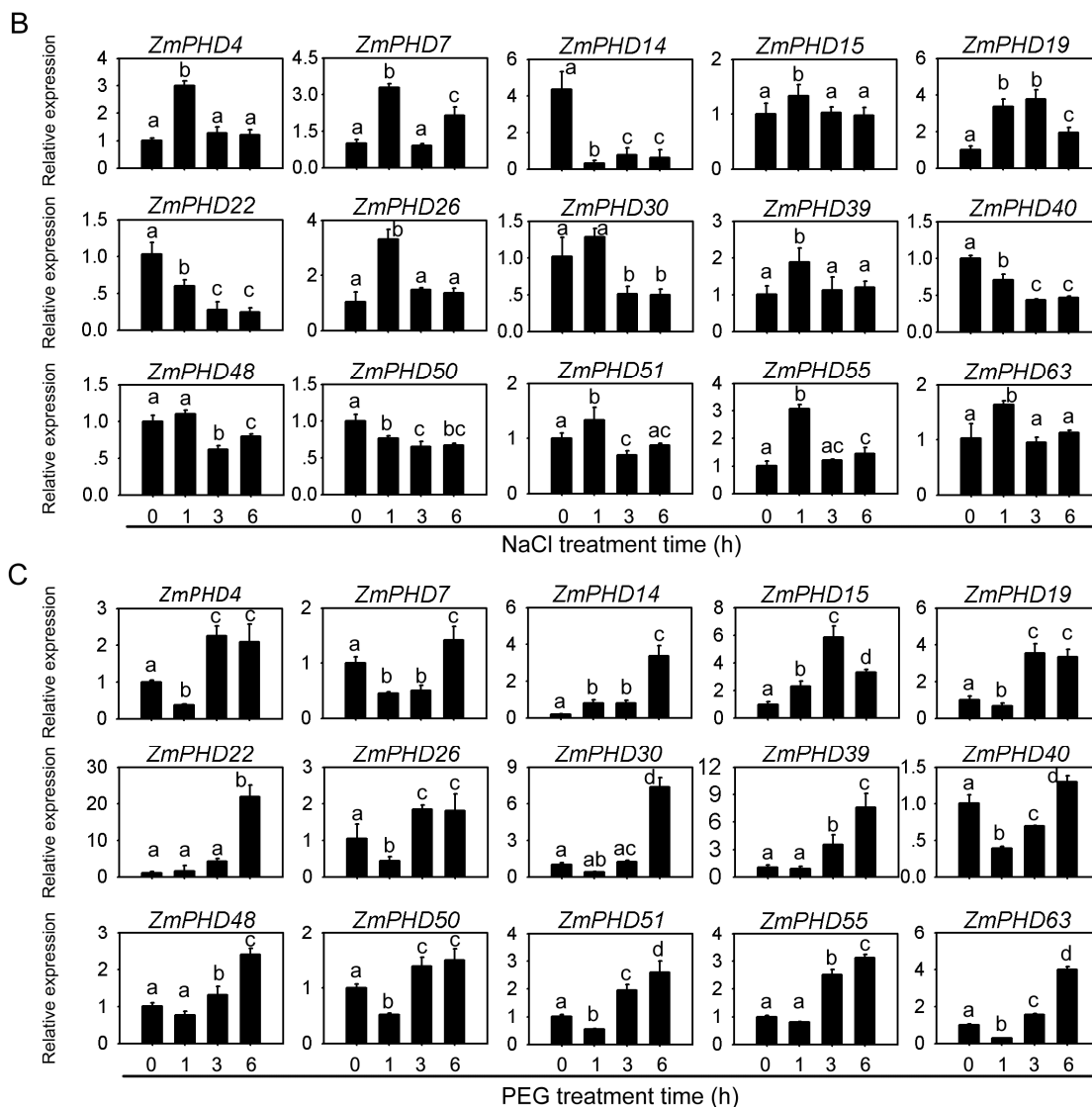


Figure 8. Cont.

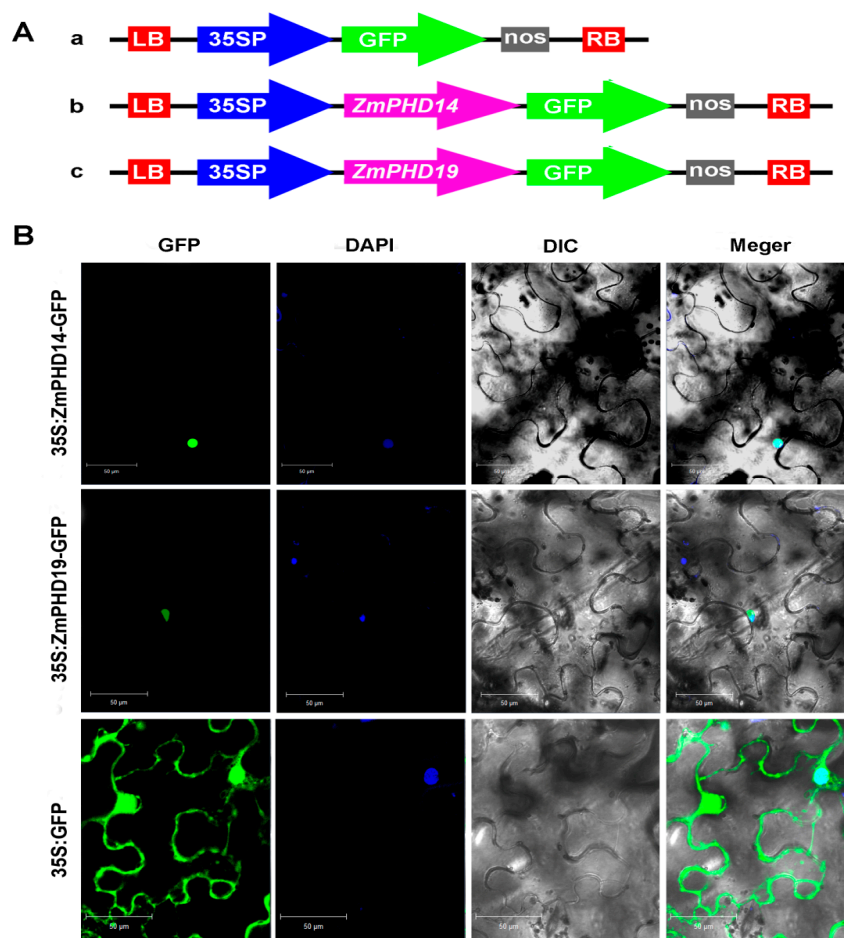


**Figure 8.** Expression patterns of *PHD* group IX genes in response to drought, NaCl and ABA treatments. The relative expression level of 15 *PHD* group IX genes was examined by the qRT-PCR and normalized with the reference gene *ZmGAPDH*. (A) Relative expression of 15 *PHD* genes under ABA treatment at 0, 1, 3, 6 h; (B) Relative expression of 15 *PHD* genes under NaCl treatment at 0, 1, 3, 6 h; and (C) Relative expression of 15 *PHD* genes under PEG6000 treatment at 0, 1, 3, 6 h. The error bars represent standard deviations (SD), y-axes are scales of relative expression level and x-axes are the time course of treatments for each condition. Different lowercase letters indicate significant differences at  $p < 0.05$  (Duncan’s test).

### 2.9. Subcellular Localization of *ZmPHD14* and *ZmPHD19*

It is reported that a large proportion of PHD-finger proteins are located in nucleus, while some of them are in membrane [33,34]. Six GmPHD proteins (GmPHD1–GmPHD6) were targeted to the nucleus and the PHD domains were required for their nuclear location [21]. Group IX members had a close relationship and similar structures with these GmPHD proteins, hence they might also locate in the nucleus. *ZmPHD14* and *ZmPHD19* showed great expression level changes under NaCl,

ABA and PEG treatments, and implied they probably participated in the ABA-mediated way of adversity response. To further study these two proteins characteristics, subcellular localization analysis was performed. The *ZmPHD14-GFP* and *ZmPHD19-GFP* fusion constructs and the GFP control driven by CaMV 35S promoter were introduced into tobacco epidermal cells. As shown in Figure 9B, the GFP signal was consistently observed throughout the whole cell, whereas *ZmPHD14-GFP* and *ZmPHD19-GFP* fusion proteins were restricted to the nucleus as confirmed by DAPI staining. In addition, in the onion epidermal cells, the GFP signals of *ZmPHD14-GFP* and *ZmPHD19-GFP* were only found in the nucleus, while the control GFP signals were widely distributed (Figure S2). These results both indicated that *ZmPHD14* and *ZmPHD19* were nuclear proteins, which was similar to *GmPHDs* [21].



**Figure 9.** Subcellular localization of *ZmPHD14-GFP* and *ZmPHD19-GFP* fusion protein. (A) Schematic representation of the 35S:GFP, 35S:*ZmPHD14-GFP* and 35S:*ZmPHD19-GFP* fusion constructs used for transient expression; and (B) Fusion proteins were transiently expressed under control of the CaMV35S promoter in tobacco leaves and observed under a laser scanning confocal microscope. Green color is GFP protein signal, and blue color represents DAPI stained for nucleus. Bars = 50 µm.

### 3. Discussion

PHD-finger proteins have been identified in different plant species as involved in various biological processes, including the regulation of chromatin structure and transcription, response to adversity

conditions, phosphate deficiency, DNA replication and so on. Although the number of PHD-finger genes that have been identified is increasing in different species, most of putative PHD family members remain unrecognized and uncharacterized and few comprehensive analyses of the PHD family in evolution had been conducted. In the present study, we performed a comprehensive genome-wide analysis of PHD family in maize. A total of 67 maize non-redundant *PHD* genes were identified. Previous study demonstrated that the maize genome size is approximately 18 times larger than *Arabidopsis*, while the maize gene number is about 1.3 times higher than *Arabidopsis* [23]. In *Arabidopsis*, a total of 45 *PHD* genes were identified from its genome [35]. However, the number of maize *PHD* genes was about 1.39 times higher than that in *Arabidopsis*, which was agreement with previous research (1.3) [23] and partially accounted for the support of PHD conservation in these two species during the evolutionary process.

Gene structure analysis indicated that most maize *PHD* genes within the same group shared similar organization, including exon/intron distribution and motif components, while divergence also existed among different groups, which might relate to functional diversity of maize PHD members (Figures 1 and 3). In addition, of the 13 *ZmPHD* paralogous pairs, most shared conserved gene structure as well as the motif composition, while several pairs exhibited certain degrees of divergence. For example, the *ZmPHD65* contained five introns, whereas its counterpart *ZmPHD66* only possessed four introns. This divergence was also observed between *ZmPHD14* and *ZmPHD63* (Figure 1), which might be attributed to the single intron loss or gain during the evolution process in gene structure. In summary, the same group maize *PHD* genes were conserved in evolution, and also accompanying the variance of gene organization in some degree, suggesting that some maize PHD members were functionally diversified through differential expansion. In general, the analysis of maize *PHD* genes was kept in line with the phylogenetic analysis, and the consistency of gene structure and phylogenetic analysis also demonstrated the reliability of the phylogenetic analysis as well as the conserved evolutionary relationships among *ZmPHD* proteins.

Gene duplications are one of the primary driving forces in the evolution of genomes and genetic systems, including segmental duplication, tandem duplication, transposition events and whole-genome duplication [36], which had been demonstrated to play crucial roles in gene family expansion in many species, such as the HD-Zip and HSF family in maize, the Cyclophilin family in soybean and the WRKY family in cotton [22,37–39]. It was observed that many *PHD* genes in maize had one or more paralog genes, which implied that the expansion of the maize PHD family might be due to the gene duplications. It was believed that segmental duplication often occurred in a gene family that evolved slowly, like the *MYB* gene family [29], whereas, tandem duplication in local genomic clusters with low retention is common in the large and rapidly evolving gene family, like the Nucleotide binding site-leucine-rich repeats (NBS-LRR) disease resistance family [29]. Our analysis revealed that the number of maize *PHD* genes involved in segments duplication (12 pairs) was much larger than those arranged in chromosomal tandem duplication events (1 pair) (Table 2). Therefore, we hypothesize that the maize *PHD* gene family is a slowly evolving gene family, and that the segment duplications have played a key role in the expansion of the maize *PHD* gene family, which was also observed in some other maize gene families, such as the mitogen-activated protein kinase (MAPK) and CCCH-type zinc fingers [24,40]. It is hard to achieve the amplification of transcriptional regulating gene family through single-gene duplication events, which indicate the significance of

genome duplications in the process of expanding the regulatory gene repertoire [41]. During the last 150 million years, more than 90% increases in regulatory genes were attributed to genome duplication in the *Arabidopsis* lineage [42]. Studies for the maize genome indicated that its genome had undergone at least two rounds of genome duplication, one ancient event was before rice diverged from the common ancestor of maize and sorghum at about 50 Mya, the other one was a recent one at approximately 11.9 Mya [43]. The duplication dates of maize *PHD* duplicated gene pairs were ranged from 7.38 to 74.53 million years (Mya) (Table 2), and the dates of four gene pairs were close to the first whole genome duplication Event (WGD) at 50 Mya, including *ZmPHD15/55*, *ZmPHD40/55*, *ZmPHD48/39* and *ZmPHD48/51*. Similarly, the dates of *ZmPHD14/63*, *ZmPHD22/30*, *ZmPHD37/56*, and *ZmPHD40/50* were close to the recent WGD date, which indicated that the large-scale genome duplication events also had played crucial roles in the expansion of maize *PHD* family. This phenomenon was also reported in *Gossypium raimondii*, where the average duplication date of *GrTCP* genes was very close to the recent whole-genome duplication date of *G. raimondii*, suggesting that large-scale genome duplication events might also contribute to the expansion of the *GrTCP* family [44]. Generally, the value of  $Ka/Ks > 1$  indicated positive selection that accelerated the evolution; the  $Ka/Ks$  ratio = 1 signified neutral selection; while  $Ka/Ks$  ratio  $< 1$  stood for negative selection or purifying selection [45]. By estimating the value of  $Ka/Ks$  for 13 duplicated gene pairs in maize, we found that the duplicated *ZmPHD* genes were under strong purifying selection during the history of evolution with an average  $Ka/Ks$  ratio of 0.286 (Table 2).

Previous study indicated that rice diverged from the common ancestor of maize, sorghum and rice at approximately 50 million years ago, which was followed by the divergence of maize and sorghum at 11.9 million years ago [31]. Our studies showed that 71 orthologous gene pairs were identified between maize and sorghum, whereas, only 54 orthologous gene pairs were found between maize and rice (Figure 5, Tables S2 and S3), which indicated that *ZmPHDs* were more closely allied with *SbPHDs* than *OsPHDs*; this phenomenon was also observed in maize peroxidase family [46]. Some *PHD* genes in maize had more than one counterpart in sorghum or rice, suggesting that *PHD* genes may undergo differential amplification in these three species. Similarly, 18 sorghum and nine rice *PHD* members shared one-more orthologous relationships with maize *PHD* genes, such as *Sb10g005910.1* to *ZmPHD56*, -67 and *LOC\_Os01g06540.1* to *ZmPHD39*, -48 (Figure 5, Tables S2 and S3). These results indicated the conserved evolutionary relationship among maize, sorghum and rice, as well as the expansion of *PHD* family resulted from the whole-genome duplicated events.

According to the microarray expression profile analysis, these newly identified *PHD* genes exhibited distinct temporal and spatial expression patterns in different tissues or organs. This result indicated that these genes might be expressed under specific environments or at specific developmental stages. In addition, some duplicated *ZmPHD* gene pairs also showed diverse expression profiles between each other, which revealed that functional diversification of the surviving duplicated genes was a main feature of the long-term evolution [47]. Many *PHD* proteins had been detected in plenty of species to be involved in numerous physiological and biochemical processes, and their main functions was the focus on the regulation of chromatin structure and transcription, whereas, *PHD* members in response to abiotic stress in plant was rarely reported. Up to now, we found six soybean *PHD* proteins that conferred *Arabidopsis* salt tolerance through ABA-mediated signal transduction pathway [21]. Protein structure comparison and phylogenetic analysis revealed that these *GmPHDs* share close

relationships with maize group IX PHD members (Figure S1), indicating the potential function of *ZmPHDs* in response to abiotic stress in maize, which was supported by qPCR-PCR results that 15 maize *PHD* genes exhibited different degree response to ABA, NaCl and PEG treatments. Previous study indicated three outcomes of duplicated genes: neofunctionalization, subfunctionalization and pseudogenization [45]. Studies for 10 pairs of paralogous genes showed that tandem duplicated genes, *ZmPHD65* and *-66*, were not expressed during the whole treatment stage under three stress treatments. In addition, microarray expression profile analysis also revealed that both of them had no expressing among all tissues, which indicated that *ZmPHD65*, *-66* might be expressed under special conditions or at specific developmental stages or had undergone pseudogenization; this phenomenon was also detected in the study of soybean cyclophilin family. Of segmental duplicated genes, most exhibited similar expression profiles under certain stress treatment, such as *ZmPHD14/63*, *ZmPHD40/50* following PEG treatment, *ZmPHD39/51* following NaCl treatment (Figure 8), which indicated that the duplicated genes might have redundant functions in response to specific stress. However, the divergence was also existent between duplicated genes, suggesting neofunctionalization or subfunctionalization of duplicated genes, which was the major feature of most duplicated genes [47].

ABA plays a key role in plant adaptation to adverse environmental conditions. Lots of functional genes have been reported to be involved in plant abiotic stress response through ABA-mediated signal pathways, such as *OsMYB2*, *ZmMPK7* [48,49]. In the present study, six genes (*ZmPHD7*, *-14*, *-15*, *-19*, *-30*, *-39*) were strongly induced under ABA treatment (>3-fold) (Figure 8), indicating they might participate in the ABA signal pathway, which was supported by the promoter *cis*-element analysis (Figure 6). In addition, *ZmPHD14* and *ZmPHD19* also were highly induced or repressed under NaCl and PEG stress treatments. Previous studies showed that most PHD proteins were located in the nucleus. Subcellular localization analysis demonstrated that *ZmPHD14* and *ZmPHD19* were also nuclear proteins (Figure 9 and Figure S2), which was in accord with *GmPHDs* [21]. According to these results, we speculate that *ZmPHD14* and *ZmPHD19* and *GmPHDs* may share similar functions in response to abiotic stress conditions, and these remain to be further confirmed experimentally.

## 4. Experimental Section

### 4.1. Collection and Classification of PHD-Finger Transcription Factors

Identification and annotation of *PHD*-finger genes were performed in maize. First of all, we downloaded the Hidden Markov Model (HMM) profile of *PHD*-finger proteins from Pfam (Available online: <http://pfam.sanger.ac.uk/search>). The consensus protein sequences of the *PHD*-finger Hidden Markov Model profile were employed as a query to search against maize genome database with BLASTP program (Available online: <ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST>) [50]. Based on the results of further analysis for sequence alignment, chromosomal location and sequence identification numbers, all redundant protein sequences were discarded. To confirm the putative *PHD* genes in maize genome, the amino acid sequences of all the proteins were searched for the presence of *PHD* domains by Pfam and SMART (Available online: <http://smart.embl-heidelberg.de/>).



#### 4.2. Phylogenetic and Gene Structure Analysis

Multiple sequence alignments for the 67 maize PHD-finger protein sequences were conducted using ClustalX version 1.83 with default settings [51]. Based on alignments, then an unrooted phylogenetic tree was generated by MEGA 4.0 (Available online: <http://www.megasoftware.net/>) using neighbor-joining (NJ) method with bootstrap analysis (1000 replicates) [52,53]. Another unrooted tree was constructed using the same method with the alignment of 17 maize PHDs and six soybean PHD protein sequences. Multiple sequence alignments were performed with BOXSHADE (Available online: [http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)) to highlight the conserved or similar amino acids residue location. The exon-intron organization analysis of the *PHD* genes was performed with Gene Structure Display Server (Available online: <http://gsds.cbi.pku.edu.cn/>) by comparing the CDS sequences with their corresponding DNA sequences. To identify the conserved motifs, the maize PHD proteins sequences were submitted to online Multiple Expectation maximization for Motif Elicitation (MEME) program (Available online: <http://meme-suite.org/tools/meme>) with the following parameters: any number of repetitions, the optimum width from 15 to 200 residues and maximum number of motifs 20. In addition, the indentified motifs were annotated with the Pfam and SMART programs [44].

#### 4.3. The Analysis of Chromosome Location and Gene Duplication

Maize *PHD* genes were placed on 10 maize chromosomes according to their positions given in the TIGR maize database (Available online: [http://maize.jcvi.org/repeat\\_db.shtml](http://maize.jcvi.org/repeat_db.shtml)). The distribution of *ZmPHD* genes on the maize chromosomes was drawn by MapInspect (Available online: <http://mapinspect.software.informer.com/>) and modified manually with annotation. The tandem duplicated gene pairs and segmental duplicated gene pairs were identified according to previous studies [44,54]. The software MEGA 4.0 and DNAMAN 5.22 (Available online: <http://dnaman.software.informer.com/>) were used to analyses the homology of duplicated *ZmPHD* genes [55]. To further analyze gene duplication events, the nonsynonymous substitution rate ( $K_a$ ) and synonymous substitution rate ( $K_s$ ) were calculated by DnaSP 5.0 (Available online: <http://www.ub.edu/dnasp>) [56]. The ratios of  $K_a/K_s$  between each paralogs were estimated to detect the selection pressure in evolution. The  $K_s$  value was used to count the dates of every duplicated events occurred in maize, based on a rate of  $\lambda$  substitutions per synonymous site per year. The duplicated time ( $T$ ) =  $K_s/2\lambda \times 10^{-6}$  Mya ( $\lambda = 6.5 \times 10^{-9}$  for grasses) [30,57].

#### 4.4. Interspecies Microsynteny Analysis

To detect the syntenic regions among maize, sorghum and rice, the MCScanX (Available online: <http://chibba.pgml.uga.edu/mcscan2/>) was used [58]. First of all, the whole-genome protein sequences of maize, sorghum and rice were searched against themselves using BLASTP with an  $E$ -value cutoff  $1 \times 10^{-10}$  and identity  $>75\%$ . Subsequently, the MCScanX and associated downstream tools using default parameters were used for detecting the collinear blocks according to the previous study [58]. Finally, we identified the duplicated *PHDs* in these syntenic regions using in-house perl script,

and the relationships of PHD orthologous genes among the three species were plotted using Circos software (Available online: <http://circos.ca/>) [59].

#### 4.5. *Cis-Elements in the Promoter Regions of Stress-Responsive PHD Class IX Genes*

To predict *cis*-acting regulatory DNA elements (*cis*-elements) in promoter regions of maize *PHD* group IX genes, the PLACE website (Available online: <http://www.dna.affrc.go.jp/PLACE/signalscan.html>) was adopted to identify putative *cis*-elements in the 2000 bp genomic DNA sequences upstream of the initiation codon (ATG) [60].

#### 4.6. *Microarray-Based Expression Analysis of Maize PHD Genes*

To analyze the spatial and temporal expression patterns of *PHD* genes during maize development, transcriptome data of the genome-wide genes expression atlas of maize inbred line B73 made by the NimbleGen microarray technology was downloaded from Plexdb (ZM37) (Available online: [http://www.plexdb.org/modules/PD\\_browse/experiment\\_browser.php?experiment=ZM37](http://www.plexdb.org/modules/PD_browse/experiment_browser.php?experiment=ZM37)). We got the transcriptome data from the article of Maize Gene Atlas Developed by RNA Sequencing and Comparative Evaluation of Transcriptomes Based on RNA Sequencing and Microarrays [61], the average expression value of a gene had to be greater than 0 FPKM (Fragments Per Kilobase Exon model per Million mapped fragments) in at least one of the 18 tissues, the RNA-Seq data were processed by adding a number 1. Then the microarray data of *PHD* genes were imported into R and Bioconductor (Available online: <http://www.bioconductor.org/>) for expression analysis.

#### 4.7. *Plant Materials and Stress Treatment*

The maize inbred line B73 was used to examine the expression patterns of *PHD*-finger genes in all experiments. Plants grew in a greenhouse at 28 °C with a photoperiod of 14-h light/10-h dark and relative humidity 60% ± 5%. The soil was mixed containing soil/vermiculite/perlite as the volume proportion of 4:1:1. Distilled water was used to water the plant before treatment. Three-weeks-old B73 seedlings were treated under salt, PEG6000 and ABA stress. For ABA treatment, three-week-old seedling leaves were sprayed with 100 µM ABA solution, followed by sampling at 0, 1, 3, and 6 h [62,63]. Seedlings at the same stage were watering with 20% PEG6000 and 150 mM sodium chloride (NaCl) solution, followed by sampling at 0, 1, 3 and 6 h, which simulated the drought and salt stress, respectively. Seedlings without treatment were used as the control. For all the samples, three biological replicates were conducted.

#### 4.8. *RNA Isolation and qRT-PCR Analysis*

Total RNA of all the samples was isolated using the Trizol reagent (Invitrogen, Foster city, CA, USA) according to the manufacturer's instructions and the RNAs were treated with the DNase (Invitrogen). The first-strand cDNA was then synthesized using M-MLV reverse transcriptase (Invitrogen). The gene-specific primers were designed to amplify 80–150 bp PCR products by Primer Express 3.0 software (Applied Biosystems, Foster City, CA, USA), and their specificity were examined through Primer Blast on the NCBI (Table S1). Real-time PCR was performed on an ABI 7300 Real-Time

system (Applied Biosystems). The total reaction volume was 20  $\mu\text{L}$  contained 10  $\mu\text{L}$  of SYBR Green Master Mix reagent (Applied Biosystems), 1.5  $\mu\text{L}$  of cDNA sample, 1  $\mu\text{L}$  of each gene-specific primer, and 6.5  $\mu\text{L}$ . The PCR condition was set as follows: 95  $^{\circ}\text{C}$  for 10 min, 40 cycles of 95  $^{\circ}\text{C}$  for 15 s and 60  $^{\circ}\text{C}$  for 1 min. Each sample was performed four technical replicates. The expression level of maize *GAPDH* gene was served as an internal control. The relative expression level was calculated as  $2^{-\Delta\Delta\text{Ct}}$  ( $\Delta\text{Ct} = \text{Ct}_{\text{Target}} - \text{Ct}_{\text{GAPDH}}$ ,  $\Delta\Delta\text{Ct} = \Delta\text{Ct}_{\text{treatment}} - \Delta\text{Ct}_{\text{0 h}}$ ). The relative expression level ( $2^{-\Delta\Delta\text{Ct}^0 \text{ h}}$ ) in the control plants without treatment was normalized to 1 as described previously [64]. Statistical analyses were performed using SDS software 1.3.1 (Applied Biosystems), and SPSS 19.0 (Available online: <http://www.spss.com.cn/>).

#### 4.9. Determination of Subcellular Localization of *ZmPHD14* and *ZmPHD19*

Full-length cDNA of *ZmPHD14* and *ZmPHD19* were PCR-amplified and sequenced (Figure 9A), the primers were shown in Table S5. These two cDNA sequences were cloned between *Xba*I and *Bam*HI sites of the pCAMBIA1305-GFP vector (pCAMBIA1305-GFP is modified from pCAMBIA1305 vector). The resulting 35S:*ZmPHD14*-GFP, 35S:*ZmPHD19*-GFP and GFP control vector were transiently expressed in *Nicotiana benthamiana* leaves via *Agrobacterium*-mediated infiltration as described previously [65]. Two days later, the fluorescence of the infected leaf tissue was observed under a Zeiss LSM700 (Zeiss, Jena, Germany) confocal microscope, the DNA dye 4,6-diamidino-2-phenylindole (DAPI) was used to visualize the nucleus. For the subcellular localization in onion, the *Agrobacterium*-mediated method was used. The onion cells were soaked in prepared *Agrobacterium* solution ( $\text{OD}_{600} = 1$ ) for about 35 min and then placed on Murashige and Skoog (MS) medium. After 24 h of incubation at 25  $^{\circ}\text{C}$  in the dark, GFP fluorescence was observed with a fluorescence microscope (LEICA, Wetzlar, Germany).

## 5. Conclusions

We identified 67 non-redundant PHD members in the maize genome. Further analyses for phylogenetic relationships, gene structure and duplication, indicates that the conservation of the maize PHD family in evolution was accompanied with divergence to a certain degree. In addition, the expansion of the PHD family occurred, resulting from both large-scale genome duplication and small-scale duplicated events, such as segmental duplication and tandem duplication. Expression analysis demonstrated that group IX genes, especially, *ZmPHD14* and *ZmPHD19* are likely involved in the abiotic stress response. From an applied perspective, the identification of *ZmPHDs* will serve a very useful reference for more detailed functional analysis, and be useful in selecting appropriate candidate genes for further study of *PHD* genes in maize.

## Supplementary Materials

Supplementary materials can be found at <http://www.mdpi.com/1422-0067/16/10/23517/s1>.

## Acknowledgments

We thank the members of the Key Laboratory of Crop Biology of Anhui Province for their assistance in this study. This work was supported by grants from National Basic Research Program of China (2014CB138200) and the National Natural Science Foundation of China (91435110, 31301324).

## Author Contributions

Qianqian Wang, Jinyang Liu, Haiyang Jiang and Beijiu Cheng conceived and designed this research; Qianqian Wang and Jinyang Liu performed the experiment; Qianqian Wang, Jinyang Liu, Yu Wang and Yang Zhao analyzed the data; Qianqian Wang, Jinyang Liu, Haiyang Jiang and Beijiu Cheng contributed reagents/materials/analysis tools; Qianqian Wang, Jinyang Liu, Haiyang Jiang and Beijiu Cheng wrote the manuscript. All authors read and approved the manuscript.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Takatsuji, H. Zinc-finger transcription factors in plants. *Cell. Mol. Life Sci.* **1998**, *54*, 582–596.
2. Aasland, R.; Gibson, T.J.; Stewart, A.F. The PHD finger: Implications for chromatin-mediated transcriptional regulation. *Trends Biochem. Sci.* **1995**, *20*, 56–59.
3. Borden, K.L.; Freemont, P.S. The ring finger domain: A recent example of a sequence-structure family. *Curr. Opin. Struct. Biol.* **1996**, *6*, 395–401.
4. Townsley, F.M.; Thompson, B.; Bienz, M. Pygopus residues required for its binding to legless are critical for transcription and development. *J. Biol. Chem.* **2004**, *279*, 5177–5183.
5. Schindler, U.; Beckmann, H.; Cashmore, A.R. HAT3.1, a novel *Arabidopsis* homeodomain protein containing a conserved cysteine-rich region. *Plant J.* **1993**, *4*, 137–150.
6. Champagne, K.S.; Kutateladze, T.G. Structural insight into histone recognition by the ING PHD fingers. *Curr. Drug Targets* **2009**, *10*, 432–441.
7. Martin, D.G.; Baetz, K.; Shi, X.; Walter, K.L.; MacDonald, V.E.; Wlodarski, M.J.; Gozani, O.; Hieter, P.; Howe, L. The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. *Mol. Cell. Biol.* **2006**, *26*, 7871–7879.
8. Wysocka, J.; Swigut, T.; Xiao, H.; Milne, T.A.; Kwon, S.Y.; Landry, J.; Kauer, M.; Tackett, A.J.; Chait, B.T.; Badenhorst, P. A PHD finger of nurf couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **2006**, *442*, 86–90.
9. Reddy, T.V.; Kaur, J.; Agashe, B.; Sundaresan, V.; Siddiqi, I. The duet gene is necessary for chromosome organization and progression during male meiosis in *Arabidopsis* and encodes a PHD finger protein. *Development* **2003**, *130*, 5975–5987.
10. Woo, H.R.; Pontes, O.; Pikaard, C.S.; Richards, E.J. VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes Dev.* **2007**, *21*, 267–277.
11. Yang, X.; Makaroff, C.A.; Ma, H. The *Arabidopsis* MALE MEIOCYTE DEATH1 gene encodes a PHD-finger protein that is required for male meiosis. *Plant Cell* **2003**, *15*, 1281–1295.

12. Barlow, J.H.; Faryabi, R.B.; Callén, E.; Wong, N.; Malhowski, A.; Chen, H.T.; Gutierrez-Cruz, G.; Sun, H.W.; McKinnon, P.; Wright, G. Identification of early replicating fragile sites that contribute to genome instability. *Cell* **2013**, *152*, 620–632.
13. Liang, C.; Zhang, X.; Song, S.; Tian, C.; Yin, Y.; Xing, G.; He, F.; Zhang, L. Identification of UHRF1/2 as new *N*-methylpurine DNA glycosylase-interacting proteins. *Biochem. Biophys. Res. Commun.* **2013**, *433*, 415–419.
14. Liu, X.; Gao, Q.; Li, P.; Zhao, Q.; Zhang, J.; Li, J.; Koseki, H.; Wong, J. UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nat. Commun.* **2013**, *4*, doi:10.1038/ncomms2562.
15. Saleh, A.; Alvarez-Venegas, R.; Yilmaz, M.; Le, O.; Hou, G.; Sadler, M.; Al-Abdallat, A.; Xia, Y.; Lu, G.; Ladunga, I. The highly similar arabidopsis homologs of trithorax ATX1 and ATX2 encode proteins with divergent biochemical functions. *Plant Cell* **2008**, *20*, 568–579.
16. Merkel, D.J.; Wells, S.B.; Hilburn, B.C.; Elazzouzi, F.; Pérez-Alvarado, G.C.; Lee, B.M. The C-terminal region of cytoplasmic polyadenylation element binding protein is a ZZ domain with potential for protein–protein interactions. *J. Mol. Biol.* **2013**, *425*, 2015–2026.
17. Baker, L.A.; Allis, C.D.; Wang, G.G. PHD fingers in human diseases: Disorders arising from misinterpreting epigenetic marks. *Mutat. Res./Fundam. Mol. Mech. Mutagen.* **2008**, *647*, 3–12.
18. Geng, Y.; Gao, Y.; Ju, H.; Yan, F. Diagnostic and prognostic value of plasma and tissue ubiquitin-like, containing PHD and ring finger domains 1 in breast cancer patients. *Cancer Sci.* **2013**, *104*, 194–199.
19. Coscoy, L.; Ganem, D. PHD domains and E3 ubiquitin ligases: Viruses make the connection. *Trends Cell Biol.* **2003**, *13*, 7–12.
20. Gozani, O.; Karuman, P.; Jones, D.R.; Ivanov, D.; Cha, J.; Lugovskoy, A.A.; Baird, C.L.; Zhu, H.; Field, S.J.; Lessnick, S.L. The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor. *Cell* **2003**, *114*, 99–111.
21. Wei, W.; Huang, J.; Hao, Y.J.; Zou, H.F.; Wang, H.W.; Zhao, J.Y.; Liu, X.Y.; Zhang, W.K.; Ma, B.; Zhang, J.S. Soybean GMPHD-type transcription regulators improve stress tolerance in transgenic *Arabidopsis* plants. *PLoS ONE* **2009**, *4*, e7209.
22. Zhao, Y.; Zhou, Y.; Jiang, H.; Li, X.; Gan, D.; Peng, X.; Zhu, S.; Cheng, B. Systematic analysis of sequences and expression patterns of drought-responsive members of the *HD-Zip* gene family in maize. *PLoS ONE* **2011**, *6*, e28488.
23. Schnable, P.S.; Ware, D.; Fulton, R.S.; Stein, J.C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T.A. The B73 maize genome: Complexity, diversity, and dynamics. *Science* **2009**, *326*, 1112–1115.
24. Peng, X.; Zhao, Y.; Cao, J.; Zhang, W.; Jiang, H.; Li, X.; Ma, Q.; Zhu, S.; Cheng, B. CCCH-type zinc finger family in maize: Genome-wide identification, classification and expression profiling under abscisic acid and drought treatments. *PLoS ONE* **2012**, *7*, e40120.
25. Capili, A.D.; Schultz, D.C.; Rauscher, F.J.; Borden, K.L. Solution structure of the PHD domain from the KAP-1 corepressor: Structural determinants for PHD, RING and LIM zinc-binding domains. *EMBO J.* **2001**, *20*, 165–177.

26. Callebaut, I.; Courvalin, J.C.; Mornon, J.P. The BAH (bromo-adjacent homology) domain: A link between DNA methylation, replication and transcriptional regulation. *FEBS Lett.* **1999**, *446*, 189–193.
27. Wang, Y.; Reddy, B.; Thompson, J.; Wang, H.; Noma, K.I.; Yates, J.R.; Jia, S. Regulation of Set9-mediated H4K20 methylation by a PWWP domain protein. *Mol. Cell* **2009**, *33*, 428–437.
28. Qiu, C.; Sawada, K.; Zhang, X.; Cheng, X. The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat. Struct. Mol. Biol.* **2002**, *9*, 217–224.
29. Cannon, S.B.; Mitra, A.; Baumgarten, A.; Young, N.D.; May, G. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* **2004**, *4*, doi:10.1186/1471-2229-4-10.
30. Gaut, B.S.; Morton, B.R.; McCaig, B.C.; Clegg, M.T. Substitution rate comparisons between grasses and palms: Synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10274–10279.
31. Swigoňová, Z.; Lai, J.; Ma, J.; Ramakrishna, W.; Llaca, V.; Bennetzen, J.L.; Messing, J. Close split of sorghum and maize genome progenitors. *Genome Res.* **2004**, *14*, 1916–1923.
32. Narusaka, Y.; Nakashima, K.; Shinwari, Z.K.; Sakuma, Y.; Furihata, T.; Abe, H.; Narusaka, M.; Shinozaki, K.; Yamaguchi-Shinozaki, K. 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.
33. Goto, E.; Ishido, S.; Sato, Y.; Ohgimoto, S.; Ohgimoto, K.; Nagano-Fujii, M.; Hotta, H. c-MIR, a human E3 ubiquitin ligase, is a functional homolog of herpesvirus proteins MIR1 and MIR2 and has similar activity. *J. Biol. Chem.* **2003**, *278*, 14657–14668.
34. Mansouri, M.; Bartee, E.; Gouveia, K.; Hovey Nerenberg, B.T.; Barrett, J.; Thomas, L.; Thomas, G.; Mcfadden, G.; Früh, K. The PHD/LAP-domain protein M153R of myxomavirus is a ubiquitin ligase that induces the rapid internalization and lysosomal destruction of CD4. *J. Virol.* **2003**, *77*, 1427–1440.
35. Feng, Y.; Liu, Q.P.; Xue, Q.Z. Comparative phylogenetic analysis of the rice and *Arabidopsis* PHDfinger proteins. *Acta Genet. Sin.* **2004**, *31*, 1284–1293.
36. Moore, R.C.; Purugganan, M.D. The early stages of duplicate gene evolution. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15682–15687.
37. Lin, Y.X.; Jiang, H.Y.; Chu, Z.X.; Tang, X.L.; Zhu, S.W.; Cheng, B.J. Genome-wide identification, classification and analysis of heat shock transcription factor family in maize. *BMC Genom.* **2011**, *12*, doi:10.1186/1471-2164-12-76.
38. Mainali, H.R.; Chapman, P.; Dhaubhadel, S. Genome-wide analysis of Cyclophilin gene family in soybean (*Glycine max*). *BMC Plant Biol.* **2014**, *14*, doi:10.1186/s12870-014-0282-7.
39. Dou, L.; Zhang, X.; Pang, C.; Song, M.; Wei, H.; Fan, S.; Yu, S. Genome-wide analysis of the *WRKY* gene family in cotton. *Mol. Genet. Genom.* **2014**, *289*, 1103–1121.
40. Liu, Y.; Zhang, D.; Wang, L.; Li, D. Genome-wide analysis of mitogen-activated protein kinase gene family in maize. *Plant Mol. Biol. Rep.* **2013**, *31*, 1446–1460.
41. Van de Peer, Y.; Maere, S.; Meyer, A. The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* **2009**, *10*, 725–732.



42. Maere, S.; de Bodt, S.; Raes, J.; Casneuf, T.; van Montagu, M.; Kuiper, M.; van de Peer, Y. Modeling gene and genome duplications in eukaryotes. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5454–5459.
43. Lai, J.; Ma, J.; Swigoňová, Z.; Ramakrishna, W.; Linton, E.; Llaca, V.; Tanyolac, B.; Park, Y.J.; Jeong, O.Y.; Bennetzen, J.L. Gene loss and movement in the maize genome. *Genome Res.* **2004**, *14*, 1924–1931.
44. Ma, J.; Wang, Q.; Sun, R.; Xie, F.; Jones, D.C.; Zhang, B. Genome-wide identification and expression analysis of TCP transcription factors in gossypium raimondii. *Sci. Rep.* **2014**, *4*, doi:10.1038/srep06645.
45. Lynch, M.; Conery, J.S. The evolutionary fate and consequences of duplicate genes. *Science* **2000**, *290*, 1151–1155.
46. Wang, Y.; Wang, Q.; Zhao, Y.; Han, G.; Zhu, S. Systematic analysis of maize Class III peroxidase gene family reveals a conserved subfamily involved in abiotic stress response. *Gene* **2015**, *566*, 95–108.
47. Blanc, G.; Wolfe, K.H. Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* **2004**, *16*, 1679–1691.
48. Zong, X.J.; Li, D.P.; Gu, L.K.; Li, D.Q.; Liu, L.X.; Hu, X.L. Abscisic acid and hydrogen peroxide induce a novel maize group C MAP kinase gene, *ZmMPK7*, which is responsible for the removal of reactive oxygen species. *Planta* **2008**, *229*, 485–495.
49. Yang, A.; Dai, X.; Zhang, W.H. A R2R3-type myb gene, *OsMYB2*, is involved in salt, cold, and dehydration tolerance in rice. *J. Exp. Bot.* **2012**, *63*, 2541–2556.
50. Bienz, M. The PHD finger, a nuclear protein-interaction domain. *Trends Biochem. Sci.* **2006**, *31*, 35–40.
51. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. The clustal\_x windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **1997**, *25*, 4876–4882.
52. Clewley, J.P.; Arnold, C. Megalign. In *Sequence Data Analysis Guidebook*; Springer: New York, NY, USA, 1997; pp. 119–129.
53. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425.
54. Yang, S.; Zhang, X.; Yue, J. X.; Tian, D.; Chen, J. Q. Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mol. Genet. Genom.* **2008**, *280*, 187–198.
55. Zhao, P.; Li, Q.; Li, J.; Wang, L.; Ren, Z. Genome-wide identification and characterization of R2R3MYB family in solanum lycopersicum. *Mol. Genet. Genom.* **2014**, *289*, 1183–1207.
56. Rozas, J.; Sánchez-DelBarrio, J.C.; Messeguer, X.; Rozas, R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **2003**, *19*, 2496–2497.
57. Quraishi, U.M.; Abrouk, M.; Murat, F.; Pont, C.; Foucrier, S.; Desmaizieres, G.; Confolent, C.; Rivière, N.; Charmet, G.; Paux, E. Cross-genome map based dissection of a nitrogen use efficiency ortho-metaqtl in bread wheat unravels concerted cereal genome evolution. *Plant J.* **2011**, *65*, 745–756.

58. Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.H.; Jin, H.; Marler, B.; Guo, H. *MCSscanx*: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, doi:10.1093/nar/gkr1293.
59. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An information aesthetic for comparative genomics. *Genome Res.* **2009**, *19*, 1639–1645.
60. Higo, K.; Ugawa, Y.; Iwamoto, M.; Korenaga, T. Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res.* **1999**, *27*, 297–300.
61. Sekhon, R.S.; Briskine, R.; Hirsch, C.N.; Myers, C.L.; Springer, N.M.; Buell, C.R.; de Leon, N.; Kaeppeler, S.M. Maize gene atlas developed by rna sequencing and comparative evaluation of transcriptomes based on RNA sequencing and microarrays. *PLoS ONE* **2013**, *8*, e61005.
62. Lu, G.; Gao, C.; Zheng, X.; Han, B. Identification of OsbZIP72 as a positive regulator of ABA response and drought tolerance in rice. *Planta* **2009**, *229*, 605–615.
63. Xiang, Y.; Huang, Y.; Xiong, L. Characterization of stress-responsive *CIPK* genes in rice for stress tolerance improvement. *Plant Physiol.* **2007**, *144*, 1416–1428.
64. Schmittgen, T.D.; Livak, K.J. Analyzing real-time PCR data by the comparative ct method. *Nat. Protoc.* **2008**, *3*, 1101–1108.
65. Chung, E.; Seong, E.; Kim, Y.C.; Chung, E.J.; Oh, S.K.; Lee, S.; Park, J.M.; Joung, Y.H.; Choi, D. A method of high frequency virus induced gene silencing in Chili pepper (*Capsicum annuum* L. Bukang). *Mol. Cells* **2004**, *17*, 377–380.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).