Molecular Plant
Research Article

CellPress
PARTNER JOURNAL

Genetic Contribution of Paleopolyploidy to Adaptive Evolution in Angiosperms

Shengdan Wu[1,2], Baocai Han[1] and Yuannian Jiao[1,2,*]

[1]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

[2]University of Chinese Academy of Sciences, Beijing 100049, China

***Correspondence: Yuannian Jiao (jiaoyn@ibcas.ac.cn)**

https://doi.org/10.1016/j.molp.2019.10.012

ABSTRACT

Ancient whole-genome duplications (WGDs or polyploidy) are prevalent in plants, and some WGDs occurred during the timing of severe global environmental changes. It has been suggested that WGDs may have contributed to plant adaptation. However, this still lacks empirical evidence at the genetic level to support the hypothesis. Here, we investigated the survivors of gene duplicates from multiple ancient WGD events on the major branches of angiosperm phylogeny, and aimed to explore genetic evidence supporting the significance of polyploidy. Duplicated genes co-retained from three waves of independent WGDs (~120 million years ago [Ma], ~66, and <20 Ma) were investigated in 25 selected species. Gene families functioning in low temperature and darkness were commonly retained gene duplicates after the eight independently occurring WGDs in many lineages around the Cretaceous-Paleocene boundary, when the global cooling and darkness were the two main stresses. Moreover, the commonly retained duplicates could be key factors which may have contributed to the robustness of the critical stress-related pathways. In addition, genome-wide transcription factors (TFs) functioning in stresses tend to retain duplicates after waves of WGDs, and the coselected gene duplicates in many lineages may play critical roles during severe environmental stresses. Collectively, these results shed new light on the significant contribution of paleopolyploidy to plant adaptation during global environmental changes in the evolutionary history of angiosperms.

Key words: whole-genome duplication, paleopolyploidy, adaptive evolution, phylogenomic, Cretaceous-Paleocene boundary, gene regulatory network

Wu S., Han B., and Jiao Y. (2020). Genetic Contribution of Paleopolyploidy to Adaptive Evolution in Angiosperms. Mol. Plant. **13**, 59–71.

INTRODUCTION

Angiosperms (or flowering plants) are the most diverse and abundant in the plant kingdom, with about 350,000 known species on Earth. Charles Darwin described the rapid rise and early diversification of angiosperms from the middle to late Cretaceous period as "an abominable mystery" (Friedman, 2009). Currently, angiosperms constitute the dominant vegetation of the Earth's surface, covering regions from tropical to polar terrestrial zones, as well as aquatic habitats. The success is speculated because of, to some extent, prevalent whole-genome duplication (WGD) events in the evolutionary history of angiosperms (Levin, 1983; Soltis et al., 2009; Van de Peer et al., 2009, 2017). WGD has long been recognized as an important evolutionary force for speciation, adaptation, and diversification (Wood et al., 2009; Soltis and Soltis, 2016).

Within recent two decades, tremendous efforts have shown that WGDs are far more prevalent than previously thought in the evolutionary history of flowering plants (Bowers et al., 2003; Blanc and Wolfe, 2004; Cui et al., 2006; Soltis et al., 2009; Jiao et al., 2011; Renny-Byfield and Wendel, 2014; Van de Peer et al., 2017). Two ancestral WGDs were identified before the diversification of extant angiosperms and seed plants, respectively (Jiao et al., 2011). Two major clades in angiosperms, eudicots and monocots, both experienced paleopolyploidization events early in their evolutionary history, named gamma (γ) and tau (τ) (Jaillon et al., 2007; Tang et al., 2010; Jiao et al., 2012, 2014; Vekemans et al., 2012; Ming et al., 2015). In addition, WGDs also occurred in the common ancestors of many species-rich groups, such as Asteraceae, Brassicaceae, Cucurbitaceae, Fabaceae, and Poaceae (Cannon et al., 2015; Edger et al., 2015; Huang et al., 2016; McKain et al., 2016; Ren et al., 2018; Wang et al., 2018). Especially, WGDs recurrently occurred in many lineages. For example, three rounds of WGDs (γ-β-α) occurred in the

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and IPPE, CAS.

evolutionary history of *Arabidopsis thaliana* after a split from monocots (Bowers et al., 2003); and the lineage of *Musa* (bananas) independently experienced three rounds of WGDs after a split with Poaceae (D'Hont et al., 2012).

Furthermore, previous studies found that the timing of WGDs are not randomly distributed across the phylogeny of angiosperms, indicating possible roles of WGDs under environmental selection. A wave of ancient WGDs occurred around the Cretaceous-Paleocene (K-Pg) boundary independently in many plant lineages, suggesting that WGDs potentially helped species survive the extinction event (Fawcett et al., 2009; Vanneste et al., 2014). It also has been proposed that polyploidization was associated with C_4 grassland expansion during the late Miocene, as well as with adaptation to recent glaciation maxima (Estep et al., 2014; Novikova et al., 2018). Therefore, WGD has been speculated to be associated with extinction events and other extreme environmental changes. However, evidence at a genetic level supporting the contribution and significance of WGDs to adaptation remains largely unexplored.

It is well acknowledged that polyploidy simultaneously duplicates tens of thousands of genes by adding one extra set of genomes, which provides a large amount of raw genetic materials for evolution (Adams and Wendel, 2005; Doyle et al., 2008; Hegarty and Hiscock, 2008; Soltis et al., 2015; Van de Peer et al., 2017). During the subsequent fractionation and diploidization processes, a large proportion of genes will quickly return to a single-copy state (Lynch and Conery, 2000), while the retained ones are considered of particular importance to genetic innovation through neofunctionalization and subfunctionalization (Ohno, 1970; Force et al., 1999). In addition, the duplicated genes might also result in changes of gene regulatory networks (GRNs) (Conant, 2010; De Smet and Van de Peer, 2012), which could potentially contribute to plant adaptation.

To explore the significance of WGDs, here we comprehensively tracked the evolutionary history of global gene families in 25 plant genomes, and investigated the genetic modifications after independent WGDs. Firstly, 25 sequenced plant genomes representing major lineages of angiosperms were used to reconstruct global gene families, and phylogenomic analyses were performed to identify the gene families that retained duplicated genes after ancient WGDs. Then, we identified gene families with retained duplicates after independent WGDs from certain periods with extreme environmental changes, looking for potential selection signatures at the genetic level. Finally, by reconstructing GRNs from RNA sequencing (RNA-seq) data and integrating previously known pathways, we provided evidence showing how retained duplications have contributed to reshaping the GRNs in response to environmental stresses.

RESULTS

Identification of Paralogous Genes Retained after WGDs

To identify genetic contribution of WGDs, we investigated 21 well-acknowledged polyploidization events in the evolutionary history of angiosperms (Figure 1). We selected 25 sequenced plant genomes (Supplemental Table 1) and constructed putative gene families from their protein-coding sequences. In total, 66,509 orthogroups were constructed by OrthoMCL (Li et al., 2003). Among them, 12,077 orthogroups with four or more genes and including at least one gene from outgroups (*Physcomitrella patens*, *Selaginella moellendorffii*, and *Amborella trichopoda*) were used to reconstruct maximum likelihood phylogenetic trees (see Methods). Then, phylogenomic analyses were performed to look for duplication events from each gene family phylogeny as described previously (Jiao et al., 2011).

To largely ensure that the duplications were generated from WGDs, tandem duplications were removed firstly based on the chromosomal positions of the duplicated genes, which could screen out some ancient small-scale duplications (see Methods). Furthermore, synteny analyses of each species were able to provide synteny support for a large proportion the duplications identified in this study (Figure 1 and Supplemental Figure 1). In addition, some duplications were hard to classify due to a lack of branching species between two consecutive WGD events, such as ρ or σ in Poales, and α/β^{M} or γ^{M} in *Musa acuminata*. We employed a synonymous substitution per synonymous site (K_S) approach to distinguish the gene survivors from certain duplication events (see Methods). Together, we were able to gather the gene families with surviving duplicates from each WGD event (Figure 1).

Gene Retention Pattern after WGDs Occurred in Three Periods

WGDs could generate a large number of duplicated genes in one event, which provides a tremendous amount of raw genetic material for evolution. If the independent WGDs helped species survive environmental changes, similar functional gene duplicates would be retained in these different species as they might have been selected by common environmental pressures.

The previously identified and dated 21 WGDs were denoted on the species tree of selected land pants, 14 of which could be classified into three waves based on periods of the timing of occurrence (Figure 1 and Supplemental Table 2). The most ancient wave was around ~120 million years ago (Ma) when γ (Jaillon et al., 2007; Jiao et al., 2012; Vekemans et al., 2012) and τ (Jiao et al., 2014; Ming et al., 2015) events occurred in the early evolutionary history of eudicots and monocots, respectively. The second one is the well-known wave around the K-Pg boundary, when a large number of WGD events occurred (Paterson et al., 2004; Tuskan et al., 2006; Rensing et al., 2008; Fawcett et al., 2009; Schmutz et al., 2010; D'Hont et al., 2012; The Tomato Genome Consortium, 2012; Singh et al., 2013; Vanneste et al., 2014). The relatively recent wave was within 20 Ma during which four independent WGDs occurred in the evolutionary history of *Glycine max*, *Panicum virgatum*, *Tarenaya hassleriana*, and *Zea mays* (Blanc and Wolfe, 2004; Schmutz et al., 2010; Cheng et al., 2013; Lu et al., 2013).

We found certain gene families indeed survived gene duplicates from several independent WGDs in many species (Figure 2A), which were likely the signal of selection from the specific stress environments. Sixty-six gene families commonly retained gene duplicates in three periods (Supplemental Figure 2), which were

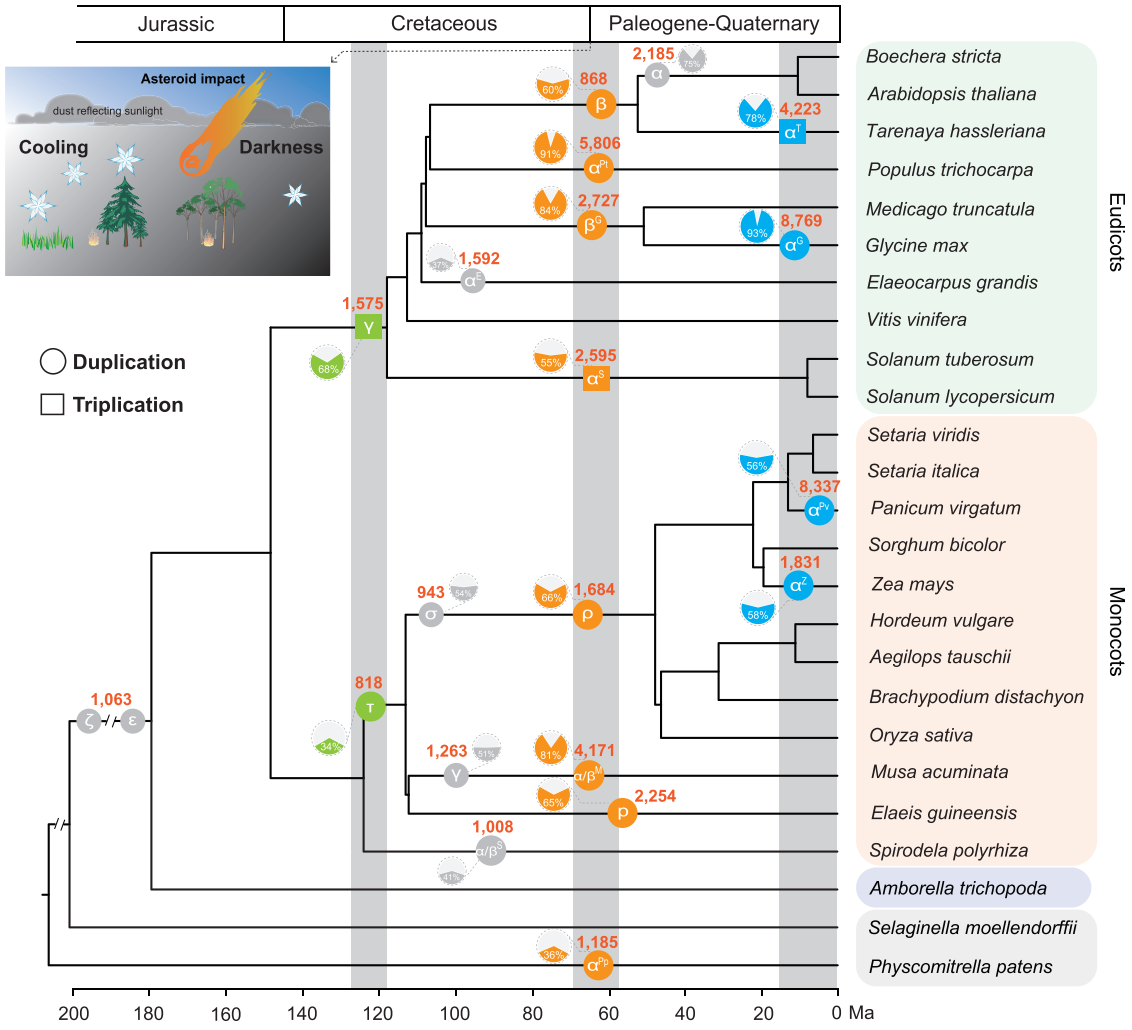


Figure 1. Inferring Number of Gene Families Surviving Duplicates after Each WGD in the Evolutionary History of Land Plants.
Phylogenetic tree showed the topology and divergence times of 25 plants in this study. The evolutionary relationships of the 25 species were based on current accepted topology (Angiosperm Phylogeny Website). Divergence time of each node of the species tree was obtained from the TimeTree Website (http://timetree.org/). Well-acknowledged whole-genome duplication (circle) and triplication (square) events were positioned onto the branches of the phylogeny. Three periods (~120, ~66, and <20 Ma) with prolific WGDs were recognized, and are denoted in green, orange, and blue, respectively. The number of gene families with duplicates retention following each WGD are shown around the corresponding circle or square. The proportion of the duplications verified by synteny evidence were generated from WGDs, which were indicated in the dashed circles. A sketch map in the upper left shows the major environmental stresses during the Cretaceous-Paleogene extinction period.

mainly protein kinases, transporters, and protein binding gene families (Supplemental Table 3). Three hundred and twenty gene families retained duplicates from the most ancient wave of WGDs (γ and τ), which were enriched for genes functioning in response to water deprivation and salt stress (Figure 2B). These survivors may be, at least in part, selected by the arid climate around 120 Ma of the Cretaceous (Heimhofer et al., 2005). The second wave of WGDs were during the K-Pg boundary with severe environmental changes, including global cooling, darkness, acid rain, and wildfires (Nichols and Johnson, 2008; Schulte et al., 2010). Four hundred and ninety-three gene families retained gene duplicates from at least six independent WGDs during the K-Pg boundary (Figure 2A), which were enriched for many stress-related gene ontology (GO) terms including cold, heat, osmotic, salt stress, water deprivation, and wounding (Figure 2B), as well as several other biological processes associated with stress response (e.g., the abscisic acid signaling pathway, cellular response to phosphate starvation, defense response, and response to karrikin) (Supplemental Figure 3). We also investigated the other five lineages without paleopolyploidization events during the K-Pg boundary and found that they retained small-scale duplications in 12 gene families (Supplemental Figure 4). However, these gene families mainly encode enzymes or transporters in plant metabolic processes, which are not directly related to environmental adaptation (Supplemental Table 4). The most recent wave of WGDs were within 20 Ma and retained duplicates from 844 gene families (Figure 2A), of which the functional category enrichments are responses to salt stress, cold stress, water deprivation, and wounding (Figure 2B). The recorded environmental changes were low CO_2 concentrations and relatively cool temperatures during that period (Zachos

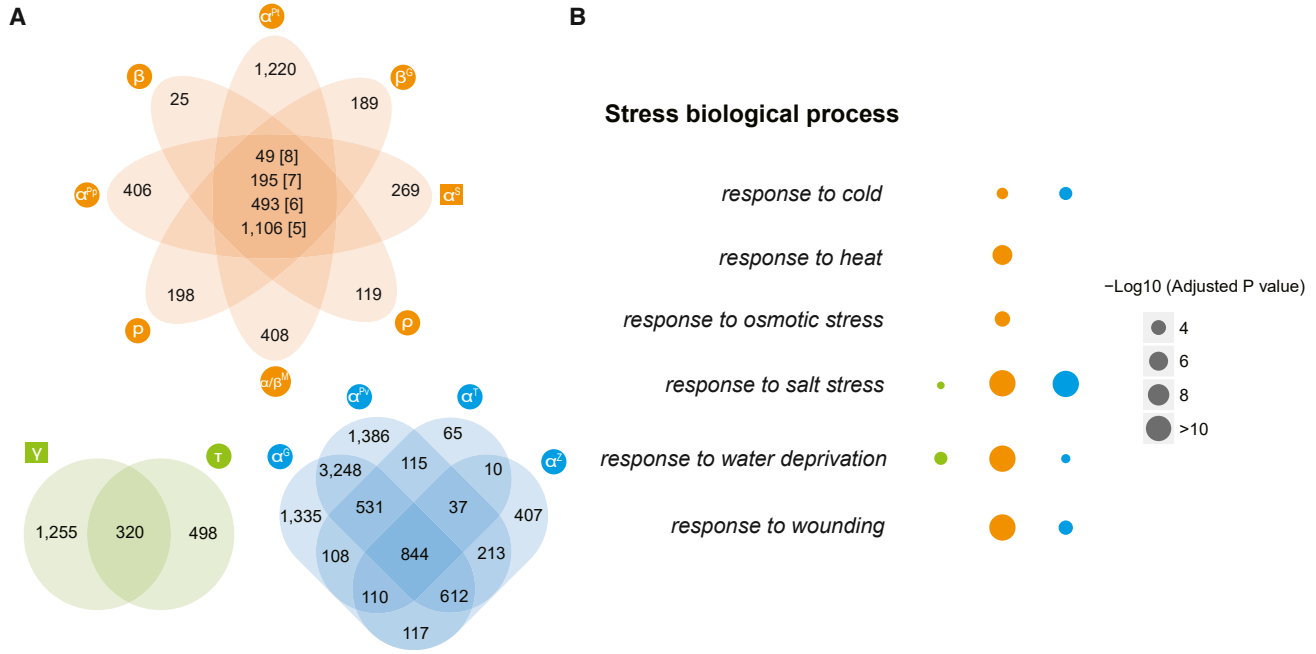


Figure 2. Retention Patterns of Stress-Related Gene Families in Three Periods.
(A) Venn diagram showing the shared and specific gene families surviving duplicates after multiple WGDs in certain periods. Numbers represent the number of gene families with gene duplications. Numbers in square bracket indicate number of WGDs with sharing gene families surviving duplicates. **(B)** The significantly enriched GO terms of stress-related biological processes for the shared gene families retained gene duplicates in three periods. The three columns with different colors are corresponding to the WGDs in three periods as in **(A)**.

et al., 2008). Our enriched GO terms might partially explain the environmental changes, but also suggest the presence of other environmental selection for different lineages.

Biased Retention of Transcription Factor Gene Families after WGDs

TFs play a critical role in the transcriptional regulation of genes involved in many biological processes (e.g., growth, development, and stress responses) (de Mendoza et al., 2013). Previous studies demonstrated that TFs are the vastly over-retained genes after WGDs (Maere et al., 2005; Freeling, 2009).

We examined the retention pattern of gene duplicates of TFs after three waves of WGDs based on the retention value (*R* value, see Methods). In general, the majority of the TF genes tend to be retained after WGDs (Figure 3), which is consistent with previous analyses (Maere et al., 2005; Freeling, 2009). However, we found that not all TF gene families were over-retained, and that different families of TF showed certain retention preferences (Figure 3). For example, high-retention gene families, including ARF, C2H2, C3H, CO-like, ERF, G2-like, GRAS, HD-ZIP, HSF, LBD, MYB, NAC, Trihelix, WRKY, bHLH, and bZIP gene families, tend to repeatedly retain duplicates after WGDs independent of the evolutionary periods and diverse lineages (Figure 3). Many of the high-retention TF gene families are involved in diverse development processes and response to abiotic and biotic stresses (Khan et al., 2018). However, some TFs were lowly retained after many WGDs, such as FAR1, HB-PHD, HRT-like, LFY, LSD, NF-X1, S1Fa-like, STAT, SAP, and Whirly, suggesting the conservative function and dosage of these TFs (Figure 3). Most low-retention TF gene families were functioning in conserved biological processes. For instance, *LFY* controls the switch from vegetative to reproductive development (William et al., 2004) and *LSD1* negatively regulates plant cell death pathway (Dietrich et al., 1997).

Duplicated TF genes, which were co-retained after specific waves of WGDs, are considered as critical genetic contributions for species surviving environmental changes. Co-retained TF genes at ~120 Ma were mainly involved in plant growth, development, morphogenesis, and stress response (Supplemental Table 5). For example, the retained duplicate genes and their functional divergence in four orthogroups of the MADS-box gene family likely contributed to the morphological novelty of floral organs in both core eudicots and monocots (Zhao et al., 2017). Two orthogroups of heat stress transcription factor (HSF) play roles in responding to heat stress. However, co-retained TFs at the K-Pg boundary were mainly involved in responding to various abiotic stresses (Supplemental Table 6). Orthogroups of the C2H2, ERF, and RAV families were involved in the response to low temperatures. Orthogroups of the HD-ZIP family were involved in the shade-avoidance syndrome and dehydration stress responses, respectively. Orthogroup of the WRKY family were involved in the response to low phosphate stress (Supplemental Table 6).

Contribution of WGDs on the Complexity of GRNs

As WGDs could potentially rewire the GRNs (Conant, 2010; De Smet and Van de Peer, 2012), we aimed to explore the contribution of WGDs on reshaping networks during adaptation to environmental changes following the K-Pg boundary (Alvarez et al., 1980; Nichols and Johnson, 2008; Schulte et al., 2010).

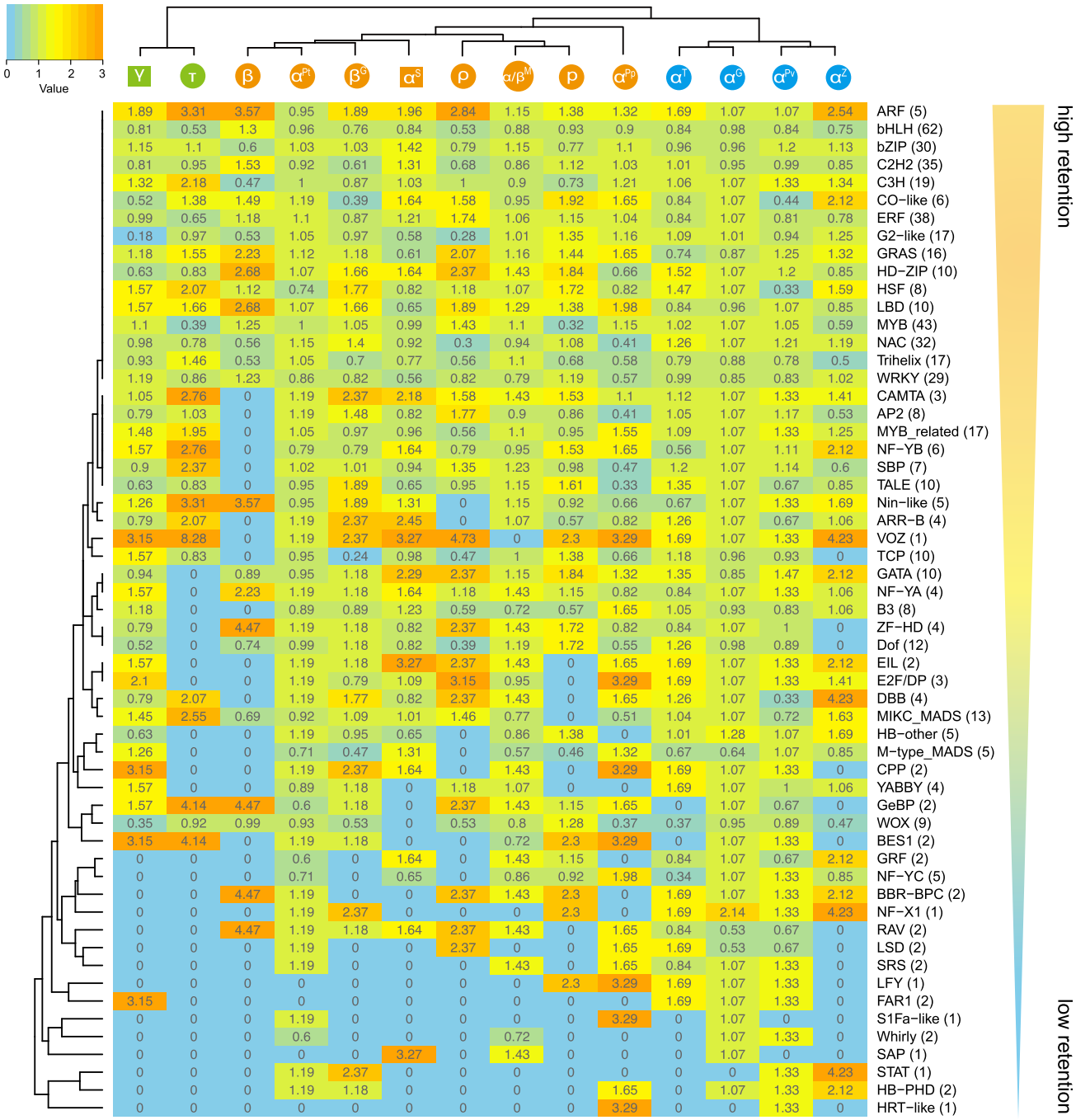


Figure 3. Biased Retention Patterns of Transcription Factor Gene Families after WGDs.

TFs (rows) were clustered based on their retention values, and WGDs (columns) were grouped according to their occurring timing. Gene families to the top of the heatmap were the high retention ones after WGDs, while TFs to the bottom of the heatmap were the low retention ones. Color key on the upper left denotes the retention values of the TFs. The number in each cell of the heatmap represents the retention value of each TF after corresponding WGD. The numbers in parentheses after the TF names represent the total number of orthogroups belonging to the TF gene families.

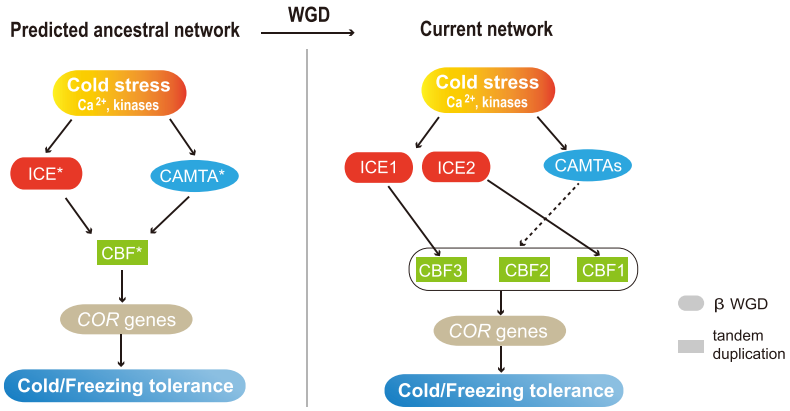
Global cooling (or low temperatures) was a major environmental stress during the mass extinction period (Schulte et al., 2010), and the C-repeat/DREB binding factor (CBF)-dependent signaling pathway is the well-known major cold signaling pathway (Chinnusamy et al., 2007; Shi et al., 2015, 2018). Currently, the core components of the CBF-dependent signaling pathway have been deciphered in *A. thaliana* (Shi et al., 2015). *CBF* genes, as key components in the pathway, are regulated by upstream ICE and CAMTA TFs (Shi et al., 2015; Zhao et al., 2015), and are able to trigger the expression of many cold-responsive (*COR*) genes under cold stress (Chinnusamy et al., 2007).

By tracking the evolutionary history of key gene families in the CBF pathway, we found that the CBF, ICE, CAMTA, and

A

Ortho ID	Cold pathway genes	β	α^{Pt}	β^{G}	α^{S}	ρ	α/β^{M}	ρ	α^{Pp}
3343	ICE	●	●	x	x	●	●	x	●
610	CAMTA	x	●	●	x	●	●	●	x
568	CBF	x	●	●	●	●	●	●	x
962	SIZ	x	●	●	●	●	●	●	●
110	OST	x	●	●	●	●	●	x	x
1242	EIN	x	●	●	●	●	●	x	x
926	FRY	x	●	x	●	●	●	●	x

B



C

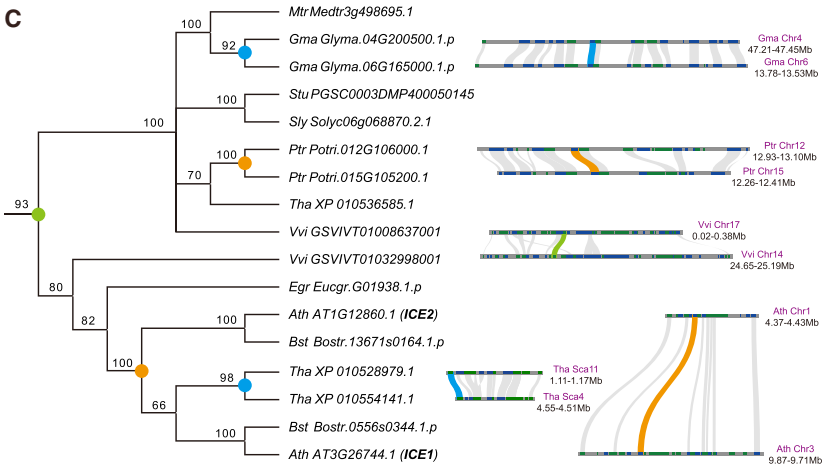


Figure 4. The Duplication Pattern of Key Genes in Cold-Responsive Pathway after WGDs around the K-Pg Boundary.

(A) Summary of duplicates retention status of known important gene families in CBF-dependent signaling pathway in angiosperms after eight WGDs at the K-Pg boundary. The ICE, CAMTA, and CBF are the key transcription factor gene families, and the SIZ, OST, EIN, and FRY are other related gene families involved in CBF-dependent signaling pathway. "x" denotes no retention and solid dots indicate gene retentions.

(B) An illustration of expansion and remodeling of CBF-dependent signaling pathway following WGD in *A. thaliana*. *ICE1* and *ICE2* were duplicated from β WGD. *CBF1*, *CBF2*, and *CBF3* were generated by tandem duplications.

(C) Phylogeny of the ICE gene family showed the duplications in its evolutionary history. Solid circles indicate duplications occurred in different periods. Numbers on branches show the bootstrap supporting values. Syntenic blocks with *ICE* genes were placed on the right of the phylogenetic tree.

other related families (SIZ, EIN, and so forth) presented as duplicated states in many different lineages (Figure 4A). The *ICE1-ICE2* were duplicated from β WGD in *Arabidopsis* (Figure 4B and 4C). The *ice1* loss-of-function mutant is sensitive to cold stress, which leads to significant reduction of survival rate than the wild type (Chinnusamy et al., 2003). Overexpression of *ICE2* greatly enhanced the cold tolerance in transgenic plants (Fursova et al., 2009). In *Oryza sativa*, *CBF* genes were also retained as duplicated copies after the ρ WGD, which also play an important role in cold stress (Supplemental Figure 5). Therefore, duplicated genes that were retained from WGDs occurred during the K-Pg boundary in difference lineages have largely contributed to the copy number (probably dosage at first) and the complexity of the current CBF-dependent signaling network functioning in cold stress tolerance in eudicots and monocots (Figure 4B).

We also performed a comparison of the network of CBF pathway members following the polyploidization events of certain lineages. Coexpression networks have been widely used for identifying functional related genes (Obayashi and Kinoshita, 2010; You et al., 2016; Obayashi et al., 2018). For investigation of network evolution in *Arabidopsis* lineage, *Vitis* is an ideal outgroup that experienced nonadditional WGD after the γ event. We constructed the cold-specific coexpression networks for *A. thaliana* and *Vitis vinifera* using 162 and 60 RNA-seq data, respectively (see Methods). For the duplicated *ICE* genes from β WGD, we examined the coexpression networks of *AthICE1* and *AthICE2* in *Arabidopsis* and their orthologous *VviICE* in *Vitis* (Figure 5). Most of the coexpressed genes in *VviICE* module have orthologous genes clustered in the *AthICE1* and *AthICE2* modules, and the corresponding orthologous genes in *Arabidopsis* could be divided into three sets: one set specifically coexpressed with *AthICE1*, one set specifically coexpressed with *AthICE2*, and one set coexpressed with both *AthICE1 and AthICE2* (Figure 5), indicating subfunctionalization of the duplicated *ICE* genes after WGD. Moreover, the module of *AthICE1* and *AthICE2* in *Arabidopsis* is twice as large as the module of *VviICE* by recruiting additional genes into the network after β and α WGDs, which might potentially increase the cold stress tolerance.

Darkness (or low light) was another major environmental stress encountered by species during the mass extinction period, due to the atmospheric dust reflecting sunlight over a long period (Schulte et al., 2010). We investigated the key components in

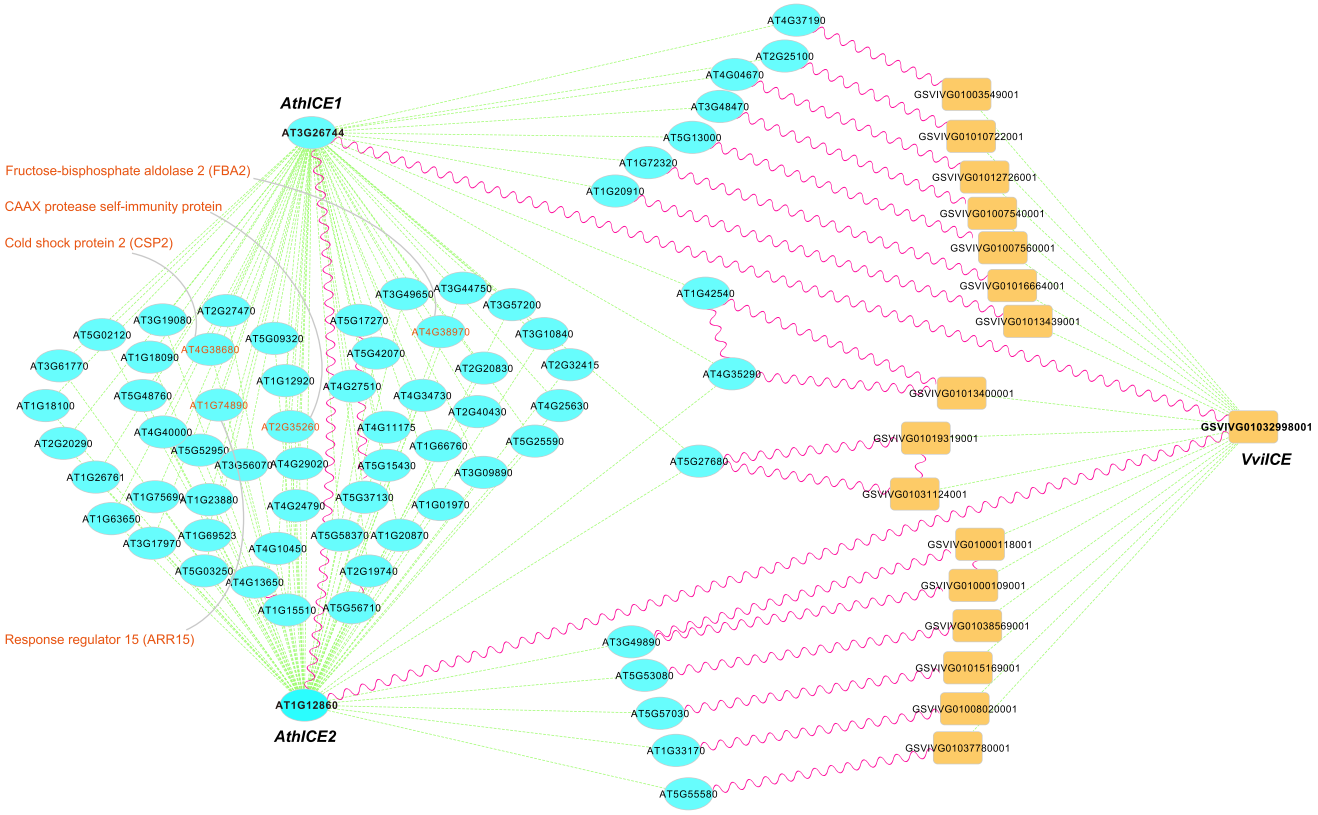


Figure 5. Comparison of Coexpression Networks between Duplicates of *ICE1* and *ICE2* in *A. thaliana* and the Orthologous *ICE* in *V. vinifera*.

ICE1 and *ICE2* were generated by β WGD. The red sinewave lines link corresponding orthologous pairs as they clustered in the same orthogroup. Green dashed lines between two nodes indicate positive coexpression relationships. Four genes in the *Arabidopsis* coexpression network, which have been previously demonstrated function in cold treatment response, were highlighted with annotation information.

the shade avoidance pathway in plants (Jiao et al., 2007; Ruberti et al., 2012), and also found that several key genes were duplicated from the WGDs in multiple lineages (Figure 6A). In *A. thaliana*, the *ATHB2* and *HAT1* in the HD-ZIP II gene family, which function in shade avoidance response, were derived from the β WGD (Figure 6B and 6C). Molecular genetic analyses revealed that *ATHB2* is rapidly induced by low red:far red light in *Arabidopsis*, and the *athb2* loss-of-function mutant displays significant reduction of hypocotyl elongation and shade avoidance ability compared with the wild type (Carabelli et al., 2013). By using *Arabidopsis* as an example, a putative model for network evolution from pre-WGD to post-WGD is illustrated in Figure 6B. Despite that the predicted ancestral network is somewhat uncertain, our results showed clear evidence for the expansion of a shade avoidance pathway after WGDs, which may enhance the perception of light signals and better adapt to low light environment.

To test the possible link between WGDs and plant adaptation, we compared the specific retention pattern of regulatory genes in response to the cold and dark stresses after the three waves of WGDs. The regulatory genes of the cold stress pathway have a higher chance to be retained after the recent two waves of WGDs (~66 and <20 Ma), when the global cooling has been recorded during these two periods (Nichols and Johnson, 2008; Zachos et al., 2008; Schulte et al., 2010) (Supplemental Figure 6). The global darkness was only reported during the K-Pg boundary (Nichols and Johnson, 2008; Schulte et al., 2010). Genes in the shade avoidance pathway have a particularly higher retention after the WGDs around the K-Pg boundary than the retention after the other two waves of WGDs (Supplemental Figure 6). In addition, we further investigated another stress pathway (Na^+ tolerance), despite high Na^+ not being the main global stress during the K-Pg boundary. The Salt Overly Sensitive (SOS) signaling pathway has functions in maintaining ion homeostasis under high Na^+ tolerance (Ji et al., 2013; Supplemental Figure 7A). Duplicates of the core members of the SOS pathway, such as SOS3, ScaBP8, SOS2, and SOS1, were only biasedly retained after the examined WGDs (Supplemental Figures 6 and 7B). Therefore, the preferential retention of the key members in stress-related networks after multiple independent WGDs may serve as critical evidence supporting the contribution of WGDs to the adaptation of species during the global environmental changes.

DISCUSSION

The Nature of Periodic Occurrence of Ancient WGDs in Angiosperms

To bridge the gap of the genetic contribution of paleopolyploidizations with adaptive evolution in general, we need to explore

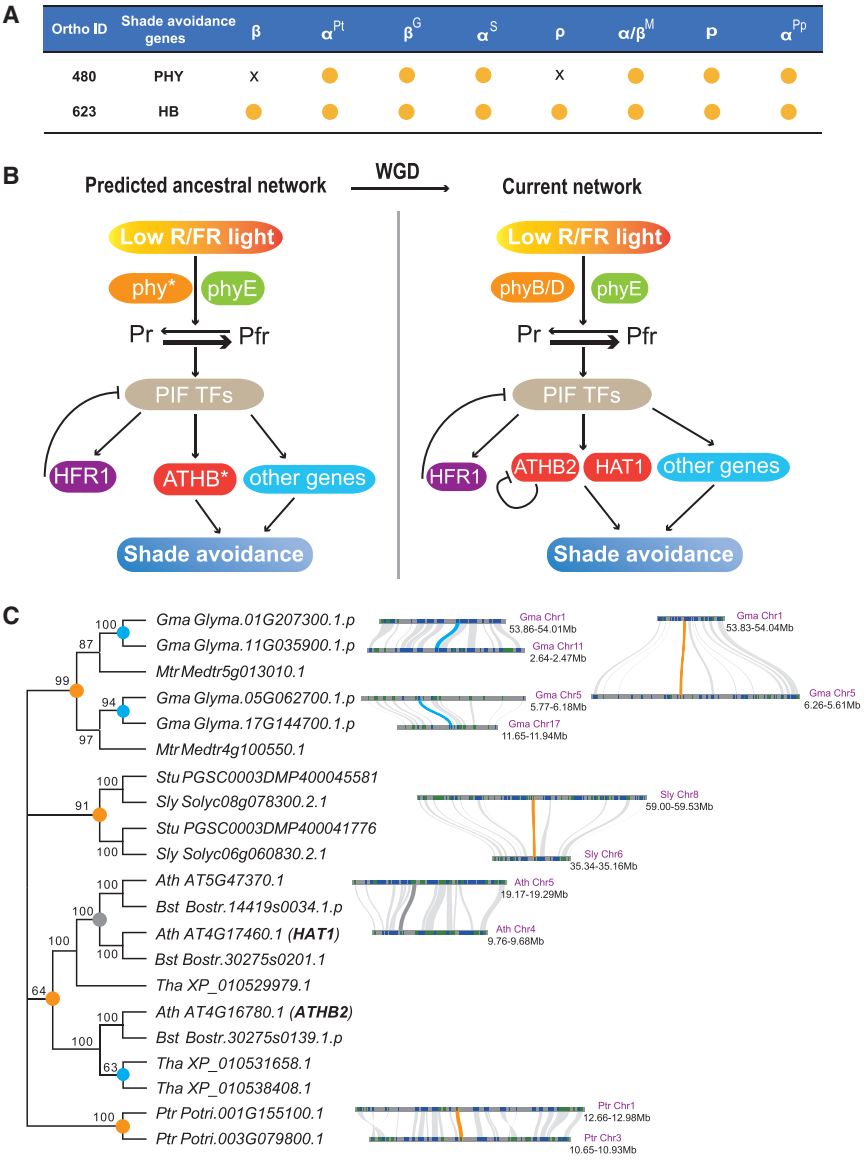


Figure 6. The Retention Pattern of Key Genes in Shade Avoidance Pathway after WGDs around the K-Pg Boundary.

(A) Summary of retention status of the PHY and HB gene families after eight independent WGDs around the K-Pg boundary. PHY and HB are the two major gene families in shade avoidance pathway. "x" denotes no retention and solid dots indicate gene retentions.

(B) An illustration of expansion and remodeling of shade avoidance pathway following WGD by comparing a predicted ancestral network with the current network in *A. thaliana*. *ATHB2* and *HAT1* were generated by β WGD.

(C) Phylogeny of the HD-ZIP II gene family showed the duplications in its evolutionary history. Solid circles indicate duplications that occurred in different periods. Numbers on branches show the bootstrap supporting values. Syntenic blocks with HD-ZIP II genes were placed on the right of the phylogenetic tree.

the empirical adaptive genetic signatures of many WGDs during the evolutionary history of angiosperms. Polyploids are very common in nature. However, the nascent polyploid individuals tend to encounter internal and external obstacles, including increased rates of chromosome segregation errors, small effective population size, competition with progenitor diploid species, and so on (Comai, 2005; Arrigo and Barker, 2012). Several studies suggested that polyploidy is usually an evolutionary dead end (Stebbins, 1950; Mayrose et al., 2011). Recently formed polyploid plants have to find certain ecological niches that are different from corresponding diploid species to survive (Stebbins, 1950; Levin, 1983; Ramsey, 2011; te Beest et al., 2012; Visger et al., 2016). Polyploid plants could move to a new, but stressful environment with no competition with their ancestral diploids, or survive after a strong environmental selection that swapped the diploid ancestors for polyploids (Otto and Whitton, 2000; Brochmann et al., 2004; Ramsey, 2011; te Beest et al., 2012; Chao et al., 2013; Parisod and Broennimann, 2016). Therefore, the paleopolyploidization events in angiosperms clustered and co-occurred with past global environmental changes, which might have played significant roles in the establishment of polyploids (Van de Peer et al., 2009, 2017).

Challenges to Infer the Evolutionary Significance of WGDs

Because of the recurrent occurrences and subsequent large-scale gene losses, the remaining signals of genetic contribution from ancient WGDs became complicated and ambiguous (Doyle et al., 2008; Schnable et al., 2011; Wendel et al., 2016). In addition, the environmental selection pressure usually do not last as long as tens of million years. The increased novel genetic contribution of ancient WGDs might have been lost after the environmental conditions changed. Moreover, hybridization and recombination could also remove the critical genetic information that helped species survive severe environmental changes during a particular period. Therefore, it poses challenges to infer the significance of WGDs in the evolutionary history of angiosperms. Here, we have to include many high-quality completely sequenced genomes sharing one ancient WGD in our analysis to avoid missing critical genes due to incomplete and/or improper genome assembly and annotation. More importantly, although these critical genes might be lost in some species, we might still be able to piece together a broad picture from simultaneous consideration of many species. Finally, we investigated several independent WGDs that occurred at same period to look for shared duplicated genes. Therefore, we were able to identify critical genetic signals for species surviving dramatic environmental changes, and propose the likely evolutionary consequences of WGDs for plant adaptation.

Genetic Evidence Sheds New Light on the Contribution of WGDs to Adaptation

In addition, to illustrate the genetic impact of individual WGDs, we also need to consider certain waves of ancient WGDs independently occurring in different lineages. The severe environmental changes should have posed similar selection for all species on Earth. Previous studies have demonstrated the biased retention for genes related to regulation and development (Maere et al., 2005; Freeling, 2009). After comprehensive investigation of gene families, we found that certain functional genes were duplicated after independent WGDs occurred during the same period, which provides likely evidence supporting global environmental selection on the paleopolyploids from different lineages. For example, in response to low-temperature and low-light environmental changes during the K-Pg boundary, the second wave of ancient WGDs have contributed to reshape CBF-dependent signaling (Figures 4 and 5) and shade avoidance pathways (Figure 6). The duplicates in ICE and CBF gene families are recruited in the pathway which indeed enhanced cold tolerance in plants (Shi et al., 2015).

METHODS

Genome Data

We selected 25 sequenced plant genomes representing major lineages of angiosperms and having clear WGD records during their evolutionary history. The studied species included 10 eudicots (*A. thaliana*, *Boechera stricta*, *Eucalyptus grandis*, *G. max*, *Medicago truncatula*, *Populus trichocarpa*, *Solanum lycopersicum*, *Solanum tuberosum*, *Tarenaya hassleriana*, and *V. vinifera*), 12 monocots (*Brachypodium distachyon*, *O. sativa*, *P. virgatum*, *Sorghum bicolor*, *Setaria italic*, *Spirodela polyrhiza*, *Setaria viridis*, *Z. mays*, *Aegilops tauschii*, *Elaeis guineensis*, *Hordeum vulgare*, and *M. acuminata*), one living representative of a lineage that represents the extant earliest diverging lineage of flowering plants also named basal angiosperm (*A. trichopoda*), one lycophyte (*S. moellendorffii*) and one moss (*P. patens*). Genome data of *A. tauschii*, *E. guineensis*, and *M. acuminata* were downloaded from their project websites (Supplemental Table 1), and the other genome data were mainly downloaded from Phytozome (version 11) (Goodstein et al., 2012).

Gene Family Classification and Phylogenetic Analysis

We classified protein-coding genes into putative gene families or subfamilies using the OrthoMCL method (version 2.0.9) (Li et al., 2003) with an inflation parameter of 1.5, and obtained 66 509 orthogroups in total. The orthogroups with less than four genes and/or without at least one gene from outgroups were filtered out, and the remaining 12,077 orthogroups were processed to phylogenetic analysis. The taxonomic distribution of the 12,077 orthogroups by considering the last common ancestor of genes in each orthogroup was shown in Supplemental Figure 8. For the construction of global gene family trees, the amino acid sequences of each orthogroup were aligned with MAFFT (Katoh et al., 2005). Then the corresponding nucleic acid sequences were forced onto amino acid alignments using PAL2NAL (Suyama et al., 2006). To remove poorly aligned regions, the nucleic acid alignments were refined using trimAl 1.4 (Capella-Gutirrez et al., 2009) with the option "automated1." Phylogenetic trees were conducted using maximum likelihood method in RAxML 8.2.11 (Stamatakis, 2014) with the fast bootstrap option, 100 replicates under GTRGAMMA model.

Identifying Gene Duplication Events

To accurately identify gene duplications, we followed the same standard for gene trees and species tree reconciliation as proposed by Jiao et al. (2011). That is, two child branches need to have genes from at least one common species, and the parental node and one of the child nodes should both have bootstrap values equal or greater than 50%. Since the phylogenetic relationships of 25 species sampled in this study were clear (Angiosperm Phylogeny Website), we directly adopted the currently accepted topology as species tree. Genes from outgroups (*P. patens*, *S. moellendorffii*, and *A. trichopoda*) were used to root trees.

Firstly, we used Notung 2.9 (Stolzer et al., 2012), a gene tree-species tree reconciliation program, to batch reconciling all the nodes of gene trees with corresponding nodes in species tree. Parsimony-based optimization criterion was employed in Notung to minimize the duplication/loss cost. We ran the analysis based on the duplication-loss events model. In addition to "—reconcile" mode, the "—rearrange" mode was also performed with parameters setting as "—threshold 50%." This option could rearrange weakly supported edges (such as bootstrap <50%) and reduce the duplication uncertainty of inference (Notung 2.9 manual). Secondly, after carefully checking thousands of reconciled trees resulting from Notung, we further applied the standard that "two child branches need have genes from at least one common species," and removed some low-confidence duplications from all the reconciled trees.

Eliminating Tandem Duplications

We defined two genes located within five genes as tandem duplicates. If a duplication node contains two genes (gene1, gene2), or two child branches ((gene1, gene2), (gene3, gene4)), either of the two genes located proximal to each other were treated as from tandem duplications. Based on the above criteria, we removed tandem duplication events for all the reconciled gene trees.

K_S Calculation of Duplication Node and Circumscription of Individual WGDs

We used K_S values to date two WGDs occurring on the same branch that could not be circumscribed using the phylogenetic method. K_S estimates for pairwise comparisons (one gene in the m branch and the other gene in the n branch) were obtained using the Nei-Gojobori method (Nei and Gojobori, 1986) implemented in the yn00 program of the PAML package (Yang, 1997). The sum of K_S values for all pairwise comparisons were then divided by the number of K_S estimates (m*n). Thus, we got a weighted K_S value for a duplication event. According to the K_S ranges of ρ and σ events indicated by a sequential K_S curve of syntenic gene pairs of *O. sativa*, as obtained using Plant Genome Duplication Database online tools, we roughly defined that duplication nodes with $K_S \geq 1.0$ belong to the σ event, and that duplication nodes with $K_S < 1.0$ belong to the ρ event (Supplemental Figure 9). $K_S \geq 0.7$ belonging to the γ^M event and $K_S < 0.7$ belonging to the α/β^M event were defined based on the circumscription of a previous study (D'Hont et al., 2012).

Syntenic Conservation Analysis of Retained Paralogous Genes

The above procedures allow us to obtain the paralogous genes potential related to each WGD. To further validate if the duplicate genes are still located on syntenic blocks, we performed collinearity analysis of WGD-derived species. The intragenome syntenic blocks were detected by using MCScanX based on the default parameters (Wang et al., 2012). We then scored the percentage of paralogous genes with syntenic evidence out of the total genes through phylogenomic timing.

GO Annotation of Gene Orthogroups and Functional Enrichment Analysis

To annotate orthogroups, we used the full GO term of *A. thaliana* genes as the annotation of orthogroups if they have *Arabidopsis* genes. Otherwise, we searched the InterPro domain of protein sequences using InterProScan (Zdobnov and Apweiler, 2001) and got the full GO term annotation. Statistical enrichment of GO terms was evaluated by comparing the sample (common retained orthogroups) with the background (all annotated orthogroups) based on Fisher's exact test and adjusted *P* values according to the Benjamini and Hochbery (false discovery rate) method (Ashburner et al., 2000).

Retention Analysis of Transcription Factors

We used the gene families of transcription factors (TFs) of *A. thaliana* downloaded from PlantTFDB 4.0 (Jin et al., 2017) to annotate orthogroups. For each WGD event, the retained orthogroups of each TF family were identified. Owing to the sequence divergence after duplications, some TF families were usually classified into multiple orthogroups. To eliminate the influence of the size of orthogroups in one TF gene family, we calculated a normalized value (retention value, *R* value) to reflect the retention pattern of each TF after corresponding WGD. *R* value is calculated from the formula as follows:

$$R_{value} = \frac{\text{Number of orthogroups with retention in specific TF}}{\text{Total number of orthogroups in specific TF}} \Big/ \frac{\text{Number of all TF orthogroups with retention}}{\text{Total number of TF orthogroups}},$$

where $\frac{\text{Number of orthogroups with retention in specific TF}}{\text{Total number of orthogroups in specific TF}}$ is the probability of retention of a specific TF family after corresponding WGD and $\frac{\text{Number of all TF orthogroups with retention}}{\text{Total number of TF orthogroups}}$ is the probability of retention of all TF families after the corresponding WGD.

Coexpression Network Construction and Comparison

A total of 222 cold-related RNA-seq samples (162 for *A. thaliana* in Supplemental Table 7 and 60 for *V. vinifera* in Supplemental Table 8) were downloaded from the NCBI's Sequence Read Archive database (https://www.ncbi.nlm.nih.gov/sra/). These datasets were then investigated with quality control, clean reads mapping, FPKM (Fragments Per Kilobase of transcript per Million mapped reads) calculation, as in You et al. (2016). After removing non-expression genes of each sample (FPKM 0.14 and 0.36 were selected as cutoffs for *A. thaliana* and *V. vinifera*, respectively; detailed method as described in You et al., 2016), the remaining expressed genes were used to calculate coexpression relationships using Pearson's correlation coefficient (PCC). Subsequently, Mutual rank (MR) (calculated as the geometric mean of the PCC rank from gene A to gene B and the rank of gene B to gene A), was used to construct the coexpression network (Obayashi and Kinoshita, 2010; You et al., 2016). MR was demonstrated to be more effective to get credible coexpression gene pairs than PCC (Obayashi and Kinoshita, 2010). Therefore, we constructed MR-based coexpression networks for each species. Then, we selected the top 300 coexpressed genes (a threshold used by You et al. [2016] and Obayashi et al. [2018]) in *Arabidopsis* and in *Vitis* for network comparison. In such cases, the coexpression networks of *AthICE1* and *AthICE2* (formed by β WGD) in *Arabidopsis* and their orthologous *VviICE* in *Vitis* were used to assess the evolutionary pattern of these key genes in the CBF signaling pathway.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Molecular Plant Online*.

FUNDING

This research was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31000000). We also thank the start-up funding from State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, the Chinese Academy of Sciences.

AUTHOR CONTRIBUTIONS

Y.J. and S.W. designed the study. S.W., B.H., and Y.J. performed the data analyses and wrote the paper.

ACKNOWLEDGMENTS

No conflict of interest declared.

Received: June 5, 2019
Revised: October 16, 2019
Accepted: October 23, 2019
Published: November 1, 2019

REFERENCES

Adams, K.L., and Wendel, J.F. (2005). Polyploidy and genome evolution in plants. Curr. Opin. Plant Biol. **8**:135–141.

Alvarez, L.W., Alvarez, W., Asaro, F., and Michel, H.V. (1980). Extraterrestrial cause for cretaceous-tertiary extinct. Science **208**:1095–1108.

Arrigo, N., and Barker, M.S. (2012). Rarely successful polyploids and their legacy in plant genomes. Curr. Opin. Plant Biol. **15**:140–146.

Ashburner, M., Ball, C., Blake, J.A., Botstein, D., Butler, H., Cherry, J., Davis, A., Dolinski, K., Dwight, S., Eppig, J., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. **25**:25–29.

Blanc, G., and Wolfe, K.H. (2004). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell **16**:1667–1678.

Bowers, J.E., Chapman, B.A., Rong, J.K., and Paterson, A.H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature **422**:433–438.

Brochmann, C., Brysting, A.K., Alsos, I.G., Borgen, L., Grundt, H.H., Scheen, A.C., and Elven, R. (2004). Polyploidy in arctic plants. Biol. J. Linn. Soc. Lond. **82**:521–536.

Cannon, S.B., McKain, M.R., Harkess, A., Nelson, M.N., Dash, S., Deyholos, M.K., Peng, Y., Joyce, B., Stewart, C.N., et al. (2015). Multiple polyploidy events in the early radiation of nodulating and nonnodulating legumes. Mol. Biol. Evol. **32**:193–210.

Capella-Gutirrez, S., Silla-Martínez, J.M., and Gabaldn, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics **25**:1972–1973.

Carabelli, M., Turchi, L., Ruzza, V., Morelli, G., and Ruberti, I. (2013). Homeodomain-leucine zipper II family of transcription factors to the limelight. Plant Signal. Behav. **8**:e25447.

Chao, D.Y., Dilkes, B., Luo, H., Douglas, A., Yakubova, E., Lahner, B., and Salt, D.E. (2013). Polyploids exhibit higher potassium uptake and salinity tolerance in *Arabidopsis*. Science **341**:658–659.

Cheng, S., van den Bergh, E., Zeng, P., Zhong, X., Xu, J., Liu, X., Hofberger, J., de Bruijn, S., Bhide, A.S., Kuelahoglu, C., et al. (2013). The *Tarenaya hassleriana* genome provides insight into reproductive trait and genome evolution of crucifers. Plant Cell **25**:2813–2830.

Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B.H., Hong, X., Agarwal, M., and Zhu, J.K. (2003). *ICE1*: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. Genes Dev. **17**:1043–1054.

Chinnusamy, V., Zhu, J., and Zhu, J.K. (2007). Cold stress regulation of gene expression in plants. Trends Plant Sci. **12**:444–451.

Comai, L. (2005). The advantages and disadvantages of being polyploid. Nat. Rev. Genet. **6**:836–846.

Conant, G.C. (2010). Rapid reorganization of the transcriptional regulatory network after genome duplication in yeast. Proc. Biol.Sci. **277**:869–876.

Cui, L., Wall, P.K., Leebens-Mack, J.H., Lindsay, B.G., Soltis, D.E., Doyle, J.J., Soltis, P.S., Carlson, J.E., Arumuganathan, K., Barakat, A., et al. (2006). Widespread genome duplications throughout the history of flowering plants. Genome Res. **16**:738–749.

D'Hont, A., Denoeud, F., Aury, J.M., Baurens, F.C., Carreel, F., Garsmeur, O., Noel, B., Bocs, S., Droc, G., Rouard, M., et al. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. Nature **488**:213–217.

de Mendoza, A., Sebé-Pedrós, A., Šestak, M.S., Matejčić, M., Torruella, G., Domazet-Lošo, T., and Ruiz-Trillo, I. (2013). Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. Proc. Natl. Acad. Sci. U S A **110**:E4858–E4866.

De Smet, R., and Van de Peer, Y. (2012). Redundancy and rewiring of genetic networks following genome-wide duplication events. Curr. Opin. Plant Biol. **15**:168–176.

Dietrich, R.A., Richberg, M.H., Schmidt, R., Dean, C., and Dangl, J.L. (1997). A novel zinc finger protein is encoded by the *Arabidopsis LSD1* gene and functions as a negative regulator of plant cell death. Cell **88**:685–694.

Doyle, J.J., Flagel, L.E., Paterson, A.H., Rapp, R.A., Soltis, D.E., Soltis, P.S., and Wendel, J.F. (2008). Evolutionary genetics of genome merger and doubling in plants. Annu. Rev. Genet. **42**:443–461.

Edger, P.P., Heidel-Fischer, H.M., Bekaert, M., Rota, J., Glöckner, G., Platts, A.E., Heckel, D.G., Der, J.P., and Wafula, E.K. (2015). The butterfly plant arms-race escalated by gene and genome duplications. Proc. Natl. Acad. Sci. U S A **112**:8362–8366.

Estep, M.C., McKain, M.R., Vela Diaz, D., Zhong, J., Hodge, J.G., Hodkinson, T.R., Layton, D.J., Malcomber, S.T., Pasquet, R., and Kellogg, E.A. (2014). Allopolyploidy, diversification, and the Miocene grassland expansion. Proc. Natl. Acad. Sci. U S A **111**:15149–15154.

Fawcett, J.A., Maere, S., and Van de Peer, Y. (2009). Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. Proc. Natl. Acad. Sci. U S A **106**:5737–5742.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. Genetics **151**:1531–1545.

Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. Annu. Rev. Plant Biol. **60**:433–453.

Friedman, W.E. (2009). The meaning of Darwin's 'abominable mystery'. Am. J. Bot. **96**:5–21.

Fursova, O.V., Pogorelko, G.V., and Tarasov, V.A. (2009). Identification of *ICE2*, a gene involved in cold acclimation which determines freezing tolerance in *Arabidopsis thaliana*. Gene **429**:98–103.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., et al. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. **40**:D1178–D1186.

Hegarty, M.J., and Hiscock, S.J. (2008). Genomic clues to the evolutionary success of polyploid plants. Curr. Biol. **18**:R435–R444.

Heimhofer, U., Hochuli, P.A., Burla, S., Dinis, J.M.L., and Weissert, H. (2005). Timing of Early Cretaceous angiosperm diversification and possible links to major paleoenvironmental change. Geology **33**:141.

Huang, C.H., Zhang, C., Liu, M., Hu, Y., Gao, T., Qi, J., and Ma, H. (2016). Multiple polyploidization events across Asteraceae with two nested events in the early history revealed by nuclear phylogenomics. Mol. Biol. Evol. **33**:2820–2835.

Jaillon, O., Aury, J.M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature **449**:463–467.

Ji, H., Pardo, J.M., Batelli, G., Van Oosten, M.J., Bressan, R.A., and Li, X. (2013). The Salt Overly Sensitive (SOS) pathway: established and emerging roles. Mol. Plant **6**:275–286.

Jiao, Y., Lau, O.S., and Deng, X.W. (2007). Light-regulated transcriptional networks in higher plants. Nat. Rev. Genet. **8**:217–230.

Jiao, Y., Leebens-Mack, J., Ayyampalayam, S., Bowers, J.E., McKain, M.R., McNeal, J., Rolf, M., Ruzicka, D.R., Wafula, E., Wickett, N.J., et al. (2012). A genome triplication associated with early diversification of the core eudicots. Genome Biol. **13**:R3.

Jiao, Y., Li, J., Tang, H., and Paterson, A.H. (2014). Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. Plant Cell **26**:2792–2802.

Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E., Tomsho, L.P., Hu, Y., Liang, H., Soltis, P.S., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. Nature **473**:97–100.

Jin, J., Tian, F., Yang, D.C., Meng, Y.Q., Kong, L., Luo, J., and Gao, G. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. **45**:D1040–D1045.

Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. **33**:511–518.

Khan, S.A., Li, M.Z., Wang, S.M., and Yin, H.J. (2018). Revisiting the role of plant transcription factors in the battle against abiotic stress. Int. J. Mol. Sci. **19**:1634.

Levin, D.A. (1983). Polyploidy and novelty in flowering plants. Am. Nat. **122**:1–25.

Li, L., Stoeckert, C.J., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. **13**:2178–2189.

Lu, F., Lipka, A.E., Glaubitz, J., Elshire, R., Cherney, J.H., Casler, M.D., Buckler, E.S., and Costich, D.E. (2013). Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. PLoS Genet. **9**:e1003215.

Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. Science **290**:1151–1155.

Maere, S., De Bodt, S., Raes, J., Casneuf, T., Van Montagu, M., Kuiper, M., and Van de Peer, Y. (2005). Modeling gene and genome duplications in eukaryotes. Proc. Natl. Acad. Sci. U S A **102**:5454–5459.

Mayrose, I., Zhan, S.H., Rothfels, C.J., Magnuson-Ford, K., Barker, M.S., Rieseberg, L.H., and Otto, S.P. (2011). Recently formed polyploid plants diversify at lower rates. Science **333**:1257.

McKain, M.R., Tang, H., McNeal, J.R., Ayyampalayam, S., Davis, J.I., depamphilis, C.W., Givnish, T.J., Pires, J.C., Stevenson, D.W., and Leebens-Mack, J.H. (2016). A phylogenomic assessment of ancient polyploidy and genome evolution across the Poales. Genome Biol. Evol. **8**:1150–1164.

Ming, R., VanBuren, R., Wai, C.M., Tang, H., Schatz, M.C., Bowers, J.E., Lyons, E., Wang, M.L., Chen, J., Biggers, E., et al. (2015). The pineapple genome and the evolution of CAM photosynthesis. Nat. Genet. **47**:1435–1442.

Nei, M., and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3**:418.

Nichols, D.J., and Johnson, K.R. (2008). Plants and the K-T Boundary (England: Cambridge University Press).

Novikova, P.Y., Hohmann, N., and Van de Peer, Y. (2018). Polyploid *Arabidopsis* species originated around recent glaciation maxima. Curr. Opin. Plant Biol. **42**:8–15.

Obayashi, T., Aoki, Y., Tadaka, S., Kagaya, Y., and Kinoshita, K. (2018). ATTED-II in 2018: a plant coexpression database based on investigation of the statistical property of the mutual rank index. Plant Cell Physiol. **59**:e3.

Obayashi, T., and Kinoshita, K. (2010). Coexpression landscape in ATTED-II: usage of gene list and gene network for various types of pathways. J. Plant Res. **123**:311–319.

Ohno, S. (1970). Evolution by Gene Duplication (Berlin: Springer-Verlag).

Otto, S.P., and Whitton, J. (2000). Polyploid incidence and evolution. Annu. Rev. Genet. **34**:401–437.

Parisod, C., and Broennimann, O. (2016). Towards unified hypotheses of the impact of polyploidy on ecological niches. New Phytol. **212**:540–542.

Paterson, A.H., Bowers, J.E., and Chapman, B.A. (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc. Natl. Acad. Sci. U S A **101**:9903–9908.

Ramsey, J. (2011). Polyploidy and ecological adaptation in wild yarrow. Proc. Natl. Acad. Sci. U S A **108**:7096–7101.

Ren, R., Wang, H., Guo, C., Zhang, N., Zeng, L., Chen, Y., Ma, H., and Qi, J. (2018). Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. Mol. Plant **11**:414–428.

Renny-Byfield, S., and Wendel, J.F. (2014). Doubling down on genomes: polyploidy and crop plants. Am. J. Bot. **101**:1711–1725.

Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P., Lindquist, E.A., Kamisugi, Y., et al. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. Science **319**:64–69.

Ruberti, I., Sessa, G., Ciolfi, A., Possenti, M., Carabelli, M., and Morelli, G. (2012). Plant adaptation to dynamically changing environment: the shade avoidance response. Biotechnol. Adv. **30**:1047–1058.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., et al. (2010). Genome sequence of the palaeopolyploid soybean. Nature **463**:178–183.

Schnable, J.C., Springer, N.M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proc. Natl. Acad. Sci. U S A **108**:4069–4074.

Schulte, P., Alegret, L., Arenillas, I., Arz, J.A., Barton, P.J., Bown, P.R., Bralower, T.J., Christeson, G.L., Claeys, P., Cockell, C.S., et al. (2010). The chicxulub asteroid impact and mass extinction at the Cretaceous-Paleogene boundary. Science **327**:1214–1218.

Shi, Y., Ding, Y., and Yang, S. (2015). Cold signal transduction and its interplay with phytohormones during cold acclimation. Plant Cell Physiol. **56**:7–15.

Shi, Y., Ding, Y., and Yang, S. (2018). Molecular regulation of CBF signaling in cold acclimation. Trends Plant Sci. **23**:623–637.

Singh, R., Ong-Abdullah, M., Low, E.L., Manaf, M.A.A., Rosli, R., Nookiah, R., Ooi, L.C., Ooi, S., Chan, K., Halim, M.A., et al. (2013). Oil palm genome sequence reveals divergence of interfertile species in Old and New Worlds. Nature **500**:335–339.

Soltis, D.E., Albert, V.A., Leebens-Mack, J., Bell, C.D., Paterson, A.H., Zheng, C., Sankoff, D., Depamphilis, C.W., Wall, P.K., and Soltis, P.S. (2009). Polyploidy and angiosperm diversification. Am. J. Bot. **96**:336–348.

Soltis, P.S., Marchant, D.B., Van de Peer, Y., and Soltis, D.E. (2015). Polyploidy and genome evolution in plants. Curr. Opin. Genet. Dev. **35**:119–125.

Soltis, P.S., and Soltis, D.E. (2016). Ancient WGD events as drivers of key innovations in angiosperms. Curr. Opin. Plant Biol. **30**:159–165.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics **30**:1312–1313.

Stebbins, G.L. (1950). Variation and Evolution in Plants (New York: Columbia University Press).

Stolzer, M., Lai, H., Xu, M., Sathaye, D., Vernot, B., and Durand, D. (2012). Inferring duplications, losses, transfers and incomplete lineage sorting with nonbinary species trees. Bioinformatics **28**:i409–i415.

Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. **34**:W609.

Tang, H., Bowers, J.E., Wang, X., and Paterson, A.H. (2010). Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. Proc. Natl. Acad. Sci. U S A **107**:472–477.

te Beest, M., Le Roux, J.J., Richardson, D.M., Brysting, A.K., Suda, J., Kubesova, M., and Pysek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. Ann. Bot. **109**:19–45.

The Tomato Genome Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature **485**:635–641.

Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science **313**:1596–1604.

Van de Peer, Y., Maere, S., and Meyer, A. (2009). The evolutionary significance of ancient genome duplications. Nat. Rev. Genet. **10**:725–732.

Van de Peer, Y., Mizrachi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. Nat. Rev. Genet. **18**:411–424.

Vanneste, K., Baele, G., Maere, S., and Van de Peer, Y. (2014). Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous-Paleogene boundary. Genome Res. **24**:1334–1347.

Vekemans, D., Proost, S., Vanneste, K., Coenen, H., Viaene, T., Ruelens, P., Maere, S., Van de Peer, Y., and Geuten, K. (2012). Gamma paleohexaploidy in the stem lineage of core eudicots: significance for *MADS-box* gene and species diversification. Mol. Biol. Evol. **29**:3793–3806.

Visger, C.J., Germain-Aubrey, C.C., Patel, M., Sessa, E.B., Soltis, P.S., and Soltis, D.E. (2016). Niche divergence between diploid and autotetraploid *Tolmiea*. Am. J. Bot. **103**:1396–1406.

Wang, J., Sun, P., Li, Y., Liu, Y., Yang, N., Yu, J., Ma, X., Sun, S., Xia, R., Liu, X., et al. (2018). An overlooked paleotetraploidization in Cucurbitaceae. Mol. Biol. Evol. **35**:16–26.

Wang, Y., Tang, H., Debarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.H., Jin, H., Marler, B., Guo, H., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. **40**:e49.

Wendel, J.F., Jackson, S.A., Meyers, B.C., and Wing, R.A. (2016). Evolution of plant genome architecture. Genome Biol. **17**:37.

William, D.A., Su, Y., Smith, M.R., Lu, M., Baldwin, D.A., and Wagner, D. (2004). Genomic identification of direct target genes of *LEAFY*. Proc. Natl. Acad. Sci. U S A **101**:1775–1780.

Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B., and Rieseberg, L.H. (2009). The frequency of polyploid speciation in vascular plants. Proc. Natl. Acad. Sci. U S A **106**:13875–13879.

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13**:555–556.

You, Q., Zhang, L., Yi, X., Zhang, K., Yao, D., Zhang, X., Wang, Q., Zhao, X., Ling, Y., Xu, W., et al. (2016). Co-expression network analyses identify functional modules associated with

development and stress response in *Gossypium arboreum*. Sci. Rep. **6**:38436.

Zachos, J.C., Dickens, G.R., and Zeebe, R.E. (2008). An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. Nature **451**:279–283.

Zdobnov, E.M., and Apweiler, R. (2001). InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17**:847–848.

Zhao, C., Lang, Z., and Zhu, J.K. (2015). Cold responsive gene transcription becomes more complex. Trends Plant Sci. **20**:466–468.

Zhao, T., Holmer, R., de Bruijn, S., Angenent, G.C., van den Burg, H.A., and Schranz, M.E. (2017). Phylogenomic synteny network analysis of MADS-Box transcription factor genes reveals lineage-specific transpositions, ancient tandem duplications, and deep positional conservation. Plant Cell **29**:1278–1292.