




Functional Divergence in Orthologous Transcription Factors: Insights from AtCBF2/3/1 and OsDREB1C

Deyin Deng ^{1,†}, Yixin Guo ^{1,†}, Liangyu Guo ¹, Chengyang Li ¹, Yuqi Nie ¹, Shuo Wang ¹, and Wenwu Wu ^{1,2,*}

¹State Key Laboratory of Subtropical Silviculture, College of Forestry and Biotechnology, Zhejiang A&F University, Hangzhou 311300, China

²Zhejiang International Science and Technology Cooperation Base for Plant Germplasm Resources Conservation and Utilization, Zhejiang A&F University, Hangzhou 311300, China

[†]These authors share the first authorship.

*Corresponding author: E-mail: wwwu@zafu.edu.cn.

Associate editor: Irina Arkhipova

Abstract

Despite traditional beliefs of orthologous genes maintaining similar functions across species, growing evidence points to their potential for functional divergence. C-repeat binding factors/dehydration-responsive element binding protein 1s (CBFs/DREB1s) are critical in cold acclimation, with their overexpression enhancing stress tolerance but often constraining plant growth. In contrast, a recent study unveiled a distinctive role of rice OsDREB1C in elevating nitrogen use efficiency (NUE), photosynthesis, and grain yield, implying functional divergence within the CBF/DREB1 orthologs across species. Here, we delve into divergent molecular mechanisms of OsDREB1C and AtCBF2/3/1 by exploring their evolutionary trajectories across rice and *Arabidopsis* genomes, regulomes, and transcriptomes. Evolutionary scrutiny shows discrete clades for OsDREB1C and AtCBF2/3/1, with the Poaceae-specific DREB1C clade mediated by a transposon event. Genome-wide binding profiles highlight OsDREB1C's preference for GCCGAC compared to AtCBF2/3/1's preference for A/GCCGAC, a distinction determined by R12 in the OsDREB1C AP2/ERF domain. Cross-species multiomic analyses reveal shared gene orthogroups (OGs) and underscore numerous specific OGs uniquely bound and regulated by OsDREB1C, implicated in NUE, photosynthesis, and early flowering, or by AtCBF2/3/1, engaged in hormone and stress responses. This divergence arises from gene gains/losses (~16.7% to 25.6%) and expression reprogramming (~62.3% to 66.2%) of OsDREB1C- and AtCBF2/3/1-regulated OGs during the extensive evolution following the rice–*Arabidopsis* split. Our findings illustrate the regulatory evolution of OsDREB1C and AtCBF2/3/1 at a genomic scale, providing insights on the functional divergence of orthologous transcription factors following gene duplications across species.

Key words: C-repeat binding factor (CBF), OsDREB1C, growth–defense tradeoff, evolutionary divergence, orthologous gene, abiotic stress.

Introduction

Determining the functional conservation and evolution of orthologous genes is a cornerstone in molecular and evolutionary biology. Traditional views have long held that orthologs, arising from speciation events, maintain similar functions across species, underscoring evolutionary parallels (Koonin 2005). However, advancements in genomic sequencing and functional analysis reveal that orthologous genes can diverge functionally, acquiring species-specific roles (Gharib and Robinson-Rechavi 2011; Laurent et al. 2020). The divergence often reflects adaptation to varying environmental pressures, facilitated by subfunctionalization and neofunctionalization, processes that often occur after gene duplications (Force et al. 1999). Furthermore,

evolutionary dynamics of gene expression complicate this landscape, extending the roles of orthologous genes beyond functional conservation to encompass spatio-temporal variations (Liao et al. 2006).

Among varying environmental pressures, low temperature stands out as a critical stressor that impedes plant growth and development (Shi et al. 2018; Guo et al. 2024). In response, plants have evolved sophisticated adaptive mechanisms, such as cold acclimation, a strategy often observed in temperate plants that enhances freezing tolerance through prior exposure to low, nonfreezing temperatures (Jaglo-Ottosen et al. 1998; Jia et al. 2016; Zhao et al. 2016; Shi et al. 2018). At the molecular level, cold acclimation involves the coregulation of numerous cold-responsive (COR) genes, centrally mediated by C-repeat binding factors

Received: January 13, 2024. **Revised:** April 19, 2024. **Accepted:** May 06, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Open Access

(CBFs), also known as dehydration-responsive element binding protein 1s (DREB1s; Ding et al. 2020; Kidokoro et al. 2022).

CBF/DREB1 orthologs, identified in many angiosperms, are widely recognized as pivotal transcription factors in cold tolerance (Benedict et al. 2006; Welling and Palva 2008; Menon et al. 2015; Jiang et al. 2022). Our study traced the evolutionary trajectory of these genes, revealing adaptive expansions during periods of global cooling, occurring independently in eudicots and monocots (Nie et al. 2022). In the representative eudicot *Arabidopsis thaliana* genome, there are six *CBF/DREB1* members, featuring tandemly arrayed *AtCBF2*, *AtCBF3*, and *AtCBF1* (briefly *AtCBF2/3/1*) and dispersedly distributed *AtCBF4*, *AtCBF5/DDF1*, and *AtCBF6/DDF2* (Nakano et al. 2006; Nie et al. 2022). In contrast, the representative monocot *Oryza sativa* (rice) genome harbors ten *CBF/DREB1* members, including five organized on two sets of tandem arrays (i.e. *OsDREB1J* and *OsDREB1I*; *OsDREB1B*, *OsDREB1H*, and *OsDREB1A*) and five dispersedly distributed *OsDREB1C*, *OsDREB1D*, *OsDREB1E*, *OsDREB1F*, and *OsDREB1G* (Mao and Chen 2012; Nie et al. 2022). Under chilling temperatures, the tandemly arrayed genes *AtCBF2/3/1* in *Arabidopsis* and *OsDREB1B/H/A* in rice are highly induced, leading to the activation of numerous *COR* genes to enhance freezing tolerance (Liu et al. 1998; Dubouzet et al. 2003; Mao and Chen 2012). Additionally, some *CBF/DREB1* members, including *AtCBF4*, *OsDREB1C*, and *OsDREB1G*, respond moderately to cold (Haake et al. 2002; Mao and Chen 2012; Nie et al. 2022), and others to different abiotic stresses, exemplified by the induction of *AtCBF4*, *AtCBF5*, and *AtCBF6* in *Arabidopsis* and *OsDREB1C* and *OsDREB1F* in rice under osmotic stress conditions, such as drought and salt (Magome et al. 2004; Mao and Chen 2012). This multifaceted response indicates that while certain *CBF/DREB1* orthologs are primarily conserved in cold responses in *Arabidopsis* and rice, other members have evolved divergent regulatory roles in response to other stresses.

Genetically, constitutive overexpression of *CBF/DREB1* genes has demonstrated efficacy in enhancing cold, drought, and salinity tolerance across various angiosperms (Gilmour et al. 2004; Kasuga et al. 2004; Ito et al. 2006; Reis et al. 2014; Jin et al. 2017; Li et al. 2022). However, the enhanced stress tolerance often comes at the cost of growth constraints, leading to variable degrees of growth retardation and delayed flowering (Dubouzet et al. 2003; Gilmour et al. 2004; An et al. 2016; Liao et al. 2017; Baker et al. 2022). These studies collectively indicate a conserved role of *CBF/DREB1* genes in a functional tradeoff: striking a balance between the activation of stress tolerance and the suppression of plant growth. Nevertheless, a few isolated studies have reported that the overexpression of specific *CBF/DREB1* genes, such as *OsDREB1F* and *ZjDREB1.4*, may yield moderate or negligible effects on plant growth (Wang et al. 2008; Feng et al. 2019). Moreover, a recent groundbreaking study revealed a remarkable outcome: overexpression of *OsDREB1C* in rice confers substantial

benefits for plant growth by boosting nitrogen use efficiency (NUE) usage, promoting early flowering, and increasing grain yields (Wei et al. 2022). Therefore, in contrast to the majority of *CBF/DREB1* genes, notably *Arabidopsis AtCBF2/3/1*, specific orthologous genes, like rice *OsDREB1C*, seem to have evolved distinct functional roles in promoting plant growth and development. However, the molecular mechanisms governing the functional divergence that leads to disparate overexpression phenotypes between the orthologous transcription factors *AtCBF2/3/1* and *OsDREB1C* remain poorly understood.

In this study, we assembled representative species from 2 basal angiosperms, 4 eudicots, and 14 monocots. Given an intricate set of *CBF/DREB1* genes in monocots, we intentionally selected a diverse range of 14 monocot species, comprising 2 from the basal monocot Alismatales, 1 species each from Orchidaceae, Musaceae, Bromeliaceae, and Cyperaceae, and 8 from Poaceae (Fig. 1a, supplementary table S1, Supplementary Material online). Initially, we identified *CBF/DREB1* coorthologs in these species and conducted a thorough phylogenetic analysis, with a focus on the *DREB1C* clade in monocots and the *CBF* clade in eudicots. Subsequently, we compared gene binding profiles and motifs of *OsDREB1C* and *AtCBF2/3/1* across the rice and *Arabidopsis* genomes. Utilizing sequence alignment, protein-DNA docking, and experimental validation, we scrutinized the key residues responsible for divergent DNA-binding specificities between *OsDREB1C* and *AtCBF2/3/1*. Moreover, through the integration of cross-species multiomic data, we revealed both shared and numerous specific genes directly regulated by *OsDREB1C* and/or by *AtCBF2/3/1*. In conclusion, we highlighted two keys, gene gains/losses and expression reprogramming, pivotal to the evolutionarily functional divergence of the orthologous transcription factors *OsDREB1C* in rice and *AtCBF2/3/1* in *Arabidopsis*.

Results

Evolutionary Divergence of *OsDREB1C* and *AtCBF2/3/1* Clades

In the above 20 selected species, we identified 154 *CBF/DREB1* coorthologs characterized by a conserved AP2/ERF domain and 2 flanking *CBF/DREB1*-specific signatures, PKRxAGR and DSAWR (supplementary fig. S1, Supplementary Material online). Subsequently, we constructed the phylogeny of these *CBF/DREB1* proteins (Fig. 1b). The phylogenetic analysis distinctly revealed the proliferation of *CBF/DREB1* genes through multiple duplications after the divergence of eudicots and monocots, with a remarkable expansion in the Poaceae family. Consistent with previous studies (Nakano et al. 2006; Nie et al. 2022), the six *AtCBF* genes in *Arabidopsis* aligned within the eudicot *CBF* group, in which the tandemly arrayed *AtCBF2/3/1* clustered closely together, proximal to *AtCBF4* and distinct from *AtCBF5* and *AtCBF6* (Fig. 1c). In the monocot group, the ten rice *OsDREB1* genes formed six clades: *DREB1F*, *DREB1E/1G*,

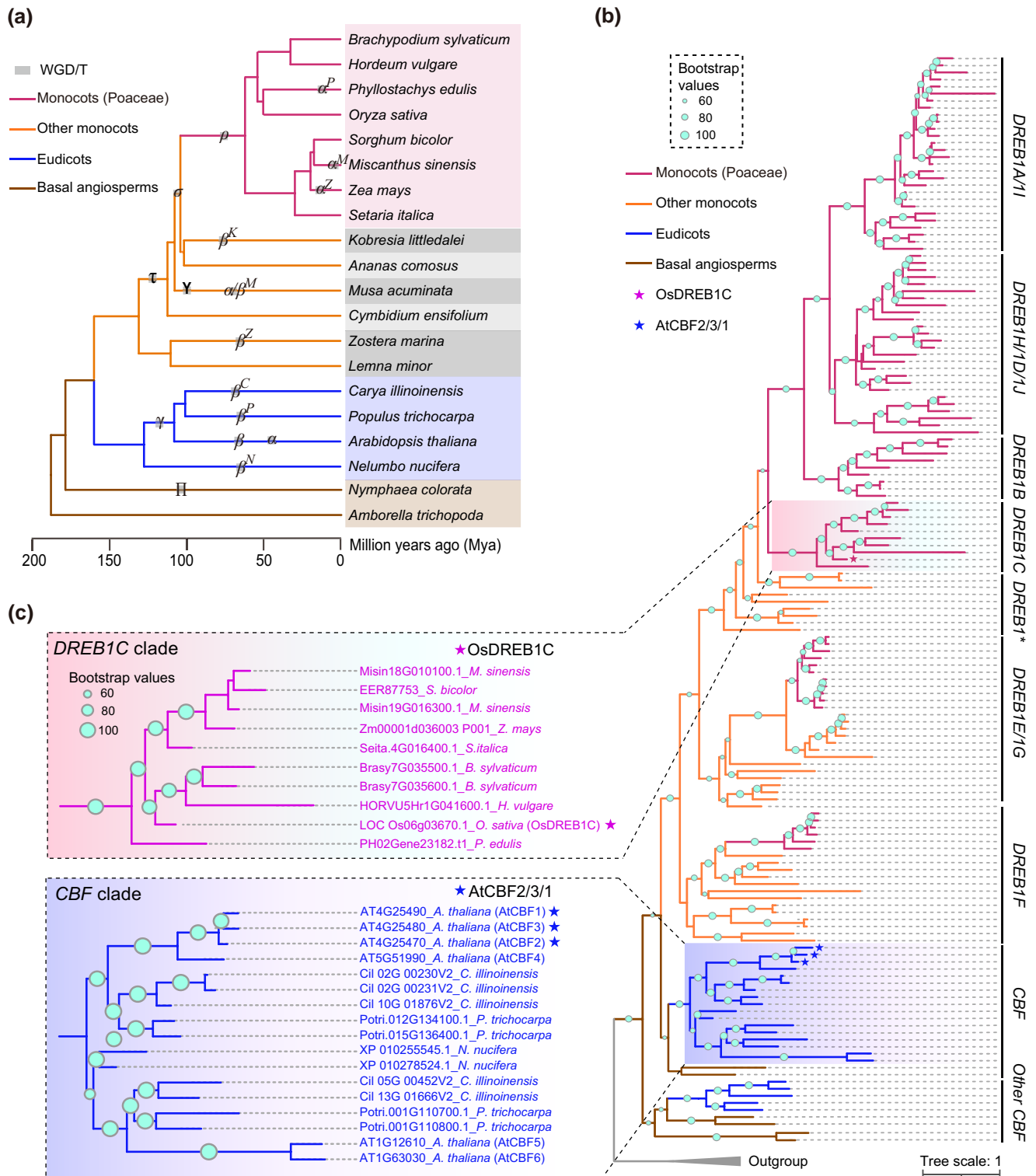


Fig. 1. Evolved clades of CBF/DREB1 genes in eudicots and monocots. a) Phylogenetic tree of 20 selected plants in this study. The evolutionary relationships of the plants, including 2 basal angiosperms, 4 eudicots, and 14 monocots, were obtained from the TimeTree website (Kumar et al. 2022). The tree incorporates well-acknowledged WGD/T events (supplementary table S2, Supplementary Material online) shown on the branches. b and c) Diverse clades of CBF/DREB1 genes b) and an enlarged view of two clades containing monocot DREB1C and eudicot CBF genes c). The CBF/DREB1 phylogeny was constructed using IQ-TREE v2.1.3 (Minh et al. 2020) through the ML method, with close relative homologs from the DREB III subfamily selected as outgroups for rooting (see Materials and Methods). The clade marked as DREB1* signifies orthologs of the Poaceae DREB1A//H/D//J//B/C in basal monocots. The sizes of circles on branches indicate support values derived from 1,000 ultrafast bootstrap samplings. Stars denote OsDREB1C in rice and the tandemly arrayed AtCBF2, AtCBF3, and AtCBF1 (abbreviated as AtCBF2/3/1) in *Arabidopsis*. WGD/T, whole-genome duplication/triplication.

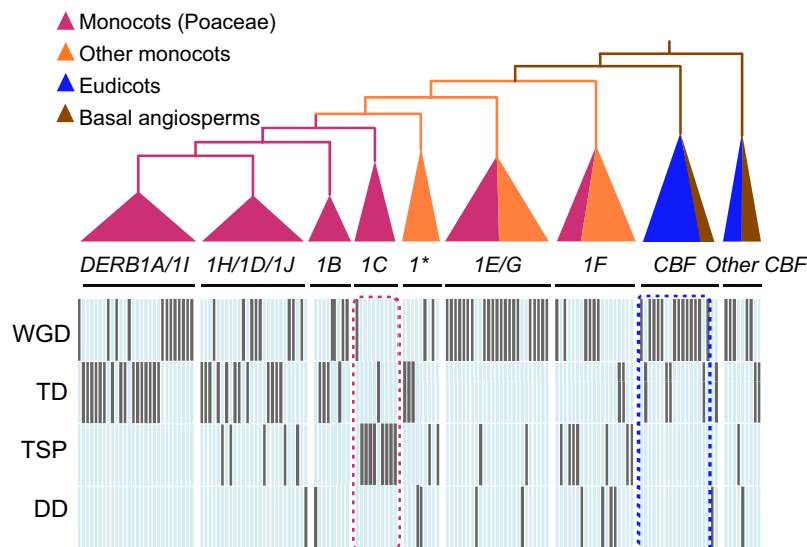


Fig. 2. A contrasting duplication mechanism in the *DREB1C* clade compared to other *CBF/DREB1* clades. Duplication mechanisms of *CBF/DREB1* genes. WGD, whole-genome duplication; TD, tandem duplication; TSP, transposed duplication; DD, dispersed duplication. The nine condensed clades were extracted from Fig. 1b, with dashed boxes indicating *DREB1C* genes in monocots and *CBF* genes in eudicots.

and the Poaceae-specific *DREB1C*, *DREB1B*, *DREB1H/1D/1J*, and *DREB1A/1I* (Fig. 1b). The *DREB1F* clade emerged as the monocot basal lineage proximal to the eudicot *CBF* clade, and the *DREB1C* clade positioned as the basal sister lineage to *DREB1B*, *DREB1H/1D/1J*, and *DREB1A/1I* clades in the Poaceae. Across these *CBF/DREB1* clades, there are clear variations marked by distinct conserved amino acids in the DNA-binding AP2 domain, confirming their evolutionary divergences (supplementary fig. S1, Supplementary Material online). These observations underscore the evolutionary divergence of *AtCBF2/3/1* and *OsDREB1C* into distinct clades, with *OsDREB1C* belonging to the basal clade of the Poaceae-specific expanded *DREB1C/1B/1H/1D/1J/1A/1I* lineages.

Transposition of *DREB1C* before the Poaceae Divergence

Next, we delved into duplication mechanism that gave rise to *CBF/DREB1* genes in the aforementioned species. In alignment with previous studies (Guo et al. 2022; Nie et al. 2022), *CBF/DREB1* genes predominantly emerged through whole-genome duplication (WGD) and tandem duplication (TD) in both monocots and eudicots (Fig. 2). Regarding *AtCBF2/3/1*, it has been demonstrated that *AtCBF2* and *AtCBF4* originated from the β -WGD preceding the split of Cleomaceae and Brassicaceae, while the tandemly arrayed *AtCBF2/3/1* resulted from TD events approximately 29.2 million years ago (Mya) in ancient Brassicaceae (Nie et al. 2022). In contrast, the genesis of *DREB1C* in Poaceae seemed exclusively linked to transposed duplication (TSP; Fig. 2). This observation led to the inference that a transposition event preceded the diversification of the Poaceae, leading to the emergence of the extant *DREB1C* genes. Due to the extensively long evolutionary history (~101 million years) since the Poaceae

emerged (Huang et al. 2022), it is difficult to determine the specific type of transposable elements involved, including those characterized by terminal inverted repeats and target site duplications.

Genome-Wide Landscapes of *OsDREB1C*- and *AtCBF2/3/1*-Bound Genes

The functional divergence of *OsDREB1C* and *AtCBF2/3/1* is most prominently reflected by their respective binding genes across genomes. To delineate their binding landscapes, we conducted an in-depth analysis on experimental data sets derived from chromatin immunoprecipitation sequencing (ChIP-seq) and DNA affinity purification sequencing (DAP-seq) for *AtCBF2/3/1* in the *Arabidopsis* genome (O'Malley et al. 2016; Song et al. 2021) and for *OsDREB1C* in the rice genome (Wei et al. 2022; see Materials and Methods). For each of *AtCBF1*, *AtCBF2*, and *AtCBF3*, comparative analysis of ChIP-seq and DAP-seq data revealed both significant overlaps and notable distinctions in their respective binding genes (supplementary fig. S2A, Supplementary Material online). To fortify the reliability of our analysis, we exclusively retained genes identified by both ChIP-seq and DAP-seq experiments, thereby establishing a high-confidence set of *CBF/DREB1C*-binding genes.

Upon conducting a comparative assessment of three distinct sets of *AtCBF1*-, *AtCBF2*-, and *AtCBF3*-binding genes, a substantial degree of overlap was evident, as depicted in supplementary fig. S2B, Supplementary Material online. Notably, approximately 95% of the genes bound by *AtCBF3* were also bound by *AtCBF1* and *AtCBF2*. This observation agrees well with the documented redundant functions of *AtCBF1*, *AtCBF2*, and *AtCBF3* (Gilmour et al. 2004; Song et al. 2021). Consequently, we consolidated the tandemly arrayed *AtCBF2*, *AtCBF3*, and

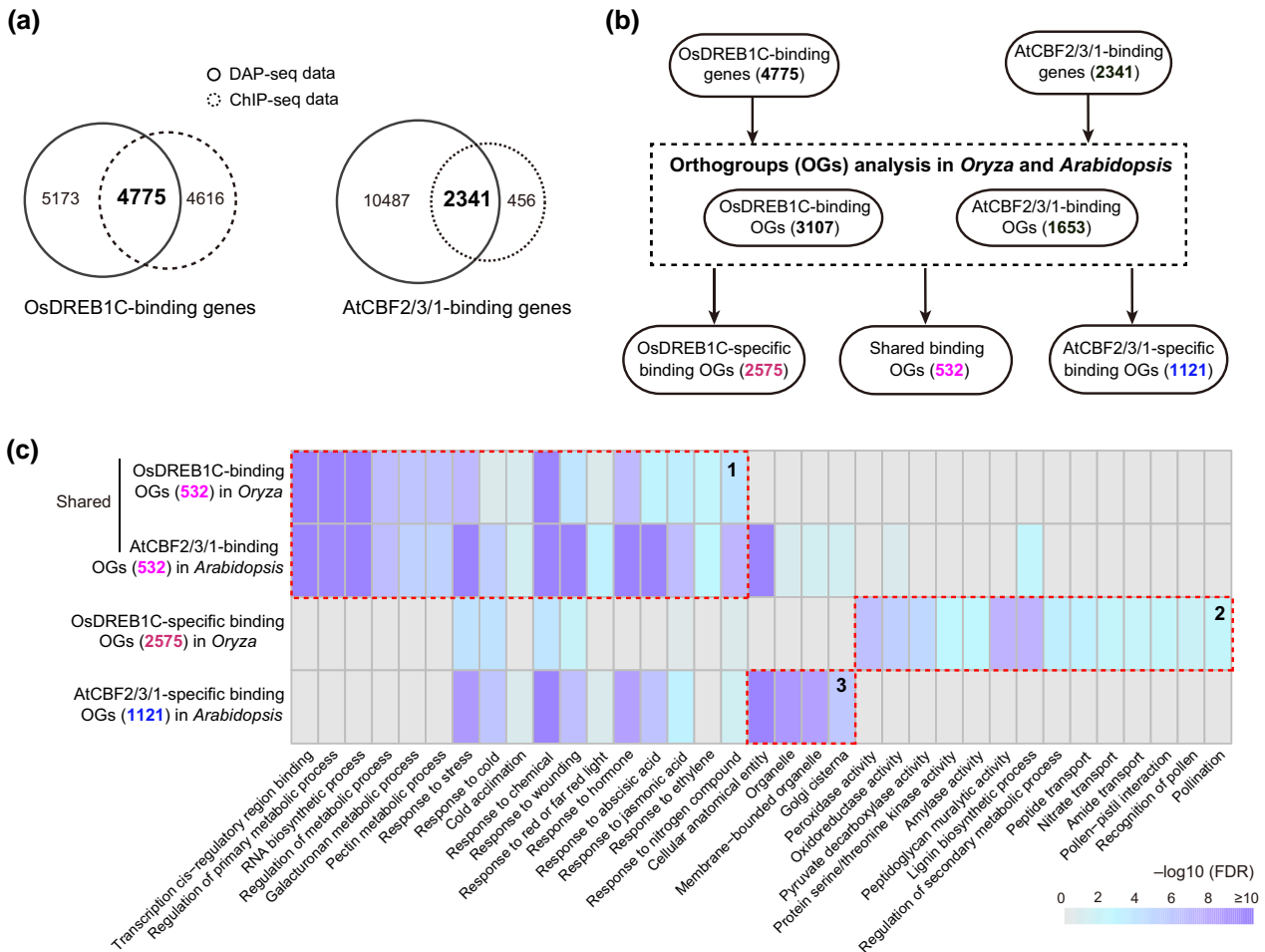


Fig. 3. Genome-wide binding profiles of OsDREB1C and AtCBF2/3/1 and their conserved and specific binding genes and functions in rice and *Arabidopsis*. a) Identification of OsDREB1C and AtCBF2/3/1 binding genes analyzed through both ChIP-seq and DAP-seq. Considering high overlaps and functional redundancy among AtCBF2, AtCBF3, and AtCBF1 (supplementary fig. S2, Supplementary Material online), these three tandemly arrayed genes were collectively termed AtCBF2/3/1 for comparison with OsDREB1C. b) Orthologous comparisons of OsDREB1C and AtCBF2/3/1 binding genes. After OG identification in rice and *Arabidopsis* (see Materials and Methods), these binding genes were categorized as OsDREB1C-specific, AtCBF-specific, or OsDREB1C- and AtCBF2/3/1-shared binding OGs in both species. c) GO term enrichments of genes from shared binding OGs, OsDREB1C-specific, and AtCBF2/3/1-specific bound OGs. Enrichments of genes from shared, OsDREB1C-specific, or AtCBF-specific OGs are demarcated by dotted boxes. The significance ($-\log_{10}$ [false discovery rate, FDR]) is represented in varying shades from gray to purple, indicating the level of term enrichment analyzed from GO enrichment analysis.

AtCBF1 into a unified entity referred to as AtCBF2/3/1 in *Arabidopsis* for subsequent comparative analyses with OsDREB1C in rice. Therefore, we obtained a total of 4,775 OsDREB1C-binding genes in the rice genome and 2,341 AtCBF2/3/1-binding genes in the *Arabidopsis* genome (Fig. 3a, supplementary table S4, Supplementary Material online).

Shared and Specific Bound Genes of OsDREB1C and AtCBF2/3/1

Given complex relationships of coorthologs and in-paralogs between and within plant genomes (Koonin 2005; Guo et al. 2022), direct comparisons between genes bound by OsDREB1C in rice and those bound by AtCBF2/3/1 in *Arabidopsis* carry inherent risks. Following a methodology similar to our prior study (Wang, Shen, et al.

2023), we conducted an analysis of orthogroups (OGs), defined as sets of genes derived from a single gene in the last common ancestor (Emms and Kelly 2019). In total, 13,513 OGs were identified in rice and/or *Arabidopsis* (supplementary fig. S3A, Supplementary Material online; see Materials and Methods). Subsequently, we mapped genes bound by OsDREB1C and AtCBF2/3/1 to these OGs, revealing 2,575 OsDREB1C-specific, 1,121 AtCBF2/3/1-specific, and 532 OsDREB1C- and AtCBF2/3/1-shared binding OGs (Fig. 3b, supplementary fig. S3B, Supplementary Material online).

Within the shared binding OGs in rice and *Arabidopsis*, genes were commonly enriched in transcription *cis*-regulatory region binding, RNA biosynthetic process, regulation of metabolic process, galacturonan and pectin metabolic process, and responses to stress and hormone (Fig. 3c). In OsDREB1C-specific bound OGs in rice, genes

were predominantly enriched in enzyme activities, lignin biosynthetic process, nitrogen transport and metabolic processes, and pollen–pistil transport and pollination, aligning with the enhanced NUE usage and early flowering phenotype observed in *OsDREB1C* overexpression lines (Wei et al. 2022). For AtCBF2/3/1-specific bound OGs in *Arabidopsis*, genes were enriched in cellular anatomical entities, membrane-bounded organelles, and Golgi cisterna. Notably, it appeared that AtCBF2/3/1-binding genes exhibited greater enrichments in hormone-related processes (e.g. abscisic acid and jasmonic acid) and stress responses (e.g. cold and salt) than *OsDREB1C*-binding genes, although the latter also displayed significance in these terms (Fig. 3c). Collectively, these findings underscore the conservative binding of over 500 OGs by both *OsDREB1C* and AtCBF2/3/1, likely descending from the last common ancestor of rice and *Arabidopsis*. Simultaneously, there are thousands of OGs specifically bound by *OsDREB1C* or AtCBF2/3/1, indicative of divergent functions emerging after the split of rice and *Arabidopsis*.

Genome-Wide Binding Motifs of *OsDREB1C* and AtCBF2/3/1

The specific binding of CBF/DREB1 proteins to the C-repeat/dehydration-responsive element (CRT/DRE) with a core motif G/ACCGAC is a well-established relation (Stockinger et al. 1997; Liu et al. 1998). To investigate the genome-wide DNA-binding motifs of *OsDREB1C* and AtCBF2/3/1, we conducted a motif enrichment analysis using the aforementioned DAP/ChIP-seq experimental data. Consistent with previous findings (Song et al. 2021; Wei et al. 2022), a significant enrichment of the motif G/ACCGAC was observed in AtCBF2/3/1-binding regions, while the corresponding motif in *OsDREB1C*-binding regions was identified as GCCGAC (Fig. 4a). The preference for binding to GCCGAC was also noted in previous studies on *OsDREB1A* and ZjDREB1.4, albeit at small scales (Dubouzet et al. 2003; Feng et al. 2019). Moreover, count and proportion analyses confirmed a high occurrence of GCCGAC and an exceedingly low incidence of pure ACCGAC in *OsDREB1C*-binding regions, in stark contrast to comparable proportions of these sequences in AtCBF2/3/1-binding regions (Fig. 4b).

Structural Hints in Binding Divergence of *OsDREB1C* and AtCBF2/3/1

To explore the cause of their binding divergence, we aligned the DNA-binding AP2/ERF domain sequences of CBF/DREB1 proteins in *Arabidopsis* and rice (Fig. 4c). Experimental evidence from prior studies has pinpointed two residues (V14 and E19) as pivotal for the DNA-binding divergence of DREB1 and DREB2 (Sakuma et al. 2002). Unfortunately, the two residues remain identical between *OsDREB1* and AtCBF proteins and thus seem impossible to determine their DNA-binding divergence. Instead, there are eight specific residues (G8, P9, A10, R12, S25, A33, A37, and G53) in *OsDREB1C* that differ

from those in AtCBF2/3/1 proteins (Fig. 4c). Notably, P9, S25, A33, A37, and G53 are also distinct in *OsDREB1A*. Given both *OsDREB1A* and *OsDREB1C* can bind the same DNA motif, GCCGAC (Dubouzet et al. 2003), these residues might not be crucial for the recognition of GCCGAC. Therefore, our focus shifted to the remaining three residues, G8, A10, and R12, present in both *OsDREB1C* and *OsDREB1A* but distinct from those in AtCBF2/3/1 proteins, due to their potential significance in specific binding to GCCGAC.

Conducting molecular docking simulations of *OsDREB1C*-binding GCCGAC (see Materials and Methods), we observed a structural conformation consistent with the NMR structure of AtERF100 (previously designated AtERF1; Allen et al. 1998). The AP2/ERF domain of *OsDREB1C* forms a three-stranded β -sheet and an approximately parallel α -helix, allowing the β -sheet to fit into the major groove of the DNA (Fig. 4d). Notably, A10 and R12 are positioned in close proximity to the initial guanine (G) of the sequence GCCGAC, suggesting a potential direct interaction with G or an influence on the structural configuration to facilitate binding to GCCGAC (Fig. 4e and f). Furthermore, in comparison to A10, when R12 was subjected to an *in silico* mutation to mimic the AtCBF2/3/1 counterpart, the interaction seemed compromised (supplementary fig. S4, Supplementary Material online).

Functional Role of Residue R12 in *OsDREB1C* Preferentially Binding GCCGAC

To validate the role of R12, we conducted an electrophoretic mobility shift assay (EMSA) with DNA fragments containing unlabeled GCCGAC, ACCGAC, or TTTTAC sequences as competitors to determine their binding specificity to AtCBF1, *OsDREB1C*, and *OsDREB1C*-m2 (R12K; Fig. 5a). As expected, AtCBF1 exhibited efficient binding to ACCGAC, with reduced binding in the presence of unlabeled GCCGAC or ACCGAC (Fig. 5b). *OsDREB1C* demonstrated robust bind to GCCGAC, significantly diminished upon the addition of unlabeled GCCGAC fragments compared to ACCGAC (Fig. 5c). Consistent with the genome-wide binding motifs (Fig. 4a), the EMSA data affirm that AtCBF1 binds GCCGAC and ACCGAC in an unbiased manner, while *OsDREB1C* exhibits a preference for GCCGAC. Intriguingly, the binding of *OsDREB1C*-m2 (R12K) to GCCGAC is substantially impaired, further decreasing after the introduction of unlabeled ACCGAC or GCCGAC fragments (Fig. 5d). This experimental validation aligns with the above simulated crystal structure depicting a compromised interaction between *OsDREB1C* and GCCGAC following the *in silico* mutation of R12K (supplementary fig. S4B, Supplementary Material online). Additionally, the examination of DNA binding for *OsDREB1C*-m1 (A10_) revealed a binding preference to GCCGAC akin to wild-type (WT) *OsDREB1C* (supplementary fig. S5, Supplementary Material online), showing a negligible role of A10 in binding the initial G

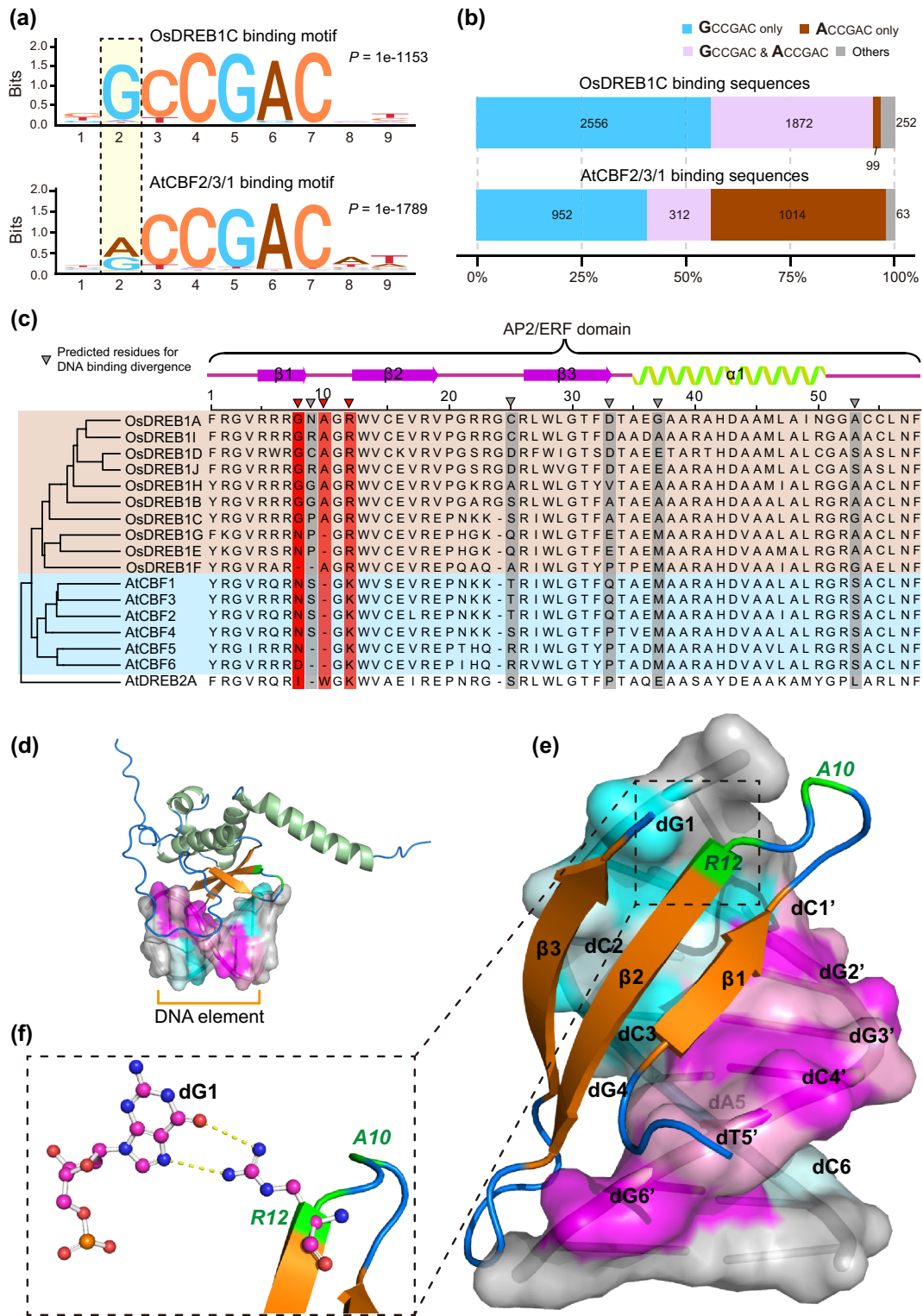


Fig. 4. DNA-binding specificities of OsDREB1C and AtCBF2/3/1 and potential determining residues. a) Genome-wide binding motifs of OsDREB1C and AtCBF2/3/1. The statistical significance (P value) for each motif was calculated from HOMER2 (Aasted et al. 2015). The main specific binding site between the motifs is highlighted within the dashed box. b) Number and proportion of different motif sequences in OsDREB1C- and AtCBF2/3/1-binding regions. c) Multiple sequence alignment of DNA-binding AP2/ERF domains of CBF/DREB1 proteins in *Arabidopsis* and rice. The secondary structures of the AP2/ERF domain are depicted above the panel, and residues potential to determine the DNA-binding divergence of OsDREB1C and AtCBF2/3/1 are highlighted. d to f) Molecular docking of the OsDREB1C-GCCGAC interaction. The panoramic interaction interface of OsDREB1C and GCCGAC d), interaction depicted as a cartoon e), and predicted H bond interactions of R12 and dG1 f). Residues (A10 and R12) predicted to interact with dG1 or optimize the structure configuration to bind GCCGAC are highlighted in green and labeled. DNA is represented as a cartoon with a transparent surface. The DNA nucleotides interacting with OsDREB1C are labeled and colored in mazarine/aquamarine for one strand and magenta/pink for the complementary strand. β sheet: orange; α helix: pale green; and loops: marine. R, arginine; A, alanine.

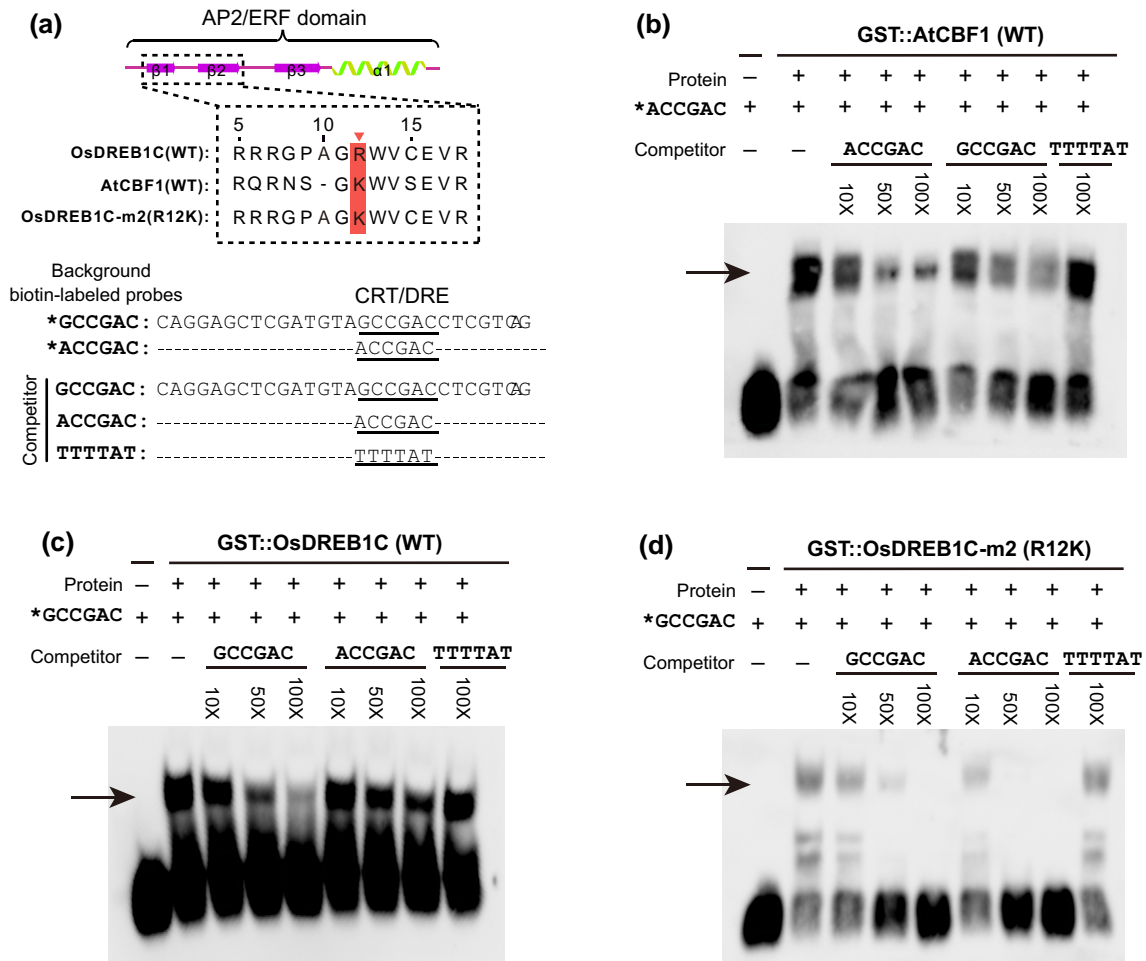


Fig. 5. The residue contributes to the preference of OsDREB1C binding to GCCGAC. a) Schematic representation of AtCBF1, OsDREB1C, and OsDREB1C-m2 proteins, along with DNA-binding fragments containing GCCGAC or its variants. Two sites (A10 and R12) in the AP2/ERF domain (Fig. 4e) of OsDREB1C were selected and mutated to the counterparts of AtCBF2/3/1, producing two variant proteins: OsDREB1C-m1 (A10; [supplementary fig. S5, Supplementary Material](#) online) and OsDREB1C-m2 (R12K). The 29-bp fragment from the OsNR2 exon with a core GCCGAC sequence and its variants are depicted below. b to d) Gel mobility shift assay illustrating DNA-binding affinities of AtCBF1 b), OsDREB1C c), and OsDREB1C-m2 d). The arrow shows the location of the protein–probe complex band. Competitor fragments, including 10-fold (10×), 50-fold (50×), or 100-fold (100×) excess amounts of unlabeled ACCGAC, GCCGAC, or TTTTAT, were employed. In contrast to OsDREB1C-m2, OsDREB1C-m1 exhibited a binding preference for GCCGAC, consistent with WT OsDREB1C ([supplementary fig. S5, Supplementary Material](#) online).

of GCCGAC ([supplementary fig. S4A, Supplementary Material](#) online). Consequently, it becomes evident that the residue (R12) in the OsDREB1C AP2/ERF domain plays a pivotal role in the preferential binding of OsDREB1C to GCCGAC, potentially contributing to gene variations bound and regulated by OsDREB1C and AtCBF2/3/1 across their respective genomes.

Shared and Specific Genes Bound and Regulated by OsDREB1C and AtCBF2/3/1

The mere binding of transcription factors does not inherently translate into the regulation of gene expression. To discern the genes under the regulatory influence of OsDREB1C and AtCBF2/3/1, we scrutinized differentially expressed genes (DEGs) through RNA-seq analyses on OsDREB1C-OE lines (Wei et al. 2022) and *cbfs* mutants

(Song et al. 2021), comparing them with their respective WT (see Materials and Methods). This yielded a total of 1,769 OsDREB1C-regulated genes in rice and 1,194 AtCBF2/3/1-regulated genes in *Arabidopsis* (Fig. 6a). Parallel to the above categorization (Fig. 3b), these regulated genes were further classified into 1,279 OsDREB1C-specific, 722 AtCBF2/3/1-specific, and 172 OsDREB1C- and AtCBF2/3/1-shared regulated OGs. Furthermore, a holistic analysis, integrating RNA-seq with the above detailed DAP/ChIP-seq analysis, revealed 245 OsDREB1C-specific bound and regulated OGs in rice, 216 AtCBF2/3/1-specific bound and regulated OGs in *Arabidopsis*, and approximately 52 to 65 OGs that were shared, bound, and regulated by both OsDREB1C and AtCBF2/3/1 across the rice and *Arabidopsis* genomes (Fig. 6a). This observation suggests that a considerable number of orthologous genes are directly regulated by

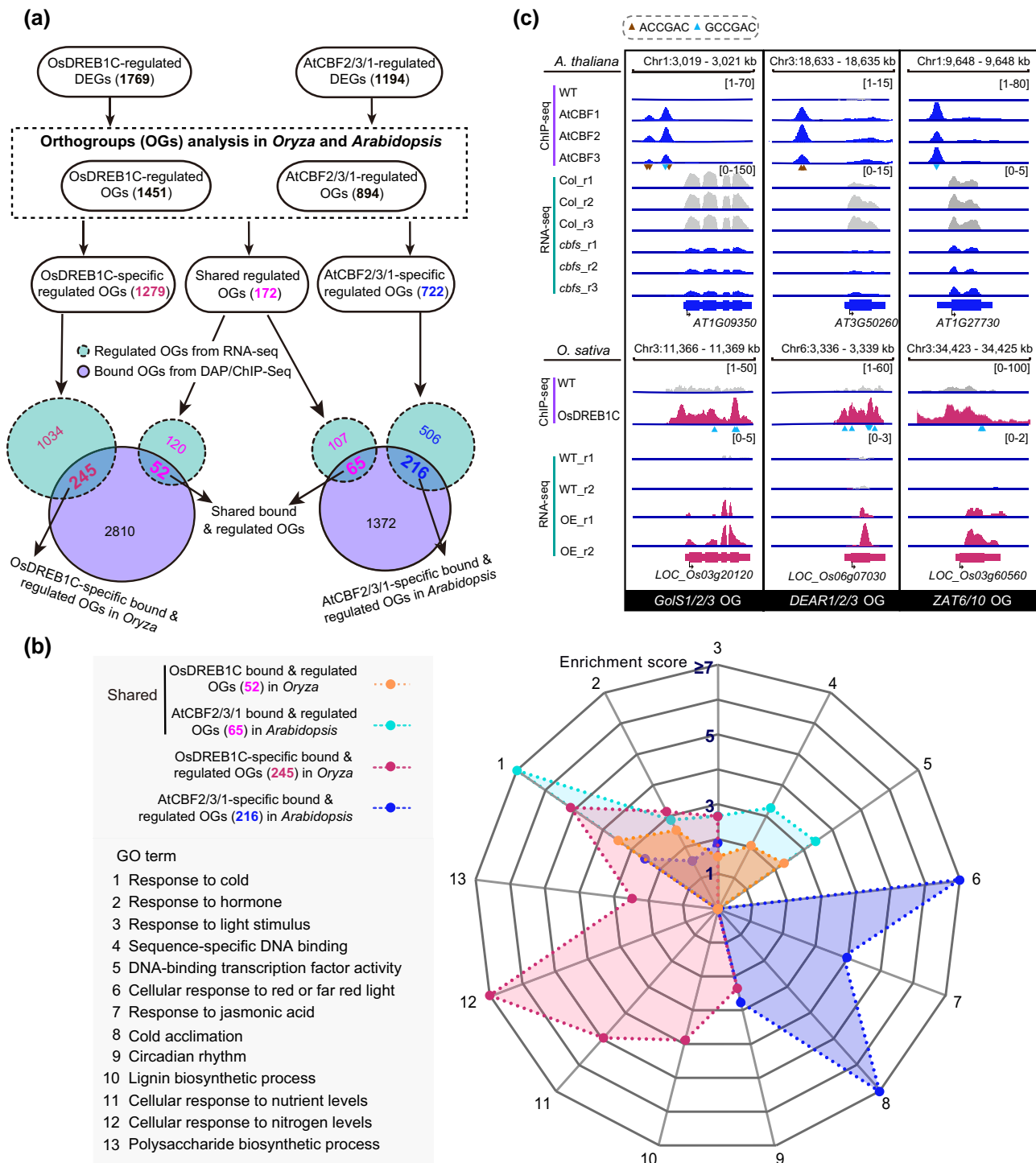


Fig. 6. Conserved and specificity of AtCBF2/3/1- and OsDREB1C-regulated genes in the rice and *Arabidopsis* genomes. a) A pipeline for comparing OsDREB1C and AtCBF2/3/1 bound and regulated genes. DEGs were identified from RNA-seq experiments involving *OsDREB1C-OE* lines (Wei et al. 2022) and *cbfs* mutants (Song et al. 2021) compared to their respective WT (see Materials and Methods). These DEGs were categorized as OsDREB1C- and AtCBF2/3/1-shared regulated OGs, OsDREB1C-, or AtCBF2/3/1-specific regulated OGs in either rice or *Arabidopsis*. The lower left panel presents a Venn diagram illustrating shared or special genes bound and regulated by OsDREB1C in rice, while the lower right panel displays another Venn diagram illustrating shared or special genes bound and regulated by AtCBF2/3/1 in *Arabidopsis*, between ChIP/DAP-seq and RNA-seq analysis. b) Radar chart depicting selected enrichment terms of shared and specific genes directly regulated by OsDREB1C and AtCBF2/3/1. Each segment in the chart indicates enrichment terms from a set of genes explained in the upper left panel. c) Three illustrative examples of shared COR genes bound and regulated by OsDREB1C and AtCBF2/3/1. For clarity, representative genes from *GolS1/2/3*, *DEAR1/2/3*, and *ZAT6/10* OGs in rice and *Arabidopsis* were selected to demonstrate OsDREB1C and AtCBF2/3/1 binding peaks and expression changes in the triple mutant *cbfs* or *OsDREB1C* overexpression lines, as calculated from the ChIP-seq and RNA-seq data (see Materials and Methods). Three additional examples of COR genes are provided in [supplementary fig. S6A, Supplementary Material online](#).

OsDREB1C and AtCBF2/3/1, likely inherited from their last common ancestor. Moreover, a multitude of genes are subject to direct and specific regulation either by OsDREB1C in rice or by AtCBF2/3/1 in *Arabidopsis*.

Moreover, functional enrichment analyses revealed compelling associations within the distinct regulatory roles of OsDREB1C-specific and AtCBF2/3/1-specific directly regulated OGs. OsDREB1C-specific OGs exhibited a pronounced connection with cellular responses to nutrient and nitrogen, as well as polysaccharide and lignin biosynthetic processes, imparting substantial influence on plant growth (Fig. 6b). Conversely, AtCBF2/3/1-specific directly regulated OGs showed enrichment in cellular responses to red or far-red light, response to jasmonic acid, and cold acclimation, aligning with previously established functions of AtCBF2/3/1 (Jia et al. 2016; Zhao et al. 2016; Song et al. 2021). Those OGs shared and regulated by OsDREB1C and AtCBF2/3/1 demonstrated enrichment in analogous biological processes, encompassing responses to cold, hormone, and light stimulus, as well as sequence-specific DNA binding and transcription activity, albeit more pronounced in *Arabidopsis* than in rice (Fig. 6b). In the context of the cold response, for example, representative OGs (e.g. *GolS1/2/3*, *DEAR1/2/3*, *ZAT6/10*, *RD26*, *ERD4*, and *CYS1/4*) exhibited shared binding and upregulation by both OsDREB1C and AtCBF2/3/1, although many of their binding sites had changed (Fig. 6c, supplementary fig. S6A, Supplementary Material online). Furthermore, various CBF/DREB1 genes, including OsDREB1C, are induced by cold stress across diverse angiosperm lineages (Wang, Zhang, et al. 2023; supplementary fig. S6B, Supplementary Material online). According to the maximum parsimony principle, these findings suggest the establishment of a cold-responsive regulatory network mediated by CBF/DREB1 in the last common ancestor of monocots and eudicots, consistent with the previous study (Nie et al. 2022). However, after the split of rice and *Arabidopsis*, OsDREB1C and AtCBF2/3/1 evolved independently to target and regulate distinct sets of genes, primarily contributing to the evolutionarily functional divergence of OsDREB1C and AtCBF2/3/1.

Gains and Losses of OsDREB1C- and AtCBF2/3/1-Regulated Genes

Following their evolutionary divergence, rice and *Arabidopsis* evolved independently, experiencing great alterations in genetic materials through at least three large-scale WGDs and numerous small-scale duplications (SSDs) accompanied by biased gene retention and loss (Qiao et al. 2019; Wu et al. 2020). This intricate process led to the emergence of intertwined coorthologs and in-paralogs between and within species. The repercussions of these genetic changes may influence the divergence of OsDREB1C- and AtCBF2/3/1-regulated genes.

In detail, we delved into the evolution of genes within 297 OsDREB1C and 281 AtCBF2/3/1 directly regulated OGs across the rice and *Arabidopsis* genomes (see

Materials and Methods). Among these regulated OGs, approximately 12.1% to 17.1% exhibited conservative regulation in the same direction by OsDREB1C in rice and by AtCBF2/3/1 in *Arabidopsis* (Fig. 7a). In contrast, around 25.6% of OsDREB1C directly regulated OGs underwent evolutionary changes specific to rice or were lost in *Arabidopsis*. Similarly, approximately 16.7% of AtCBF2/3/1 directly regulated genes experienced evolution specific to *Arabidopsis* or were lost in rice. For example, well-known representative AtCBF2/3/1-regulated genes associated with cold tolerance, such as *COR15A/15B*, *KIN1/2*, and *RD29A*, are conspicuously absent in the rice genome (Fig. 7a). These findings underscore that over the course of extensive evolution, biased gene retention and loss, occurring independently in rice and *Arabidopsis*, have dynamically contributed to the gains and losses shaping the divergence of OsDREB1C- and AtCBF2/3/1-regulated genes.

Expression Divergence of Genes Regulated by OsDREB1C and AtCBF2/3/1

The divergence may be further magnified by expression reprogramming among orthologous genes, notwithstanding their tendency to share similar functions across species (Gabaldón and Koonin 2013). In a detailed exploration, we investigated the expression dynamics of orthologous genes within the directly regulated OGs by OsDREB1C and AtCBF2/3/1 between rice and *Arabidopsis* (Fig. 7a). Notably, approximately 57.2% of OsDREB1C directly regulated OGs exhibited specific regulation by OsDREB1C in rice, while their counterparts in *Arabidopsis* showed no discernible regulation by AtCBF2/3/1. For example, representative genes, such as early-flowering genes (*FTL1* and *MADS1*), nitrogen transporter and reductase genes (*NR2*, *DNRT1.1B*, and *NRT2.4*), crucial photosynthesis-related genes (*RbcS3* and *ppc2b*), and those associated with grain size and yield (*Atg8a* and *PINA/C*), were directly bound by OsDREB1C on their promoters and/or exons, resulting in significant upregulation in OsDREB1C overexpression lines (Fig. 7b, supplementary fig. S7A, Supplementary Material online). Conversely, their closely related orthologs in *Arabidopsis* exhibited no discernible binding by AtCBF2/3/1 and their expression remained unaltered after AtCBF2/3/1 knockout in the triple mutant *cbfs*.

Likewise, around 60.1% of AtCBF2/3/1 directly regulated OGs were exclusively regulated by AtCBF2/3/1 in *Arabidopsis* (Fig. 7a). Representative genes associated with cold or related tolerance, such as *COR47*, *COR413-PM1*, *HVA22D/E*, *COR413Im1*, *RD2*, *AIRP2*, *DI21*, *UGT79B2*, and *PLC1*, were directly bound by AtCBF2/3/1 on their promoters, with significant downregulation after AtCBF2/3/1 knockout in the triple mutant *cbfs*. In contrast, their closely related orthologs in rice exhibited no regulation by OsDREB1C (Fig. 7c, supplementary fig. S7B, Supplementary Material online). Furthermore, a small proportion (~5.1% to 6.1%) of OGs were regulated in opposite orientations by OsDREB1C in rice and by AtCBF2/3/1 in *Arabidopsis*

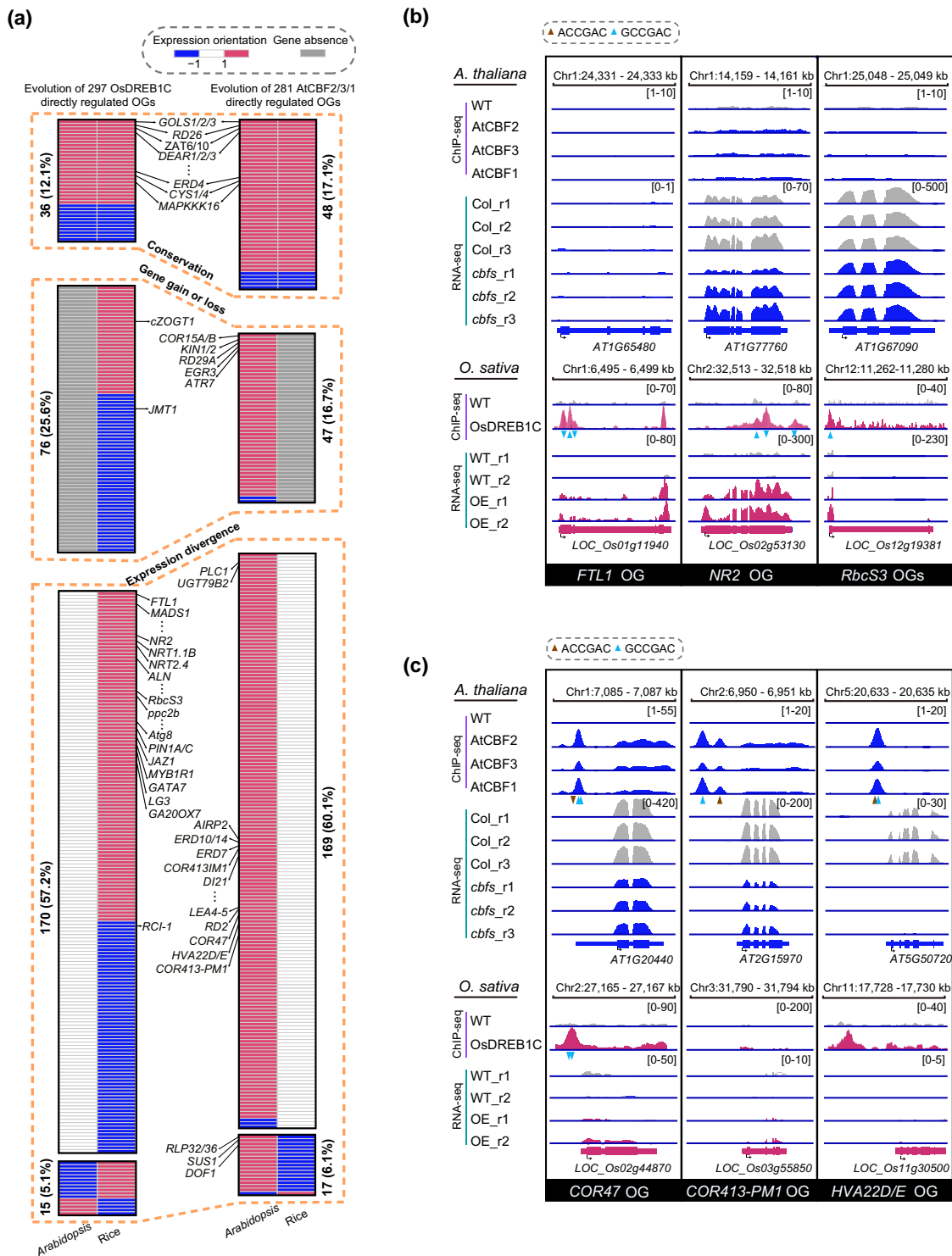


Fig. 7. Gene gains/losses and expression divergence of *OsDREB1C*- and *AtCBF2/3/1*-regulated OGs. a) A schematic diagram illustrating conservation, gains/losses, and expression divergence of *OsDREB1C*- and/or *AtCBF2/3/1*-regulated OGs across *Arabidopsis* and rice. The 297 and 281 OGs directly regulated by *OsDREB1C* and *AtCBF2/3/1* were identified through DAP/ChIP-seq and RNA-seq analyses (Fig. 6a). Each OG is represented by a selected gene indicating binding and expression changes regulated by *OsDREB1C* or *AtCBF2/3/1*. The chart displays OGs with representative genes in rice and *Arabidopsis* based on their expression changes: upregulation ($\log_2FC > 1$), downregulation ($\log_2FC < -1$), and non-differential expression ($-1 < \log_2FC < 1$). The absence of OGs in rice or *Arabidopsis* is also indicated. Well-documented growth- and stress-related genes are highlighted. The details of the OGs are provided in [supplementary table S5, Supplementary Material](#) online. b and c) Representatives of expression divergence in growth-related OGs regulated by *OsDREB1C* but not *AtCBF2/3/1* (e.g. *RbcS3*, *NR2*, and *FTL1*) b) and stress-related OGs regulated by *AtCBF2/3/1* but not *OsDREB1C* (e.g. *COR47*, *COR413-PM1*, and *HVA22D/E*) c). The mapping coverage from the ChIP-seq and RNA-seq is visualized by IGV v2.15.2 (Robinson et al. 2011), with additional representatives of growth- and stress-related OGs provided in [supplementary fig. S7, Supplementary Material](#) online.

(Fig. 7a). Consequently, these findings demonstrate that over half of the targeted genes between rice and *Arabidopsis* experienced divergent evolution in regulation by OsDREB1C and AtCBF2/3/1, thereby amplifying the functional divergence of OsDREB1C and AtCBF2/3/1.

Discussion

Functional divergence of orthologous genes challenges the traditional belief in their functional conservation across species. This divergence, most likely attributed to subfunctionalization and neofunctionalization after gene duplications across species, underscores the complexity and adaptability of orthologous genes beyond mere evolutionary parallels. Orthologous *OsDREB1C* and *AtCBF2/3/1* evolved independently in distinct clades across different species. *OsDREB1C* emerged from a TSP before Poaceae divergence (~101 Mya; Huang et al. 2022), while *AtCBF2/3/1* experienced TD approximately 29.2 Mya in the ancient Brassicaceae (Nie et al. 2022). The genome-wide binding analysis illuminated a preference of OsDREB1C for GCCGAC, in contrast to the A/GCCGAC preference exhibited by AtCBF2/3/1. Experimental validation through protein–DNA docking and EMSA experiments confirmed that R12 in the AP2/ERF domain plays a critical role in the interaction of OsDREB1C with the initial G of GCCGAC. The divergence in DNA-binding motifs contributed substantially to the observed differences in the genes bound and regulated by OsDREB1C and AtCBF2/3/1. Cross-species analyses unveiled that a limited subset of ~12.1% to 17.1% of OGs exhibited conservative regulation by both OsDREB1C and AtCBF2/3/1. In contrast, ~16.7% to 25.6% of OGs were either gained or lost in a species-specific manner, regulated by OsDREB1C in rice or by AtCBF2/3/1 in *Arabidopsis*. This phenomenon primarily resulted from biased gene retention and loss following WGDs and SSDs during the extensive evolution after species divergence. Moreover, the functional divergence of OsDREB1C and AtCBF2/3/1 was substantially magnified by expression reprogramming of these OGs, ranging from 62.3% to 66.2% differentially regulated by OsDREB1C and AtCBF2/3/1. In summary, we conclude that two keys underpin the functional divergence of OsDREB1C and AtCBF2/3/1: (i) biased gene gains and losses and (ii) expression reprogramming of OsDREB1C- and AtCBF2/3/1-regulated genes during the extensive evolutionary trajectory following the split of *Arabidopsis* and rice.

Binding Regions and Transcriptional Outcomes

In comparison to the preferential binding of gene promoters by AtCBF2/3/1, the identified representative genes suggest that OsDREB1C may display a propensity for binding gene exon regions, also supported by a previous study (Wei et al. 2022). Subsequently, we quantified the distribution of the OsDREB1C-binding sequence GCCGAC across the rice genome. Remarkably, a significant higher

proportion was detected in exon regions compared to other genomic regions, including promoters, intergenic, and introns (supplementary fig. S8, Supplementary Material online). Additionally, the detected ChIP-seq binding peaks of OsDREB1C were broader than those of AtCBF2/3/1 (Figs. 6c and 7b and c). This discrepancy was likely attributed to significantly longer average fragment lengths in the DREB1C cDNA library (~300 bp) as compared to those in the AtCBF2/3/1 libraries (~255 bp). Despite DREB1C exhibiting broader peaks, motif analysis revealed solely the GCCGAC motif, highlighting the targeted specificity of DREB1C's interaction with the motif (Fig. 4a). The transcriptional outcomes of OsDREB1C and AtCBF2/3/1 were also examined. Notably, 96.1% of OGs directly regulated by AtCBF2/3/1 were upregulated, in stark contrast to the comparable proportions of upregulated (55.9%) and downregulated (44.1%) OGs by OsDREB1C (Fig. 7a, supplementary table S5, Supplementary Material online). The higher prevalence of upregulated genes by AtCBF2/3/1 was consistently reported in other two independent studies (Jia et al. 2016; Zhao et al. 2016). These discrepancies in DNA-binding motifs and regions, coupled with differential transcriptional outcomes, contribute to expression reprogramming in OsDREB1C- and AtCBF2/3/1-regulated genes. Moreover, it is crucial to note that DNA-binding and gene expression changes were detected in rice lines with *OsDREB1C* overexpression (Wei et al. 2022), which might inadvertently regulate genes beyond its specific targets, potentially leading to an overestimation of genes regulated by OsDREB1C. This scenario might exaggerate the discrepancies between genes regulated by OsDREB1C and those by AtCBF2/3/1.

Contrasting Regulatory Roles in Gibberellin Signaling Pathway

Gibberellin (GA) is well established for its pivotal role in plant growth and floral transition, achieved through the inhibition of the growth-inhibitory DELLA proteins. Achard et al. (2008) demonstrated that AtCBF1 upregulates the expression of *RGL3* (encoding a DELLA protein), along with *GA2ox3* and *GA2ox6* (encoding GA-2 oxidase, involved in inactivation of GAs). This upregulation leads to a reduction in active GA levels and an accumulation of DELLA, resulting in dwarfism in *AtCBF1* overexpression plants (Achard et al. 2008). We also examined the binding and expression patterns of these genes utilizing the aforementioned ChIP-seq and RNA-seq data related to OsDREB1C and AtCBF2/3/1. Consistently, AtCBF2/3/1 binding peaks were observed in the promoters of *GA2ox6* and *RGL3* genes, with all three genes exhibiting decreased expression after AtCBF2/3/1 knockout in the triple mutant *cbfs* when compared to the WT (supplementary fig. S9, Supplementary Material online). In contrast, the expression of *GA2ox3* and *GA2ox6* orthologs in rice showed no significant change, and the expression of the *RGL3* ortholog in rice (*SLR1*) was relatively downregulated in rice overexpressing *OsDREB1C*. Intriguingly, in contrast to

AtCBF2/3/1, which exhibited neither binding nor regulation of the expression of GA20ox7, an essential gene in GA biosynthesis, OsDREB1C was observed to bind and significantly upregulate its expression ([supplementary fig. S9, Supplementary Material](#) online). These observations propose that OsDREB1C may have evolved a diverse range of functions, including the activation of the GA signaling pathway, in stark contrast to the repression by AtCBF2/3/1. However, it is imperative to conduct additional experiments to corroborate and validate the positive regulatory role of OsDREB1C in the GA signaling pathway.

Discrepancies in Growth Phenotypes

Moreover, studies investigating the overexpression of the same *CBF/DREB1* gene have yielded inconsistent or discrepant results. For example, [Ito et al. \(2006\)](#) observed growth retardation under nonstress conditions in rice overexpressing *OsDREB1A*, while another study noted mild growth retardation, particularly during bolting timing, when *OsDREB1A* was overexpressed in *Arabidopsis* ([Dubouzet et al. 2003](#)). Additionally, constitutive expression of *OsDREB1B* in tobacco did not induce growth retardation or observable phenotypic changes ([Gutha and Reddy 2008](#)). In contrast, transgenic rice plants overexpressing *OsDREB1B* exhibited growth retardation in other studies ([Dubouzet et al. 2003](#); [Ito et al. 2006](#)). The reasons for the discrepancies remain elusive. Nonetheless, considering these reports alongside our results, it is speculated that the seemingly inconsistent growth phenotypes may be, at least in part, related to the host plants, particularly given the differences between eudicots and monocots. Alternatively, the diverse outcomes could be attributed to varying levels of *CBF/DREB1* overexpression in these studies. In any case, the relationships among *CBF/DREB1* genes are intricate and warrant further in-depth exploration.

Materials and Methods

Plant Genomes, Gene Annotations, and Protein Sequences

To unravel the evolutionary trajectory of *CBF/DREB1* genes, with a focus on OsDREB1C in monocot rice, we selected 20 representative species, including 2 basal angiosperms, 4 eudicots, and 14 monocots, and downloaded their genomes, gene annotations, and protein sequences from publicly available databases, with detailed information provided in [supplementary table S1, Supplementary Material](#) online.

Timing of WGD Events

For γ -, β -, α -, τ -, σ -, ρ -, and other WGD events, we directly obtained their inferred times from previous reports ([supplementary table S2, Supplementary Material](#) online). To ascertain the timeline of the α^M -WGD event in *Miscanthus sinensis*, we applied the formula $t = Ks/2r$. Utilizing the average synonymous substitution rate per site per year (r) in monocots ([Gaut et al. 1996](#)) and

synonymous substitution (Ks) values derived from collinear homologous pairs in the *M. sinensis* genome, we determined the date of α^M -WGD to be approximately 5.92 ± 0.01 Mya, with 95% confidence interval.

CBF/DREB1 Identification and Phylogenetic Analysis

Initially, we identified AP2/ERF genes by searching the hidden Markov model (HMM) profile of the AP2/ERF domain (PF00847) against the above species protein sequences using HMMER3.1 with E value = 1×10^{-10} ([Eddy 2011](#)). Utilizing *Arabidopsis* and rice *CBF/DREB1* protein sequences as templates ([Nie et al. 2022](#)), we specifically constructed the HMM profile of *CBF/DREB1* proteins to identify their orthologs within AP2/ERF protein sequences. Finally, a total of 154 *CBF/DREB1* genes were identified from the above 20 selected plant species ([supplementary table S3, Supplementary Material](#) online).

For an in-depth exploration of the evolutionary relationships among *CBF/DREB1* genes, we conducted multiple alignments of their protein sequences using MAFFT v7.453 ([Katoh and Standley 2013](#)). Subsequently, a phylogenetic tree was constructed by IQ-TREE v2.1.3 ([Minh et al. 2020](#)), using the maximum likelihood (ML) analysis with 1,000 bootstrap replicates. The best model, JTT + F + R6, was selected from 545 candidates ([Minh et al. 2020](#)). To establish the phylogenetic root, five close homologs of *CBF/DREB1* from DREB III genes were chosen as the outgroup.

Gene Duplication and Transposition Analysis

In accordance with previously outlined procedures ([Nie et al. 2022](#)), we applied MCScanX ([Wang et al. 2012](#)) and DupGen_finder ([Qiao et al. 2019](#)) to discern various duplication modes, such as whole-genome/segmental duplication (WGD), TD, TSP, and dispersed duplications (DDs), responsible for generating paralogous genes in each species. Specifically, we extracted *CBF/DREB1* genes with their respective duplication modes.

Genome-Wide Binding Analysis of *CBF/DREB1* Proteins

We curated multiple experimental data sets, encompassing genome-wide binding profiles for *CBF/DREB1* proteins, including AtCBF1, AtCBF2, and AtCBF3 by ChIP-seq (NCBI PRJNA732005) and DAP-seq (PRJNA257556; [O'Malley et al. 2016](#); [Song et al. 2021](#)) and OsDREB1C by ChIP-seq (PRJNA841272) and DAP-seq (PRJNA841281; [Wei et al. 2022](#)). The sequencing reads were aligned to their respective reference genomes—*A. thaliana* (TAIR10) and *O. sativa* Nipponbare (MSU/TIGR V7)—using Bowtie2 with default parameters ([Langmead and Salzberg 2012](#)). Subsequent analyses involved peak calling (q value < 0.01 and fold enrichment > 1.5) utilizing MACS v2.7.1 ([Zhang et al. 2008](#)), followed by annotation of significant peaks based on q value with their nearest genes employing ChIPseeker v1.20 ([Yu et al. 2015](#)). These proximate genes were considered potential targets bound by the corresponding *CBF/DREB1* proteins.

The replicate numbers for DAP-seq and ChIP-seq experiments are presented in [supplementary figs. 10A and 11A](#). To assess experiment reproducibility, we conducted an irreproducibility discovery rate (IDR) analysis ([Li et al. 2011](#)) using the ENCODE IDR pipeline ([Landt et al. 2012](#)). For the replicates of OsDREB1C DAP-seq samples, the analysis demonstrated that both the self-consistency ratio “ $\max(N_1, N_2)/\min(N_1, N_2)$ ” and the rescue ratio “ $\max(N_p, N_t)/\min(N_p, N_t)$ ” are less than 2, exhibiting a significant consistency of peaks (around 80%) at $\text{IDR} \leq 0.05$ between the two true replicates ([supplementary fig. 10B and C, Supplementary Material online](#)). This underscores the high repeatability of OsDREB1C DAP-seq replicates. No IDR analysis was conducted due to the absence of replicates for the OsDREB1C ChIP-seq sample ([Wei et al. 2022](#)). To ensure high confidence in identifying OsDREB1C-binding genes, we retained only those identified by both ChIP-seq and DAP-seq experiments ([supplementary table S4, Supplementary Material online](#)).

For the DAP-seq and ChIP-seq samples of AtCBF2, AtCBF3, and AtCBF1, despite the lack of replicates ([O’Malley et al. 2016; Song et al. 2021](#)) and considering their significant functional redundancy ([Park et al. 2015; Jia et al. 2016; Liu et al. 2018](#)), we performed IDR analysis ([Li et al. 2011](#)). The analysis of AtCBF2/3/1 ChIP-seq samples showed that all values of self-consistency ratios and rescue ratios are less than 2, displaying a substantial consistency of peaks (~52% to 66%) at $\text{IDR} \leq 0.05$ between the true samples ([supplementary fig. 11B to H, Supplementary Material online](#)). For AtCBF2/3/1 DAP-seq samples, despite self-consistency ratios of AtCBF3 versus AtCBF2 and AtCBF1 greater than 2, all rescue ratios are less than 2, also revealing a substantial degree of consistent peaks (~53% to 60%) at $\text{IDR} \leq 0.05$ between the true samples ([supplementary fig. 11B to H, Supplementary Material online](#)). According to the standard of the ratio between peaks enriched in replicates and pseudo-replicates ([Landt et al. 2012; Schmitz et al. 2022](#)), the analyses demonstrate high and accepted reproducibility of AtCBF2/3/1 ChIP-seq and DAP-seq samples, respectively. Given the observed high consistent overlap in binding peaks and the functional redundancy of AtCBF2, AtCBF3, and AtCBF1 ([Song et al. 2021](#)), we collectively treated them as a union, denoted as AtCBF2/3/1, for comparison to OsDREB1C. Similarly, to ensure a high-confidence level in identifying AtCBF2/3/1 binding genes, we retained only those genes identified by both ChIP-seq and DAP-seq experiments ([supplementary table S4, Supplementary Material online](#)).

OG Analysis

Given intertwined coorthologs and in-paralogs arising from recurrent gene gain/loss during long evolution after the divergence of eudicots and monocots (~200 Mya), direct comparisons of genes bound by OsDREB1C and AtCBFs carry inherent risks. Instead, we opted for an OG analysis to trace the evolutionary trajectories of genes in the genomes. To ensure robustness and generalization

across eudicots and monocots, we selected a diverse set of species, including the basal angiosperm (*Amborella trichopoda*), four representative eudicots (*A. thaliana*, *Nelumbo nucifera*, *Carya illinoensis*, and *Populus trichocarpa*), and six representative monocots (*O. sativa*, *Ananas comosus*, *Lemna minor*, *Phyllostachys edulis*, *Setaria italica*, and *Zea mays*). We utilized OrthoFinder v2.5.4 (inflation parameter = 1.2; [Emms and Kelly 2019](#)) to classify all coding genes into OGs. Filtering out OGs lacking representation in both rice and *Arabidopsis*, we identified a total of 13,513 OGs in either or both species. Within these OGs, we categorized those containing genes from both species as ancestor-descendent OGs (~68%, or termed conserved or shared OGs) and those containing genes exclusively from one species as rice- or *Arabidopsis*-specific OGs (~24% in rice and 9% in *Arabidopsis*; [supplementary fig. S3A, Supplementary Material online](#)). Subsequently, we classified the above OsDREB1C- and AtCBF2/3/1-binding genes into OsDREB1C-specific, AtCBF2/3/1-specific, and shared binding OGs in rice and *Arabidopsis* ([supplementary fig. S3B, Supplementary Material online](#)).

Functional Annotation and Enrichment Analysis

To unravel the functional implications of our identified target genes, we conducted comprehensive Gene Ontology (GO) term annotations. The GO term annotations for *A. thaliana* genes were directly obtained from The Arabidopsis Information Resource (TAIR). Subsequently, we employed BLASTP with $E = 1 \times 10^{-10}$ to compare rice protein sequences with *Arabidopsis* proteins ([Altschul et al. 1997](#)), attributing GO terms from the most similar hits to the corresponding rice genes. This systematic approach allowed us to unravel the biological processes, molecular functions, and cellular components associated with the identified target genes.

Genome-Wide Analysis of DNA-Binding Motifs

To elucidate the sequence specificity underlying DNA binding, we conducted an analysis of the DNA-binding motifs for transcription factors—AtCBF2, AtCBF3, and AtCBF1 from *Arabidopsis* and OsDREB1C from rice. Leveraging the above ChIP/DAP-seq analysis, we extracted DNA-binding regions from the respective genomes. Subsequently, we conducted motif calling using HOMER2 with default parameters ([Aasted et al. 2015](#)). Strikingly, the analysis revealed a shared enrichment of the A/GCCGAC motif in the DNA-binding regions of AtCBF2, AtCBF3, and AtCBF1. In contrast, OsDREB1C exhibited a more specific enrichment pattern, with the GCCGAC motif prevailing in its DNA-binding regions.

Structural Modeling and Protein–DNA Docking

In pursuit of understanding the residues crucial for OsDREB1C-binding GCCGAC, we employed a structural modeling and protein–DNA docking analysis. The OsDREB1C 3D structure (AF-Q9LWV3-F1-model_v4) was obtained from the AlphaFold Protein Structure Database

(Varadi et al. 2022). Utilizing this structure and the GCCGAC sequence, we generated 100 protein–DNA interaction models through the HDOCK server (Yan et al. 2017) and carefully selected the optimal model. By analyzing structural position effects, we predicted specific residues (A10 and R12) crucial in the OsDREB1C–GCCGAC interaction. Furthermore, we introduced mutations, OsDREB1C-m1 (A10_) and OsDREB1C-m2 (R12K), by aligning the amino acids with those of AtCBF2/3/1. Using ColabFold v1.5.2-patch (Mirdita et al. 2022) and HDOCK (Yan et al. 2017), we investigate whether these mutations influenced OsDREB1C binding to GCCGAC (supplementary fig. S4, Supplementary Material online). The resulting protein–DNA binding structures were visualized by PyMOL v2.5.5 (Schrödinger).

EMSA

To experimentally validate the impact of the two sites (A10 and R12) on DNA binding, we conducted EMSA employing biotin-labeled probes and GST-tag fusion proteins following the Chemiluminescent EMSA Kit (GS009, Beyotime). Initially, the full-length coding sequences of AtCBF1, OsDREB1C, OsDREB1C-m1 (A10_), and OsDREB1C-m2 (R12K) were cloned into the pGEX-4T-1 vector. These constructs were then expressed in *Escherichia coli* BL21 (DE3), and high-quality purified GST-tag fusion proteins were obtained using the GST-tag Protein Purification Kit (P2262, Beyotime). For the binding assays, a 29-bp fragment containing GCCGAC from the OsNR2 exon was labeled using the EMSA Probe Biotin Labeling Kit (GS008, Beyotime). Unlabeled fragments with GCCGAC, ACCGAC, or TTTTCT were used for competition assays in the context of OsDREB1C, OsDREB1C-m1, and OsDREB1C-m2 binding, whereas the mutated and labeled ACCGAC fragment was used for AtCBF1 binding. The binding reactions were incubated at room temperature (25 °C) for 30 min and subsequently separated on 4% nondenaturing polyacrylamide gels.

Genome-Wide Analysis of OsDREB1C- and AtCBF2/3/1-Regulated Genes

In pursuit of understanding the transcriptional landscape, we obtained RNA-seq data for *OsDREB1C* overexpression lines (*OsDREB1C-OE*) and WT plants (NCBI PRJNA724935; Wei et al. 2022). After removing adapter and low-quality reads, we aligned clean reads to the *O. sativa* Nipponbare reference genome (MSU/TIGR V7) using HISAT2 (Kim et al. 2015) and calculated gene expression levels using StringTie v1.3.4 (Pertea et al. 2016). Subsequently, we employed the DESeq2 package (Love et al. 2014) to identify significantly DEGs between *OsDREB1C-OE* and WT, with a threshold of $P < 0.05$ and $|\log_2 \text{fold change (FC)}| > 1$. This analysis yielded a total of 1,769 DEGs regulated by OsDREB1C in rice. Regarding the genes regulated by AtCBF2/3/1 in *Arabidopsis*, we directly obtained 1,194 AtCBF2/3/1-regulated genes with a threshold of $P < 0.05$ and $|\log_2 \text{FC}| > 1$ from a previous report (Song et al. 2021).

To discern the uniqueness and overlap in gene regulation between OsDREB1C in rice and AtCBF2/3/1 in *Arabidopsis*, we classified these DEGs into OsDREB1C-specific, AtCBF2/3/1-specific, and conserved regulated OGs.

Integrative Analysis of OsDREB1C and AtCBF2/3/1 Binding and Regulation

In our quest for comprehensive insights, we conducted an integrated ChIP/DAP- and RNA-seq analysis, incorporating stringent criteria to ensure study robustness and repeatability. Specifically, genes identified to be bound by OsDREB1C or AtCBF2/3/1 were retained only if they demonstrated binding in both ChIP-seq and DAP-seq experiments. This meticulous selection process resulted in the classification of genes into OsDREB1C-specific, AtCBF2/3/1-specific, and shared binding OGs. Building upon this refined data set, we embarked on an integrative analysis, juxtaposing the bound genes from the ChIP/DAP-seq analysis with the DEGs identified through the RNA-seq analysis. This integrative approach, conducted at the OG level, facilitated the identification of OsDREB1C- or AtCBF2/3/1-specific bound and regulated OGs in rice or *Arabidopsis*, as well as shared bound and regulated OGs in both species.

Cross-Species Comparison of OsDREB1C- and AtCBF2/3/1-Regulated OGs

By the above integrative analysis, we obtained OGs bound and regulated by OsDREB1C in rice and by AtCBF2/3/1 in *Arabidopsis*. To ensure the accuracy in these retrieved OGs, reciprocal BLASTP searches ($E < 1 \times 10^{-10}$, identity $\geq 30\%$) were conducted between rice and *Arabidopsis* and 14 *Arabidopsis* OGs and 9 rice OGs were subsequently dissected due to low intergene identities within the OGs. Accordingly, as shown in Fig. 6a, we obtained a total of 297 (245 + 52) OsDREB1C bound and regulated OGs in rice and 281 (216 + 65) AtCBF2/3/1 bound and regulated OGs in *Arabidopsis*. The refined OGs were then curated to examine conservation, gene gain or loss, and expression divergence of OsDREB1C and AtCBF2/3/1 directly regulated OGs across the rice and *Arabidopsis* genomes. First, focusing regulatory outcomes, we identified representative genes in OGs that not only contained at least one binding peak for OsDREB1C or AtCBF2/3/1 ($q < 0.01$ and fold enrichment > 1.5) but were also differentially regulated in the same orientation ($P < 0.05$ and $\log_2 \text{FC} > 1$ or < -1) by both regulators. These conservatively regulated OGs across rice and *Arabidopsis* genomes were defined, despite the OGs in one of the two species may encompass genes indirectly regulated by OsDREB1C or AtCBF2/3/1. Subsequently, further exploration encompassed OGs unique to either rice or *Arabidopsis*, wherein genes were bound and regulated exclusively by the corresponding CBFs/DREB1s (OsDREB1C or AtCBF2/3/1). This classification demarcated gene gains in one species or losses in the other. Moreover, genes in OGs bearing at least one binding peak for OsDREB1C or AtCBF2/3/1, yet

demonstrating differential regulation or regulation in opposing orientations by OsDREB1C and AtCBF2/3/1, were considered instances of expression divergence or reprogramming in the regulated OGs. The three categorized OGs are provided in [supplementary table S5, Supplementary Material](#) online.

Supplementary material

[Supplementary material](#) is available at *Molecular Biology and Evolution* online.

Acknowledgments

We thank Professors Mingbing Zhou and Xinchun Lin at Zhejiang A&F University for their suggestions on the evolutionary analysis and Professor Zeng Tao at Zhejiang University for his support in conducting the EMSA experiments. We also appreciate Professor Ning Jiang at Michigan State University for her assistance in identifying the transposition of *DREB1C*. Additionally, we thank ChatGPT for its invaluable assistance in language editing, which significantly enhanced the readability of this manuscript.

Author Contributions

W.W. conceptualized and designed the study. D.D., L.G., and Y.N. executed the analysis, and Y.G., L.C., and S.W. designed and conducted the experimental work. D.D., S.W., and W.W. contributed to the presentation and interpretation of the data. D.D. and W.W. collaborated to write and polish the manuscript.

Funding

This work is supported by the National Natural Science Foundation of China (grant number 31871233).

Data Availability

The data underlying this article are available in the article and in its online [Supplementary Material](#).

References

- Aasted CM, Yucel MA, Cooper RJ, Dubb J, Tsuzuki D, Becerra L, Petkov MP, Borsook D, Dan I, Boas DA. Anatomical guidance for functional near-infrared spectroscopy: atlasViewer tutorial. *Neurophotonics*. 2015;2(2):20801. <https://doi.org/10.1117/1.NPh.2.2.020801>.
- Achard P, Gong F, Cheminant S, Alioua M, Hedden P, Genschik P. The cold-inducible CBF1 factor-dependent signaling pathway modulates the accumulation of the growth-repressing DELLA proteins via its effect on gibberellin metabolism. *Plant Cell*. 2008;20(8):2117–2129. <https://doi.org/10.1105/tpc.108.058941>.
- Allen MD, Yamasaki K, Ohme-Takagi M, Tateno M, Suzuki M. A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *Embo J*. 1998;17(18):5484–5496. <https://doi.org/10.1093/emboj/17.18.5484>.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
- An D, Ma Q, Yan W, Zhou W, Liu G, Zhang P. Divergent regulation of CBF regulon on cold tolerance and plant phenotype in cassava overexpressing *Arabidopsis* CBF3 gene. *Front Plant Sci*. 2016;7:1866. <https://doi.org/10.3389/fpls.2016.01866>.
- Baker CR, Stewart JJ, Amstutz CL, Ching LG, Johnson JD, Niyogi KK, Adams WW, Demmig Adams B. Genotype-dependent contribution of CBF transcription factors to long-term acclimation to high light and cool temperature. *Plant Cell Environ*. 2022;45(2):392–411. <https://doi.org/10.1111/pce.14231>.
- Benedict C, Skinner JS, Meng R, Chang Y, Bhalerao R, Huner NP, Finn CE, Chen TH, Hurry V. The CBF1-dependent low temperature signalling pathway, regulon and increase in freeze tolerance are conserved in *Populus* spp. *Plant Cell Environ*. 2006;29(7):1259–1272. <https://doi.org/10.1111/j.1365-3040.2006.01505.x>.
- Ding Y, Shi Y, Yang S. Molecular regulation of plant responses to environmental temperatures. *Mol Plant*. 2020;13(4):544–564. <https://doi.org/10.1016/j.molp.2020.02.004>.
- Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J*. 2003;33(4):751–763. <https://doi.org/10.1046/j.1365-313x.2003.01661.x>.
- Eddy SR. Accelerated profile HMM searches. *Plos Comput Biol*. 2011;7(10):e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
- Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20(1):238. <https://doi.org/10.1186/s13059-019-1832-y>.
- Feng W, Li J, Long S, Wei S. A DREB1 gene from zoysiagrass enhances *Arabidopsis* tolerance to temperature stresses without growth inhibition. *Plant Sci*. 2019;278:20–31. <https://doi.org/10.1016/j.plantsci.2018.10.009>.
- Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*. 1999;151(4):1531–1545. <https://doi.org/10.1093/genetics/151.4.1531>.
- Gabaldón T, Koonin EV. Functional and evolutionary implications of gene orthology. *Nat Rev Genet*. 2013;14(5):360–366. <https://doi.org/10.1038/nrg3456>.
- Gaut BS, Morton BR, McCaig BC, Clegg MT. 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 U S A*. 1996;93(19):10274–10279. <https://doi.org/10.1073/pnas.93.19.10274>.
- Gharib WH, Robinson-Rechavi M. When orthologs diverge between human and mouse. *Brief Bioinform*. 2011;12(5):436–441. <https://doi.org/10.1093/bib/bbr031>.
- Gilmour SJ, Fowler SG, Thomashow MF. *Arabidopsis* transcriptional activators CBF1, CBF2, and CBF3 have matching functional activities. *Plant Mol Biol*. 2004;54(5):767–781. <https://doi.org/10.1023/B:PLAN.0000040902.06881.d4>.
- Guo L, Wang S, Jiao X, Ye X, Deng D, Liu H, Li Y, Van de Peer Y, Wu W. Convergent and/or parallel evolution of RNA-binding proteins in angiosperms after polyploidization. *New Phytol*. 2024;242(3):1377–1393. <https://doi.org/10.1111/nph.19656>.
- Guo L, Wang S, Nie Y, Shen Y, Ye X, Wu W. Convergent evolution of AP2/ERF III and IX subfamilies through recurrent polyploidization and tandem duplication during eudicot adaptation to paleoenvironmental changes. *Plant Commun*. 2022;3(6):100420. <https://doi.org/10.1016/j.xplc.2022.100420>.
- Gutha LR, Reddy AR. Rice *DREB1B* promoter shows distinct stress-specific responses, and the overexpression of cDNA in tobacco

- confers improved abiotic and biotic stress tolerance. *Plant Mol Biol.* 2008;**68**(6):533–555. <https://doi.org/10.1007/s11103-008-9391-8>.
- Haake V, Cook D, Riechmann JL, Pineda O, Thomashow MF, Zhang JZ. Transcription factor CBF4 is a regulator of drought adaptation in *Arabidopsis*. *Plant Physiol.* 2002;**130**(2):639–648. <https://doi.org/10.1104/pp.006478>.
- Huang W, Zhang L, Columbus JT, Hu Y, Zhao Y, Tang L, Guo Z, Chen W, McKain M, Bartlett M, et al. A well-supported nuclear phylogeny of Poaceae and implications for the evolution of C₄ photosynthesis. *Mol Plant.* 2022;**15**(4):755–777. <https://doi.org/10.1016/j.molp.2022.01.015>.
- Ito Y, Katsura K, Maruyama K, Taji T, Kobayashi M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. Functional analysis of rice DREB1/CBF-type transcription factors involved in cold-responsive gene expression in transgenic rice. *Plant Cell Physiol.* 2006;**47**(1):141–153. <https://doi.org/10.1093/pcp/pci230>.
- Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF. *Arabidopsis* CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science.* 1998;**280**(5360):104–106. <https://doi.org/10.1126/science.280.5360.104>.
- Jia Y, Ding Y, Shi Y, Zhang X, Gong Z, Yang S. The *cbfs* triple mutants reveal the essential functions of CBFs in cold acclimation and allow the definition of CBF regulons in *Arabidopsis*. *New Phytol.* 2016;**212**(2):345–353. <https://doi.org/10.1111/nph.14088>.
- Jiang H, Shi Y, Liu J, Li Z, Fu D, Wu S, Li M, Yang Z, Shi Y, Lai J, et al. Natural polymorphism of ZmICE1 contributes to amino acid metabolism that impacts cold tolerance in maize. *Nat Plants.* 2022;**8**(10):1176–1190. <https://doi.org/10.1038/s41477-022-01254-3>.
- Jin R, Kim BH, Ji CY, Kim HS, Li HM, Ma DF, Kwak S. Overexpressing *lbCBF3* increases low temperature and drought stress tolerance in transgenic sweetpotato. *Plant Physiol Bioch.* 2017;**118**:45–54. <https://doi.org/10.1016/j.plaphy.2017.06.002>.
- Kasuga M, Miura S, Shinozaki K, Yamaguchi-Shinozaki K. A combination of the *Arabidopsis* DREB1A gene and stress-inducible *rd29A* promoter improved drought- and low-temperature stress tolerance in tobacco by gene transfer. *Plant Cell Physiol.* 2004;**45**(3):346–350. <https://doi.org/10.1093/pcp/pch037>.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;**30**(4):772–780. <https://doi.org/10.1093/molbev/mst010>.
- Kidokoro S, Shinozaki K, Yamaguchi-Shinozaki K. Transcriptional regulatory network of plant cold-stress responses. *Trends Plant Sci.* 2022;**27**(9):922–935. <https://doi.org/10.1016/j.tplants.2022.01.008>.
- Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;**12**(4):357–360. <https://doi.org/10.1038/nmeth.3317>.
- Koonin EV. Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet.* 2005;**39**(1):309. <https://doi.org/10.1146/annurev.genet.39.073003.114725>.
- Kumar S, Suleski M, Craig JM, Kasprowitz AE, Sanderford M, Li M, Stecher G, Hedges SB. TimeTree 5: an expanded resource for species divergence times. *Mol Biol Evol.* 2022;**39**(8):c174. <https://doi.org/10.1093/molbev/msac174>.
- Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* 2012;**22**(9):1813–1831. <https://doi.org/10.1101/gr.136184.111>.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;**9**(4):357–359. <https://doi.org/10.1038/nmeth.1923>.
- Laurent JM, Garge RK, Teufel AI, Wilke CO, Kachroo AH, Marcotte EM. Humanization of yeast genes with multiple human orthologs reveals functional divergence between paralogs. *Plos Biol.* 2020;**18**(5):e3000627. <https://doi.org/10.1371/journal.pbio.3000627>.
- Li C, Sun Y, Li J, Zhang T, Zhou F, Song Q, Liu Y, Brestic M, Chen TH, Yang X. ScCBF1 plays a stronger role in cold, salt and drought tolerance than StCBF1 in potato (*Solanum tuberosum*). *J Plant Physiol.* 2022;**278**:153806. <https://doi.org/10.1016/j.jplph.2022.153806>.
- Li Q, Brown JB, Huang H, Bickel PJ. Measuring reproducibility of high-throughput experiments. *Ann Appl Stat.* 2011;**5**(3):1752–1779. <https://doi.org/10.1214/11-AOAS466>.
- Liao B, Scott NM, Zhang J. Impacts of gene essentiality, expression pattern, and gene compactness on the evolutionary rate of mammalian proteins. *Mol Biol Evol.* 2006;**23**(11):2072–2080. <https://doi.org/10.1093/molbev/msl076>.
- Liao X, Guo X, Wang Q, Wang Y, Zhao D, Yao L, Wang S, Liu G, Li T. Overexpression of *MsDREB6.2* results in cytokinin-deficient developmental phenotypes and enhances drought tolerance in transgenic apple plants. *Plant J.* 2017;**89**(3):510–526. <https://doi.org/10.1111/tpj.13401>.
- Liu J, Shi Y, Yang S. Insights into the regulation of C-repeat binding factors in plant cold signaling. *J Integr Plant Biol.* 2018;**60**(9):780–795. <https://doi.org/10.1111/jipb.12657>.
- Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell.* 1998;**10**(8):1391–1406. <https://doi.org/10.1105/tpc.10.8.1391>.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;**15**(12):550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Magome H, Yamaguchi S, Hanada A, Kamiya Y, Oda K. dwarf and delayed-flowering 1, a novel *Arabidopsis* mutant deficient in gibberellin biosynthesis because of overexpression of a putative AP2 transcription factor. *Plant J.* 2004;**37**(5):720–729. <https://doi.org/10.1111/j.1365-313X.2003.01998.x>.
- Mao D, Chen C. Colinearity and similar expression pattern of rice *DREB1s* reveal their functional conservation in the cold-responsive pathway. *PLoS One.* 2012;**7**(10):e47275. <https://doi.org/10.1371/journal.pone.0047275>.
- Menon M, Barnes WJ, Olson MS. Population genetics of freeze tolerance among natural populations of *Populus balsamifera* across the growing season. *New Phytol.* 2015;**207**(3):710–722. <https://doi.org/10.1111/nph.13381>.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 2020;**37**(8):2461. <https://doi.org/10.1093/molbev/msaa131>.
- Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods.* 2022;**19**(6):679–682. <https://doi.org/10.1038/s41592-022-01488-1>.
- Nakano T, Suzuki K, Fujimura T, Shinshi H. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol.* 2006;**140**(2):411–432. <https://doi.org/10.1104/pp.105.073783>.
- Nie Y, Guo L, Cui F, Shen Y, Ye X, Deng D, Wang S, Zhu J, Wu W. Innovations and stepwise evolution of CBFs/DREB1s and their regulatory networks in angiosperms. *J Integr Plant Biol.* 2022;**64**(11):2111–2125. <https://doi.org/10.1111/jipb.13357>.
- O'Malley RC, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR. Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell.* 2016;**165**(5):1280–1292. <https://doi.org/10.1016/j.cell.2016.04.038>.
- Park S, Lee CM, Doherty CJ, Gilmour SJ, Kim Y, Thomashow MF. Regulation of the *Arabidopsis* CBF regulon by a complex low-temperature regulatory network. *Plant J.* 2015;**82**(2):193–207. <https://doi.org/10.1111/tpj.12796>.
- Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie

- and Ballgown. *Nat Protoc.* 2016;**11**(9):1650–1667. <https://doi.org/10.1038/nprot.2016.095>.
- Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, Zhang S, Paterson AH. Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. *Genome Biol.* 2019;**20**(1):38. <https://doi.org/10.1186/s13059-019-1650-2>.
- Reis RR, Andrade Dias Brito Da Cunha B, Martins PK, Martins MTB, Alekcevetch JC, Chalfun-Júnior A, Andrade AC, Ribeiro AP, Qin F, Mizoi J, et al. Induced over-expression of *AtDREB2A* CA improves drought tolerance in sugarcane. *Plant Sci.* 2014;**221**–**222**:59–68. <https://doi.org/10.1016/j.plantsci.2014.02.003>.
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol.* 2011;**29**(1):24–26. <https://doi.org/10.1038/nbt.1754>.
- Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K. DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem Biophys Res Co.* 2002;**290**(3):998–1009. <https://doi.org/10.1006/bbrc.2001.6299>.
- Schmitz RJ, Marand AP, Zhang X, Mosher RA, Turck F, Chen X, Axtell MJ, Zhong X, Brady SM, Megraw M, et al. Quality control and evaluation of plant epigenomics data. *Plant Cell.* 2022;**34**(1):503–513. <https://doi.org/10.1093/plcell/koab255>.
- Shi Y, Ding Y, Yang S. Molecular regulation of CBF signaling in cold acclimation. *Trends Plant Sci.* 2018;**23**(7):623–637. <https://doi.org/10.1016/j.tplants.2018.04.002>.
- Song Y, Zhang X, Li M, Yang H, Fu D, Lv J, Ding Y, Gong Z, Shi Y, Yang S. The direct targets of CBFs: in cold stress response and beyond. *J Integr Plant Biol.* 2021;**63**(11):1874–1887. <https://doi.org/10.1111/jipb.13161>.
- Stockinger EJ, Gilmour SJ, Thomashow MF. *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc Natl Acad Sci U S A.* 1997;**94**(3):1035–1040. <https://doi.org/10.1073/pnas.94.3.1035>.
- Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 2022;**50**(D1):D439–D444. <https://doi.org/10.1093/nar/gkab1061>.
- Wang Q, Guan Y, Wu Y, Chen H, Chen F, Chu C. Overexpression of a rice *OsDREB1F* gene increases salt, drought, and low temperature tolerance in both *Arabidopsis* and rice. *Plant Mol Biol.* 2008;**67**(6):589–602. <https://doi.org/10.1007/s11103-008-9340-6>.
- Wang S, Shen Y, Deng D, Guo L, Zhang Y, Nie Y, Du Y, Zhao X, Ye X, Huang J, et al. Orthogroup and phylotranscriptomic analyses identify transcription factors involved in the plant cold response: a case study of *Arabidopsis* BBX29. *Plant Commun.* 2023;**4**(6):100684. <https://doi.org/10.1016/j.xplc.2023.100684>.
- Wang S, Zhang Y, Ye X, Shen Y, Liu H, Zhao X, Guo L, Cao L, Du Y, Wu W. A phylotranscriptomic dataset of angiosperm species under cold stress. *Sci Data.* 2023;**10**(1):399. <https://doi.org/10.1038/s41597-023-02307-8>.
- Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, et al. MCSscanx: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012;**40**(7):e49. <https://doi.org/10.1093/nar/gkr1293>.
- Wei S, Li X, Lu Z, Zhang H, Ye X, Zhou Y, Li J, Yan Y, Pei H, Duan F, et al. A transcriptional regulator that boosts grain yields and shortens the growth duration of rice. *Science.* 2022;**377**(6604):i8455. <https://doi.org/10.1126/science.abi8455>.
- Welling A, Palva ET. Involvement of CBF transcription factors in winter hardiness in birch. *Plant Physiol.* 2008;**147**(3):1199–1211. <https://doi.org/10.1104/pp.108.117812>.
- Wu S, Han B, Jiao Y. Genetic contribution of paleopolyploidy to adaptive evolution in angiosperms. *Mol Plant.* 2020;**13**(1):59–71. <https://doi.org/10.1016/j.molp.2019.10.012>.
- Yan Y, Zhang D, Zhou P, Li B, Huang S. HDOCK: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res.* 2017;**45**(W1):W365–W373. <https://doi.org/10.1093/nar/gkx407>.
- Yu G, Wang L, He Q. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics.* 2015;**31**(14):2382–2383. <https://doi.org/10.1093/bioinformatics/btv145>.
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;**9**(9):R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
- Zhao C, Zhang Z, Xie S, Si T, Li Y, Zhu JK. Mutational evidence for the critical role of CBF transcription factors in cold acclimation in *Arabidopsis*. *Plant Physiol.* 2016;**171**(4):2744–2759. <https://doi.org/10.1104/pp.16.00533>.