RESEARCH



Unveiling the evolutionary and transcriptional landscape of ERF transcription factors in wheat genomes: a genome-wide comparative analysis



Liwen Wang^{2†}, Hongjun Zhao^{3†}, Runfang Li¹, Rumei Tian¹, Kaihua Jia¹, Yongchao Gong¹, Song Hou¹, Nana Li¹ and Yanyan Pu^{1*}

Abstract

Ethylene response factors (ERFs), belonging to the AP2/ERF superfamily, play vital roles in plant growth, development, and stress responses. The evolutionary and expression features of the members of the ERF gene family have not yet been extensively analyzed through comprehensive comparative genomics across various diploid, tetraploid, and hexaploid wheat genomes. In this study, we identified a total of 2,967 ERF genes across one diploid, two tetraploid, and five hexaploid wheat genomes using the characteristics of conserved domains of ERF proteins. Phylogenetic analysis revealed that ERF genes clustered into two main groups. Analyses of expansion of the ERF gene family indicated that the members of IIIc and IX (sub)groups were observed to show the expansion in tetraploid and hexaploid wheat compared to diploid wheat. Tandem duplication was identified as a key mechanism for ERF gene family expansion, with varying proportions across different wheat genomes. Ancient evolutionary evidence was traced using *Amborella trichopoda* as a reference, revealing the retention of gene copies in both tetraploid and hexaploid wheat. Then, we analyzed the expression of ERF genes, and reported the stress tolerant and disease resistant ERF genes in hexaploid wheat. These findings provide valuable insights into the evolutionary dynamics and functional features of ERF genes in wheat, paving the way for genetic breeding and molecular improvement of wheat species.

Keywords Genome-wide, Ethylene response factor, Triticum lineage, Tandem duplication, Expression divergence

[†]Liwen Wang and Hongjun Zhao contributed equally to this work.

*Correspondence: Yanyan Pu puyanyan12@126.com ¹Institute of Crop Germplasm Resources, Shandong Academy of Agricultural Sciences, Jinan 250100, China ²School of Life Sciences, Qilu Normal University, Jinan 250200, China ³Institute of industrial Crops, Shandong Academy of Agricultural Sciences, Jinan 250100, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Background

Wheat (Triticum aestivum L.), a cereal plant belonging to the Poaceae family and the genus Triticum, has been a staple food for humanity for millennia. Wheat is also rich in micronutrients and dietary fiber, containing minerals, vitamins, and fats [1, 2]. Over the course of long-term development, wheat has become one of the most widely distributed, extensively cultivated, second-highest vielding, most traded, and nutritionally valuable cereal crops worldwide. It is the result of human domestication of its wild ancestor during the Neolithic period, with a cultivation history spanning more than 10,000 years. Wheat originated in a small region of the Fertile Crescent and later spread to diverse environments around the world [3]. Common wheat is an allohexaploid species consisting of the A, B, and D subgenomes. It originated through two consecutive polyploidization events involving species from the Triticum and Aegilops genera, leading to the formation of tetraploid wheat (AABB) and hexaploid wheat (AABBDD) [4]. Allopolyploidization has enabled wheat to adapt to a wide range of growing environments. Wheat cultivation now spans from 67°N in Northern Europe to 45°S in southern Argentina, and from 150 m below sea level in the Turpan Basin to 4,100 m above sea level on the Qinghai-Tibet Plateau. Since 2013, the genomes of ancient wheat species T. urartu (AA), Aegilops tauschii (DD), and candidate species for the B genome ancestor have been assembled [5, 6]. Compared to other major cereal crops, the wheat genome is relatively large; the diploid emmer wheat is approximately 5 Gb, the tetraploid wild wheat exceeds 10 Gb, and the hexaploid common wheat is about 16 Gb. A series of reference-quality pseudomolecule genome assemblies have been published, including those for common wheat, durum wheat, diploid wheat ancestor species (T. urartu) [7], tetraploid ancestor species (T. turgidum ssp. dicoccoides) [8], and other allohexaploid cultivar such as Aikang58 [9], Alchemy, ArinaLrFor [10], Attraktion [11], Cadenza, CDC Landmark, CDC Stanley, Chinese Spring, Chuanmai104 [12], Claire, CS42, Fielder, Jagger, Julius, Kariega, KN9204, LongReach Lancer, Mace, Norin 61, Paragon, PI190962 (Spelt), Renan [13], Robigus, Sonmez, SY Mattis, Weebil and Zang1817 [14]. The genomic data resources of wheat with different ploidy offer researchers a crucial foundation for investigating the evolutionary mechanisms of gene families from a whole-genome perspective.

ERF (Ethylene Response Factor) is a specific type of transcription factor in plants, belonging to the AP2/ERF superfamily [15]. Members of the ERF family are widely involved in regulating plant growth and development, stress responses, and various signaling processes, especially in plant responses to biotic and abiotic stresses, such as drought, salinity, low temperature, and pathogen

infections [16]. ERF family members possess a highly conserved AP2/ERF domain, consisting of approximately 60-70 amino acids, responsible for specific binding to DNA [17]. Based on the cis-acting elements to which ERF transcription factors bind, they can be classified into two major categories [18]. The first category binds to the GCC box, previously known as the AGC box or PR box, with a conserved sequence of AGCCGCC, mainly found in the promoters of many basic PR genes, acting as an ethylene response element. The second category binds to DRE/CRT (DRE: dehydration-responsive element; CRT: C repeat), with a conserved sequence of CCGAC, present in many cold-responsive or drought-induced genes. The ERF family was first associated with the ethylene signaling pathway and is involved in regulating plant developmental processes and stress responses [19]. Ethylene is an important plant hormone, and ERF regulates the expression of ethylene-responsive genes by binding to specific cis-acting elements such as the GCC box [20]. During the plant response to environmental stresses such as drought, high salinity, and low temperature, ERF transcription factors activate or inhibit the expression of stress-responsive genes, thereby enhancing plant stress tolerance [16]. The wheat ERF transcription factor, TaERF3, was reported to promote significantly enhanced tolerance to salt and drought stresses in wheat [21]. In addition to participating in stress responses, ERF transcription factors are also related to plant growth and developmental processes, such as regulating flowering, fruit ripening, and leaf senescence. ERF transcription factors play a broad role in plant growth, development, and stress responses, making them of significant research and practical value [22].

Here, we conducted a comprehensive genome-wide comparative genomics analysis of the ERF gene family across eight wheat genomes. Through evolutionary analysis, we elucidated the formation of the ERF gene family and traced the evolutionary trajectories of ERF genes in the A (sub)genome of wheat. Additionally, we examined the impact of tandem duplication (TD) events on the generation of ERF genes by analyzing TD events genomewide. Furthermore, we explored the functional divergence of ERF genes under different salt stress conditions using transcriptomic profiling in T. aestivum (AABBDD, Chinese Spring). These findings offer novel insights into the evolutionary history and expression characteristics of the ERF gene family, thereby laying the groundwork for genetic breeding and molecular improvement of diploid, tetraploid, and hexaploid wheat species.

Methods

Data resources

Wheat diploid (AA), tetraploid (AABB), and hexaploidy (AABBDD) genomes including *T. urartu* (AA, G1812, GCA_003073215.2), *T. turgidum* ssp. *Durum* (durum

wheat) (AABB, Svevo, GCA_900231445.1), T. turgidum ssp. dicoccoides (wild emmer wheat) (AABB, Zavitan, GCF_002162155.2), T. aestivum (AABBDD, Chinese Spring, GCA_002220415.3), T. aestivum (AABBDD, Fielder, GCA_907166925.1), T. aestivum (AABBDD, Kariega, GCA_910594105.1), and T. aestivum (AABBDD, Renan, GCA_937894285.1) were retrieved from NCBI Genomes (https://www.ncbi.nlm.nih.gov/datasets/geno me/. The T. aestivum (AABBDD, KN9204) genome was retrieved from the Genome Warehouse (https://ngdc. cncb.ac.cn/gwh/) with the accession number: GWHB-JWI0000000 [23]. Arabidopsis genome was retrieved from TAIR (Araport11_blastsets) (https://www.arabidop sis.org/) [24]. A. trichopoda genome (AMTR1.0.57) was retrieved from Ensembl Plants (https://plants.ensembl.o rg) [25]. The profile Hidden Markov Models (HMMs) of AP2 domain (PF00847.25) and B3 DNA binding domain (PF02362.26) were retrieved Pfam 37.1 (23,794 entries, 751 clans) (http://pfam-legacy.xfam.org/) [26]. The RNA -seq expression data of T. aestivum (AABBDD, Chinese Spring) was retrieved from the Sequence Read Archive (SRA) database with accession numbers: PRJNA293629 [27].

Identification of ERF transcription factors

All the wheat genes with evidence of alternative splicing were identified and the longest proteins served as the representatives in eight wheat diploid, tetraploid, and hexaploidy genomes. ERF candidate proteins in eight wheat genomes were classified using HMMER (v3.2.1foss-2018b) program with "trusted cutoff" as threshold [28]. The MAFFT v7.520 performed multiple sequence alignments (MSA) with the highly conserved ERF candidate proteins [29]. Then, the MSA of ERF proteins were used to construct species-specific profile HMM of ERF transcription factors using the "hmmbuild" module in HMMER (v3.2.1-foss-2018b). With species-specific profile HMM of ERF transcription factors, the ERF proteins were classified within eight wheat diploid, tetraploid, and hexaploid genomes using HMMER (v3.2.1-foss-2018b). To ensure the accuracy of the identified ERF transcription factors, InterProScan was employed to confirm the conserved domains of ERF within eight wheat diploid, tetraploid, and hexaploidy genomes [30].

Phylogeny of ERF transcription factors

MAFFT v7.520 was employed to perform multiple sequence alignments (MSA) with ERF protein sequences within eight wheat genomes [29]. Then, IQ-Tree v2.3.1 with the parameters: -bb 1000 -redo -alrt 1000 -m MFP -nt AUTO was used to perform a phylogenetic analysis [31].

The identification of tandemly duplicated ERF genes

To identify tandemly duplicated ERF genes, the diamond v0.9.24.125 were employed to perform all-against-all BLAST of protein sequences with the parameters: -e 1e-5 --outfmt 6 --more-sensitive --threads 62–quiet, and then the paralogous gene pairs were identified within single wheat genome [32]. Combined with the information of gene location on chromosomes or assembled scaffolds in wheat genomes, the paralogous gene pairs anchored on closer location in the same genomic regions, tandemly duplicated genes were identified in eight wheat genomes, and one unrelated gene was allowed within a tandem array.

The identification of orthologous ERF gene pairs

Orthologous gene pairs between eight wheat genomes and *A. trichopoda* were detected using MCScanX software with the parameters (MATCH_SIZE=5 and E_ VALUE=1e-10). And the parameter "MATCH_SIZE" meant that five continuous orthologous gene pairs were detected from MCScanX analysis [33]. Then, the orthologous gene pairs that contained two ERF genes were obtained between *A. trichopoda* and single wheat genome. Further, the whole-genome duplicated ERF genes were identified in eight wheat diploid, tetraploid, and hexaploidy genomes.

Expression analysis of ERF genes in *Triticum aestivum* (AABBDD, Chinese spring)

Using the available RNA-seq data downloaded from SRA, the raw data of different samples were trimmed and cleaned using Trimmomatic v0.39 [34], and the cleaned short-reads were mapped to the hexaploid wheat Chinese Spring reference genome through HISAT (version 2.2.1) [35]. StringTie version 2.0.6 was used to calculate the fragments per kilobase of exon model per killion mapped fragments (FPKM) of mapped short reads to target genes or transcripts in different cotton genomes and FPKM values were normalized by log2 [36]. The significant expression differences of ERF genes under different conditions with the same time points were classified using the DESeq2 package with the parameters: |log(FC)| >1 and P-value = < 0.05 [37]. The expression profile of the ERF genes was implemented using the heatmap module from the R package.

Results

Identification of ERF genes in wheat genomes

The identification of diverse wheat genome ploidy has significantly advanced the analysis of gene families associated with key traits or phenotypes. While over 40 wheat genomes, encompassing various species, subspecies, and cultivars, have been released with assembled sequences, only a few possess complete genomic data suitable for



Fig. 1 Distribution and Phylogenetic tree of ERF genes in Arabidopsis and eight wheat genomes. (A) The distribution of AP2, ERF, and RAV genes in various wheat genomes; (B) phylogenetic tree with Maximum Likelihood was constructed by IQTree with 1000 replications

| Organism Name | Sample | Karyotype | Genome Size | Total | AP2 | ERF | RAV | Publication |
|------------------------------------|---------|-----------|----------------|---------|-------|-------|-------|---------------------------|
| - | | | (bp) | Genes | genes | genes | genes | |
| Triticum urartu | G1812 | AA | 4,849,000,000 | 35,768 | 19 | 130 | 4 | Ling HQ et al., 2018 |
| Triticum turgidum ssp. Durum | Svevo | AABB | 9,964,000,000 | 63,993 | 16 | 217 | 5 | Maccaferri M et al., 2019 |
| Triticum turgidum ssp. Dicoccoides | Zavitan | AABB | 10,677,000,000 | 66,764 | 37 | 282 | 8 | Zhu T et al., 2019 |
| Triticum aestivum | Chinese | AABBDD | 14,566,502,436 | 103,785 | 45 | 466 | 14 | Shimizu K. K et al., 2021 |
| | Spring | | | | | | | |
| Triticum aestivum | Fielder | AABBDD | 14,702,880,414 | 120,708 | 48 | 487 | 15 | Sato K. et al., 2021 |
| Triticum aestivum | Kariega | AABBDD | 14,677,204,660 | 116,576 | 46 | 528 | 14 | Athiyannan N. et al., |
| | | | | | | | | 2022 |
| Triticum aestivum | KN9204 | AABBDD | 14,467,938,304 | 147,600 | 30 | 425 | 13 | Shi X. et al., 2022 |
| Triticum aestivum | Renan | AABBDD | 14,195,643,615 | 109,549 | 42 | 432 | 12 | Aury JM et al., 2022 |

 Table 1
 Statistics of ERF genes in various wheat genomes

comprehensive genome-wide analysis of the ERF gene family. For this study, eight wheat genomes were selected from NCBI Genomes, comprising one wild diploid (AA), two tetraploid (AABB), and five hexaploid (AABBDD) genomes, to classify the ERF genes within these genomes. The AP2/ERF superfamily in plants, characterized by the presence of AP2/ERF domains, is divided into three families: AP2, ERF, and RAV [19]. AP2 family proteins feature two repeated AP2/ERF domains, ERF family proteins contain a single AP2/ERF domain, and RAV family proteins include both a B3 domain and a single AP2/ERF domain [38]. Leveraging the conserved nature of ERF proteins in plants, a profile HMM of the AP2 domain was employed to identify members of the AP2/ERF superfamily across the eight wheat genomes. Subsequently, members of the AP2 and RAV families were excluded, and the ERF family members were classified (Fig. 1A). After meticulous manual curation, we identified a total of 2,967 ERF genes across the eight wheat genomes: 130 in T. urartu (AA, G1812), 217 in T. turgidum ssp. durum (durum wheat) (AABB, Svevo), 282 in T. turgidum ssp. dicoccoides (wild emmer wheat) (AABB, Zavitan), 466 in *T. aestivum* (AABBDD, Chinese Spring), 487 in *T. aestivum* (AABBDD, Fielder), 528 in *T. aestivum* (AABBDD, Kariega), 425 in *T. aestivum* (AABBDD, KN9204), and 432 in *T. aestivum* (AABBDD, Renan), which showed expansion of ERF gene family (122) in the Arabidopsis genome [15](Supplemental table S1, Table 1).

Phylogeny of ERF genes among various wheat genomes

The availability of Arabidopsis genomic data for over two decades has significantly enhanced its status as a key model plant, thereby facilitating in-depth research into ERF genes [39]. Toshitsugu N. et al. summarized the characteristics of the ERF gene family and identified 122 ERF genes in the previous Arabidopsis genome [15]. They further classified the entire ERF gene family in the Arabidopsis genome into 12 subgroups based on the characteristics of conserved domains of ERF proteins, ultimately distributing them into 'Groups I to IV', 'Groups V to X', and 'Groups VI-L and Xb-L'. With the updated Arabidopsis TAIR11 genome, a total of 124 ERF genes were identified, which is more than the previous study [24]. However, the classification of the newly identified Arabidopsis ERF genes followed the previous criteria. Using a dataset of 3,091 ERF protein sequences (2,967 from eight wheat species and 124 from Arabidopsis), we constructed a phylogenetic tree to investigate the phylogeny of all ERF genes in eight diploid, tetraploid, and hexaploid wheat genomes. All ERF genes were clustered into two main groups: 'Groups I to IV' and 'Groups V to X,' with 'Groups VI-L and Xb-L' including in the 'Groups V to X.' (Fig. 1B).

Based on the classification criteria for ERF genes in Arabidopsis, the 'Groups I to IV' group contains 1,282 ERF genes, representing 43.17% of the total ERF genes in the eight wheat genomes. Specifically, there are 133 in Group I (a and b), 164 in Group II (a, b, and c), 762 in Group III (a, b, c, d, and e), and 223 in Group IV (a, b, and unclassified ERF genes clustered with subgroups a and b, designated as subgroup c). The III group of the ERF gene family in tetraploid and hexaploid wheat genomes shows the expansion compared to diploid wheat genomes, particularly in the IIIc subgroup. Previous studies have shown that members of the IIIc subgroup

 Table 2 The groups of the members of the ERF gene family

Page 5 of 11

exhibit stress-responsive gene expression, such as under low-temperature, salt, and/or drought conditions [40], indicating enhanced stress response functions in tetraploid and hexaploid wheat species. The 'Groups V to X' contains 1,507 ERF genes, representing 50.79% of the total ERF genes in the eight wheat genomes. Specifically, there are 202 in Group V (a and b), 77 in Group VI, 245 in Group VII (a), 237 in Group VIII (a and b), 503 in Group IX (a, b, and c), and 243 in Group X (a, b, and c). The IX group of the ERF gene family in tetraploid and hexaploid wheat genomes also shows the expansion compared to diploid wheat genomes. Members of Group IX are often linked to defensive gene expressions in response to pathogen infection, and their expressions are differentially induced by defense-related phytohormones such as ethylene, jasmonate, and salicylic acid [41]. The remaining ERF genes are classified as unclassified ERF genes (Table 2).

Tandem duplication events in various wheat genomes

Tandem duplication is a significant mechanism in plants that leads to an increase in the number of gene family

| Group | | | Arabidopsis | AA AABB | | AABBDD | | | | | |
|----------------------|------|-------|-------------|---------|-------|---------|-----------------------|---------|---------|--------|-------|
| | | | | G1812 | Svevo | Zavitan | Chinese Spring | Fielder | Kariega | KN9204 | Renan |
| Groups I to IV | I | la | 2 | 2 | 0 | 4 | 0 | 0 | 0 | 6 | 0 |
| | | lb | 8 | 8 | 8 | 11 | 17 | 16 | 24 | 15 | 22 |
| | Ш | lla | 6 | 1 | 2 | 4 | 7 | 6 | 6 | 6 | 5 |
| | | llb | 7 | 5 | 9 | 7 | 9 | 15 | 15 | 13 | 14 |
| | | llc | 2 | 0 | 2 | 3 | 6 | 12 | 6 | 5 | 6 |
| | Ш | Illa | 3 | 2 | 4 | 4 | 6 | 6 | 6 | 5 | 5 |
| | | IIIb | 3 | 4 | 9 | 10 | 15 | 15 | 14 | 13 | 15 |
| | | lllc | 6 | 14 | 22 | 42 | 82 | 80 | 88 | 56 | 68 |
| | | llld | 4 | 7 | 13 | 13 | 21 | 20 | 20 | 18 | 19 |
| | | llle | 6 | 3 | 2 | 4 | 6 | 11 | 10 | 5 | 5 |
| | IV | IVa | 5 | 5 | 4 | 5 | 0 | 0 | 11 | 11 | 0 |
| | | IVb | 5 | 1 | 4 | 7 | 20 | 6 | 6 | 5 | 13 |
| | | IVc | 4 | 5 | 16 | 5 | 16 | 14 | 20 | 33 | 16 |
| Groups V to X | v | Va | 4 | 5 | 6 | 10 | 17 | 18 | 19 | 15 | 14 |
| | | Vb | 1 | 4 | 7 | 8 | 12 | 15 | 18 | 21 | 13 |
| | VI | VI | 8 | 5 | 9 | 12 | 14 | 19 | 8 | 4 | 6 |
| | VII | Vlla | 5 | 12 | 21 | 31 | 25 | 45 | 42 | 37 | 32 |
| | VIII | VIIIa | 8 | 4 | 13 | 19 | 35 | 30 | 28 | 27 | 29 |
| | | VIIIb | 7 | 6 | 5 | 5 | 6 | 7 | 6 | 7 | 10 |
| | IX | IXa | 3 | 3 | 6 | 6 | 6 | 6 | 18 | 9 | 0 |
| | | IXb | 6 | 4 | 6 | 8 | 3 | 15 | 3 | 12 | 11 |
| | | IXc | 8 | 19 | 24 | 35 | 63 | 64 | 84 | 50 | 48 |
| | Х | Ха | 6 | 1 | 9 | 8 | 13 | 17 | 23 | 10 | 14 |
| | | Xb | 1 | 8 | 8 | 15 | 23 | 21 | 27 | 25 | 21 |
| | | Хс | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Groups VI-L and Xb-L | VI-L | | 3 | 1 | 3 | 4 | 3 | 5 | 0 | 3 | 16 |
| | Xb-L | | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Unclassified | | | 0 | 1 | 4 | 2 | 41 | 24 | 26 | 14 | 30 |
| Total | | | 124 | 130 | 217 | 282 | 466 | 487 | 528 | 425 | 432 |

Table 3Tandem duplication events of the members of ERFgene family

| Karyotype | Cultivar | ERF No. TD Percer genes genes (TD genes) %) | | Percentage (TD genes %) | No. tan- dem |
|------------|-------------------|---|-----|-------------------------------|--------------------|
| | | | | | array |
| Diploid | G1812 | 130 | 23 | 17.69 | 11 |
| Tetraploid | Svevo | 217 | 25 | 11.52 | 12 |
| | Zavitan | 282 | 71 | 25.18 | 29 |
| Hexaploid | Chinese Spring | 466 | 117 | 25.11 | 45 |
| | Fielder | 487 | 137 | 28.13 | 59 |
| | Kariega | 528 | 145 | 27.46 | 54 |
| | KN9204 | 425 | 80 | 18.82 | 35 |
| | Renan | 432 | 112 | 25.93 | 50 |

members, thereby expanding the gene family; this expansion is often accompanied by an enhancement of the functions associated with the gene family. Here, we identified tandemly duplicated ERF genes to provide evidence for the expansion of the ERF gene family across eight wheat genomes. By combining paralogous gene pairs and their genomic locations, we identified a total of 23, 25, 71, 117, 137, 145, 80, and 112 tandemly duplicated genes in wild diploid wheat G1812, tetraploid wheat Svevo and Zavitan, and hexaploid wheat Chinese Spring, Fielder, Kariega, KN9204, and Renan, respectively. These represent 17.69%, 11.52%, 25.18%, 25.11%, 28.13%, 27.46%, 18.82%, and 25.93% of the total protein-coding genes in the corresponding wheat genomes (Table 3). Among the eight wheat genomes, tetraploid wheat Svevo had the lowest proportion, indicating the weakest influence of tandem duplication on the ERF gene family. Among the cultivated hexaploid wheat genomes, wheat cultivar Fielder had the highest proportion, while cultivar KN9204 had the lowest, suggesting that Fielder may have experienced the strongest influence of tandem duplication events on the ERF gene family. The 137 tandemly duplicated ERF genes in wheat cultivar Fielder were distributed into 59 tandem arrays ranging from 2 to 6 genes, while the 80 tandemly duplicated ERF genes in cultivar KN9204 were distributed into 35 tandem arrays ranging from 2 to 4 genes (Supplemental Table S2). Interestingly, both wheat cultivars Kariega and Renan had one eightgene tandem array, despite having lower proportions than wheat cultivar Fielder.

Evolution of ERF genes in wheat genomes

Hexaploid wheat, commonly known as bread wheat (*T. aestivum*), originated through two major allopolyploidization events. The first event occurred around 800,000 years ago, leading to the formation of tetraploid wild emmer wheat (*T. turgidum* ssp. *dicoccoides*), which involved the hybridization of *T. urartu* (providing the A subgenome) and a species related to *Aegilops speltoides* (providing the B subgenome). The second event, which resulted in hexaploid wheat, involved the hybridization of tetraploid wild emmer wheat with *A. tauschii* (providing the D subgenome) around 8,500 to 9,000 years ago [42] (Fig. 2). In the first event, the wild emmer wheat (Zavitan) genome received 142 of its 282 ERF genes from the



Fig. 2 Two major allopolyploidization events in genus Triticum and the distribution of ERF genes in the subgenomes in various wheat genomes

Table 4Summary of tandemly duplicated genes in differentwheat subgenomes

| Samples | Total | No. ERF genes in subgenomes | | | | | | |
|-------------------|--------------|-----------------------------|-----------------------|-----------------------|---------|--|--|--|
| | ERF genes | AA | BB | DD | Unknown | | | |
| G1812 | 130 | 130(TD:24, 18.46%) | / | / | / | | | |
| Svevo | 217 | 109(TD:14, 12.84%) | 108(TD:12, 11.11%) | / | / | | | |
| Zavitan | 282 | 142(TD:32, 22.54%) | 140(TD:39, 27.86%) | / | / | | | |
| Chinese Spring | 466 | 147(TD:31, 21.09%) | 153(TD:38, 24.84%) | 164(TD:49, 29.88%) | 2 | | | |
| Fielder | 487 | 159(TD:38, 23.9%) | 160(TD:51, 31.88%) | 168(TD:51, 30.36%) | / | | | |
| Kariega | 528 | 170(TD:44, 25.88%) | 181(TD:59, 32.6%) | 177(TD:49, 27.68%) | / | | | |
| KN9204 | 425 | 139(TD:55, 39.57%) | 143(TD:60, 41.96%) | 139(TD:54, 38.85%) | 4 | | | |
| Renan | 432 | 150(TD:47, 31.33%) | 130(TD:31, 23.85%) | 152(TD:37, 24.34%) | / | | | |

A subgenome, indicating an expansion of the ERF gene family compared to the 130 ERF genes found in wild diploid wheat G1812 (Table 4). In contrast, the durum wheat (Svevo) genome, which contains 217 ERF genes, received 109 from the A subgenome, showing a contraction of the ERF gene family compared to the total number of ERF genes in wild diploid wheat (AA) G1812. In the second event, the numbers of ERF genes donated from the A and B subgenomes in three hexaploid wheat genomes (Chinese Spring, Fielder, and Kariega) increased compared to those in the A and B subgenomes of the two tetraploid wheat genomes and the A genome of wild diploid wheat (Fig. 2). However, the number of ERF genes in the A subgenome of hexaploid wheat KN9204 and the B subgenome of hexaploid wheat Renan was lower than that in wild emmer wheat. Further, we detected the tandemly duplicated ERF genes in all kinds of subgenomes among various wheat ploidy and found that more than 20% of ERF genes in A, B, and D subgenomes were generated via tandem duplication events. Interestingly, although the numbers of ERF genes in A and B subgenomes of hexaploid wheat Chinese Spring showed higher but had a lower proportion than those in tetraploid wild emmer wheat.

Ancient evolutionary evidence in various wheat species

To trace the ancient evolutionary evidence of ERF genes in wheat genomes, A. trichopoda, recognized as the most basal extant flowering plant, was used to identify orthologous ERF pairs between A. trichopoda and eight wheat genomes separately [43]. The results showed that a total of 0, 2, 8, 15, 16, 6, 19, and 11 ERF genes were identified in wild diploid wheat G1812, tetraploid wheat Svevo and Zavitan, and hexaploid wheat Chinese Spring, Fielder, Kariega, KN9204, and Renan, respectively. These represent 0%, 0.92%, 2.84%, 3.22%, 3.29%, 1.14%, 4.47%, and 2.55% of the total ERF genes in the respective wheat genomes, with 0, 1, 5, 9, 8, 3, 9, and 6 orthologous ERF genes identified in the A. trichopoda genome (Table 5). These findings suggest that ancient ERF genes in the A. trichopoda genome are retained as two- or threegene copies in tetraploid and hexaploid wheat genomes. Regrettably, no orthologous ERF pairs were detected between A. trichopoda and wild diploid wheat G1812. However, this does not imply a loss of ancient evolutionary evidence; rather, it may be due to the stringent parameters used during the collinear analysis with MCScanX [33]. Except, the analyses of Ka/Ks ratios between orthologous gene pairs revealed that the four orthologous gene pairs were under purifying selection. but there are no significant differences between the two of them.

Furthermore, the retention of ERF genes from wild diploid wheat G1812 (AA) in tetraploid and hexaploid wheat species was assessed through collinear analyses between G1812 and the A subgenomes of polyploid wheat species. In tetraploid genomes, a total of 69 and 116 orthologous ERF genes were detected in tetraploid wheat Svevo and Zavitan, respectively, representing 48.46% and 71.54% of the total ERF genes in the diploid wheat G1812 genome (Table 5). This indicates that the A subgenome in tetraploid wheat Zavitan retained more collinear ERF genes from the diploid wheat genome, suggesting a closer

Table 5 The orthologous genes of the members of ERF gene family

| Tuble by the orthologous genes of the members of Entry gene furnity | | | | | | | | |
|---|----------------|-----------|----------------|-------------------------|----------------|-------------------------|--|--|
| Karyotype | Cultivar | ERF genes | No. orthologs1 | Percentage (orthologs1) | No. orthologs2 | Percentage (orthologs2) | | |
| Diploid (AA) | G1812 | 130 | 0 | 0 | / | / | | |
| Tetraploid (AABB) | Svevo | 217 | 2 | 0.92 | 63 | 48.46 | | |
| | Zavitan | 282 | 8 | 2.84 | 93 | 71.54 | | |
| Hexaploid (AABBDD) | Chinese Spring | 466 | 15 | 3.22 | 88 | 67.69 | | |
| | Fielder | 487 | 16 | 3.29 | 93 | 71.54 | | |
| | Kariega | 528 | 6 | 1.14 | 39 | 30 | | |
| | KN9204 | 425 | 19 | 4.47 | 85 | 65.38 | | |
| | Renan | 432 | 11 | 2.55 | 80 | 61.54 | | |

Note: orthologs 1 represents the orthologous gene pairs between A. trichopoda and wheat genomes; orthologs 2 represents the orthologous gene pairs between G1812 and A subgenome in wheat genomes

evolutionary relationship between the A subgenome of the tetraploid wheat Zavitan genome and diploid wild wheat. In hexaploid genomes, a total of 88, 93, 39, 85, and 80 orthologous ERF genes were detected in hexaploid wheat Chinese Spring, Fielder, Kariega, KN9204, and Renan, respectively, representing 67.69%, 71.54%, 30%, 65.38%, and 61.54% of the total ERF genes in the diploid wheat G1812 genome. The analyses of *K*a/*K*s ratios between orthologous gene pairs revealed that all orthologous gene pairs were under purifying selection. However, no significant differences were detected among the five types of orthologous gene pairs.

Expression analysis of ERF genes in *Triticum aestivum* (AABBDD, Chinese Spring)

To investigate the expression profile of wheat ERF genes under salt stress, we analyzed the expression of ERF genes in leaves of hexaploid wheat Chinese Spring (Triticum aestivum, AABBDD) under different treatments using an in-house pipeline. The treatments included control and salt stress conditions, each sampled at 6, 12, 24, and 48 h, resulting in a total of eight samples. Utilizing RNA-seq data from these samples, we conducted an expression analysis of protein-coding genes and identified 418 expressed ERF genes in the Chinese Spring genome (Supplemental Table S3). Among these, 319 ERF genes showed significant expression changes under salt stress compared to control conditions across the four time points, comprising 217 up-regulated and 102 downregulated genes (Fig. 3A). Further analysis revealed that 86 and 14 ERF genes were consistently up-regulated and down-regulated, respectively, across all four-time points under salt stress (Fig. 3B, C). The 86 up-regulated genes were distributed across the A, B, and D subgenomes (Supplemental Figure S1), while the 14 down-regulated genes were confined to chromosomes 4 and 5 of the A, B, and D subgenomes.

Phylogenetic analysis indicated that the IIIc subgroup of the ERF family in hexaploid wheat Chinese Spring has expanded through tandem duplication events. These genes are responsive to various stresses, including low temperature, salt, and drought (Fig. 3D). Among the 82 ERF genes in the IIIc subgroup, 13 were expressed at different time points, with seven showing up-regulation under salt stress compared to control conditions, while six exhibited the opposite expression pattern (Fig. 4A). Additionally, the IX subgroup, often associated with defense gene expression in response to pathogen infection and influenced by phytohormones such as ethylene and jasmonate, showed significant expression under salt stress. Out of 72 ERF genes in the IX subgroup, 19 were highly expressed under salt stress at each time point, which might indicate that salt stress can induce the expression of ERF genes in wheat Chinese Spring (Fig. 4B).

Discussion

Novel sequencing technology facilities the release of complete genomes and functional genes with key traits or phenotypes. According to the characteristics of conserved domains of ERF gene family, the previous Arabidopsis genome identified 122 ERF genes, but 124 ERF genes in the available TAIR11 genome. Out of 122 ERF genes in previous Arabidopsis genome, three Arabidopsis genes (AT1G63040, AT1G25470, and AT5G67000) were removed from ERF gene family; the gene AT1G63040 were lost in TRIR11 genome, and the genes AT1G25470 and AT5G67000 did not contain the conserved domains of ERF gene family. Five newly identified ERF genes in TAIR11 genome, including AT2G39250, AT2G41710, AT3G54990, AT5G10510, and AT5G60120. The classification of the newly identified Arabidopsis ERF genes followed the previous criteria of 122 ERF genes in the previous Arabidopsis genome (Supplemental figure S2). So, the gene AT2G41710 was clustered into the IVb subgroup, and the rest four Arabidopsis genes were clustered with the IVa and IVb subgroups, but separately. In this analysis, we named IVc subgroup in the ERF gene family in various wheat genomes.

The ERF gene family exhibits intricate evolutionary patterns such as wheat transitions from diploid to tetraploid and hexaploid forms. For instance, the A subgenome of tetraploid wheat Zavitan retains a substantial number of collinear ERF genes from diploid wheat G1812, indicating a close evolutionary relationship. In contrast, hexaploid wheat Fielder's A subgenome shows an even closer relationship with diploid wild wheat. These differences likely stem from distinct selective pressures, genetic recombination events, and environmental adaptations experienced by each variety. The formation of hexaploid wheat through two allopolyploidization events also impac the ERF gene family. The first event led to an expansion of ERF genes in wild emmer wheat, while the second event saw an increase in ERF genes in hexaploid wheat compared to tetraploid wheat. However, some hexaploid varieties like KN9204 and Renan show a reduction in ERF genes in certain subgenomes, possibly due to gene loss or selective retention. Functionally, the IIIc subgroup of the ERF family, which has expanded through tandem duplication in hexaploid wheat, is responsive to multiple stresses, suggesting a role in enhancing stress tolerance. Similarly, the IX subgroup, associated with defense responses and regulated by phytohormones, shows significant expression under salt stress, indicating that both biotic and abiotic stresses can induce ERF gene expression. Understanding these evolutionary and



Fig. 3 Statistics of the expressed ERF genes under salt stress. (A) Statistics of the differentially expressed ERF genes in different time points; (B) Venn graphics of up-regulated ERF genes in different time points; (C) Venn graphics of down-regulated ERF genes in different time points; (D) The phylogeny of the members of IIIc and IX (sub)group in wheat Chinese Spring genome

functional dynamics can provide valuable insights for breeding wheat varieties with improved stress resilience.

This analysis explored that TD events played an important role in the expansion of the ERF gene family in various wheat genomes. The ERF gene family in wheat Zavitan and Fielder were strongly influenced by the TD events compared to other wheat species with identical ploidy in tetraploid and hexaploid wheat genomes respectively. The impact of TD events on the ERF gene family in wheat was further investigated, revealing that the effects of tandem duplication events on the B subgenomes of tetraploid (Zavitan) and hexaploid (Fielder) wheat genomes, as well as the DD subgenome in the hexaploid wheat Fielder genome, were more pronounced than those on the AA subgenomes, regardless of whether the wheat genomes were tetraploid or hexaploid. Furthermore, the retention of ERF genes in the AA subgenomes in tetraploid and hexaploidy wheat genomes was detected via collinear analyses with the diploid wheat G1812 genome. The ERF genes in the AA subgenomes in tetraploid wheat Zavitan and hexaploid wheat Fielder genomes retained the most ancient ERF genes from diploid wheat genome respectively. These results illustrated the independent evolution of tetraploid and hexaploid wheat species after species formation in the two major stages of wheat allopolyploidization events.



Fig. 4 Expression profile of the expressed ERF genes under salt stress. (A) Expression heatmap of ERF genes from IIIc subgroup of ERF gene family; (B) Expression heatmap of ERF genes from IX group of ERF gene family

Conclusions

Wheat is one of the most vital food crops, supplying approximately one-fifth of the calories and proteins needed to nourish the ever-growing global population. Using profile HMM of the AP2/ERF domain, a total of 2,967 ERF genes were identified across eight wheat genomes, including one wild diploid, two tetraploid, and five hexaploid genomes. Phylogenetic analysis revealed that ERF genes clustered into two groups, inconsistent with previous classifications in Arabidopsis. The IIIc and IX subgroups showed the expansion in tetraploid and hexaploid wheat compared to diploid wheat, indicating enhanced stress response and defense mechanisms. Tandem duplication was identified as a key mechanism for ERF gene family expansion, with varying proportions across different wheat genomes. Ancient evolutionary evidence has been utilized to trace the evolutionary history of wheat, with A. trichopoda serving as a reference, revealing retained gene copies in both tetraploid and hexaploid wheat. The retention of ERF genes from wild diploid wheat G1812 in tetraploid and hexaploid wheat species was detected through collinear analysis. The A subgenome in tetraploid wheat Zavitan and hexaploid wheat Fielder showed closer evolutionary relationships with wild diploid wheat, while the A subgenome in Kariega exhibited the opposite pattern. The study also examined the evolution of ERF genes through allopolyploidization events, revealing expansion and contraction patterns in different subgenomes. The expression profiles of ERF genes in hexaploid wheat Chinese Spring were analyzed under salt stress, revealing the consistently up- and down-regulated across all time points. The IIIc subgroup responded to multiple stresses, while group IX ERF genes showed high expression under salt stress, indicating their involvement in both biotic and abiotic stress responses. These findings provide insights into the evolutionary dynamics and functional roles of ERF genes in wheat.

Supplementary Information

The online version contains supplementary material available at https://doi.or g/10.1186/s12864-025-11671-1.

Supplementary Material 1 Supplementary Material 2

Acknowledgements

Not applicable.

Author contributions

YP and LW conceived the project. LW and HZ analyzed the data and prepared the manuscript. RL and RT performed experiments and analyzed the data. KJ performed phylogenetic analysis and revised the draft manuscript. YG collected genomic data, performed comparative analysis of ERF genes and revised the draft manuscript. SH performed the tandem duplication analysis of ERF genes and revised the draft manuscript. NL and YP revised the manuscript. All authors read and approved of the final manuscript.

Funding

This work was supported by the Key R&D Program of Shandong Province, China (2023LZGC001 and 2024LZGC001), Shandong Province Science and Technology-based Small and Medium Enterprises Innovation Capability Enhancement Project (2023TSGC0360), and 2024 Jinan City Agricultural Science and Technology Research Project (GG202403).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 24 February 2025 / Accepted: 2 May 2025 Published online: 19 May 2025

References

- Shiferaw B, Smale M, Braun H-J, Duveiller E, Reynolds M, Muricho G. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Secur. 2013;5:291–317.
- Lafiandra D, Riccardi G, Shewry PR. Improving cereal grain carbohydrates for diet and health. J Cereal Sci. 2014;59:312–26.
- Salamini F, Özkan H, Brandolini A, Schäfer-Pregl R, Martin W. Genetics and geography of wild cereal domestication in the near East. Nat Rev Genet. 2002;3:429–41.
- 4. IWGSC. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. Science. 2014;345:1251788.
- Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, et al. Aegilops tauschii draft genome sequence reveals a gene repertoire for wheat adaptation. Nature. 2013;496:91–5.
- Luo M-C, Gu YQ, You FM, Deal KR, Ma Y, Hu Y, et al. A 4-gigabase physical map unlocks the structure and evolution of the complex genome of Aegilops Tauschii, the wheat D-genome progenitor. Proc Natl Acad Sci. 2013;110:7940–5.
- Ling H-Q, Ma B, Shi X, Liu H, Dong L, Sun H, et al. Genome sequence of the progenitor of wheat A subgenome Triticum urartu. Nature. 2018;557:424–8.
- Avni R, Nave M, Barad O, Baruch K, Twardziok SO, Gundlach H, et al. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science. 2017;357:93–7.
- Wang Z, Zhao G, Yang Q, Gao L, Liu C, Ru Z, et al. Helitron and CACTA DNA transposons actively reshape the common wheat - AK58 genome. Genomics. 2022;114:110288.
- Shimizu KK, Copetti D, Okada M, Wicker T, Tameshige T, Hatakeyama M, et al. De Novo genome assembly of the Japanese wheat cultivar Norin 61 highlights functional variation in flowering time and Fusarium-Resistant genes in East Asian genotypes. Plant Cell Physiol. 2021;62:8–27.
- 11. Kale SM, Schulthess AW, Padmarasu S, Boeven PHG, Schacht J, Himmelbach A, et al. A catalogue of resistance gene homologs and a chromosome-scale reference sequence support resistance gene mapping in winter wheat. Plant Biotechnol J. 2022;20:1730–42.
- 12. Liu Z, Yang F, Deng C, Wan H, Tang H, Feng J, et al. Chromosome-level assembly of the synthetic hexaploid wheat-derived cultivar chuanmai 104. Sci Data. 2024;11:670.
- Aury J-M, Engelen S, Istace B, Monat C, Lasserre-Zuber P, Belser C et al. Longread and chromosome-scale assembly of the hexaploid wheat genome achieves high resolution for research and breeding. Gigascience. 2022;11.
- 14. Guo W, Xin M, Wang Z, Yao Y, Hu Z, Song W, et al. Origin and adaptation to high altitude of Tibetan semi-wild wheat. Nat Commun. 2020;11:5085.
- 15. Nakano T, Suzuki K, Fujimura T, Shinshi H. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. Plant Physiol. 2006;140:411–32.
- Xu Z-S, Chen M, Li L-C, Ma Y-Z. Functions of the ERF transcription factor family in plants. Botany. 2008;86:969–77.
- Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. 2017;45:D1040–5.

- Nole-Wilson S, Krizek BA. DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. Nucleic Acids Res. 2000;28:4076–82.
- Wang K, Guo H, Yin Y. AP2/ERF transcription factors and their functions in Arabidopsis responses to abiotic stresses. Environ Exp Bot. 2024;222:105763.
- Wang X, Wen H, Suprun A, Zhu H. Ethylene signaling in regulating plant growth, development, and stress responses. Plants. 2025;14.
- Rong W, Qi L, Wang A, Ye X, Du L, Liang H, et al. The ERF transcription factor TaERF3 promotes tolerance to salt and drought stresses in wheat. Plant Biotechnol J. 2014;12:468–79.
- 22. Ma Z, Hu L, Jiang W. Understanding AP2/ERF transcription factor responses and tolerance to various abiotic stresses in plants: A comprehensive review. Int J Mol Sci. 2024;25.
- Shi X, Cui F, Han X, He Y, Zhao L, Zhang N, et al. Comparative genomic and transcriptomic analyses uncover the molecular basis of high nitrogen-use efficiency in the wheat cultivar Kenong 9204. Mol Plant. 2022;15:1440–56.
- Cheng C-Y, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD. Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. Plant J. 2017;89:789–804.
- Bolser D, Staines DM, Pritchard E, Kersey P. Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomics data. Methods Mol Biol. 2016;1374:115–40.
- Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: the protein families database in 2021. Nucleic Acids Res. 2021;49:D412–9.
- Taneja M, Upadhyay SK. Molecular characterization and differential expression suggested diverse functions of P-type II Ca(2+)ATPases in Triticum aestivum L. BMC Genomics. 2018;19:389.
- Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. Nucleic Acids Res. 2018;46:W200–4.
- Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 2019;20:1160–6.
- Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterPro-Scan 5: genome-scale protein function classification. Bioinformatics. 2014;30:1236–40.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37:1530–4.
- 32. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12:59–60.
- Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40:e49.
- 34. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 2014;30:2114–20.
- Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12:357–60.
- Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33:290–5.
- Love MI, Huber W, Anders S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
- Choi J-W, Choi HH, Park Y-S, Jang M-J, Kim S. Comparative and expression analyses of AP2/ERF genes reveal copy number expansion and potential functions of ERF genes in Solanaceae. BMC Plant Biol. 2023;23:48.
- Reiser L, Bakker E, Subramaniam S, Chen X, Sawant S, Khosa K et al. The Arabidopsis information resource in 2024. Genetics. 2024;227.
- Xu Y, Jiang J, Zeng L, Liu H, Jin Q, Zhou P, et al. Genome-wide identification and analysis of ERF transcription factors related to abiotic stress responses in Nelumbo nucifera. BMC Plant Biol. 2024;24:1057.
- Ma N, Sun P, Li Z-Y, Zhang F-J, Wang X-F, You C-X, et al. Plant disease resistance outputs regulated by AP2/ERF transcription factor family. Stress Biol. 2024;4:2.
- 42. Wang Z, Wang W, He Y, Xie X, Yang Z, Zhang X, et al. On the evolution and genetic diversity of the bread wheat D genome. Mol Plant. 2024;17:1672–86.
- 43. Amborella Genome Project. The Amborella genome and the evolution of flowering plants. Science. 2013;342:1241089.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.