



OPEN

Genome-wide identification and evolutionary analysis of the *AP2/EREBP*, *COX* and *LTP* genes in *Zea mays* L. under drought stress

Amaal Maghraby¹✉ & Mohamed Alzalaty²

AP2 (APETALA2)/EREBP (ethylene-responsive element-binding protein), cytochrome c oxidase (*COX*) and nonspecific lipid transfer proteins (*LTP*) play important roles in the response to drought stress. This is the first study to identify the *COX* gene in *Zea mays* L. via genome-wide analysis. The qRT-PCR results indicated that *AP2/EREBP*, *COX* and *LTP* were downregulated, with fold changes of 0.84, 0.53 and 0.31, respectively, after 12 h of drought stress. Genome-wide analysis identified 78 *AP2/EREBP*, 6 *COX* and 10 *LTP* genes in *Z. mays* L. Domain analysis confirmed the presence of the AP2 domain, Cyt_c_Oxidase_Vb domain and nsLTP1 in the *AP2/EREBP*, *COX* and *LTP* proteins, respectively. The *AP2/EREBP* protein family (AP2) includes five different domain types: the AP2/ERF domain, the EREBP-like factor (EREBP), the ethylene responsive factor (ERF), the dehydration responsive element binding protein (DREB) and the SHN SHINE. Synteny analysis of the *AP2/EREBP*, *COX* and *LTP* genes revealed collinearity orthologous relationships in *O. sativa*, *H. vulgare* and *A. thaliana*. *AP2/EREBP* genes were found on the 10 chromosomes of *Z. mays* L. *COX* genes were found on chromosomes 1, 3, 4, 5, 7 and 8. *LTP* genes were found on chromosomes 1, 3, 6, 8, 9 and 10. In the present study, the *Ka/Ks* ratios of the *AP2/EREBP* paralogous pairs indicated that the *AP2/EREBP* genes were influenced primarily by purifying selection, which indicated that the *AP2/EREBP* genes received strong environmental pressure during evolution. The *Ka/Ks* ratios of the *COX-3/COX-4* paralogous pairs indicate that the *COX-3/COX-4* genes were influenced primarily by Darwinian selection (driving change). For the *LTP* genes, the *Ka/Ks* ratios of the *LTP-1/LTP-10*, *LTP-5/LTP-3* and *LTP-4/LTP-8* paralogous pairs indicate that these genes were influenced primarily by purifying selection, while the *Ka/Ks* ratios of the *LTP-2/LTP-6* paralogous pairs indicate that these genes were influenced primarily by Darwinian selection. The duplication time of the *AP2/EREBP* paralogous gene pairs in *Z. mays* L. ranged from approximately 9.364 to 100.935 Mya. The duplication time of the *COX-3/COX-4* paralogous gene pair was approximately 5.217 Mya. The duplication time of the *LTP* paralogous gene pairs ranged from approximately 19.064 to 96.477 Mya. The major focus of research is to identify the genes that are responsible for drought stress tolerance to improve maize for drought stress tolerance. The results of the present study will improve the understanding of the functions of the *AP2/EREBP*, *COX* and *LTP* genes in response to drought stress.

Keywords *AP2/EREBP*, *COX*, *LTP*, Genome-wide identification, Evolutionary analysis, Drought stress

Abbreviations

AP2/EREBP *APETALA2*/Ethylene-responsive element-binding protein
COX Cytochrome c oxidase
LTP Nonspecific lipid transfer protein

¹Botany and Microbiology Department, Faculty of Science, Cairo University, Giza, Egypt. ²Department of Plant Genetic Transformation, Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center (ARC), Giza, Egypt. ✉email: amaalmaghraby@cu.edu.eg

Ka/Ks Ratio of nonsynonymous/synonymous
 RMSD Root mean square deviation

Drought is defined as dryness (lack of water, moisture deficit, shortage of precipitation) for a period of time that affects organisms (plants, animals, humans) in the affected area^{1–4}. Drought is the most destructive type of hydrological hazard⁵. The development of drought-tolerant seed varieties can help farmers produce crops under drought stress. In maize (*Z. mays* L.), drought stress is one of the major environmental stress effects on yield reduction, and drought can affect maize at any stage of development. Breeding programmes are used to improve the drought tolerance of maize hybrids. Molecular biology techniques were used to improve breeding efficiency by identifying genes related to drought stress⁶. Climate changes such as water scarcity have negative effects on plant growth and yield production⁷. Crop plants are exposed to several types of environmental stress, which affects their growth and development throughout their life cycle. Drought activates gene expression pathways in plants to protect cells against water deficit⁸. *APETALA2* (*AP2*)/ethylene-responsive element-binding protein (*EREBP*) is a transcription factor that has essential regulatory functions for protecting plants during stress. *AP2/EREBP* is one of the largest transcription factor families in plants^{9–11}. *AP2s/EREBPs* are involved in regulating gene expression during abiotic stress¹² and plant growth and development^{11,13}. The *AP2/EREBP* protein family is classified into: the *APETALA2* (*AP2*), ethylene-responsive factor (*ERF*), dehydration-responsive element binding protein (*DREB*), and related to *ABI3/VP1* (*RAV*). In cotton (*Gossypium raimondii*), the genes are distributed on all chromosomes. In *Gossypium hirsutum*, the *ERF* and *DREB* genes play important roles in stress responses¹¹. GhERF12 protein play crucial roles in organ development and differentiation in *G. hirsutum*¹⁴. Cytochrome c oxidase (*COX*) catalyzes the transfer of electrons from reduced cytochrome c (*CYTc*) to the final acceptor O_2 to H_2O for ATP production^{15,16}. Restriction mapping and DNA sequencing were used to study the functional relationships of the *COX* mitochondrial genes. The results did not reveal a transcript of *cox1*; rather, a reduced level of a *cox2* transcript and two different *cox3* transcripts were detected. Results indicate that genomic rearrangements of the both 5' and 3' flanking regions of the *cox1* gene leads to impaired of *cox1* transcription¹⁷. The nonspecific lipid transfer protein (*LTPs*) are involved in different biological processes and play key roles in plant¹⁸. *LTP2* play critical role for abiotic¹⁹ and biotic stresses tolerance in plants²⁰. *LTPs* are a large protein family present in all plants and are expressed in many different tissues. *LTPs* play important roles in signaling via their structure, which contains a N-terminal signal peptide that delivers proteins to the plasma membrane²¹. In maize, *LTPs* are differentially regulated by drought and salt treatments²². In rice, the expression of *LTP* is strongly induced under drought and salinity stresses²³. In wheat, *TaLTP1.2* and *TaLTP1.13* are upregulated during drought²⁴. In *Lotus japonica*, The *LjLTP* genes are expressed in aerial tissues under drought stress²⁵. In the Moss *Physcomitrium* (*Physcomitrella*) patens, the expression of 8 *LTPgs* was investigated during several abiotic stresses. Three *LTPg* genes are significantly upregulated, which leads to the downregulation of the PpLTPg genes²⁶.

Methods

Identification of the *AP2/EREBP*, *COX* and *LTP* genes in *Z. mays* L.

The genomes of *Z. mays* L., *Oryza sativa*, *Hordeum vulgare* and *Arabidopsis thaliana* were downloaded from the Phytozome database²⁷. The *AP2/EREBP* (accession number: NP_001183842.1), *COX* (accession number: NP_001288395.1) and *LTP* (accession number: ABA33850.1) proteins were used as query proteins from the NCBI database (<https://www.ncbi.nlm.nih.gov/>)²⁸ (Sheet 1 Online Resource SI 1) to screen *AP2/EREBP*, *COX* and *LTP* protein members in the genomes of *Z. mays* L. from the Phytozome database (<https://phytozome.jgi.doe.gov/>)²⁷ with an E-value $\leq 1e^{-30}$ and $\geq 50\%$ identity for *AP2/EREBP* proteins (<https://phytozome-next.jgi.doe.gov/blast-results/694644>), whereas the Phytozome database parameters for *COX* proteins (<https://phytozome-next.jgi.doe.gov/blast-results/694640>) and *LTP* proteins (<https://phytozome-next.jgi.doe.gov/blast-results/694645>) had an E-value $\leq 1e^{-30}$.

Characterization of the *AP2/EREBP*, *COX* and *LTP* proteins in *Z. mays* L.

Circoletto (<http://tools.bat.infospire.org/circoletto/>)²⁹ visualized the sequence identity of the *AP2/EREBP*, *COX* and *LTP* proteins. The physical and chemical properties of the *AP2/EREBP*, *COX* and *LTP* proteins, including the molecular weight, isoelectric point, total number of negatively charged residues, total number of atoms, instability and grand average hydropathicity (GRAVY), were computed using the ExPASy ProtParam Tool³⁰.

Phylogenetic, chromosomal distribution, evolutionary analysis and synteny analysis of the *AP2/EREBP*, *COX* and *LTP* genes

Multiple sequence alignments of the *AP2/EREBP*, *COX* and *LTP* proteins from *Z. mays* L. were performed via the MUSCLE method. Molecular evolutionary genetic analysis (MEGA-11)³¹ was subsequently conducted on a phylogenetic tree with a maximum likelihood of 1000 bootstrap replicates based on the WAG with Freqs. (+ F) Model. The Itools online website³² was used to modify and visualize the tree.

According to the position information of the *AP2/EREBP*, *COX* and *LTP* genes on the chromosome, a karyotype map of the *AP2/EREBP*, *COX* and *LTP* genes was drawn using TBtools³³. The output image was used to show all the *AP2/EREBP*, *COX* and *LTP* genes on the chromosome.

The rates of synonymous (*Ks*) and nonsynonymous (*Ka*) substitutions were calculated by TBtools³³ to investigate selection pressure. The divergence time of the gene pairs was estimated using the synonymous mutation rate of substitutions per synonymous site per million years ago (Mya) as follows: " $T = Ks/2\lambda$ ", with a λ value of 6.05×10^{-9} ³⁴.

The duplicated genes were identified as paralogous if the alignment covered $\geq 70\%$ of the longer gene and if the identity of the aligned region was $\geq 70\%$ ³⁵; additionally, the genes were identified by the MEGA-11³¹ gene

duplication wizard. Paralogous gene pair (tandem and segmental genes) collinearity analysis was visualized as a Circos plot through TBtools³³.

TBtools³³ were used to determine the syntenic relationships of the *AP2/EREBP*, *COX* and *LTP* genes in *Z. mays* L. against *O. sativa*, *H. vulgare* and *A. thaliana*.

Conserved domain, conserved motif, gene structure and Promoter analyses of the *AP2/EREBP*, *COX* and *LTP* genes

The NCBI conserved domain tool³⁶ was used to search against the Pfam v34.0–19,178 PSSMs database for *AP2/EREBP*, *COX* and *LTP* proteins. The InterPro tool³⁷ was used to analyze the domains of the *AP2/EREBP*, *COX* and *LTP* proteins. MEME 5.5.2³⁸ was used to compute the conserved motifs of the *AP2/EREBP*, *COX* and *LTP* proteins. Pfam³⁹ was used for motif description. The gene structures obtained from the GFF file were downloaded from the phytozome of the *Z. mays* L. genome and subsequently illustrated using TBtools³³.

The promoter sequences of the *AP2/EREBP*, *COX* and *LTP* genes in *Z. mays* L. 1500 bp upstream of the TSS of each *AP2/EREBP*, *COX* and *LTP* gene were retrieved from the *Z. mays* L. genome sequence file and downloaded from the Phytozome database²⁷. Cis-regulatory elements (CREs) were also analyzed in Plant CARE⁴⁰. A graphical representation of the CRE elements present in the promoter region of the gene was generated via TBTool³³.

Subcellular localization, nuclear localization signal, transmembrane helices, phosphorylation sites and three-dimensional (3-D) structure prediction

Subcellular localization predictor (CELLO) version 2.5 (<http://cello.life.nctu.edu.tw/>)⁴¹ was used to predict the subcellular localization of the proteins, and TBtools was used to visualize the results³³. NLSDB⁴² was used to search for nuclear localization signal potentials. The TMHMM server version 2.0⁴³ confirmed the presence of transmembrane helical domains (TMs) in the *AP2/EREBP*, *COX* and *LTP* proteins. The NetPhos 3.1 server⁴⁴ was used to predict the phosphorylation sites of the *AP2/EREBP*, *COX* and *LTP* proteins. The I-TASSER⁴⁵ program predicted the three-dimensional (3-D) structure of the *AP2/EREBP*, *COX* and *LTP* proteins.

Prediction of miRNAs targeting the *AP2/EREBP*, *COX* and *LTP* genes

The psRNA target database⁴⁶ and miRBase⁴⁷ were used to predict miRNAs. IPknot⁴⁸ was used to predict RNA secondary structures with pseudoknots for the *AP2/EREBP*, *COX* and *LTP* proteins.

Gene Ontology enrichment and functional relationship analysis of the *AP2/EREBP*, *COX* and *LTP* genes

ShinyGO 0.77⁴⁹ was used for Gene Ontology enrichment analysis. We performed a gene ontology (GO) annotation analysis by submitting all the *AP2/EREBP*, *COX* and *LTP* gene sequences to the eggNOG database⁵⁰ and Phytozome database²⁷. The GO annotation data were processed in SRPLOT⁵¹ to construct the gene ontology chord for the functional relationships of the *AP2/EREBP*, *COX* and *LTP* genes.

Maize plant growth and drought treatment

These experiments were conducted in the Department of Botany and Microbiology, Faculty of Science at Cairo University. The seeds used in this study were obtained from the Agricultural Research Center (ARC), located in Giza, Egypt. The seeds originated from Egypt and were certified to be of the white three-way cross 310 variety. Drought treatment was conducted in 2 groups. The first group consisted of control plants, whereas the other consisted of stressed plants. Thirty seeds were planted in small pots in a growth room for 14 days. The second group of stressed plants was subjected to continuous water withholding for 12 h, while the first group was treated with Hoagland's solution as a control. The plants were harvested after 12 h of drought stress. Three control plants and three stressed plants were subjected to RNA extraction and sequencing.

RNA isolation, qRT–PCR expression analysis and sequencing

This study identified *AP2/EREBP* (forward primer: AACCCAAGAACACGCTTCCT and reverse primer: AAGCCACATCCCATCCCAAC), *COX* (forward primer: TCCGTAGTTGGGATTCGTCG and reverse primer: CTG GATTGGTTTCTAGTTTCTTTGA) and *LTP* genes (forward primer: ATAGGAACGTACGCACGCAG and reverse primer: ATGCAAGTCGTGATCATGCG). Total RNA was isolated from the leaves of 15-day-old *Z. mays* L. seedlings using a GeneTireX kit. The residual DNA was removed using RNase-free recombinant DNase I (Thermo Scientific, Litwania). First-strand cDNA was synthesized in a 20 μ L reaction mixture using a Grisp reverse transcription kit (<https://grisp.pt/>) with approximately two micrograms of DNA-free total RNA from each sample. qRT–PCR was performed to quantify the relative transcription levels of the *AP2/EREBP*, *COX* and *LTP* genes expressed in the leaves. qPCR was performed with a CFX Connect Real-Time PCR System (Bio-Rad, Singapore) under the following conditions: 94 °C for 5 min; 40 cycles of 94 °C for 10 s, 58 °C for 20 s, and 72 °C for 30 s; a plate read; a melt curve of 65–95 °C with an increment of 0.5 °C for 10 s; and subsequent sequencing. The Ct (cycle threshold) value was used as a measure of the starting copy number of the target gene⁵². The relative gene expression level was calculated using the $2^{-\Delta\Delta Ct}$ method⁵³. Actin (Act) was used as an internal reference gene. The forward primer used was CTGAGGTTCTATTCCAGCCATCC, and the reverse primer used was CCA CCACTGAGGACAACATTACC.

Ethical approval

The authors declare that the experimental research work involving the growth of plants in this study, was conducted in compliance with relevant institutional, national, and international guidelines and legislation.

Results

Identification of the AP2/EREBP, COX and LTP genes in *Z. mays* L.

A total of 78 AP2/EREBP, 6 COX and 10 LTP candidate genes were retrieved from the *Z. mays* L. genome and were named according to their chromosomal positions from AP2-EREBP-1 to AP2-EREBP-78, COX-1 to COX-6 and LTP-1 to LTP-10 for the AP2/EREBP, COX and LTP genes, respectively (Table S1 Online Resource SI 1).

Characterization of the AP2/EREBP, COX and LTP proteins in *Z. mays* L.

The sequence identities of 78 AP2/EREBP, 6 COX and 10 LTP proteins are shown by the color-by-E-value ratio (blue, $\leq 60\%$; green, $\leq 80\%$; orange, $\leq 90\%$), as shown in Fig. 1. Analysis of protein physical and chemical properties revealed that the length of the AP2/EREBP family amino acids in *Z. mays* L. ranged from 154 (AP2-EREBP-28) to 452 (AP2-EREBP-18). The length of the COX family amino acids ranged from 88 (COX-3) to 483 (COX-1). The length of the LTP family amino acids ranged from 106 (LTP-5) to 247 (LTP-1). The molecular weights (MWs) of AP2/EREBP ranged from 16837.95 (AP2-EREBP-28) to 48127.18 (AP2-EREBP-18). The molecular weights of COX ranged from 9972.12 (COX-3) to 51,834.35 (COX-1). The molecular weights of the LTPs ranged from 11062.05 (LTP-5) to 24921.63 (LTP-1). The isoelectric point (PI) of AP2/EREBP ranged from 4.63 (AP2-EREBP-30) to 10.31 (AP2-EREBP-10). The isoelectric point of COX ranged from 4.46 (COX-4) to

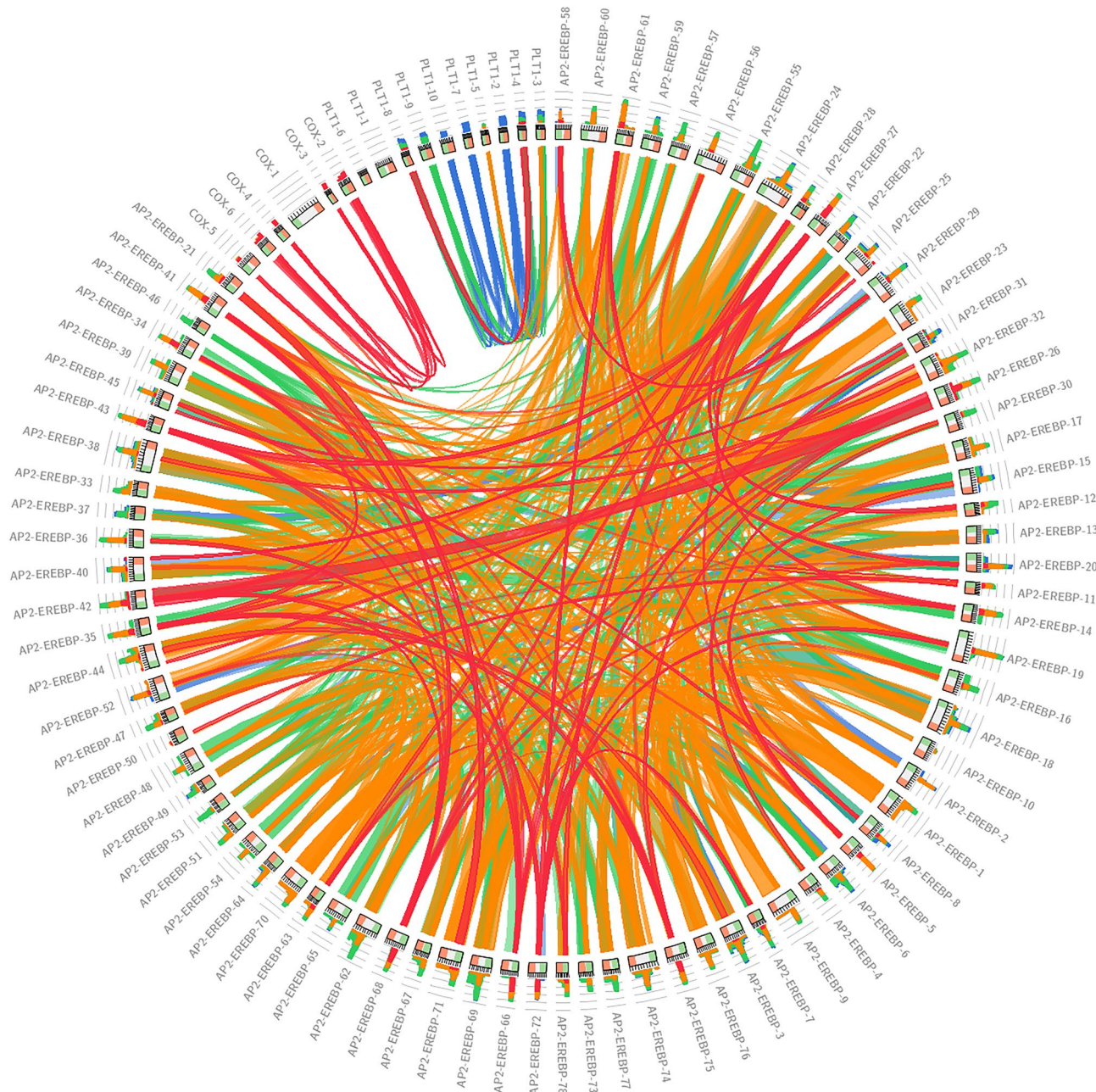


Figure 1. Sequence identity of the AP2/EREBP, COX and LTP proteins.

10.25 (COX-1). The isoelectric point (p-I) of *LTP* ranged from 4.69 (*LTP*-7) to 9.73 (*LTP*-5). The total number of atoms in *AP2/EREBP* ranged from 2329 (*AP2-EREBP*-28) to 6674 (*AP2-EREBP*-24). The total number of atoms in *COX* ranged from 1361 (*COX*-3) to 7348 (*COX*-1). The total number of atoms in the *LTP* ranged from 1564 (*LTP*-5) to 3509 (*LTP*-1). The average hydropathicity value (GRAVY) of *AP2/EREBP* ranged from -0.787 (*AP2-EREBP*-13) to -0.408 (*AP2-EREBP*-58). The average hydropathicity value of *COX* ranged from -0.559 (*COX*-6) to 0.231 (*COX*-1). The average hydropathicity value of *LTP* ranged from 0.147 (*LTP*-10) to 0.683 (*LTP*-6) (Table S1 Online Resource SI 1).

Phylogenetic, chromosomal distribution, evolutionary analysis and synteny analysis of the *AP2/EREBP*, *COX* and *LTP* genes

A phylogenetic tree was constructed using maximum likelihood with 1000 bootstrap replicates, and the *AP2/EREBP*, *COX* and *LTP* protein sequences were used to analyze the possible evolutionary history of *Z. mays* L. In the resulting phylogenetic tree, the *AP2/EREBP* proteins were classified into three distinct clades. The *AP2/EREBP* protein family (*AP2*) includes five different domain types according to the Phytozome-13 website²⁷: the *AP2/ERF* domain, *EREBP*-like factor (*EREBP*), ethylene responsive factor (*ERF*), dehydration responsive element binding protein (*DREB*) and *SHN SHINE* (Fig. 2 and Table S2 Online Resource SI 1). *COX* and *LTP* proteins were classified into three distinct clades (Fig. S1. Online Resource SI 2).

Based on the information available on the Phytozome-13 website²⁷, the *AP2/EREBP*, *COX* and *LTP* genes were physically drawn on the chromosomes in the *Z. mays* L. genome. *AP2/EREBP* genes were found on 10 chromosomes of *Z. mays* L. *COX* genes were found on chromosomes 1, 3, 4, 5, 7 and 8. *LTP* genes were found on chromosomes 1, 3, 6, 8, 9 and 10 (Fig. 3).

The selective pressure on the *AP2/EREBP*, *COX* and *LTP* genes was investigated by calculating the nonsynonymous/synonymous ratio (*Ka/Ks*). A *Ka/Ks* ratio > 1 suggested positive selection, a *Ka/Ks* ratio = 1 indicated neutral selection, and a *Ka/Ks* ratio < 1 suggested purifying selection⁵⁴. In the present study, the *Ka/Ks* ratios of the *AP2/EREBP* paralogous pairs were less than 1, which indicates that the *AP2/EREBP* genes were influenced

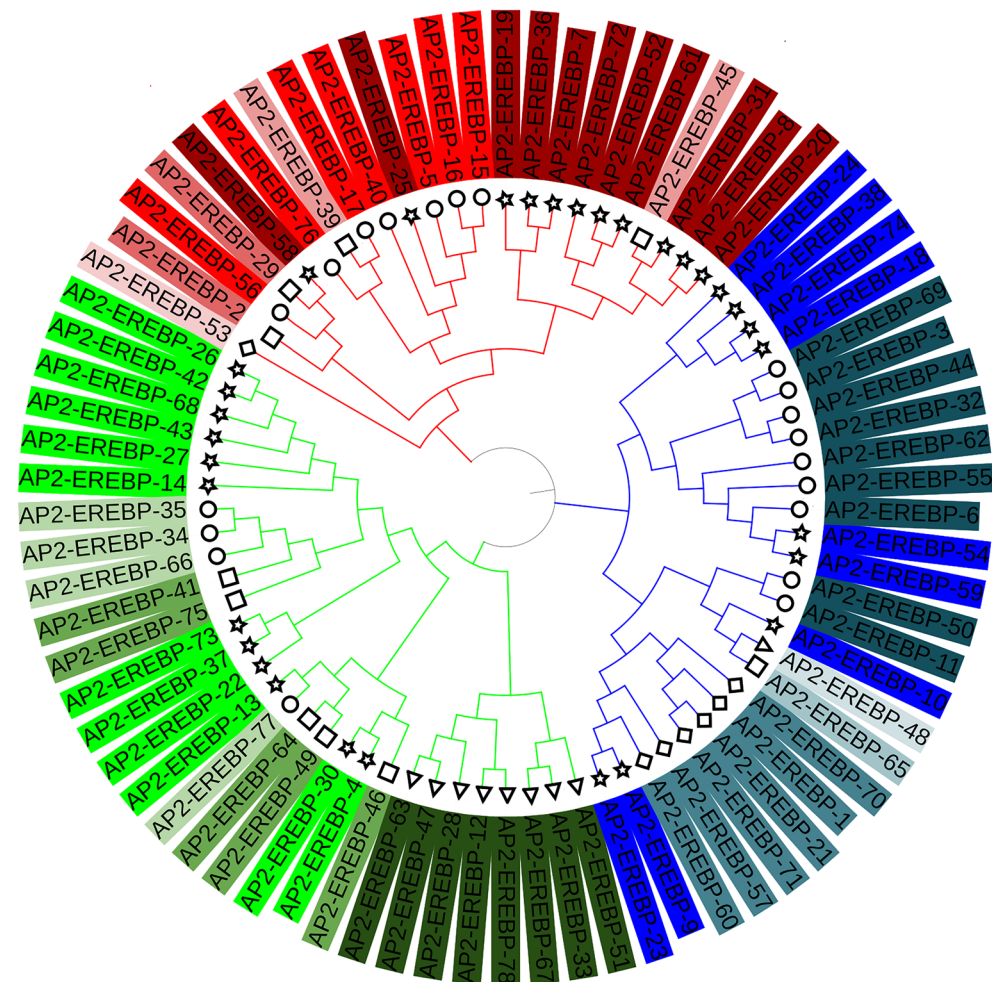


Figure 2. Maximum likelihood phylogenetic tree of the *AP2/EREBP* protein family in *Z. mays* L.; the proteins are labeled with ☆ for the *AP2/ERF* domain, ○ for the *EREBP*-like factor (*EREBP*), □ for the ethylene responsive factor (*ERF*), ◇ for the dehydration responsive element binding protein (*DREB*) and △ for the *SHN SHINE*.

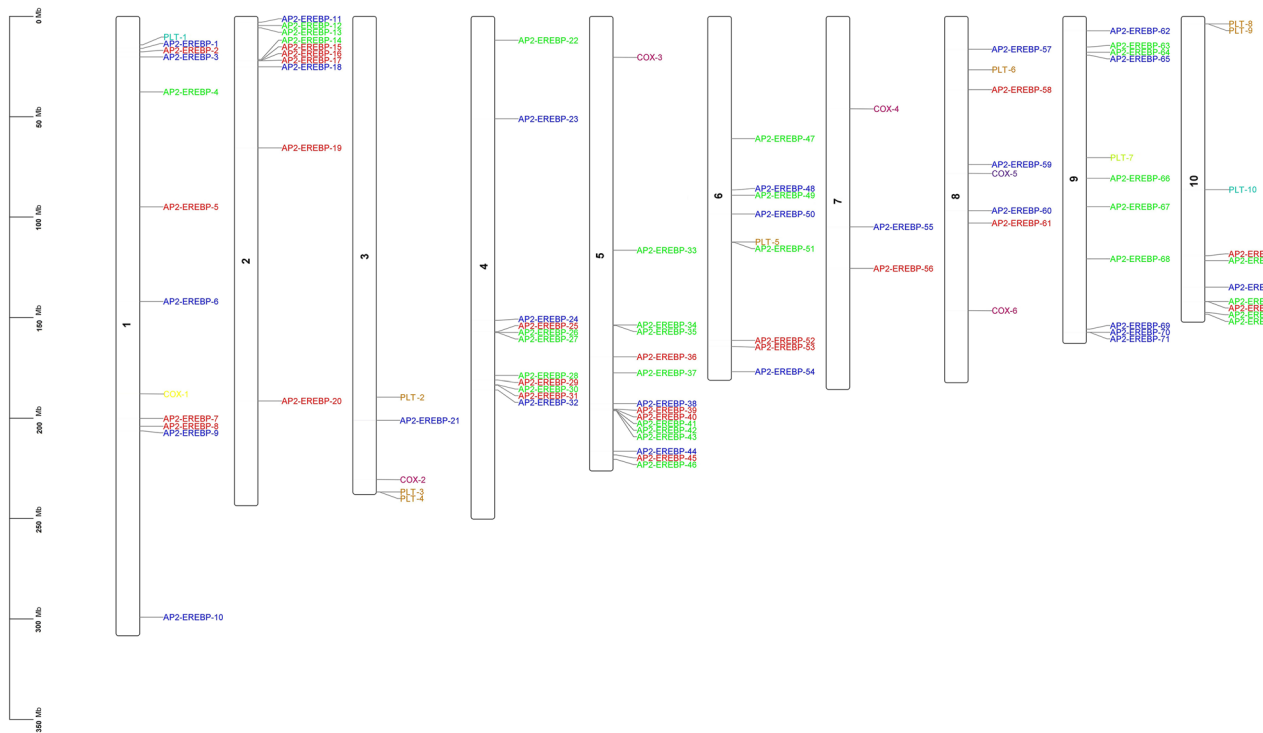


Figure 3. Distribution of the *AP2/EREBP*, *COX* and *LTP* genes on *Z. mays* L. chromosomes.

primarily by purifying selection, which suggests that the *AP2/EREBP* genes received strong environmental pressure during evolution. The Ka/Ks ratios of the *COX-3/COX-4* paralogous pairs were greater than 1, which indicated that the *COX-3/COX-4* genes were involved in positive or Darwinian selection (driving change). For the *LTP* genes, the Ka/Ks ratios of the *LTP-1/LTP-10*, *LTP-5/LTP-3* and *LTP-4/LTP-8* paralogous pairs were less than 1, while those of the *LTP-2/LTP-6* paralogous pairs were greater than 1 (Table 1).

The duplication time of the *AP2/EREBP* paralogous gene pairs in *Z. mays* L. ranged from approximately 9.364 to 100.935 Mya. The duplication time of the *COX-3/COX-4* paralogous gene pair was approximately 5.217 Mya. The duplication time of the *LTP* paralogous gene pairs ranged from approximately 19.064 to 96.477 Mya (Fig. 4 and Table 1).

The *AP2/EREBP*, *COX* and *LTP* genes were analyzed for interspecies collinearity to determine the orthologous relationships of *Z. mays* L. with *O. sativa*, *H. vulgare* and *A. thaliana*. Collinearity analysis revealed robust orthologs of the *AP2/EREBP*, *COX* and *LTP* genes among *Z. mays* L. compared with those of the other three plant species (Fig. 5 and Table S3 Online Resource SI 1).

Conserved domain, conserved motif and gene structure and Promoter analyses of the *AP2/EREBP*, *COX* and *LTP* genes

Domain analysis was carried out for all 78 *AP2/EREBP*, 6 *COX* and 10 *LTP* proteins, and domain analysis confirmed the presence of the AP2 domain (Fig. 6), Cyt_c_Oxidase_Vb domain (Fig. S2 Online Resource SI 2) and the nsLTP1 domain (Fig. S3 Online Resource SI 2) on the *AP2/EREBP*, *COX* and *LTP* proteins, respectively. Motif analysis indicated that the phylogenetic relationships were similar to the conserved motif distributions within the clade. For instance, the motif distributions of the *AP2/EREBP*, *COX* and *LTP* proteins exhibited similar motifs within the clade, with few differences. The *AP2/EREBP* motif distributions for *AP2-EREBP-26*, *AP2-EREBP-42*, *AP2-EREBP-68*, *AP2-EREBP-43*, *AP2-EREBP-27*, *AP2-EREBP-14*, *AP2-EREBP-75*, *AP2-EREBP-41*, *AP2-EREBP-66*, *AP2-EREBP-35*, *AP2-EREBP-34*, *AP2-EREBP-22*, *AP2-EREBP-73*, *AP2-EREBP-37*, *AP2-EREBP-30*, *AP2-EREBP-4*, *AP2-EREBP-46*, *AP2-EREBP-13*, *AP2-EREBP-77*, *AP2-EREBP-64* and *AP2-EREBP-49* proteins had conserved motif numbers 1, 2, 3, and 8. The *AP2-EREBP-67*, *AP2-EREBP-33*, *AP2-EREBP-51*, *AP2-EREBP-12*, *AP2-EREBP-78*, *AP2-EREBP-28*, *AP2-EREBP-63* and *AP2-EREBP-47* proteins carried conserved motif numbers 1, 2, 3, and 9. The remaining *AP2/EREBP* proteins carried conserved motif numbers 1, 2 and 3 (Fig. 6 and Sheet 2 Online Resource SI 1). The *COX* motif distributions for the *COX-4*, *COX-3*, *COX-2* and *COX-6* proteins revealed conserved motif numbers 1 and 3 (Fig. 6 and Sheet 3 Online Resource SI 1). Most of the *LTP* proteins presented conserved motif numbers of 1, 2 and 3 (Fig. 6 and Sheet 4 Online Resource SI 1). The exon–intron structure is an important source of plant biodiversity and gene family evolution. The gene structure results revealed that 21 of the 78 *AP2/EREBP* genes had introns (Fig. 6). All the *COX* genes had introns, while 8 *LTP* genes had introns (Fig. S2 and Fig. S3 Online Resource SI 2).

The *AP2/EREBP*, *COX* and *LTP* gene sequences (1500 bp upstream of the start codon) (Table S4 Online Resource SI 1) were selected for cis-element analysis using the PlantCARE web tool to identify their biological functions (stress response, growth and development). The promoter regions of the *AP2/EREBP*, *COX* and *LTP*

Locus 1	locus 2	Ka	Ks	Ka/Ks	Time
AP2-EREBP-15	AP2-EREBP-16	0.201887685	0.502399279	0.401847083	38.292628
AP2-EREBP-36	AP2-EREBP-7	0.435204658	0.87977765	0.494675737	67.05622333
AP2-EREBP-52	AP2-EREBP-61	0.07647816	0.207878061	0.367899141	15.84436439
AP2-EREBP-8	AP2-EREBP-20	0.385987339	0.570267544	0.676853071	43.46551404
AP2-EREBP-74	AP2-EREBP-18	0.05168891	0.193976752	0.266469614	14.78481342
AP2-EREBP-69	AP2-EREBP-3	0.077309887	0.187624472	0.412045858	14.30064575
AP2-EREBP-44	AP2-EREBP-32	0.057691935	0.122859443	0.469576728	9.364286818
AP2-EREBP-54	AP2-EREBP-59	0.110298946	0.164779455	0.669373168	12.55940971
AP2-EREBP-50	AP2-EREBP-11	0.338025374	0.624790478	0.541021968	47.62122548
AP2-EREBP-21	AP2-EREBP-71	0.334683458	0.545162686	0.61391483	41.55203397
AP2-EREBP-9	AP2-EREBP-23	0.103291176	0.277490994	0.372232534	21.15022823
AP2-EREBP-33	AP2-EREBP-67	0.176957458	1.324271753	0.133626242	100.935347
AP2-EREBP-47	AP2-EREBP-63	0.128995603	0.294776002	0.437605512	22.46768307
AP2-EREBP-4	AP2-EREBP-30	0.449320876	0.611974544	0.734214978	46.64440122
AP2-EREBP-49	AP2-EREBP-64	0.056604591	0.171835289	0.329411909	13.09720188
AP2-EREBP-37	AP2-EREBP-73	0.290799369	0.632410291	0.459827067	48.2020039
AP2-EREBP-34	AP2-EREBP-35	0.187778474	0.553114724	0.339492814	42.15813446
AP2-EREBP-42	AP2-EREBP-26	0.083190211	0.207698132	0.400534227	15.83065032
COX-3	COX-4	0.100244299	0.068451386	1.464459755	5.217331218
LTP-1	LTP-10	0.80611789	1.265784761	0.636852264	96.477497
LTP-2	LTP-6	0.837292901	0.570546713	1.467527341	43.48679218
LTP-5	LTP-3	0.397424829	0.419284855	0.947863543	31.95768715
LTP-4	LTP-8	0.111949387	0.25012764	0.447569039	19.06460671

Table 1. Paralogous pairs of *AP2/EREBP*, *COX* and *LTP* genes and the *Ka/Ks* ratio.

genes in *Z. mays* L. contain a large number of plant hormone response elements. Most *AP2/EREBP*, *COX* and *LTP* proteins contain defense and stress response elements, abscisic acid-responsive elements, methyl jasmonate (MeJA)-responsive elements, salicylic acid and the MYB binding site (MBS) element, which are involved in the drought response (Fig. 7).

Subcellular localization, nuclear localization signal, transmembrane helices, phosphorylation sites and three-dimensional (3-D) structure prediction

Subcellular localization analysis revealed that most of the *AP2/EREBP* proteins were located in the nucleus. *COX* proteins were predicted to be expressed in different organelles; for instance, *COX-1* was predicted to be expressed in the plasma membrane, whereas *COX-2* and *COX-6* were predicted to be expressed in chloroplasts. Most of the *LTP* proteins were located in the extracellular space. A heatmap was constructed to predict the subcellular localization of the *AP2/EREBP*, *COX* and *LTP* proteins, as shown in Fig. S4 Online Resource SI 2 and Table S5 Online Resource SI 1.

Sixteen putative nuclear localization signals (NLSs) were predicted for 21 *AP2/EREBP* proteins, whereas no nuclear localization signals (NLSs) were predicted for *COX* or *LTP* proteins (Table S6 Online Resource SI 1).

The TMHMM results predicted the transmembrane helices in *AP2-EREBP-10*, *AP2-EREBP-21*, *COX-1* and all 10 *LTP* proteins (Fig. S5, S6 and S7 Online Resource SI 2).

The phosphorylation site prediction results for the *AP2/EREBP*, *COX* and *LTP* proteins for kinases are shown in Table S7 Online Resource SI 1.

To study the putative functions of the *AP2/EREBP*, *COX* and *LTP* proteins in *Z. mays* L., we selected a protein from each clade. The *AP2-EREBP-24*, *AP2-EREBP-51*, *AP2-EREBP-53*, *COX-1*, *COX-2*, *COX-5*, *LTP-1*, *LTP-3* and *LTP-7* proteins were modeled with I-TASSER software to construct 3-D structures. The 3-D structures were constructed according to similar structural templates and crystal structures obtained from the Protein Data Bank (Fig. 8). C-scores were used to estimate the confidence of the constructed protein model for the *AP2-EREBP-24*, *AP2-EREBP-51*, *AP2-EREBP-53*, *COX-1*, *COX-2*, *COX-5*, *LTP-1*, *LTP-3* and *LTP-7* proteins. The closest structural similarity protein models were selected as the best-predicted models for the *AP2/EREBP*, *COX* and *LTP* proteins, with C-scores ranging (Table 2). Due to their structural similarity, proteins that are structurally close to the target in the PDB often have similar functions. The C-scores suggested that the structures of the *AP2/EREBP*, *COX* and *LTP* proteins were constructed with high accuracy.

Prediction of miRNAs targeting the *AP2/EREBP*, *COX* and *LTP* proteins

A total of 187 microRNAs were predicted to target the *AP2/EREBP* genes, 48 microRNAs were predicted to target *COX* proteins, and 30 microRNAs were predicted to target *LTP* genes. The microRNA targeting relationships for the *AP2/EREBP*, *COX* and *LTP* genes are shown in Table S8 Online Resource SI 1.

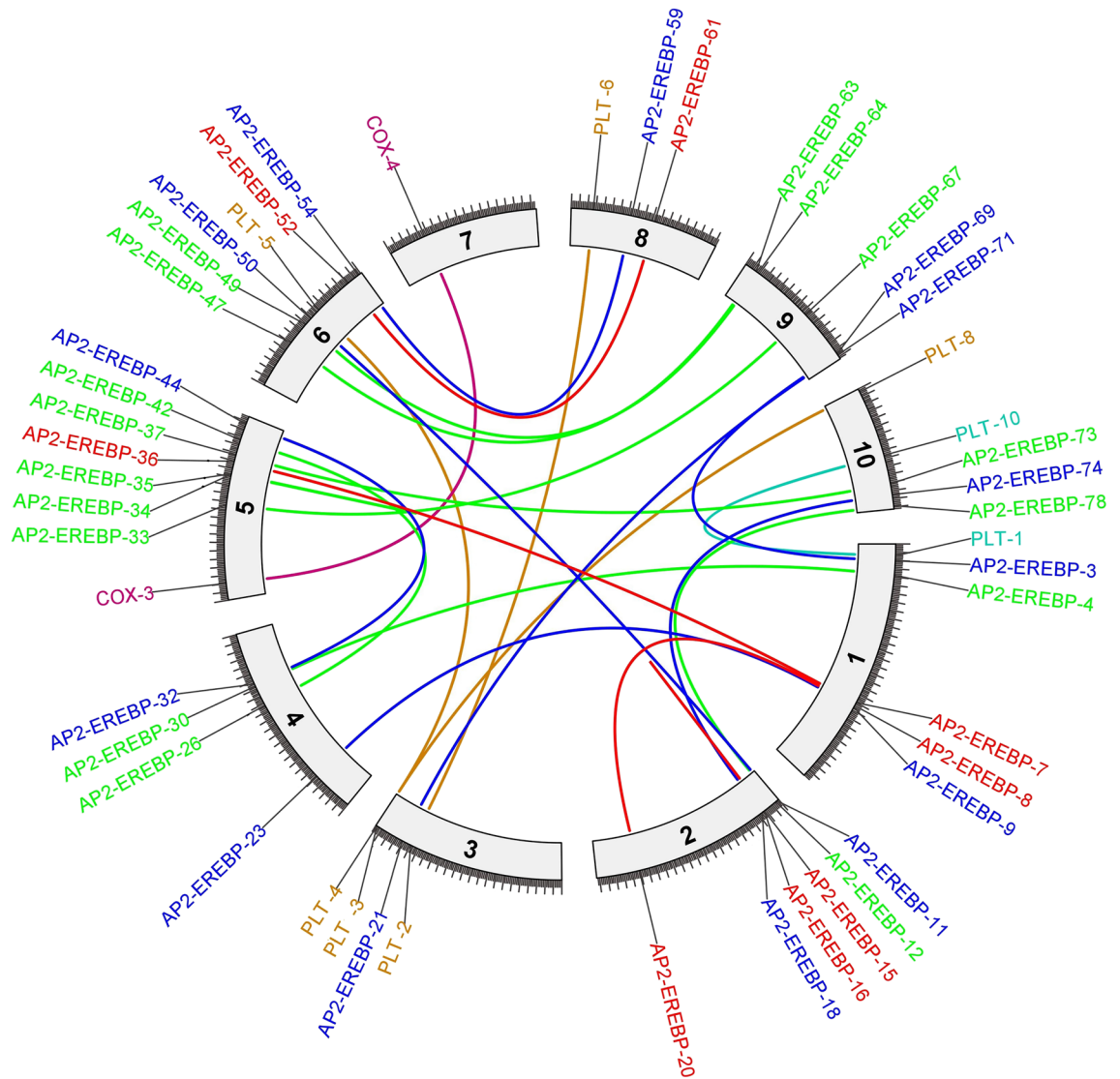


Figure 4. Segmental and tandem duplication of *AP2/EREBP*, *COX* and *LTP* among the *Z. mays* L. chromosomes.

The results from the prediction of RNA secondary structures with pseudoknots for the *AP2/EREBP* (*AP2-EREBP-24*, *AP2-EREBP-51* and *AP2-EREBP-53*); *COX* (*COX-1*, *COX-2* and *COX-5*); and *LTP* (*LTP-1*, *LTP-3* and *LTP-7*) proteins are shown in Fig. S8: Fig. S16 Online Resource SI 2.

Gene ontology enrichment and functional relationship analysis

To further determine the functions of the *AP2/EREBP*, *COX* and *LTP* genes, we performed enrichment analysis and gene ontology (GO) analysis based on biological processes and molecular functions. GO terms help us understand the function of genes at the molecular level (Figs. S17, S18 and S19 Online Resource SI 2). GO terms for the *AP2/EREBP*, *COX* and *LTP* genes confirmed the functional role of *AP2/EREBP*, *COX* and *LTP* as stress responsive genes (Fig. S20, S21 and S22 Online Resource SI 2).

In the present study, qRT-PCR analysis revealed that the *AP2/EREBP*, *COX* and *LTP* proteins were expressed in leaves, and drought decreased the expression levels of *AP2/EREBP*, *COX* and *LTP* by 0.84, 0.53 and 0.31, respectively, after 12 h of drought stress (Sheet 1 Online Resource SI 1). Domain structure, promoter and gene ontology enrichment analyses confirmed the functional role of the *AP2/EREBP*, *COX* and *LTP* proteins in stress responses.

Discussion

Genome-wide analysis identified 78 *AP2/EREBP* genes in *Z. mays* L. Phylogenetic classification revealed that the 78 *AP2/EREBP* proteins could be divided into three distinct clades, which included the *AP2/ERF* domain, the *EREBP*-like factor (*EREBP*), the ethylene responsive factor (*ERF*), the dehydration responsive element binding protein (*DREB*) and the SHN SHINE according to the description available on Phytozome-13²⁷. Cheng *et al.*³⁵ identified 229 *AP2/ERF* genes in the maize genome. In addition, phylogenetic analysis revealed that the

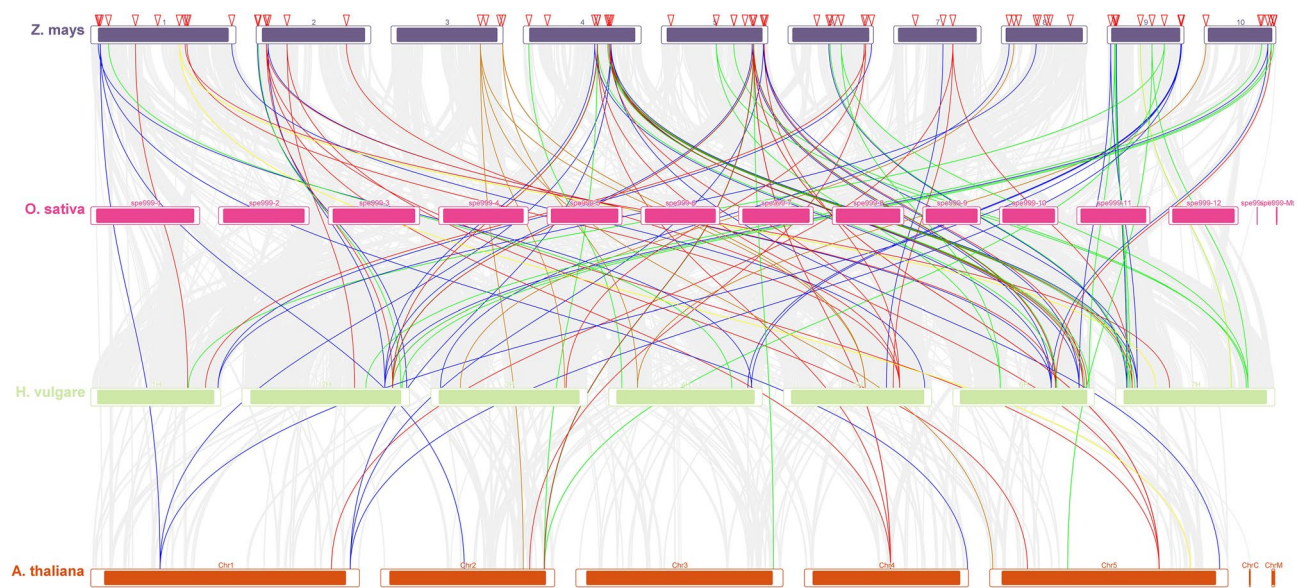


Figure 5. The collinear relationships of the *AP2/EREBP*, *COX* and *LTP* genes are shown as colored lines in the phylogenetic tree.

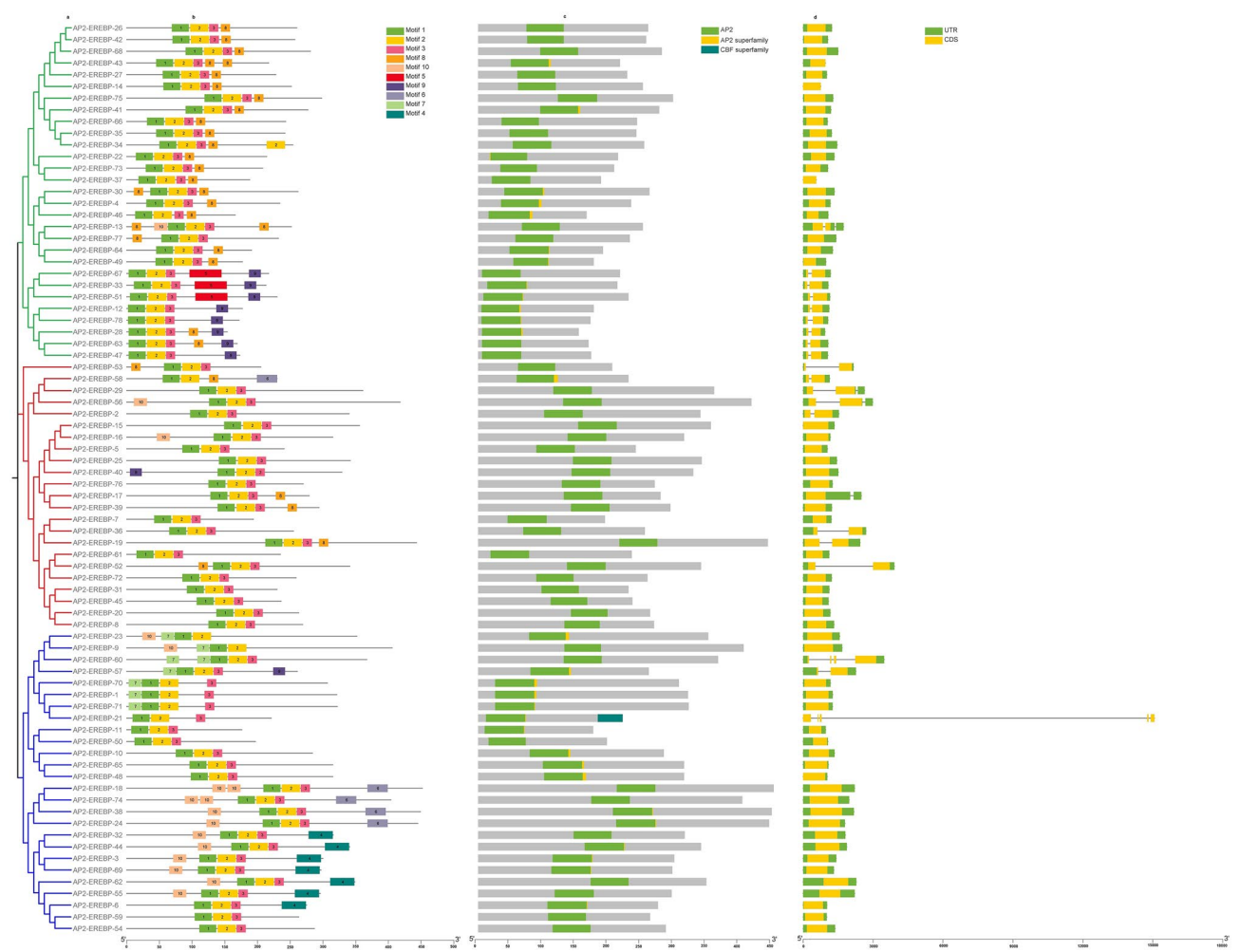


Figure 6. AP2/EREBP proteins. (a) Rectangular phylogenetic tree. (b) Conserved motifs were predicted using MEME. (c) Protein domains. (d) Gene structure.



Figure 7. Cis-acting elements in the promoter regions (1500 bp upstream of the start codon) of the AP2/EREBP, COX and LTP genes.

ZmAP2/ERF family members could be divided into five clades, namely, 27 AP2 (*APETALA2*), 105 ERF (*ethylene responsive factor*), 89 DREB (*dehydration responsive element binding*), 5 RAV (*related to ABI3/VP*) and a soloist. In this study, we identified 6 COXs in *Z. mays* L. To date, no comprehensive investigation of the COX gene in maize has been reported via genome-wide analysis. We also identified 10 LTP genes in *Z. mays* L. Wei and Zhong²² identified 63 LTP genes in maize, which were divided into six types, whereas Fang *et al.*⁵⁶ identified 65 LTP genes in maize. Our qRT–PCR results indicated that AP2/EREBP COX and LTP were downregulated, with fold changes

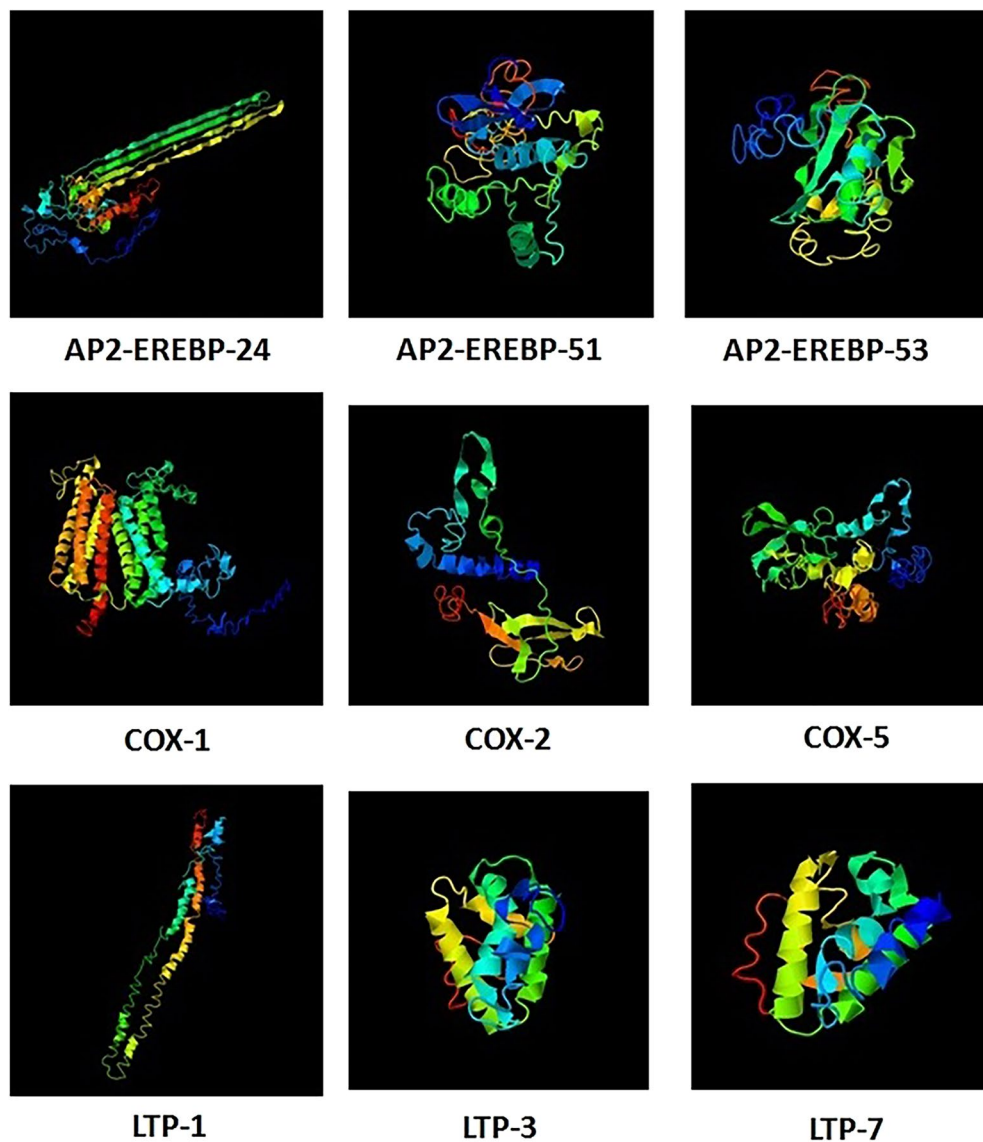


Figure 8. Structural analysis of the AP2-EREBP-24, AP2-EREBP-51, AP2-EREBP-53, COX-1, COX-2, COX-5, LTP-1, LTP-3 and LTP-7 proteins.

Protein	C-Score	TM-Score	RMSD (Å)	Best Identified Structural Analogs in PDB				
				PDB Hit	TM-Score a	RMSD a	IDEN a	Cov
AP2-EREBP-24	-1.98	0.48±0.15	11.8±4.5	5fmwA	0.896	2.59	0.094	0.957
AP2-EREBP-51	-3.82	0.30±0.10	14.9±3.6	6fokA	0.438	5.86	0.053	0.748
AP2-EREBP-53	-2.77	0.40±0.13	11.8±4.5	5wx9A	0.561	2.35	0.315	0.634
COX-1	-1.50	0.53±0.15	10.8±4.6	6a2jA	0.623	1.12	0.154	0.631
COX-2	-2.04	0.47±0.15	9.5±4.6	6t15d	0.690	1.18	0.283	0.727
COX-5	-4.30	0.26±0.08	16.4±3.0	6rc9A1	0.435	6.24	0.063	0.771
LTP-1	-4.31	0.26±0.08	16.5±3.0	5owvC	0.536	4.62	0.044	0.757
LTP-3	-1.19	0.57±0.15	6.8±4.1	4xuwA	0.718	0.88	0.446	0.748
LTP-7	-0.89	0.60±0.14	6.1±3.8	4xuwA	0.720	1.08	0.319	0.765

Table 2. Modeling parameters for the AP2/EREBP, COX and LTP proteins.

of 0.84, 0.53 and 0.31, respectively, under drought stress. Sharoni *et al.*⁵⁷ reported the same results for the AP2, DREB, and ERF genes in the IR77298-14-1-2-B-10 line; these genes were highly activated in leaves under severe stress treatment and downregulated under severe stress treatment. Similarly, Trindade *et al.*⁵⁸ reported the same results in *Medicago truncatula* for *COX5b*, which strongly downregulated under water deficit conditions. Wei and Zhong²² reported the same results in maize, in four *ZmLTP* which were downregulated under drought stress. These results suggest that, under drought stress, *ZmLTP* genes may exhibit tissue-specific expression because *ZmLTP1.2* was significantly downregulated in ovarian tissue but upregulated in the leaf meristem. In the present study, Pfam domain analysis confirmed the presence of the AP2 domain, Cyt_c_Oxidase_Vb and nsLTP1 on the AP2/EREBP, COX and LTP proteins, respectively. Motif and gene structure analyses indicated that genes with closer phylogenetic relationships exhibited more similar genetic structures. The promoter regions of the AP2/EREBP, COX and LTP genes contain defense and stress response elements, abscisic acid-responsive elements, methyl jasmonate (MeJA)-responsive elements, salicylic acid and the MYB binding site (MBS) element, which are involved in the drought response. The AP2/EREBP genes were found on 10 chromosomes of *Z. mays* L., the same results were found by Cheng *et al.*⁵⁵. COX genes were found on chromosomes 1, 3, 4, 5, 7 and 8. LTP genes were found on chromosomes 1, 3, 6, 8, 9 and 10. The Ka/Ks ratios of the AP2/EREBP paralogous pairs were less than 1, which indicates that the AP2/EREBP genes were influenced primarily by purifying selection, which means that the AP2/EREBP genes received strong environmental pressure during evolution. The Ka/Ks ratios of the COX-3/COX-4 paralogous pairs were greater than 1, which indicated that the COX-3/COX-4 genes were involved in positive or Darwinian selection (driving change). For the LTP genes, the Ka/Ks ratios of the LTP-1/LTP-10, LTP-5/LTP-3 and LTP-4/LTP-8 paralogous pairs were less than 1, which indicates that these genes were involved in purifying selection, while the Ka/Ks ratios of the LTP-2/LTP-6 paralogous pairs were greater than 1. Synteny analysis of the AP2/EREBP, COX and LTP proteins revealed collinearity orthologous relationships in *O. sativa*, *H. vulgare* and *A. thailana*. Gene Ontology enrichment analysis confirmed the functional role of stress-responsive AP2/EREBP, COX and LTP.

Conclusion

This is the first study to identify the COX gene in *Z. mays* L. by genome-wide analysis. Domain structure, promoter and gene ontology enrichment analyses confirmed the functional role of the AP2/EREBP, COX and LTP proteins in stress responses. The results of the present study could improve the understanding of how AP2/EREBP, COX and LTP are mechanistically linked to drought stress responses in maize and could be used for the genetic improvement of maize.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Received: 20 January 2024; Accepted: 18 March 2024

Published online: 31 March 2024

References

- Warrick, R. A. *Drought Hazard in the United States: A Research Assessment*. Monograph no. NSF/RA/E-75/004, 199. (University of Colorado, Institute of Behavioral Science, 1975).
- Huschke, R. E. (ed.) *Glossary of Meteorology* 638 (American Meteorological Society, 1959).
- Rosenberg, N. J., ed. Drought in the Great Plains—Research on impacts and strategies. In *Proceedings of the Workshop on Research in Great Plains Drought Management Strategies*, University of Nebraska, Lincoln 225. (Water Resources Publications, 1979).
- Yevjevich, V., Hall, W. A. & Salas, J. D., eds. Drought research needs. In *Proceedings of the Conference on Drought Research Needs* 276. (Colorado State University, 1977).
- The Food and Agriculture Organization (FAO) <https://www.fao.org>
- Maazou, A. R. S., Tu, J., Qiu, J. & Liu, Z. Breeding for drought tolerance in maize (*Zea mays* L.). *Am. J. Plant Sci.* 7(14), 1858 (2016).
- Xie, Z., Nolan, T. M., Jiang, H. & Yin, Y. AP2/ERF transcription factor regulatory networks in hormone and abiotic stress responses in *Arabidopsis*. *Front. Plant Sci.* 10, 228. <https://doi.org/10.3389/fpls.2019.00228> (2019).
- Kizis, D., Lumberras, V. & Pagès, M. Role of AP2/EREBP transcription factors in gene regulation during abiotic stress. *FEBS Lett.* 498(2–3), 187–189. [https://doi.org/10.1016/s0014-5793\(01\)02460-7](https://doi.org/10.1016/s0014-5793(01)02460-7) (2001).
- Chen, L. *et al.* Expansion and stress responses of AP2/EREBP superfamily in *Brachypodium distachyon*. *Sci. Rep.* 6, 21623. <https://doi.org/10.1038/srep21623> (2016).
- Pande, A. *et al.* Role of AP2/EREBP transcription factor family in environmental stress tolerance. *Cell Cell. Life Sci. J.* <https://doi.org/10.23880/CCLSJ-16000120> (2018).
- Liu, C. & Zhang, T. Expansion and stress responses of the AP2/EREBP superfamily in cotton. *BMC Genom.* 18, 118. <https://doi.org/10.1186/s12864-017-3517-9> (2017).
- Dietz, K. J., Vogel, M. O. & Viehhauser, A. AP2/EREBP transcription factors are part of gene regulatory networks and integrate metabolic, hormonal and environmental signals in stress acclimation and retrograde signaling. *Protoplasma* 245(1–4), 3–14. <https://doi.org/10.1007/s00709-010-0142-8> (2010).
- Li, B. *et al.* Two novel AP2/EREBP transcription factor genes TaPARG have pleiotropic functions on plant architecture and yield-related traits in common wheat. *Front. Plant Sci.* 7, 1191. <https://doi.org/10.3389/fpls.2016.01191> (2016).
- Zhou, Y. *et al.* Overexpression of a cotton gene that encodes a putative transcription factor of AP2/EREBP family in *Arabidopsis* affects growth and development of transgenic plants. *PLoS ONE* 8(10), e78635. <https://doi.org/10.1371/journal.pone.0078635> (2013).
- Jett, K. A. & Leary, S. C. Building the CuA site of cytochrome c oxidase: A complicated, redox-dependent process driven by a surprisingly large complement of accessory proteins. *J. Biol. Chem.* 293, 4644–4652 (2017).
- Mansilla, N., Racca, S., Gras, D. E., Gonzalez, D. H. & Welchen, E. The complexity of mitochondrial complex IV: An update of Cytochrome c Oxidase biogenesis in plants. *Int. J. Mol. Sci.* 19, 662 (2018).
- Ikeda, T. M. & Tsunewaki, K. Deficiency of cox1 gene expression in wheat plants with *Aegilops columnaris* cytoplasm. *Curr. Genet.* 6, 509–514. <https://doi.org/10.1007/s002940050163> (1996).

18. D'Agostino, N. *et al.* Identification of non-specific Lipid Transfer Protein gene family members in *Solanum lycopersicum* and insights into the features of Sola 13 protein. *Sci. Rep.* **9**, 1607. <https://doi.org/10.1038/s41598-018-38301-z> (2019).
19. Mikołajczak, K., Ogródowicz, P., Surma, M., Adamski, T. & Kuczyńska, A. Introgression of LTP2 gene through marker assisted backcross in barley (*Hordeum vulgare* L.). *Electron. J. Biotechnol.* **24**, 9–11 (2016).
20. Guo, L., Yang, H., Zhang, X. & Yang, S. Lipid transfer protein 3 as a target of MYB96 mediates freezing and drought stress in Arabidopsis. *J. Exp. Bot.* **64**, 1755–1767 (2013).
21. Salminen, T. A., Blomqvist, K. & Edqvist, J. Lipid transfer proteins: Classification, nomenclature, structure, and function. *Planta* **244**, 971–997. <https://doi.org/10.1007/s00425-016-2585-4> (2016).
22. Wei, K. & Zhong, X. Non-specific lipid transfer proteins in maize. *BMC Plant Biol.* **14**, 281 (2014).
23. Guo, C., Ge, X. & Ma, H. The rice OsDIL gene plays a role in drought tolerance at vegetative and reproductive stages. *Plant Mol. Biol.* **82**, 239–253 (2013).
24. Yu, G. *et al.* Identification of wheat non-specific lipid transfer proteins involved in chilling tolerance. *Plant Cell Rep.* **33**, 1757–1766 (2014).
25. Tapia, G., Morales-Quintana, L., Parra, C., Berbel, A. & Alcorta, M. Study of nsLTPs in *Lotus japonicus* genome reveal a specific epidermal cell member (LjLTP10) regulated by drought stress in aerial organs with a putative role in cutin formation. *Plant Mol. Biol.* **82**, 485–501 (2013).
26. Edstam, M. M. *et al.* Characterization of the GPI-anchored lipid transfer proteins in the moss *Physcomitrella patens*. *Plant Physiol. Biochem.* **75**, 55–69 (2014).
27. Goodstein, D. M. *et al.* Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178–D1186 (2012).
28. Resource Coordinators, N. C. B. I. *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **46**, D8–D13 (2018).
29. Darzentas, N. Circoletto: Visualizing sequence similarity with Circos. *Bioinformatics* **26**(20), 2620–2621. <https://doi.org/10.1093/bioinformatics/btq484> (2010).
30. Gasteiger, E., *et al.* Protein Identification and Analysis Tools on the ExPASy Server 571–607 (2005)
31. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
32. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).
33. Chen, C. *et al.* TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **13**, 1194–1202 (2020).
34. Yang, S., Zhang, X., Yue, J.-X., Tian, D. & Chen, J.-Q. Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mol. Genet. Genom.* **280**, 187–198. <https://doi.org/10.1007/s00438-008-0355-0> (2008).
35. Gu, Z., Cavalcanti, A., Chen, F. C., Bouman, P. & Li, W. H. Extent of gene duplication in the genomes of Drosophila, nematode, and yeast. *Mol. Biol. Evol.* **19**(3), 256–262 (2002).
36. Lu, S. *et al.* CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **48**(D1), D265–D268. <https://doi.org/10.1093/nar/gkz991> (2020).
37. Paysan-Lafosse, T. *et al.* InterPro in 2022. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkac993> (2022).
38. Bailey, T. L., Johnson, J., Grant, Ch. E. & Noble, W. S. The MEME suite. *Nucleic Acids Res.* **43**(W1), W39–W49. <https://doi.org/10.1093/nar/gkv416> (2015).
39. Mistry, J. *et al.* The protein families database in 2021. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkaa913> (2020).
40. Rombauts, S., Déhais, P., Van Montagu, M. & Rouzé, P. PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Res.* **27**, 295–296. <https://doi.org/10.1093/nar/27.1.295> (1999).
41. Yu, C. S., Chen, Y. C., Lu, C. H. & Hwang, J. K. Prediction of protein subcellular localization. *Proteins Struct. Funct. Bioinform.* **64**, 643–651 (2006).
42. Nair, R., Carter, P. & Rost, B. NLSdb: Database of nuclear localization signals. *Nucleic Acids Res.* **31**(1), 397–399. <https://doi.org/10.1093/nar/gkg001> (2003).
43. Krogh, A., Larsson, B., Von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
44. Blom, N., Gammeltoft, S. & Brunak, S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* **294**(5), 1351–1362 (1999).
45. Zheng, W. *et al.* Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell Rep. Methods* **1**, 100014 (2021).
46. Dai, X., Zhuang, Z. & Zhao, P. X. psRNATarget: A plant small RNA target analysis server (2017 release). *Nucl. Acids Res.* **46**(W1), W49–W54. <https://doi.org/10.1093/nar/gky316> (2018).
47. Kozomara, A. & Griffiths-Jones, S. miRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **42**, D68–73 (2014).
48. Sato, K. & Kato, Y. Prediction of RNA secondary structure including pseudoknots for long sequences. *Brief. Bioinform.* **23**(1), bbab395. <https://doi.org/10.1093/bib/bbab395> (2022).
49. Ge, S. X., Jung, D. & Yao, R. ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).
50. Cantalapiedra, C. P., Hernandez-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**(12), 5825–5829. <https://doi.org/10.1093/molbev/msab293> (2021).
51. Tang, D. *et al.* SRplot: A free online platform for data visualization and graphing. *PLoS One* **18**(11), e0294236. <https://doi.org/10.1371/journal.pone.0294236> (2023).
52. Chen, Y. *et al.* Selection of reference genes for quantitative real-time PCR normalization in creeping bentgrass involved in four abiotic stresses. *Plant Cell Rep.* **34**, 1825–1834. <https://doi.org/10.1007/s00299-015-1830-9> (2015).
53. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. *Methods* **25**, 402–408 (2001).
54. Anton, N., Makova, K. D. & Li, W. H. The Ka/Ks ratio test for assessing the protein-coding potential of genomic regions: An empirical and simulation study. *Genome Res.* **12**, 198–202 (2002).
55. Cheng, C. *et al.* Wide-range portrayal of AP2/ERF Transcription factor family in maize (*Zea mays* L.) development and stress responses. *Genes (Basel)* **14**(1), 194. <https://doi.org/10.3390/genes14010194> (2023).
56. Fang, C. *et al.* Systematic investigation of lipid transfer proteins involved in male fertility and other biological processes in maize. *Int. J. Mol. Sci.* **24**(2), 1660. <https://doi.org/10.3390/ijms24021660> (2023).
57. Sharoni, A. M. *et al.* Comparative transcriptome analysis of AP2/EREBP gene family under normal and hormone treatments, and under two drought stresses in NILs setup by Aday selection and IR64. *Mol. Genet. Genom.* **287**(1), 1–19. <https://doi.org/10.1007/s00438-011-0659-3> (2012).
58. Trindade, I., Capita, C., Dalmay, T., Feveiro, M. P. & Santos, D. M. miR398 and miR408 are upregulated in response to water deficit in *Medicago truncatula*. *Planta* **3**, 705–771 (2010).

Author contributions

A.M. devised the study, M.A. and A.M. collected the data, conducted the analysis, and discussed the data. A.M. wrote the first draft of the manuscript. Both authors wrote, read, and approved the final version of the manuscript.

Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-57376-5>.

Correspondence and requests for materials should be addressed to A.M.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2024