

RESEARCH

Open Access



De novo assembly provides new insights into the evolution of *Elaeagnus angustifolia* L.

Yunfei Mao¹, Xueli Cui¹, Haiyan Wang¹, Xin Qin¹, Yangbo Liu¹, Yijun Yin¹, Xiafei Su¹, Juan Tang², Fengling Wang², Fengwang Ma³, Naibin Duan⁴, Donglin Zhang⁵, Yanli Hu¹, Wenli Wang¹, Shaochong Wei¹, Xiaoliu Chen¹, Zhiquan Mao¹, Xuesen Chen¹ and Xiang Shen^{1*}

Abstract

Background: *Elaeagnus angustifolia* L. is a deciduous tree in the family *Elaeagnaceae*. It is widely used to study abiotic stress tolerance in plants and to improve desertification-affected land because of its ability to withstand diverse types of environmental stress, such as drought, salt, cold, and wind. However, no studies have examined the mechanisms underlying the resistance of *E. angustifolia* to environmental stress and its adaptive evolution.

Methods: Here, we used PacBio, Hi-C, resequencing, and RNA-seq to construct the genome and transcriptome of *E. angustifolia* and explore its adaptive evolution.

Results: The reconstructed genome of *E. angustifolia* was 526.80 Mb, with a contig N50 of 12.60 Mb and estimated divergence time of 84.24 Mya. Gene family expansion and resequencing analyses showed that the evolution of *E. angustifolia* was closely related to environmental conditions. After exposure to salt stress, GO pathway analysis showed that new genes identified from the transcriptome were related to ATP-binding, metal ion binding, and nucleic acid binding.

Conclusion: The genome sequence of *E. angustifolia* could be used for comparative genomic analyses of *Elaeagnaceae* family members and could help elucidate the mechanisms underlying the response of *E. angustifolia* to drought, salt, cold, and wind stress. Generally, these results provide new insights that could be used to improve desertification-affected land.

Keywords: *Elaeagnus angustifolia* L., Genome, Evolutionary response mechanism, Desertification-affected land

Background

The world's population is increasing rapidly and is projected to reach 9.6 billion by 2050 [1]. Hence, global food production will need to increase 38 and 57% by 2025 and 2050, respectively, to maintain the current level of food supply [2]. However, the world's irrigated land is decreasing by 1–2% annually [3], and soil degradation due to salinization is one of the major causes of this reduction

in irrigated land [4]. More than 1125 million hectares of land worldwide are salt-affected, of which approximately 76 million hectares are affected by human-induced salinization and sodification [4]. Salinity stress is a major abiotic stress affecting plant growth and crop productivity [5]. Soil salinization is a major cause of land degradation, and it can make land unsuitable for crop cultivation [6]. In recent years, biological measures have been shown to be some of the most effective approaches for ameliorating salt-affected soil [7].

E. angustifolia L., also known as Russian olive, is a deciduous tree belonging to the family *Elaeagnaceae* (Fig. 1). It is native to central and western Africa and is

*Correspondence: shenx@sdau.edu.cn

¹ College of Horticultural Science and Engineering/State Key Laboratory of Crop Biology, Shandong Agricultural University, Tai'an, China
Full list of author information is available at the end of the article



© The Author(s) 2022. **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 Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



Fig. 1 *E. angustifolia* in the south campus of Shandong Agricultural University, Tai'an, Shandong, China

distributed in the United States, Canada, the Mediterranean coast, southern Russia, Iran, and India. It is widely distributed in China and occurs in several provinces including Xinjiang, Gansu, Ningxia, and Shandong [8]. The fruit of *E. angustifolia* is rich in sugars, flavonoids, and other substances that can regulate the circulation of blood and immune function in humans; the branches, leaves, and flowers have anti-aging properties and can be used to treat burns, bronchitis, dyspepsia, and neurasthenia [9]. *E. angustifolia* is tolerant of drought, salt, cold, and wind stress. It is prolific and highly adaptable, as it can grow in a variety of climates and soils [10]. *E. angustifolia* can grow and reproduce normally in soil with a salt content of 0.8–1.2% [11]. As a nitrogen-fixing, actinorrhizal plant, *E. angustifolia* is likely an early successional, pioneer species that can colonize nitrogen-poor soils such as sandy, eroded mineral soils and wetlands [12]. Consequently, *E. angustifolia* has often been used for the reforestation of arid and salinized zones [13].

Although the development of genome sequencing has aided the domestication and improvement of many species [14], relatively few studies have used genomic tools to study *E. angustifolia*. Ghodhbane-Gtari et al. [15] reported the 11.3-Mbp draft genome sequence of *Frankia* sp. strain BMG5.11 in *E. angustifolia*, which had a G+C content of 69.9% and 9926 candidate protein-encoding genes. Lin et al. [16] conducted a genome-wide transcriptome analysis and found that high salt concentration inhibited the growth and photosynthesis of *E. angustifolia*, which was caused by the down-regulation of genes encoding key enzymes involved in photosynthesis and genes related to important structures in the photosystems and light-harvesting complexes. However, no studies to date have examined the reference genome of *E. angustifolia*. Furthermore, little is known about the mechanisms of adaptive evolution and transcription of *E. angustifolia* under salt stress.

Here, we used PacBio, Hi-C assisted assembly, resequencing, and other technologies to explore the adaptive evolution of *E. angustifolia*. We used the transcriptome to explore the mechanism by which *E. angustifolia* responds to salt stress. The results of this study provide new insights that could be used to aid the planting of *E. angustifolia*, increase food income, and promote recovery from global land desertification.

Materials and methods

Sample collection and DNA extraction

Samples from *E. angustifolia* trees (obtained from Xinjiang Province, NCBI Taxonomic ID, 36777) were collected from the south campus of Shandong Agricultural University, Tai'an, Shandong, China (36.17101°N, 117.16074°E, 134.0 masl) (Fig. 1) for genomic DNA sequencing, Hi-C assisted assembly, and genome evolution and transcriptome analyses. Twelve wild *E. angustifolia* samples were collected for genome resequencing in Gansu, China (Table 1). After collection, samples were immediately immersed in liquid nitrogen and stored until DNA extraction. DNA was extracted using the Cetyltrimethyl Ammonium Bromide (CTAB) method. The quality of the extracted genomic DNA was assessed using 1% agarose gel electrophoresis [17], and the concentration was quantified using a Qubit fluorimeter (Invitrogen Co., Carlsbad, CA, USA).

Genomic sequencing

The library was constructed after evaluating the quantity and quality of DNA samples. Two 270-bp paired-end libraries were prepared according to the Illumina protocol and sequenced on the Illumina HiSeq 2000 platform (Biomarker Technologies Co., Ltd., Beijing, China). Whole-genome sequencing was performed using the PacBio Sequel sequencer (Biomarker Technologies Co., Ltd., Beijing, China). Twelve 20-kb single-molecule

Table 1 Wild *E. angustifolia* samples used in this study for whole-genome resequencing

Population	Code ID	Sample ID	Group	Country	Province	City	Coordinates	Altitude (m)
<i>E. angustifolia</i>	R01	1583	G1	China	Gansu	Gulang	37°29'25"N; 102°54'37"E	2026.8
	R02	1619	G1	China	Gansu	Jinchang	38°21'25"N; 102°8'24"E	1704.6
	R03	1684	G1	China	Gansu	Linze	38°58'1"N; 99°54'52"E	1900.3
	R04	1697	G2	China	Gansu	Jia Yuguan	39°31'28"N; 98°51'8"E	1531.1
	R05	1726	G2	China	Gansu	Jiuquan	39°33'37"N; 98°51'37"E	1457.4
	R06	1758	G2	China	Gansu	Jiuquan	39°41'40"N; 98°35'55"E	1433.6
	R07	1784	G2	China	Gansu	Jiuquan	39°41'44"N; 98°35'57"E	1473.0
	R08	1811	G2	China	Gansu	Zhangye	39°41'42"N; 99°35'33"E	1450.1
	R09	1857	G3	China	Gansu	Zhangye	39°47'14"N; 98°14'36"E	1695.0
	R10	2063	G3	China	Gansu	Dunhuang	40°7'55"N; 94°54'48"E	1756.2
	R11	2106	G3	China	Gansu	Guazhou	40°22'3"N; 99°54'52"E	1133.0
	R12	2216	G3	China	Gansu	Lanzhou	36°29'21"N; 103°33'22"E	1902.8

The numbers (R01, R02, R03, R04, R05, R06, R07, R08, R09, R10, R11, R12) represent 12 wild *E. angustifolia* samples. G1, G2, and G3 were short for Group 1, 2, and 3, respectively

real-time sequencing (SMRT) bell libraries were constructed using a PacBio DNA Template Prep Kit 1.0 (Vazyme Biotech Co., Ltd., Nanjing, Jiangsu, China). The templates were size-selected, and the BluePippin devices (Sage Science, Inc., Annoron, Beijing, China) were used to enrich long snippets (> 10 kb). The PacBio Sequel platform processed a total of 39 single cells per molecule in real time (SMRT).

Genome assembly and genome annotation

A kmer map of $k=19$ was constructed using the two 270-bp libraries and used to evaluate genome size, repeat sequence ratio, and heterozygosity. The formula used was $G = (N \text{ k-mer} - N_{\text{error k-mer}}) / D$, where G is genome size, $N \text{ k-mer}$ is the number of k -mers, $N_{\text{error k-mer}}$ is the number of k -mers with the depth of 1, and D is the k -mer depth.

All low-quality sequences shorter than 500 bp were removed through the PacBio sequencing platform (Bio-marker Technologies Co., Ltd., Beijing, China). The results of Canu's [18] assembly and the wtdbg (<https://github.com/ruanjue/wtdbg>) assembly were merged using Quickmerge [19]. The homologous contigs of the merged genomes were determined by comparison of the two genome assemblies, and the short-read data were used to error correct the merged genome with Pilon [20].

State programs [21] were used to construct a database of repetitive sequences for the *E. angustifolia* genome. EVM v1.1.1 [20] software was then used to create a consensus repeat library. Genscan [22] and other programs were used for ab initio gene model prediction. GeMoMa v1.3.1 [23] was used for homology-based gene prediction. Hisat v2.0.4 [24] and Stringtie v1.2.3 [24] were used for assembly based on reference transcripts. TransDecoder

v2.0 [25] and GeneMarkS-T v5.1 [26] were used for gene prediction. PASA v2.0.2 [27] was used for the prediction of unigene sequences based on transcriptome data without a reference assembly. Finally, EVM v1.1.1 was used to integrate the prediction results obtained by the above three methods. Based on the Rfam [28] database and miRBase [29] database and Infernal 1.1 [30] for rRNA and microRNA prediction, tRNAscan-SE v1.3.1 [31] was used to identify tRNAs. BLAST v2.2.31 [22] with an E-value cutoff of $1E-5$ was used to align the predicted gene sequences with functional databases such as Gene Ontology (GO) [32] and Kyoto Encyclopedia of Genes and Genomes (KEGG) [33].

The Core Eukaryotic Gene Mapping Approach (CEGMA) v2.5 [34] database contains 458 conserved core genes in eukaryotes. CEGMA was used to evaluate the completeness of the final genome assembly. The embryophyta_odb9 database in BUSCO v2 [35] contains 1440 conserved core genes in terrestrial plants. We used BUSCO software to evaluate the completeness of the *E. angustifolia* genome assembly.

Hi-C analysis and assembly

Formaldehyde was used to fix the fresh *E. angustifolia* tissue samples. The DNA was digested with the restriction enzyme *HindIII*; after adding biotin-labeled bases, the repaired DNA was circularized, de-crosslinked, and purified, followed by fragmentation into 300–700-bp fragments. Streptavidin magnetic beads were used to capture the DNA fragments showing interaction relationships for library construction. Qubit2.0 (Life Technologies, CA, USA) and Agilent 2100 (Agilent Technologies) were used to detect the concentration and insert size of the library,

and qPCR was used to quantify the effective concentration of the library.

The Illumina high-throughput sequencing platform was used to sequence the Hi-C library to produce a large number of high-quality reads. BWA (version: 0.7.10-r789; alignment mode: aln; other parameters default) [36] was used to compare the sequencing paired-end data with the sequences of the assembled genome. The contigs of the draft genome were split into simulated 500-kb contigs, and LACHESIS software [37] was used to cluster these contigs into groups.

Genome evolution and gene family expansion

We used the single-copy protein sequences of seven other species, *Solanum lycopersicum*, *Arabidopsis thaliana* (Linn) Heynh, *Populus trichocarpa* Torr & Gray, *Glycine max* (Linn) Merr, *Oryza sativa* Linn, *Amborella trichopoda* Baill, and *Ziziphus jujuba* M, to build phylogenetic trees using PHYML software [38]. The divergence times between *E. angustifolia* and the seven other sequenced species were estimated using MrBayes60 and the Mcmctree [39] programs implemented in the Phylogenetic Analysis by Maximum Likelihood (PAML) 61 software package. Calibration times were obtained from the TimeTree database (<http://www.timetree.org/>) with '(*A. trichopoda*, *O. sativa*)' <199>173', (*A. thaliana*, *P. trichocarpa*)' <117>98', (*Z. jujuba*, *E. angustifolia*)' <117>79', and (*G. max*, *E. angustifolia*)' <113>89'. We then calculated the four-fold synonymous third-codon transversion (4DTv) values.

OrthoMCL [40] software was used to cluster the protein sequences of *E. angustifolia* and seven other sequenced species. CAFE [41] was used to analyze gene family contraction and expansion. Functional enrichment analysis was used to identify the function of genes over-represented in our genome assembly. GO enrichment analysis and KEGG annotation were performed using R scripts [42]. CodeML [43] was used in PAML to detect the rapidly evolving genes in *E. angustifolia* with a single copy shared by all species.

Genome resequencing

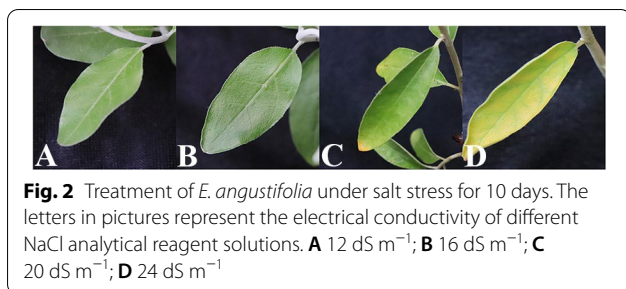
The code IDs of 12 wild *E. angustifolia* samples were R01–R12 (Table 1). The raw reads obtained by sequencing were filtered, low-quality reads with adapters were removed, and clean reads were obtained for subsequent information analysis. We used single nucleotide polymorphisms (SNPs) and small insertions and deletions (small Indels) to detect differences between our 12 wild *E. angustifolia* and the reference genome. The demographic history of 12 wild *E. angustifolia* was inferred using a hidden Markov model approach as implemented in the Pairwise Sequentially Markovian Coalescent (PSMC) model

based on the SNP distribution [44]. We scaled results to real time, assuming 2 years per generation and a neutral mutation rate of 7.31×10^{-9} (CI 95% Poisson distribution: $5.20 \times 10^{-9} \sim 8.00 \times 10^{-9}$) per generation [45]. Genes with SNPs and Indels were analyzed by functional annotation. Polymorphic genes in the 12 wild *E. angustifolia* samples were compared with BLAST, GO, KEGG, and other functional databases to evaluate their function. MEGA X [46] software was used to construct phylogenetic trees with the neighbor-joining method. Admixture [47] software was used to analyze the group structure of the samples. The number of subgroups (K value) was preset to 1–10 for clustering; the clustering results were cross-validated, and the optimal number of clusters was determined according to the valley value of the cross-validation error rate. We artificially divided the 12 wild *E. angustifolia* species into three groups (G1–G3) based on factors such as latitude and longitude of the sampling sites (Table 1), and we designated the outgroup *Z. jujuba* as G0. PCA was performed on 12 wild *E. angustifolia* species and *Z. jujuba*.

Salinity experiment

A greenhouse experiment was conducted in 2018 at State Key Laboratory of Crop Biology, Shandong Agriculture University, Tai'an, China. In March 2018, we mixed 200 seeds (collected from the *E. angustifolia* field in the breeding experimental base of Shandong Agricultural University) and wet sand into woven bags and placed them in an outdoor leeward shelter for lamination. We prepared 45 clay pots with an outer diameter of 29 cm, an inner diameter of 25 cm, and a depth of 18 cm. Each pot contained 15 kg of sterilized soil [48]. Three seeds were planted in each pot after one-third of the seeds had germinated in April 2018.

Irrigation of all pots was carried out for 2 months using normal urban water (salinity of 0.6 dS m^{-1}) at field capacity. When seedlings were well established two months after planting, four concentrations (electrical conductivity of 12 dS m^{-1} , 16 dS m^{-1} , 20 dS m^{-1} , and 24 dS m^{-1}) of NaCl analytical reagent (Keepway Technologies Corporation, Beijing, China) solution were used for the salt stress treatment; 5 pots of *E. angustifolia* were randomly selected for each treatment following the methods of Zeng et al. [49]. The other 25 pots were watered using normal urban water. Observations of the leaves of the four test groups were made every day to determine the level of salt stress (Fig. 2). Salinity stress continued for 10 days when the salinity toxicity in the leaves was over 20 dS m^{-1} . Salinity treatment was initiated with 25 pots by NaCl analytical reagent (Keepway Technologies Corporation, Beijing, China) solution with 20 dS m^{-1} . We randomly



divided 25 pots into 5 treatments (Group 1, 2, 3, 4, 5). The leaves of *E. angustifolia* in Group 1 were collected immediately (control group including T01&T02&T03); the leaves in Group 2 were collected 1 h after treatment (salinity group including T04&T05&T06); the leaves in Group 3 were collected 6 h after treatment (salinity group including T07&T08&T09); the leaves in Group 4 were collected 12 h after treatment (salinity group including T10&T11&T12); and the leaves in Group 5 were collected 24 h after treatment (salinity group including T13&T14&T15). All fresh leaf samples were immediately dissected and submerged in RNA later[®] solution and stored in a - 80 °C freezer for RNA extraction. During the experiment, the maximum temperature was 31.0 ± 5.0 °C, and the minimum temperature was 19.5 ± 4.5 °C.

RNA extraction, library construction, and RNA sequencing

The total RNA of each sample was extracted from the leaves of *E. angustifolia* using the RNA plant Plus Reagent (Vazyme Biotech Co., Ltd., Nanjing, Jiangsu, China). The RNA integrity and concentration were assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA). mRNA was isolated by the NEBNext Poly (A) mRNA Magnetic Isolation Module (NEB, E7490). The cDNA library was constructed using the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, E7530) and NEBNext Multiplex Oligos for Illumina (NEB, E7500) per the manufacturer's instructions. Briefly, the enriched mRNA was fragmented into approximately 200-nt RNA inserts, which were used to synthesize first-strand cDNA and second-strand cDNA. End-repair/dA-tail and adaptor ligation were performed on the double-stranded cDNA. The suitable fragments were isolated by Agencourt AMPure XP beads (Beckman Coulter, Inc.) and enriched by PCR amplification. Finally, the constructed cDNA libraries of *E. angustifolia* were sequenced on a flow cell using the Illumina HiSeq[™] sequencing platform.

Transcriptome analysis using reference genome-based reads mapping

Low-quality reads, such as those containing adaptors, with unknown nucleotides > 5%, or Q20 < 20% (percentage of sequences with sequencing error rates < 1%), were removed using perl scripts. The clean reads that were filtered from the raw reads were mapped to the *E. angustifolia* genome (OGSv3.2) using Tophat2 [50]. The aligned records from the aligners in BAM/SAM format were further examined to remove potential duplicate molecules. Gene expression levels were estimated using FPKM values (fragments per kilobase of exon per million fragments mapped) by Cufflinks software [51].

Identification of differentially expressed genes

DESeq2 [52] and Q-value were used to identify differentially expressed genes between Group 1 and Groups 2, 3, 4, and 5. Differences in gene abundance between these samples were calculated based on the ratio of the FPKM values. The false discovery rate (FDR) control method was used to identify the P-value threshold to evaluate the significance of differences. Here, only genes with an absolute value of log₂ ratio ≥ 2 and FDR significance score < 0.01 were used in subsequent analyses.

Sequence annotation

Differentially expressed genes were compared against various protein databases by BLASTX, including the National Center for Biotechnology Information (NCBI) non-redundant protein (Nr) database and Swiss-Prot database with a cut-off E-value of 10⁻⁵. Furthermore, genes were searched against the NCBI non-redundant nucleotide sequence (Nt) database using BLASTn with a cut-off E-value of 10⁻⁵. Genes were retrieved based on the best BLAST hit (highest score) along with their protein functional annotation.

The Nr BLAST results were imported into the Blast2GO program [53]. GO annotations for the genes were obtained by Blast2GO. In this analysis, all of the annotated genes were mapped to GO terms in the database, and the number of genes associated with each term was counted. A perl script was then used to plot the GO functional classification for the unigenes to visualize the distribution of gene functions. The obtained annotations were enriched and refined using the TopGo (R package). The gene sequences were also aligned to the Clusters of Orthologous Group (COG) database to predict and classify functions [54]. KEGG pathways were assigned to the assembled sequences by perl scripts.

Results and discussion

Genome assembly of *E. angustifolia*

One band was visible following 1% agarose gel electrophoresis, and the concentration of DNA extracted was approximately 474.3 mg L⁻¹. A total of 5,125,675 subreads were obtained by filtering low-quality data, and a total of 44.27 Gb raw PacBio sequel reads were obtained, with an average length of 8.64 kb (Additional file 1: Table S1). The subread N50 was 12,635 bp, and the average length was 8636 bp (Additional file 1: Table S2). After merging and correcting the assemblies, the final estimated genome size was 526.80 Mb, and the contig N50 was 12.60 Mb (Additional file 1: Table S3).

The total number of k-mers was 153,631,375,991, with a k-mers peak at a depth of 111 (Fig. 3A), and the final genome had a heterozygosity estimate of 1.47%. We constructed a specific repeat sequence database for the prediction of repeat sequences for specific species, and the prediction yielded approximately 263.44 Mb of repeats without overlap, accounting for 50.01% of the total length of repeat sequences (Additional file 1: Table S4). A total of 31,730 (Additional file 1: Table S5) protein-coding genes

and 127 miRNAs were annotated by integrating different methods (Additional file 1: Table S6). There were 30,771 genes available for transcriptome sequencing, accounting for 96.98% of all genes (Fig. 3B). A total of 96.89% of the genes could be annotated to NR and other databases (Additional file 1: Table S7). Conserved CEGMA analyses indicated that 97.38% of the core protein-coding genes were recovered in our assembled genome (Additional file 1: Table S8). BUSCO revealed 1290 complete gene models out of 1,440 (89.58%); 23 fragmented and 184 complete genes were found in duplicate (Additional file 1: Table S9).

Hi-C assisted assembly

We obtained 39.56 Gb clean reads, with a sequencing coverage of 75 × and Q30 ratio of >95.46% (Additional file 1: Table S10). The ratio of reads mapped to the assembled genome was 90.68%, and the ratio of unique mapped read pairs was 61.13%, indicating that the Hi-C data were suitable for subsequent analysis (Additional file 1: Table S11). A total of 80.79 M pairs of reads from the genome were obtained in this experimental library.

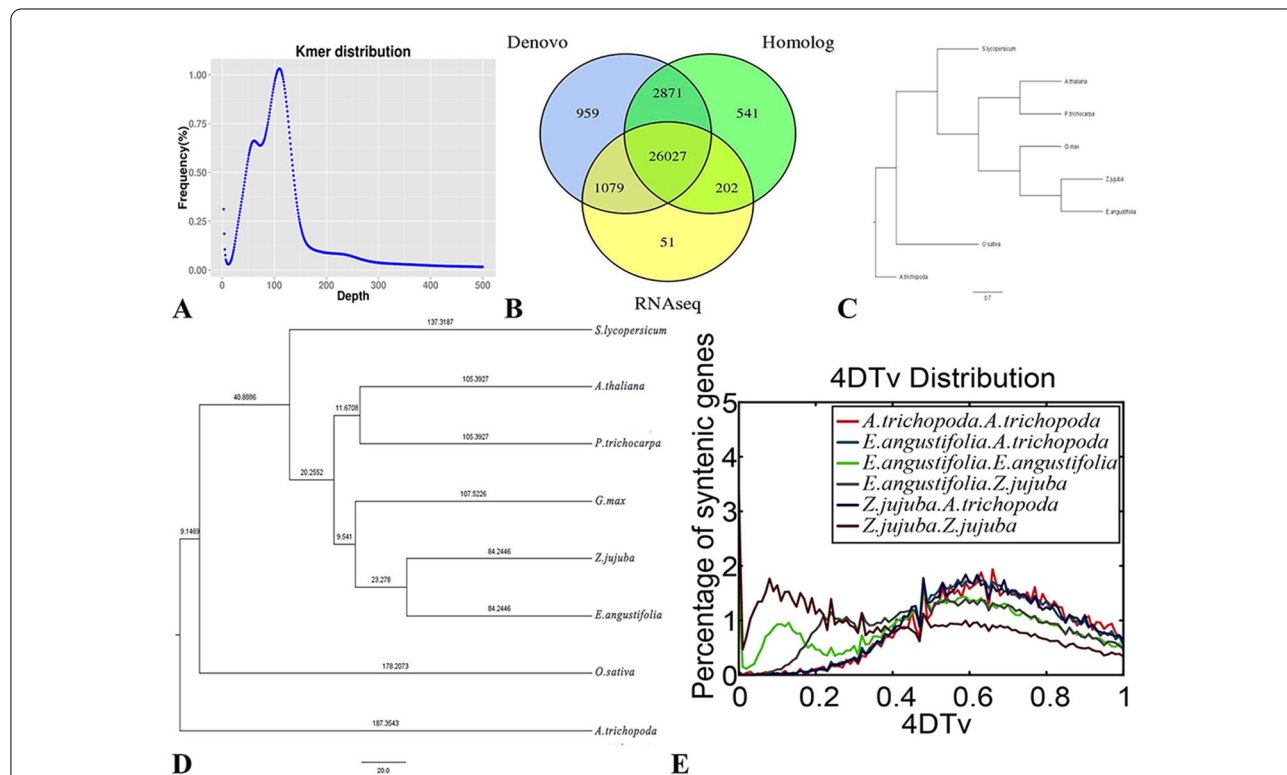


Fig. 3 Genome assembly and genome evolution. **A** K-mer distribution of genome sequencing reads of *E. angustifolia*. K-mers (K = 19) were extracted from the paired-end library with an insert size of 270 bp. The total 23-mer count is 51,064,317,165. The peak 19-mer depth is 111, and the genome size was calculated as 51,064,317,165/111 = 456.24 Mb. **B** The integrated gene is derived from the distribution map of three prediction methods. **C** Phylogenetic tree of eight species **D** Estimation of the divergence time between *E. angustifolia* and other species. The number in each branch is the divergent time from the present (in million years ago). **E** Distribution of 4DTV values among *Z. jujuba*, *E. angustifolia*, and *A. trichopoda* genomes

Among them, 72.98 M pairs were valid Hi-C data, accounting for 90.32% of the data from the genome, and the ratio of invalid interaction pairs was 9.68% (Additional file 1: Table S12).

E. angustifolia has 14 chromosomes. The genome was also visualized at the chromosome level. After Hi-C assembly, a total of 510.71 Mb of genomic sequences were mapped to the chromosomes, accounting for 96.94% of the total length of the sequences, and the corresponding number of sequences was 269, accounting for 45.83% of the total number of sequences. Among the sequences located on the chromosomes, the sequence length that could determine the order and direction of chromosomes was 473.91 Mb, accounting for 92.8% of the total length of the sequences located on the chromosomes, and the number of corresponding sequences was 104, accounting for 38.66% of the total number of sequences located on the chromosomes (Table 2).

Genome evolution of *E. angustifolia*

Comparative genome analysis harnesses the power of sequence comparisons within and between species to infer evolutionary history and provide information on the function of specific DNA sequences [55]. The deep-level evolutionary history of *E. angustifolia* has not yet been clarified. In our study, a phylogenetic tree (Fig. 3C, D) revealed that the origins of *A. trichopoda*, *O. sativa*, and *S. lycopersicum* were dated to 187.35,

178.21, and 137.32 million years ago (Mya), respectively. The origins of *A. thaliana* in the family *Brassicaceae* and *P. trichocarpa* in the family *Salicaceae* were dated to 137.32 Mya; *G. max* (107.52 Mya) in the family *Leguminosae* was in the same branch as *Z. jujuba* in the family *Rhamnaceae* and *E. angustifolia* in the family *Elaeagnaceae*. The origins of *Z. jujuba* and *E. angustifolia* were both dated to 84.24 Mya. These findings are similar to those of Harkess et al. [56].

To further analyze the evolutionary divergence of *E. angustifolia* and other species, 4DTV rates were calculated (Fig. 3E). The highest 4DTV value of *A. trichopoda* was 0.68, indicating that no recent genome-wide duplication has occurred. The 4DTV value peak of *Z. jujuba* was 0.08, and that of *E. angustifolia* was 0.12, indicating that both *Z. jujuba* and *E. angustifolia* have undergone recent genome-wide replication. Both 4DTV values of the autopolyploid of *Z. jujuba* and *E. angustifolia* peaked at 0.48, indicating that the time of the polyploid splitting event of *Z. jujuba* and *E. angustifolia* was similar to the time before the divergence of *Rosaceae* and *Elaeagnaceae*. The orthologs between *E. angustifolia* and *Z. jujuba* and between *E. angustifolia* and *A. trichopoda* indicated that the 4DTV values peaked at 0.25 and 0.48, respectively, suggesting that the divergence between *E. angustifolia* and *A. trichopoda* occurred earlier.

Expanded gene families related to stress adaptation in *E. angustifolia*

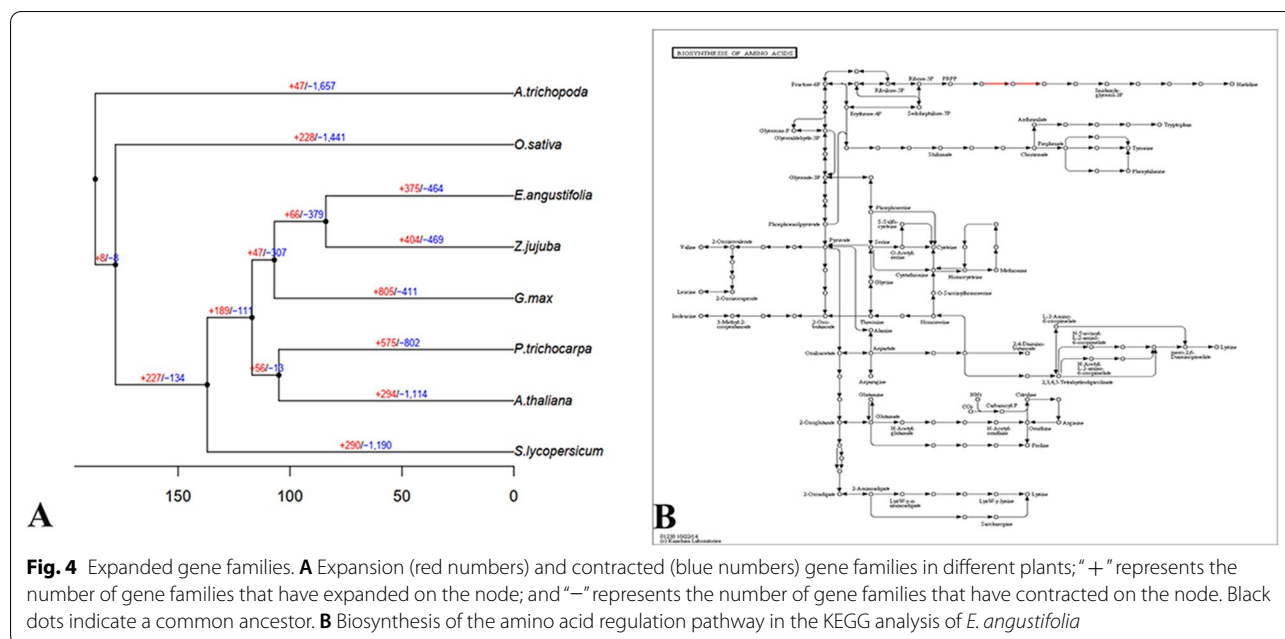
The identification of expanded gene families provides valuable insights into the biological innovation and adaptive evolution of *E. angustifolia*. A total of 27,553 of 31,730 genes of *E. angustifolia* could be classified into 13,309 gene families, of which 433 gene families were unique to *E. angustifolia* (Additional file 1: Table S13). Compared with *Z. jujuba*, *E. angustifolia* had fewer expanded gene families (375 vs. 404) and fewer contracted gene families (464 vs. 469) (Fig. 4A, Additional file 2: Table S14). Enrichment analysis of the 839 expanded and contracted gene families revealed that they were enriched in the pathways strictosidine synthase and xylanase inhibitor N-terminal. Strictosidine synthase has been detected in some major crops and is thought to make crops resistant to salt stress [57]. Xylanase inhibitor N-terminal is thought to be related to thermal stability [58]. The amplification of the above taxonomic genes might be related to the environmental conditions experienced by *E. angustifolia*. High temperature and drought have driven adaptive evolution in *E. angustifolia* [8] and increased evolutionary differences between *E. angustifolia* and other species.

The KEGG results revealed a scattered functional distribution of these genes, and 1 variant gene of the 330 genes was related to the biosynthesis of amino acids

Table 2 The genomic sequence of *E. angustifolia* mapped to the chromosomes

Group	Sequence count	Sequence length (bp)
Lachesis group 0	28	27,993,989
Lachesis group 1	41	99,096,873
Lachesis group 2	23	57,474,495
Lachesis group 3	12	48,501,719
Lachesis group 4	28	30,089,891
Lachesis group 5	7	26,462,114
Lachesis group 6	28	18,656,685
Lachesis group 7	10	24,577,875
Lachesis group 8	8	26,266,316
Lachesis group 9	8	24,034,829
Lachesis group 10	9	27,579,915
Lachesis group 11	19	32,396,042
Lachesis group 12	26	37,035,969
Lachesis group 13	22	30,539,364
Total sequences clustered (ratio %)	269 (45.83)	510,706,076 (96.94)
Total sequences ordered and oriented (ratio %)	104 (38.66)	473,912,606 (92.8)

The statistics were all based on sequence lengths of 1 Kb or more



(Additional file 1: Table S15). Further analysis of this pathway (Fig. 4B) revealed that the expression of the two pathways was increased during the synthesis of PRPP and imidazole-glycerol-3p. Previous studies indicate that both PRPP and imidazole-glycerol-3p are related to defense [59]. The GO (Additional file 1: Table S15) results showed that 28 variant genes were related to catalytic activity, 36 to metabolic process, and 29 to cellular process. Many variant genes were related to response to stimulus, which might indicate previous natural selection on *E. angustifolia* in response to harsh environments.

Drawing conclusions regarding the reduction in the adaptation of gene families was difficult given that we only analyzed gene families that were partly amplified from *E. angustifolia*. We identified 80 rapidly evolving genes in *E. angustifolia* through comparison of genes with 7 other species (Additional file 3: Table S16). The functional predictions showed that several genes were related to pseudouridine-metabolizing bifunctional protein and RNA pseudouridine synthase. Previous research has shown that pseudouridine-metabolizing bifunctional protein is related to human urinary tract pathogenic *Escherichia coli*, the principal agent of urinary tract infections in humans [60]. Genomic studies of humans and other mammals have shown that there is no gene encoding pseudouridine-metabolizing bifunctional protein, and the ability to metabolize pseudouridine has been lost [60]. This pattern of evolution in *E. angustifolia* might also be related to human and mammalian activities. The gene encoding carbon catabolite repressor protein 4 is also rapidly evolving in *E. angustifolia*. The carbon

catabolite repressor 4-CCR4 associated factor1 complex is the major enzyme complex responsible for catalyzing mRNA deadenylation. The degradation of messenger RNA caused by poly (A) tail shortening (deadenylation) is a central mechanism for the biological regulation of gene expression. This is a mechanism that plants have evolved to reprogram gene expression to maintain homeostasis in constantly changing environments [61]. Previous studies investigating carbon catabolite repressor 4-CCR4 associated factor1 function in plants have focused on the role of these genes in mediating biotic stress responses, such as resistance to pathogens [62]. This study has shown that the regulation of the gene encoding carbon catabolite repressor protein 4 has evolved rapidly in *E. angustifolia* to improve its tolerance to abiotic and biotic stress [63]. This is also the first study showing that a gene regulating carbon catabolite repressor protein 4 has been discovered in *E. angustifolia*. Genes related to biotic and abiotic stress resistance were also among some of the 80 rapidly evolving genes, such as peptidyl-prolyl cis-trans isomerase [64].

Resequencing analysis of 12 wild *E. angustifolia* samples in Gansu Province, China

The resequencing analysis of 12 wild *E. angustifolia* samples in Gansu Province, China (Table 1, Additional file 6: Fig. S1) generated 86.58 Gp clean reads, with a Q20 of 96.64% and Q30 of 91.79%. The average coverage depth was 11 ×, and the genome coverage was 92.89% (Additional file 1: Tables S17–S19). In diploid genomes, runs of homozygosity are uninterrupted homozygous segments

in the genome [65], and they can be used to quantify the level of inbreeding either in individuals or populations [66]. In this study, homozygous types accounted for most of the other 11 individuals except for R11 (Table 3). This suggests that hybridization might occur between different close relatives in nature, which is consistent with the observation that hybridization is common in plants [67]. We also detected 74,790 CDS and 1,791,719 genome-wide Indels (Additional file 1: Table S20).

The Pairwise Sequentially Markovian Coalescent (PSMC) model was used to estimate the historical effective population size based on genome-wide data of 12 wild *E. angustifolia* species (Fig. 5A). In this study, there was no significant difference in the effective population size among 12 wild *E. angustifolia* species 3.0 Mya, which indicated that there were no evolutionary differences among the 12 *E. angustifolia* 3.0 Mya. This was also similar to the time of differentiation of *A. trichopoda*, the sister of all angiosperms [56]. From 3.0 Mya to 350 thousand years ago (Kya), the effective population size of 12 wild *E. angustifolia* changed slightly. Notably, from 1.5 Mya to 150 Kya, the species occurred in Africa and showed no signs of divergence. The effective population size of the genome data of each race in this period was the same [44]. The effective population size of 12 wild *E. angustifolia* changed greatly from 350 to 23 Kya. The largest change was observed in the effective population size of wild *E. angustifolia* R11 (2106), especially in the period 55–35 Kya. The effective population size reached a maximum of approximately 60×10^4 . At this time, the global temperature continued to rise, which might promote the rapid growth of the population of *E. angustifolia* [68]. At

23 Kya, the effective population size of 12 wild *E. angustifolia* decreased drastically, which corresponded to the last glacial period (the duration is approximately 26.5 to 19 Kya). The global temperature continued to decrease, and the glaciers at the poles and mountains began to extend to low latitudes and altitudes [69]. This drastic change in the environment might have led to a reduction in the effective population size of *E. angustifolia*.

We used the R01 result to evaluate the function of the mutated genes (Additional file 1: Tables S21, S22, Fig. 5B, C). Substitutions mainly occurred in DNA binding (GO:0003677), Cellular Component: Nucleus (GO:0005634), DNA binding (GO:0003677) related to cytochrome [70], and Cellular Component: Nucleus (GO:0005634) related to defensive T cells [71]. This variation might be related to the environmental conditions experienced by *E. angustifolia*, including water shortages at high altitude [72].

The results of the phylogenetic tree (Fig. 5D) and PCA (Fig. 5E) were not consistent with expectation, with the exception of G1. This might be related to climate change [73] and landform change [74]. Alternatively, differences in the altitude, latitude, and longitude of the sampling sites might explain this pattern (Table 1); various other factors might cause the results to deviate from expectation.

Genetic mechanisms underlying salt tolerance in *E. angustifolia*

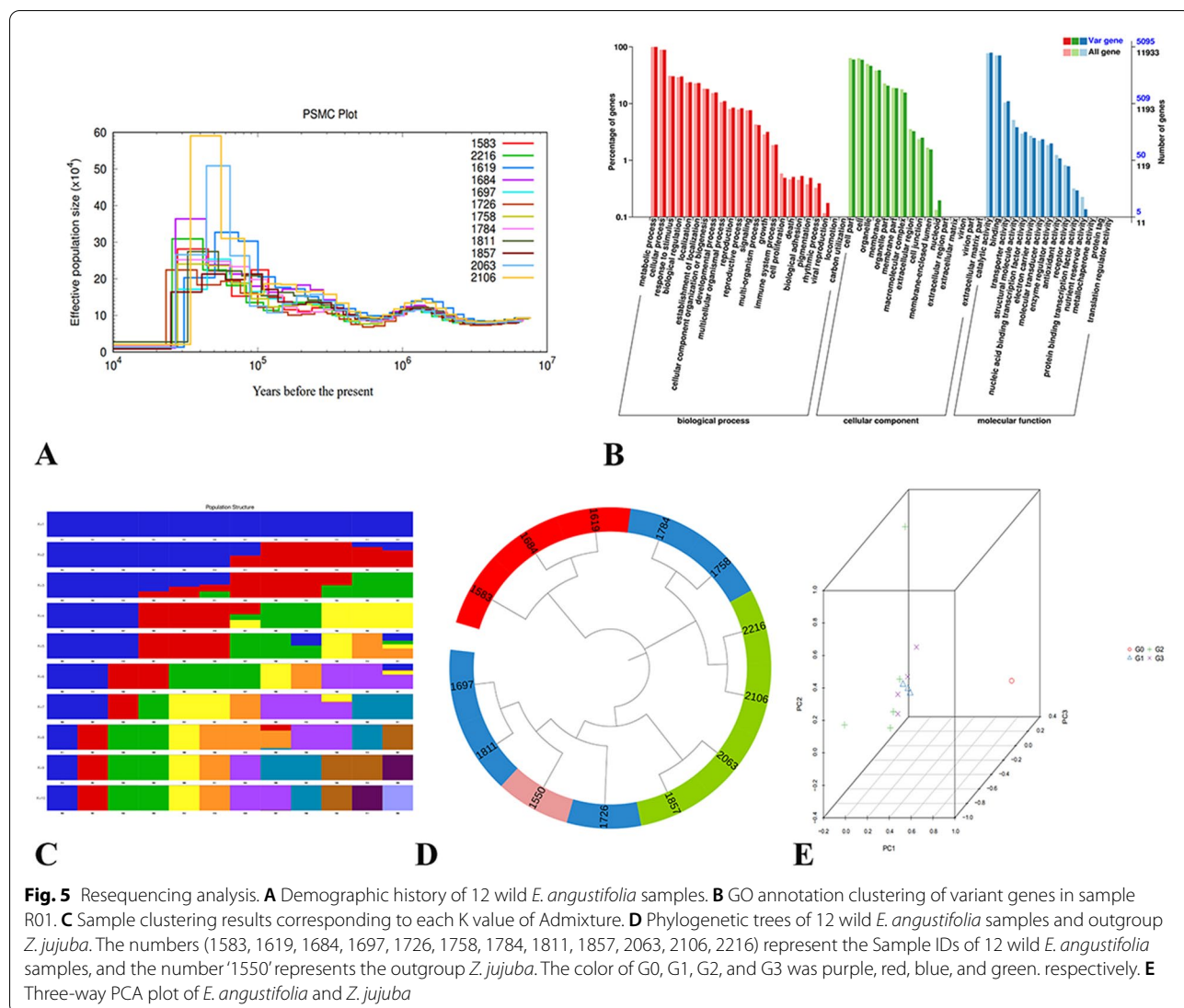
To investigate the genetic mechanisms underlying salt tolerance in *E. angustifolia*, we performed a salinity experiment. We obtained 113.85 Gb clean reads. There

Table 3 Number of SNPs for wild *E. angustifolia* samples

Code ID	Sample ID	SNP number	Transition	Transversion	Ti/Tv	Heterozygosity	Homozygosity	Het-ratio (%)
R01	1583	3,097,175	2,110,179	986,996	2.13	1,412,964	1,684,211	45.62
R02	1619	3,223,761	2,197,723	1,026,038	2.14	1,492,655	1,731,106	46.30
R03	1684	3,094,736	2,108,611	986,125	2.13	1,417,978	1,676,758	45.81
R04	1697	3,021,352	2,057,327	964,025	2.13	1,417,892	1,603,460	46.92
R05	1726	3,045,841	2,075,585	970,256	2.13	1,256,764	1,789,077	41.26
R06	1758	3,193,315	2,178,336	1,014,979	2.14	1,442,262	1,751,053	45.16
R07	1784	3,225,579	2,198,583	1,026,996	2.14	1,405,876	1,819,703	43.58
R08	1811	3,108,956	2,117,042	991,914	2.13	1,537,532	1,571,424	49.45
R09	1857	3,022,369	2,058,271	964,098	2.13	1,311,017	1,711,352	43.37
R10	2063	3,028,547	2,060,171	968,376	2.12	1,379,687	1,648,860	45.55
R11	2106	3,223,275	2,197,798	1,025,477	2.14	1,617,799	1,605,476	50.19
R12	2216	3,115,254	2,122,083	993,171	2.13	1,369,366	1,745,888	43.95

In a diploid, if a certain SNP site on homologous chromosomes is the same base, the SNP site is referred to as a homozygous SNP site. If the SNP loci on homologous chromosomes contain different types of bases, the SNP locus is referred to as a heterozygous SNP locus. Greater numbers of homozygous SNPs correspond to greater differences between samples and the reference genome. Greater numbers of heterozygous SNPs correspond to higher degrees of heterozygosity

number number of SNPs detected; *Transition* number of SNPs of the conversion type; *Transversion* number of SNPs of transversion type; *Ti/Tv* the ratio of SNPs of conversion to transversion type; *Heterozygosity* number of heterozygous SNPs; *Homozygosity* number of homozygous SNPs; *Het-ratio* the ratio of heterozygous SNPs



were a total of 6.40 Gb clean reads for each sample, and the Q30 base percentage was 90.74% and higher (Additional file 1: Table S23). The clean reads of each sample were aligned to the reference genome, and the alignment efficiency ranged from 90.87 to 92.30% (Additional file 1: Table S24). Based on the comparison results, we carried out an alternative splicing prediction analysis and gene structure optimization analysis. We also compared the genome sequencing results of *E. angustifolia* and identified 4,404 new genes in the transcriptome of *E. angustifolia* (Additional file 4: Table S25), and we obtained 3953 functional annotations (Additional file 1: Table S26). We analyzed the GO pathways functions of the new genes (Additional file 5: Table S27). The results showed that 3391 genes were involved in the regulation of Molecular Function, 3121 genes were involved

in the regulation of Biological Process, and 2613 genes were involved in regulation of Cellular Component. In Molecular Function, 256 genes were involved in regulating ATP binding (GO:0005524), 158 genes were involved in regulating metal ion binding (GO:0046872), and 131 genes were involved in regulating zinc ion binding (GO:0008270), 106 genes were involved in regulating the nucleic acid binding (GO:0003676), and 104 genes were involved in regulating the structural constituent of ribosome (GO:0003735). Previous studies have shown that AtABC36-overexpressing plants are more resistant to drought and salt stress [75], overexpression of metal ion binding peptides–phytochelatins and metallothionein genes increases tolerance to stress [76], and overexpression of GsZFP1 (zinc ion binding) in alfalfa increases salt tolerance [77]. ‘Nucleic acid binding’ was the most

significantly enriched GO term in the MF category for both Supreme and Parish up-regulated genes under salt treatment, suggesting that this process might play an important role in salt tolerance in both cultivars [78]; the structural constituent of ribosome has been shown to play an important role in salinity tolerance [79]. In Biological Process, there were 236 genes involved in the regulation of oxidation–reduction process (GO:0055114), 108 genes involved in the regulation of protein phosphorylation (GO:0006468), and 94 genes involved in the regulation of transcription, DNA-templated (GO:0006355). Previous transcriptome profiling experiments have indicated that oxidation–reduction processes, protein phosphorylation, and DNA-templated are related to salt tolerance in plants [80–83]. Xiong et al. [84] suggested that the homeostasis of the oxidation–reduction process is important for salt tolerance in plants. Hsu et al. [85] identified novel salt stress-responsive protein phosphorylation sites from membrane isolates of abiotic-stressed plants by membrane shaving followed by Zr4+ -IMAC enrichment, and the identified phosphorylation sites play an important role in the salt stress response in plants. Wu et al. [78] showed that genes that were down-regulated under salt treatment are involved in “regulation of transcription, DNA-templated.” In Cellular Component, there were 668 genes involved in the regulation of the integral component of membrane (GO:0016021), 201 genes involved in the regulation of nucleus (GO:0005634), 153 genes in the regulation of membrane (GO:0016020), 116 genes involved in regulating cytoplasm (GO:0005737), and 101 genes involved in regulating plasma membrane (GO:0005886). In conclusion, the GO pathways functions of the new genes in *E. angustifolia* showed that the functions of the new genes in the Molecular Function and Biological Process categories were mainly involved in the regulation of salt stress.

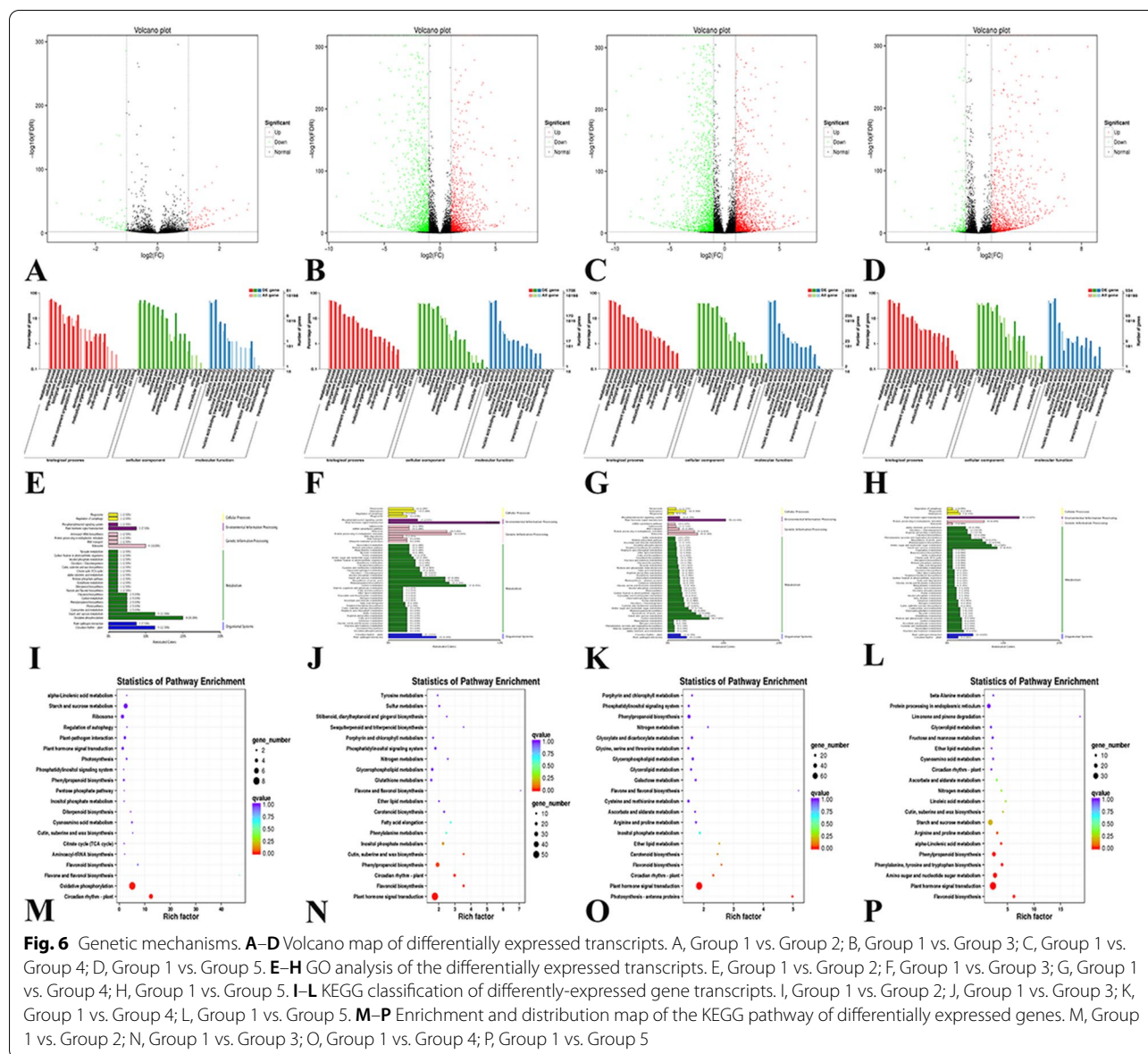
According to Additional file 1: Table S28 and Fig. 6A–D, the number of transcripts with significant expression differences between control Group 1 and experimental Group 2 was 137. The expression of 82 and 55 transcripts was significantly up-regulated and down-regulated, respectively. Similarly, the number of transcripts with significant expression differences between control Group 1 and experimental Group 3 was 2670, of which 1260 were up-regulated and 1410 were down-regulated. The number of transcripts with significant expression differences between control Group 1 and experimental Group 4 was 3619, of which 1668 were up-regulated and 1951 were down-regulated. The number of transcripts with significant expression differences between control Group 1 and experimental Group 5 was 1404, of which 1193 were up-regulated and 211 were down-regulated. A large number of differentially expressed transcripts in the two

databases indicate that salt stress induces a large number of gene expression changes in *E. angustifolia*, which reflects the complexity of the mechanism underlying the response of *E. angustifolia* to salt stress.

The GO pathway annotations (Fig. 6E–H) of the four experimental groups compared with the control group revealed a pattern consistent with the differential expression among the four experimental groups, and salt stress had a significant effect on proteins in the functional categories Biological Process, Molecular Function, and Cellular Component [86]; the effects on different physiological processes also differed. After salt stress, metabolic process, cellular process, and single-organism process were greatly affected within Biological Process, and binding and catalytic activity were greatly affected within Molecular Function. Salt stress affected the expression of related genes during growth and development as well as the activity of certain proteins [87], which had a greater impact on the growth and development of *E. angustifolia*.

We also performed KEGG pathway analysis of differentially expressed gene transcripts (Fig. 6I–L). The metabolic process and signal pathways that regulate and change under salt stress in plant bodies could be clearly observed according to the KEGG pathway analysis. Comparison of the experimental group and control group revealed that these changes were mostly concentrated in sugar, amino acid metabolism, and plant hormone signal transduction. Salt stress had obvious effects on Environmental Information Process, Genetic Information Processing, and Metabolism and Organismal Systems. After 1 h of salt stress, there were significant changes in oxidative phosphorylation, ribosome, and starch and sucrose metabolism, and KEGG pathways (Fig. 6M) were enriched in oxidative phosphorylation and circadian rhythm-plant. After 6 h and 12 h of salt stress, there were significant changes in plant hormone signal transduction, protein processing in endoplasmic reticulum, biosynthesis of amino acids, phenylpropanoid biosynthesis, carbon metabolism, and plant-pathogen interaction, and KEGG pathways (Fig. 6N, O) were enriched in plant hormone signal transduction. After 24 h of salt stress, there were significant changes in amino sugar and nucleotide sugar metabolism and other metabolic processes, and KEGG pathways (Fig. 6P) were enriched in plant hormone signal transduction.

Several studies have shown that partial KEGG pathway changes are related to salt resistance. Circadian rhythms synchronize intracellular calcium dynamics and ATP production for growth [88] and have a large effect on plant immunity during plant–pathogen interactions [89]. Oxidative phosphorylation plays a role in plant salt tolerance by coordinating ROS scavenging pathways to regulate intracellular ROS levels, prevent cell damage,



and control ROS signal transduction [90]. The significance of compatible solutes such as sugars, amino acids, and tertiary amines in salt stress has been well documented [91]. They not only provide an essential source of energy and nutrients for plants under salt stress but also act as osmotic adjustment substances to balance the osmotic potential appended by high salinity [92]. The accumulation of soluble sugars and sucrose can improve salt tolerance [93]. The loss of integrity of the ribosome through the removal of a putative ribosome maturation factor or a ribosomal protein confers salt tolerance to *E. coli* cells [94]. The endoplasmic reticulum is associated with salt tolerance in tomato [95]. Metabolic adaptation is crucial for abiotic stress resistance in plants, and the

accumulation of specific amino acids has been suggested to increase tolerance to salt [96]. Enrichment of carbon metabolism and the biosynthesis of amino acids suggests that the synthesis of compatible solutes is fortified to alleviate osmotic stress in plant seedlings under salt treatment [26]. Variation in gene expression under salt stress is also regulated by phytohormones [26]. Changes in the transcription level of hormone genes affect drought and salt stress in plants [97]. de Bruxelles et al. [98] showed that hormones are responsible for changes in the expression of salt-induced genes.

Conclusions

In this study, the genome of *E. angustifolia* L. was obtained using PacBio technology and Hi-C assisted assembly technology. We estimated the origin of *E. angustifolia* and evaluated its evolutionary relationships with 8 other species. We used comparative genomics to study the adaptive evolution of *E. angustifolia* in Gansu, China. Genes and pathways of salt resistance of *E. angustifolia* were identified through transcriptome analysis. Our findings could be used to aid future comparative genomic analyses of *E. angustifolia* and enhance our understanding of the response of *E. angustifolia* to drought, salt, cold, and wind stress. Our findings also have implications for the planting of *E. angustifolia* and recovery from global land desertification. Several limitations of our study require consideration, given that the findings might be affected by the analytical method used, test conditions, and other environmental conditions; follow-up studies are needed to confirm our findings.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13007-022-00915-w>.

Additional file 1: Table S1. Filtering raw data of Pac-bio sequencing. **Table S2.** Length distribution of subreads of Pac-bio sequencing. **Table S3.** Genome assembly evaluation statistics. **Table S4.** Repeating sequence statistics. **Table S5.** Gene information statistics. **Table S6.** Non-coding RNA information. **Table S7.** Gene function annotation statistics. **Table S8.** Statistics of the completeness of the assembled genome. **Table S9.** Statistics of the BUSCO of the assembled genome. **Table S10.** Sequencing data volume statistics. **Table S11.** Clean data and genome alignment results statistics. **Table S12.** Hi-C sequencing data validation. **Table S13.** Classification statistics of gene families. **Table S15.** Function prediction of *E. angustifolia*. **Table S17.** Evaluation statistics of sequencing data of wild *E. angustifolia* samples. **Table S18.** Comparison results of wild *E. angustifolia* samples. **Table S19.** Coverage depth and coverage ratio statistics of wild *E. angustifolia* samples. **Table S20.** InDel statistics of whole genome and coding region of wild *E. angustifolia* samples. **Table S21.** Statistical table of variation genes of wild *E. angustifolia* samples. **Table S22.** Annotated list of some variant genes in sample R01. **Table S23.** Statistics of transcriptome sequencing data. **Table S24.** Results of the comparison between the sequencing transcriptome samples and the genome data. **Table S26.** Statistics of functional annotation results of new genes. **Table S28.** Statistics on the number of different genes in the transcription.

Additional file 2: Table S14. Functional gene categories enriched for *E. angustifolia* expansion and contraction families.

Additional file 3: Table S16. 80 rapidly evolving genes in *E. angustifolia*.

Additional file 4: Table S25. New genes discovered in *E. angustifolia* partially.

Additional file 5: Table S27. Go pathways functions of new genes.

Additional file 6: Figure S1. 12 wild *E. angustifolia* samples.

Acknowledgements

Not applicable.

Author contributions

YM and XShen planned and designed the research; YM, XLCui, HW, XQ, YL, YY, XSu, JT and FW performed experiments, conducted fieldwork, analyzed data; YM, FM, ND, DZ, YH, WW, SW, XLChen, ZM, XSC and XShen wrote the manuscript. Every author contributed equally. All authors read and approved the final manuscript.

Funding

This study was supported by Project supported by the National Natural Science Foundation of China (32072520); Fruit innovation team project of Shandong Province (CN) (SDAIT-06-07); Natural Science Foundation of Shandong Province (CN) (ZR2020MC132) and Industrialization project of improved varieties in Shandong Province (CN) (2019LZGC007).

Availability of data and materials

This Whole Genome Shotgun project has been deposited in NCBI. Raw sequencing reads and genome assembly are available at GenBank as BioProject PRJNA647537 (<https://dataview.ncbi.nlm.nih.gov/object/PRJNA647537?reviewer=5dse95jg6ae9tao2s0ettacev>). Raw sequencing data (Hi-C, trans data, survey, PB) have been deposited in SRA (Sequence Read Archive) database as SRR12563589–SRR12563593 [99–102]. Resequencing has been deposited in SRA (Sequence Read Archive) database as SRR12578558–SRR12578569 [103]. Transcriptome has been deposited in SRA (Sequence Read Archive) database as SRR12569921–SRR12569935 [104]. Gene annotation data of *E. angustifolia* genome has been deposited in FigShare (<https://doi.org/10.6084/m9.figshare.12957110.v1>).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹College of Horticultural Science and Engineering/State Key Laboratory of Crop Biology, Shandong Agricultural University, Tai'an, China. ²Biomarker Technologies Corporation, Beijing, China. ³College of Horticulture, Northwest Agriculture and Forestry University, Yangling, China. ⁴Germplasm Resource Center of Shandong Province, Shandong Academy of Agricultural Sciences, Jinan, China. ⁵Department of Horticulture, University of Georgia, Athens, USA.

Received: 9 February 2022 Accepted: 26 May 2022

Published online: 18 June 2022

References

- UN News Centre. World population projected to reach 9.6 billion by 2050—UN report. 2013. <http://www.un.org/en/development/desa/news/population/2015-report.html>. Accessed 30 Jul 2015.
- Wild A. Soils, land and food: managing the land during the twenty-first century. Cambridge: Cambridge University Press; 2003. p. 256.
- FAO. World Agricultural Center, FAOSTAT agricultural statistic database gateway. Rome: FAO; 2014.
- Hossain MS. Present scenario of global salt affected soils, its management and importance of salinity research. *Int Res J Biol Sci.* 2019;1:1–3.
- Gupta B, Huang BR. Mechanism of salinity tolerance in plants: physiological, biochemical, and molecular characterization. *Int J Genom.* 2014;2014: 701596. <https://doi.org/10.1155/2014/701596>.
- Xiao Y, Zhao G, Li T, Zhou X, Li J. Soil salinization of cultivated land in Shandong Province, China—dynamics during the past 40 years. *Land Degrad Dev.* 2019;30:426–36.

7. Stoto MA. Population health measurement: applying performance measurement concepts in population health settings. *EGEMS*. 2014;2:1132. <https://doi.org/10.13063/2327-9214.1132>.
8. Wang BS, Qu HY, Ma J, Sun XL, Wang D, Zheng QS. Protective effects of *elaeagnus angustifolia* leaf extract against myocardial ischemia/reperfusion injury in isolated rat heart. *J Chem*. 2014;3: 693573. <https://doi.org/10.1155/2014/693573>.
9. Vitas AI, Garcia-Jalon VAEI. Occurrence of *Listeria monocytogenes* in fresh and processed foods in Navarra (Spain). *Int J Food Microbiol*. 2004;90:349–56. [https://doi.org/10.1016/S0168-1605\(03\)00314-3](https://doi.org/10.1016/S0168-1605(03)00314-3).
10. Klich MG. Leaf variations in *Elaeagnus angustifolia* related to environmental heterogeneity. *Environ Exp Bot*. 2000;44:171–83. [https://doi.org/10.1016/S0098-8472\(00\)00056-3](https://doi.org/10.1016/S0098-8472(00)00056-3).
11. Li LX, Zhu TT, Liu J, Zhao C, Li YY, Chen M. An orthogonal test of the effect of NO_3^- , PO_4^{3-} , K^+ , and Ca^{2+} on the growth and ion absorption of *Elaeagnus angustifolia* L. seedlings under salt stress. *Acta Physiol Plant*. 2019;41:179. <https://doi.org/10.1007/s11738-019-2969-8>.
12. Caru M, Mosquera G, Bravo L, Guevara R, Sepulveda D, Cabello A. Infectivity and effectivity of *Frankia* strains from the Rhamnaceae family on different actinorhizal plants. *Plant Soil*. 2003;251:219–25.
13. Zhang X, Huang G, Bian X, Zhao Q. Effects of root interaction and nitrogen fertilization on the chlorophyll content, root activity, photosynthetic characteristics of intercropped soybean and microbial quantity in the rhizosphere. *Plant Soil Environ*. 2013;59:80–8. <https://doi.org/10.17221/613/2012-PSE>.
14. Zhang X, Liu L, Chen B, Qin Z, Xiao Y, Zhang Y, Yao R, Liu H, Yang H. Progress in understanding the physiological and molecular responses of populus to salt stress. *Int J Mol Sci*. 2019;20:1312. <https://doi.org/10.3390/ijms20061312>.
15. Ghodhbane-Gtari F, Swanson E, Gueddou A, Simpson S, Morris K, Thomas WK, Gtari M, Tisa LS. Draft genome sequence for *Frankia* sp. strain BMG5.11, a nitrogen-fixing bacterium isolated from *Elaeagnus angustifolia*. *Microbiol Resour Ann*. 2020. <https://doi.org/10.1128/MRA.00824-20>.
16. Lin J, Li JP, Yuan F, Yang Z, Wang BS, Chen M. Transcriptome profiling of genes involved in photosynthesis in *Elaeagnus angustifolia* L. under salt stress. *Photosynthetica*. 2018;56:998–1009. <https://doi.org/10.1007/s11099-018-0824-6>.
17. Aboul-Maaty NAF, Oraby HAS. Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. *Bull Natl Res Cent*. 2019;43:25. <https://doi.org/10.1186/s42269-019-0066-1>.
18. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhou B, Flavell A, Leroy P, Morgante M, Panaud O. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet*. 2007;8:973–82. <https://doi.org/10.1038/nrg2165>.
19. Tarailo-Graovac M, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinform*. 2009. <https://doi.org/10.1002/0471250953.bi0410s25>.
20. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Genome Biol*. 2008;9:R7. <https://doi.org/10.1186/gb-2008-9-1-r7>.
21. Xu Z, Wang H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res*. 2007;35:265–8. <https://doi.org/10.1093/nar/gkm286>.
22. Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA. *J Mol Biol*. 1997;268:78–94. <https://doi.org/10.1006/jmbi.1997.0951>.
23. Keilwagen J, Hartung F, Grau J. GeMoMa: homology-based gene prediction utilizing intron position conservation and RNA-seq data. *Methods Mol Biol*. 2019;1962:161–77. https://doi.org/10.1007/978-1-4939-9173-0_9.
24. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc*. 2016;11:1650–67. <https://doi.org/10.1038/nprot.2016.095>.
25. Wang JP, Tian SL, Sun XL, Cheng XC, Duan NB, Tao JH, Shen GN. Construction of pseudomolecules for the Chinese chestnut (*Castanea mollissima*) genome. *G3 Genes Genom Genet*. 2020;10:3565–74. <https://doi.org/10.1534/g3.120.401532>.
26. Tang S, Lomsadze A, Borodovsky M. Identification of protein coding regions in RNA transcripts. *Nucleic Acids Res*. 2015;43: e78. <https://doi.org/10.1093/nar/gkv227>.
27. Campbell MA, Haas BJ, Hamilton JP, Mount SM, Buell CR. Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. *BMC Genom*. 2006;7:327. <https://doi.org/10.1186/1471-2164-7-327>.
28. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A. Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res*. 2005;33:121–4. <https://doi.org/10.1093/nar/gki081>.
29. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*. 2006;34:140–4. <https://doi.org/10.1093/nar/gkj112>.
30. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*. 2013;29:2933–5. <https://doi.org/10.1093/bioinformatics/btt509>.
31. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997;25:0955–64. <https://doi.org/10.1093/nar/25.5.955>.
32. Blanco E, Parra G, Guigó R. Using geneid to identify genes. *Curr Protoc Bioinform*. 2007. <https://doi.org/10.1002/0471250953.bi0403s18>.
33. Korf I. Gene finding in novel genomes. *BMC Bioinform*. 2004;5:59. <https://doi.org/10.1186/1471-2105-5-59>.
34. Parra G, Bradnam K, Korf I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics*. 2007;23:1061–7. <https://doi.org/10.1093/bioinformatics/btm071>.
35. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31:3210–2. <https://doi.org/10.1093/bioinformatics/btv351>.
36. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
37. Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat Biotechnol*. 2013;31:1119–25. <https://doi.org/10.1038/nbt.2727>.
38. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science*. 2002;295:1306–11. <https://doi.org/10.1126/science.1067799>.
39. Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet*. 2013;14:390–403. <https://doi.org/10.1038/nrg3454>.
40. Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res*. 2003;13:2178–89. <https://doi.org/10.1101/gr.1224503>.
41. de Bie T, Cristianini N, Demuth JP, Hahn MW. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics*. 2006;22:1269–71. <https://doi.org/10.1093/bioinformatics/btl097>.
42. Götz S, Arnold R, Sebastián-León P, Martín-Rodríguez S, Tischler P, Jehl MA, Dopazo J, Rattei T, Conesa A. B2G-FAR, a species-centered GO annotation repository. *Bioinformatics*. 2011;27:919–24. <https://doi.org/10.1093/bioinformatics/btr059>.
43. Schabauer H, Valle M, Pacher C, Stockinger H, Stamatakis A, Robinson-Rechavi M, Yang ZH, Salamin N. SlimCodeML: an optimized version of CodeML for the branch-site model. *Piscataway: 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum*; 2012. p. 706–14.
44. Li H, Durbin R. Inference of human population history from individual whole-genome sequences. *Nature*. 2011;475:493–6. <https://doi.org/10.1038/nature10231>.
45. Krasovec M, Chester M, Ridout K, Filatov DA. The mutation rate and the age of the sex chromosomes in *Silene latifolia*. *Curr Biol*. 2018;28:1832–8. <https://doi.org/10.1016/j.cub.2018.04.069>.
46. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35:1547–9. <https://doi.org/10.1093/molbev/msy096>.
47. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19:1655–64. <https://doi.org/10.1007/s00262-014-1555-6>.

48. Wang YF, Fu FY, Li JJ, Wang GS, Wu MM, Zhan J, Chen XS, Mao ZQ. Effects of seaweed fertilizer on the growth of *Malus hupehensis* Rehd. seedlings, soil enzyme activities and fungal communities under replant condition. *Eur J Soil Biol.* 2016;75:1–7. <https://doi.org/10.1016/j.ejsobi.2016.04.003>.
49. Zeng L, Tu XL, Dai H, Han FM, Lu BS, Wang MS, Nanaei HA, Tajabadipour A, Mansouri M, Li XL, Ji LL, Irwin DM, Zhou H, Liu M, Zheng HK, Esmaili-zadeh A, Wu DD. Whole genomes and transcriptomes reveal adaptation and domestication of pistachio. *BMC Genome Biol.* 2019;20:79. <https://doi.org/10.1186/s13059-019-1686-3>.
50. Kim D, Perrea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 2013;14:R36. <https://doi.org/10.1186/gb-2013-14-4-r36>.
51. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010;28:511–5. <https://doi.org/10.1038/nbt.1621>.
52. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
53. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics.* 2005;21:3674–6. <https://doi.org/10.1093/bioinformatics/bti610>.
54. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 2000;28:33–6. <https://doi.org/10.1093/nar/28.1.33>.
55. Li F, Bian L, Ge J, Han F, Liu Z, Li X, Liu Y, Lin Z, Shi H, Liu C, Chang Q, Lu B, Zhang S, Hu J, Xu D, Shao C, Chen S. Chromosome-level genome assembly of the East Asian common octopus (*Octopus sinensis*) using PacBio sequencing and Hi-C technology. *Mol Ecol Resour.* 2020;20:1572–82. <https://doi.org/10.1111/1755-0998>.
56. Harkess A, Kochko AD, Chanderbali A, Meyers BC, Walts B, Fogliani B, Guo CC, Zheng CF, de Pamphilis CW, Job C, Sankoff D, Job D, Soltis DE, Paoli ED, Ibarra-Laclette E, Lyons E, Wafula E, Chen F, Li GL, Tang HB, Ma H, Kong HZ, Estill J, Leebens-Mack J, Burnette JM, Talag J, Palmer JD, Harholt J, Xue JY, Chen JQ, Liu J, Zhai JX, Park JS, Der JP, Acosta JJ, Liu K, Li L, Carretero-Paulet L, Rajjou L, Herrera-Estrella L, Tomsho L, Kirst M, Villegente M, Axtell M, Altman NS, Farrell NP, Soltis PS, Ralph P, Ulvskov P, Bruenn RA, Wing RA, Sederoff R, Detemann RO, Ammi'Raju' JSS, Shanid S, Kim ST, Ayyampalayam S, Arikitt S, Pissis SP, Chamala S, Wanke S, Schuster SC, Rounsley SD, Wessler SC, Lan TY, Chang TH, Yeh TF, Burtet-Sarramegna V, Poncet V, Albert VA, Chiang V, Barbazuk WB, Mei WB, Yu XX, Zhang XY, Qi XS, Sun YH, Jiao YN. The amborella genome and the evolution of flowering plants. *Science.* 2013;342:1241089. <https://doi.org/10.1126/science.1241089>.
57. Aghaei K, Komatsu S. Crop and medicinal plants proteomics in response to salt stress. *Front Plant Sci.* 2013;4:8. <https://doi.org/10.3389/fpls.2013.00008>.
58. Yin X, Li JF, Wang JQ, Tang CD, Wu MC. Enhanced thermostability of a mesophilic xylanase by N-terminal replacement designed by molecular dynamics simulation. *J Sci Food Agr.* 2013;93:3016–23. <https://doi.org/10.1002/jsfa.6134>.
59. Li XP, Fernández-Ortuño D, Grabke A, Schnabel G. Resistance to fludioxonil in *Botrytis cinerea* isolates from blackberry and strawberry. *Phytopathology.* 2014;104:724–32. <https://doi.org/10.1094/PHYTO-11-13-0308-R>.
60. Preumont A, Snoussi K, Stroobant V, Collet JF, Schaftingen EV. Molecular identification of pseudouridine-metabolizing enzymes. *J Biol Chem.* 2008;283:25238–46. <https://doi.org/10.1074/jbc.M804122200>.
61. Walley JW, Kelley DR, Nestorova G, Hirschberg DL, Dehesh K. *Arabidopsis* deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. *Plant Physiol.* 2010;152:866–75. <https://doi.org/10.1104/pp.109.149005>.
62. Sarowar S, Oh HW, Cho HS, Baek KH, Seong ES, Joung YH, Choi GJ, Lee S, Choi D. *Capsicum annuum* CCR4-associated factor CaCAF1 is necessary for plant development and defence response. *Plant J.* 2007;51:792–802. <https://doi.org/10.1111/j.1365-313X.2007.03174.x>.
63. Zhang L, Li MH, Li QQ, Chen CQ, Qu M, Li MY, Wang Y, Shen XH. The catabolite repressor/activator Cra is a bridge connecting carbon metabolism and host colonization in the plant drought resistance-promoting bacterium *Pantoea alhagi* LTYR-11Z. *Appl Environ Microb.* 2018;84:e00054-e118. <https://doi.org/10.1128/AEM.00054-18>.
64. Lin WL, Bonin M, Boden A, Wieduwild R, Murawala P, Wermke M, Andrade H, Bornhäuser M, Zhang YX. Peptidyl prolyl cis/trans isomerase activity on the cell surface correlates with extracellular matrix development. *Commun Biol.* 2019;2:58. <https://doi.org/10.1038/s42003-019-0315-8>.
65. Xie R, Shi L, Liu J, Deng T, Wang L, Liu Y, Zhao F. Genome-wide scan for runs of homozygosity identifies candidate genes in three pig breeds. *Animals.* 2019;9:518. <https://doi.org/10.3390/ani9080518>.
66. Lemes RB, Nunes K, Carnavalli JEP, Kimura L, Mingroni-Netto RC, Meyer D, Otto PA. Inbreeding estimates in human populations: applying new approaches to an admixed Brazilian isolate. *PLoS ONE.* 2018;13:e0196360. <https://doi.org/10.1371/journal.pone.0196360>.
67. Vigueira CC, Olsen KM, Caicedo AL. The red queen in the corn: agricultural weeds as models of rapid adaptive evolution. *Heredity.* 2012;110:303–11. <https://doi.org/10.1038/hdy.2012.104>.
68. Johnson RN, O'Meally D, Chen Z, Etherington GJ, Ho SYW, Nash WJ, Grueber CE, Cheng YY, Whittington CM, Dennison S, Peel E, Haerty W, O'Neill RJ, Colgan D, Russell TL, Albuquerque-Planas DE, Attenbrow V, Bragg JG, Brandies PA, Chong AYY, Deakin JE, Palma FD, Duda Z, Eldridge MDB, Ewart KM, Hogg CJ, Frankham GJ, Georges A, Gillett AK, Govendir M, Greenwood AD, Hayakawa T, Helgen KM, Hobbs M, Holleley CE, Heider TN, Jones EA, King A, Madden D, Graves JAM, Morris KM, Neaves LE, Patel HR, Polkinghorne A, Renfree MB, Robin C, Salinas R, Tsangaras K, Waters PD, Waters SA, Wright B, Wilkins MR, Timms P, Belov K. Adaptation and conservation insights from the koala genome. *Nat Genet.* 2018;50:1102–11. <https://doi.org/10.1038/s41588-018-0153-5>.
69. Chase BM, Meadows ME. Late Quaternary dynamics of southern Africa's winter rainfall zone. *Earth Sci Rev.* 2007;84:103–38. <https://doi.org/10.1016/j.earscirev.2007.06.002>.
70. Yang WH, Hammes SR. *Xenopus laevis* CYP17 regulates androgen biosynthesis independent of the cofactor cytochrome b₅. *J Biol Chem.* 2005;280:10196–201. <https://doi.org/10.1074/jbc.M411886200>.
71. Akeus P, Langenes V, von Mentzer A, Yrild U, Sjöling Å, Saksena P, Raghavan S, Järbrink M. Altered chemokine production and accumulation of regulatory T cells in intestinal adenomas of APC^{Min/+} mice. *Cancer Immunol Immun.* 2014;63:807–19.
72. Zalesny RSJ, Stange CM, Rirr BA. Survival, height growth, and phyto-extraction potential of hybrid poplar and Russian olive (*Elaeagnus angustifolia* L.) established on soils varying in salinity in North Dakota, USA. *Forests.* 2019;10:672. <https://doi.org/10.3390/f10080672>.
73. Xiang L, Gao X, Peng YH, Liang J. Coupling the occurrence of correlative plant species to predict the habitat suitability for lesser white-fronted goose (*Anser erythropus*) under climate change: a case study in the middle and lower reaches of the Yangtze River. *J Res Ecol.* 2020;11:140–9. <https://doi.org/10.5814/j.issn.1674-764x.2020.02.002>.
74. Rishworth GM, Cawthra HC, Dodd C, Perissinotto R. Peritidal stromatolites as indicators of stepping-stone freshwater resources on the Palaeo-Agulhas Plain landscape. *Quat Sci Rev.* 2020;235: 105704. <https://doi.org/10.1016/j.quascirev.2019.03.026>.
75. Kim DY, Jin JY, Alejandro S, Martinoia E, Lee Y. Overexpression of AtABCG36 improves drought and salt stress resistance in *Arabidopsis*. *Physiol Plantarum.* 2010;139:170–80. <https://doi.org/10.1111/j.1399-3054.2010.01353.x>.
76. Xu H, Song P, Gu WB, Yang ZR. Effects of heavy metals on production of thiol compounds and antioxidant enzymes in *Agaricus bisporus*. *Ecotox Environ Safe.* 2011;74:1685–92. <https://doi.org/10.1016/j.ecoenv.2011.04.010>.
77. Tang LL, Cai H, Ji W, Luo X, Wang ZY, Wu J, Wang XD, Cui L, Wang Y, Zhu YM, Bai X. Overexpression of GsZFP1 enhances salt and drought tolerance in transgenic alfalfa (*Medicago sativa* L.). *Plant Physiol Bioch.* 2013;71:22–30. <https://doi.org/10.1016/j.plaphy.2013.06.024>.
78. Wu PP, Cogill S, Qiu YJ, Li ZG, Zhou M, Hu Q, Chang ZH, Noorai RE, Xia XX, Sasaki C, Raymer P, Luo H. Comparative transcriptome profiling provides insights into plant salt tolerance in seashore paspalum (*Paspalum vaginatum*). *BMC Genom.* 2020;21:131. <https://doi.org/10.1186/s12864-020-6508-1>.

79. Shabala S, Wu H, Bose J. Salt stress sensing and early signalling events in plant roots: current knowledge and hypothesis. *Plant Sci.* 2015;241:109–19. <https://doi.org/10.1016/j.plantsci.2015.10.003>.
80. Bushman BS, Amundsen KL, Warnke SE, Robins JG, Johnson PG. Transcriptome profiling of Kentucky bluegrass (*Poa pratensis* L.) accessions in response to salt stress. *BMC Genom.* 2016;17:48. <https://doi.org/10.1186/s12864-016-2379-x>.
81. Zhang J, Jiang DC, Liu BB, Luo WC, Lu J, Ma T, Wan DS. Transcriptome dynamics of a desert poplar (*Populus pruinosa*) in response to continuous salinity stress. *Plant Cell Rep.* 2014;33:1565–79. <https://doi.org/10.1007/s00299-014-1638-z>.
82. Zhao S, Zhang Q, Liu M, Zhou H, Ma C, Wang P. Regulation of plant responses to salt stress. *Int J Mol Sci.* 2021;22:4609. <https://doi.org/10.3390/ijms22094609>.
83. Peng Z, He SP, Gong WF, Sun JL, Pan ZE, Xu FF, Lu YL, Du XM. Comprehensive analysis of differentially expressed genes and transcriptional regulation induced by salt stress in two contrasting cotton genotypes. *BMC Genom.* 2014;15:760. <https://doi.org/10.1186/1471-2164-15-760>.
84. Xiong HC, Guo HJ, Xie YD, Zhao LS, Gu JY, Zhao SR, Li JH, Liu LX. RNAseq analysis reveals pathways and candidate genes associated with salinity tolerance in a spaceflight-induced wheat mutant. *Sci Rep.* 2017;7:2731. <https://doi.org/10.1038/s41598-017-03024-0>.
85. Hsu JL, Wang LY, Wang SY, Lin CH, Ho KC, Shi FK, Chang IF. Functional phosphoproteomic profiling of phosphorylation sites in membrane fractions of salt-stressed *Arabidopsis thaliana*. *Proteome Sci.* 2009;7:42. <https://doi.org/10.1186/1477-5956-7-42>.
86. Cai GH, Wang G, Wang L, Liu Y, Pan JW, Li DQ. A maize mitogen-activated protein kinase kinase, ZmMKK1, positively regulated the salt and drought tolerance in transgenic *Arabidopsis*. *J Plant Physiol.* 2014;171:1003–16. <https://doi.org/10.1016/j.jplph.2014.02.012>.
87. Wang C, Lu WJ, He XW, Wang F, Zhou YL, Guo XL, Guo XQ. The cotton mitogen-activated protein kinase kinase 3 functions in drought tolerance by regulating stomatal responses and root growth. *Plant Cell Physiol.* 2016;57:1629–42. <https://doi.org/10.1093/pcp/pcw090>.
88. Yue X, Gao XQ, Zhang XS. Circadian rhythms synchronise intracellular calcium dynamics and ATP production for facilitating *Arabidopsis* pollen tube growth. *Plant Signal Behav.* 2015;10: e1017699. <https://doi.org/10.1080/15592324.2015.1017699>.
89. Cheng DD, Sun XB, Zhao M, Chow WS, Sun GY, Hu YB, Liu MJ, Zhang ZS. Light suppresses bacterial population through the accumulation of hydrogen peroxide in tobacco leaves infected with *Pseudomonas syringae* pv. tabaci. *Front Plant Sci.* 2016;7:521. <https://doi.org/10.3389/fpls.2016.00512>.
90. Jia HH, Hao LL, Guo XL, Liu SC, Yan Y, Guo XQ. A Raf-like MAPKKK gene, GhRaf19, negatively regulates tolerance to drought and salt and positively regulates resistance to cold stress by modulating reactive oxygen species in cotton. *Plant Sci.* 2016;252:267–81. <https://doi.org/10.1016/j.plantsci.2016.07.014>.
91. Munns R, Tester M. Mechanisms of salinity tolerance. *Annu Rev Plant Biol.* 2008;59:651–81. <https://doi.org/10.1146/annurev.arplant.59.032607.092911>.
92. Tang XL, Mu XM, Shao HB, Wang HL, Brestic M. Global plant-responding mechanisms to salt stress: physiological and molecular levels and implications in biotechnology. *Crit Rev Biotechnol.* 2014;32:425–37. <https://doi.org/10.3109/07388551.2014.889080>.
93. Ashorf M, Akran NA. Improving salinity tolerance of plants through conventional breeding and genetic engineering: an analytical comparison. *Biotechnol Adv.* 2009;27:744–52. <https://doi.org/10.1016/j.biotechadv.2009.05.026>.
94. Hase Y, Tarusawa T, Muto A, Himeno H. Impairment of ribosome maturation or function confers salt resistance on *Escherichia coli* cells. *PLoS ONE.* 2013;8: e65747. <https://doi.org/10.1371/journal.pone.0065747>.
95. Fu C, Liu XX, Yang WW, Zhao CM, Liu J. Enhanced salt tolerance in tomato plants constitutively expressing heat-shock protein in the endoplasmic reticulum. *Genet Mol Res.* 2016. <https://doi.org/10.4238/gmr.15028301>.
96. Hildebrandt TM. Synthesis versus degradation: directions of amino acid metabolism during *Arabidopsis* abiotic stress response. *Plant Mol Biol.* 2018;98:121–35. <https://doi.org/10.1007/s11103-018-0767-0>.
97. Yan HR, Jia HH, Chen XB, Hao LL, An HL, Guo XQ. The cotton WRKY transcription factor GhWRKY17 functions in drought and salt stress in transgenic *Nicotiana benthamiana* through ABA signaling and the modulation of reactive oxygen species production. *Plant Cell Physiol.* 2014;55:2060–76. <https://doi.org/10.1093/pcp/pcu133>.
98. de Bruxelles GL, Peacock WJ, Dennis ES, Dolferus R. Abscisic acid induces the alcohol dehydrogenase gene in *Arabidopsis*. *Plant Physiol.* 1996;111:381–91. <https://doi.org/10.1104/pp.111.2.381>.
99. SDAU. Hi-C of *Elaeagnus angustifolia* L. Bethesda: GenBank; 2020.
100. SDAU. Trans data of *Elaeagnus angustifolia* L. Bethesda: GenBank; 2020.
101. SDAU. Survey of *Elaeagnus angustifolia* L. Bethesda: GenBank; 2020.
102. SDAU. PB of *Elaeagnus angustifolia* L. Bethesda: GenBank; 2020.
103. SDAU. Resequencing of *Elaeagnus angustifolia* L. Bethesda: GenBank; 2020.
104. SDAU. transcriptome of *Elaeagnus angustifolia* L. Bethesda: GenBank; 2020.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

