SCIENCE CHINA
Life Sciences

A high-quality genome sequence of alkaligrass provides insights into halophyte stress tolerance

Wenting Zhang, Jie Liu, Yongxue Zhang, Jie Qiu, Ying Li, Baojiang Zheng, Fenhong Hu, Shaojun Dai and Xuehui Huang

Citation: SCIENCE CHINA Life Sciences; doi: 10.1007/s11427-020-1662-x

View online: http://engine.scichina.com/doi/10.1007/s11427-020-1662-x

Published by the Science China Press

SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

https://doi.org/10.1007/s11427-020-1662-x

A high-quality genome sequence of alkaligrass provides insights into halophyte stress tolerance

Wenting Zhang[1†], Jie Liu[1†], Yongxue Zhang[1,2], Jie Qiu[1], Ying Li[2], Baojiang Zheng[2], Fenhong Hu[1*], Shaojun Dai[1*] & Xuehui Huang[1*]

[1]*Shanghai Key Laboratory of Plant Molecular Sciences, College of Life Sciences, Shanghai Normal University, Shanghai 200234, China;*
[2]*Key Laboratory of Saline-alkali Vegetation Ecology Restoration (Northeast Forestry University), Ministry of Education, College of Life Sciences, Northeast Forestry University, Harbin 150040, China*

Received January 16, 2020; accepted March 1, 2020; published online March 12, 2020

Alkaligrass (*Puccinellia tenuiflora*) is a monocotyledonous halophytic forage grass widely distributed in Northern China. It belongs to the Gramineae family and shares a close phylogenetic relationship with the cereal crops, wheat and barley. Here, we present a high-quality chromosome-level genome sequence of alkaligrass assembled from Illumina, PacBio and 10× Genomics reads combined with genome-wide chromosome conformation capture (Hi-C) data. The ~1.50 Gb assembled alkaligrass genome encodes 38,387 protein-coding genes, and 54.9% of the assembly are transposable elements, with long terminal repeats being the most abundant. Comparative genomic analysis coupled with stress-treated transcriptome profiling uncovers a set of unique saline- and alkaline-responsive genes in alkaligrass. The high-quality genome assembly and the identified stress related genes in alkaligrass provide an important resource for evolutionary genomic studies in Gramineae and facilitate further understanding of molecular mechanisms underlying stress tolerance in monocotyledonous halophytes. The alkaligrass genome data is freely available at http://xhhuanglab.cn/data/alkaligrass.html.

alkaligrass, *Puccinellia tenuiflora*, genome assembly, saline-alkaline tolerance

Citation: Zhang, W., Liu, J., Zhang, Y., Qiu, J., Li, Y., Zheng, B., Hu, F., Dai, S., and Huang, X. (2020). A high-quality genome sequence of alkaligrass provides insights into halophyte stress tolerance. Sci China Life Sci 63, https://doi.org/10.1007/s11427-020-1662-x

INTRODUCTION

Saline-alkaline soil imposes ion toxicity, osmotic stress, and pH stress to plants, severely limiting crop distribution and agricultural productivity worldwide (Zhu, 2016; Mahajan and Tuteja, 2005). Studies of saline-alkaline responsive mechanisms in plants are necessary toward crop genetic improvements. Halophytes can survive on saline-alkaline soil due to their unique structure and metabolic pathways to cope with these stresses (Zhang et al., 2012; Tuteja et al., 2007). Alkaligrass (*Puccinellia tenuiflora*) is a monocotyledonous halophyte widely distributed in the saline-alkaline land in Northern China, and is considered as an outstanding pasture with high protein content and good palatability for livestock. Alkaligrass belongs to the Gramineae family and has a close genetic relationship with wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*) (Wang et al., 2007; Zhang et al., 2013). Through its adaptation to the local environment, alkaligrass has developed strong tolerance to various stress conditions, such as stress of up to 600 mmol L^{-1} NaCl and 150 mmol L^{-1} Na_2CO_3 (pH 11.0) (Zhang et al., 2013), as well as drought and chilling stresses (Meng et al., 2016). Alkaligrass has evolved high ability of ion homeostasis and compartmentation, such as selectivity of K^+ over Na^+ (Wang et al., 2009), restriction of unidirectional

†Contributed equally to this work
*Corresponding authors (Xuehui Huang, email: xhhuang@shnu.edu.cn; Shaojun Dai, email: daishaojun@hotmail.com; Fenhong Hu, email: fenhonghu@shnu.edu.cn)

Na^+ influx (Peng et al., 2004), secretion of Na^+ onto leaf surface (Sun et al, 2005), as well as accumulation of organic acids and inorganic anions (Guo et al., 2010). Therefore, as a cool-season perennial plant, alkaligrass is used for turf, forage and reclamation on saline-alkaline soil, and also serves as a model plant for studying saline-alkaline stress responses (Zhao et al., 2016). Investigation of the fine-tuned physiological and molecular mechanism in alkaligrass in response to saline-alkaline is critical for crop improvement (e.g., through genome editing strategy).

Diverse saline-alkali responsive strategies in alkaligrass, such as the enhanced ion transportation, osmotic homeostasis, and reactive oxygen species (ROS) scavenging, have been reported (Peng et al., 2004; Shi et al., 2002; Sun et al., 2005). Moreover, high-throughput transcriptome and proteome analyses have revealed that a number of genes and proteins in leaves, roots and callus from alkaligrass are involved in the response to saline-alkaline (e.g., NaCl, Na_2CO_3, and $NaHCO_3$) stresses (Zhang et al., 2013; Yu et al., 2011; Yu et al., 2013; Zhao et al., 2016; Yin et al., 2019; Suo et al., 2019; Zhang et al., 2019). The patterns of gene expression, protein abundance, and post-translational modification implied that the signal transduction, ion and osmotic homeostasis, ROS scavenging, transcription and protein synthesis, as well as energy and secondary metabolisms are sophisticated in certain organs under various stresses. However, to date only a few genes have been cloned and characterized due to the lack of genome information and genetic transformation system. Most of these cloned genes encode antiporters/channel proteins (e.g., PutPMP3-1/2, PutAKT1, PutHKT2;1, PtNHA1, and PutNHX), ROS scavenging enzymes (e.g., PutCAX1, PutAPX), and metal-binding protein (e.g., PutMT2), indicating that unique Na^+ compartmentalization, K^+ transport, and ROS homeostasis in alkaligrass are critical for stress tolerance (Ardie et al., 2010; Ardie et al., 2009; Guan et al., 2015; Kobayashi et al., 2012; Liu et al., 2009; Zhang et al., 2008; Zhang et al., 2014).

A high-quality genome sequence of alkaligrass is much needed to facilitate the understanding of the fine-tuned gene functions and its unique gene family feature. The fast-evolving sequencing and genome assembling technologies have allowed the generation of a large number of reference genomes in plants during the past decade (Shen et al., 2019; Ma and Cao, 2018; Zhao et al., 2018), including several from halophytes. The genome assemblies of *Thellungiella parvula* (Dassanayake et al., 2011) and *Eutrema salsugineum* (formerly *Thellungiella halophila*) (Wu et al., 2012; Yang et al., 2013), two halophytic relatives of *Arabidopsis thaliana*, have revealed a complex regulatory network of salinity tolerance, which involves gene duplication and transcriptional/post-transcriptional regulation. Comparative transcriptome analyses show that *SALT OVERLY SENSITIVE1* (*SOS1*) and ABA responsive MAPK cascades contribute to stress tolerance in these halophytes (Oh et al., 2010; Lee et al., 2013). In addition, genome assemblies of a monocot halophyte *Oryza coarctata* and a C4 halophyte *Suaeda aralocaspica* also provide insights into salt tolerance mechanisms in halophytes (Mondal et al., 2018; Wang et al., 2019).

In this study, we present a high-quality chromosome-level genome assembly of alkaligrass ($2n$=14), which contains 38,387 predicted protein-coding genes. Transcriptome profiling analysis uncovered novel signaling and metabolic pathways in alkaligrass involved in coping with saline-alkaline stress. Our results provide critical clues for molecular genetic studies of stress responses in halophytes and have potential applications in improving the crop saline-alkaline tolerance.

RESULTS

Genome sequencing and *de novo* assembly

Alkaligrass is a tufted perennial halophyte in the genus *Puccinellia*, which consists of species with diverse morphology features and genome sizes (Murray et al., 2005; Koce et al., 2008). The identity of alkaligrass was confirmed by checking its morphology and taxonomy characteristics under stereo microscope (Figure S1 in Supporting Information). The nuclear genome size of alkaligrass was estimated to be ~1.5 gigabase (Gb) using flow cytometry (Figure S2, Table S1 in Supporting Information). We sequenced the genome of alkaligrass using both Illumina HiSeq X-Ten sequencing platform and PacBio single-molecule real-time (SMRT) sequencing technology. In total, ~119 Gb PacBio reads (equivalent to ~80× genome coverage) and ~101 Gb Illumina reads were used to assemble the genome. Analysis of 17-mer sequences revealed a high level of heterozygosity of the alkaligrass genome (1.64%) (Figure S3, Table S2 in Supporting Information); therefore, we also used the genome sequence of *Aegilops tauschii* ($2n$=14; a closely related species of alkaligrass) to remove potential heterozygous contig sequences using ALLMAPS (Tang et al., 2015). Afterwards, the remaining contigs were assembled into scaffolds by ARKS (Coombe et al., 2018) and LINKS (Warren et al., 2015) using 10× Genomics data, and finally a total of 1,715 scaffolds were generated with a total size of 1.50 and an N50 size of 1.46 Mb (Table S3 in Supporting Information).

To assemble the scaffolds into pseudochromosomes, a chromosome conformation capture (Hi-C) library was constructed and sequenced, resulting in 275.2 Gb (184-fold) raw data. Using the Hi-C data, the assembled scaffolds were successfully clustered into seven groups corresponding to the seven alkaligrass chromosomes (see Figure S4 for Hi-C contact matrix and Figure S5 for the assembly workflow). The lengths of the pseudochromosomes ranged

from 163 to 258 Mb (Table S4 in Supporting Information). The assembled sizes were slightly larger than the estimated genome size of alkaligrass, probably due to the relatively high heterozygosity of the alkaligrass genome. Read coverage statistics showed that ~97.81% of Illumina short-insert reads could be aligned back to the final assembly, covering 94.86% of the assembly (Table S5 in Supporting Information). The assembled genome was further evaluated by BUSCO (Benchmarking Universal Single Copy Orthologs; Simão et al., 2015), indicating that 92.9% of the core conserved plant genes were completely captured in the assembly (Table S6 in Supporting Information). These metrics indicated the high quality of the alkaligrass genome assembly.

Genome annotation

De novo screening of repetitive sequences revealed an overall repeat content of ~65.37% in the alkaligrass genome (Table S7 in Supporting Information), of which ~84% are transposable elements (TEs). The retrotransposons (class I elements) constituted 45.34% of the genome, with 42.76% being long terminal repeats (LTRs; 30.79% gypsy-type and 11.97% copia-type) and 2.39% non-long terminal repeats (non-LTRs; 2.33% LINE and 0.06% SINE). Hence, LTR-retrotransposons were the most abundant elements in the alkaligrass genome, the same as those in most plant genomes. The ratio of gypsy-like LTRs to copia-like LTRs was 2.57, similar to that found in other species in Pooideae, a subfamily of the BOP clade in Gramineae (Table S8 in Supporting Information). To assess the evolutionary role of TEs, it is important to estimate when the TEs were integrated into the genome. We estimated that the majority of intact LTRs were amplified 0 to 0.5 million years ago (Mya) in the alkaligrass genome (Figure S6A in Supporting Information). The majority of LTRs in alkaligrass appeared to insert into the genome more recently when compared to other *Triticeae* species.

Furthermore, protein-coding genes were predicted in the alkaligrass genome assembly using a combination of *ab initio* prediction, homology search, and transcript mapping. The hybrid gene prediction approach generated 38,387 gene models in the alkaligrass genome, with an average gene length of 4,065 bp, an average coding sequence (CDS) length of 1,142 bp, and an average of 5 exons per gene. Among these predicted genes, 30,781 (93.1%) were homologs of *Arabidopsis thaliana* genes (Table S9 in Supporting Information), and 34,119 (92.5%) had functional annotation information (Table S10 in Supporting Information). Across the alkaligrass genome the gene density ranged from 0 to 57 genes per million bases (Mb) (Figure 1), and as expected the density of protein-coding genes generally showed an opposite pattern to that of the repetitive elements.

Evolutionary history of alkaligrass

To infer the phylogeny for alkaligrass, protein sequences of 10 species from the PACMAD clade (*Cenchrus americanus*, *Setaria italic*, *Sorghum bicolor* and *Zea mays*) and the BOP clade (*Oryza sativa*, *Brachypodium distachyon*, *Puccinellia tenuiflora*, *Aegilops tauschii*, *Triticum urartu* and *Hordeum vulgare*) in monocots, as well as two eudicots (*Arabidopsis thaliana* and *Glycine max*), were compared. Based on the 102 single-copy orthologous genes identified, we constructed a phylogenetic tree of these 12 plant species (Figure 2A). The divergence between the PACMAD clade and the BOP clade was estimated to occur ~47 Mya, in congruence with the previous report (International Brachypodium Initiative, 2010).

We further investigated whether any whole genome duplication (WGD) events have occurred during alkaligrass evolution. We identified 19,511 paralogous gene pairs that covered 50.8% of the alkaligrass genome assembly, including 11,225 inter-chromosomal paralogous gene pairs. We used synonymous substitution rate (K_s) values of the inter-chromosomal paralogous gene pairs to calculate the age distribution of the duplication events and identified two peaks at K_s of ~0.06 and ~0.80. We then performed the same analysis using the genome data of other five BOP species to determine whether these WGD events in alkaligrass were species-specific or shared with others. The result showed that K_s of ~0.80 peak was shared in all Poaceae species (Figure 2B), suggesting that alkaligrass underwent the same WGD in the ancestor of Poaceae species, about 65.5 Mya. Meanwhile, the K_s of ~0.06 weak peak observed in alkaligrass and *H. vulgare*, obscured in *A. tauschii* and *O. sativa*, dissappered in *T. urartu* and *B. distachyon* was not considered as alkaligrass-specific WGD, possibly due to imperfections in gene annotation or retro-transposon genes in these genomes, which needs to be investigated in details in future study.

Based on the collinear blocks between the genomes of alkaligrass and three *Triticeae* species, as well as *O. sativa* and *C. americanus*, we calculated the density distribution of the K_s values for the paired genes within each syntenic genomic block (Figure 2C; Table S11 in Supporting Information). The peak of K_s was ~0.5 for orthologous gene pairs between alkaligrass and rice and ~0.6 between alkaligrass and *C. americanus*, indicating that the more ancient WGD event occurred before the divergence of the PACMAD clade and the BOP clade.

Comparative genomics analysis

To investigate gene family expansion in alkaligrass, we compared proteomes of alkaligrass and 12 other plants, and identified a total of 21,803 orthologous gene families that

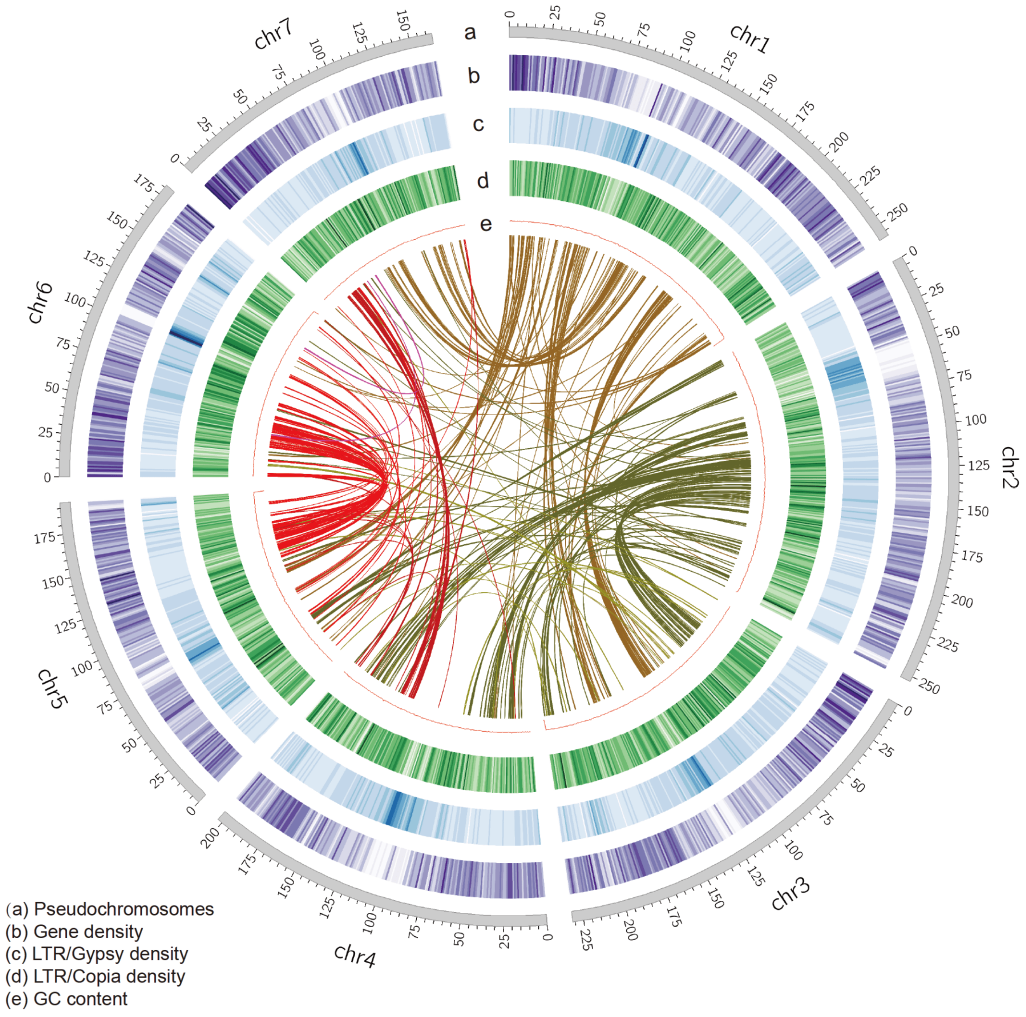

Figure 1 Landscape of the alkaligrass genome. Circos plot of the alkaligrass genome assembly. Circles from the outside inwards: (a) pseudochromosomes, (b) gene density, (c) LTR/Gypsy density, (d) LTR/Copia density, and (e) GC content. These density metrics were calculated with 1 Mb sliding windows. Genome syntenic blocks are illustrated with colored lines.

consist of 340,014 genes, with 5,947 orthogroups containing proteins from all species. All Gramineae species included in the analysis shared 719 single-copy and 1,075 multiple-copy orthogroups (Figure 3A). In total, 6,691 gene families were expanded in alkaligrass, whereas 5,059 were contracted (Figure 3A). Compared to other BOP clade species, much more gene families were expanded in alkaligrass. The expanded gene families in alkaligrass were mainly enriched in functional categories such as binding (GO: 0005488) and protein binding (GO: 0005515), suggesting possible roles in relation to transcription or translation (Figure 3B). In addition, the most expanded family was the F-box/FBD/LRR-repeat protein which implies improved resistance to stress in alkaligrass (Figure S7 in Supporting Information).

By comparing alkaligrass with five other BOP species including *B. distachyon*, *A. tauschii*, rice, *T. urartu* and barley, we found that 58.3% (8,006/13,726) of the gene families in alkaligrass were shared by all these six species, while only 1.2% (159) were specific to alkaligrass covering 257 genes (Figure 3C). Among the genes in alkaligrass specific families, 146 were annotated to have *Arabidopsis* homologs, and most proteins encoded by those genes were located in plasma membrane and endomembrane system, implying that saline-alkali sensing, signal transduction, ion transport, vesicular trafficking and cell structure dynamics would be crucial for alkaligrass salinity tolerance.

A number of synteny blocks were detected within the genome of the alkaligrass and between genomes of alkaligrass and other species (Table S11 in Supporting Information; Figure 1). In total, 832 and 822 synteny blocks with an average size of 1.77 and 1.35 Mb were identified in alkaligrass when compared with rice and *T. urartu* genomes, respectively (Table S11 in Supporting Information). Synteny blocks between alkaligrass and *T. urartu* genomes showed that approximately one third of the genes (12,201) in alkaligrass had collinear relationships with *T. urartu* genes

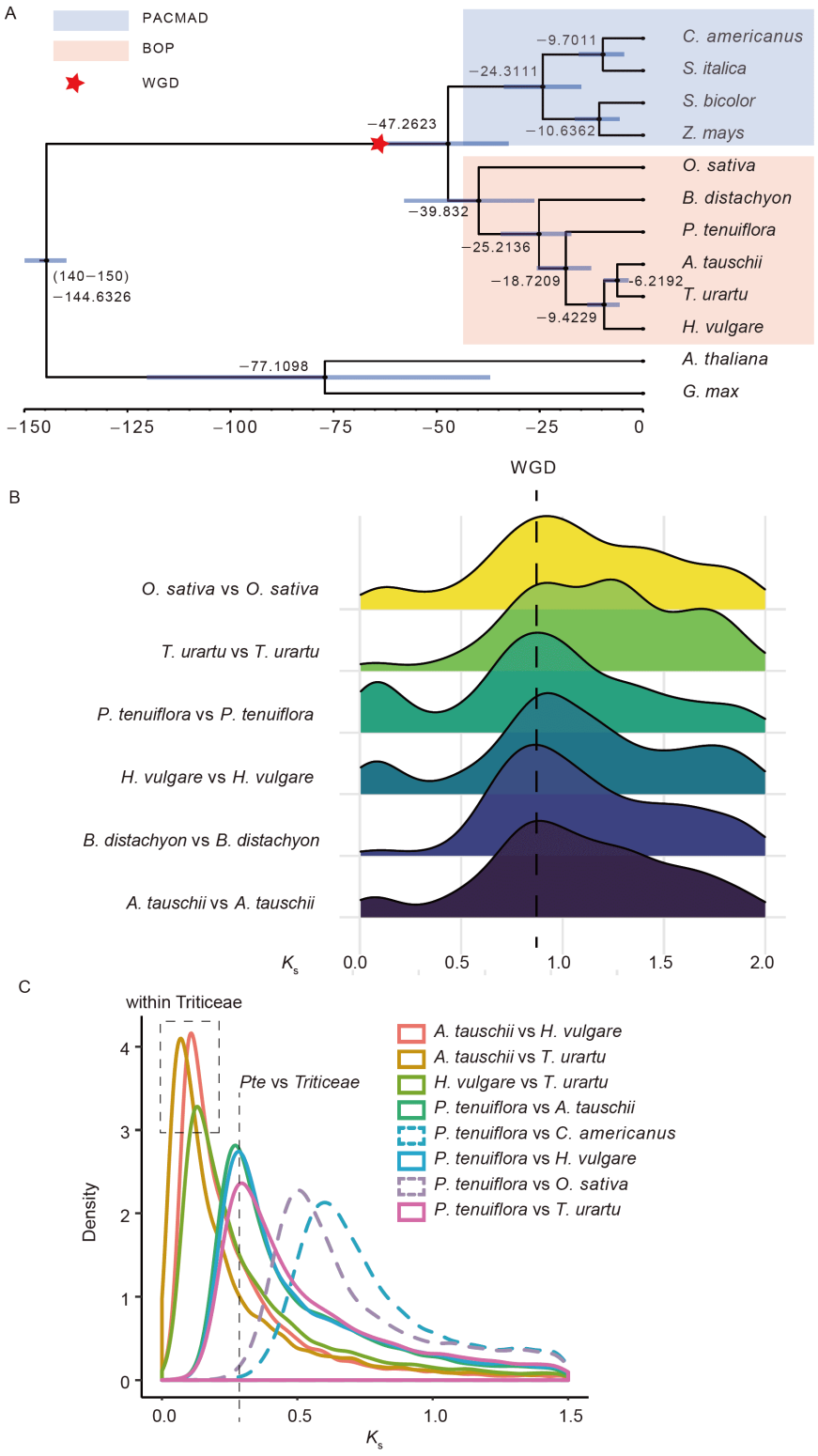

Figure 2 Evolution of the alkaligrass genome. A, Phylogenetic tree of 12 species constructed based on 102 single-copy genes, with *A. thaliana* as the outgroup. Divergence times were estimated using the divergence time of monocot-dicot (140–150 Mya) as the calibration point. Blue bars on the nodes are the estimated range of divergence times (Mya). B, Distribution of synonymous substitution rates (K_s) among collinear paralogs in six BOP plants. C, Frequency distributions of K_s of collinear orthologs among four BOP plants (*A. tauschii*, *T. urartu*, *H. vulgare*, and *O. sativa*), one PACMAD plant (*C. americanus*) and alkaligrass.

(Figure S8 in Supporting Information). Despite highly collinear relationship, several inversion and transposition events were observed between the genomes of alkaligrass and *T. urartu* (e.g., those for Pte_chr4-Tur_chr7, Pte_chr2-

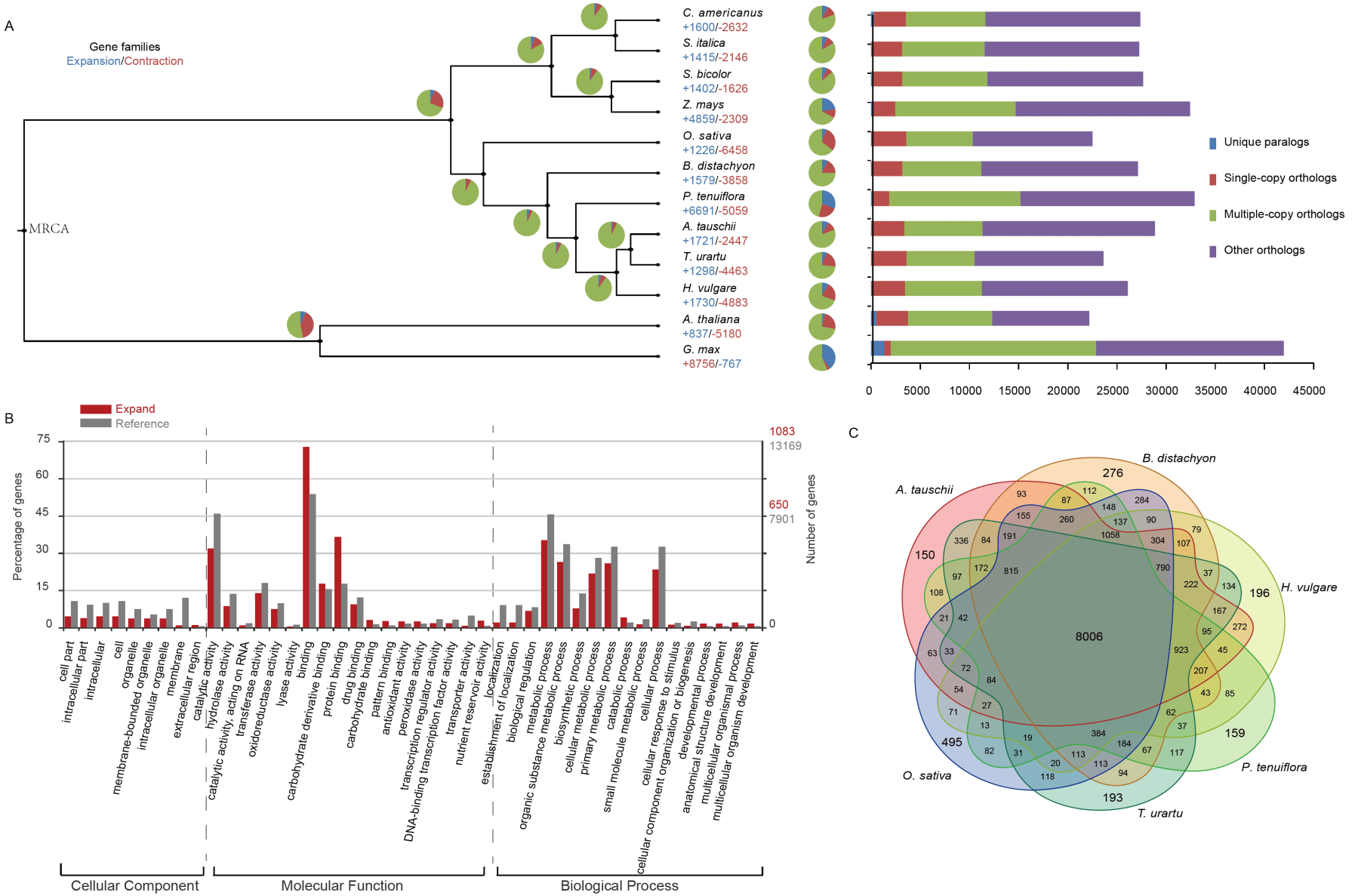

Figure 3 Comparative analysis of gene families in alkaligrass. A, Gene family expansions and contractions (left panel) and proportion of expanded/contracted/remained gene families in each plant species (pie charts). MRCA, most recent common ancestor. The right panel displays the distribution of single-copy, multiple-copy, unique and other orthologs. B, Gene ontology (GO) functional classification. "Expand" indicates GO term annotations of expanded genes. "Reference" indicates GO term annotations of all alkaligrass genes. C, Venn diagram showing the number of gene families shared among six Poaceae species.

Tur_chr6, Pte_chr5-Tur_chr1, and Pte_chr6-Tur_chr3) (Figure S9 in Supporting Information). Chromosome-scale rearrangements were also observed between alkaligrass and rice. For example, alkaligrass chromosome 3 seems to correspond to a fusion of rice chromosomes 4 and 7 (Figure S9A in Supporting Information).

Genome-wide saline-alkali response revealed from transcriptome analysis

To examine the genome-wide response to saline-alkali stress in alkaligrass, we performed a series of deep transcriptome sequencing of 50-day seedlings under NaCl, Na_2CO_3, and $NaHCO_3$ treatments (Table S12 in Supporting Information), using the Illumina sequencing platform. We analyzed the expression profiles in roots and leaves in response to various levels of salinity and alkali stresses (Figure 4A). Hierarchical clustering of all samples using Pearson's correlation based on FPKM values showed a high correlation within root or leaf samples, but poor correlation between root and leaf samples. The expression levels of two representative NaCl-responsive genes (*PutANN1* and *PutBIK1*) were evaluated by the RT-qPCR analysis (Figure 4B). *PutANN1* and *PutBIK1* encoded a calcium-dependent annexin and a plasma membrane-localized Ser/Thr protein kinase, respectively. Both genes were significantly induced by NaCl in roots when compared with those in leaves of alkaligrass, which implied their important roles in roots for saline-alkaline sensing and signaling (Wang et al., 2015; Zhang et al., 2016).

The transcriptomic data revealed that genes encoding receptor-like kinase FER and FLS2, as well as their partners, were induced by saline/alkali in the roots of alkaligrass (Figure 4C). Genes that encode FER and its relative HERK1, as well as their interacting proteins (i.e., LLG1, RALF1, LRX2, AGB1, and ROPGEF1) were up-regulated in roots for sensing and triggering certain saline-/alkali-responses. The downstream members (i.e., GTPROP5, ABI2, and SnRK2.2) in the ABA signaling were also induced in roots under salinity and alkaline stresses. Moreover, genes encoding FLS2 and its interacting proteins (i.e., EFR, BAK1, BIK1, SCD1, BSK1, and GRP7) were all regulated by saline/alkali in roots. Additionally, genes that encode ROS-generated protein Rboh A/B/E/J, diverse ROS-mediated proteins in kinase signaling pathways (i.e., MEKK1, MKK1/2/4/5, and MPK3/4), proteins in Ca^{2+} signaling pathway (ANN1 and NORTIA), as well as transcription factors WRKY22/25/29, were detected to be up-regulated in response to saline-alkali stress (Figure 4C). Among them, two representative NaCl-induced genes, *PutMEKK* and *PutFLS2*, were evaluated and confirmed by the RT-qPCR analysis (Figure 4D).

In addition, four FER orthologs were identified in alkaligrass, while only one existed in the *Arabidopsis* genome. RNA-Seq data showed that the four FER homologs exhibited consistent expression pattern in roots in response to salinity (Figure 4E). The diverse expression levels of *FER*s were probably due to the sequence divergence in their promoters, with identity of sequences 1 kb upstream of transcriptional start site of these four *FER* copies ranging from 41% to 47%. As difference in *cis*-elements is an important factor to influence expression levels, the *cis*-elements in the promoter sequences of the four *FER*s identified with FIMO (Grant et al., 2011) may be potential targets for genome editing to fine-tune the expression levels and improve salinity stress resistance (Supplemental file 2 in Supporting Information).

DISCUSSION

High-quality genome is necessary for molecular genetic studies in alkaligrass

The saline-alkali soil mainly contains diverse salinity and alkali contents such as NaCl, Na_2CO_3, $NaHCO_3$, Na_2SO_4 and NaOH, which inhibit plant development and growth (Wang et al., 2007). Alkaligrass (*P. tenuiflora*) has strong tolerance to saline, alkali, drought, and chilling stresses (Zhang et al., 2013), and can grow normally under 600 mmol L^{-1} NaCl and 150 mmol L^{-1} Na_2CO_3 (pH 11.0) for 6 days (Zhang et al., 2013), as well as under 800 mmol L^{-1} $NaHCO_3$ (pH 9.0) for 7 days (Yin et al., 2019), showing greater tolerance than rice (maximum of 40 mmol L^{-1} Na_2CO_3 or 100 mmol L^{-1} NaCl for 10 days) and barley (maximum of 250 mmol L^{-1} NaCl) (Lv et al., 2013; Colmer et al., 2005). Previous studies have focused on salinity-responsive mechanisms in alkaligrass owing to its extreme salinity tolerance and high value for cereal genetic improvement. However, due to the lack of genome information, transcriptome and proteome studies are mainly dependent on the genome databases of other related plants, and the identification of interesting trait-associated genes and proteins could be very imprecise (Yu et al., 2011; Yu et al., 2013; Zhang et al., 2013; Zhao et al., 2016; Meng et al., 2016; Yin et al., 2019; Suo et al., 2019). In this study, we generated high-quality genome and transcriptome data for alkaligrass, which provide a fundamental resource for alkaligrass functional genomic studies, and for facilitating genetic mapping of various traits in alkaligrass. The detailed information of genome structure, gene annotation and gene expression provide valuable candidate genes in the QTL regions, which will contribute to genome editing and genetic complementation experiments, and other molecular genetic analyses. This would be helpful in molecular breeding for improving crops with higher yield and stress tolerance.

Receptor kinase FER and FLS2 are involved in salinity response in alkaligrass

It has been well addressed that plasma membrane-localized

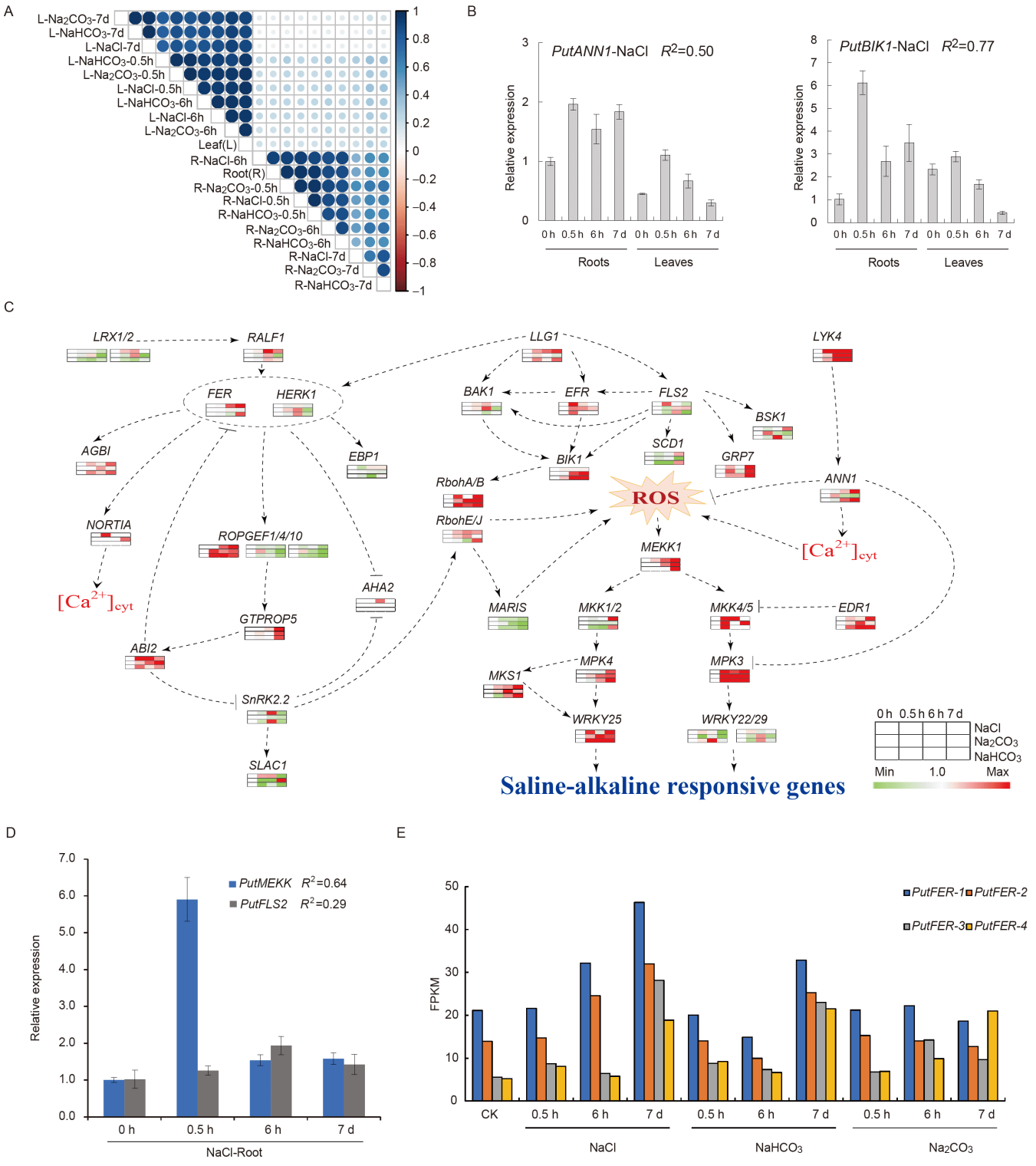

Figure 4 Salt-responsive genes and pathways in alkaligrass. A, Visualization of correlation matrix of root (R) and leaf (L) samples under the treatments of Na_2CO_3, $NaHCO_3$, and NaCl for 0.5 h, 6 h and 7 d, based on gene expression profiles. B, Relative expression levels of *PutANN1* and *PutBIK* in root and leaf under NaCl treatment evaluated by RT-qPCR analysis. C, Receptor-like kinase FER and FLS2 related pathways in roots were involved in salinity and alkali tolerance. Transcriptome profiling revealed that *FER* and *HERK1*, as well as their interacting protein encoding genes (*LLG1*, *RALF1*, *LRX2*, *AGB1*, and *ROPGEF1*) were induced by saline-alkaline. The downstream members (i.e., GTPROP5, ABI2, and SnRK2.2) in ABA signaling were also up-regulated under certain stress conditions. *FLS2* and its related genes (*EFR*, *BAK1*, *BIK1*, *SCD1*, *BSK1*, and *GRP7*) were all up-regulated. In addition, genes encoding ROS-generated protein Rboh A/B/E/J, as well as the members of kinase signaling (MEKK1, MKK1/2/4/5, and MPK3/4), Ca^{2+} signaling (ANN1 and NORTIA), and transcription factors WRKY22/25/29 were conditionally up-regulated in response to saline-alkaline stress. D, Relative expression levels of *PutMEKK* and *PutFLS2* in root under NaCl treatment. E, Expression profiles of four alkaligrass *FER* homologs in root under different treatments. R^2 in B and D represents the correlation values between the real-time PCR and RNA-Seq data. See Supplemental file 1 in Supporting Information for gene abbreviations.

FER receptor kinase and its downstream partners control plant reproduction, growth, and development, as well as participate in defense-related responses (Li et al., 2015). However, limited evidence is available to support that FERs are involved in salinity response. *Arabidopsis* FER interacts with RALF22/23 and leucine-rich repeat extensins (LRX), functioning as a module to transduce cell-wall signals, and to elicit cell-specific calcium transients for maintaining cell wall integrity in response to salt stress (Zhao et al., 2018; Feng et al., 2018). *Arabidopsis* FER also interacts with G protein beta subunit (AGB1) and RALF1 to regulate stomatal movement and salinity response (Yu and Assmann, 2018). Furthermore, FLS2 interacts with other downstream proteins to modulate immunity response (Shen et al., 2017). Importantly, *Arabidopsis* LLG1 works as a co-receptor of FER and FLS2, participating in the regulation of development and innate immunity in plants (Shen et al., 2017). However, whether these well-known development and immunity-related genes are involved in salinity response in halophytes remains unclear. Our results revealed that four *PutFERs* and a large number of genes in FER and FLS signaling pathways were induced by salinity/alkali in roots of alkaligrass (Figure 4D). This implies that *FER* and its related genes (i.e., *HERK1*, *NOTIA*, and *MARIS*), as well as *FLS2* and its interacted partners (i.e., *BAK1*, *EFR*, *BIK1*, *SCD1*, *BSK1*, and *GRP7*) synergistically modulate G protein-mediated Ca^{2+} signaling, ABA signaling, ROS homeostasis and MAPK cascade signaling to regulate salt-responsive gene expression. Functions of these candidate genes in salt responses need to be further validated and investigated by molecular genetics and biochemical analyses.

CONCLUSION

The halophytic alkaligrass (*P. tenuiflora*) is an outstanding pasture for soil improvement. Its genome assembly allows us to draw the complete catalog of genes for underlying saline-alkaline tolerant mechanisms in monocots, and also provides a crucial basis for understanding salinity adaptation of pasture and facilitating the molecular genetic improvement of crops.

MATERIALS AND METHODS

Plant materials and growth conditions

Alkaligrass seedlings were cultured in a modified Hoagland solution in a chamber under fluorescent light (300 μmol m^{-2} s^{-1}, 13 h light/11 h darkness) at 25°C and 75% humidity for 50 days (Yu et al., 2011). For transcriptome analysis, the plants were treated with 50 mmol L^{-1} Na_2CO_3, 100 mmol L^{-1} NaCl, and 100 mmol L^{-1} $NaHCO_3$ (pH 9) for 0.5 h, 6 h, and 7 d, respectively. After treatment, leaves and roots were harvested, frozen immediately in liquid nitrogen and stored at –80°C.

Genome size estimation

The genome size of alkaligrass was estimated using flow cytometry as described in Dolezel et al. (2007). Samples were prepared by homogenizing young leaves of *H. vulgare* cv. Morex, which has a genome size of 5.1 Gb (Mayer et al., 2012) and was used as an internal standard, and four independent individuals of alkaligrass on ice in Galbraith's buffer (5 mmol L^{-1} sodium metabi-sulfite and 5 μL β-mercaptoethanol complemented) with 50 μg mL^{-1} propidium iodide. After filtration with 40 μm nylon cell strainer (Falcon, BD Biosciences, San Jose, CA, USA), samples were analyzed on a MoFlo XDP Cell Sorter (excitation 488 nm, emission 620 nm; Beckman Coulter, Hialeah, FL, USA). The flow cytometry data were analyzed with the FlowJo software (version 10.6.1; FlowJo LLC, Ashland, OR, USA).

Genome assembly

Genomic DNA was extracted from leaf tissues of alkaligrass (basic plant characteristics shown in Figure S1 in Supporting Information) using the cetyltrimethylammonium bromide (CTAB) method. Paired-end (PE) library with insert size around 400 bp was constructed according to the manufacturer's instructions (Illumina, San Diego, CA, USA) and sequenced on an Illumina HiSeq X-ten system. PacBio SMRT libraries were constructed following the standard SMRT bell construction protocol and sequenced on an RS II platform (PacBio, Menlo Park, CA, USA).

PacBio long reads were error corrected using Falcon (length_cutoff=15,000, length_cutoff_pr=16,000, pa_HPCdaligner_option='-v -B100 -t12 -w8 -M24 -e.70 -k18 -h280 -l2800 -s1000', ovlp_HPCdaligner_option='-v -B120 -k20 -h480 -e.96 -l2800 -s1000 -T8'). *De novo* assembly of alkaligrass genome was performed with the error-corrected PacBio reads using wtdbg (wtdbg-1.2.8 -k 0 -p 17 -S 2; wtdbg-cns -j 500 -c 0 -k 7) (v1.8.7; Chin et al., 2016). Arrow (smrtlink5.1.0; Chin et al., 2013) and Pilon (v1.22; Walker et al., 2014) were employed to polish the assembly using PacBio and Illumina reads. The redundancy in the polished contig sequences were removed based on gene synteny compared with the genome of *Aegilops tauschii* (a closely related species of alkaligrass) using ALLMAPS (v0.8.12; Tang et al., 2015). 10× Genomics data was used to scaffold the assembled contigs using the ARKS-LINKS pipeline with the parameters of ARKS (Coombe et al., 2018) set to '-c 5 -j 0.55 -m 50-10000 -k 30 -r 0.05 -e 3000 -z 500 -d 0' and the parameters of LINKS (Warren et al., 2015) set to '-l 5 –a 0.9

-z 500'. Based on high-throughput chromosome conformation capture (Hi-C) data, scaffolds were anchored to seven pseudo-molecules using LACHESIS (Burton et al., 2013) with parameters 'CLUSTER MIN RE SITES=100; CLUSTER MAX LINK DENSITY=2.5; CLUSTER NON-INFORMATIVE RATIO=1.4; ORDER MIN N RES IN TRUNK=60; ORDER MIN N RES IN SHREDS=60'.

Genome annotation

A custom repeat library was built by screening the alkaligrass genome assembly using RepeatModeler (version open-1.0.11) (Tarailo-Graovac and Chen, 2009) (http://repeatmasker.org/). RepeatMasker (V4.0.7) was used to mask and classify repeat sequences in the genome. LTR-RTs were first annotated using LTR harvest (genome tools V1.5.11; Ellinghaus et al., 2008), LTR_FINDER (V1.07; Xu and Wang, 2007), and LTR_FINDER_parallel (Ou and Jiang, 2019). The intact LTR-RTs and the insertion time were then refined and estimated by LTR retriever (V2.1; Ou and Jiang, 2018) and plotted using ggplot in Rstudio1.1.383.

A hybrid strategy combining *ab initio* predictions, homologous gene evidence, and transcript support (RNA-Seq) was applied in gene predictions. Four *ab initio* gene finders, GeneMark-ET (V4.46) (Lomsadze et al., 2014), Augustus (V3.3.2) (Stanke et al., 2006), Fgenesh (version) (Salamov and Solovyev, 2000), and SNAP (version 2006-07-28) (Korf, 2004) were used. RNA-Seq reads generated from five tissues (spike, leaf, flower, stem, and root), as well as salt-treated samples, were *de novo* assembled using Trinity (Grabherr et al., 2011). Protein sequences of *A. tauschii*, *T. urartu*, and *H. vulgare* were used to obtain homology evidence supporting the predicted gene structures. All gene structures predicted by the above procedures were integrated into consensus gene models using MAKER (V3.1.2) (Holt and Yandell, 2011).

Protein domains were identified by comparing protein sequences of alkaligrass predicted genes against various domain databases using InterProScan v5 (Jones et al., 2014). Gene ontology (GO) terms were assigned for each gene based on its corresponding InterPro entries. Functional annotation of protein-coding genes was achieved using BLASTp (E-value<1×10^{-5}) against the *Arabidopsis thaliana* protein sequence database.

Phylogenetic tree construction

The protein sequences from *A. thaliana*, rice, *S. bicolor*, *Z. mays*, *S. italica*, *B. distachyon*, *H. vulgare*, *T. urartu*, *A. tauschii*, *C. americanus*, and *G. max* were downloaded from Ensembl (http://ensemblgenomes.org/info/genomes) and GigaDB (http://gigadb.org/). For genes with multiple transcripts, only the longest transcripts in the coding region were included in the analysis. OrthoFinder (V2.2.7) (Emms and Kelly, 2015) was used to identify paralogs and orthologs with an all-vs-all BLASTp search performed by Diamond (V0.9.23) (Buchfink et al., 2015). Single-copy orthologs were used to construct the phylogenetic tree of all tested species with MAFFT (V7.419) (Katoh et al., 2002) and RAXML (V8.2.12; parameters '-m PROTGAMMAAUTO -nb 100') (Stamatakis, 2014) using the EasySpeciesTree python script (https://github.com/Davey1220/EasySpeciesTree). Finally, the phylogeny tree was visualized using Figtree (V1.4.4) (Rambaut, 2009).

Species divergence time estimation

The MCMCTree program (http://abacus.gene.ucl.ac.uk/software/paml.html) implemented in Phylogenetic Analysis with Maximum Likelihood (PAML; V4.9i) was used to infer the divergence time of the nodes on the phylogenetic tree. Parameters used for MCMCTree were as follows: a burn-in of 2,000 steps, sample number of 20,000 and sample frequency of 10. Normal prior was used for monocot-dicot divergence time (140–150 Mya). The divergence time was also calculated using the formula $T = K_s/2R$, where K_s refers to the synonymous substitutions per site, and R is the rate of divergence of nuclear genes in plants, which was set to 6.1×10^{-9} according to Lynch and Conery (2000).

Expansion and contraction of gene families

The expansion and contraction of gene families were determined by comparing the cluster size differences between the ancestor and each species using the CAFE (V4.2.1) program (Han et al., 2013). A random birth-and-death model was used to evaluate changes in gene families along each lineage of the phylogenetic tree. Using conditional likelihood as the test statistics, we calculated the corresponding P-values of each lineage, and a P-value $\leq$0.01 was considered significant.

Comparative genomic analysis

The alkaligrass genome was compared to genomes of itself and other species using the MCScan toolkit (Tang et al., 2008). To call synteny blocks, we performed all-against-all LAST (Kiełbasa et al., 2011) and chained the LAST hits with a distance cutoff of 20 genes, requiring at least five gene pairs per synteny block. Chromosome-scale synteny block plots were constructed using the python version of MCScan. For each gene pair in a syntenic block, K_s values were calculated, and values of all gene pairs were plotted to identify putative whole-genome duplication events in alkaligrass and other BOP species. The syntenic blocks between alkaligrass chromosomes were visualized using Circos (V0.69-6; Krzywinski et al., 2009).

RNA-Seq library construction and data analysis

Paired-end RNA-Seq libraries were constructed using the NEB Next® Ultra RNA Library Prep Kit (Illumina). cDNA fragments of ~300 bp in size were excised, followed by enrichments using PCR amplification for ~10 cycles. The resulting paired-end RNA-Seq libraries were sequenced on the Illumina NovaSeq 6000 system to generate 150 bp paired-end reads. RNA-Seq reads were aligned against the alkaligrass genome using HISAT2 (V2.1.0; Kim et al., 2015). The expression level (FPKM) for each protein-coding gene was calculated using Stringtie (V1.3.5; Pertea et al., 2015) with default parameters. All primers used for qRT-PCR are listed in Supplemental file 3 in Supporting Information.

Data availability

The assembled alkaligrass genome has been deposited in BIGD under Bioproject number PRJCA002121. The genome assembly and gene annotations of alkaligrass can also be accessed from http://www.xhhuanglab.cn/data/alkaligrass.html.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *The authors thank Dr. Zhangjun Fei from Cornell University for critical reading and editing of the manuscript. This work was supported by grants from the National Key Research and Development Program of China (2018YFA090060), the Foundation of Shanghai Science and Technology Committee (17391900600), the Fund of Shanghai Engineering Research Center of Plant Germplasm Resources (17DZ2252700), and the Natural Science Foundation of Heilongjiang Province (ZD2019C003).*

References

Ardie, S.W., Xie, L., Takahashi, R., Liu, S., and Takano, T. (2009). Cloning of a high-affinity K^+ transporter gene *PutHKT2;1* from *Puccinellia tenuiflora* and its functional comparison with *OsHKT2;1* from rice in yeast and *Arabidopsis*. J Exp Bot 60, 3491–3502.

Ardie, S.W., Liu, S., and Takano, T. (2010). Expression of the AKT1-type K^+ channel gene from *Puccinellia tenuiflora*, *PutAKT1*, enhances salt tolerance in *Arabidopsis*. Plant Cell Rep 29, 865–874.

Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using diamond. Nat Methods 12, 59–60.

Burton, J.N., Adey, A., Patwardhan, R.P., Qiu, R., Kitzman, J.O., and Shendure, J. (2013). Chromosome-scale scaffolding of *de novo* genome assemblies based on chromatin interactions. Nat Biotechnol 31, 1119–1125.

Chin, C.S., Peluso, P., Sedlazeck, F.J., Nattestad, M., Concepcion, G.T., Clum, A., Dunn, C., O'Malley, R., Figueroa-Balderas, R., Morales-Cruz, A., et al. (2016). Phased diploid genome assembly with single-molecule real-time sequencing. Nat Methods 13, 1050–1054.

Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E.E., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10, 563–569.

Colmer, T.D., Munns, R., and Flowers, T.J. (2005). Improving salt tolerance of wheat and barley: future prospects. Aust J Exp Agric 45, 1425–1443.

Coombe, L., Zhang, J., Vandervalk, B.P., Chu, J., Jackman, S.D., Birol, I., and Warren, R.L. (2018). ARKS: chromosome-scale scaffolding of human genome drafts with linked read kmers. BMC Bioinformatics 19, 234.

Dassanayake, M., Oh, D.H., Haas, J.S., Hernandez, A., Hong, H., Ali, S., Yun, D.J., Bressan, R.A., Zhu, J.K., Bohnert, H.J., et al. (2011). The genome of the extremophile crucifer *Thellungiella parvula*. Nat Genet 43, 913–918.

Dolezel, J., Greilhuber, J., and Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc 2, 2233–2244.

Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). Ltrharvest, an efficient and flexible software for *de novo* detection of ltr retrotransposons. BMC Bioinformatics 9, 18.

Emms, D.M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol 16, 157.

Feng, W., Kita, D., Peaucelle, A., Cartwright, H.N., Doan, V., Duan, Q., Liu, M.C., Maman, J., Steinhorst, L., Schmitz-Thom, I., et al. (2018). The FERONIA receptor kinase maintains cell-wall integrity during salt stress through Ca^{2+} signaling. Curr Biol 28, 666–675.e5.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29, 644–652.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: Scanning for occurrences of a given motif. Bioinformatics 27, 1017–1018.

Guan, Q., Wang, Z., Wang, X., Takano, T., and Liu, S. (2015). A peroxisomal APX from *Puccinellia tenuiflora* improves the abiotic stress tolerance of transgenic *Arabidopsis thaliana* through decreasing of H_2O_2 accumulation. J Plant Physiol 175, 183–191.

Guo, L.Q., Shi, D.C., and Wang, D.L. (2010). The key physiological response to alkali stress by the alkali-resistant halophyte *Puccinellia tenuiflora* is the accumulation of large quantities of organic acids and into the rhyzosphere. J Agron Crop Sci 196, 123–135.

Han, M.V., Thomas, G.W.C., Lugo-Martinez, J., and Hahn, M.W. (2013). Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. Mol Biol Evol 30, 1987–1997.

Holt, C., and Yandell, M. (2011). MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 12, 491.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30, 3059–3066.

Kiełbasa, S.M., Wan, R., Sato, K., Horton, P., and Frith, M.C. (2011). Adaptive seeds tame genomic sequence comparison. Genome Res 21, 487–493.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12, 357–360.

Kobayashi, S., Abe, N., Yoshida, K.T., Liu, S., and Takano, T. (2012). Molecular cloning and characterization of plasma membrane- and vacuolar-type Na^+/H^+ antiporters of an alkaline-salt-tolerant monocot, *Puccinellia tenuiflora*. J Plant Res 125, 587–594.

Koce, J.D., Škondrić, S., Bačič, T., and Dermastia, M. (2008). Amounts of nuclear DNA in marine halophytes. Aquat Bot 89, 385–389.

Konopka-Postupolska, D., Clark, G., Goch, G., Debski, J., Floras, K., Cantero, A., Fijolek, B., Roux, S., and Hennig, J. (2009). The role of annexin 1 in drought stress in *Arabidopsis*. Plant Physiol 150, 1394–1410.

Korf, I. (2004). Gene finding in novel genomes. BMC Bioinformatics 5, 59.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res 19, 1639–1645.

Lee, Y.P., Giorgi, F.M., Lohse, M., Kvederaviciute, K., Klages, S., Usadel,

B., Meskiene, I., Reinhardt, R., and Hincha, D.K. (2013). Transcriptome sequencing and microarray design for functional genomics in the extremophile *Arabidopsis* relative *Thellungiella salsuginea* (*Eutrema salsugineum*). BMC Genomics 14, 793.

Li, C., Yeh, F.L., Cheung, A.Y., Duan, Q., Kita, D., Liu, M.C., Maman, J., Luu, E.J., Wu, B.W., Gates, L., et al. (2015). Glycosylphosphatidylinositol-anchored proteins as chaperones and co-receptors for FERONIA receptor kinase signaling in *Arabidopsis*. eLife 4, e06587.

Liu, H., Zhang, X., Takano, T., and Liu, S. (2009). Characterization of a *PutCAX1* gene from *Puccinellia tenuiflora* that confers Ca^{2+} and Ba^{2+} tolerance in yeast. Biochem Biophys Res Commun 383, 392–396.

Lomsadze, A., Burns, P.D., and Borodovsky, M. (2014). Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. Nucleic Acids Res 42, e119.

Lv, B.S.., Li, X.W., Ma, H.Y., Sun, Y., Wei, L.X., Jiang, C.J., and Liang, Z. W. (2013). Differences in growth and physiology of rice in response to different saline-alkaline stress factors. Agron J 105, 1119–1128.

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

Ma, X., and Cao, X. (2018). Cotton variant genomes—a breakthrough in population genetics analysis. Sci China Life Sci 61, 869–870.

Mahajan, S., and Tuteja, N. (2005). Cold, salinity and drought stresses: an overview. Arch Biochem Biophys 444, 139–158.

Mayer, K.F.X., Waugh, R., Brown, J.W.S., Schulman, A., Langridge, P., Platzer, M., Fincher, G.B., Muehlbauer, G.J., Sato, K., Close, T.J., et al. (2012). A physical, genetic and functional sequence assembly of the barley genome. Nature 491, 711–716.

Meng, X., Zhao, Q., Jin, Y., Yu, J., Yin, Z., Chen, S., and Dai, S. (2016). Chilling-responsive mechanisms in halophyte *Puccinellia tenuiflora* seedlings revealed from proteomics analysis. J Proteomics 143, 365–381.

Mondal, T.K., Rawal, H.C., Chowrasia, S., Varshney, D., Panda, A.K., Mazumdar, A., Kaur, H., Gaikwad, K., Sharma, T.R., and Singh, N.K. (2018). Draft genome sequence of first monocot-halophytic species *Oryza coarctata* reveals stress-specific genes. Sci Rep 8, 13698.

Murray, B.G., De Lange, P.J., and Ferguson, A.R. (2005). Nuclear DNA variation, chromosome numbers and polyploidy in the endemic and indigenous grass flora of New Zealand. Ann Bot 96, 1293–1305.

Oh, D.H., Dassanayake, M., Haas, J.S., Kropornika, A., Wright, C., d'Urzo, M.P., Hong, H., Ali, S., Hernandez, A., Lambert, G.M., et al. (2010). Genome structures and halophyte-specific gene expression of the extremophile *Thellungiella parvula* in comparison with *Thellungiella salsuginea* (*Thellungiella halophila*) and *Arabidopsis*. Plant Physiol 154, 1040–1052.

Ou, S., and Jiang, N. (2019). Ltr_finder_parallel: Parallelization of ltr_finder enabling rapid identification of long terminal repeat retrotransposons. Mobile DNA 10, 48.

Ou, S., and Jiang, N. (2018). Ltr_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. Plant Physiol 176, 1410–1422.

Peng, Y.H., Zhu, Y.F., Mao, Y.Q., Wang, S.M., Su, W.A., and Tang, Z.C. (2004). Alkali grass resists salt stress through high [K^+] and an endodermis barrier to Na^+. J Exp Bot 55, 939–949.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33, 290–295.

Rambaut, A. (2009). FigTree 1.3. http://tree.bio.ed.ac.uk/software/figtree/.

Salamov, A.A., and Solovyev, V.V. (2000). *Ab initio* gene finding in *Drosophila* genomic DNA. Genome Res 10, 516–522.

Shen, Q., Bourdais, G., Pan, H., Robatzek, S., and Tang, D. (2017). *Arabidopsis* glycosylphosphatidylinositol-anchored protein LLG1 associates with and modulates FLS2 to regulate innate immunity. Proc Natl Acad Sci USA 114, 5749–5754.

Shen, Y., Du, H., Liu, Y., Ni, L., Wang, Z., Liang, C., and Tian, Z. (2019). Update soybean Zhonghuang 13 genome to a golden reference. Sci China Life Sci 62, 1257–1260.

Shi, D., Yin, S., Yang, G., and Zhao, K. (2002). Citric acid accumulation in an alkali-tolerant plant *Puccinellia tenuiflora* under alkaline stress. Acta Bot Sin 44, 537–540.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212.

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

Stanke, M., Schöffmann, O., Morgenstern, B., and Waack, S. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinformatics 7, 62.

Sun, G., Peng, Y., Shao, H., Chu, L., Zhao, X., Min, H., Cao, W., and Wei, C. (2005). Does *Puccinelia tenuiflora* have the ability of salt exudation? Colloids Surf B Biointerfaces 46, 197–203.

Suo, J., Zhang, H., Zhao, Q., Zhang, N., Zhang, Y., Li, Y., Song, B., Yu, J., Cao, J., Wang, T., et al. (2019). Na_2CO_3-responsive photosynthetic and ROS scavenging mechanisms in chloroplasts of alkaligrass revealed by phosphoproteomics. bioRxiv 871046.

Tang, H., Zhang, X., Miao, C., Zhang, J., Ming, R., Schnable, J.C., Schnable, P.S., Lyons, E., and Lu, J. (2015). ALLMAPS: robust scaffold ordering based on multiple maps. Genome Biol 16, 3.

Tang, H., Wang, X., Bowers, J.E., Ming, R., Alam, M., and Paterson, A.H. (2008). Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. Genome Res 18, 1944–1954.

Tarailo-Graovac, M., and Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protocols BioInf 25, 4–10.

Tuteja, N. (2007). Mechanisms of high salinity tolerance in plants. Methods Enzymol 428, 419–438.

International Brachypodium Initiative. (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463, 763–768.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE 9, e112963.

Wang, C.M., Zhang, J.L., Liu, X.S., Li, Z., Wu, G.Q., Cai, J.Y., Flowers, T. J., and Wang, S.M. (2009). *Puccinellia tenuiflora* maintains a low Na^+ level under salinity by limiting unidirectional Na^+ influx resulting in a high selectivity for K^+ over Na^+. Plant Cell Environ 32, 486–496.

Wang, L., Ma, G., Wang, H., Cheng, C., Mu, S., Quan, W., Jiang, L., Zhao, Z., Zhang, Y., Zhang, K., et al. (2019). A draft genome assembly of halophyte *Suaeda aralocaspica*, a plant that performs C4 photosynthesis within individual cells. GigaScience 8, giz116.

Wang, X., Ma, X., Wang, H., Li, B., Clark, G., Guo, Y., Roux, S., Sun, D., and Tang, W. (2015). Proteomic study of microsomal proteins reveals a key role for *Arabidopsis* Annexin 1 in mediating heat stress-induced increase in intracellular calcium levels. Mol Cell Proteomics 14, 686–694.

Wang, Y., Chu, Y., Liu, G., Wang, M.H., Jiang, J., Hou, Y., Qu, G., and Yang, C. (2007). Identification of expressed sequence tags in an alkali grass (*Puccinellia tenuiflora*) cDNA library. J Plant Physiol 164, 78–89.

Wang, Y., Yang, C., Liu, G., and Jiang, J. (2007). Development of a cDNA microarray to identify gene expression of *Puccinellia tenuiflora* under saline–alkali stress. Plant Physiol Biochem 45, 567–576.

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

Warren, R.L., Yang, C., Vandervalk, B.P., Behsaz, B., Lagman, A., Jones, S.J.M., and Birol, I. (2015). LINKS: Scalable, alignment-free scaffolding of draft genomes with long reads. GigaScience 4, 35.

Wu, H.J., Zhang, Z., Wang, J.Y., Oh, D.H., Dassanayake, M., Liu, B., Huang, Q., Sun, H.X., Xia, R., Wu, Y., et al. (2012). Insights into salt tolerance from the genome of *Thellungiella salsuginea*. Proc Natl Acad Sci USA 109, 12219–12224.

Xu, Z., and Wang, H. (2007). Ltr_finder: An efficient tool for the prediction

of full-length ltr retrotransposons. Nucleic Acids Res 35, W265–W268.
Yang, R., Jarvis, D.E., Chen, H., Beilstein, M.A., Grimwood, J., Jenkins, J., Shu, S.Q., Prochnik, S., Xin, M., Ma, C., et al. (2013). The reference genome of the halophytic plant *Eutrema salsugineum*. Front Plant Sci 4, 46.
Yin, Z., Zhang, H., Zhao, Q., Yoo, M.J., Zhu, N., Yu, J., Yu, J., Guo, S., Miao, Y., Chen, S., et al. (2019). Physiological and comparative proteomic analyses of saline-alkali $NaHCO_3$-responses in leaves of halophyte *Puccinellia tenuiflora*. Plant Soil 437, 137–158.
Yu, J., Chen, S., Wang, T., Sun, G., and Dai, S. (2013). Comparative proteomic analysis of *Puccinellia tenuiflora* leaves under Na_2CO_3 stress. Int J Mol Sci 14, 1740–1762.
Yu, J., Chen, S., Zhao, Q., Wang, T., Yang, C., Diaz, C., Sun, G., and Dai, S. (2011). Physiological and proteomic analysis of salinity tolerance in *Puccinellia tenuiflora*. J Proteome Res 10, 3852–3870.
Yu, Y., and Assmann, S.M. (2018). Inter-relationships between the heterotrimeric Gβ subunit AGB1, the receptor-like kinase FERONIA, and RALF1 in salinity response. Plant Cell Environ 41, 2475–2489.
Zhang, C.Q., Nishiuchi, S., Liu, S., and Takano, T. (2008). Characterization of two plasma membrane protein 3 genes (*PutPMP3*) from the alkali grass, *Puccinellia tenuiflora*, and functional comparison of the rice homologues, *OsLti6a/b* from rice. BMB Rep 41, 448–454.
Zhang, H., Han, B., Wang, T., Chen, S., Li, H., Zhang, Y., and Dai, S. (2012). Mechanisms of plant salt response: insights from proteomics. J Proteome Res 11, 49–67.
Zhang, H., Huang, L., Hong, Y., and Song, F. (2016). BOTRYTIS-INDUCED KINASE1, a plasma membrane-localized receptor-like protein kinase, is a negative regulator of phosphate homeostasis in *Arabidopsis thaliana*. BMC Plant Biol 16, 152.
Zhang, M., Takano, T., Liu, S., and Zhang, X. (2014). Abiotic stress response in yeast and metal-binding ability of a type 2 metallothionein-like protein (PutMT2) from *Puccinellia tenuiflora*. Mol Biol Rep 41, 5839–5849.
Zhang, X., Wei, L., Wang, Z., and Wang, T. (2013). Physiological and molecular features of *Puccinellia tenuiflora* tolerating salt and alkaline-salt stress. J Integrat Plant Biol 55, 262–276.
Zhang, Y., Zhang, Y., Yu, J., Zhang, H., Wang, L., Wang, S., Guo, S., Miao, Y., Chen, S., Li, Y., et al. (2019). NaCl-responsive ROS scavenging and energy supply in alkaligrass callus revealed from proteomic analysis. BMC Genomics 20, 990.
Zhao, C., Zayed, O., Yu, Z., Jiang, W., Zhu, P., Hsu, C.C., Zhang, L., Tao, W.A., Lozano-Durán, R., and Zhu, J.K. (2018). Leucine-rich repeat extensin proteins regulate plant salt tolerance in *Arabidopsis*. Proc Natl Acad Sci USA 115, 13123–13128.
Zhao, Q., Feng, Q., Lu, H., Li, Y., Wang, A., Tian, Q., Zhan, Q., Lu, Y., Zhang, L., Huang, T., et al. (2018). Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. Nat Genet 50, 278–284.
Zhao, Q., Suo, J., Chen, S., Jin, Y., Ma, X., Yin, Z., Zhang, Y., Wang, T., Luo, J., Jin, W., et al. (2016). Na_2CO_3-responsive mechanisms in halophyte *Puccinellia tenuiflora* roots revealed by physiological and proteomic analyses. Sci Rep 6, 32717.
Zhu, J.K. (2016). Abiotic stress signaling and responses in plants. Cell 167, 313–324.

SUPPORTING INFORMATION

Figure S1 Morphology of alkaligrass.

Figure S2 Genome size estimation in alkaligrass by flow cytometry.

Figure S3 Genome size estimation by *K*-mer analysis.

Figure S4 Hi-C map of the alkaligrass genome showing genome-wide chromatin interactions.

Figure S5 Integrated work-flow for the assembly of the alkaligrass genome.

Figure S6 Insertion time of (A) intact LTRs, (B) Copia/LTRs and (C) Gypsy/LTRs in seven BOP species.

Figure S7 Phylogenetic tree of the most expanded family (F-box/FBD/LRR-repeat protein) in alkaligrass.

Figure S8 Comparative genomic analysis.

Figure S9 Inter-genomic comparison of (A) alkaligrass vs. *Triticum urartu* (B) alkaligrass vs. *Oryza sativa*.

Table S1 Estimation of nuclear DNA amount in *P. tenuiflora* by flow cytometry

Table S2 Genome survey of alkaligrass

Table S3 Summary of the final genome assemblies of alkaligrass

Table S4 Length and gene number statistics for each chromosome

Table S5 Coverage statistics of alkaligrass genome

Table S6 Assessment of completeness of gene annotation using BUSCO

Table S7 Summary of repetitive sequence in the assembled alkaligrass genome

Table S8 LTR subclass ratio in whole genome sequences

Table S9 Assembly summary of alkaligrass genome

Table S10 Statistics of gene function annotation alkaligrass genome

Table S11 Statistics of synteny blocks within alkaligrass and other species

Table S12 Data statistics of samples used for RNA-seq

Supplemental file 1 Abbreviations of the protein names in the pathways in Figure 4C.

Supplemental file 2 FIMO analysis result of four FER promoter sequence.

Supplemental file 3 Primer sequences of the alkaligrass genes for qRT-PCR.

The supporting information is available online at http://life.scichina.com and https://link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.