



Genomic Variation Landscape of the Model Salt Cress *Eutrema salsugineum*

Xiaojuan Wang^{1*†}, Hua Rao^{2,3†}, Jianxiang Ma², Xiaodan Chen¹, Guanglin Li² and Guifang Zhao^{1*}

¹ Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an, China, ² College of Life Sciences, Shaanxi Normal University, Xi'an, China, ³ Special Economic Zone for Science and Technology Synergy, China State-Level Xixian New Area, Xi'an, China

Eutrema salsugineum has long been used as the model for examining salt and other abiotic stress in plants. In addition to the forward genetics approaches widely used in the lab, natural variations undoubtedly will provide a rich genetic resource for studying molecular mechanisms underlying the stress tolerance and local adaptation of this species. We used 90 resequencing whole genomes of natural populations of this species across its Asian and North American distributions to detect the selection signals for genes involved in salt and other stresses at the species-range level and local distribution. We detected selection signals for genes involved in salt and other abiotic tolerance at the species level. In addition, several cold-induced and defense genes showed selection signals due to local adaptation in North America-NE Russia or northern China, respectively. These variations and findings provide valuable resources for further deciphering genetic mechanisms underlying the stress tolerance and local adaptations of this model species.

OPEN ACCESS

Edited by:

Tingshuang Yi,
Kunming Institute of Botany, Chinese
Academy of Sciences, China

Reviewed by:

Tao Ma,
Sichuan University, China
Yongshuai Sun,
Xishuangbanna Tropical Botanical
Garden (CAS), China

*Correspondence:

Xiaojuan Wang
xiaojwang@nwu.edu.cn
Guifang Zhao
gfzhao@nwu.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Systematics and Evolution,
a section of the journal
Frontiers in Plant Science

Received: 25 April 2021

Accepted: 22 July 2021

Published: 17 August 2021

Citation:

Wang X, Rao H, Ma J, Chen X, Li G
and Zhao G (2021) Genomic Variation
Landscape of the Model Salt Cress
Eutrema salsugineum.
Front. Plant Sci. 12:700161.
doi: 10.3389/fpls.2021.700161

Keywords: *Eutrema salsugineum*, population genomics, abiotic stress, selection, local adaptation

INTRODUCTION

The development of salt-tolerant crops is becoming an urgent matter due to the increased presence of salinized soils around the world. Information on the genetic basis of salt-tolerance can be obtained by studying natural extremophiles (Amtmann et al., 2005). The salt cress, *Eutrema* (= *Thellungiella*) *salsugineum* (Brassicaceae), is a halophyte with a high tolerance to salt, cold, drought, and oxidative stress; furthermore, it is closely related to the model plant *Arabidopsis thaliana* (Bressan et al., 2001; Inan et al., 2004; Gong et al., 2005; Griffith et al., 2007; Lamdan et al., 2012; Lee et al., 2012; Zhang et al., 2013; Yu and Li, 2014). Like *A. thaliana*, salt cress is an excellent experimental system with a short life cycle, self-pollination, a small genome size and easy transformation by the floral-dip method (Yu and Li, 2014). Consequently, *E. salsugineum* has been widely used as a model to study genetic mechanisms of salt and other abiotic stress tolerance in plants (Gong et al., 2005; Griffith et al., 2007; Lamdan et al., 2012). For example, the functions of numerous ion antiporters and transporters involving salt tolerance differ greatly between *E. salsugineum* and *A. thaliana* (Zhu, 2003; Kant et al., 2006; Kumari et al., 2015). In addition, the two species differ markedly in antioxidant capacity, photosynthetic pathway and accumulation of conjugated polyamines in response to salt and other stresses (Stepien and Johnson, 2009; Pang et al., 2010). However, the part played by natural selection in causing these differences has yet to be investigated.

E. salsugineum is widely distributed on saline soils from central Asia to northern China and North America. Long-distance migration and formation of a geographically disjunct distribution might have occurred very recently (Wang et al., 2015, 2018) and, in turn, promoted local adaptation as observed, for instance, in *Arabidopsis thaliana* (Fournier-Level et al., 2011). In support, two commonly used ecotypes of *E. salsugineum* collected respectively from Shandong in northern China and Yukon in Canada, North America, show contrasted expressions of abiotic stress-related genes (Wong et al., 2005), which may be important in local adaptation selected by their different habitats. Annual average spring and winter temperatures are distinctly lower in Yukon (-19°C) than in northern China (4°C) (Wong et al., 2005; Griffith et al., 2007) and therefore might select for low temperature tolerance. Preliminary transcriptome analyses revealed over 39,000 SNPs differences between the salt cress populations from these two regions (Champigny et al., 2013). However, to correctly interpret these data it is necessary to characterize the evolutionary relationships between populations since many expression differences could simply reflect neutral divergence among populations (Kryvokhyzha et al., 2016).

In this study, we present analyses of the whole genomes of salt cress individuals from North America, China and Russia (Altai and Yakutsk) in order to examine genome-scale nucleotide variation across the range of the species and local adaptation. We used the obtained all high-quality genome-wide SNPs to identify genes with selection signals at the species level and local adaptation. This was facilitated by the availability of a recently developed reference genome (Wu et al., 2012; Yang et al., 2013). Re-sequencing genomes of different populations has proved highly effective for uncovering genomic signatures of selection and inferring demographic histories in model and non-model animal and plant species in model and non-model animal and plant species (Andolfatto, 2005; Olson et al., 2010; Branca et al., 2011; Huang et al., 2012; Evans et al., 2014; Li et al., 2014; Lamichhaney et al., 2015; Qiu et al., 2015; Ru et al., 2016; Sun et al., 2020; Zhao et al., 2021). Here, we first identified genes with selection signals across the species range before analyzing signals of local adaptation especially in northern China and North America-NE Russia.

MATERIALS AND METHODS

Sample Selection and Resequencing Data Collection

Genomic nucleotide data from 90 individuals of 21 populations across the species range from central Asia to North America were obtained from the previous study (Wang et al., 2018) (Supplementary Tables 1, 2). The reference genome of salt cress, *Eutrema salsugineum*, was reported (Yang et al., 2013).

Genome Mapping and SNP Calling

Clean reads from each sample above were aligned to the *Eutrema salsugineum* nuclear genome sequence v1.0 (Salt cress) (Yang et al., 2013). Genome mapping was conducted using BWA software with “mem” option and default parameters (Li and

Durbin, 2009). The Picard package (<http://picard.sourceforge.net/>) was subsequently used to check for PCR duplicates. The Genome Analysis Toolkit (GATK) (McKenna et al., 2010) was used to perform local realignment of reads to enhance the alignments in the vicinity of putative indels.

After genome mapping, The SNP calling was done for all individuals using SAMtools v1.1 (mpileup and BCFtools). Only paired aligned reads were used for SNP calling. The genotype likelihoods for each individual per site were calculated, and allele frequencies were estimated. The “mpileup” command was used to identify SNPs with the parameters “-q 30 -C 50 -S -D -Q 30 -m 2 -F 0.002 -guf.” Low-coverage depth SNPs (summing all samples) were then filtered with the vcfutils.pl in BCFtools v1.1 (Li et al., 2009) with parameters “-d 135 -D 1800” and high-quality SNPs (RMS of mapping quality ≥ 10 , the distance of adjacent SNPs in the vicinity of indel polymorphisms ≥ 5 bp, Hardy-Weinberg equilibrium (HWE) $P < 5e-3$, SNP quality ≥ 30 , $3.0 \leq$ quality by depth (each individual) ≤ 30 , SNPs with observed heterozygosity ($H_o < 0.6$) were further filtered by Perl scripts. We used the obtained about 1.76 million high-quality SNPs for subsequent analysis.

Population Phylogenetic Analyses

The software RAXML was then used to construct phylogenetic trees with the GTR-G model and 1,000 rapid bootstrapping replicates based on the Maximum Likelihood (ML) method (Stamatakis, 2015). The final Maximum Likelihood trees were viewed using FigTree (v1.4.0) (<http://tree.bio.ed.ac.uk/software/figtree/>).

Screening for Selective Sweeps Across the Species Range

To identify genomic regions that might have been subject to selection during stress tolerance, we applied genetic diversity tests to the entire data from all 90 individuals. Nucleotide diversity was calculated using the standard estimate of the scaled mutation rate: the average pairwise nucleotide diversity θ_{π} (Tajima, 1989). Tajima's D was also calculated by dividing the difference between the average pairwise nucleotide diversity and the proportion of segregating sites by the square root of its SE (Tajima, 1989). We scanned the genome for regions with the highest differences using a window size of 20 kb and a step size of 10 kb. Windows that shared the lowest 5% of θ_{π} and lowest 5% Tajima's D estimates of the entire data were identified as putatively selected regions in the salt cress. Genes located in these regions were considered putatively selected genes. The above analyses were conducted using VcfTools (Danecek et al., 2011) or PERL scripts.

Screening for Selective Sweeps for Local Adaptation

To identify genomic regions that might have been associated with local adaptation, we also applied genetic diversity tests to subpopulation data: Y1 (from northern China) and Y4 (from North America-Russia). The fixation index (F_{ST}) (Weir and Cockerham, 1984), and nucleotide diversity (θ_{π} log-ratio Y1/Y4 and Y4/Y1) were also chosen as indicators of population differentiation. We scanned the genome for regions with the

highest differences using a window size of 20 kb and a step size of 10 kb. Windows that shared the highest 5% of F_{ST} and highest 5% log-ratio estimates were recognized as positively selected regions in a given population. Genes located in these regions were considered putatively selected genes in local area.

Gene Annotation Analysis

We annotated functional categories of *E. salsugineum* genes based on the corresponding *Arabidopsis thaliana* orthologs. Functional enrichment analysis of Gene Ontology (GO) was performed using the KOBAS 2.0 web server (Wu et al., 2006). The chi-squared test was used to calculate the statistical significance of enrichment and only terms with a p -value <0.05 were considered significant.

RESULTS

Whole Genome Resequencing

We obtained re-sequenced data of the genomes of 90 *E. salsugineum* individuals spanning their worldwide geographic distributions from Wang et al. (2018). All of these data were mapped to an available reference genome (Yang et al., 2013). All high-quality single-nucleotide polymorphisms (SNPs) were further used for subsequent analysis.

To identify genomic regions that might have been subject to selection, we first combined the all populations into a single gene pool. Out of 24,039 windows of 20 kb in length sliding in 10 kb steps across the salt cress genome, 23,366 windows contain >10 SNPs and cover 97.9% of the genome (Figure 1; Supplementary Figure 1). These 23,366 windows were used to detect signatures of selective sweeps at the species level. To identify regions with selective sweep signals, we used an empirical procedure (Branca et al., 2011) and selected windows with both significantly low diversity (θ_{π}) (5% right tail, where θ_{π} is 0.00053) and an excess of low-frequency variants (low D_T) (5% right tail, where D_T is -1.0197) when compared to the empirical distribution of these two statistics. This led to the identification of a total of 4.76 Mb genomic regions (1.97% of the genome, containing 262 genes) with strong selective sweep signals (Figure 1). Among these putatively selected regions, 227 genes were annotated and classified according to terms developed by the Gene Ontology Consortium (Berardini et al., 2004) and found in the TAIR database (Yon et al., 2003). Among them, 48% of genes were assigned to broad or unspecified GO categories: biological processes unknown, other cellular, other metabolic and other biological processes. We successfully classified the remaining genes to roles in transcription and signal transduction, protein, DNA or RNA metabolism, transport, development, response to abiotic or biotic stimulus and response to stress (Figure 2; Table 1; Supplementary Table 3). Of the categories with defined functions, the larger groups of biological processes were involved in response to abiotic or biotic stresses (10.5%) including stress (5.6%) and abiotic or biotic stimulus (4.9%), protein metabolism (8.5%), and transcription (8.5%) (Figure 2; Supplementary Table 3). Cell compartments clustered by GO terms showed that most of the predicted gene products were localized in the nucleus and endoplasmic reticulum (ER)

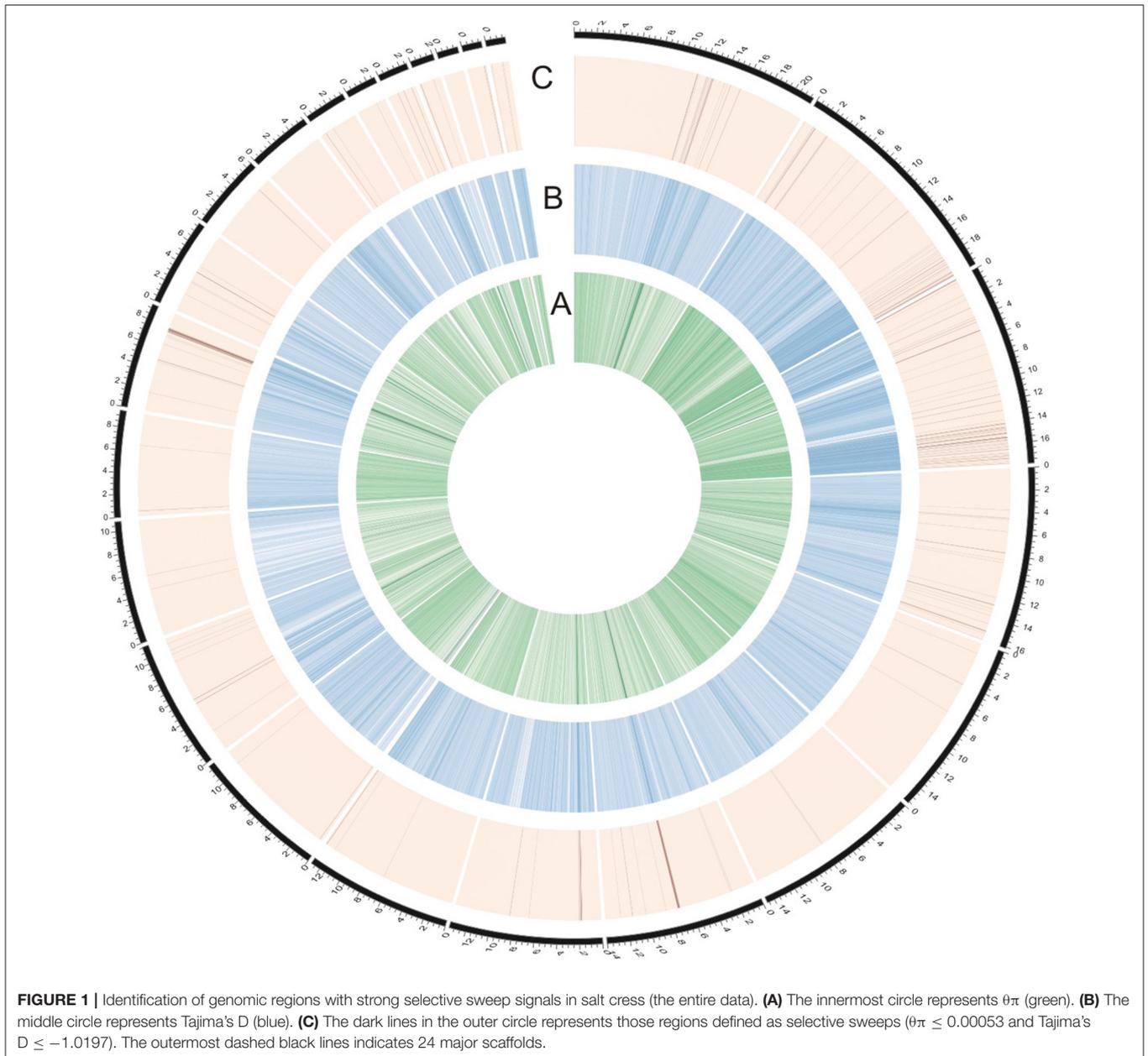
(Figure 2). We identified 16 genes that were common to abiotic and biotic stresses with selection signals related to ion homeostasis, osmotic adjustment and growth regulation (Table 1). For example, seven of these (*GLP9*, *NDPK1*, *ACO2*, *PTR3*, *PYK10*, *NUDT7*, and *MDAR2*) are involved in response to osmotic adjustment and salt stress while some others (including *ERF13*, *WRKY38*, *DAR4*, and several disease resistance protein) are closely related to cell death and defense in plant growth regulations (Table 1).

Phylogenetic Relationship

We conducted phylogenetic analyses of all sampled individuals based on the nuclear genome SNPs. Consistent with our earlier research (Wang et al., 2018), all samples of salt cress were clustered into four distinct lineages, Y1 to Y4 (Supplementary Figure 2; Supplementary Tables 1, 2). Y1 comprised all individuals from northern China (Y1), Y2 contained sampled individuals from western China (Xinjiang) while Y3 contained those from Altai. Y4 comprised all individuals from NE Russia and Canada.

Local Selection in North America-Russia and Northern China

To detect accurately the genomic footprints left by local selections between two main salt cress ecotypes defined by phylogenetic analyses (Wang et al., 2018), we applied genetic diversity tests to data from the two groups of populations, Y1 (northern China) and Y4 (North America-NE Russia). Putatively selected genes (PSGs) were identified by screening selected windows simultaneously with significantly high \log_2 [Y1 ratio (θ_{π} , $Y4/\theta_{\pi}$, Y1) and Y4 ratio (θ_{π} , $Y1/\theta_{\pi}$, Y4)] (5% right tail, where \log_2 (Y1 ratio) is 1.84 and \log_2 (Y4 ratio) is 2.4043) and significantly high F_{ST} values (5% right tail, F_{ST} threshold: 0.9098) between them (Li et al., 2014; Qiu et al., 2015; Ma et al., 2018), which also exhibited significant differences (p -value $<10^{-16}$, Mann-Whitney U -test) (Figure 3A). We identified 23 regions under selection for the population Y1 (northern China) with a total size of around 0.43 Mb (0.17% of the genome) and 17 regions for the population Y4 (North America-Russia) with a total size of around 0.33 Mb (0.13% of the genome). A total of 63 genes including 60 protein-coding genes and three hypothetical ones were annotated in the regions under selection of Y1 from northern China, while 42 protein-coding genes were annotated for those of Y4 from North America-Russia (Figure 3A; Supplementary Tables 4, 5). Functions of the genes from northern China are involved in DNA repair, plastid fission, meiosis, regulation of defense response to fungus, and regulation of sulfur metabolic process (Supplementary Table 4), while gene ontology (GO) enrichment analysis revealed that genes from North America-Russia are functionally related to response to various carbohydrates, transport, peptide biosynthetic process, and long-chain fatty acid biosynthesis process (Supplementary Table 5). Example of genes with strong selection sweep signals in northern China and North America-Russia are shown in Figure 3B.



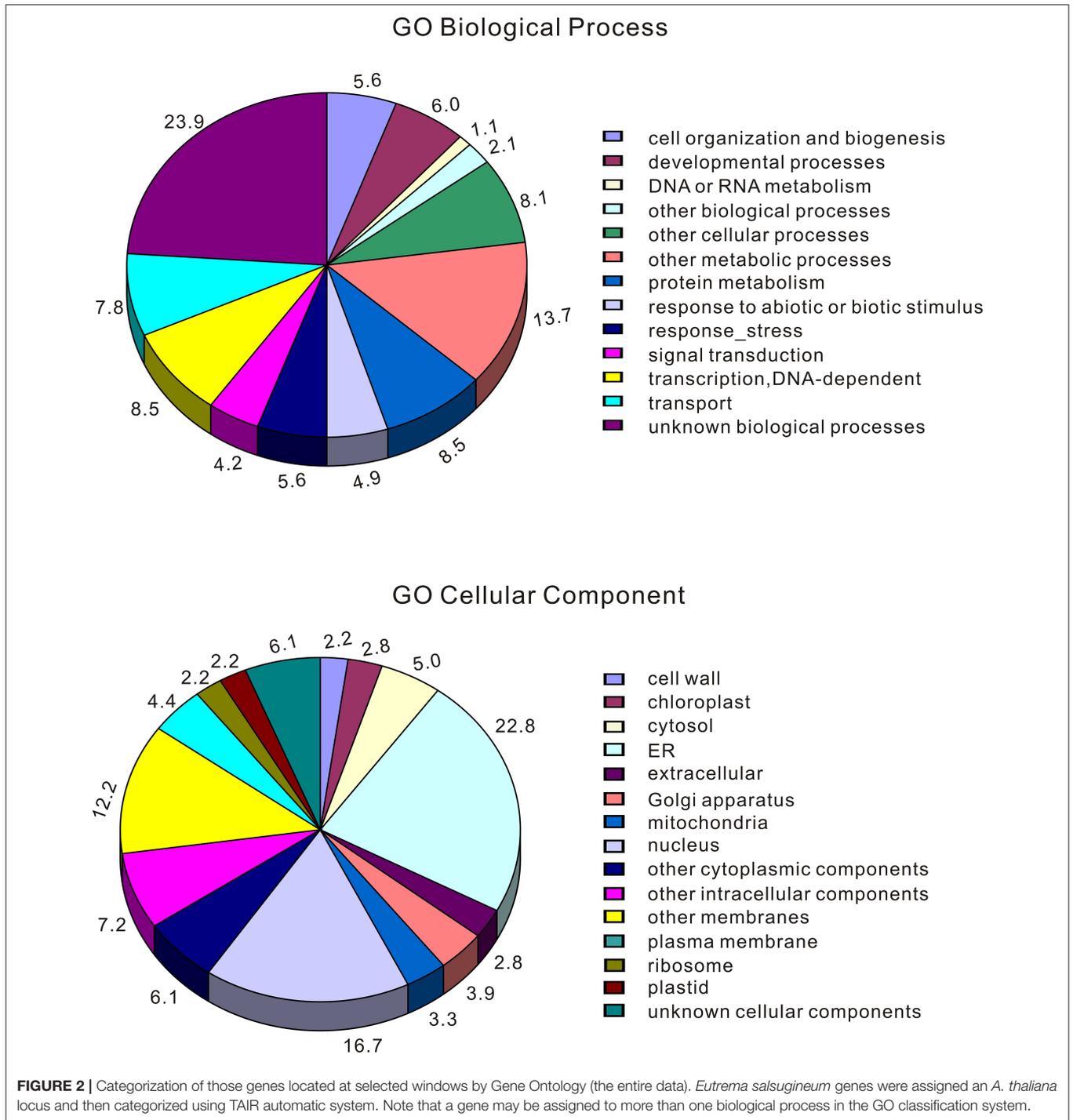
DISCUSSION

Species-Wide Selective Sweep

Salt cress can tolerate salt concentrations up to 500 mmol/L NaCl (Zhu, 2001), allowing it to survive on soils with high salt content. Such habitats would be expected to exert a high selection pressure and reduce genetic diversity in genomic regions involved in salt resistance. Since salt cress is a selfing species in which a high level of linkage disequilibrium is expected, these selective effects could also extend well-beyond the genomic regions under selection. Indeed, LD in salt cress is higher than those of other species examined (Wang et al., 2018) and its genetic diversity is extremely low and similar to those of domesticated rice and

soybean (Wang et al., 2018). Therefore, regions showing lower diversity appeared more likely to be under selection pressure (lower Tajima's D, Pearson correlation coefficient = 0.516, $p < 2.2e-16$), further suggesting that such regions resulted more from selection pressure than from genetic drift or other neutral demographic processes.

Abiotic stress tolerances of plants are likely to involve in several different physiological and developmental pathways, such as osmotic homeostasis, stress damage control and repair, and growth regulation (Zhu, 2001, 2002, 2003; Mahajan and Tuteja, 2005). Some genes among the 227 annotated genes for which selective sweeps were detected operate in pathways important in salt tolerance (Table 1; Supplementary Table 3). For example,



ACO2 is a key member of the ethylene synthesis pathway, while *ERF13* is an important ethylene-responsive transcription factor (Schellingen et al., 2014; Sogabe et al., 2014). Ethylene-mediated signaling pathways have been shown to be critically involved in enhanced salt tolerance in plants (Ryu and Cho, 2015). In addition, *MDAR2* and *NDPK1*, are related to the removal of toxic H_2O_2 (Fukamatsu et al., 2003; Lisenbee et al., 2005),

while *NUDT7* plays a vital role in regulating redox homeostasis during salt stress/defense signaling and programmed cell death in plant disease resistance (Muthuramalingam et al., 2015). Finally, several annotated genes are involved in growth regulation: for instance, *SCR*, a GRAS family transcription factor, regulates stem cell fate of the immediately surrounding cells (Moreno-Risueno et al., 2015), while *PYK10* encodes b-glucosidase in the

TABLE 1 | Gene ontologies (GOs) in the 16 putatively abiotic-biotic stress related genes of *E. salsugineum* (p -value < 0.05).

	<i>E. salsugineum</i> locus	<i>A. thaliana</i> locus	Brief_description	GO_ID	GO_Term	p -value
I	Thhalv10013014m	AT5G08020	RPA70-kDa subunit B (RPA70B)	GO: 0009651	Response to abiotic stimulus	0.00014
	Thhalv10013289m	AT5G07990	Transparent testa 7 (TT7)			
	Thhalv10010179m	AT3G54220	GRAS family transcription factor (SCR)			
II	Thhalv10026186m	AT4G14630	Germin-like protein 9 (GLP9)	GO: 0006970	Response to osmotic stress	0.0002
	Thhalv10029044m	AT4G09320	Nucleoside diphosphate kinase type 1 (NDPK1)	GO: 0009628	Response to salt stress	0.00014
	Thhalv10023595m	AT1G62380	1-aminocyclopropane-1-carboxylate oxidase 2 (ACO2) peptide			
	Thhalv10000830m	AT5G46050	Transporter 3 (PTR3)			
	Thhalv10001865m	AT3G09260	Glycosyl hydrolase superfamily protein (PYK10)			
	Thhalv10003061m	AT4G12720	MutT/nudix family protein (NUDT7)			
	Thhalv10013480m	AT5G03630	Pyridine nucleotide-disulfide oxidoreductase family protein (MDAR2)			
III	Thhalv10021496m	AT3G23180	HR-like lesion-inducing protein-like protein	GO: 0006952	Defense response	0.00038
	Thhalv10027189m	AT2G44840	Ethylene-responsive element binding factor 13 (ERF13)			
	Thhalv10015295m	AT5G22570	WRKY DNA-binding protein 38 (WRKY38)			
	Thhalv10000747m	AT5G45230	Disease resistance protein (TIR-NBS-LRR class) family			
IV	Thhalv10003061m	AT4G12720	MutT/nudix family protein (NUDT7)	GO: 0012501	Programmed cell death	0.0037
	Thhalv10012485m	AT5G17890	DA1-related protein 4 (DAR4)	GO: 0008219	Cell death	0.0054
	Thhalv10000747m	AT5G45230	Disease resistance protein (TIR-NBS-LRR class) family	GO: 0016265	Death	0.0054
	Thhalv10000769m	AT5G47260	Putative disease resistance protein	GO: 0006915 GO: 0006952	Apoptosis Defense response	0.025 0.00038

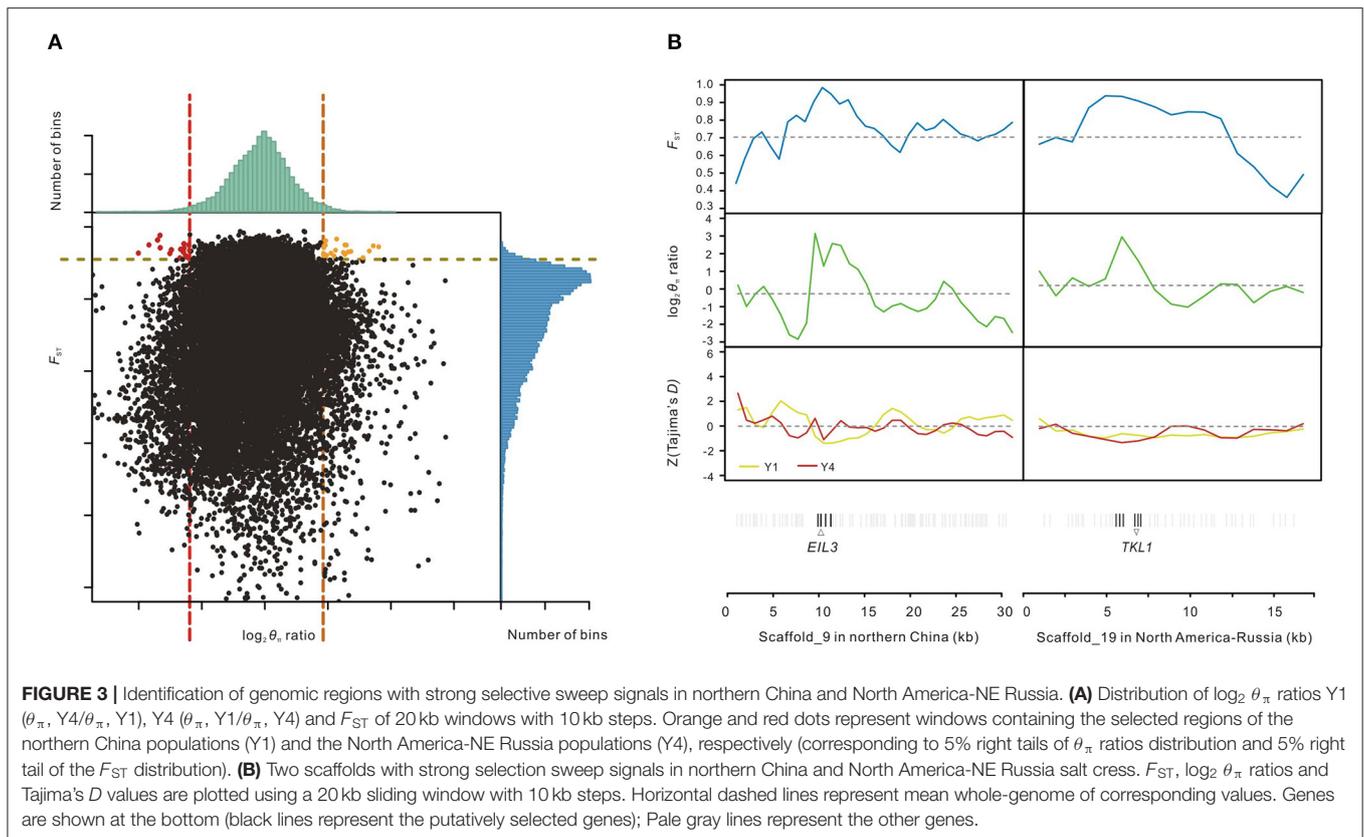
I–IV show different groups of genes with same GO terms.

sub-cellular compartments after wounding (Hara-Nishimura and Matsushima, 2003).

Local Selections in North America-Russia and Northern China

Local selections by the different habitats plays a fundamental role in the production and maintenance of genetic diversity (Savolainen et al., 2013). In the middle Pleistocene, the climate tended to become drier and cooler while desertification and salinization began to develop and expand in central Asia (Wang et al., 2015). These changes might have triggered origin of *E. salsugineum* and its divergence into northern China and NE Russia-North America via two long-distance dispersal ways (Wang et al., 2018), where may be important in local adaptation selected by their different habitats. Functional enrichment analysis of GO terms revealed a remarkable amount of divergence between salt cress populations in northern China and North America-Russia, suggesting genome-wide selection by these local habitats (**Figure 3; Supplementary Tables 4, 5**).

The acclimated freezing tolerance of salt cress was positively correlated with the average minimum habitat temperature (Yang et al., 2013), similar to those of *Arabidopsis* (Hannah et al., 2006; Zhen and Ungerer, 2008). The average minimum habitat temperatures during the coldest month of the growing season at collection sites in North America and Northeast Russia are lower 10 degrees below zero, while that at collection sites in northern China are above zero (Yang et al., 2013). The Yukon cress ecotype from North America-Russia can even tolerate temperatures as low as -19°C (Griffith et al., 2007). Low temperatures induce a number of alterations in cellular components, including the amount of unsaturated fatty acids (Mahajan and Tuteja, 2005) and changes in protein and carbohydrate composition (Lynch and Thompson, 1982). The accumulation of sucrose and other simple sugars that occurs with cold acclimation also contributes to the stabilization of membranes as these molecules protect membranes against freeze-damage (Mahajan and Tuteja, 2005). This was confirmed by the enrichment and classification of GO terms of genes exhibiting selection signals. Most of these annotated genes are involved in carbohydrate



stimulus, peptide biosynthetic and unsaturated fatty acids metabolic process (**Supplementary Table 5**). For example, *ADS2* is critical in the synthesis of unsaturated fatty acids that are an essential component for cold adaptation (Chen and Thelen, 2013) while the gene *STH1* (**Figure 2B**), as the homolog of the Salt Tolerance protein (STO) in *Arabidopsis*, may similarly regulate photomorphogenesis in light signaling in response to low temperatures (Salazar, 2010). In addition, a homolog of two tandemly duplicated genes *TKL1* (**Figure 2B**) in cucumber (*CsTK*) increases both photosynthetic rate and carboxylation efficiency under low temperature and light intensity (Bi et al., 2013). Both *STH* and *TKL1* genes exhibited the higher expressions in Yukon cress ecotype from North America-Russia than in Shandong cress ecotype from northern China when they were grown in the low-temperature common garden or in cabinet (Lee et al., 2012). These findings suggest that salt cress variants found at high latitude in North America and NE Russia could be associated with adaptation to low-temperature habitats.

In northeast China, salt cress usually grows close to vast flood plains with high salinity, extremely wet air and high chemical pollution compared with other populations (Wang et al., 2015). Contaminated soil is a great threat for plants, which can even cause damage to DNA. As a predominantly selfing plant, rapid habitat range expansion could bring about a large amount of harmful mutations (Wang et al., 2018). It is likely

that many of these selected DNA-repair associated genes could efficiently clear up the negative effects of exposure of harmful variants. Two of the over-represented gene ontology categories in northern Chinese populations were “telomere maintenance in response to DNA damage” and “DNA recombination” (**Supplementary Table 4**). Some of these genes, for example, *IPT7*, are involved in cytokinin (CK) biosynthetic process that promotes cell differentiation and regulates root length (Dello Ioio et al., 2012), while *RAD54* is an important eukaryotic-specific recombination factor that plays a critical role in repairing damaged DNA due to radiation or heavy metal contamination (Sung, 1994; Raoul Tan et al., 2003). Moreover, only did it reach northern China, salt cress expanded and reached widespread distributions (Wang et al., 2018) might due to temperate climate conditions along the inland of the Yellow River and stronger resistance to pathogens (Yeo, 2014). Some other selected genes are related to defense against pathogens including “regulation of defense response to fungus” and “regulation of sulfur metabolic process.” Two genes, *SPLAYED* and *EIL3* (**Figure 3B**), are known to play critical roles in enhancing defense against pathogens in biotic stress signaling networks (Van der Ent et al., 2008; Walley et al., 2008). In addition, the gene *WES1* plays a key role in plant sulfur metabolism such as auxin and phytoalexin camalexin biosynthesis in pathogen stress response (Wang et al., 2012). Undoubtedly, genome-wide adaptive divergence further support that salt cress

of northern China could display good resistance to biotic stress in their nature habitat (Pedras and Zheng, 2010; Yeo, 2014).

Due to its natural adaptations to various harsh climates and soil conditions, *Eutrema salsugineum* has long been used as an important model for deciphering mechanisms of salt and other abiotic stress in plants (Gong et al., 2005; Griffith et al., 2007; Lamdan et al., 2012). During adaptive evolution of the salt cress, genetic variation and natural selection are non-randomly fostering stress-related genes, gene interaction network in the whole genome, as well as prompting local adaptation, differentiation and diverse biological stresses between different accessions.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: All Illumina sequence data would be deposited in the National Center for Biotechnology Information short-read archive (project SRP135200).

AUTHOR CONTRIBUTIONS

XW and GZ participated in the design of this study. XW and HR performed the statistical analysis. XC collected important

background information. JM provided assistance for data analysis. XW wrote the manuscript. GL and GZ revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by National Natural Science Foundation of China (Grant No. 31901232), China Postdoctoral Science Foundation (Grant No. 2018M643714), and the Natural Science Foundation of Shanxi Province, China (Grant No. 2020JM-427). This work was also supported by Xi'an Fengdong Yixiang Technology Service Co., Ltd.

ACKNOWLEDGMENTS

We thank Peng Zhao and Bin Hao from Fengdong New City Artificial Intelligence Center for their help in providing computational resources.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.700161/full#supplementary-material>

REFERENCES

- Amtmann, A., Bohnert, H. J., and Bressan, R. A. (2005). Abiotic stress and plant genome evolution. Search for new models. *Plant Physiol.* 138, 127–130. doi: 10.1104/pp.105.059972
- Andolfatto, P. (2005). Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* 437, 1149–1152. doi: 10.1038/nature04107
- Berardini, T. Z., Mundodi, S., Reiser, L., Huala, E., Garcia-Hernandez, M., Zhang, P. F., et al. (2004). Functional annotation of the Arabidopsis genome using controlled vocabularies. *Plant Physiol.* 135, 745–755. doi: 10.1104/pp.104.040071
- Bi, H., Wang, M., Dong, X., and Ai, X. (2013). Cloning and expression analysis of transketolase gene in *Cucumis sativus* L. *Plant Physiol. Biochem.* 70, 512–521. doi: 10.1016/j.plaphy.2013.06.017
- Branca, A., Paape, T. D., Zhou, P., Briskine, R., Farmer, A. D., Mudge, J., et al. (2011). Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula*. *Proc. Natl. Acad. Sci. U. S. A.* 108, E864–870. doi: 10.1073/pnas.1104032108
- Bressan, R. A., Zhang, C., Zhang, H., Hasegawa, P. M., Bohnert, H. J., and Zhu, J. K. (2001). Learning from the *Arabidopsis* experience. The next gene search paradigm. *Plant Physiol.* 127, 1354–1360. doi: 10.1104/pp.010752
- Champigny, M. J., Sung, W. W., Catana, V., Salwan, R., Summers, P. S., Dudley, S. A., et al. (2013). RNA-Seq effectively monitors gene expression in *Eutrema salsugineum* plants growing in an extreme natural habitat and in controlled growth cabinet conditions. *BMC Genomics* 14:578. doi: 10.1186/1471-2164-14-578
- Chen, M., and Thelen, J. J. ACYL-LIPID DESATURASE2 is required for chilling and freezing tolerance in *Arabidopsis*. (2013). *Plant Cell.* 25, 1430–1444. doi: 10.1105/tpc.113.111179
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330
- Dello Ioio, R., Galinha, C., Fletcher, A. G., Grigg, S. P., Molnar, A., Willemsen, V., et al. (2012). A PHABULOSA/cytokinin feedback loop controls root growth in *Arabidopsis*. *Curr. Biol.* 22, 1699–1704. doi: 10.1016/j.cub.2012.07.005
- Evans, L. M., Slavov, G. T., Rodgers-Melnick, E., Martin, J., Ranjan, P., Muchero, W., et al. (2014). Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nat. Genet.* 46, 1089–1096. doi: 10.1038/ng.3075
- Fournier-Level, A., Korte, A., Cooper, M. D., Nordborg, M., Schmitt, J., and Wilczek, A. M. (2011). A map of local adaptation in *Arabidopsis thaliana*. *Science* 334, 86–89. doi: 10.1126/science.1209271
- Fukamatsu, Y., Yabe, N., and Hasunuma, K. (2003) *Arabidopsis* NDK1 is a component of ROS signaling by interacting with three catalases. *Plant Cell Physiol.* 44, 982–989. doi: 10.1093/pcp/pcg140
- Gong, Q., Li, P., Ma, S., Indu, R. S., and Bohnert, H. J. (2005). Salinity stress adaptation competence in the extremophile *Thellungiella halophila* in comparison with its relative *Arabidopsis thaliana*. *Plant J.* 44, 826–839. doi: 10.1111/j.1365-313X.2005.02587.x
- Griffith, M., Timonin, M., Wong, A. C. E., Gray, G. R., Akhter, S. R., Saldanha, M., et al. (2007). *Thellungiella*: an *Arabidopsis*-related model plant adapted to cold temperatures. *Plant Cell Environ.* 30, 529–538. doi: 10.1111/j.1365-3040.2007.01653.x
- Hannah, M. A., Wiese, D., Freund, S., Fiehn, O., Heyer, A. G., and Hincha, D. K. (2006). Natural genetic variation of freezing tolerance in *Arabidopsis*. *Plant Physiol.* 142, 98–112. doi: 10.1104/pp.106.081141
- Hara-Nishimura, I., and Matsushima, R. (2003). A wound-inducible organelle derived from endoplasmic reticulum: a plant strategy against environmental stresses? *Curr. Opin. Plant Biol.* 6, 583–588. doi: 10.1016/j.pbi.2003.09.015
- Huang, X., Kurata, N., Wei, X., Wang, Z. X., Wang, A., Zhao, Q., et al. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490, 497–501. doi: 10.1038/nature11532
- Inan, G., Zhang, Q., Li, P., Wang, Z., Cao, Z., Zhang, H., et al. (2004). Salt cress. A halophyte and cryophyte *Arabidopsis* relative model system and its applicability to molecular genetic analyses of growth and development of extremophiles. *Plant Physiol.* 135, 1718–1737. doi: 10.1104/pp.104.041723

- Kant, S., Kant, P., Raveh, E., and Barak, S. (2006). Evidence that differential gene expression between the halophyte, *Thellungiella halophila*, and *Arabidopsis thaliana* is responsible for higher levels of the compatible osmolyte proline and tight control of Na⁺ uptake in *T. halophila*. *Plant Cell Environ.* 29, 1220–1234. doi: 10.1111/j.1365-3040.2006.01502.x
- Kryukovkhyza, D., Holm, K., Chen, J., Cornille, A., Glémin, S., Wright, S. I., et al. (2016). The influence of population structure on gene expression and flowering time variation in the ubiquitous weed *Capsella bursa-pastoris* (Brassicaceae). *Mol. Ecol.* 25, 1106–1121. doi: 10.1111/mec.13537
- Kumari, A., Das, P., Parida, A. K., and Agarwal, P. K. (2015). Proteomics, metabolomics, and ionomics perspectives of salinity tolerance in halophytes. *Front. Plant Sci.* 6:537. doi: 10.3389/fpls.2015.00537
- Lamdan, N. L., Attia, Z., Moran, N., and Moshelion, M. (2012). The *Arabidopsis*-related halophyte *Thellungiella halophila*: boron tolerance via boron complexation with metabolites? *Plant Cell Environ.* 35, 735–746. doi: 10.1111/j.1365-3040.2011.02447.x
- Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martínez-Barrio, A., et al. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518, 371–375. doi: 10.1038/nature14181
- Lee, Y. P., Babakov, A., Boer, B. D., Zuther, E., and Hinch, D. K. (2012). Comparison of freezing tolerance, compatible solutes and polyamines in geographically diverse collections of *Thellungiella sp.* and *Arabidopsis thaliana* accessions. *BMC Plant Biol.* 12:131. doi: 10.1186/1471-2229-12-131
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, M., Tian, S., Yeung, C. K., Meng, X., Tang, Q., Niu, L., et al. (2014). Whole-genome sequencing of Berkshire (European native pig) provides insights into its origin and domestication. *Sci. Rep.* 4:4678. doi: 10.1038/srep04678
- Lisenbee, C. S., Lingard, M. J., and Trelease, R. N. (2005). *Arabidopsis* peroxisomes possess functionally redundant membrane and matrix isoforms of monodehydroascorbate reductase. *Plant J.* 43, 900–914. doi: 10.1111/j.1365-313X.2005.02503.x
- Lynch, D. V., and Thompson, G. A. (1982). Low temperature-induced alterations in the chloroplast and microsomal membranes of *Dunaliella salina*. *Plant Physiol.* 69, 1369–1375. doi: 10.1104/pp.69.6.1369
- Ma, T., Wang, k., Hu, Q. J., Xi, Z. X., Wan, D. S., Wang, Q., et al. (2018). Ancient polymorphisms and divergence hitchhiking contribute to genomic islands of divergence within a poplar species complex. *Proc. Natl. Acad. Sci. U. S. A.* 115, E236–E243. doi: 10.1073/pnas.1713288114
- Mahajan, S., and Tuteja, N. (2005). Cold, salinity and drought stresses: an overview. *Arch. Biochem. Biophys.* 444, 139–158. doi: 10.1016/j.abb.2005.10.018
- Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., and Cibulskis, K. (2010). The genome analysis toolkit: a map reduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110
- Moreno-Risueno, M. A., Sozzani, R., Yardimci G. G., Petricka, J. J., Vernoux, T., Blilou, I., et al. (2015). Transcriptional control of tissue formation throughout root development. *Science* 350, 426–430. doi: 10.1126/science.aad1171
- Muthuramalingam, M., Zeng, X., Iyer, N. J., Klein, P., and Mahalingam, R. (2015). A GCC-box motif in the promoter of nudix hydrolase 7 (*AtNUDT7*) gene plays a role in ozone response of *Arabidopsis ecotypes*. *Genomics* 105, 31–38. doi: 10.1016/j.ygeno.2014.10.015
- Olson, M. S., Robertson, A. L., Takebayashi, N., Silim, S., Schroeder, W. R., and Tiffin, P. (2010). Nucleotide diversity and linkage disequilibrium in balsam poplar (*Populus balsamifera*). *New Phytol.* 186, 526–536. doi: 10.1111/j.1469-8137.2009.03174.x
- Pang, Q., Chen, S., Dai, S., Chen, Y., Wang, Y., and Yan, X. (2010). Comparative proteomics of salt tolerance in *Arabidopsis thaliana* and *Thellungiella halophila*. *J. Proteome Res.* 9, 2584–2599. doi: 10.1021/pr100034f
- Pedras, M. S. C., and Zheng, Q. A. (2010). Metabolic responses of *Thellungiella halophila/salsuginea*, to biotic and abiotic stresses: metabolite profiles and quantitative analyses. *Phytochemistry* 71, 581–589. doi: 10.1016/j.phytochem.2009.12.008
- Qiu, Q., Wang, L., Wang, K., Yang, Y., Ma, T., Wang, Z., et al. (2015). Yak whole-genome resequencing reveals domestication signatures and prehistoric population expansions. *Nat. Commun.* 6:10283. doi: 10.1038/ncomms10283
- Raoul Tan, T. L., Kanaar, R., and Wyman, C. (2003). Rad54, a Jack of all trades in homologous recombination. *DNA Repair* 2, 787–794. doi: 10.1016/j.dnarep.2003.04.001
- Ru, D. F., Mao, K. S., Zhang, L., Wang, X. J., Lu, Z. Q., and Sun, Y. S. (2016). Genomic evidence for polyphyletic origins and inter-lineage gene flow within complex taxa: a case study of *Picea brachytyla* in the Qinghai-Tibet Plateau. *Mol. Ecol.* 25, 2373–2386. doi: 10.1111/mec.13656
- Ryu, H., and Cho, Y. G. (2015). Plant hormones in salt stress tolerance. *J. Plant Biol.* 58, 147–155. doi: 10.1007/s12374-015-0103-z
- Salazar, F. S. (2010). Functional and molecular characterization of STH1: similarities and differences to the homolog STO. *Universität Freiburg* 8:e23831. doi: 10.4161/psb.23831
- Savolainen, O., Lascoux, M., and Meril, J. (2013). Ecological genomics of local adaptation. *Nat. Rev. Genet.* 14, 807–820. doi: 10.1038/nrg3522
- Schelling, K., Van Der Straeten, D., Vandenbussche, F., Prinsen, E., Remans, T., Vangronsveld, J., et al. (2014). Cadmium-induced ethylene production and responses in *Arabidopsis thaliana* rely on ACS2 and ACS6 gene expression. *BMC Plant Biol.* 14:214. doi: 10.1186/s12870-014-0214-6
- Sogabe, Y., Nakamura, H., Nakagawa, T., Hasegawa, S., Asano, T., Ohta, H., et al. (2014). Visualization of wounding-induced root-to-shoot communication in *Arabidopsis*. *Plant Signal. Behav.* 6, 1037–1039. doi: 10.4161/psb.6.7.15602
- Stamatakis, A. (2015). Using RAxML to infer phylogenies. *Curr. Protoc. Bioinform.* 51, 6–14. doi: 10.1002/0471250953.bi0614s51
- Stepien, P., and Johnson, G. N. (2009). Contrasting responses of photosynthesis to salt stress in the glycophyte *Arabidopsis* and the halophyte *thellungiella*: role of the plastid terminal oxidase as an alternative electron sink. *Plant Physiol.* 149, 1154–1165. doi: 10.1104/pp.108.132407
- Sun, Y. S., Lu, Z. Q., Zhu, X. F., and Ma, H. (2020). Genomic basis of hybrid speciation within chestnut trees. *Nat. Commun.* 11:3375. doi: 10.1038/s41467-020-17111-w
- Sung, P. (1994). Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science* 265, 1241–1243. doi: 10.1126/science.8066464
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. doi: 10.1093/genetics/123.3.585
- Van der Ent, S., Verhagen, B. W., Van Doorn, R., Bakker, D., Verlaan, M. G., Pel, M. J. C., et al. (2008). MYB72 is required in early signaling steps of rhizobacteria-induced systemic resistance in *Arabidopsis*. *Plant Physiol.* 146, 1293–1304. doi: 10.1104/pp.107.113829
- Walley, J. W., Rowe, H. C., Xiao, Y., Chehab, E. W., Kliebenstein, D. J., Wagner, D., et al. (2008). The chromatin remodeler SPLAYED regulates specific stress signaling pathways. *PLoS Pathog.* 4:E1000237. doi: 10.1371/journal.ppat.1000237
- Wang, M. Y., Liu, X. T., Chen, Y., Xu, X. J., Yu, B., Zhang, S. Q., et al. (2012). *Arabidopsis* acetyl-amido synthetase GH 3.5 involvement in camalexin biosynthesis through conjugation of indole-3-carboxylic acid and cysteine and upregulation of camalexin biosynthesis genes. *J. Integr. Plant Biol.* 54, 471–485. doi: 10.1111/j.1744-7909.2012.01131.x
- Wang, X. J., Hu, Q. J., Guo, X. Y., Wang, K., Ru, D. F., German, D. A., et al. (2018). Demographic expansion and genetic load of the halophyte model plant *Eutrema salsugineum*. *Mol. Ecol.* 27, 2943–2955. doi: 10.1111/mec.14738
- Wang, X. J., Shi, D. C., Wang, X. Y., Wang, J., Sun, Y. S., and Liu, J. Q. (2015). Evolutionary migration of the disjunct salt creek *Eutrema salsugineum* (= *Thellungiella salsuginea*, Brassicaceae) between Asia and North America. *PLoS ONE* 10:e0124010. doi: 10.1371/journal.pone.0124010
- Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi: 10.1111/j.1558-5646.1984.tb05657.x
- Wong, C. E., Li, Y., Whitty, B. R., Diaz-Camino, C., Akhter, S. R., Golding, G. B., et al. (2005). Expressed sequence tags from the Yukon ecotype of *Thellungiella* reveal that gene expression in response to cold, drought and salinity shows little overlap. *Plant Mol. Biol.* 58, 561–574. doi: 10.1007/s11103-005-6163-6

- Wu, H. J., Zhang, Z., Wang, J. Y., Oh, D. H., Dassanayake, M., Liu, B., et al. (2012). Insights into salt tolerance from the genome of *Thellungiella salsuginea*. *Proc. Natl. Acad. Sci. U. S. A.* 109, 12219–12224. doi: 10.1073/pnas.1209954109
- Wu, J., Mao, X., Cai, T., Luo, J., and Wei, L. (2006). KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic Acids Res.* 34, W720–W724. doi: 10.1093/nar/gkl167
- Yang, R., Jarvis, D. E., Chen, H., Beilstein, M. A., Grimwood, J., Jenkins, J., et al. (2013). The Reference genome of the halophytic plant *Eutrema salsugineum*. *Front. Plant Sci.* 4:46. doi: 10.3389/fpls.2013.00046
- Yeo, M. (2014). Investigating disease resistance in *Eutrema salsugineum* and the establishment of a *Eutrema*-*p. Syringae* plant pathosystem. *Biology*. Open Access Dissertations and Theses. Identifier: opendissertations/8936/10012/5509193
- Yon, R. S., William, B., Berardini, T. Z., Chen, G., David, D., Doyle, A., et al. (2003). The *Arabidopsis* information resource (tair): a model organism database providing a centralized, curated gateway to *Arabidopsis* biology, research materials and community. *Nucleic Acids Res.* 31, 224–228. doi: 10.1093/nar/gkg076
- Yu, B., and Li, W. (2014). Comparative profiling of membrane lipids during water stress in *Thellungiella salsuginea* and its relative *Arabidopsis thaliana*. *Phytochemistry* 108, 77–86. doi: 10.1016/j.phytochem.2014.09.012
- Zhang, X. D., Wang, R. P., Zhang, F. J., Tao, F. Q., and Li, W. Q. (2013). Lipid profiling and tolerance to low-temperature stress in *Thellungiella salsuginea* in comparison with *Arabidopsis thaliana*. *Biol. Plantarum* 57, 149–153. doi: 10.1007/s10535-012-0137-8
- Zhao, Y. Y., Zhang, R., Jiang, K. W., Qi, J., Hu, Y., Guo, J., et al. (2021). Nuclear phylotranscriptomics/phylogenomics support numerous polyploidization events and hypotheses for the evolution of rhizobial nitrogen-fixing symbiosis in fabaceae. *Mol. Plant.* 14, 748–773. doi: 10.1016/j.molp.2021.02.006
- Zhen, Y., and Ungerer, M. C. (2008). Clinal variation in freezing tolerance among natural accessions of *Arabidopsis thaliana*. *New Phytol.* 177, 419–427. doi: 10.1111/j.1469-8137.2007.02262.x
- Zhu, J. K. (2001). Plant salt tolerance. *Trends Plant Sci.* 6, 66–71. doi: 10.1016/S1360-1385(00)01838-0
- Zhu, J. K. (2002). Salt and drought stress signal transduction in plants. *Annu. Rev. Plant Biol.* 53, 247–273. doi: 10.1146/annurev.arplant.53.091401.143329
- Zhu, J. K. (2003). Regulation of ion homeostasis under salt stress. *Curr. Opin. Plant Biol.* 6, 441–445. doi: 10.1016/S1369-5266(03)0085-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wang, Rao, Ma, Chen, Li and Zhao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.