



# Article Genetic Diversity and Phylogenetic Analysis of Zygophyllum loczyi in Northwest China's Deserts Based on the Resequencing of the Genome

Mengmeng Wei <sup>1,2,†</sup>, Jingdian Liu <sup>1,2,3,†</sup>, Suoming Wang <sup>4</sup>, Xiyong Wang <sup>1,2,5</sup>, Haisuang Liu <sup>4</sup>, Qing Ma <sup>4</sup>, Jiancheng Wang <sup>1,2,5</sup> and Wei Shi <sup>1,2,5,\*</sup>

- State Key Laboratory of Desert and Oasis Ecology, Key Laboratory of Ecological Safety and Sustainable, Development in Arid Lands, Xinjiang Institute of Ecology and Geography, Urumqi 830011, China; weimengmeng21@mails.ucas.ac.cn (M.W.); ariiiiiink@gmail.com (J.L.); wangxy@ms.xjb.ac.cn (X.W.); www-1256@ms.xjb.ac.cn (J.W.)
- <sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China
- <sup>3</sup> College of Forestry and Landscape Architecture, Xinjiang Agricultural University, Urumqi 830052, China <sup>4</sup> State Key Laboratory of Herbage Improvement and Grassland Agro-Ecosystems, College of Pastoral
- <sup>4</sup> State Key Laboratory of Herbage Improvement and Grassland Agro-Ecosystems, College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730020, China; smwang@lzu.edu.cn (S.W.); liuhsh19@lzu.edu.cn (H.L.); maq@lzu.edu.cn (Q.M.)
- <sup>5</sup> Turpan Eremophytes Botanic Garden, The Chinese Academy of Sciences, Turpan 838008, China
- \* Correspondence: shiwei@ms.xjb.ac.cn
- <sup>†</sup> These authors contributed equally to this work.

**Abstract:** In order to study the genetics of local adaptation in all main deserts of northwest China, whole genomes of 169 individuals were resequenced, which covers 20 populations of *Zygophyllum loczyi* (Zygophyllales: Zygophylaceae). We describe more than 15 million single nucleotide polymorphisms and numerous InDels. The expected heterozygosity and PIC values associated with local adaptation varied significantly across biogeographic regions. Variation in environmental factors contributes largely to the population genetic structure of *Z. loczyi*. Bayesian analysis performed with STRUCTURE defined four genetic clusters, while the results of principle component analysis were similar. Our results shows that the Qaidam Desert group appears to be diverging into two branches characterized by significant geographic separation and gene flow with two neighboring deserts. Geological data assume that it is possible that the Taklamakan Desert was the original distribution site, and *Z. loczyi* could have migrated later on and expanded within other desert areas. The above findings provide insights into the processes involved in biogeography, phylogeny, and differentiation within the northwest deserts of China.

**Keywords:** deserts; China; resequencing; genetic diversity; phylogeny; *Zygophyllum loczyi*; Zygophyllaceae

# 1. Introduction

Widely acknowledged as a type of microevolutionary phenomenon, environmental adaptation means the progressive transformation of organisms across generations [1,2]. Differential selection pressures caused by the spatial heterogeneity of the environment on natural populations may cause a species to adapt variably throughout its range [3,4]. While microevolutionary investigations pertaining to this subject are not uncommon, they are frequently carried out on model plants and cash commodities [5–8]. A growing number of interests has been focused on the environmental adaptability benefits of genomic population genetics research [9]. The environmental adaptation and genomic differentiation of *Agriophyllum squarrosum* were investigated by Ma et al. via simplified genome sequencing technology [10]. Insufficient reference genomes for organisms other than models, in addition to the lack of clarity regarding the most suitable sample preparation methods and



Citation: Wei, M.; Liu, J.; Wang, S.; Wang, X.; Liu, H.; Ma, Q.; Wang, J.; Shi, W. Genetic Diversity and Phylogenetic Analysis of *Zygophyllum loczyi* in Northwest China's Deserts Based on the Resequencing of the Genome. *Genes* **2023**, *14*, 2152. https://doi.org/10.3390/ genes14122152

Academic Editor: Mingcheng Wang

Received: 1 October 2023 Revised: 19 November 2023 Accepted: 23 November 2023 Published: 28 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). analyses for various research inquiries and evolutionary time scales, have caused a delay in the application of genomes to the study of adaptation in wild desert plants [11]. The evolutionary history of wild desert plants and their adaptation to environmental change require more consideration [12–14]. Investigations into the population genetics of desert plant differentiation and adaptation not only yield fresh insights into the study of evolution in its natural habitat, but also present a chance to identify stress-resistance genes that may have significant agricultural implications in the face of climate change [15,16].

Based in the mid-latitudes of the heartland of the Eurasian continent, the Northwest Arid Zone of China has undergone substantial plate tectonic processes [17]. The unique topography formed as a result of these geological processes is composed of expansive inland basins interspersed with towering mountain ranges. Desert basins such as the Taklamakan Desert (TKD), Gurbantunggut Desert (GTD), Badanjilin Desert (BJD), Tengger Desert, Kumtage Desert, and Qaidam Desert (QD) are prominent characteristics of this area. These desert basins are separated by towering mountain ranges [18,19]. These deserts share several inherent attributes: arid conditions characterized by infrequent precipitation, a broad annual temperature spectrum that fluctuates between extreme heat and cold, frequent occurrences of winds and sandstorms, and a vegetation community that is sparse and susceptible to damage [20]. Evidence dates back to the early Cretaceous, according to Wu et al. (1995), which suggests that deserts have existed intermittently in China since at least the Pliocene [21]. During the Early Tertiary, the majority of China's sandy regions received subtropical arid vegetation [22]. However, as a result of its extensive scale and geographical diversity, vegetation formation differed across different locations, and contemporary communities cannot be classified as either exclusively younger nor uniformly ancient [23]. Quaternary desert evolution and formation resulted from the combined effects of Ice Age climate variability and Tibetan Plateau uplift [24,25]. The Junggar flora, predominantly influenced by their Central Asian component, emerged in the Quaternary period [26,27]. Floral diversity in the Tarim Basin experienced significant expansion during the Quaternary, having its origins in the Early Tertiary [28–30]. During the Pliocene of the Late Tertiary, a temperate desert emerged in the Qaidam Basin, which underwent further development during the Quaternary [31,32]. During the Quaternary, the desert flora of Alashan underwent significant development, having originated during the Tertiary [33,34]. Populations may experience large-scale replicative gene duplication events when species distributions are negatively impacted by extreme environments [35]. The correlation between environmental stress and polyploidization events is strong, and it has been suggested that polyploidization can enhance organisms' capacity to swiftly adapt to severe environmental fluctuations [36,37]. Many plant species, including Zygophyllum loczyi (Kanitz, 1891) (Zygophyllales: Zygophyllaceae), which has adapted to arid environments, are found in every major desert basin in the region [38]. As a result of combining phylogenetic analysis and population genetic structure, one can discern the sequence of population formation and the mechanisms underlying the dispersal of widespread plants like Z. loczyi. This can provide insights into the overarching characteristics of adaptation and dispersal in the arid regions of Northwest China.

*Z. loczyi* is a C4 herbaceous plant with a life history of one to two years [39,40]. With seventeen species, two subspecies, and three varieties found in China, this genus comprises around 150 species throughout the Old World [38,41,42]. The family of Zygophyllaceae is not only widespread but also prevalent in arid and semi-arid regions, particularly deserts with seasonal dryness [41]. *Zygophyllum* species grow in stony residual dune slopes, fixed and semi-fixed sands, dry riverbeds, gravelly inter-dune flats, and steep loess walls. These species are exceptionally adapted to arid conditions and provide essential ecosystem services in arid environments such as deserts and steppes in the Gobi [43–45]. *Zygophyllum* serves as a fundamental component in arid environments due to its susceptibility to wind erosion, drought tolerance, salinity tolerance, and the capacity to thrive in infertile soils. [39,46,47]. Research on the genus has so far focused on its molecular systematics and genetic diversity [48–50], morph-anatomy [41,50–52],

seed biology [53,54], and genetic and chemical aspects of adaptation [55]. Different *Z. loczyi* phenotypes have resulted from the distinct climatic characteristics of China's desert regions, indicating that local adaptation may be extraordinarily beneficial to comprehend when thinking of plant environmental tolerance.

A total of 169 *Z. loczyi* individuals have been sequenced genetically in this investigation, which spanned four significant desert regions in northwestern China. Our analysis focused on comprehending the potential environmental adaptations of the species in relation to its evolutionary lineage and the geological background of the area.

## 2. Materials and Methods

# 2.1. Sampling and DNA Extraction

A total of 169 plant samples were collected in July 2021 and 2022 from four different desert *Z. loczyi* populations in western China. A total of 28 individuals were from TKD, 35 from GTD, 39 from BJD, and 67 from QD (Table 1). We defined these natural populations as the following four groups: (1) TKD group, (2) GTD group, (3) BJD group, and (4) QD group. Four different regions of *Z. loczyi* seeds were selected to germinate to obtain fresh samples, and the ploidy of each individual was measured by flow cytometry. During sample processing, at least 10 individuals were collected from each population. Detailed records were kept for each sample, including geographic coordinates, elevation, and other environmental conditions at the sampling sites (Figure 1).

Table 1. Population information of Z. loczyi.

Area	Pop	Latitude	Longitude	Amount
	a1	80.912446	41.430356	8
TKD	a2	77.35496	37.60674	13
IKD	a3	77.67286	37.79617	3
	a4	78.256189	37.509027	4
	b1	88.797772	44.94489	12
CTD	b2	89.472407	44.771408	7
GID	b3	89.972732	44.607383	5
	b4	83.328662	44.574587	11
	c1	97.233058	37.124617	6
	c2	95.60336	37.458975	11
OD	c3	97.334042	37.141595	15
QD	c4	95.377427	37.572502	8
	c5	95.287567	37.88953	10
	c6	91.039958	38.098013	14
	d1	100.559867	39.710435	5
	d2	100.802402	39.587463	9
רוק	d3	98.796332	39.895022	3
БJD	d4	101.515437	39.19226	3
	d5	103.137744	41.685669	17
	d6	102.925948	38.442503	10



Figure 1. Map of Z. loczyi sampling points. Background filled by elevation as color.

#### 2.2. Determination of DNA Content by Flow Cytometry

Total genomic DNA was extracted from leaf tissues using the Cetyl Tri-methyl Ammonium Bromide (CTAB) method [56]. The DNA quality and concentration were assessed using 1% agarose gel electrophoresis and a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). For resequencing, library construction and sequencing were conducted at Biomarker technologies (Beijing, China) on an Illumina platform (Illumina HiSeq 4000 PE150, Santiago CA, USA), employing a 300-bp read length.

Live samples from different distribution areas of *Z. loczyi* individuals were selected and rinsed repeatedly under running water for 30 s 3–5 times, then dried with tissue paper and set aside. Leaves were digested with both WPB disassociation solution and GLB disassociation solution, respectively, to screen for the suitability of different disassociation solutions. Ploidy was determined using DAPI solution (20 mg/L) staining under UV light during flow cytometry. Genome size was detected using PI solution (20 mg/L) staining and flow cytometry at 632 nm frequency [57–59]. We used *Populus tomentosa* (Malpighiales: Salicaceae) leaves as the reference standard.

# 2.3. Genome Resequencing, Assembly, and Annotation

After the evaluation and qualification of the genomic DNA sample, it underwent fragmentation by ultrasound-induced mechanical interruption [60]. The produced fragments were subsequently cleaned by fragment purification, end repair, 3'-end addition of A, connection of sequencing junctions, agarose gel electrophoresis to select fragment size, and PCR amplification to create a sequencing library [61]. Clean Reads were obtained after the Raw Reads were filtered to eliminate those containing adapters, exceeding 10% N content, or more than 50% bases with a quality value below 10 [62].

As the sequencing accuracy escalates in relation to the length of the sequenced reads, the quality values had been transformed into error rates and executed the base type of distribution analysis to detect the existence of AT and GC segregation [63]. Due to the fact that *Z. loczyi* is known as a wild plant and acquiring the reference genome of close relatives has a stronger challenge, *Zygophyllum*. *Xanthoxylum* is selected by us, which is also a species of the *Zygophyllum* genus, as the reference genome [44]. It is necessary to transfer the clean sequences obtained by sequencing to the reference genome. Therefore, we compared the Clean Reads with the reference genome using bwa-mem2 (v2.2) software, sorted the results using samtools (v1.9) sort comparison, and statistically calculated the sequencing depth and genome coverage of each sample based on the sorted results [64,65].

We determined the starting and ending positions of the reference genome's doubleended sequence. The CollectInsertSizeMetric.jar application from the Picard (v2.25.5) software toolset is used for calculating the insert fragment's size subsequent to the interruption of the sample DNA [66,67].

#### 2.4. SNP and Variant Detection and Annotation

SnpEff [4] is software made to identify the impact of variants and to annotate variants [68]. To ensure the reliability of SNPs, the statistical cumulative distribution of distances between neighboring SNPs is used along with the number of reads that correlate with the detected SNPs [69]. The finding of the variant locus's site and the consequence of the variant can be accomplished by utilizing the reference genome's gene position information combined with the variant locus's position.

Detection of SNPs and InDels was performed using GATK (v3.8) [70]. To ensure the accuracy of the detection results, redundant reads were filtered using samtools (v1.9) based on the alignment of cleaned reads to the reference genome [64,65]. Subsequently, the GATK HaplotypeCaller algorithm was employed for SNP and InDel variant detection. Through filtering, a final set of variant sites was obtained and stored in VCF format [71]. Using the vcfutils.pl subroutine of bcftools (var Filter-W 5-W 10), SNPs is filtered out SNPs in the 5 bp range of InDels and neighboring InDels in the 10 bp range. Cluster Size is set to 2 and Cluster Window Size to 5, indicating that the number of variants in a 5 bp window should not exceed 2. We filtered out variants with quality scores below 30, QD values below 20, FS values above 60, and/or MQ values below 40. Other variant filtering parameters followed the default values specified by GATK. Making use of the Circos (0.69-9) software, the distribution of the results for each type of mutation obtained from the assay was plotted [72].

The annotations of these genes were accessible for the purpose of analyzing the functions of the genes through the comparison of variant genes with functional databases maintained by Diamond, including NR, Swiss Prot, GO, COG, and KEGG [73–77].

## 2.5. Genetic Evolution Analysis

The population structure and admixture are inferred among our 169 samples using MEGA X (https://www.megasoftware.net/, accessed on 25 July 2023) under the Kimura 2-parameter model; clade support was calculated using 1000 bootstrap replications [78]. We also performed clustering analyses as a complimentary way to detect genetic structure. The population genetic structure of *Z. loczyi* was assessed by employing ADMIXTURE (v1.22) and utilizing high-quality SNPs [79]. The most likely number of clusters was computed with 10-fold cross-validation (CV), comparing K-values from 2 to 10.

A PCA based on SNP using the smartPCA program (https://data.broadinstitute.org/ alkesgroup/EIGENSOFT/EIG-6.1.4.tar.gz, accessed on 25 July 2023) in EIGENSOFT also be created (v6.0) (https://www.megasoftware.net/, accessed on 25 July 2023) to study genetic relatedness and clustering among populations [80]. Finally, we created a kinship heat map for estimation of kinship between any two individuals using GCTA (v1.92.1) (https://yanglab.westlake.edu.cn, accessed on 25 July 2023) [81]. The PopLDdecay has been used (v3.41) to estimate linkage disequilibrium (LD) decay based on the coefficient of determination (r<sup>2</sup>) between any two loci (https://github.com/BGI-shenzhen/PopLDdecay, accessed on 25 July 2023) [82]. The Plot\_MultiPop.pl script that comes with the software was then used to plot the decay curve.

Diverse population genetics metrics were computed utilizing the VCFtools (0.1.15) software utility, with a sliding window of 100 kb and a step size of 10 kb, the SNPs that exhibited the highest degree of consistency [71].

3.1. Quality Control of Sequencing Data

3.1.1. Genome Size and Sequencing

As *Z. loczyi* is a non-model species, we used *Z. xanthoxylum* for a reference genome (NCBI BioProject PRJNA933961). By flow cytometry, we determined that the *Z. loczyi* chromosomal ploidy is diploid, with a genome size of approximately 500 Mb. (Figure 2). A total of 1491.98 Gbp of genome-pure data were obtained by resequencing, with Q30 reaching 91.96–95.75% and an average GC content of 34.28%. The alignment rate between the sample and the reference genome was about 60.77%, while average coverage depth was average  $3.81 \times$  (Supplementary Table S1).



**Figure 2.** Four different *Z. loczyi* populations' DNA content and ploidy measured by 670-30A Dualbeam Infrared Spectrophotometer. The excess spectral absorption peaks may be a result of uneven cell staining.

3.1.2. Analysis of Base Sequencing Quality Distribution

During the execution of base sequencing quality distribution analysis, it was observed that the samples which include the final dozen bases and the first four bases show lower quality values compared to the intermediate sequencing bases. However, all of these samples carried quality values more than Q30%. To illustrate that, we transformed the quality values into error rates and graphically represented the error rate distribution as follows (Figure 3). The examination of base type distribution showed that AT and CG bases were basically not separated, the curve was gentle, and the sequencing results were normal (Figure 4).



**Figure 3.** Distribution of base error rate among part of *Z. loczyi* samples. The horizontal coordinate is the base position of the Reads, and the vertical coordinate is the single base error rate. The first 150 bp is the distribution of error rate of the first end of the sequenced Reads of the bipartite sequenced sequence, and the last 150 bp is the distribution of the error rate of the other end of the sequenced Reads.



**Figure 4.** Distribution of the proportion of each base of the bases in some samples of *Z. loczyi*. The horizontal coordinate is the base position of the Reads, and the vertical coordinate is the proportion of bases; green represents base G, blue represents base C, red represents base A, purple represents base T, and grey represents base N that was not identified in sequencing. The first 150 bp is the base distribution of the sequenced Reads of the bipartite sequencing sequence, and the last 150 bp is the base distribution of the first end of the sequenced Reads of the other end of the sequence. The first 150 bp is the base distribution of the sequenced Reads of the other end of the sequences, and the second 150 bp is the base distribution of the sequenced Reads of the other end.

## 3.1.3. Analysis of Reference Genome Comparisons

Comparison with the reference genome has shown that there is no contamination in the experimental process, and graphing based on the depth of coverage of each chromosome locus shows that the genome is covered more evenly, indicating better sequencing randomness. The uneven depth on the graph may be due to repeated sequences, PCR preference.

By detecting the start and stop positions of the bipartite sequences on the reference genome, the precise measurements of the sequenced fragments acquired subsequent to the interruption of the sample DNA could be ascertained. This analysis confirmed that the length distribution of the insert fragments followed a normal distribution, suggesting that the library construction of the sequencing data was normal.

After localization to the reference genome, the number of Reads can be discovered with the quantification of base coverage on the reference genome (Figure 5). A more uniform distribution of bases on the genome in terms of coverage depth suggests that the sequencing randomness has been enhanced. Figure 6 below illustrates the coverage distribution curve and base coverage depth distribution curve of the samples (Figure 7).

Genomewide distribution of read coverage		Genomewide distribution of read coverage			
	ada minana kaomina atao matao matao karina matao kan Arel na karina matao matao matao matao matao matao matao m	LG01	4-	physical a combines on a new concernment on cool and all the decomplete constructions of the second metric occ	LG01
	And so, and entertied many life colorest and device and deriver discussion of devices	LG02	LG02	elikiti, "anakoaniki kuserkusatan asaran masarin sasihiran, antikan da ah, san ber	LG02
1	al accellations had som fahren fat in flick anoradis freminister des ausentikalistissen	LG03	1	e la tra trible a chura, ann a la nadacha fhait an ambhait na mhar an bha ann bha na an ann t	LG03
(3) 1 1 1 1 1 1 1	sternal ano an exclusive acceleration and an ensure of prosted and access defined two tests and a	LG04	(2))	resettlen ussemen den annendersche annen franzöhnen der nach bei sichtlichten von stem sich	LG04
sity(log	parentes for an exact of the and developments are reached and the statements of the statements of	LG05	isity(log	entre et da la provincia da cara la da cara la da cara da cara conserva en aporte da la cara da cara.	LG05
ad deins	secure and a community of a device of the later, set was reached and an associated by a set	LG06	ad dein	et en transmission alles des tables, transmissioned, second addes data	LG06
an of re	were leaves an amount of warrant with the anneadar mouth of the second of the	LG07	an of re	attelete Lancon unbolestore, antalteen avolution door avoid the	LG07
Media	energlas are diferentes concernentes de arrevener en entre de concernentes en entre de	LG08	Medi	and cardelings, and transmission as the construction of the constr	LG08
-	in the maximum and with the out on a low of the second dama	LG09	-	n sans a na saidh na bha tha tha ann ann ann ann ann ann ann ann ann a	LG09
	province and the state work on the deformance and all the bridge	LG10	-	adaabka, waxaa aaniida ay karooskii daha	LG10
1	e de constituer en sus la tre la contracte en activitador de la constitución	LG11	-	rices of the monotonic activity of strandard s	LG11
	0kb 2000kb 4000kb			0kb 20000kb 40000kb	
Genomewide distribution of read coverage Genomewide distribution of read coverage					
G	Chromsome position enomewide distribution of read coverage		G	Chromsome position	
G ::	Chromsome position enomewide distribution of read coverage adulter rate dute as the anti-orten durant of the order to be the trade of t	LG01	G	Chromsome position ienomewide distribution of read coverage plates werd instruction and their records with plates and successful as the second second second second second	LG01
G ;; ;; ;;	Chromisone position enomewide distribution of read coverage shalehe call, do but a transmitted a sol whick introduct, and it is no a collision or a of have another but at the net on occurs on a column at the metral base but at the constraint the the the track occurs on a column at the metral base but	LG01 LG02	110778200 10078200	Criminante position ienomewide distribution of read coverage platine, a crist-ten end societati bala are ado, and platine, and positive to social and social advoc, incorrected models are and a cristian to addition to social, distributions	LG01 LG02
G 	Chromosine position enomewide distribution of read coverage skelder out a base attender at the mitoder coal chickenemies, som it is at as a coffee brance of the encouncies the device the operation of the test blowder but generat skeler is due to the device booms and an activity of the matter skeler is due to the device booms and an activity of the matter	LG01 LG02 LG03	110078888 10078880 10078880	Commone position ienomewide distribution of read coverage plates a studietus and socialisti di ale sa adas at di plata a sult, sutificadore sudiestinate adares tanonalistana data suti a suto accesto a adare a successiva di ale seria associalisto, itali serializzati du bin lastaniona de a tra attesta la suto	LG01 LG02 LG03
3(2)) 	Chromose position enomewide distribution of read coverage adultional abdus a threa distribution of the distribution of the distribution of the distribution of have an overable to a like the thread of the other of the distribution of the distribution abdus a distribution of the distribution of the distribution of the distribution of the distribution abdus a sector of the distribution of the distribution of the distribution of the distribution and white product the distribution of the distribution of the distribution of the distribution and white product the distribution of the distribution of the distribution of the distribution of the distribution and white product the distribution of the distribu	LG01 LG02 LG03 LG04	(2)) 222255 222255 222255 222255	Chromosome position internetwide distribution of read coverage plata successful and successful and a static static successful and a static of the successful and the same static scale at the successful district successful the static successful distribution of a static static static successful the static successful distribution of a static static static successful the static static static static static static static static static successful the static static static static static static static static static static static distributions at the static static static static static static static static distributions at the static static static static static static static static distributions at the static static static static static static static static distributions at the static	LG01 LG02 LG03 LG04
nsity(log(2))	Chromosme position enomewide distribution of read coverage aktélional aludia anna menninka karola disk kelonania semi karola sena disk herman efikase, enomendek to si blan kelo eos en solman si Vitemiera la tekas akterisé disk i aludia se kiser kelo tekase barron karba dasak hi koman enomekter momentes attele se ato per difensive dasen se e ters sont ker a senakter momentes attele se ato per difensive dasen se e ters sont ker	LG01 LG02 LG03 LG04 LG05	stity(log(2))	Chromosome position Chromosome position and the second sec	LG01 LG02 LG03 LG04 LG05
sad deinsity(log(2)) 고프트 코프트 코프트 코프트 코프트 <u>프</u> 트	Chromosme position enomewide distribution of read-coverage patrichtenetic dystanten en er die ekstematierke of felsk bieneties over it an in een effective effective enoemelie twe if the decision enough an ex Vitemiers in tender einer einder it de een titer ekste door enough en ex Vitemiers in tender enoekter ween er die een titer ekste door enough en ex vitemiers in tender enoekter ween er die een die ekste enough en ekster een ekster enoekter ween ere door en door een ekster een eksteren.	LG01 LG02 LG03 LG04 LG05 LG06	ad deinsity(log(2))	enconsome position inconservide distribution of read-coverage plates are endistribution of read-overage descent and the second s	LG01 LG02 LG03 LG04 LG05 LG05
ian of read deinsity(log(2)) 2 고 후 2 고 후 2 고 후 2 고 후 2 고 후 2 고 후 0	Chromosme position enomewide distribution of read coverage physicile cells dynamic a secretize de also an objete et al dynamic a secretize distributions also and also also an etale also also an objete et allo de also de also de also also also also also an etale also also anon duran est allo missione de also de also also also also also antico also also antico also also also also also also also also also also also also also also	LG01 LG02 LG03 LG04 LG05 LG05 LG05	an of read deinsity(log(2))	envoisione position incomerciale distribution of read-coverage plates, we reliance and second activity of the seco	LG01 LG02 LG03 LG04 LG05 LG05 LG05
Median of read deinsliy(log(2)) 한고훈 감각은 감각은 장고훈 감각은 감각은 감각은 요구한 요	Chromose position enormewide distribution of read coverage physicile cells define a second as the coverage at here, encouncel in a file of the definition of	LG01 LG02 LG03 LG04 LG05 LG05 LG05 LG07	Median of read deinsity(log(2))	encomervide distribution of read coverage plateau overfide mentes a coverage plateau overfide mentes a coverage coverage plateau overfide mentes a coverage coverage cover a cover a coverage distribution of a cover plateau overfide mentes a coverage coverage coverage coverage distribution of a cover plateau overfide mentes are done in the coverage in the coverage distribution of a cover plateau overfide mentes are done in the coverage in the coverage over a coverage plateau overfide mentes are done in the coverage in the coverage over a coverage plateau overfide mentes are done in the coverage in the coverage over a coverage and coverage overfide a coverage over a coverage over a coverage over a coverage and coverage overfide a coverage over a coverage over a coverage and coverage overfide a coverage over a coverage of a coverage over a coverage and coverage overfide a coverage over a coverage of a coverage over a coverage and coverage overfide a coverage over a coverage of a coverage over a cov	LG01 LG02 LG03 LG04 LG05 LG05 LG05 LG05
Median of read deinsity(log(2)) 고프는 프프폰 프프폰 프프폰 프프폰 프프폰 프프폰 프프폰 D	Chromewide distribution of read coverage adultational data and adultation of the distribution of the distr	L001 L002 L003 L005 L005 L005 L005 L009	Median of read deinsity(log(2))	envestide distribution of read coverage plates, so and second se	LG01 LG03 LG04 LG05 LG05 LG05 LG07 LG08 LG09
Median of read deinsity(log(2)) 고프는 고프는 고프는 고프는 고프로 코프트 <u>고프는 고프는</u> 요	Chromewide distribution of read coverage addeleteral algebra an annalytic of the distribution of read coverage addeleteral algebra an annalytic of the distribution of	L.092 L.092 L.093 L.095 L.095 L.095 L.095 L.099 L.099 L.099	Median of read deinsity(log(2))	enormowide distribution of read coverage plates a conditioned to some other and a strategies of the latest and a strategies of the solution o	LG01 LG02 LG03 LG04 LG06 LG06 LG06 LG00 LG09 LG10
다고로 고고픈 고고트 주고트 고고로 고고로 고고로 고고로 고고로 요구가 요	Chromese position enouver de distribution of read coverage elicitette distribution of the distribution	LG01 LG02 LG04 LG06 LG06 LG06 LG09 LG09 LG10	Median of read deinsity(log(2))	common position     information position     information position     information definition of read coverage     plates, and the second	LG01 LG02 LG03 LG04 LG05 LG05 LG05 LG05 LG05 LG11

**Figure 5.** Chromosome coverage depth distribution of some samples of *Z. loczyi*. The horizontal coordinate is the chromosome position and the vertical coordinate is the value obtained by taking the logarithm of the depth of coverage at the corresponding position on the chromosome.



**Figure 6.** Distribution of *Z. loczyi* insert fragments. The horizontal coordinate is the length of the inserted segment and the vertical coordinate is its corresponding number of Reads.



**Figure 7.** Distribution in depth of a selection of *Z. loczyi* samples. The above figure reflects the basic distribution of sequencing depth, with the horizontal coordinate being the sequencing depth; the left vertical coordinate being the percentage of bases corresponding to that depth, which corresponds to the red curve; and the right vertical coordinate being the percentage of bases at and below that depth, which corresponds to the blue curve.

# 3.1.4. SNP Identification and Quality Control

To provide a genome-wide overview of the dynamics underlying local adaptation, a total of 169 *Z. loczyi* individuals were collected from 20 natural populations across their current distribution in China (Figure 1). Based on these population samples, our genome resequencing approach yielded 232,724,423 high quality SNPs (allele frequency > 0.05 and integrity > 0.8) which were used for subsequent population genetic analyses (Figure 8). To ensure the reliability of the SNPs, we examined the cumulative SNP depth distribution to identify the predominant SNP types and their frequencies. Within the 25–75% interval, the SNPs displayed high depths with pronounced peaks, suggesting that the SNPs are of better quality (Figure 8).



**Figure 8.** A cumulative plot of the number of SNP Reads supported is shown on the left, and a cumulative plot of the distance between neighboring SNPs is shown on the right.

#### 3.1.5. Detection and Distribution of Variation

A total of 150,819,465 SNPs were detected, with a Het-ratio (heterozygosity/homozygosity) of 0.65% to 2.99%. The Ti/Tv (Transition/Transversion) ratio ranged from 1.38 to 1.43. These values are based on a Ti range of 419,115–607,294 and a Tv range of 295,847–437,912, which correspond to different samples (Supplementary Table S1). A comprehensive analysis of the detected SNPs revealed distinct distribution patterns among different genomic regions. Among all the SNPs identified, 18.85% were classified as intergenic, 25.79% were found in intronic regions, and 31.94% were within CDS (Figure 9). Notably, among the CDS SNPs, a significant proportion consisted of non-synonymous coding variants (15.47%) and synonymous coding variants (15.20%) (Figure 9). These findings highlight the prevalence of genetic variation within protein-coding regions, with potential functional implications associated with both non-synonymous and synonymous alterations.



**Figure 9.** Classification results based on the reference genome SNP annotation of all samples. Proportions of the various SNPs in the *Z. loczyi* and the reference genome of *Z. xanthoxylum*.

A total of 1,296,479 InDels were detected in the dataset. The heterozygosity ranged from 2866 to 12,552, while the homozygosity ranged from 360,119 to 701,259. The Het-ratio varied from 0.75% to 2.16% (Supplementary Table S1). In terms of distribution across different genomic regions, introns accounted for 0.35% of the total InDels, intergenic regions represented 0.31%, downstream non-coding regions accounted for 0.10%, upstream non-coding regions represented 0.09%, and the CDS accounted for 0.06% (Figure 10). Within the CDS category, the main subtypes of InDels were frameshifts (0.04%) and codon-insertions (0.006%) (Figure 10). These findings provide insights into the prevalence and distribution of InDels, including within protein-coding regions, suggesting potential functional implications of genetic variation in the studied population.





The SNP density across various chromosomes is depicted in Figure 11. Chromosome 1 exhibited the highest density of SNPs, with a count of 325,704 SNPs, while chromosome 9 displayed the lowest SNP density, comprising 132,516 SNPs (Figure 11). Within each chromosome, the distribution of polymorphism was uneven, encompassing both densely populated and sparsely populated regions of SNPs.



**Figure 11.** Distribution of SNPs and InDels detected in the *Z. loczyi* and the reference genome of *Z. xanthoxylum*, in the 11 chromosomes (color block = chromosome coordinates, green line = gene density distribution, orange line = SNP density distribution, purple line = InDel density distribution).

# 3.1.6. Genomic Signals of Adaptation

GO analysis was performed to elucidate gene functions across three major categories: biological processes, cellular components, and molecular functions (Figure 12). The GO analysis of biological processes revealed the involvement of genes in various essential biological activities. These processes ranged from fundamental cellular functions such as metabolism, cell cycle regulation, and signal transduction, to more specialized processes like immune response, development, and neuronal signaling. In terms of cellular components, the GO analysis provided insights into the localization and organization of gene products within cells. The variant gene COG categorization statistics revealed that the most prevalent items were T (signal transduction mechanisms), G (carbohydrate transport and metabolism), R (general function prediction only), and J (translation, ribosomal structure, and biogenesis) (Figure 13).



Figure 12. The SNPs of Z. loczyi annotation clustering according to the GO.





#### 3.2. Genetic Evolution Analysis

# 3.2.1. Genetic Diversity

Based on the population structure of *Z. loczyi*, we calculated seven genetic indices (MAF, Ae, Ao, He, Ho, PIC, and I) for each clade and population. The MAF across the four clusters ranged from 0.25 to 0.28, demonstrating relatively consistent values. The QD clade exhibited the highest genetic diversity (He = 0.365), followed by the TKD clade (He = 0.353) and the BJD clade (He = 0.333), while the GTD clade had the lowest genetic diversity (He = 0.318) (Table 2). These findings suggest that within the other three populations, there exists a non-random distribution of genotypes among individuals, possibly attributable

to selection for specific beneficial genotypes or a heterozygote advantage at polymorphic loci. In contrast, the QD population demonstrated the Ho lower than the He, implying a genotype distribution closer to random among individuals in this group, devoid of discernible selective advantages or excess heterozygosity effects.

Group	MAF	Ae	He	Nei	Poly Marker	Ao	Ho	PIC	Ι
TKD	0.27	1.000-2.000 (1.441)	0.035–0.500 (0.353)	0.036–0.533 (0.360)	45,656	1.000–2.000 (1.718)	0.036–1.000 (0.414)	0.034–0.375 (0.280)	0.090–0.693 (0.523)
GTD	0.25	1.000–2.000 (1.370)	0.028–0.500 (0.318)	0.029–0.517 (0.323)	42,241	1.000–2.000 (1.664)	0.029–1.000 (0.388)	0.028–0.375 (0.253)	0.075–0.693 (0.476)
QD	0.28	1.000–2.000 (1.500)	0.015–0.500 (0.365)	0.015–0.507 (0.368)	50,167	1.000–2.000 (1.789)	0.015–1.000 (0.362)	0.015–0.375 (0.290)	0.044–0.693 (0.540)
BJD	0.26	1.000–2.000 (1.375)	0.025–0.500 (0.333)	0.026–0.516 (0.337)	41,169	1.000–2.000 (1.647)	0.026–1.000 (0.404)	0.025–0.375 (0.265)	0.069–0.69 (0.498)

Table 2. Genetic diversity of the four deserts.

(MAF = average MAF, Ae = expected allele number, He = expected heterozygous number, Nei = Nei diversity index, Mp = number of poly markers, Ao = observed allele number, Ho = observed heterozygous number, PIC = Polymorphism information content, I = Shannon–Wiener index).

When using Nei's diversity index, the mean values for the four groups were as follows: TKD = 0.36, GTD = 0.323, QD = 0.368, and BJD = 0.337. Based on these mean values, the QD group displayed the highest Nei's diversity, while the GTD group had the lowest. In this study, all populations showed medium variation (0.25 < PIC < 0.5). We also calculated the Shannon Information Index for each of the four populations: TKD (0.523, 0.09-0.693), GTD (0.476, 0.075-0.693), QD (0.540, 0.044-0.693), and BJD (0.498, 0.069-0.69). The TKD group had the highest number of polymorphic markers (45,656), while the BJD group had the lowest number of polymorphic markers (41,169). These findings demonstrate the diversity and complexity of information across these groups. Despite the relatively low average values, the wide distribution suggests the presence of distinct sources of genetic information and unique characteristics within each group.

#### 3.2.2. Phylogenetic and Population Genomic Analyses

The optimal ancestral clustering at K = 4 was determined based on the cross-validation error rate (Figure 14). The geographic divisions observed in the population align closely with the actual geographic divisions.



**Figure 14.** (a) Clustering results of samples corresponding to each of the ADMIXTURE population genetic structure; (b) Genetic structure analysis of *Z. loczyi* based on the Bayesian model, The red dot represents the appropriate K value.

We also reconstructed the phylogenetic relationship of the 20 populations based on the same SNP dataset using the neighbor-joining method. The results are generally consistent with the population structure detailed above; however, the QD group is further divided into two subgroups (Figure 15a). Principal component analysis (PCA) further supported the existence of four distinct groups among the 20 populations (Figure 15b). Notably, although *Z. loczyi* exhibited a distinct spatial structure according to various genomic methods, a relatively small amount of genetic variation was observed. Additionally, PCA and AD-MIXTURE analyses based on the Bayesian algorithm corroborated the population structure observed in the phylogenetic tree. The optimal clustering solution for the populations was K = 4. Similarities existed in terms of population composition and geographic dispersion.



**Figure 15.** (a) Phylogenetic trees were generated for each sample by employing neighbor-joining with 1000 bootstrap replications and the Kimura 2-parameter model.; (b) the sample is clustered in two dimensions using principal component analysis (PCA), where PC1 and PC3 denote the first and third principal components, respectively. A color denotes a group, while a dot represents a sample.

# 3.2.3. Linkage Disequilibrium Decay Analysis

The LD between any two SNPs within a certain distance range (20 kb) was calculated on the same chromosome, and the strength of linkage disequilibrium was expressed as  $r^2$ . To assess the level of linkage disequilibrium in the 20 populations, genome-wide SNPs were applied to map the attenuation of the different populations. The GTD and BJD populations had lower levels of LD ( $r^2$  values) than the TKD and QD population groups (Figure 16).



**Figure 16.** Linkage disequilibrium (LD) is a measure of whether genotypic changes in two molecular markers are in step and correlated.

## 4. Discussion

The technology of resequencing sequencing contributes significantly to the investigation of the genetic information of a vast array of species, particularly non-model organisms [5–8]. Through flow karyotyping, which detects alterations in chromosome number and structure, we analyzed chromosomal polymorphisms [83,84]. SNP and InDel mutation rates can be accelerated by the polyploidy of plant chromosomes under unfavorable conditions, which can hinder the detection and analysis of these genetic variants within the genome [37,85,86]. As a result, flow cytometric karyotype analysis is of the utmost importance in plant genomics, which provides essential information for subsequent genome sequencing, SNP detection, and genome assembly by facilitating the prediction of the number and structure of chromosomal variants [87,88]. Genomic DNA sequences frequently comprise an extensive number of SNPs and InDels, which can be efficiently detected and exhaustively examined through the utilization of high-depth resequencing technology [7]. Subsequent information analysis made use of sample base error rates, base type distribution checks, maps showing the depth distribution of sample chromosome coverage, statistics on the distribution of insertion fragments, and sample depth distribution posts. Moreover, the assessment of GC content holds significance as it is considered a characteristic feature in genome organization [89]. The customary spectrum of GC content in eukaryotic genomes extends from 30% to 65% [90]. This study's GC concentrations fell well within this range, indicating that the sequencing data were accurate [91].

When PIC  $\geq$  0.5, the locus is considered highly polymorphic. For  $0.25 \leq$  PIC < 0.5, the locus is moderately polymorphic, while a PIC < 0.25 indicates low polymorphism. Based on our results, the genetic diversity observed in *Z. loczyi* falls within the range of  $0.25 \leq$  PIC < 0.5, indicating moderate genetic diversity. Adaptive genetic variation is influenced by various factors such as geology, climate, and altitude [92]. The values of He and Ho were lower in the GTD region than in the other three regions. We hypothesize that natural selection will likely favorably select environmentally acclimated individuals,

thereby causing a shift in the genotypic distribution. Particular deviations from Harry Weinberg may result from this, particularly in cases where particular genotypes possess a substantial fitness advantage or disadvantage. Nevertheless, the possibility remains that additional factors, such as genetic drift, migration, gene interactions, natural selection, and so forth, could exert an influence. These results suggest that regions with lower genetic differentiation among populations exhibit higher genetic variation [93]. Furthermore, comparing genetic diversity among populations also emphasizes the importance of genetic conservation efforts for Z. loczyi. An interesting result is that the QD group has the highest Nei's diversity index, while the GTD group has the lowest. Genetic diversity is an important indicator of a population's ability to adapt to changing environments and potential threats [94]. A higher Nei's diversity index in the QD group implies that this population may possess a wider range of genetic variation, which could potentially provide them with a greater capacity to respond to selective pressures or environmental changes. On the other hand, the lower Nei's diversity index observed in the GTD group indicates that this population has less genetic variation [28]. This could imply a reduced ability to adapt to environmental challenges due to this limited gene pool [95,96].

There is substantial evidence to suggest that the four genetic categories closely align with regions of geographical distribution. The population structure is in accordance with the species' arid evolution [21]. The overlap between the population structure of K = 4and geographic partitioning supports a genetic–geographic boundary correspondence [97]. This implies that Z. loczyi has evolved to differentiate advantageously due to variations in the natural environment and geography across the four sampling regions. In particular, probable gene flow between the BJD and QD populations was observed. This hypothetical scenario posits that although the four primary deserts exhibit conspicuous distinctions, there remains potential for genetic material exchange and interconnection among specific desert populations. Principal component analyses and our phylogenetic tree indicate that there may be some gene flow between BJD and QD [32]. Reduced genetic associations between the analyzed SNPs were indicated by the lower LD in the GTD and BJD populations, which suggested the possibility of recombination events and increased genetic diversity in these populations [98]. On the other hand, the larger LD values observed in the TKD and QD population groups suggest a more robust genetic association and increased correlation among the analyzed SNPs. This suggests that certain genomic regions may be undergoing selection or genetic linkage [99,100]. Nevertheless, gene migration represents merely one among several possible explanations [101]. Incomplete germline classification, convergent evolution, the structure of ancestral populations, and additional variables may also account for our results [102–105].

Over two million square kilometers in northern China are classified as sandy and/or desolate terrain [106]. Variations in the distribution of plant species among the four primary deserts are discernible within the ancient genus Zygophyllum [46]. Recurrent climatic fluctuations throughout the Quaternary Ice Age may have prompted plant species to seek sanctuary in regions more conducive to survival during cooler periods [107,108]. After the Ice Age, certain plant species migrated and disseminated from their refuges to other regions [1,109]. TKD in the Tarim Basin began to appear during the mid-Pleistocene (0.78-0.13 Mya) as a product of the fourth uplift of the Tibetan Plateau (3.5–1.6 Mya) [110]. By the Holocene the desert was in a phase of major expansion [34]. Therefore, during the late Pleistocene (0.13–0.01 Mya) LGM period, many large lakes and marshes existed in the TKD [111]. Furthermore, we hypothesize that Z. loczyi may have sought refuge in the Tarim Basin. As a result of subsequent environmental degradation in TKD, Z. loczyi populations gradually migrated northward and expanded into GTD [112]. Hexi Corridor wind-sand landforms emerged during the transition from the Late Pleistocene to the Holocene [111]. The subsequent developments might have played a role in the dispersal and migration of Z. loczyi populations to BJD and QD. At this time, the BJD region was not blanketed by glaciers. QD underwent an upward trend throughout the Tertiary Himalayan orogeny [113]. The onset of arid tropical vegetation is composed primarily of plant species indigenous to the southern

littoral of the Paleo-Mediterranean area [114]. Current distribution patterns may be the result of the events described above, with the QDs retaining the greatest genetic diversity.

In summary, our research provides significant contributions to the understanding of the ecological differentiation and population genetics of *Z. loczyi* populations in China. Some of these results are applicable to conservation initiatives on a practical level, and they lay the groundwork for further investigations in fields including functional genomics, ecological genetics, and population modeling. Pursuing these directions will deepen our understanding of *Zygophyllum* and inform its conservation and sustainable management. Further studies could use SSR, cpDNA, and ITS to explore historical changes in local *Z. loczyi* populations. A deeper comprehension of the origin and evolution of desert ecosystems will result from this sequence of research efforts, which will also aid in the validation of the theory attributing to desert origins.

#### 5. Conclusions

In conclusion, the resequencing of the entire genome of *Z. loczyi* at the chromosome level is presented. Population studies based on whole-genome resequencing identified three distinct genetic lineages dispersed throughout the TKD, GTD, BDJ, and QD, indicating the adaptive evolution of the species. Additionally, gene flow may occur within QD and, respectively, between the populations of TKD and BJD. Phylogenetic tree and PCA analyses indicate that the four major deserts are clearly divided, with possible causes including climate fluctuations promoted by the uplift movement of the Tibetan Plateau. The segregation of formerly dispersed desert origins of divergence is supported by our data; therefore, we hypothesize that *Z. loczyi* populations spread from one branch of the TKD to the GTD and the other branch from the TKD to the QD, which then spreads to the BJD. Understanding the implications of this paper's discovery is crucial for the preservation of other drought-tolerant desert vegetation in Northwest China and the surrounding region.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/genes14122152/s1, Table S1: Statistical information for each sample; *Zygophyllum loczyi* map depth; *Zygophyllum loczyi* raw.

**Author Contributions:** Conceptualization, W.S. and M.W.; methodology, W.S. and S.W.; software, J.L. and Q.M.; validation, X.W., J.W., and M.W.; formal analysis, X.W.; investigation, J.W. and J.L.; resources, X.W. and H.L.; data curation, M.W.; writing—original draft preparation, M.W. and J.L.; writing—review and editing, W.S.; visualization, J.L.; supervision, X.W.; project administration, W.S.; funding acquisition, W.S. and J.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (grant No. 32170386), and the West Light Foundation of The Chinese Academy of Sciences (grant No. 2021-XBQNXZ-010).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** First, we are grateful to Wang Suo-Ming's team at Lanzhou University for providing the genomic reference data. Then we appreciate Daniel Petticord from Cornell University for his help in English and grammar editing of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

# References

- Dlugosch, K.M.; Parker, I.M. Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Mol. Ecol.* 2008, 17, 431–449. [CrossRef] [PubMed]
- 2. Reed, D.H.; Frankham, R. Correlation between fitness and genetic diversity. *Conserv. Biol.* 2003, 17, 230–237. [CrossRef]

- Gentili, R.; Solari, A.; Diekmann, M.; Dupre, C.; Monti, G.S.; Armiraglio, S.; Assini, S.; Citterio, S. Genetic differentiation, local adaptation and phenotypic plasticity in fragmented populations of a rare forest herb. *PeerJ.* 2018, *6*, e4929. [CrossRef] [PubMed]
   Kawecki, T.J.; Ebert, D. Conceptual issues in local adaptation. *Ecol. Lett.* 2004, *7*, 1225–1241. [CrossRef]
- 5. Leinonen, T.; McCairns, R.J.S.; O' Hara, R.B.; Merila, J. Q(ST)-F-ST comparisons: Evolutionary and ecological insights from genomic heterogeneity. *Nat. Rev. Genet.* 2013, *14*, 179–190. [CrossRef] [PubMed]
- Mace, E.S.; Tai, S.S.; Gilding, E.K.; Li, Y.H.; Prentis, P.J.; Bian, L.L.; Campbell, B.C.; Hu, W.S.; Innes, D.J.; Han, X.L.; et al. Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nat. Commun.* 2013, 4, 2320. [CrossRef] [PubMed]
- William, J.W.T.; Yueqi, Z.; Junrey, C.A.; Aldrin, Y.C.; Jaco, D.Z.; Samantha, L.H.; Jacqueline, B. Innovative Advances in Plant Genotyping. *Plant Genotyping* 2023, 2638, 451–465. [CrossRef]
- 8. Sakhale, S.A.; Yadav, S.; Clark, L.V.; Lipka, A.E.; Kumar, A.; Sacks, E.J. Genome-wide association analysis for emergence of deeply sown rice (*Oryza sativa*) reveals novel aus-specific phytohormone candidate genes for adaptation to dry-direct seeding in the field. *Front. Plant Sci.* **2023**, *14*, 1172816. [CrossRef]
- 9. Gupta, A.; Rico-Medina, A.; Cano-Delgado, A.I. The physiology of plant responses to drought. *Science* 2020, *368*, 266–269. [CrossRef]
- 10. Qian, C.J.; Yan, X.; Fang, T.Z.; Yin, X.Y.; Zhou, S.S.; Fan, X.K.; Chang, Y.X.; Ma, X.F. Genomic Adaptive Evolution of Sand Rice (*Agriophyllum squarrosum*) and Its Implications for Desert Ecosystem Restoration. *Front. Genet.* **2021**, *12*, 656061. [CrossRef]
- 11. McCormack, J.E.; Hird, S.M.; Zellmer, A.J.; Carstens, B.C.; Brumfield, R.T. Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol. Phylogenet. Evol.* **2013**, *66*, 526–538. [CrossRef]
- 12. Savi, T.; Bertuzzi, S.; Branca, S.; Tretiach, M.; Nardini, A. Drought-induced xylem cavitation and hydraulic deterioration: Risk factors for urban trees under climate change? *New Phytol.* **2015**, 205, 1106–1116. [CrossRef] [PubMed]
- 13. Luhar, I.; Luhar, S.; Savva, P.; Theodosiou, A.; Petrou, M.F.; Nicolaides, D. Light Transmitting Concrete: A Review. *Buildings* **2021**, *11*, 480. [CrossRef]
- 14. Wang, J.; Wang, K.L.; Zhang, M.Y.; Zhang, C.H. Impacts of climate change and human activities on vegetation cover in hilly southern China. *Ecol. Eng.* **2015**, *81*, 451–461. [CrossRef]
- 15. Bally, J.; Nakasugi, K.; Jia, F.Z.; Jung, H.T.; Ho, S.Y.W.; Wong, M.; Paul, C.M.; Naim, F.; Wood, C.C.; Crowhurst, R.N.; et al. The extremophile *Nicotiana benthamiana* has traded viral defence for early vigour. *Nat. Plants* **2015**, *1*, 15165. [CrossRef] [PubMed]
- Varshney, R.K.; Shi, C.C.; Thudi, M.; Mariac, C.; Wallace, J.; Qi, P.; Zhang, H.; Zhao, Y.S.; Wang, X.Y.; Rathore, A.; et al. Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 2017, 35, 969–976. [CrossRef] [PubMed]
- 17. Yang, F.L.; Zhou, Z.Y.; Zhang, N.; Liu, N.; Ni, B. Stress field modeling of northwestern South China Sea since 5.3 Ma and its tectonic significance. *Acta Oceanol. Sin.* **2013**, *32*, 31–39. [CrossRef]
- 18. Song, M.C.; Yi, P.H.; Xu, J.X.; Cui, S.X.; Shen, K.; Jiang, H.L.; Yuan, W.H.; Wang, H.J. A step metallogenetic model for gold deposits in the northwestern Shandong Peninsula, China. *Sci. China Earth Sci.* **2012**, *55*, 940–948. [CrossRef]
- 19. Xu, H.J. Variations of Vegetation and Its Influence Factors in the Arid Region of the Central Asia from 2000 to 2012. Master's Thesis, Lanzhou University, Lanzhou, China, 2014; p. 5.
- Zhang, M.L.; Fritsch, P.W. Evolutionary response of Caragana (Fabaceae) to Qinghai-Tibetan Plateau uplift and Asian interior aridification. *Plant Syst. Evol.* 2010, 288, 191–199. [CrossRef]
- 21. Liu, Y.X. A study on origin and formation of the Chinese desert floras. *Acta Phytotaxon. Sin.* **1995**, *2*, 131–143, (In Chinese with English Abstract).
- 22. Ragab, R.; Prudhomme, C. Climate change and water resources management in arid and semi-arid regions: Prospective and challenges for the 21st century. *Biosyst. Eng.* 2002, *81*, 3–34. [CrossRef]
- 23. Ewing, S.A.; Sutter, B.; Owen, J.; Nishiizumi, K.; Sharp, W.; Cliff, S.S.; Perry, K.; Dietrich, W.; McKay, C.P.; Amundson, R. A threshold in soil formation at Earth's arid-hyperarid transition. *Geochim. Cosmochim. Acta* 2006, *70*, 5293–5322. [CrossRef]
- Tapponnier, P.; Xu, Z.Q.; Roger, F.; Meyer, B.; Arnaud, N.; Wittlinger, G.; Yang, J.S. Geology—Oblique stepwise rise and growth of the Tibet plateau. *Science* 2001, 294, 1671–1677. [CrossRef] [PubMed]
- 25. An, Z.S.; Kutzbach, J.E.; Prell, W.L.; Porter, S.C. Evolution of Asian monsoons and phased uplift of the Himalayan Tibetan plateau since Late Miocene times. *Nature* 2001, 411, 62–66. [CrossRef]
- Huang, R.; Zhang, Y.; Shi, X.; Sun, Y.W. Middleand Late Permian Floras from the Eastern Junggar Basin, Xinjiang and Their Geological Implications. J. Jilin Univ. 2023, 53, 403–417.
- 27. Kong, X.J.; An, S.Z.; Liu, H.M. Flora of Desert Seed Plants in the Northern and Southern Margin of Junggar Basin. *Xinjiang Agric. Sci.* **2019**, *56*, 457–464, (In Chinese with English Abstract).
- 28. Ma, Y.; Liu, R.; Li, Z.; Jin, J.; Zou, X.; Tan, D.; Tao, T. Holocene environmental evolution recorded by sedimentation on the southern edge of the Gurbantunggut Deser. *Arid. Land. Geogr.* **2023**, 1–20, (In Chinese with English Abstract).
- 29. Chen, H.Z.; Jin, J.; Dong, G.R. Holocene Evolution Processes of Gurbantunggut Desert and Climatic Changes. *J. Desert Res.* 2001, 21, 333–339, (In Chinese with English Abstract).
- Chang, J.; Qiu, N.S.; Xu, W. Thermal regime of the Tarim Basin, Northwest China: A review. Int. Geol. Rev. 2017, 59, 45–61. [CrossRef]

- Yin, A.; Dang, Y.Q.; Zhang, M.; Chen, X.H.; McRivette, M.W. Cenozoic tectonic evolution of the Qaidam basin and its surrounding regions (Part 3): Structural geology, sedimentation, and regional tectonic reconstruction. *Geol. Soc. Am. Bull.* 2008, 120, 847–876. [CrossRef]
- 32. Cheng, F.; Jolivet, M.; Guo, Z.J.; Wang, L.; Zhang, C.H.; Li, X.Z. Cenozoic evolution of the Qaidam basin and implications for the growth of the northern Tibetan plateau: A review. *Earth Sci. Rev.* **2021**, *220*, 103730. [CrossRef]
- Kurschner, H. Phytosociological studies in the Alashan Gobi—A contribution to the flora and vegetation of Inner Mongolia (NW China). *Phytocoenologia* 2004, 34, 169–224. [CrossRef]
- Ren, J.; Tao, L. Quantitative Studies on Floristic Similarity of Rare and Endangered Desert Plants in China. Arid. Zone Resour. Environ. 2002, 16, 103–107, (In Chinese with English Abstract).
- 35. Zhang, Y.Y.; Zhang, W.W.; Manzoor, M.A.; Sabir, I.A.; Zhang, P.F.; Cao, Y.P.; Song, C. Differential involvement of WRKY genes in abiotic stress tolerance of Dendrobium huoshanense. *Ind. Crop. Prod.* **2023**, 204, 117295. [CrossRef]
- Cai, L.M.; Xi, Z.X.; Amorim, A.M.; Sugumaran, M.; Rest, J.S.; Liu, L.; Davis, C.C. Widespread ancient whole-genome duplications in Malpighiales coincide with Eocene global climatic upheaval. *New Phytol.* 2019, 221, 565–576. [CrossRef]
- 37. Van de Peer, Y.; Ashman, T.L.; Soltis, P.S.; Soltis, D.E. Polyploidy: An evolutionary and ecological force in stressful times. *Plant Cell* **2021**, *33*, 11–26. [CrossRef]
- 38. Wu, Z.R.; Yu, D.J.; Lin, R. Flora of China; Science Press: Beijing, China, 2004; pp. 1–622.
- Lv, D.K.; Shi, J.; Ba, Y.S.; Zhao, Y. Biomass and reproductive allocation characteristics of *Zygophyllum* L. population in Ili River Valley area. *Arid. Land. Geogr.* 2013, 36, 475–481, (In Chinese with English Abstract).
- Crookston, R.K.; Moss, D.N. C-4 and C-3 carboxylation characteristics in genus Zygophyllum (zygophyllaceae). Ann. Mo. Bot. Gard. 1972, 59, 465–470. [CrossRef]
- 41. Beier, B.A.; Chase, M.W.; Thulin, M. Phylogenetic relationships and taxonomy of subfamily Zygophylloideae (Zygophyllaceae) based on molecular and morphological data. *Plant Syst. Evol.* **2003**, 240, 11–39. [CrossRef]
- Khalik, K.N.A. A numerical taxonomic study of the family Zygophyllaceae from Egypt. Acta Bot. Bras. 2012, 26, 165–180. [CrossRef]
- 43. Zeng, Y.J.; Wang, Y.R.; Zhuang, G.H.; Yang, Z.S. Seed germination responses of *Reaumuria soongorica* and *Zygophyllum xanthoxylum* to drought stress and sowing depth. *Acta Ecol. Sin.* **2004**, *24*, 1629–1634, (In Chinese with English Abstract).
- 44. Zhang, L.; Wang, S.; Su, C.; Harris, A.J.; Zhao, L.; Su, N.; Wang, J.R.; Duan, L.; Chang, Z.Y. Comparative Chloroplast Genomics and Phylogenetic Analysis of *Zygophyllum* (Zygophyllaceae) of China. *Front. Plant Sci.* **2021**, *12*, 723622. [CrossRef] [PubMed]
- Yang, Y.C.; Jia, Y.; Zhao, Y.L.; Wang, Y.L.; Zhou, T. Comparative chloroplast genomics provides insights into the genealogical relationships of endangered *Tetraena mongolica* and the chloroplast genome evolution of related Zygophyllaceae species. *Front. Genet.* 2022, 13, 1026919. [CrossRef] [PubMed]
- Zhao, Y.Z.; Zhao, L.Q.; Rui, C. Flora Intramongolica. In *Typis Intramongolicae Popularis*, 3rd ed.; Inner Mongolia Peoples Publishing House: Huhhot, China, 2019; Volume 3, pp. 1–513.
- Bellstedt, D.U.; van Zyl, L.; Marais, E.M.; Bytebier, B.; de Villiers, C.A.; Makwarela, A.M.; Dreyer, L.L. Phylogenetic relationships, character evolution and biogeography of southern African members of *Zygophyllum* (Zygophyllaceae) based on three plastid regions. *Mol. Phylogenet. Evol.* 2008, 47, 932–949. [CrossRef] [PubMed]
- Zhang, C.B.; Shi, X.S.; Wan, T.; Cao, Y.W.; Zhang, X.M. Studies on the Genetic Relationships of 7 Species of *Zygophyllum* in Inner Mongolia Based on Random Amplified Polymorphic DNA (RAPD). *Chin. J. Grassl.* 2006, 28, 86–90, (In Chinese with English Abstract).
- Wan, T.; Yan, L.; Shi, X.S.; Yi, W.D.; Zhang, X.M. Comparative Analysis of Genetic Diversity of *Zygophyllum*, *L*. and its Related Congener *Sarcozygium xanthoxylon* Bunge in Inner Mongolia. *J. Arid. Land Resour. Environ.* 2006, 20, 199–203, (In Chinese with English Abstract).
- 50. Yang, S.M.; Furukawa, L. Anatomical adaptations of three species of Chinese xerophytes (Zygophyllaceae). *J. For. Res.* 2006, 17, 247–251, (In Chinese with English Abstract). [CrossRef]
- 51. Wan, T.; Shi, X.S.; Yi, W.D.; Zhang, X.M.; Zhang, C.B. Pollen Morphologies of Seven Species of *Zygophyllum* Lin Alashan Desert. *Acta Bot. Boreali Occident. Sin.* **2006**, *26*, 1704–1708, (In Chinese with English Abstract).
- Sayed, O. Adaptational responses of *Zygophyllum* qatarense Hadidi to stress conditions in a desert environment. *J. Arid. Environ.* 1996, 32, 445–452. [CrossRef]
- Li, Y.; Qu, J.J.; An, L.Z. Germinating Physiological Conditions of *Zygophyllum xanthoxylon* Maxim. Seeds. *Plant Physiol. Newsl.* 2008, 44, 276–278, (In Chinese with English Abstract).
- Lu, N.N.; Cui, X.L.; Wang, J.H.; Zhao, B.B.; Xu, X.L.; Liu, K. Effect of Storage and Light Conditions on Seed Germination of 5 Desert Species in Zygophyllaceae. J. Desert Res. 2008, 28, 1130–1135, (In Chinese with English Abstract).
- 55. Lefevre, I.; Correal, E.; Lutts, S. Cadmium tolerance and accumulation in the noxious weed *Zygophyllum fabago*. *Can. J. Bot. Rev. Can. Bot.* 2005, *83*, 1655–1662. [CrossRef]
- 56. Allen, G.C.; Flores-Vergara, M.A.; Krasynanski, S.; Kumar, S.; Thompson, W.F. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat. Protoc.* **2006**, *1*, 2320–2325. [CrossRef] [PubMed]
- 57. Borst, P. Ethidium DNA agarose gel electrophoresis: How it started. *Iubmb Life* 2005, 57, 745–747. [CrossRef]
- Simbolo, M.; Gottardi, M.; Corbo, V.; Fassan, M.; Mafficini, A.; Malpeli, G.; Lawlor, R.T.; Scarpa, A. DNA Qualification Workflow for Next Generation Sequencing of Histopathological Samples. *PLoS ONE* 2013, *8*, e62692. [CrossRef] [PubMed]

- 59. Dolezel, J.; Lucretti, S.; Schubert, I. Plant chromosome analysis and sorting by flow cytometry. *Crit. Rev. Plant Sci.* **1994**, *13*, 275–309. [CrossRef]
- Gavrieli, Y.; Sherman, Y.; Ben-Sasson, S.A. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. J. Cell Biol. 1992, 119, 493–501. [CrossRef] [PubMed]
- Takagi, H.; Abe, A.; Yoshida, K.; Kosugi, S.; Natsume, S.; Mitsuoka, C.; Uemura, A.; Utsushi, H.; Tamiru, M.; Takuno, S. QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* 2013, 74, 174–183. [CrossRef]
- 62. Ward, M.K.; Meade, A.W. Dealing with Careless Responding in Survey Data: Prevention, Identification, and Recommended Best Practices. *Annu. Rev. Psychol.* 2023, 74, 577–596. [CrossRef]
- 63. Stanke, M.; Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **2003**, *19*, 215–225. [CrossRef]
- 64. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, 25, 1754–1760. [CrossRef] [PubMed]
- 65. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M.; et al. Twelve years of SAMtools and BCFtools. *Gigascience* **2021**, *10*, giab008. [CrossRef]
- Mehta, R.L.; Pascual, M.T.; Soroko, S.; Savage, B.R.; Himmelfarb, J.; Ikizler, T.A.; Paganini, E.P.; Chertow, G.M.; PICARD. Spectrum
  of acute renal failure in the intensive care unit: The PICARD experience. *Kidney Int.* 2004, 66, 1613–1621. [CrossRef] [PubMed]
- 67. Lohmann, M.; Anzanello, M.J.; Fogliatto, F.S.; da Silveira, G.C. Grouping workers with similar learning profiles in mass customization production lines. *Comput. Ind. Eng.* **2019**, *131*, 542–551. [CrossRef]
- Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.Y.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w (1118); iso-2; iso-3. *Fly* 2012, *6*, 80–92. [CrossRef] [PubMed]
- Gudbjartsson, D.; Helgason, H.; Gudjonsson, S.A.; Zink, F.; Oddson, A.; Gylfason, A.; Besenbacher, S.; Magnusson, G.; Halldorsson, B.V.; Hjartarson, E.; et al. Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* 2015, 47, 435–444. [CrossRef] [PubMed]
- McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010, 20, 1297–1303. [CrossRef]
- 71. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef]
- 72. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An information aesthetic for comparative genomics. *Genome Res.* 2009, 19, 1639–1645. [CrossRef]
- 73. Deng, Y.Y.; Li, J.Q.; Wu, S.F.; Zhu, Y.P.; Chen, Y.W.; He, F.C. Integrated nr database in protein annotation system and its localization. *Comput. Eng. Ital.* **2006**, *32*, 71–74.
- Zaru, R.; Orchard, S.; UniProt, C. UniProt Tools: BLAST, Align, Peptide Search, and ID Mapping. Curr. Protoc. 2023, 3, e697. [CrossRef] [PubMed]
- Xavier, B.B.; Das, A.J.; Cochrane, G.; De Ganck, S.; Kumar-Singh, S.; Aarestrup, F.M.; Goossens, H.; Malhotra-Kumar, S. Consolidating and Exploring Antibiotic Resistance Gene Data Resources. *J. Clin. Microbiol.* 2016, 54, 851–859. [CrossRef] [PubMed]
- Tatusov, R.L.; Galperin, M.Y.; Natale, D.A.; Koonin, E.V. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 2000, 28, 33–36. [CrossRef] [PubMed]
- 77. Koonin, E.V.; Fedorova, N.D.; Jackson, J.D.; Jacobs, A.R.; Krylov, D.M.; Makarova, K.S.; Mazumder, R.; Mekhedov, S.L.; Nikolskaya, A.N.; Rao, B.S.; et al. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 2004, *5*, R7. [CrossRef] [PubMed]
- Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 2018, 35, 1547–1549. [CrossRef] [PubMed]
- Yoshihara, K.; Shahmoradgoli, M.; Martinez, E.; Vegesna, R.; Kim, H.; Torres-Garcia, W.; Trevino, V.; Shen, H.; Laird, P.W.; Levine, D.A. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 2013, *4*, 2612. [CrossRef]
- 80. Herrando, P.S.; Tobler, R.; Huber, C.D. smartsnp, an r package for fast multivariate analyses of big genomic data. *Methods Ecol. Evol.* **2021**, *12*, 2084–2093. [CrossRef]
- Yang, J.A.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A Tool for Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.* 2011, *88*, 76–82. [CrossRef]
- Zhang, C.; Dong, S.S.; Xu, J.Y.; He, W.M.; Yang, T.L. PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 2019, 35, 1786–1788. [CrossRef]
- Dolezel, J.; Kubalakova, M.; Paux, E.; Bartos, J.; Feuillet, C. Chromosome-based genomics in the cereals. *Chromosome Res.* 2007, 15, 51–66. [CrossRef]

- Safar, J.; Bartos, J.; Janda, J.; Bellec, A.; Kubalakova, M.; Valarik, M.; Pateyron, S.; Weiserova, J.; Tuskova, R.; Cihalikova, J. Dissecting large and complex genomes: Flow sorting and BAC cloning of individual chromosomes from bread wheat. *Plant J.* 2004, *39*, 960–968. [CrossRef] [PubMed]
- 85. Schiessl, S.; Huettel, B.; Kuehn, D.; Reinhardt, R.; Snowdon, R. Post-polyploidisation morphotype diversification associates with gene copy number variation. *Sci. Rep.* 2017, *7*, 41845. [CrossRef] [PubMed]
- Pellino, M.; Hojsgaard, D.; Schmutzer, T.; Scholz, U.; Horandl, E.; Vogel, H.; Sharbel, T.F. Asexual genome evolution in the apomictic Ranunculus auricomus complex: Examining the effects of hybridization and mutation accumulation. *Mol. Ecol.* 2013, 22, 5908–5921. [CrossRef] [PubMed]
- Tiwari, V.K.; Wang, S.C.; Danilova, T.; Koo, D.; Vrana, J.; Kubalakova, M.; Hribova, E.; Rawat, N.; Kalia, B.; Singh, N.; et al. Exploring the tertiary gene pool of bread wheat: Sequence assembly and analysis of chromosome 5M(g) of Aegilops geniculata. *Plant J.* 2015, *84*, 733–746. [CrossRef] [PubMed]
- Boutte, J.; Maillet, L.; Chaussepied, T.; Letort, S.; Aury, J.M.; Belser, C.; Boideau, F.; Brunet, A.; Coriton, O.; Deniot, G.; et al. Genome Size Variation and Comparative Genomics Reveal Intraspecific Diversity in Brassica rapa. *Front. Plant Sci.* 2020, 11, 577536. [CrossRef]
- Singh, R.; Ming, R.; Yu, Q.Y. Comparative Analysis of GC Content Variations in Plant Genomes. *Trop. Plant Biol.* 2016, *9*, 136–149.
   [CrossRef]
- Song, X.H.; Yang, T.B.; Yan, X.H.; Zheng, F.K.; Xu, X.Q.; Zhou, C.Q. Comparison of microsatellite distribution patterns in twenty-nine beetle genomes. *Gene* 2020, 757, 144919. [CrossRef]
- Browne, P.D.; Nielsen, T.K.; Kot, W.; Aggerholm, A.; Gilbert, M.T.P.; Puetz, L.; Rasmussen, M.; Zervas, A.; Hansen, L.H. GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. *Gigascience* 2020, *9*, giaa008. [CrossRef]
- 92. Jarosz, D.F.; Lindquist, S. Lindquist, Hsp90 and Environmental Stress Transform the Adaptive Value of Natural Genetic Variation. *Science* 2010, 330, 1820–1824. [CrossRef]
- Chung, M.Y.; Son, S.; Herrando-Moraira, S.; Tang, C.Q.; Maki, M.; Kim, Y.D.; Lopez, P.J.; Hamrick, J.L.; Chung, M.G. Incorporating differences between genetic diversity of trees and herbaceous plants in conservation strategies. *Conserv. Biol.* 2020, 34, 1142–1151. [CrossRef]
- 94. Zhang, X.; Chen, G.; Ma, Y.P.; Ge, J.; Sun, W.B. Genetic diversity and population structure of *Buddleja crispa* Bentham in the Himalaya-Hengduan Mountains region revealed by AFLP. *Biochem. Syst. Ecol.* **2015**, *58*, 13–20. [CrossRef]
- 95. Ding, Y.L.; Shi, Y.T.; Yang, S.H. Advances and challenges in uncovering cold tolerance regulatory mechanisms in plants. *New Phytol.* **2019**, 222, 1690–1704. [CrossRef] [PubMed]
- 96. Davis, M.B.; Shaw, R.G. Range shifts and adaptive responses to Quaternary climate change. *Science* **2001**, 292, 673–679. [CrossRef] [PubMed]
- Colwell, R.K.; Lees, D.C. The mid-domain effect: Geometric constraints on the geography of species richness. *Trends Ecol. Evol.* 2000, 15, 70–76. [CrossRef] [PubMed]
- 98. Rundle, H.D.; Nosil, P. Ecological speciation. Ecol. Lett. 2005, 8, 336–352. [CrossRef]
- Rundle, H.D.; Nagel, L.; Boughman, J.W.; Schluter, D. Natural selection and parallel speciation in sympatric sticklebacks. *Science* 2000, 287, 306–308. [CrossRef]
- Flint-Garcia, S.A.; Thornsberry, J.M.; Buckler, E.S. Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 2003, 54, 357–374. [CrossRef]
- 101. Beye, A.; Billot, C.; Ronfort, J.; McNally, K.L.; Diouf, D.; Glaszmann, J.C. Traces of Introgression from cAus into Tropical Japonica Observed in African Upland Rice Varieties. *Rice* 2023, *16*, 12. [CrossRef]
- 102. Martin, J.; Ponstingl, H.; Lefranc, M.P.; Archer, J.; Sargan, D.; Bradley, A. Comprehensive annotation and evolutionary insights into the canine (*Canis lupus familiaris*) antigen receptor loci. *Immunogenetics* **2018**, *70*, 223–236. [CrossRef]
- Hu, Y.; Hu, Y.; Zhou, W.; Wei, F. Conservation Genomics and Metagenomics of Giant and Red Pandas in the Wild. *Annu. Rev. Anim. Biosci.* 2023, 12, 7.1–7.21. [CrossRef]
- 104. Poulicard, N.; Pagan, I.; Gonzalez-Jara, P.; Mora, M.A.; Hily, J.M.; Fraile, A.; Pinero, D.; Garcia-Arenal, F. Repeated loss of the ability of a wild pepper disease resistance gene to function at high temperatures suggests that thermoresistance is a costly trait. *New Phytol.* 2023. *early view*. [CrossRef] [PubMed]
- 105. Barrios-Leal, D.Y.; Menezes, R.S.T.; Zappi, D.; Manfrin, M.H. Unravelling the genetic diversity and population dynamics of three Tacinga species (*Cactaceae: Opuntioideae*) in the Caatinga. *Bot. J. Linnean Soc.* **2023**, 203, boad054. [CrossRef]
- 106. Wang, X.M.; Chen, F.; Hasi, E.; Li, J.C. Desertification in China: An assessment. Earth-Sci. Rev. 2008, 88, 188–206. [CrossRef]
- 107. Soltis, D.E.; Morris, A.B.; McLachlan, J.S.; Manos, P.S.; Soltis, P.S. Comparative phylogeography of unglaciated eastern North America. *Mol. Ecol.* 2006, *15*, 4261–4293, (In Chinese with English Abstract). [CrossRef] [PubMed]
- 108. Bashalkhanov, S.; Johnson, J.S.; Rajora, O.P. Postglacial phylogeography, admixture, and evolution of red spruce (*Picea rubens* Sarg.) in Eastern North America. *Front. Plant Sci.* **2023**, *14*, 1272362. [CrossRef] [PubMed]
- 109. Nie, Z.L.; Hodel, R.; Johnson, G.; Ren, C.; Meng, Y.; Ickert-Bond, S.M.; Liu, X.-Q.; Zimmer, E.; Wen, J. Climate-influenced boreotropical survival and rampant introgressions explain the thriving of New World grapes in the north temperate zone. *J. Integr. Plant Biol.* 2023, 65, 1183–1203. [CrossRef] [PubMed]

- 110. Wen, J.; Zhang, J.Q.; Nie, Z.L.; Zhong, Y.; Sun, H. Evolutionary diversificatons of plants on the Qinghai-Tibetan Plateau. *Front. Genet.* **2014**, *5*, 4. [CrossRef]
- Nottebaum, V.; Lehmkuhl, F.; Stauch, G.; Lu, H.Y.; Yi, S. Late Quaternary aeolian sand deposition sustained by fluvial reworking and sediment supply in the Hexi Corridor—An example from northern Chinese drylands. *Geomorphology* 2015, 250, 113–127. [CrossRef]
- Bush, A.B.G.; Little, E.C.; Rokosh, D.; White, D.; Rutter, N.W. Investigation of the spatio-temporal variability in Eurasian Late Quaternary loess-paleosol sequences using a coupled atmosphere-ocean general circulation model. *Quat. Sci. Rev.* 2004, 23, 481–498. [CrossRef]
- 113. Mehta, P.K. Tectonic significance of the young mineral dates and the rates of cooling and uplift in the Himalaya. *Tectonophysics* **1980**, *62*, 205–217. [CrossRef]
- 114. Pant, N.C.; Singh, P.; Jain, A.K. A Re-look at the Himalayan metamorphism. Episodes 2020, 43, 369–380. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.