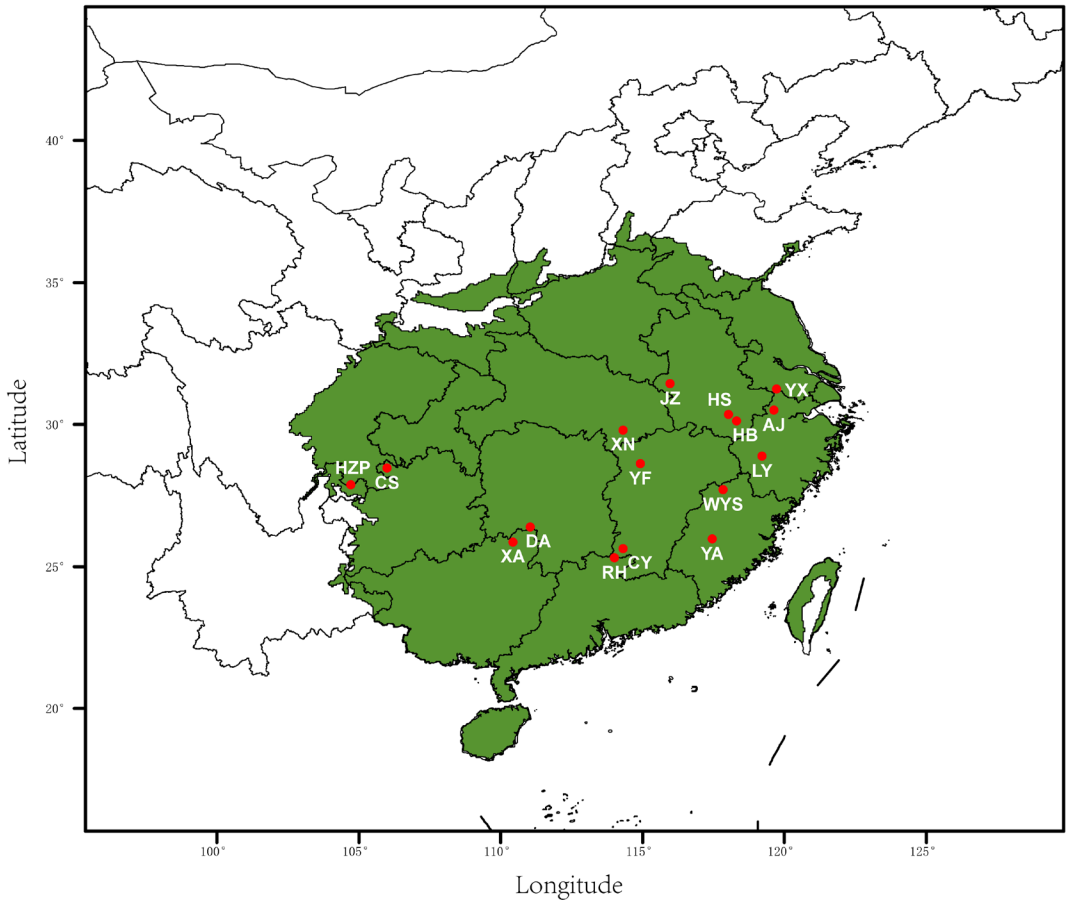
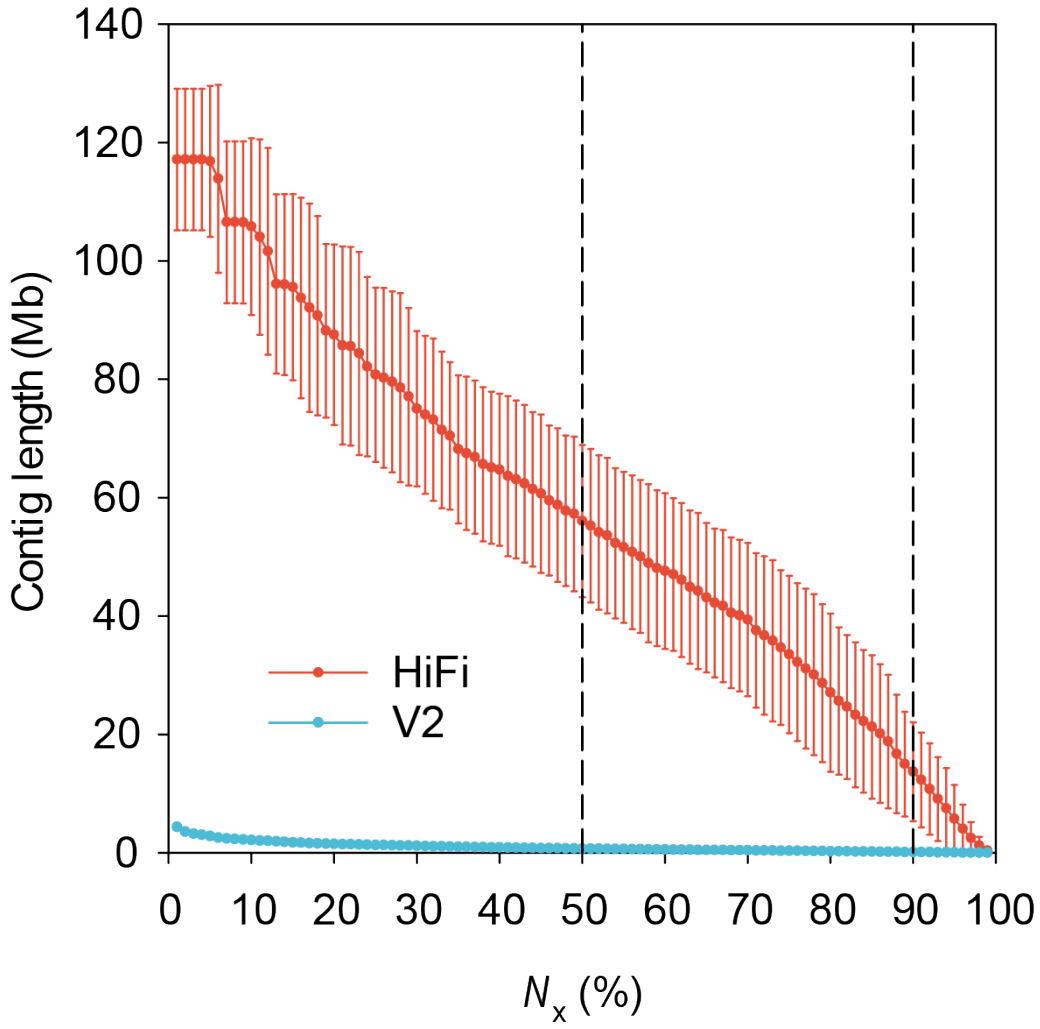
Haplotype-based pangenomes reveal genetic variations and climate adaptations in moso bamboo populations

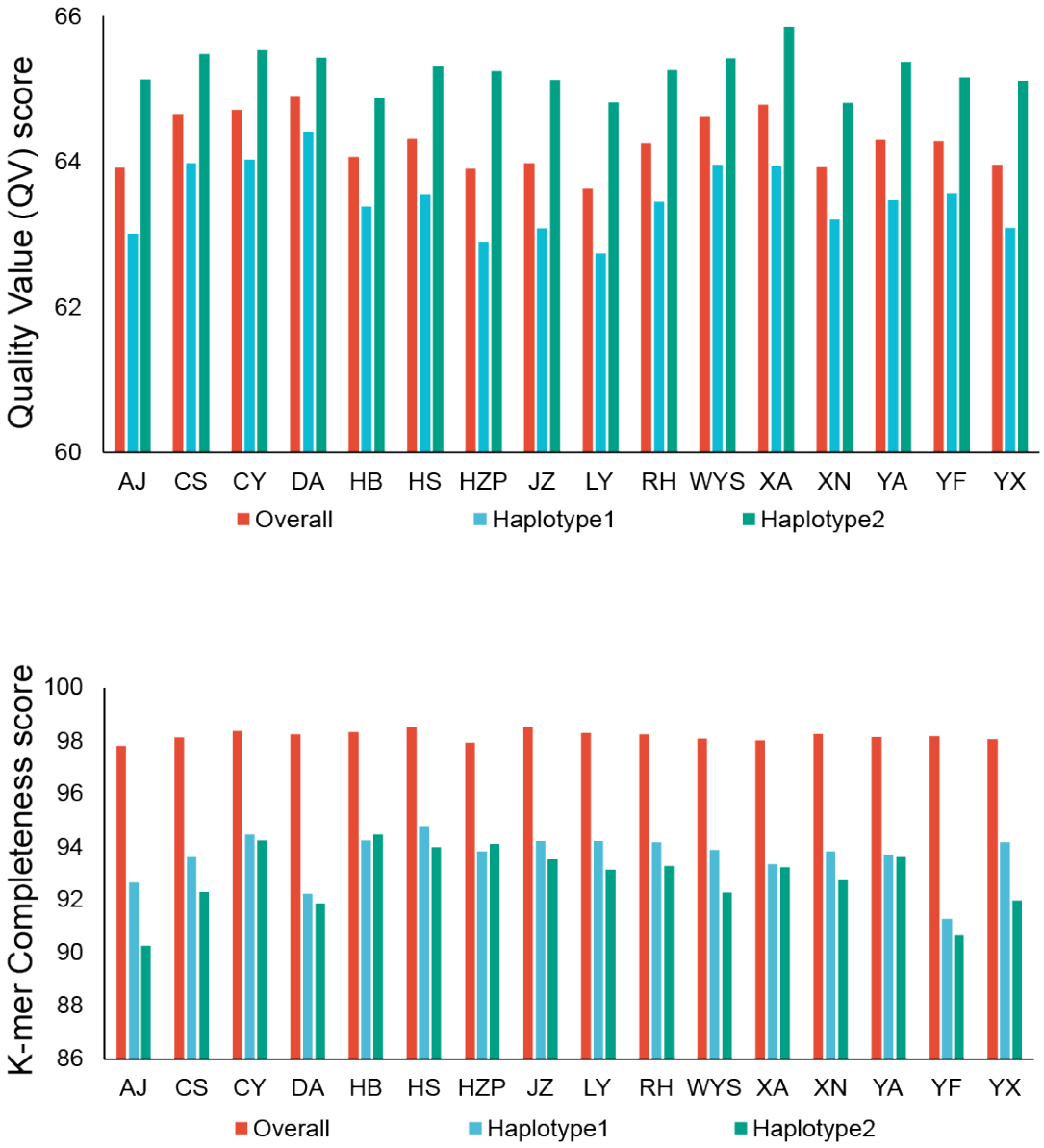
Hou *et al.*



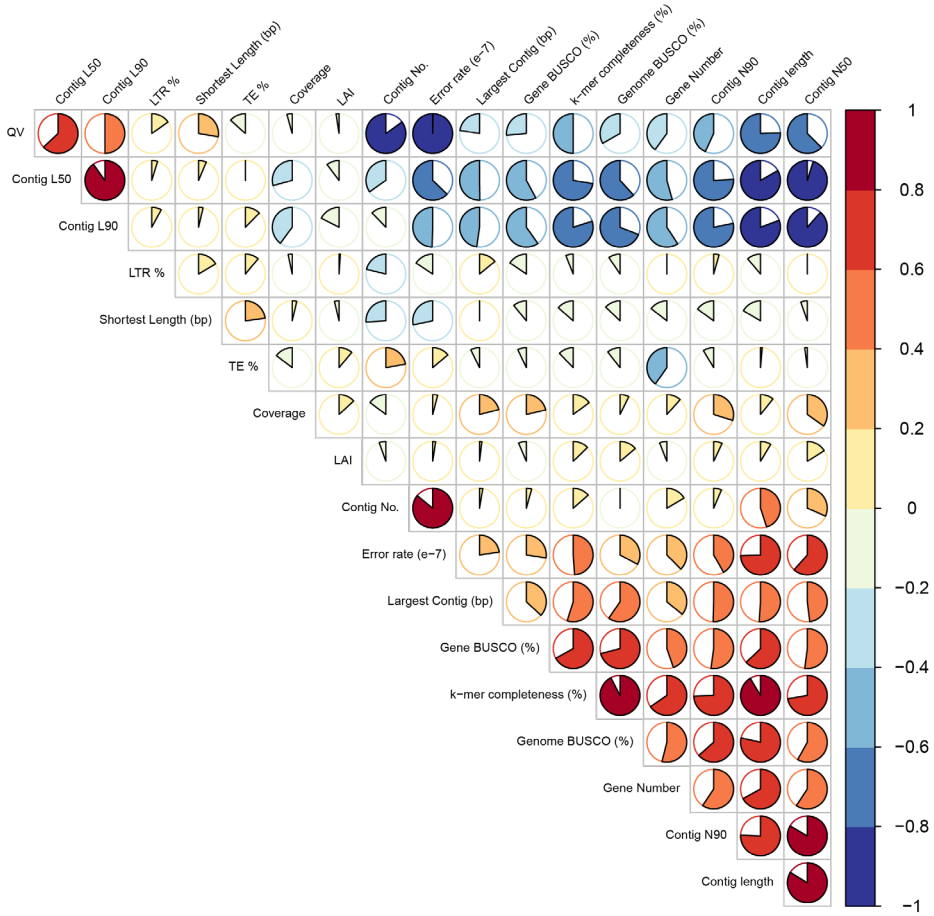
Supplementary Fig. 1 Geographic locations of the 16 representative moso bamboo accessions across China. The specific coordinates of each accession are denoted by red dots. The suitable habitat range of moso bamboo across China was delineated based on Global Biodiversity Information Facility (GBIF) data (green shaded area).



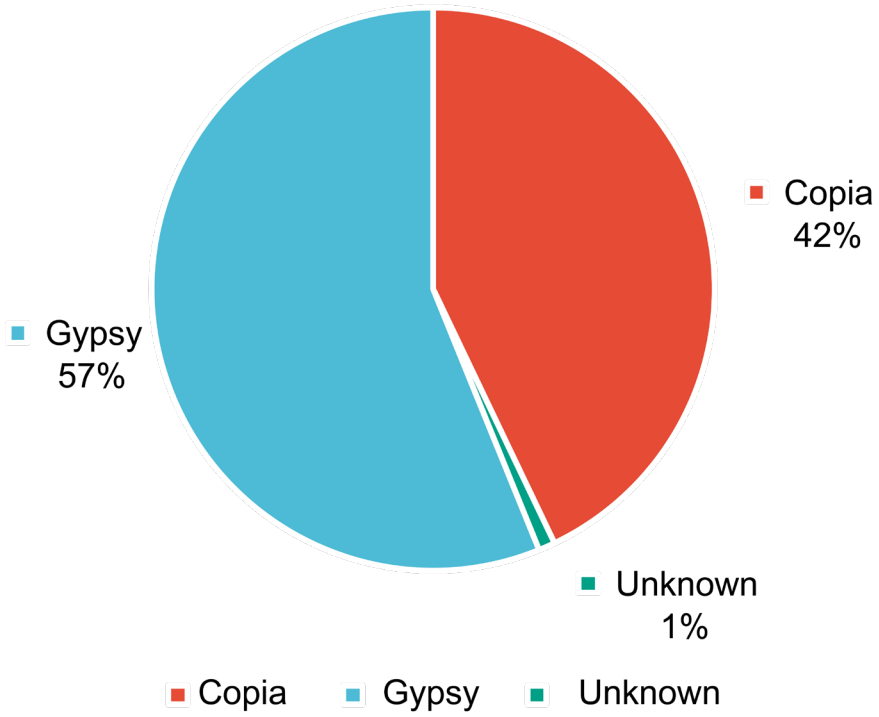
Supplementary Fig. 2 Contig Nx sizes of all genome assemblies. Contig N_x sizes are shown for the haplotype assemblies generated in this study (red) and the previously published assemblies (blue). The red line depicts the average contig N_x values across the genome assemblies of the 16 moso bamboo accessions sequenced with PacBio HiFi reads in this study. The blue line shows the contig N_x sizes for the earlier genome assemblies of a different set of moso bamboo accessions based on previous sequencing.



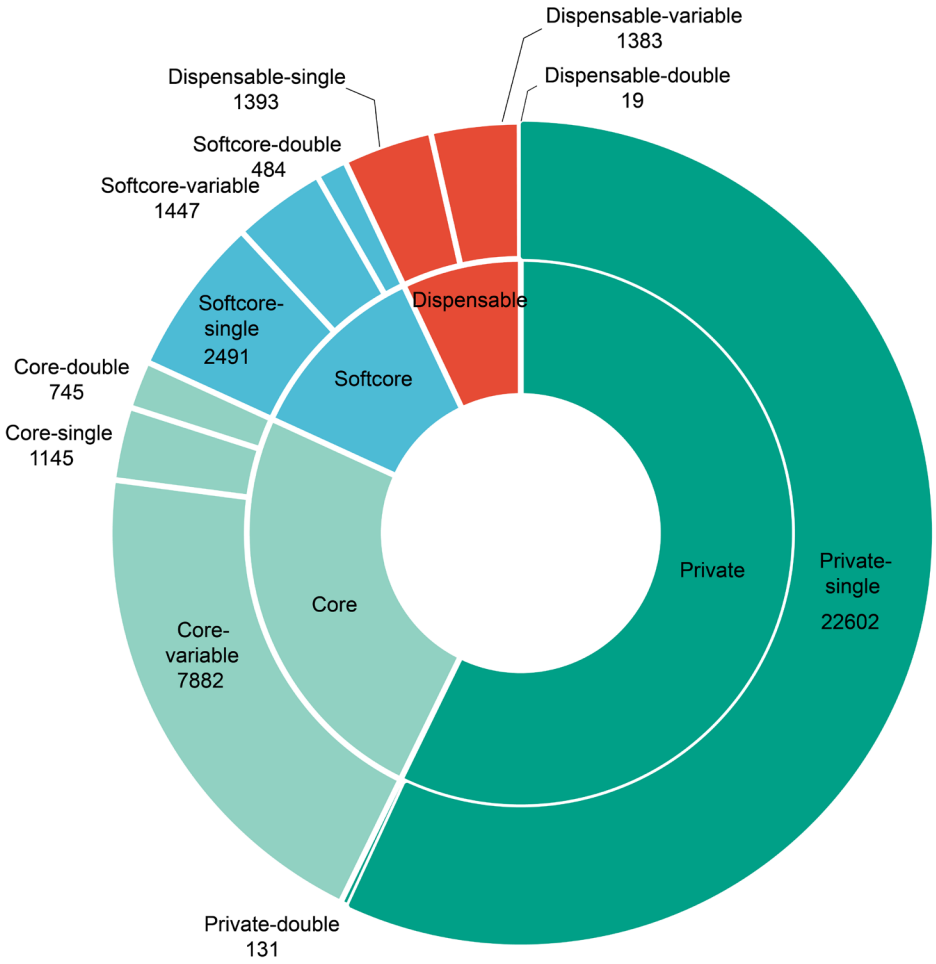
Supplementary Fig. 3 K-mer assessment of the genome. The top panel displays the average quality value (QV), and the bottom panel illustrates the k-mer completeness.



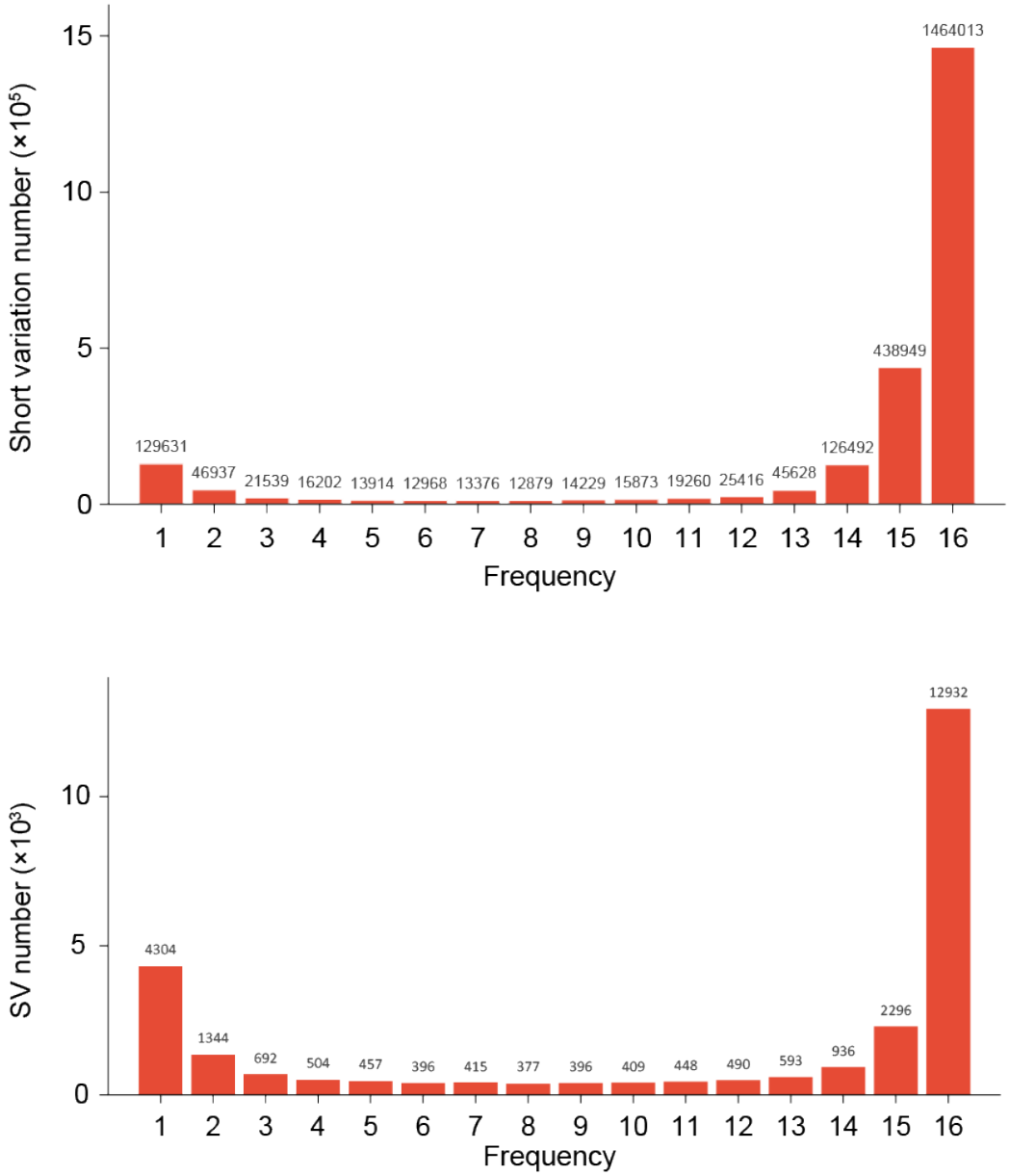
Supplementary Fig. 4 Correlations between assembly quality evaluation metrics. We assessed the quality of the genomes mainly through 3 aspects—continuity, base accuracy, and structural accuracy—and conducted correlation analyses on their parameters. Red represents a positive correlation, blue represents a negative correlation, and darker colors indicate stronger correlations. The circle area represents the parameter value, with larger areas indicating larger parameter values.



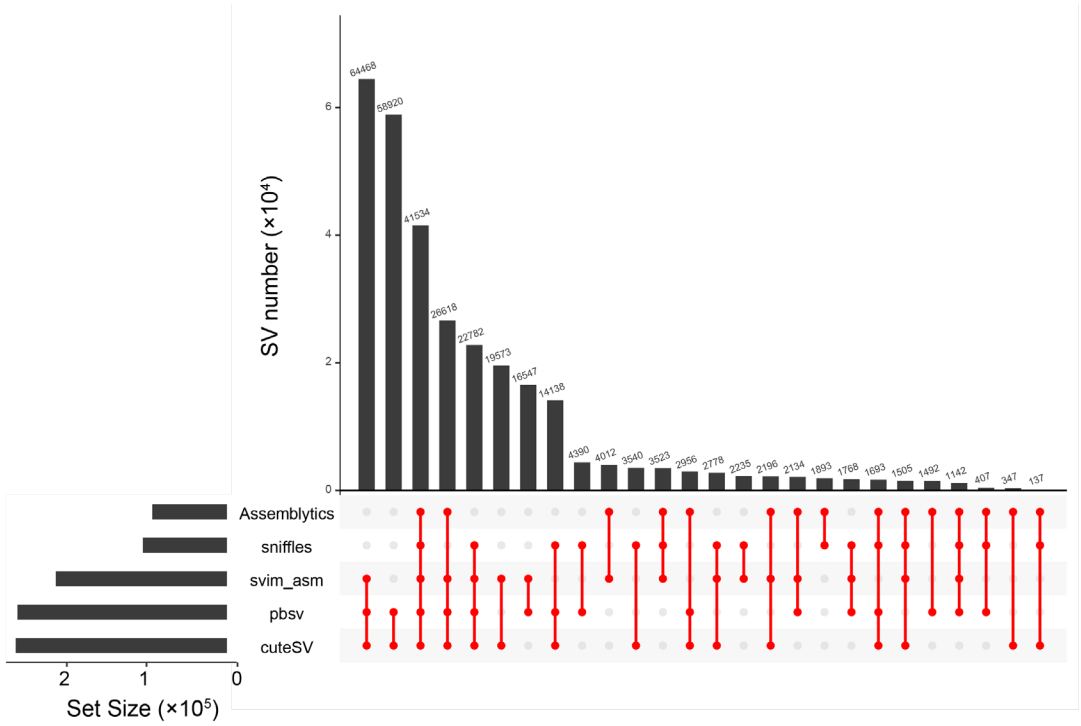
Supplementary Fig. 5 Pie chart illustrating the relative abundances of different LTR retrotransposon families in the moso bamboo genome.



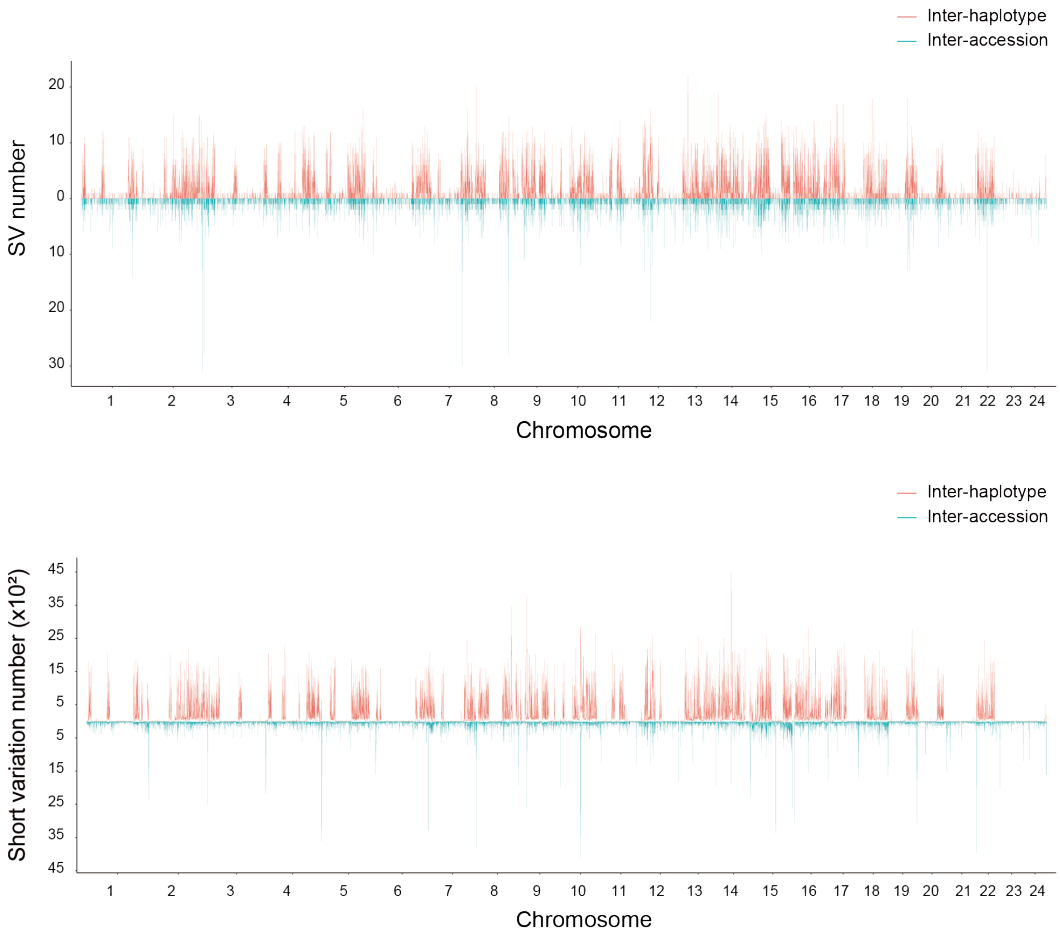
Supplementary Fig. 6 Pie chart comparing LTR retrotransposon types between the two haplotypes within each moso bamboo genome.



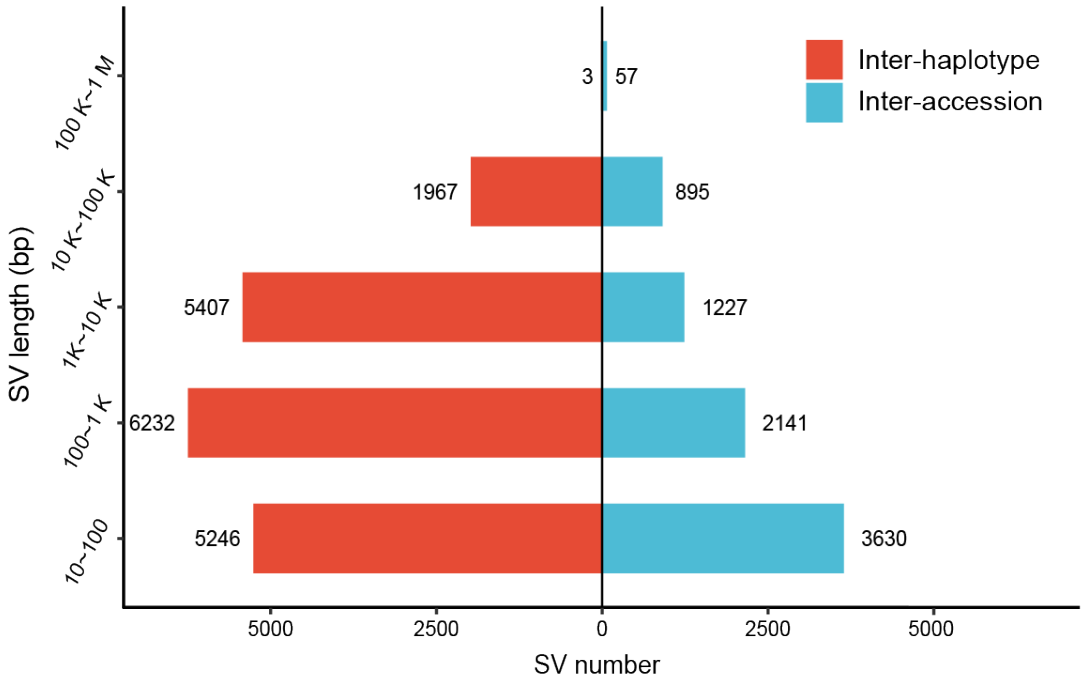
Supplementary Fig. 7 Accession frequency distribution of short variations and SVs. The x-axis represents the frequency of occurrence in different accessions. The upper plot shows the number of short variations, and the lower plot shows the number of SVs. In this study, short variations refer to SNPs and InDels.



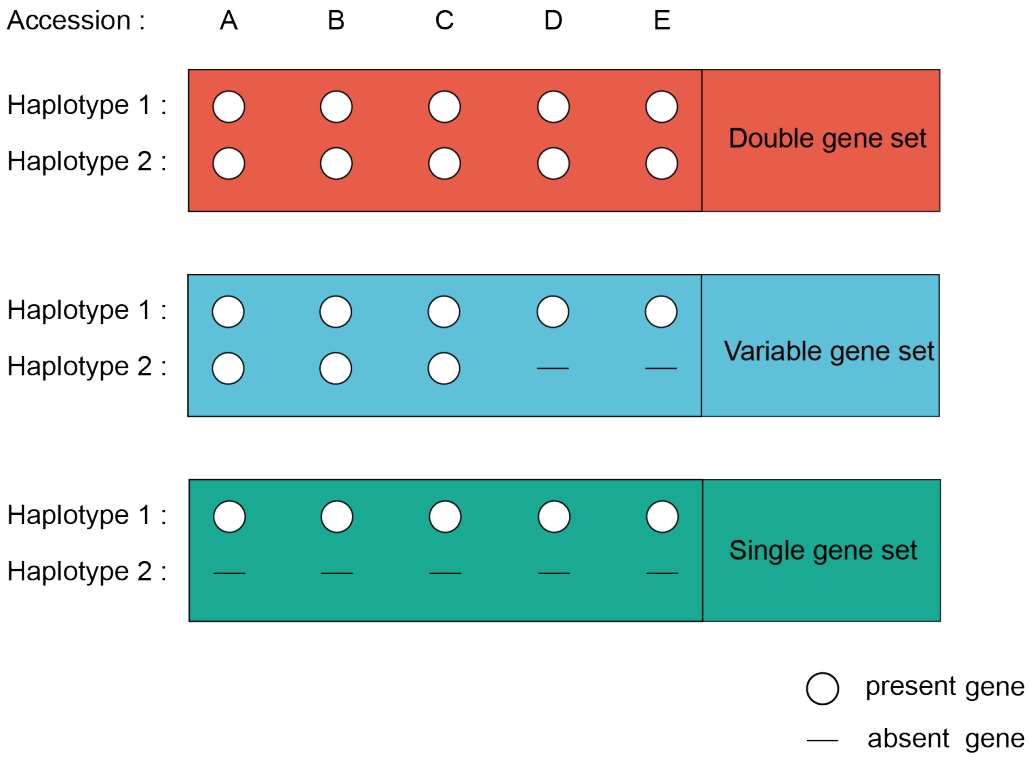
Supplementary Fig. 8 Upset plot showing the number of SVs detected by different methods. We used five SV callers (Assemblytics, sniffles, svim_asm, pbsv, and cuteSV) based on two strategies (read alignment and whole-genome alignment). The upper bar chart shows the intersections of callers, with height indicating size. The lower left bar chart shows caller sets, with length indicating size. The matrix shows intersections containing sets via marks at corresponding row-column points. This revealed overlapping relationships between sets.



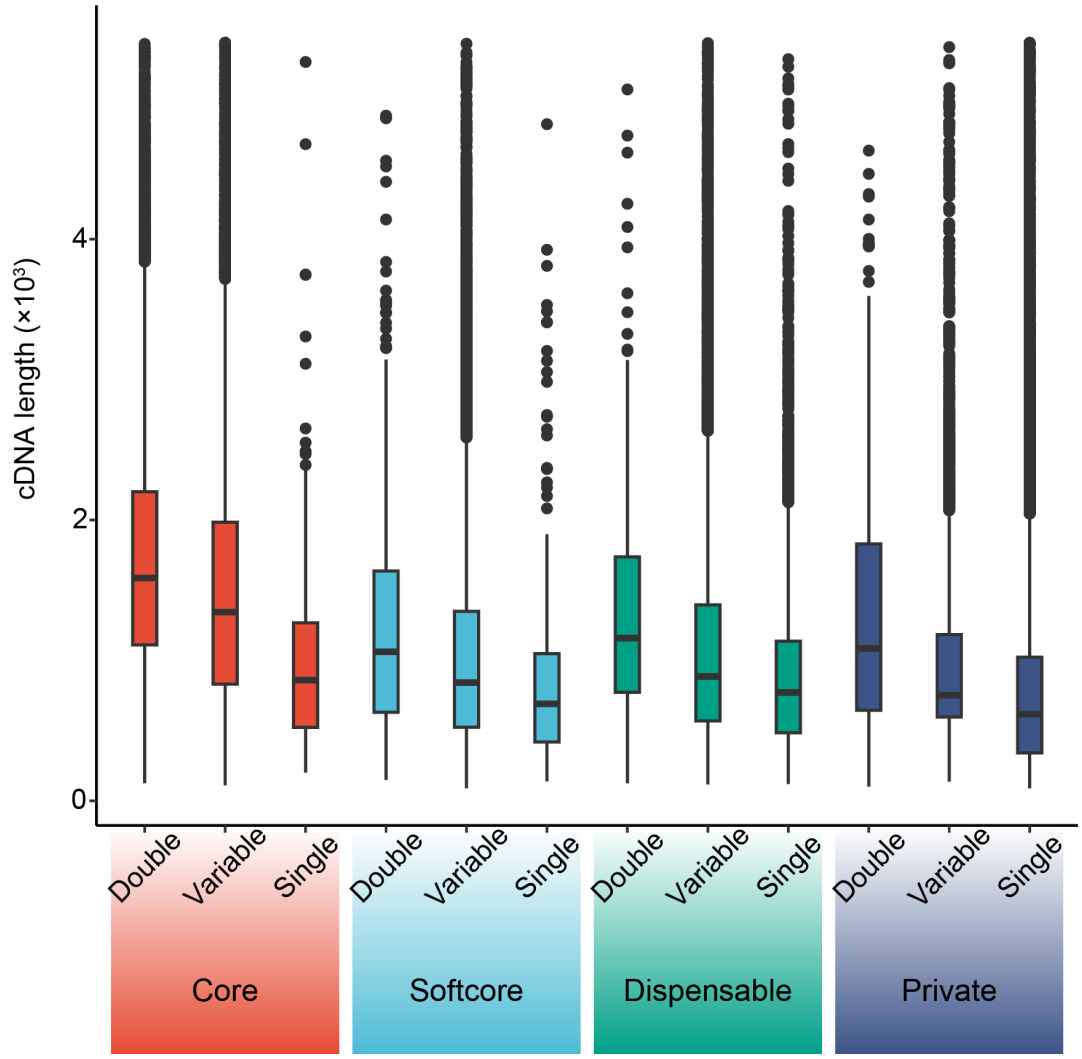
Supplementary Fig. 9 Frequency distributions of both structural variations (SVs) and short variations (SNPs and InDels) across the moso bamboo genome. The upper panel depicts the distribution of SVs, while the lower panel illustrates the distribution of short variations. The x-axis indicates the chromosome numbers, and the y-axis indicates the number of variations on each chromosome. The blue line represents the frequency of inter-accession variations, which are differences between individual moso bamboo accessions, while the red line represents the frequency of inter-haplotype variations, which are differences between the two haplotypes within each accession.



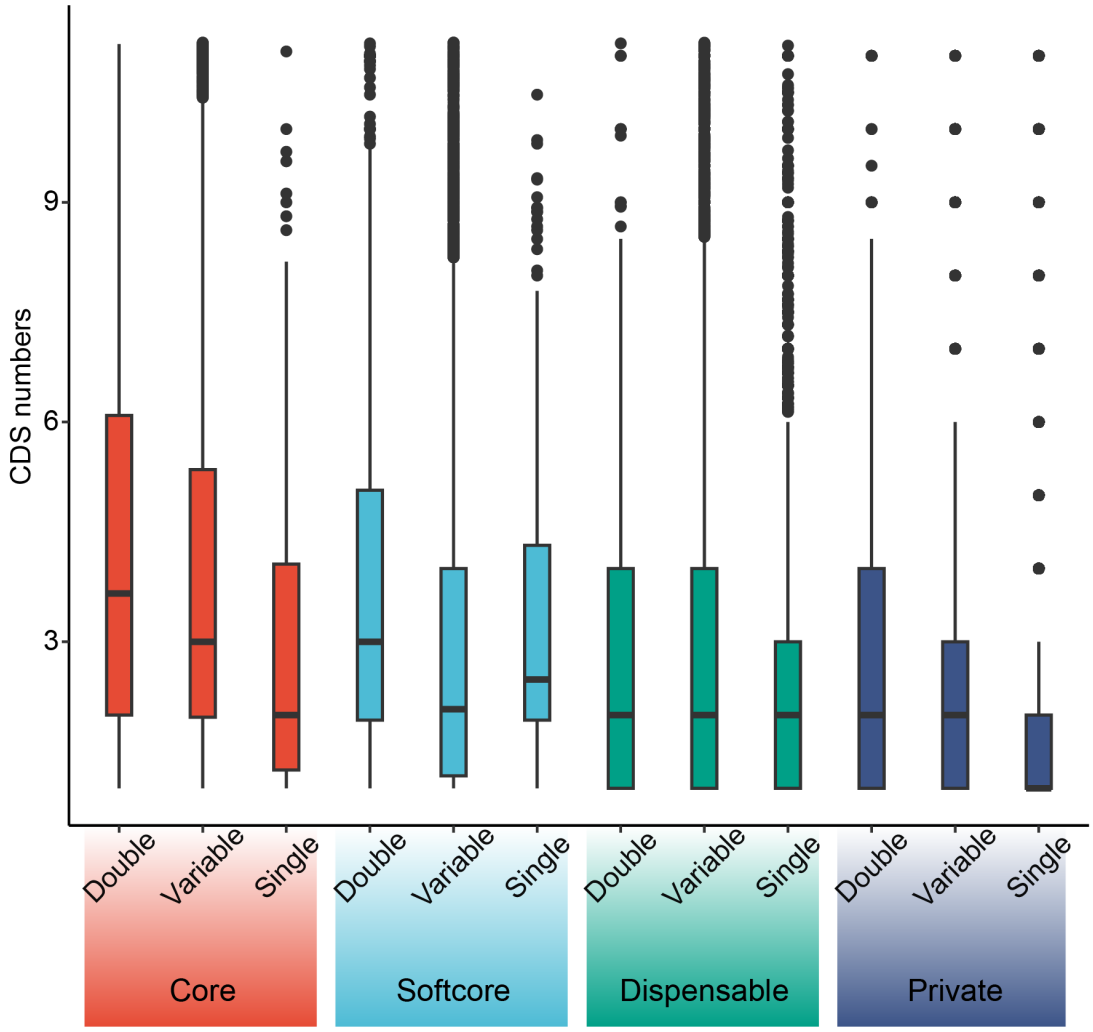
Supplementary Fig. 10 Numbers of SVs in different length categories. Based on different SV lengths (y-axis), red represents inter-haplotype SVs, and blue represents inter-accession SVs.



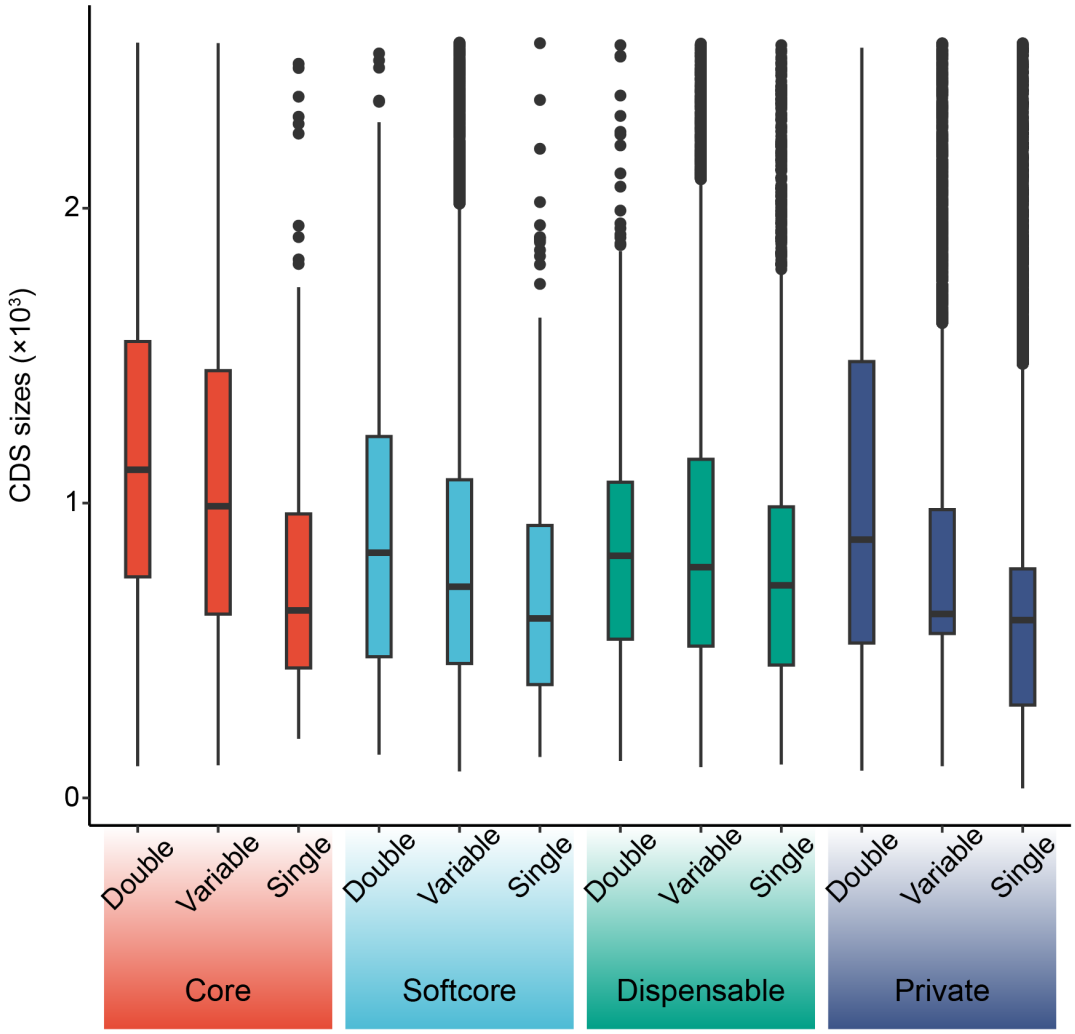
Supplementary Fig. 11 Schematic diagram illustrating the classification based on alleles. Gene sets were divided into three groups based on allele composition: double-allele gene sets (the allele was detected in paired haplotypes of all accessions), single-allele gene sets (the allele was detected in only one of the paired haplotypes of all accessions), and variable-allele gene sets (the allele was detected in paired haplotypes of some accessions). A–E represent accessions. A circle represents a present gene, while a horizontal line represents an absent gene.



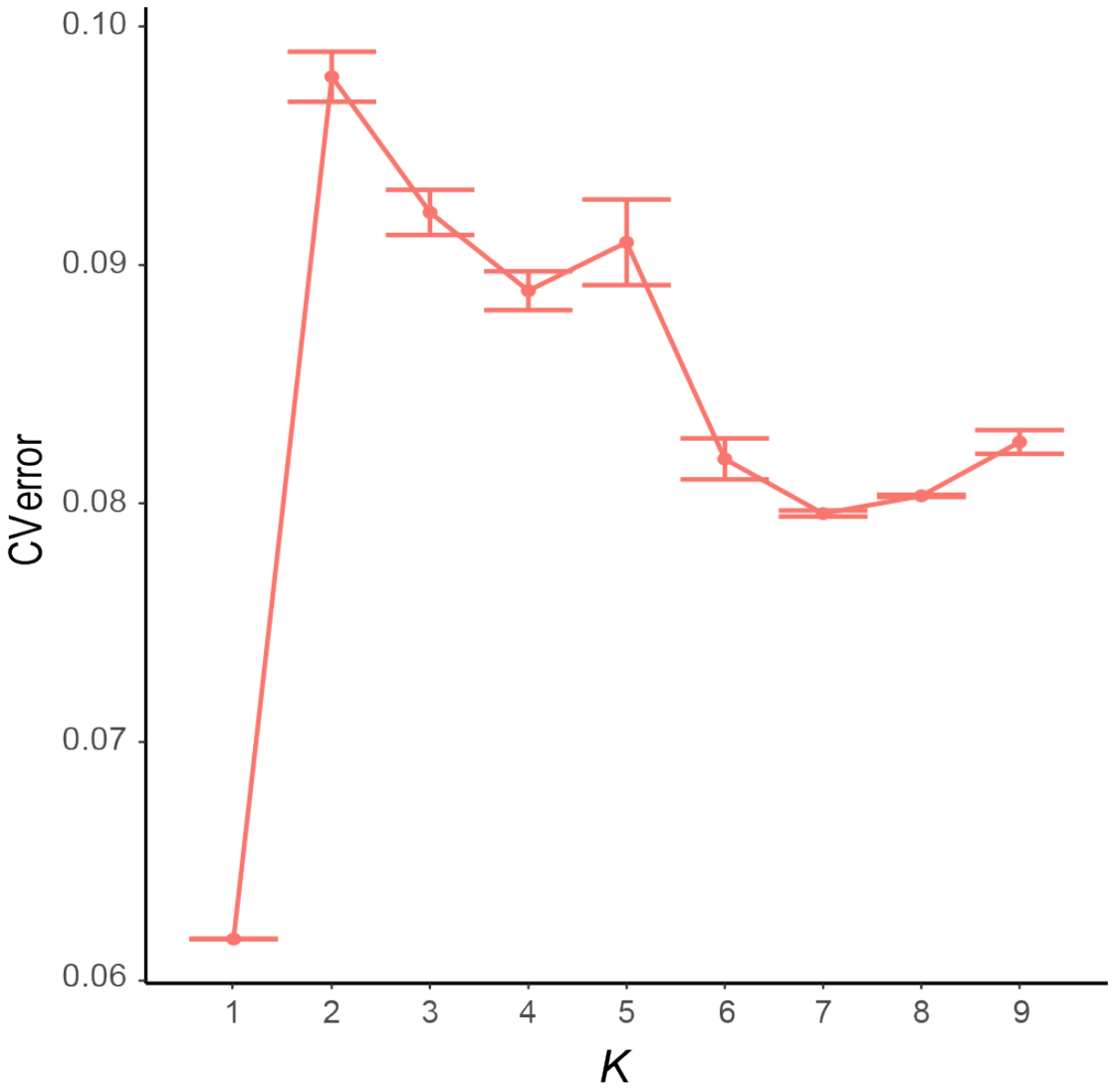
Supplementary Fig. 12 Box plot of cDNA length in the moso bamboo pangenome. The y-axis shows the CDS number, and the x-axis represents the gene set type. The box plots show the medians (centerlines), interquartile ranges (boxes), and 1.5 times the interquartile ranges (whiskers).



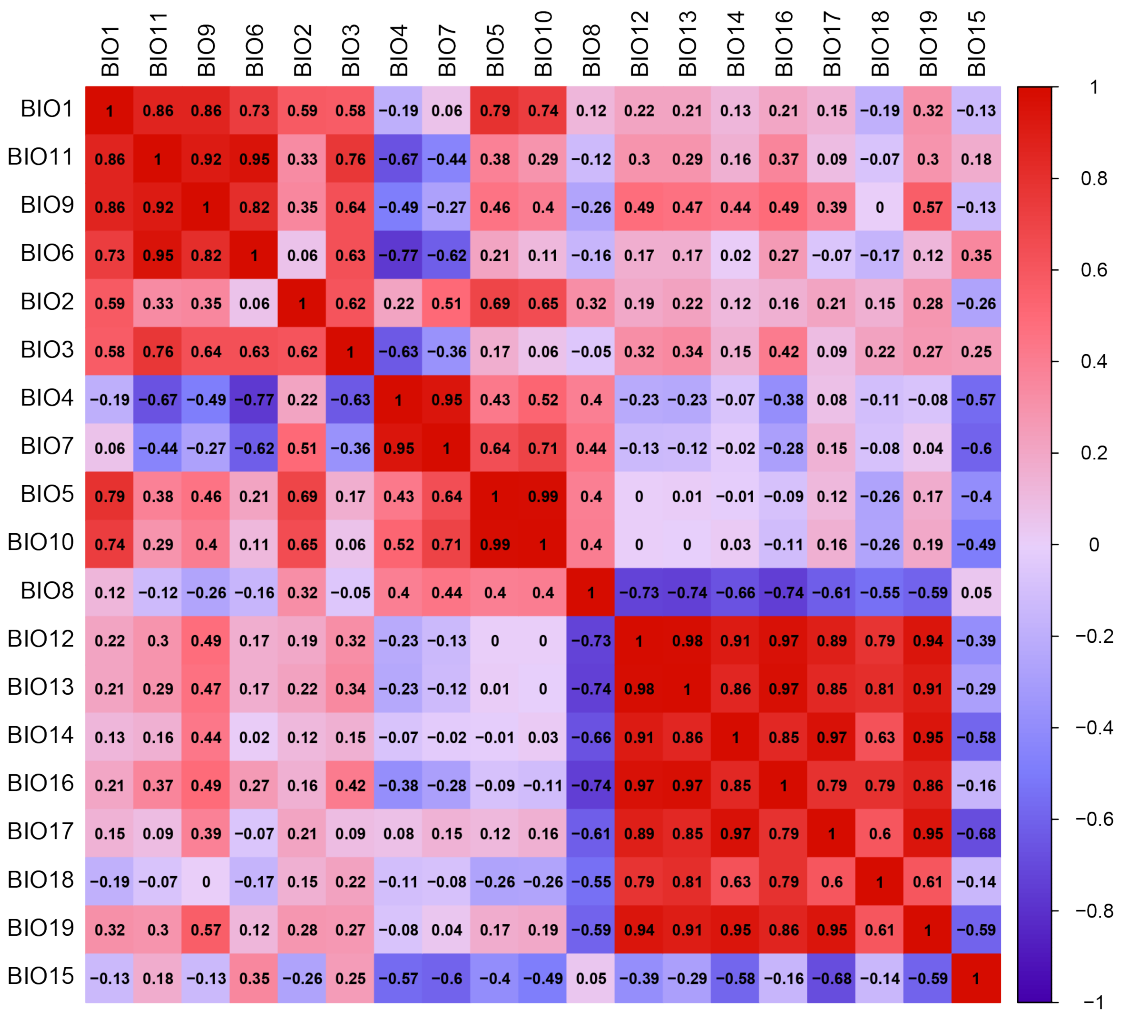
Supplementary Fig. 13 Box plot of CDS numbers in the moso bamboo pangenome. The y-axis shows the CDS number, and the x-axis represents the gene set type. The box plots show the medians (centerlines), interquartile ranges (boxes), and 1.5 times the interquartile ranges (whiskers).



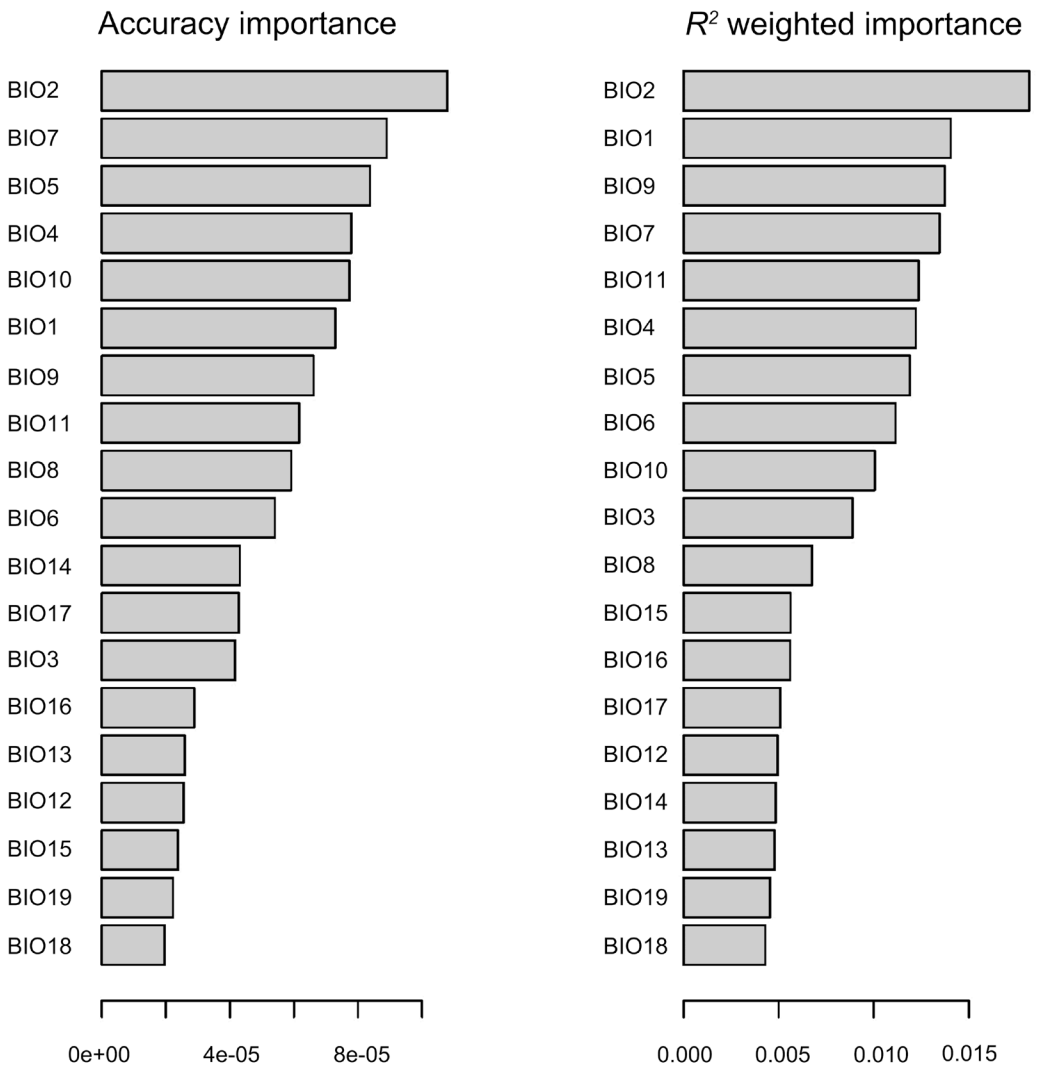
Supplementary Fig. 14 Box plot of CDS sizes in the moso bamboo pangenome. The y-axis shows CDS sizes, and the x-axis represents the gene set type. The box plots show the medians (centerlines), interquartile ranges (boxes), and 1.5 times the interquartile ranges (whiskers).



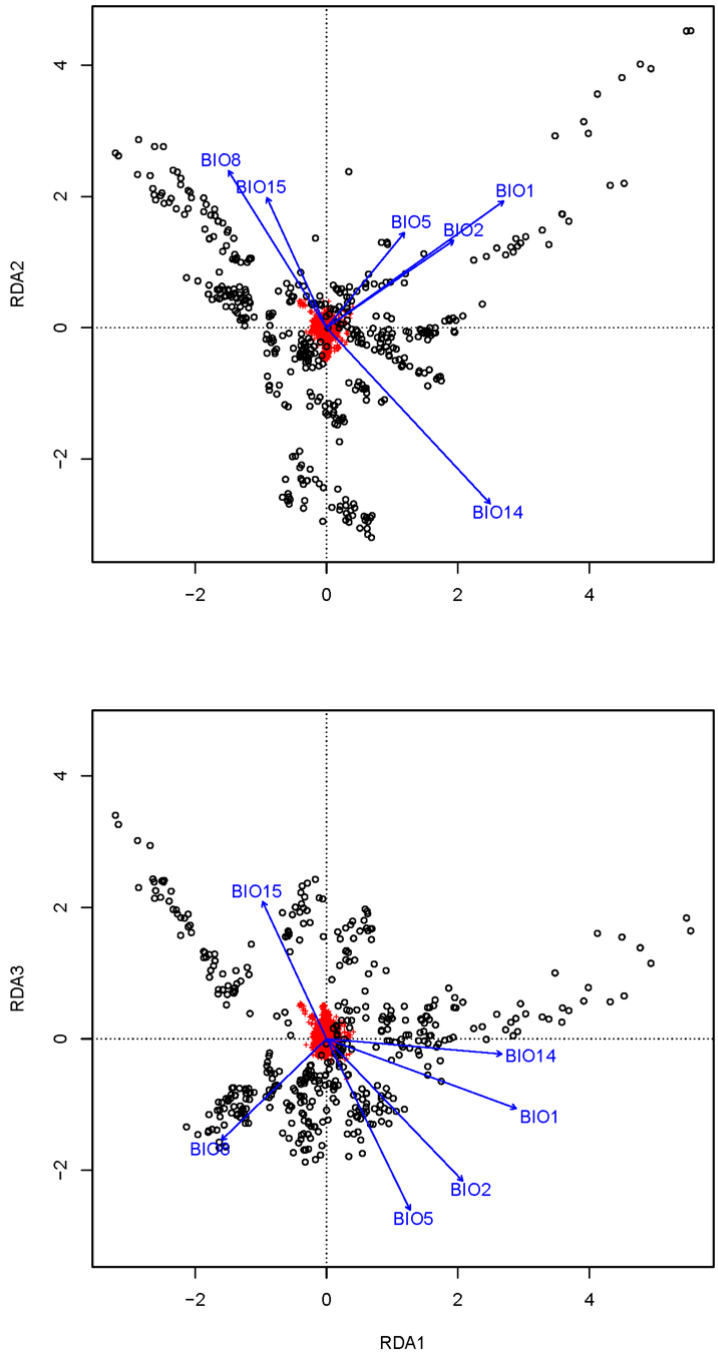
Supplementary Fig. 15 Cross-validation (CV) error values for different values of K in an ADMIXTURE analysis.



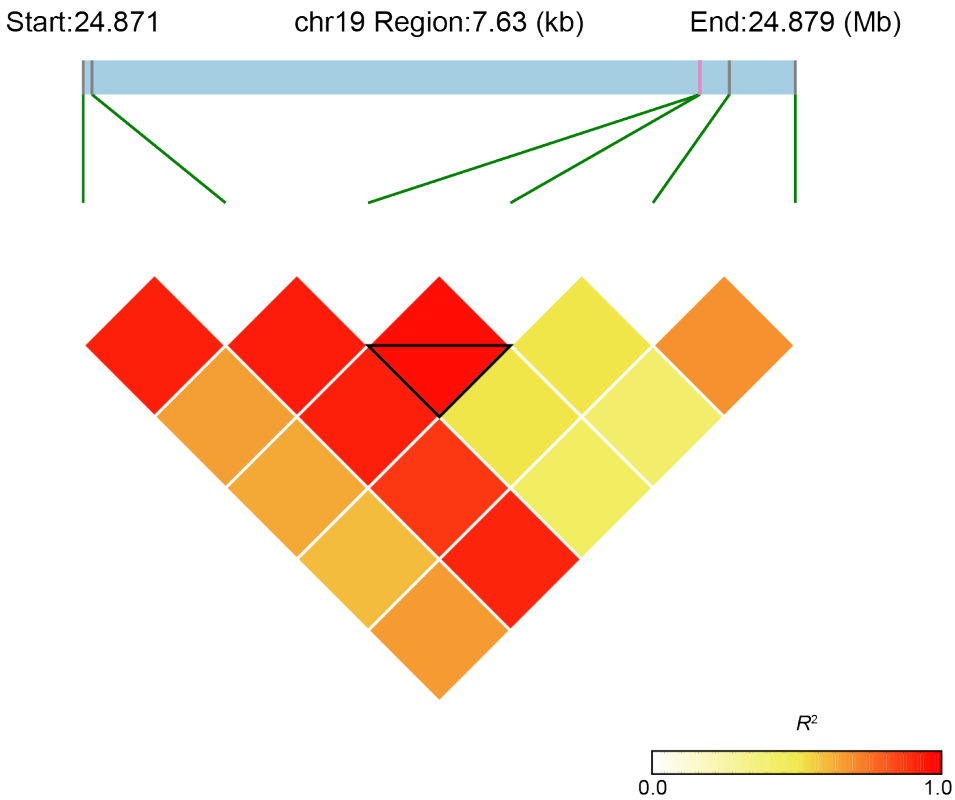
Supplementary Fig. 16 Correlations of bioclimatic variables. Red represents a positive correlation, and blue represents a negative correlation.



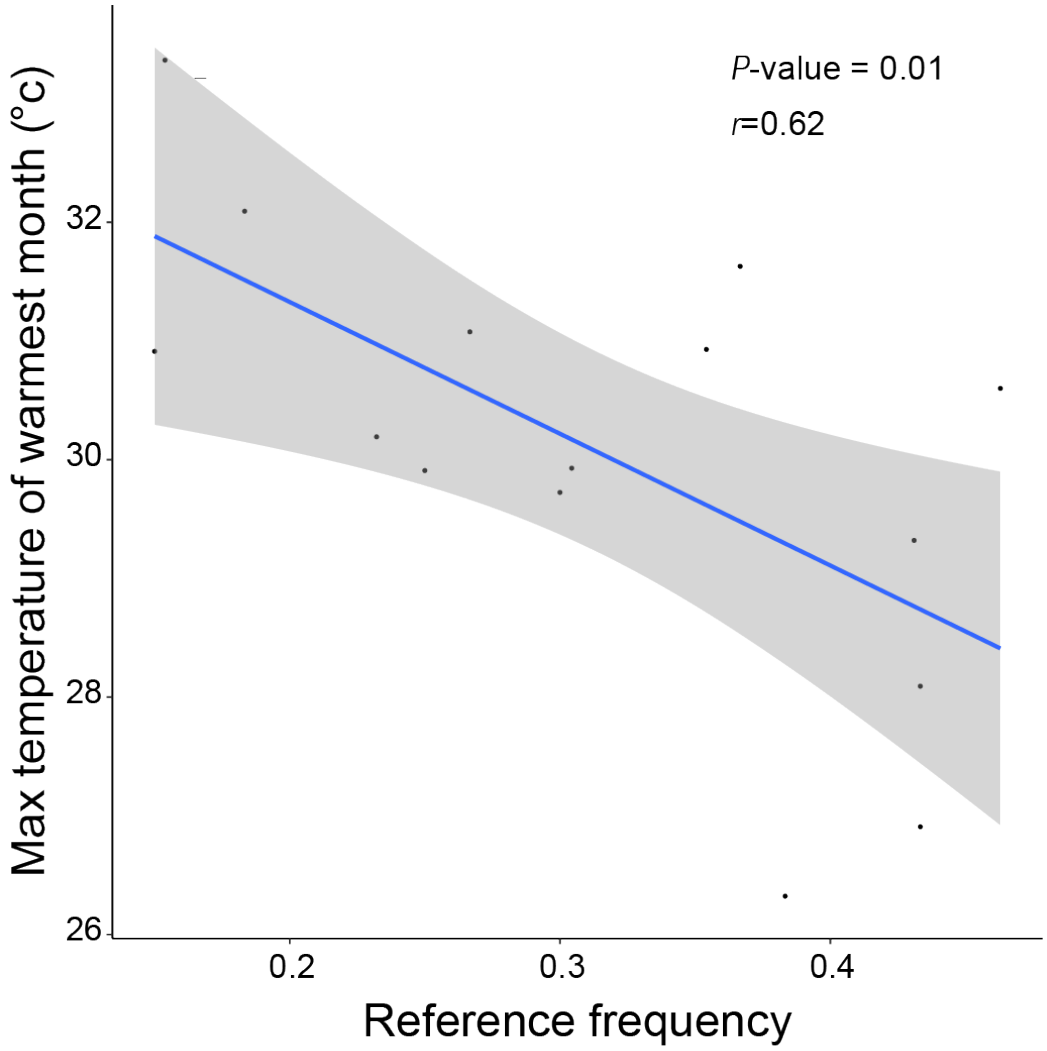
Supplementary Fig. 17 Gradient forest (GF) ranking of bioclimatic variables. The left panel shows the accuracy importance, and the right panel shows the R^2 weight importance.



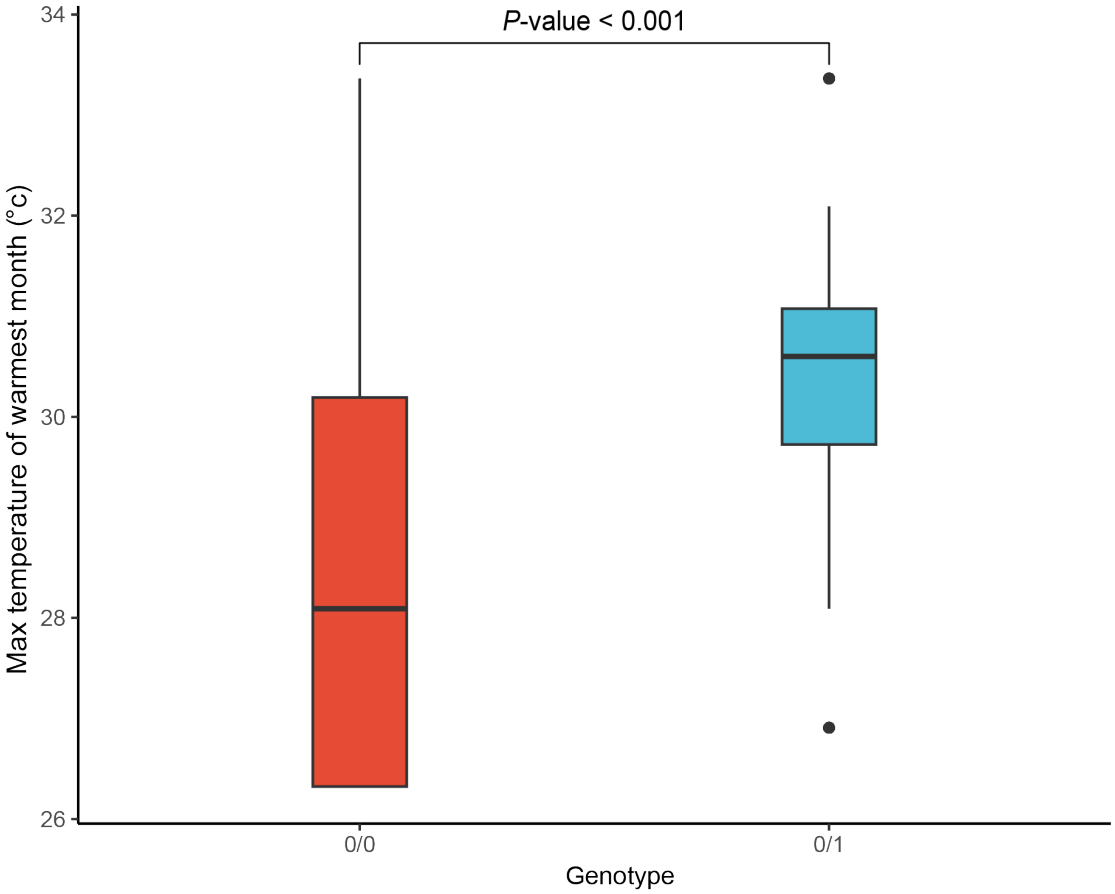
Supplementary Fig. 18 Redundancy analysis (RDA) plot. The variations are represented by red dots (in the center of each plot), and each moso bamboo accession is depicted as a black circle. The blue vectors represent the environmental variables. The upper panels represent axes 1 and 2, and the lower panels represent axes 1 and 3.



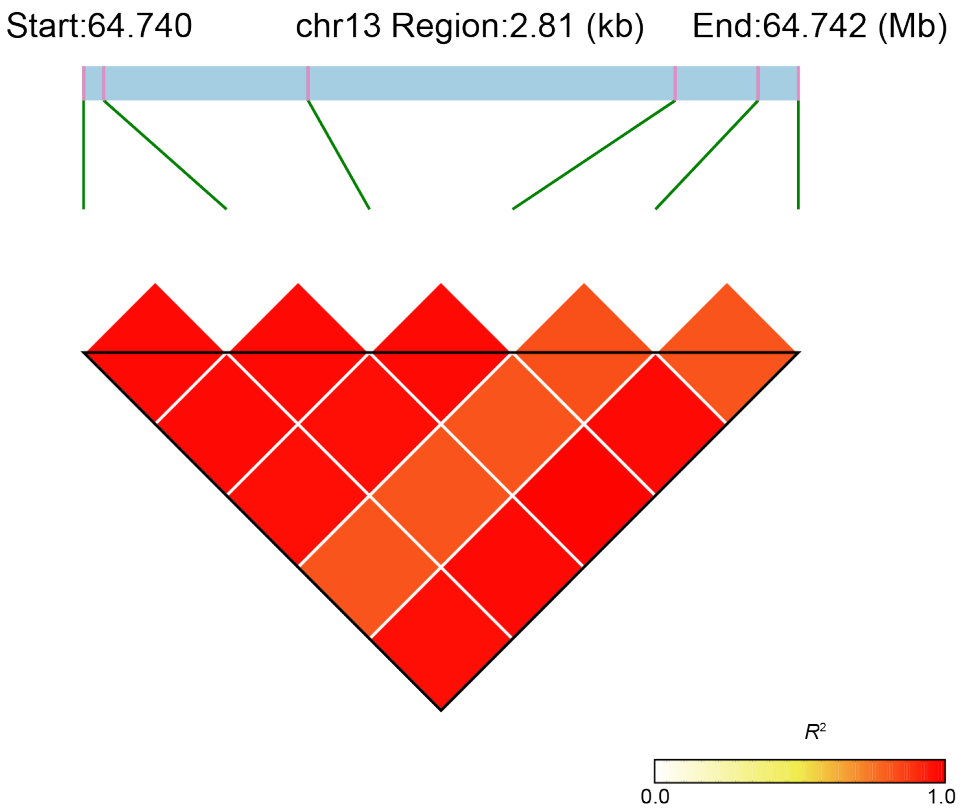
Supplementary Fig. 19 Linkage disequilibrium (LD) heatmap of the sets of SNPs associated with the maximum temperature of the warmest month (BIO5). Regional plot of LD on chromosome 19. The main triangle shows LD between variants using the R^2 color scheme, with darker red indicating greater LD (scale 0.0 to 1.0). Each square represents the R^2 value for a pair of variants at particular positions, with darker shading reflecting higher R^2 values. Positional coordinates are depicted along the triangle edges, spanning the depicted region on chromosome 19.



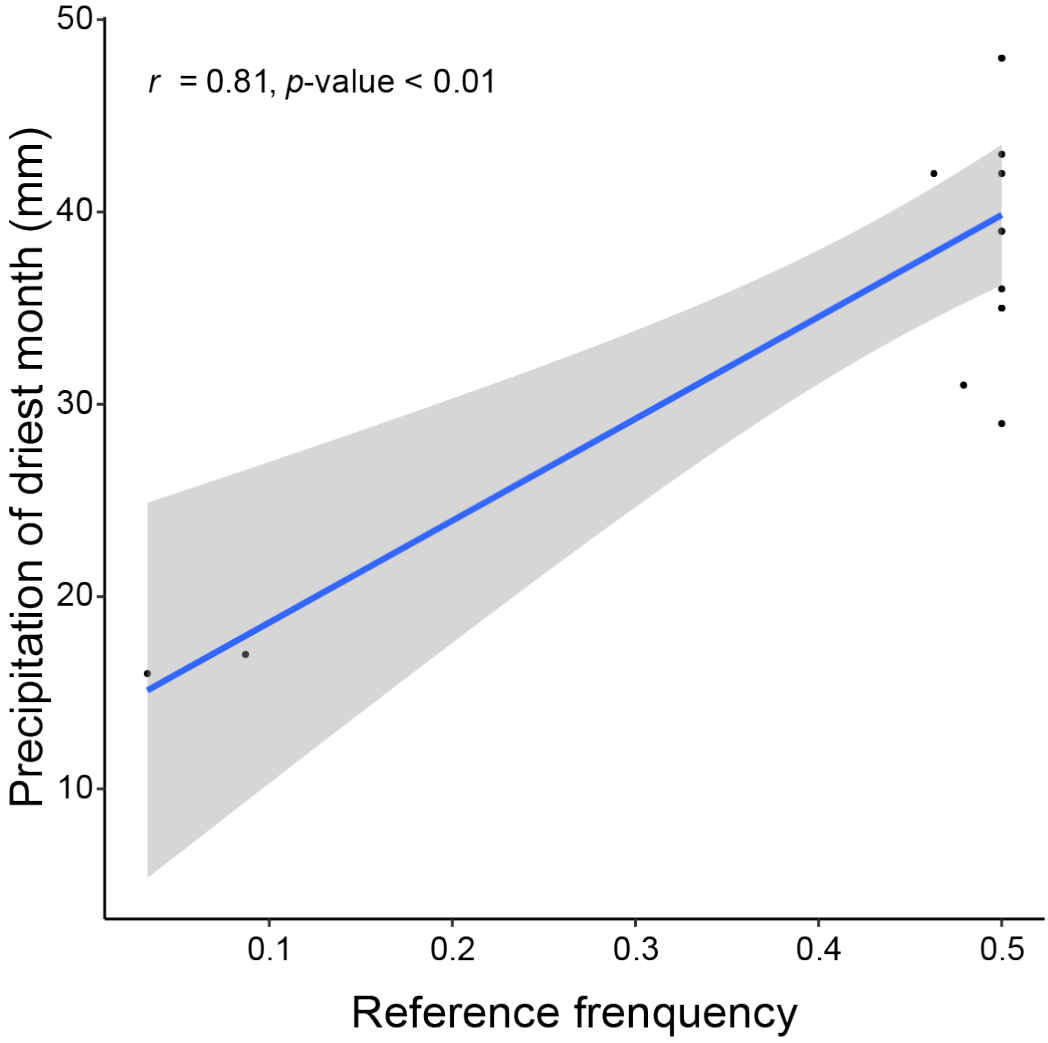
Supplementary Fig. 20 Correlation between the alternative allele frequency of chr19_24871064_SNP and the maximum temperature of the warmest month (BIO5). Scatter plot showing the relationship between the reference frequency and BIO5. Data points representing individual accessions are depicted as black circles. The blue line indicates the linear regression fit, with the surrounding shaded area representing the 95% confidence interval. A significant positive correlation was observed between ALT frequency and BIO5 (r = 0.62, two-sided Pearson's correlation test with P-values = 0.01 <0.05).



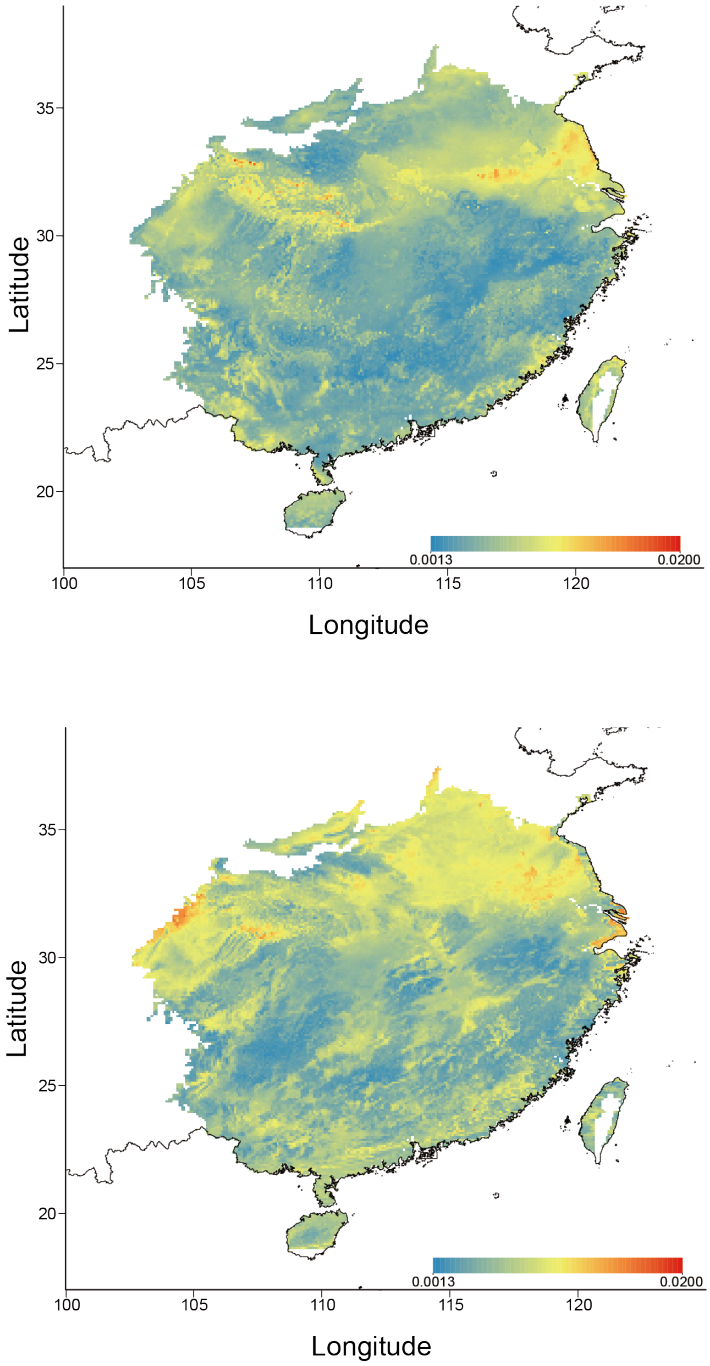
Supplementary Fig. 21 Comparison of BIO5 across different genotypes of chr18_28562210_SV. Boxplot comparing BIO15 between genotypes 0/0 and 0/1. Boxplots showing the distribution of BIO15 for genotype 0/0 (red, $n = 82$) and genotype 0/1 (blue, n = 343). Statistical significance was determined using a two-sided Wilcoxon rank-sum test, with *P*-values of 8.9e-15 <0.001.



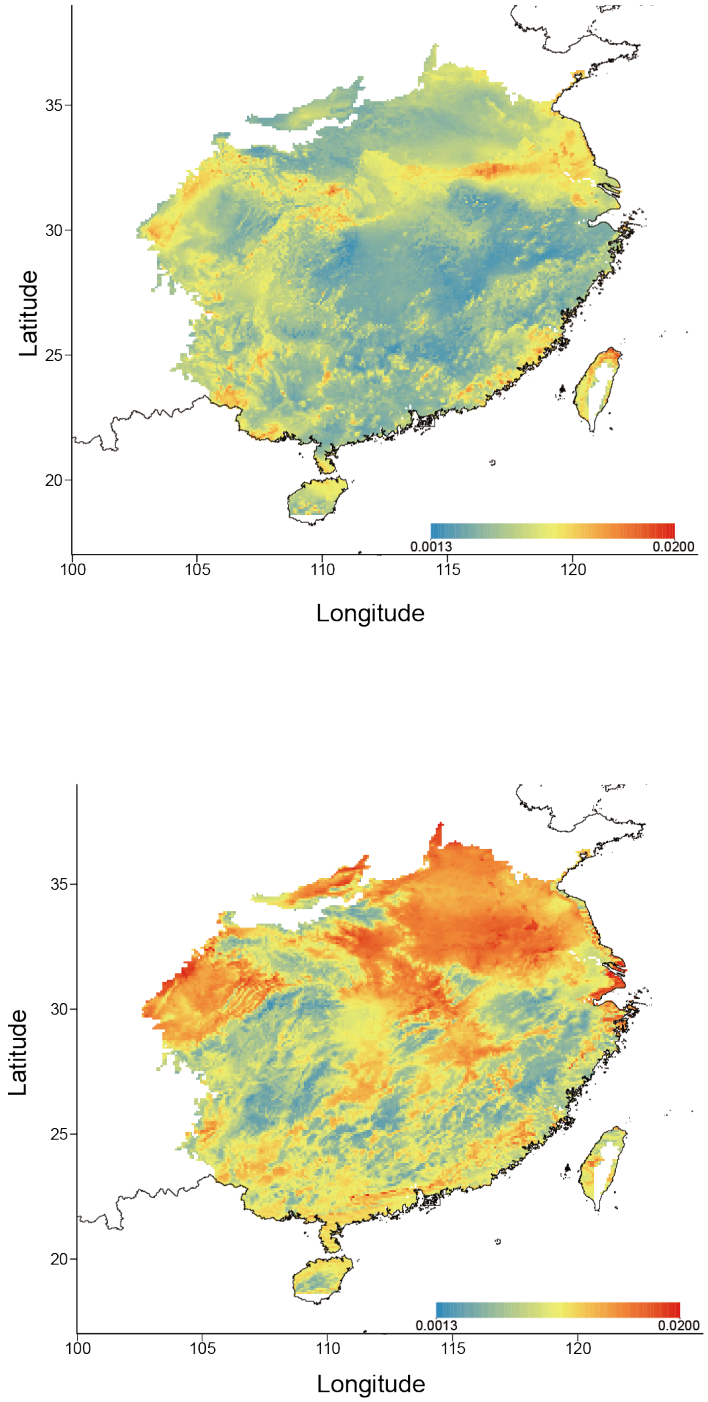
Supplementary Fig. 22 Linkage disequilibrium (LD) heatmap of the set of SNPs associated with the precipitation of the driest month (BIO14). Regional plot of LD on chromosome 13. The main triangle shows LD between variants using the R^2 color scheme, with darker red indicating greater LD (scale 0.0 to 1.0). Each square represents the R^2 value for a pair of variants at particular positions, with darker shading reflecting higher R^2 values. Positional coordinates are depicted along the triangle edges, spanning the depicted region on chromosome 13.



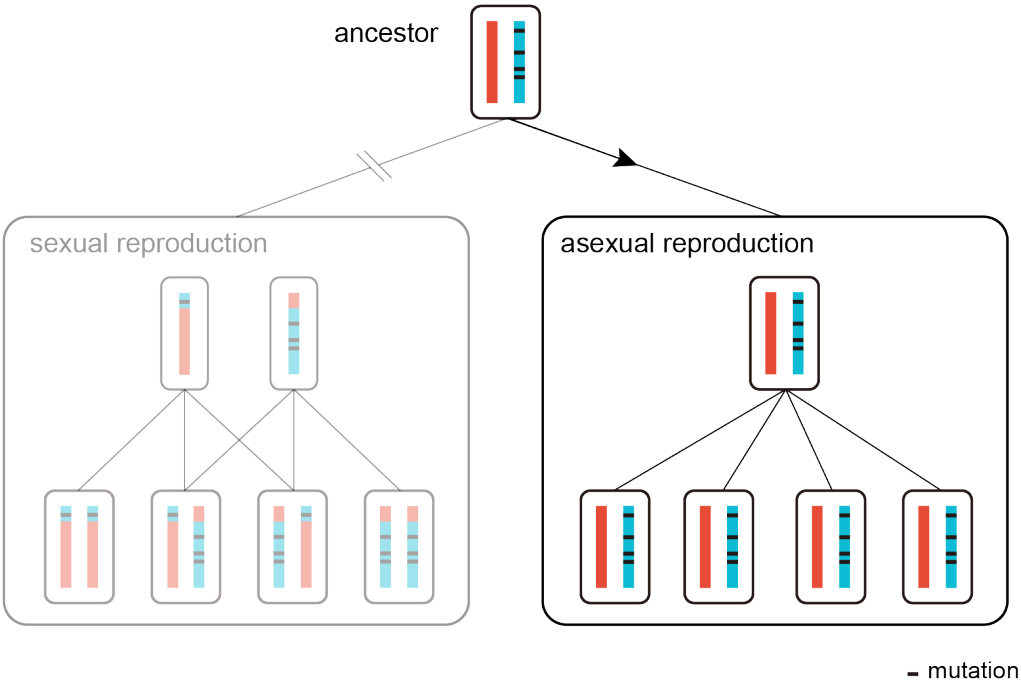
Supplementary Fig. 23 Correlation between the alternative allele frequency of chr13_64739621_SNP and the precipitation of the driest month (BIO14). Scatter plot examining the relationship between reference frequency and BIO14. Data points representing individual accessions are depicted as black circles. The blue line indicates the linear regression fit, with the surrounding shaded area representing the 95% confidence interval. A significant positive correlation was observed between the reference frequency and BIO14 (r = 0.81, two-sided Pearson's correlation test with P-values = 0.00002 <0.01).



Supplementary Fig. 24 Map of the predicted forward genomic offset and reverse genomic offset averaged across four climate models across the distribution of moso bamboo under SSP126. The upper panel shows the forward genomic offset, and the lower panel shows the reverse genomic offset.



Supplementary Fig. 25 Map of the predicted forward genomic offset and reverse genomic offset averaged across four climate models across the distribution of moso bamboo under SSP585. The upper panel shows the forward genomic offset, and the lower panel shows the reverse genomic offset.



Supplementary Fig. 26 Schematic diagrams illustrating sexual reproduction and asexual reproduction in moso bamboo. Color (red and blue) represents two different haplotypes. Although moso bamboo possesses the ability for both sexual (left) and asexual (right) reproduction, it predominantly relies on asexual reproduction through rhizome growth and vegetative propagation. This hypothesis, grounded in the understanding of the reproductive biology of moso bamboo, provides a plausible explanation for the observed pattern of genetic variation. Considering the rarity of somatic mutations, the lack of meiotic recombination due to the predominance of asexual reproduction, and the potential for haplotype divergence in the ancestral population, we provided a schematic diagram for understanding the predominance of inter-haplotype variations in moso bamboo.

Supplementary Table 1. Summary statistics of PacBio HiFi reads for the 16 representative moso bamboo accessions.

Accession	HB	AJ	RH	LY	HZP	YA	CS	YF	JZ	XN	DA	XA	YX	WYS	CY	HS
Polymerase reads	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	1E+07	2E+07
Polymerase reads bases (bp)	8E+11	1E+12	1E+12	1E+12	1E+12	9E+11	1E+12	1E+12	1E+12	1E+12	9E+11	9E+11	1E+12	1E+12	1E+12	1E+12
Average polymerase reads length (bp)	76,899	94,195	99,195	222,705	89,838	96,326	222,160	220,750	90,469	88,764	87,609	86,804	215,238	217,030	97,253	77,542
Polymerase reads length N50 (bp)	158,935	19,3965	207,696	420,055	195,713	205,878	427,707	424,832	193,151	196,566	194,660	194,983	421,984	410,912	209,832	181,749
Subreads reads	6E+07	7E+07	6E+07	8E+07	6E+07	6E+07	6E+07	7E+07	7E+07	7E+07	5E+07	6E+07	7E+07	6E+07	7E+07	8E+07
Subreads bases (bp)	8E+11	1E+12	1E+12	1E+12	1E+12	9E+11	1E+12	1E+12	1E+12	1E+12	9E+11	9E+11	1E+12	1E+12	1E+12	1E+12
Average subreads length (bp)	13,863	14,595	15,844	13,822	16,198	16,347	15,614	14,754	14,612	14,875	16,813	16,683	14,722	16,428	15,640	14,977
Subreads length N50 (bp)	15,688	16,754	17,464	14,771	18,470	18,424	17,541	15,962	15,443	16,304	18,775	19,024	16,055	17,882	17,519	16,612
GC mean	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.45	0.44	0.44	0.45	0.44	0.44	0.46	0.44	0.45
CCS reads	4E+06	4E+06	4E+06	4E+06	3E+06	4E+06	4E+06	4E+06	4E+06	4E+06	3E+06	3E+06	4E+06	4E+06	4E+06	4E+06
CCS bases (bp)	6E+10	7E+10	7E+10	6E+10	6E+10	6E+10	7E+10	7E+10	6E+10	6E+10	5E+10	6E+10	7E+10	7E+10	7E+10	7E+10
Average CCS length (bp)	15,850	16,239	17,657	14,766	17,977	18,345	17,401	16,059	15,096	16,157	17,855	18,609	15,977	17,318	17,494	16,904
Max. CCS length (bp)	49,767	50,109	50,422	50,459	50,140	49,990	50,086	50,779	48,434	49,222	50,577	50,821	49,735	50,105	50,395	49,878
CCS length N50 (bp)	16,177	17,353	18,094	14,929	18,835	19,081	18,175	16,496	15,114	16,591	18,646	19,366	16,107	17,689	18,316	16,929
Depth (X) based on 1.91 Gb	32.433	34.216	37.768	32.973	32.418	33.694	34.067	34.265	32.613	31.5	27.992	31.848	35.003	33.976	36.265	36.747

Supplementary Table 2. Estimated switch errors for phased genomes of the 16 representative moso bamboo accessions.

Accession	Switch error (%)
AJ	6.43
CS	6.89
CY	2.36
DA	9.06
HB	3.58
HS	3.39
HZP	3.56
JZ	4.52
LY	4.40
RH	4.78
WYS	6.12
XA	5.73
XN	5.54
YA	6.49
YF	8.40
YX	5.82

Supplementary Table 3. Summary statistics of Hi-C data for accessions CY, HB, and HZP.

Accession	Raw Reads Number	Raw Bases Number	Clean Reads Number	Clean Bases Number	Q30 (%)	Depth (X) based on 1.91 Gb
CY	1,755,395,808	263,309,371,200	1,719,807,574	257,971,136,100	91.11	135.06
HZP	1,816,619,620	272,492,943,000	1,734,429,002	260,164,350,300	89.25	136.21
HB	1,747,116,446	262,067,466,900	1,626,471,218	243,970,682,700	89.08	127.73
Average	1,773,043,958	265,956,593,700	1,693,569,265	254,035,389,700	89.81	139.24
Total	5,319,131,874	797,869,781,100	5,080,707,794	762,106,169,100	-	417.73

Supplementary Table 4. Summary statistics of allele pairs.

Accession	Haplotype 1	Haplotype 2	Gene with two alleles	Allele pairs	Single alleles in Haplotype 1	Single alleles in Haplotype 2
AJ	55,190	52,271	90,490	45,245	9,945	7,026
CS	53,812	51,062	83,194	41,597	12,215	9,465
CY	55,903	55,465	98,128	49,064	6,839	6,401
DA	54,118	52,947	91,294	45,647	8,471	7,300
HB	54,656	55,112	96,322	48,161	6,495	6,951
HS	58,130	56,413	99,660	49,830	8,300	6,583
HZP	55,781	56,456	97,722	48,861	6,920	7,595
JZ	53,426	54,871	92,648	46,324	7,102	8,547
LY	53,416	51,582	90,444	45,222	8,194	6,360
RH	54,220	52,994	92,362	46,181	8,039	6,813
WYS	55,658	54,087	93,674	46,837	8,821	7,250
XA	55,207	52,029	86,020	43,010	12,197	9,019
XN	53,867	53,984	91,764	45,882	7,985	8,102
YA	55,876	54,896	97,164	48,582	7,294	6,314
YF	53,988	52,498	83,760	41,880	12,108	10,618
YX	55,533	53,514	95,444	47,722	7,811	5,792

Supplementary Table 5. Comparison of variations between each accession and the reference genome haplotypes.

	Short variation			SV		
Accession	**Same as reference (inter-haplotype)**	**Only in query (inter-accession)**	**Only in reference (inter-accession)**	**Same as reference (inter-haplotype)**	**Only in query (inter-accession)**	**Only in reference (inter-accession)**
AJ	2,027,150	83,147	106,297	17,509	1,656	1,528
CS	2,021,790	78,516	113,659	17,486	1,634	1,551
DA	2,048,481	86,352	82,940	17,147	1,505	1,890
HB	2,028,758	83,939	105,328	17,071	1,576	1,966
HS	2,065,266	108,353	62,670	17,383	2,183	1,654
HZP	2,072,642	82,486	61,725	17,431	1,626	1,606
JZ	2,062,377	83,934	71,672	17,385	1,605	1,652
LY	2,060,083	85,260	73,755	17,282	1,609	1,755
RH	2,026,487	76,197	107,457	17,319	1,700	1,718
WYS	2,032,146	76,796	101,897	17,442	1,589	1,595
XA	2,039,072	83,367	94,910	17,337	1,722	1,700
XN	2,021,787	85,512	109,003	17,273	1,600	1,764
YA	2,050,853	83,801	83,331	17,402	1,683	1,635
YF	2,004,919	106,908	120,330	17,223	2,110	1,814
YX	2,022,490	81,956	112,175	17,386	1,622	1,651

Supplementary Table 6. Estimated frequency of short variations.

Accession	All short variants number	Reference size (bp)	Number of all short variants per bp	Inter-accession number	Number of inter-accession short variants per bp
AJ	2,110,297	1,962,743,480	930.08	83,147	23,605.70
CS	2,100,306	1,962,743,480	934.50	78,516	24,998.01
DA	2,134,833	1,962,743,480	919.39	86,352	22,729.57
HB	2,112,697	1,962,743,480	929.02	83,939	23,382.97
HS	2,173,619	1,962,743,480	902.98	108,353	18,114.34
HZP	2,155,128	1,962,743,480	910.73	82,486	23,794.87
JZ	2,146,311	1,962,743,480	914.47	83,934	23,384.37
LY	2,145,343	1,962,743,480	914.89	85,260	23,020.68
RH	2,102,684	1,962,743,480	933.45	76,197	25,758.80
WYS	2,108,942	1,962,743,480	930.68	76,796	25,557.89
XA	2,122,439	1,962,743,480	924.76	83,367	23,543.41
XN	2,107,299	1,962,743,480	931.40	85,512	22,952.84
YA	2,134,654	1,962,743,480	919.47	83,801	23,421.48
YF	2,111,827	1,962,743,480	929.41	106,908	18,359.18
YX	2,104,446	1,962,743,480	932.67	81,956	23,948.75
Total	31,870,825	1,962,743,480	923.86	975,223	23,104.86

Supplementary Table 7. Number of SNPs between three haplotype-resolved accessions.

Assembly 1	Assembly 2	SNP number
CY hap1	HZP hap1	785,577
CY hap1	HZP hap2	2,792,068
CY hap2	HZP hap2	787,969
CY hap2	HZP hap1	2,794,863
CY hap1	HB hap1	433,970
CY hap1	HB hap2	3,085,151
CY hap2	HB hap2	468,651
CY hap2	HB hap1	3,137,431
HZP hap1	HB hap1	683,455
HZP hap1	HB hap2	2,907,917
HZP hap2	HB hap2	667,226
HZP hap2	HB hap1	2,883,394

* The number of SNPs was obtained directly from the genome alignment.

Supplementary Table 8. Standard deviation of genome-wide distribution patterns for inter-haplotype and inter-accession variations.

Variations type	Inter-haplotype variation	Inter-accession variation	Inter-haplotype SV	Inter-accession SV
Standard Deviation	236.7887292	24.15605428	2.07802511	1.09571224

Supplementary Table 9. Statistical significance of differences in gene length between moso bamboo pangenome groups.

Group 1	Group 2	*P*-values
core-double	core-single	4.15e-24
softcore-double	softcore-single	9.98e-07
dispensable-double	dispensable-single	4.25e-09
private-double	private-single	7.00e-69
core-double	dispensable-double	3.19e-14
core-variable	dispensable-variable	7.73e-248
core-single	dispensable-single	0.58
core-double	private-double	5.30e-84

**P*-values were calculated using a two-sided Wilcoxon rank-sum test.

Supplementary Table 10. Statistical significance of differences in expression level (TPM) between moso bamboo pangenome groups.

Group 1	Group 2	*P*-values
core-double	core-single	2.00e-16
softcore-double	softcore-single	3.47e-14
dispensable-double	dispensable-single	1.32e-23
private-double	private-single	0.033
core-double	dispensable-double	3.12e-25
core-variable	dispensable-variable	0
core-single	dispensable-single	1.03e-08
core-double	private-double	3.99e-77

**P*-values were calculated using a two-sided Wilcoxon rank-sum test.

Supplementary Table 11. Statistical significance of differences in the tissue specificity index (Tau) between moso bamboo pangenome groups.

Group 1	Group 2	*P*-value
core-double	core-single	1.52e-05
softcore-double	softcore-single	0.003
dispensable-double	dispensable-single	0.06
private-double	private-single	0.46
core-double	dispensable-double	7.19e-16
core-variable	dispensable-variable	4.21e-204
core-single	dispensable-single	0.002
core-double	private-double	3.74e-30

**P*-values were calculated using a two-sided Wilcoxon rank-sum test.

Supplementary Table 12. RDA models of inertia proportions of variations.

RDA models	Candidate Inertia	Proportion of candidate variations	Total Inertia	Proportion of total variations
F~geno	136.71	0.13	13,110	0.00494
F~env	372.34	0.35	38,090	0.01435
F~ geno|env	52.78	0.05	12,410	0.004676
F~ env|geno	288.42	0.27	37,390	0.01409
F~env+geno	425.13	0.40	50,500	0.019
Total inertia	1050.00	1.00	2,654,000	1

Supplementary Table 13. Software packages, versions, and parameters used in this study.

Software	Version	Parameters	Function
Genome assembly			
CCS algorithm	v6.2.0	default parameters	Generate CCS reads
Hifiasm	v0.16.1-r375	default parameters	Genome assembling
BUSCO	v5.4.3	-m genome	Genome evaluation
LTR_retriever	v2.9.0	default parameters	LTR prediction
Meryl	v1.41	k = 20	Generate Meryl databases
Merqury	v1.3	default parameters	Genome evaluation
Genome annotation			
RepeatModeler	v2.0.3	-LTRStruct	de novo repeat library
RepeatMasker	v4.1.2-p1	default parameters	de novo repeat library
GeneWise	v2.4.1	-max_gene_length 23707 -segmentSize 1000000 -overlapSize 100000 -coverage_ratio 0.4 -evalue 1e-9	Gene prediction
Trimmomatic	v0.38	default parameters	RNA-seq quality control
HISAT2	v2.1.0	-min-intronlen 20 -max-intronlen 20000 -dta -score-min L,0.0,-0.4	Mapping to genome
AUGUSTUS	v3.4.0	default parameters	ab initio prediction
BUSCO	v5.4.3	default parameters	Annotation evaluation
DIAMOND		--sensitive --max-target-seqs 200 --evalue 1e-5	Mapping to database
Switch errors calculation			
pbmm2	v1.5.0	--preset HIFI --sort	Mapping to genome
DeepVariant	v1.4.0	--model_type PACBIO	Variation detection
Whatshap	v1.1	--ignore-read-groups	SNP phasing
nucmer	v4.0.0rc1	-mum -l 1000 -c 200 -g 200	Genome alignment
show-snps	v4.0.0rc1	default parameters	SNP calling
Bwa mem	v0.7.17	default parameters	Mapping to genome
DeepVariant	v1.4.0	-SP5M	Variation detection (Hi-C)
Allele identification			
BLASTP	v2.9.0+	-evalue 1e-10	Protein alignment
Minimap2	v2.24-r1122	-ax asm5	Genome alignment
Allele-specific expression			
HISAT2	v2.1.0	-k 1	Mapping to genome

StringTie	v1.3.5	default parameters	TPM calculation
DESeq2	v1.34.0	default parameters	Differential expression analysis
SNP and InDel identification			
nucmer	v4.0.0rc1	--mum -l 100 -c 200 -g 200	Genome alignment
show-snps	v4.0.0rc1	-Clr	SNP calling
svim-asm	v1.0.2	--min_sv_size 1 --min_sv_size 50	InDel calling
pbmm2	v1.5.0	--preset HIFI --sort	Mapping to genome
DeepVariant	v1.4.063	--model_type PACBIO	Variation calling
Structural variation identification			
Minimap2	v2.24-r1122	--paf-no-hit -ax map-hifi	Mapping to genome
cuteSV	v1.0.13	-l 50 -s 5 --max_cluster_bias_INS 1000 --diff_ratio_merging_INS 0.9 --max_cluster_bias_DEL 1000 --diff_ratio_merging_DEL 0.5	SV calling
Ngmlr	v0.2.7	default parameters	Mapping to genome
sniffles	v1.0.12	-s 5 -l 50	SV calling
pbmm2	v1.5.0	--preset HIFI	Mapping to genome
pbsv	v2.8.0	default parameters	SV calling
Minimap2	v2.24-r1122	--paf-no-hit -a -x asm5 --cs	Mapping to genome
svim-asm	v1.0.2	--min_sv_size 50	SV calling
Nucmer	v4.0.0rc1	--maxmatch -l 100 -c 500	Mapping to genome
Assemblytics	v1.2.1	default parameters	SV calling
Graph-based pangenome construction			
vg construct	v1.38	-a -S	Construct vg from a reference
vg gbwt	v1.38	-P	Manipulate GBWTs
vg snarls	v1.38	default parameters	Generate snarls from graph
vg index	v1.38	default parameters	Generate vg index
vg giraffe	v1.38	default parameters	Short read mapping to vg
SNP and InDel calling based on resequenced reads			
BWA	v0.7.17	-M	Mapping to genome
GATK GVCF	v4.2.0	-ERC GVCF --native-pair-hmm-threads 100	SNP calling
PLINK	v1.9	-indep-pairphase 100 10 0.2	LD pruning
ADMIXTURE	v1.3.0	-cv -j4	Assess population structure

Identification of climate-associated variations			
vcftools	v0.1.13	--minDP 3 --max-missing 0.2 --maf 0.05	Filter SNP
LEA	v3.6.0	K = 1	LFMM2 analysis
Calculation of RONA and genomic offset			
pyRona	v0.36	-P 0.01	RONA analysis