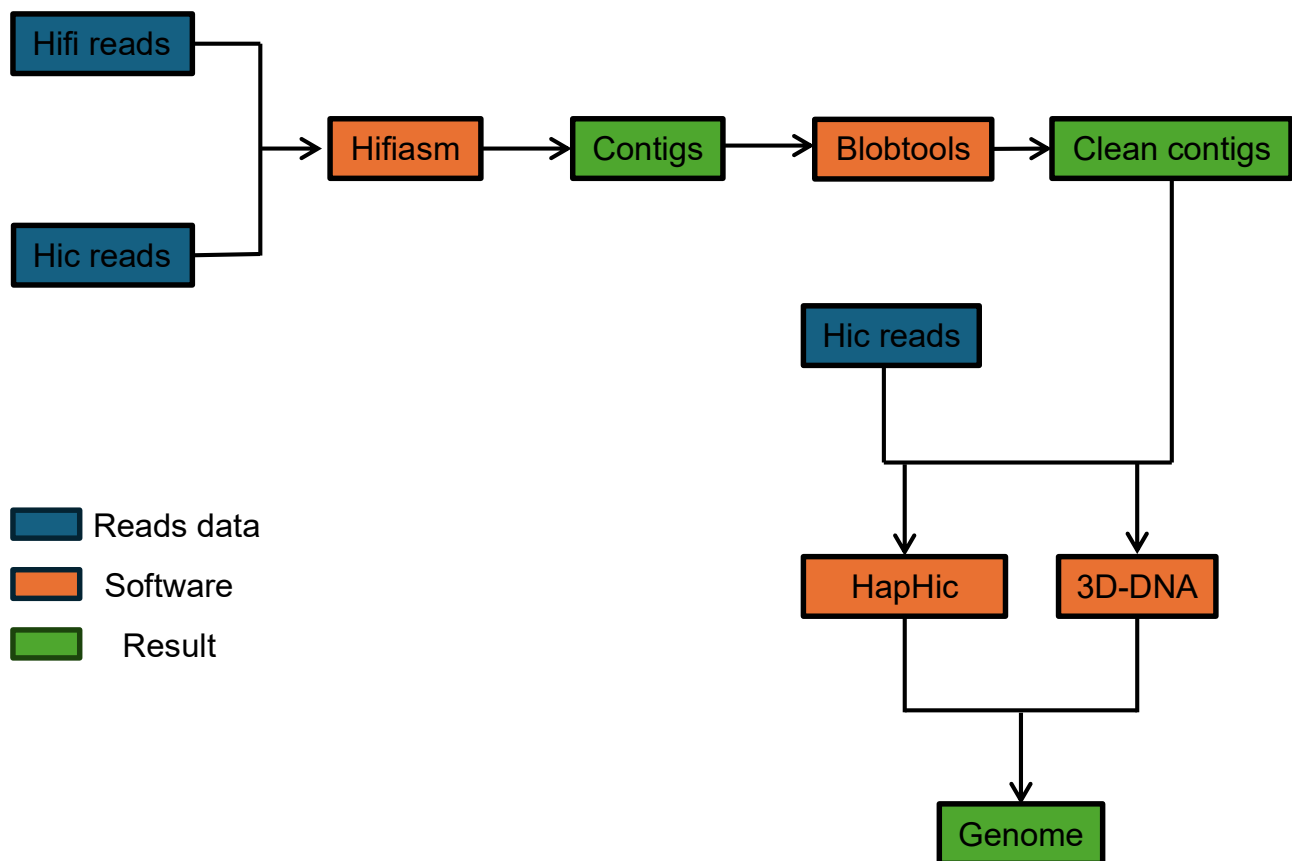
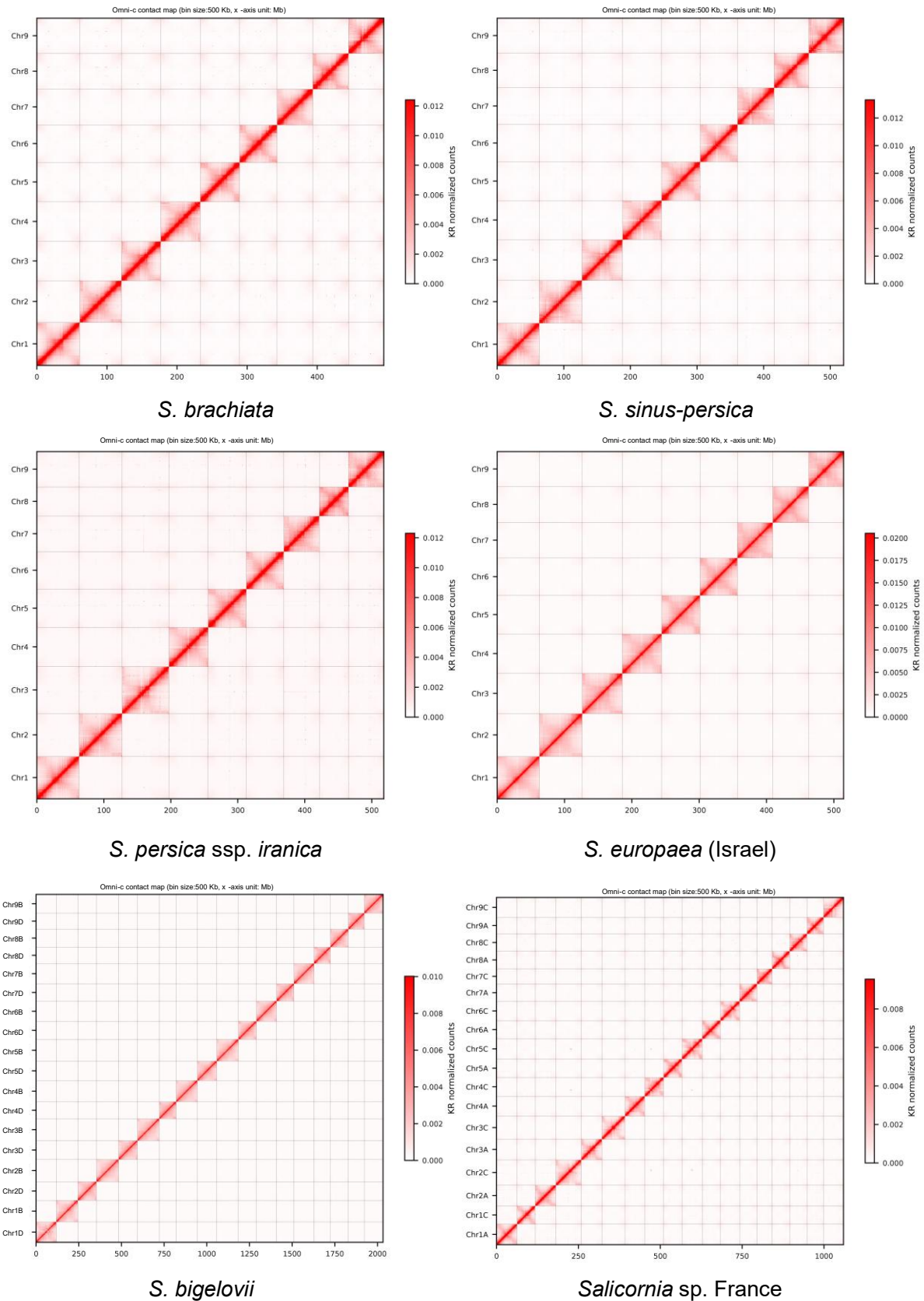


Chromosome-scale genomes and population resequencing resolve subgenome diversity and halophyte adaptation in *Salicornia*

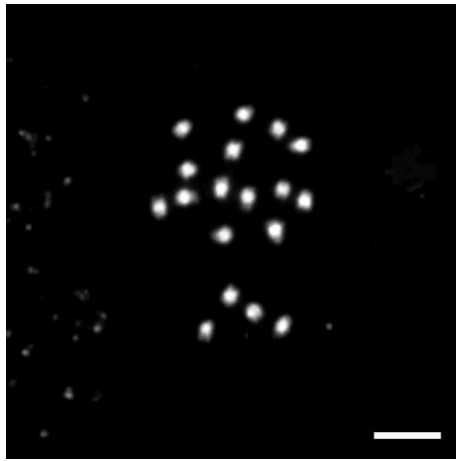
Wang and Adhikari *et al.*



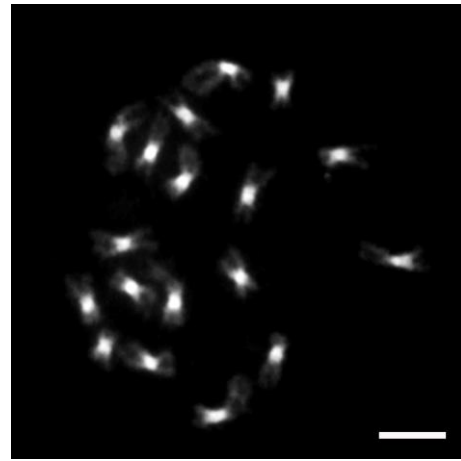
Supplementary Fig. 1. The genome assembly pipeline used in this study. PacBio HiFi reads were assembled into contigs with Hifiasm, contaminant contigs were removed with BlobTools, and the cleaned contigs were scaffolded into chromosome-scale pseudomolecules using Hi-C reads with 3D-DNA and HapHic.



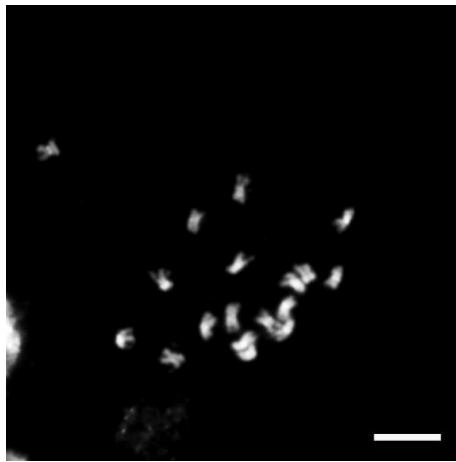
Supplementary Fig. 2. Omni-C contact heatmaps of the six chromosome-scale *Salicornia* genome assemblies. Heatmap intensity represents the frequency of Omni-C contacts between genomic loci, and the block-diagonal signal indicates correctly scaffolded pseudomolecules.



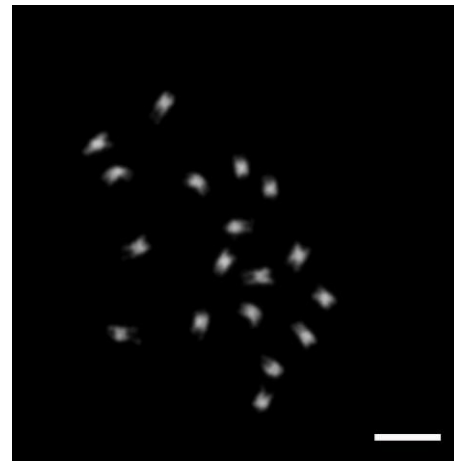
S. brachiata



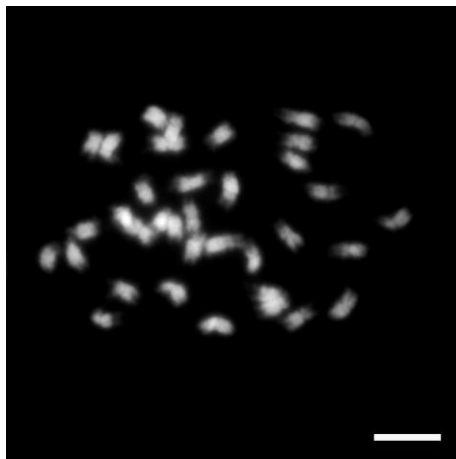
S. persica* ssp. *iranica



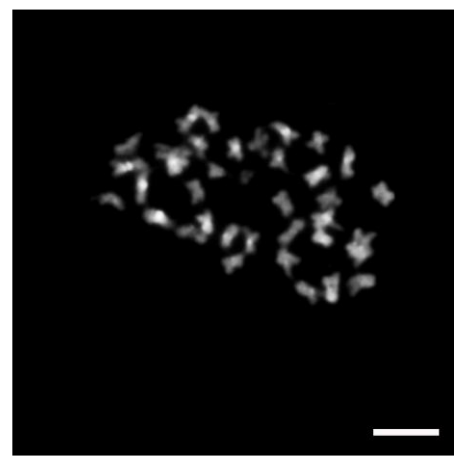
S. sinus-persica



***S. europaea* (Israel)**

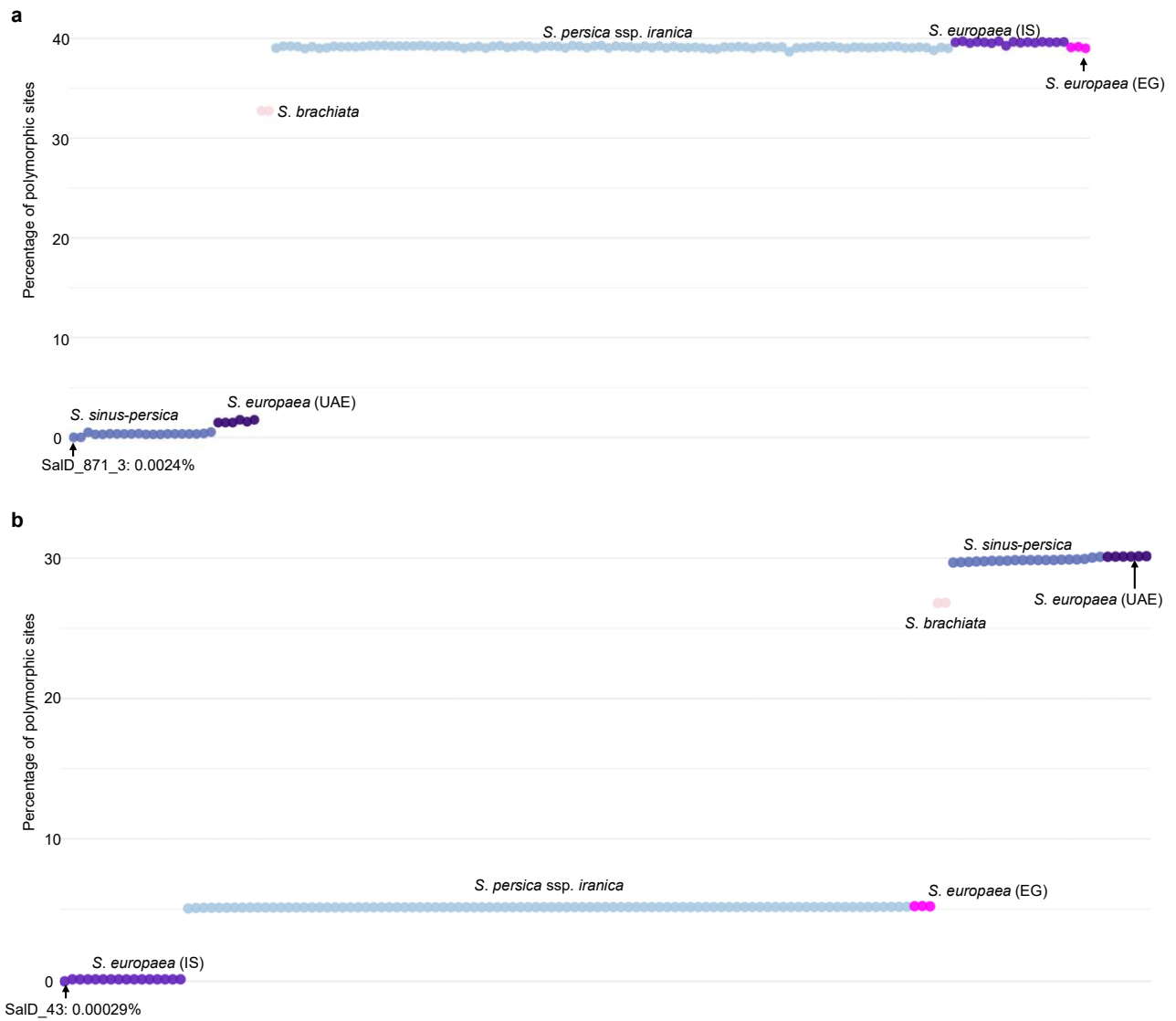


S. bigelovii

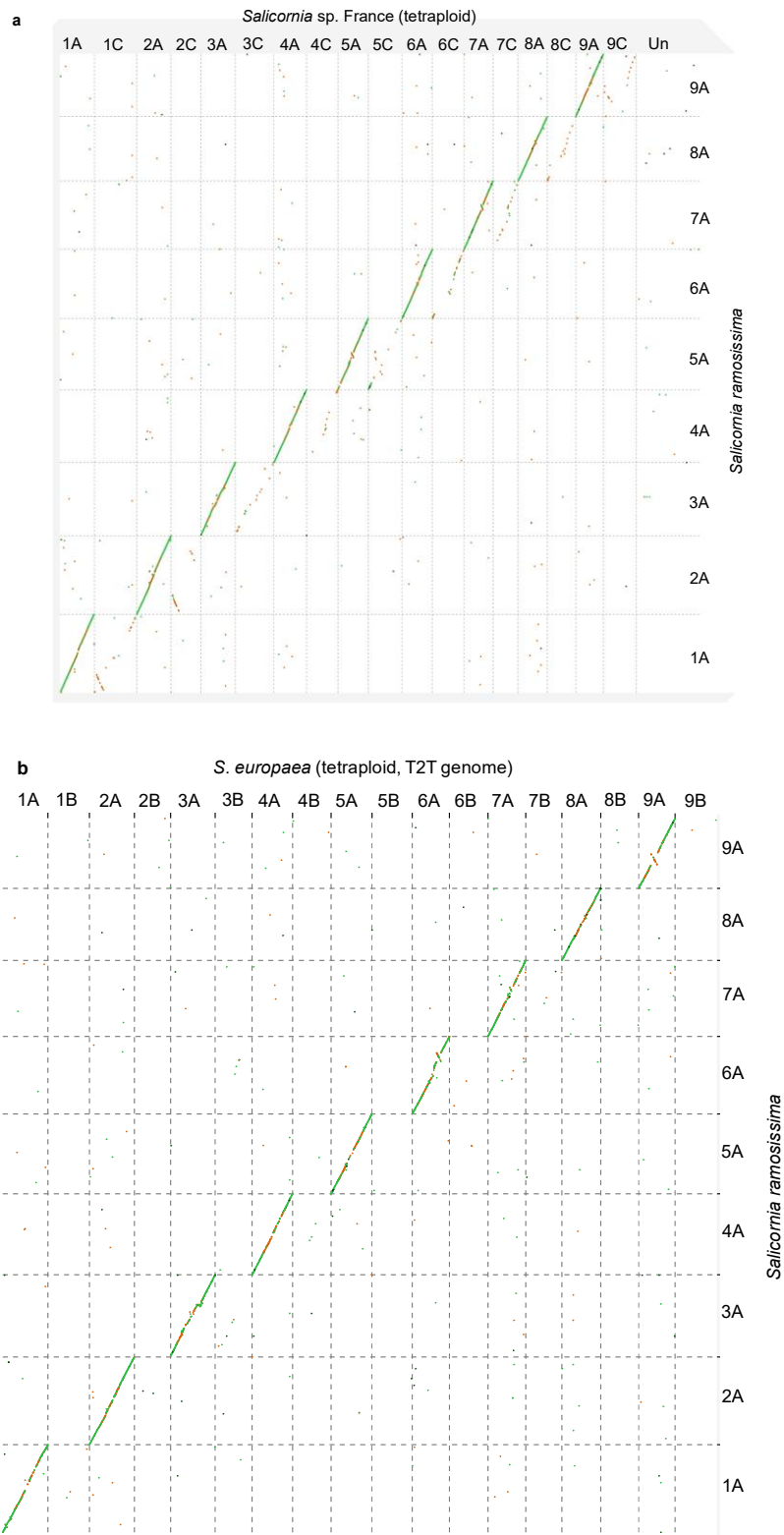


S. procumbens

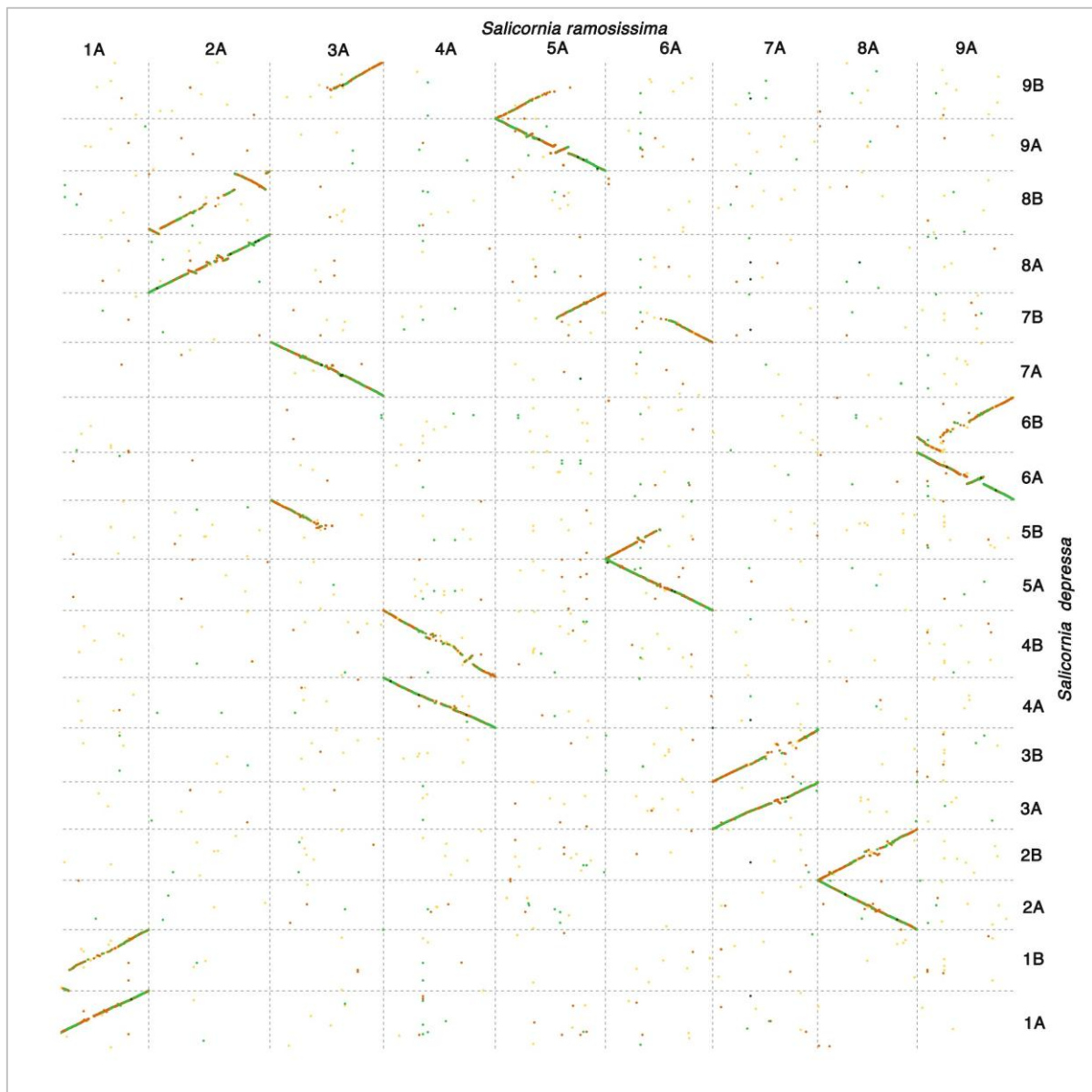
Supplementary Fig. 3. Chromosome counts on metaphase chromosomes of six *Salicornia* species. Species shown are *S. brachiata*, *S. sinus-persica*, *S. europaea* (Israel), *S. persica* ssp. *iranica*, *S. bigelovii* and *S. procumbens*. Chromosomes were stained with 4',6-diamidino-2-phenylindole (DAPI). Scale bars, 5 μ m. A minimum of five complete metaphase spreads were captured per species with consistent counts, except for *S. procumbens*, for which only one complete metaphase spread was recovered across all prepared slides.



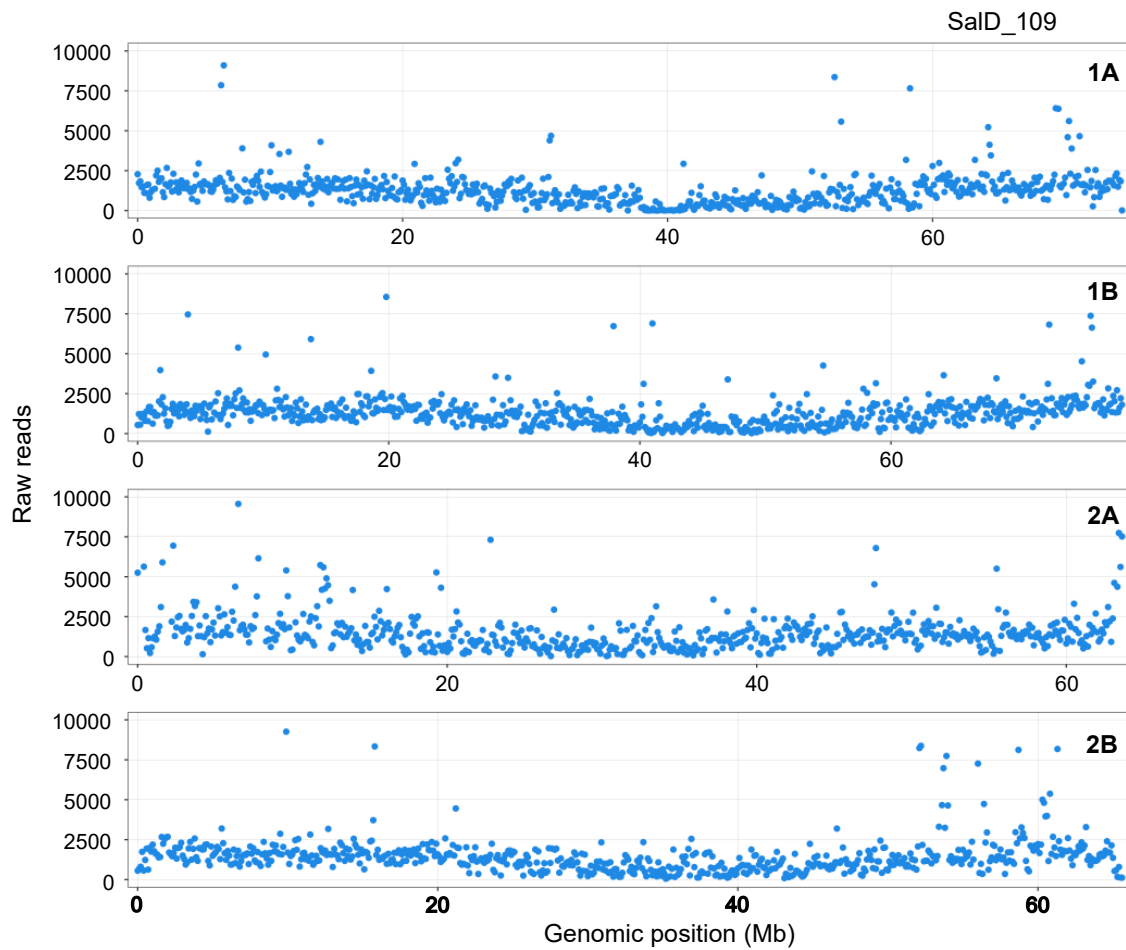
Supplementary Fig. 4. Assembly validation using Illumina short reads. a Percentage of polymorphic sites for each diploid *Salicornia* species accession compared with the SalD_871_3 reference assembly of *S. sinus-persica*. Whole-genome sequencing reads of an independently resequenced SalD_871_3 plant revealed a very low error rate of 0.0024%. **b** Percentage of polymorphic sites for each diploid *Salicornia* species accession compared with the SalD_43 reference assembly of *S. europaea* from Israel. Whole-genome sequencing reads of an independently resequenced SalD_43 plant revealed a very low error rate of 0.00029%.



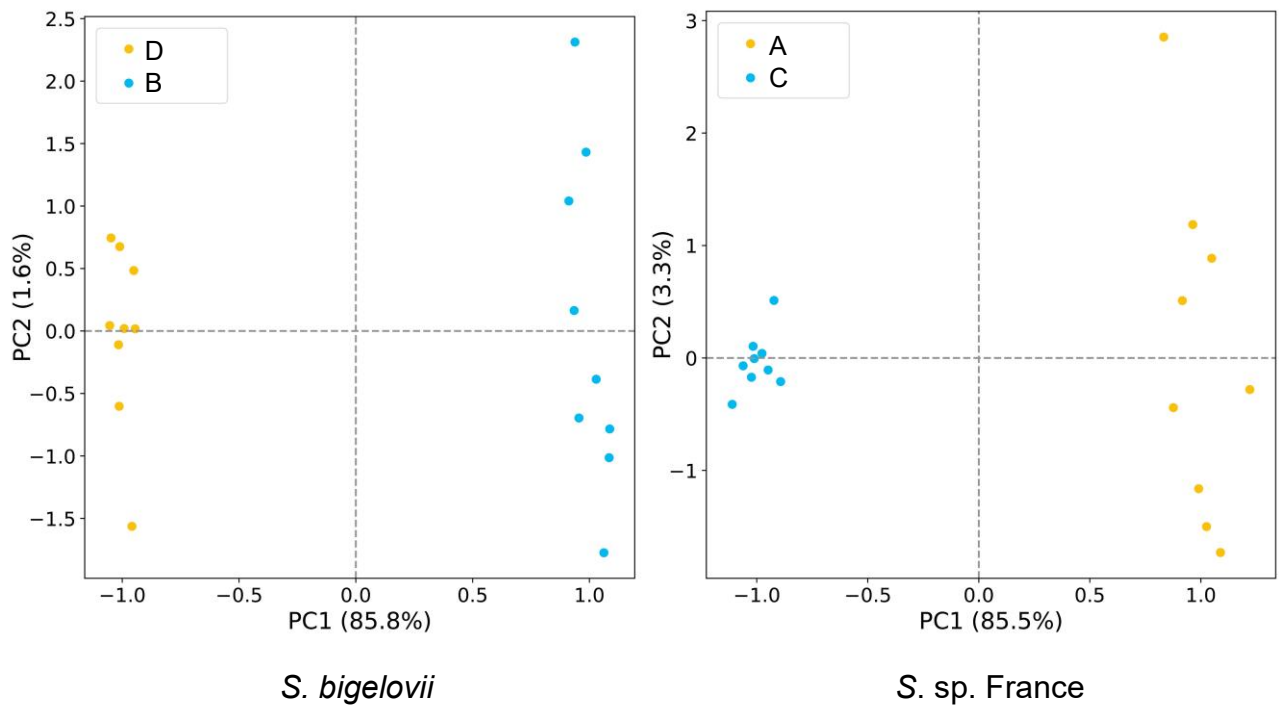
Supplementary Fig. 5. Dot plot showing homology among different genome assemblies. a Genome-wide synteny and collinearity dot plot between *Salicornia* sp. France (tetraploid) and *S. ramosissima*; **b** Genome-wide synteny and collinearity dot plot between *S. europaea* (tetraploid T2T genome, China) and *S. ramosissima* (UK).



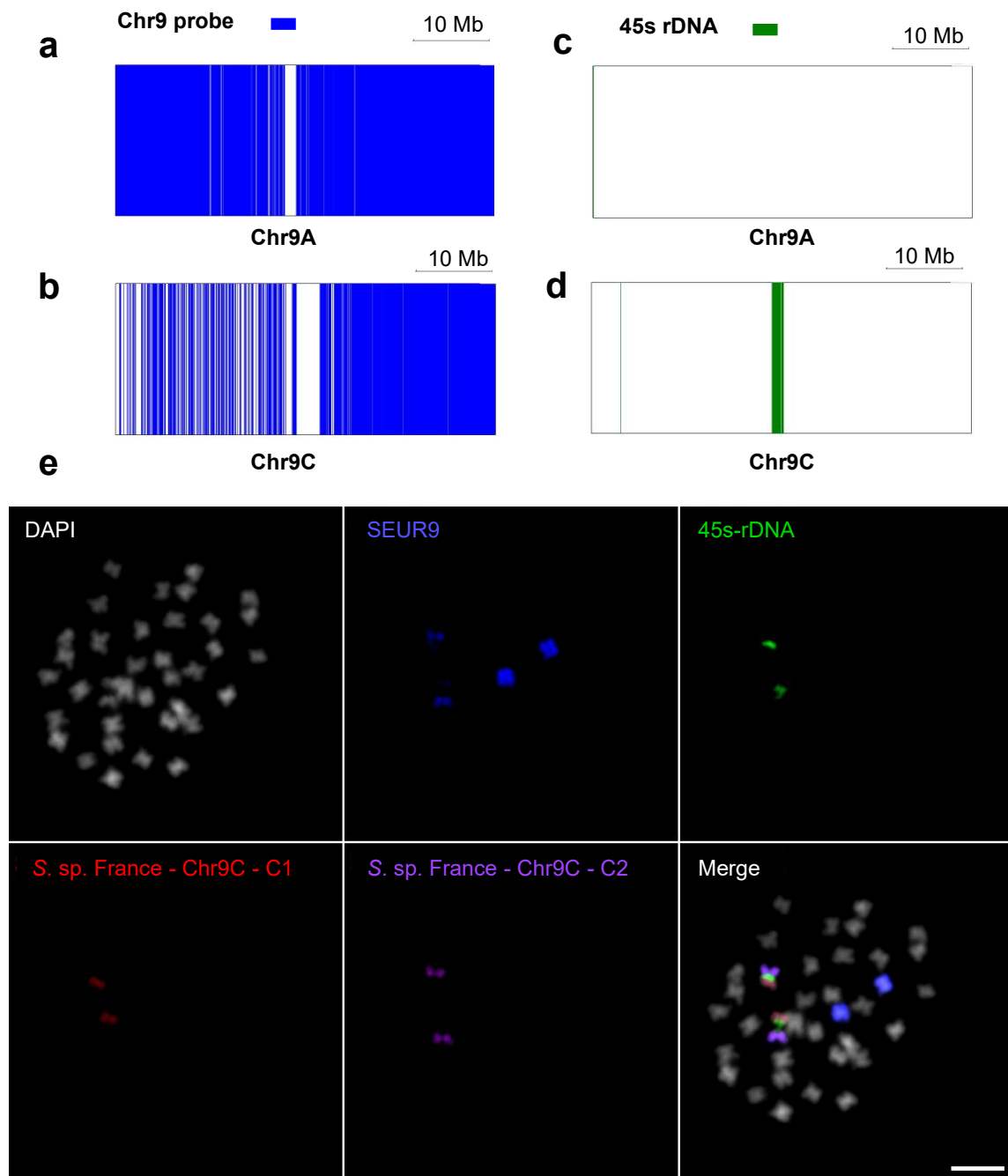
Supplementary Fig. 6. Dot plot showing discrepancies in chromosome numbering between *Salicornia depressa* and *S. ramosissima*. Chromosome nomenclature follows that of the respective published genome assemblies.



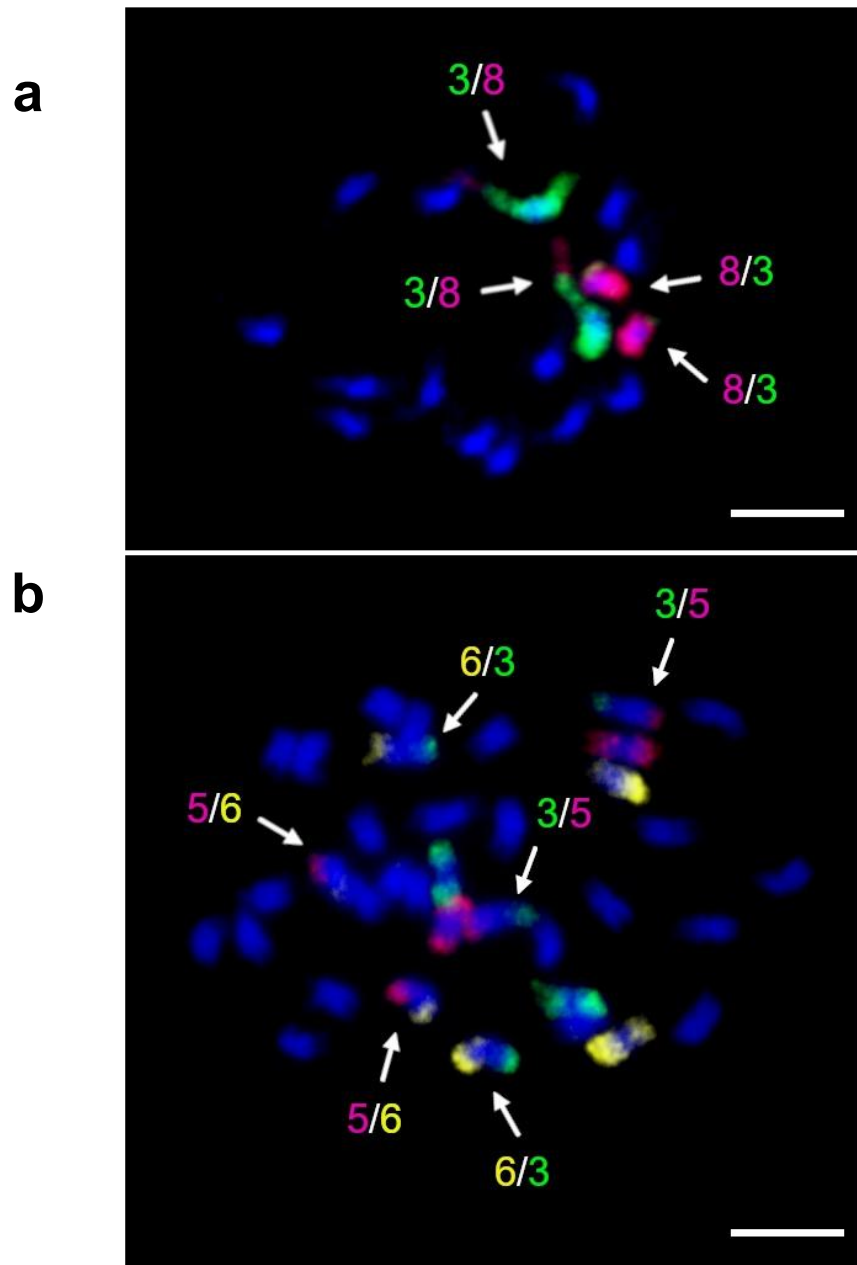
Supplementary Fig. 7. Genomic relationships between *S. bigelovii* and *S. depressa*. Read-depth coverage plots of the *S. bigelovii* accession assembled in this study (SalD_109) mapped onto the *S. depressa* chromosomes. Reads derived from the B subgenome of *S. bigelovii* mapped to the B subgenome of *S. depressa*, whereas reads from the D subgenome of *S. bigelovii* aligned with the A subgenome chromosomes of *S. depressa*.



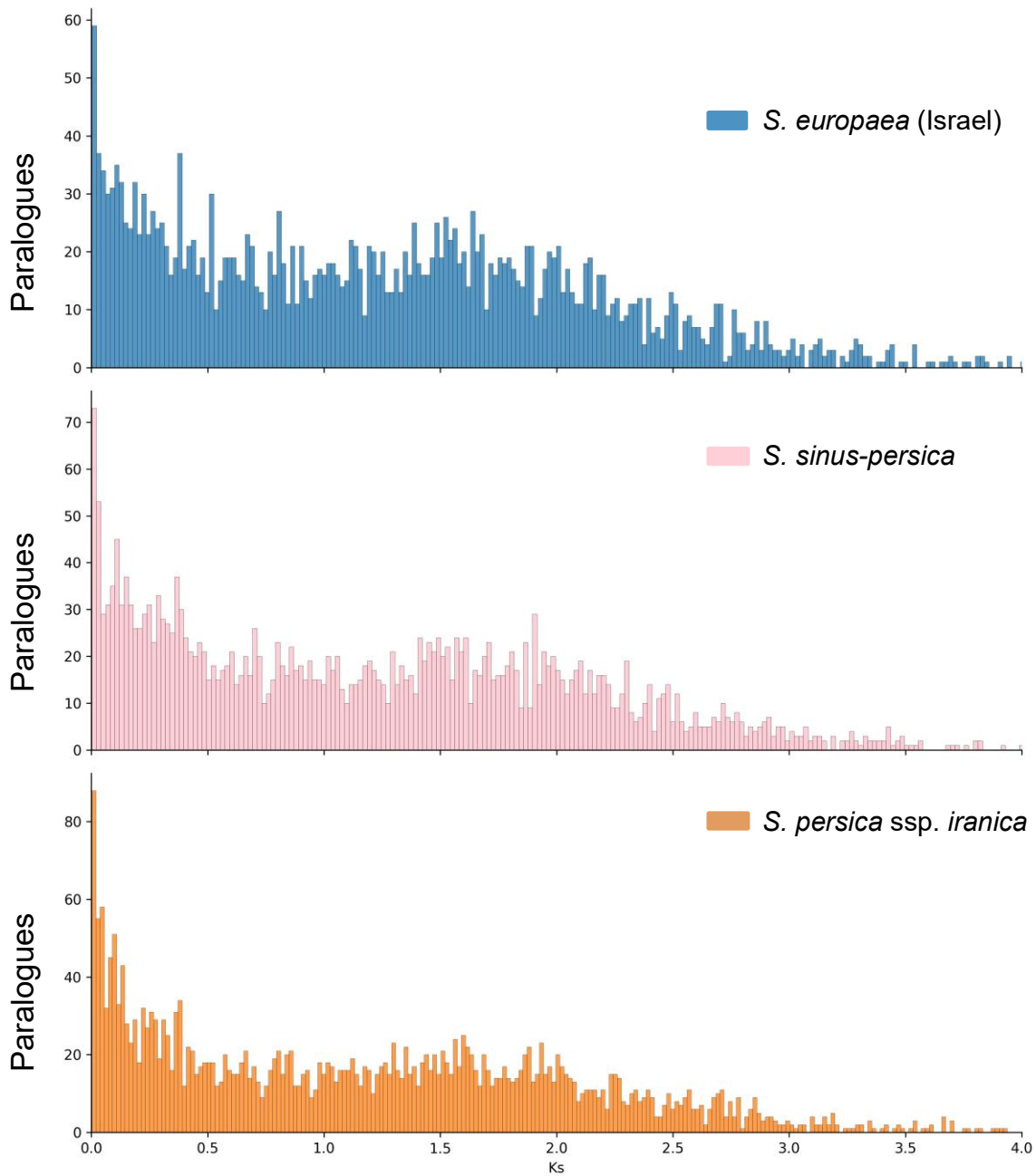
Supplementary Fig. 8. Principal component analysis (PCA) of differential k-mers. Points indicate chromosomes of the subgenome, represented by colours. Results are shown for *Salicornia bigelovii* (subgenomes B and D) and *Salicornia* sp. France (subgenomes A and C).



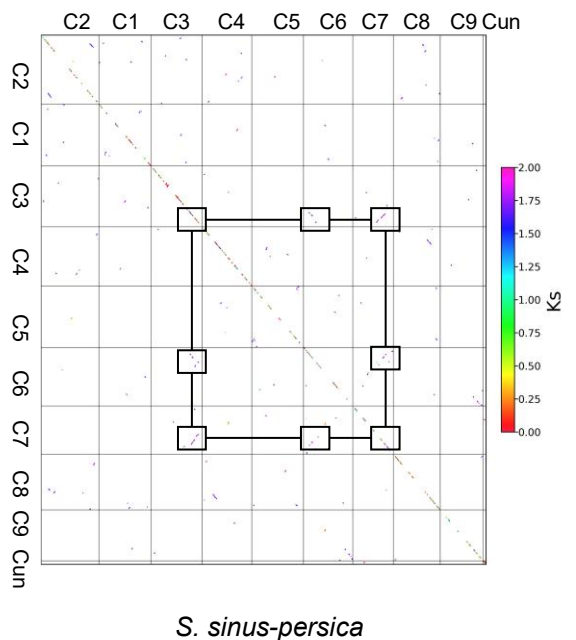
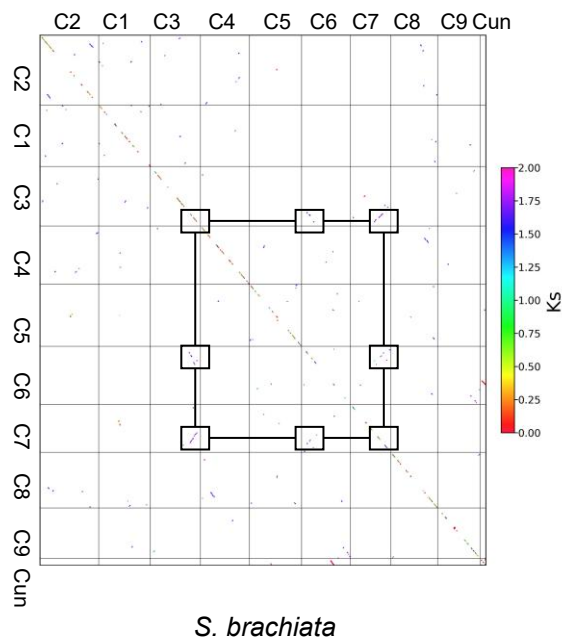
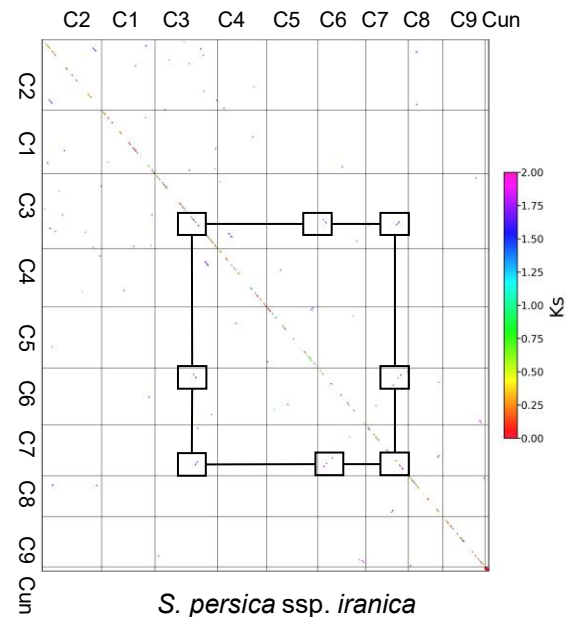
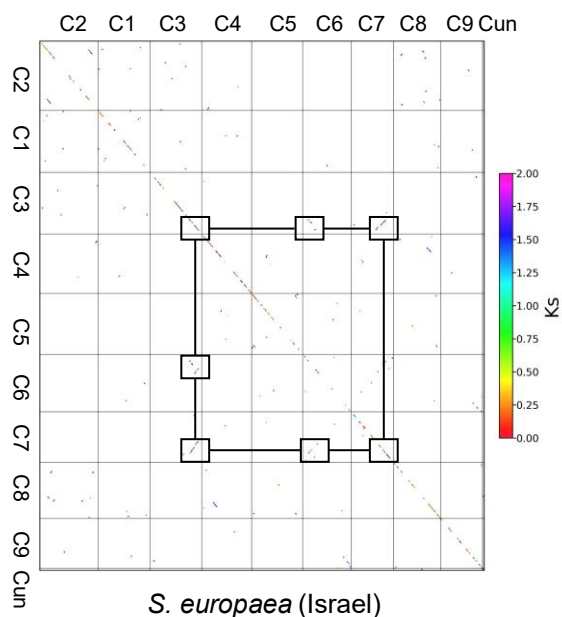
Supplementary Fig. 9. Oligo-FISH mapping of chromosome 9C in *Salicornia* sp. France. Horizontal bars are linear representations of assembled chromosomes with probe-aligned oligonucleotide positions shown as vertical marks; scale bars, 10 Mb. **a, b** Mapping of the Chr9 painting probe (blue) to chromosome 9A (**a**) and chromosome 9C (**b**). **c, d** Mapping of the 45S rDNA probe (green) to chromosome 9A (**c**) and chromosome 9C (**d**). **e** DAPI-stained (white) metaphase spread of *Salicornia* sp. France hybridized with the original Chr9 painting probe (blue), probe C1 (red), marking a region frequently spatially separated from the remainder of Chr9C, probe C2 (magenta), marking a portion of the Chr9C arm, and a 45S rDNA probe (green). Scale bar, 5 μm. Independent hybridizations ranged from one to ten per probe combination and genotype; only one complete metaphase spreads (n=1) with condensed chromosomes was captured despite >10 spread preparations.



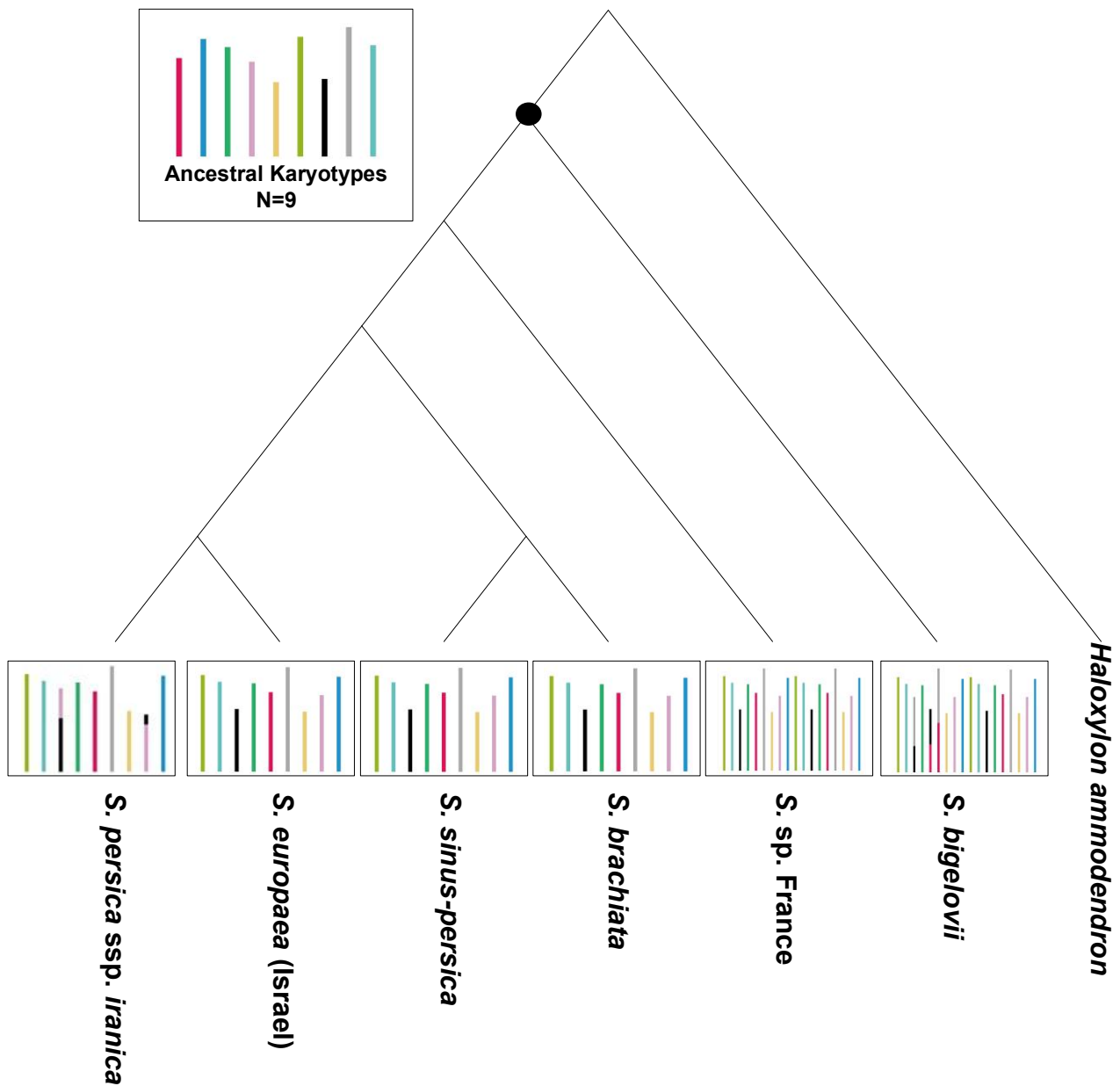
Supplementary Fig. 10. Reciprocal translocations revealed by oligo-FISH on mitotic metaphase chromosomes of *Salicornia* species. **a** Reciprocal translocation between chromosomes three and eight in *S. persica* ssp. *iranica*, detected with oligo-FISH probes SEUR3 (green) and SEUR8 (pink). **b** Reciprocal translocations between chromosomes three and five, chromosomes five and six, and chromosomes six and three in *S. bigelovii*, detected with oligo-painting probes SEUR3 (green), SEUR5 (pink) and SEUR6 (yellow). Translocated chromosomes are indicated by white arrows. Chromosomes were stained with DAPI (blue). Scale bars, 5 μm. Independent hybridizations ranged from one to ten per probe combination and genotype; a minimum of five complete metaphase spreads (n=5) were captured with consistent results.



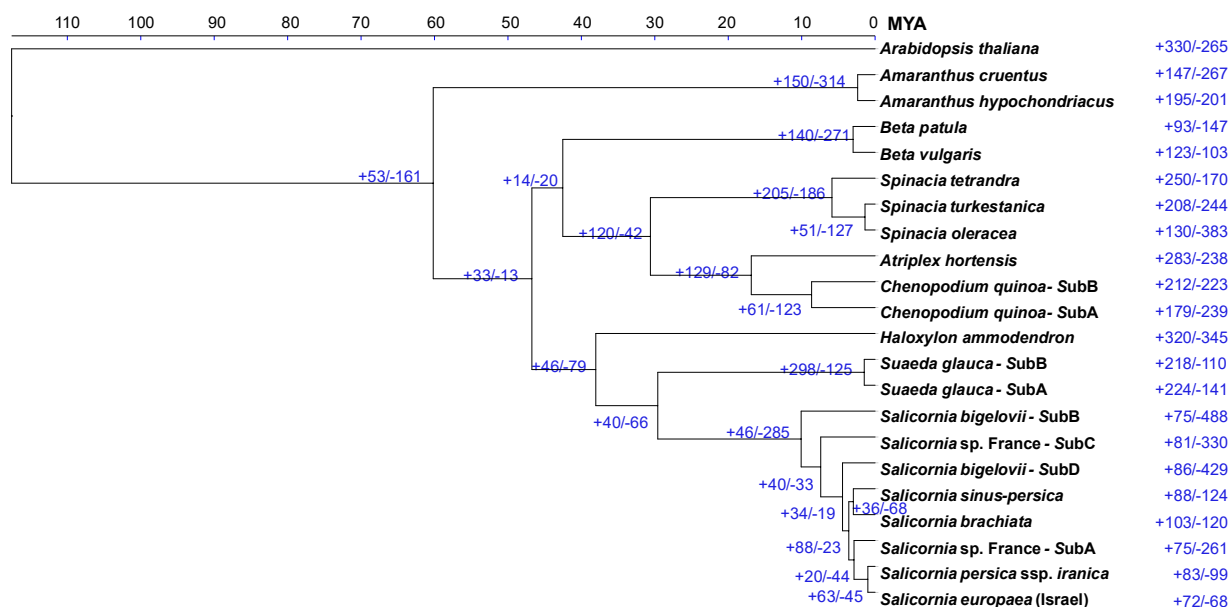
Supplementary Fig. 11. Distribution of synonymous substitutions per synonymous site (K_s). *Salicornia europaea* (Israel) paralogues (sapphire); *Salicornia sinus-persica* paralogues (pink); *Salicornia persica* ssp. *iranica* paralogues (orange).



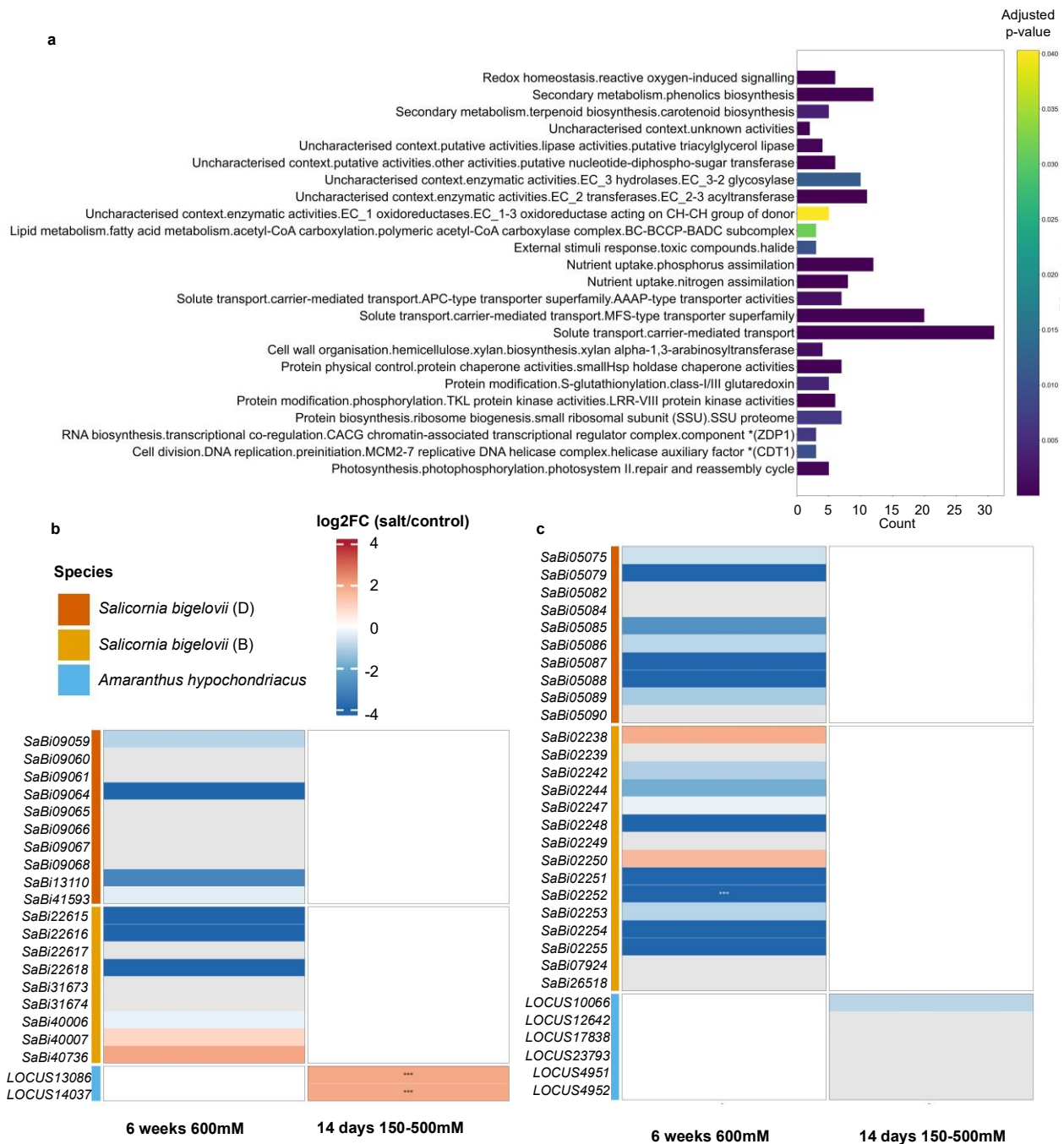
Supplementary Fig. 12. Dot plots of paralogues in four diploid *Salicornia* genomes. The whole-genome triplication event (WGT- γ) is highlighted, with 1:3 relationships shown in black boxes. C denotes chromosome.



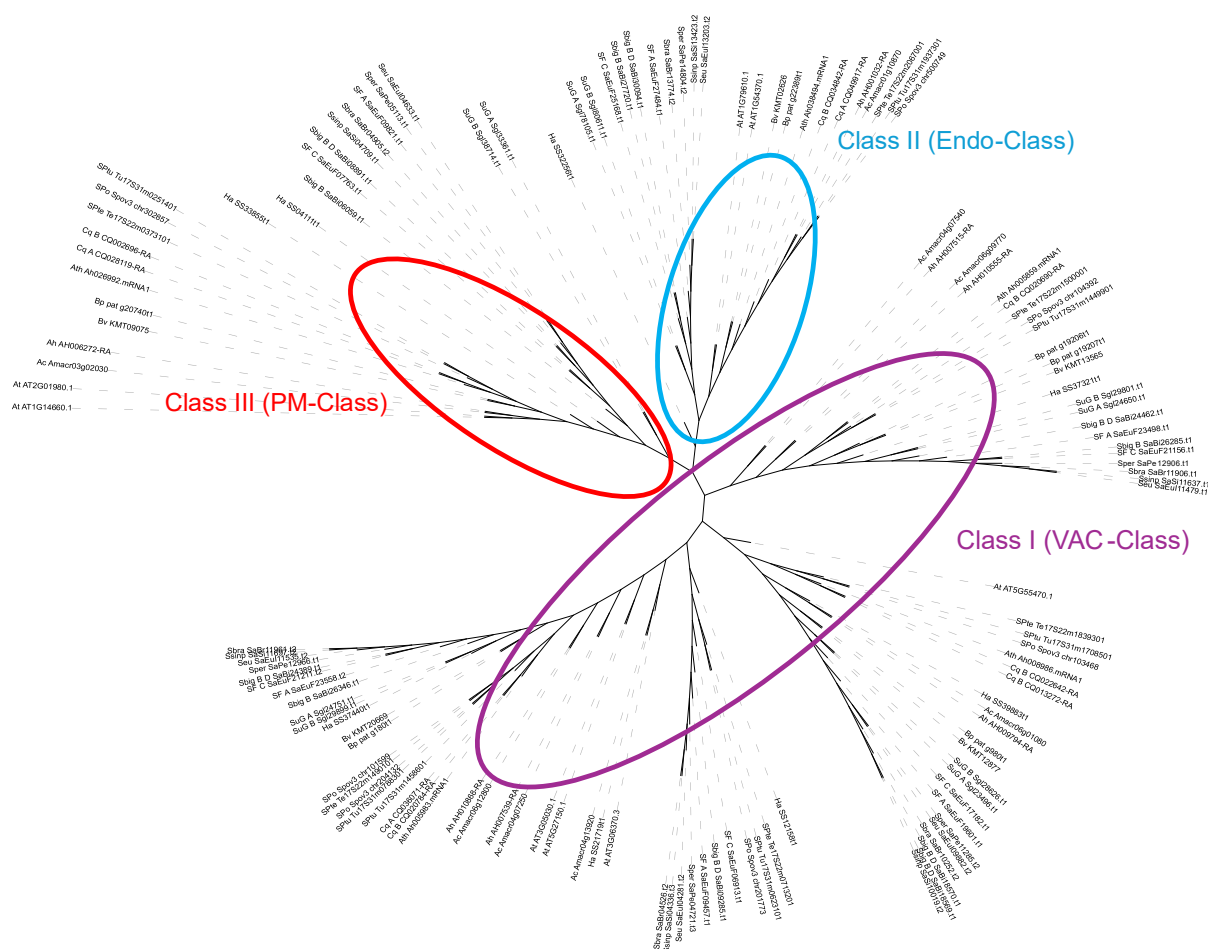
Supplementary Fig. 13. *Salicornia* ancestral karyotype. Genomes of *Salicornia europaea* (Israel), *Salicornia persica* ssp. *iranica*, *Salicornia* sp. France, *Salicornia brachiata*, *Salicornia sinus-persica*, *Salicornia bigelovii* and *Haloxylon ammodendron* were used for ancestral genome reconstruction. Chromosomes of *S. europaea* (Israel), *S. persica* ssp. *iranica*, *Salicornia* sp. France, *S. brachiata*, *S. sinus-persica* and *S. bigelovii* are painted following the colour palette of the *Salicornia* ancestor with nine protochromosomes.



Supplementary Fig. 14. Orthogroup-based phylogeny of the Amaranthaceae family. Phylogenetic tree of multi-species single-copy gene protein sequences based on the maximum likelihood method implemented in RAxML. Numbers on the phylogenetic tree represent gene family expansions (positive values) and contractions (negative values). The numbers adjacent to each tip indicate the number of gene families expanded or contracted in each single species.



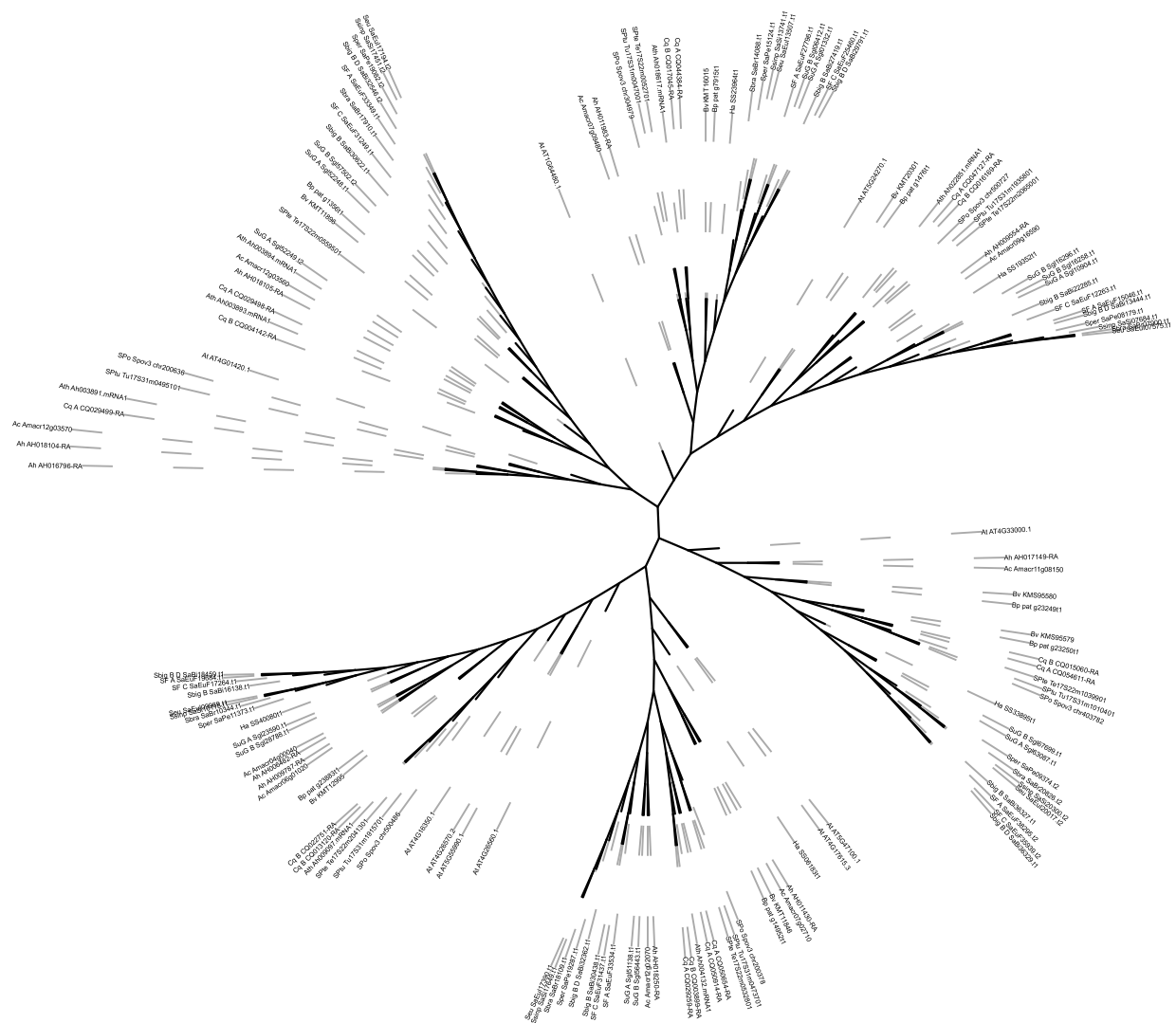
Supplementary Fig. 15. Functional enrichment and cross-species salt response of lineage-expanded gene families in *Salicornia*. **a** MapMan functional-category enrichment analysis of expanded gene families in the *Salicornia* genus. Bar length denotes the number of genes in each category, and bar colour corresponds to the adjusted *P* value. **b, c** Cross-species salt response of two lineage-expanded gene families: log₂ fold change (salt versus control) of chalcone synthase (CHS) orthologues (**b**) and Cupin_1 orthologues (**c**) in *Salicornia bigelovii* (6 weeks, 0 versus 600 mM NaCl; D and B subgenomes shown separately) and *Amaranthus hypochondriacus* (14-day stepped salinity, 0.15 to 0.5 M NaCl); *n* = 3 per condition. Differentially expressed genes (DEGs) were called with DESeq2 (Benjamini–Hochberg-adjusted *P* < 0.05, absolute log₂ fold change > 0.585); three asterisks denote adjusted *P* < 0.001. Log₂ fold change is capped at plus or minus 4. White indicates that the gene is absent in that species; light grey indicates that it is present but filtered for low counts. Log₂ fold change values are within-species and are not directly comparable across species.



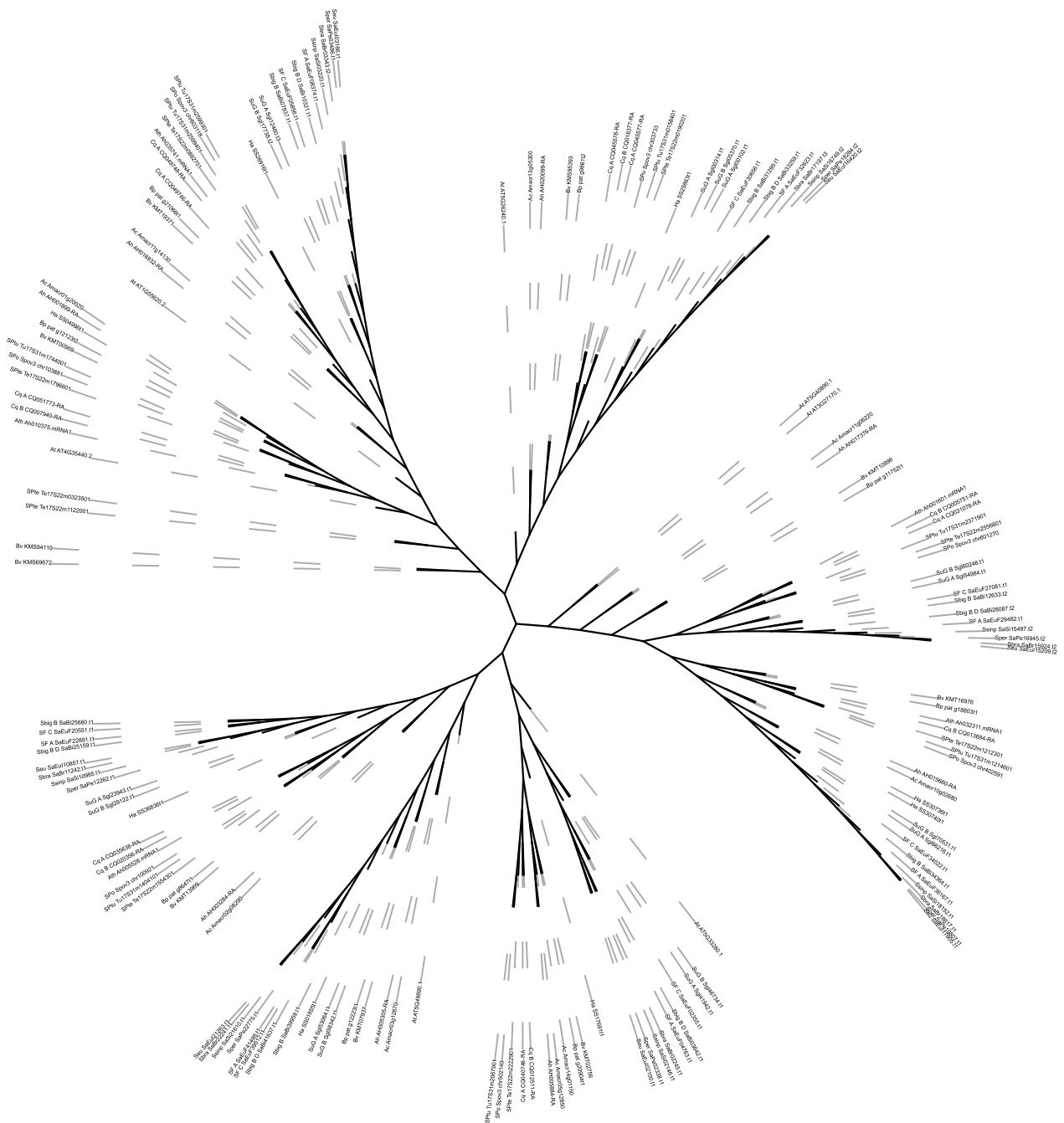
Supplementary Fig. 16. Maximum-likelihood phylogeny of NHX proteins from 18 plant species. Protein sequences were aligned with MUSCLE; the phylogeny was inferred with IQ-TREE using the best-fit substitution model selected by ModelFinder and node support estimated by 1,000 ultrafast bootstrap replicates. The tree is presented unrooted in iTOL with branch lengths omitted for clarity. Species abbreviations used in the tree are as follows: At, *Arabidopsis thaliana*; Ac, *Amaranthus cruentus*; Ah, *Amaranthus hypochondriacus*; Ath, *Atriplex hortensis*; Bp, *Beta patula*; Bv, *Beta vulgaris*; Cq_A and Cq_B, *Chenopodium quinoa* A and B subgenomes; Ha, *Haloxylon ammodendron*; SPo, *Spinacia oleracea*; SPte, *Spinacia tetrandra*; SPtu, *Spinacia turkestanica*; SuG_A and SuG_B, *Suaeda glauca* A and B subgenomes; Sbig_B and Sbig_D, *Salicornia bigelovii* B and D subgenomes; Seu, *Salicornia europaea* (Israel); Sper, *Salicornia persica* ssp. *iranica*; SF_A and SF_C, *Salicornia* sp. France A and C subgenomes; Sbra, *Salicornia brachiata*; Ssinp, *Salicornia sinuspersici*.



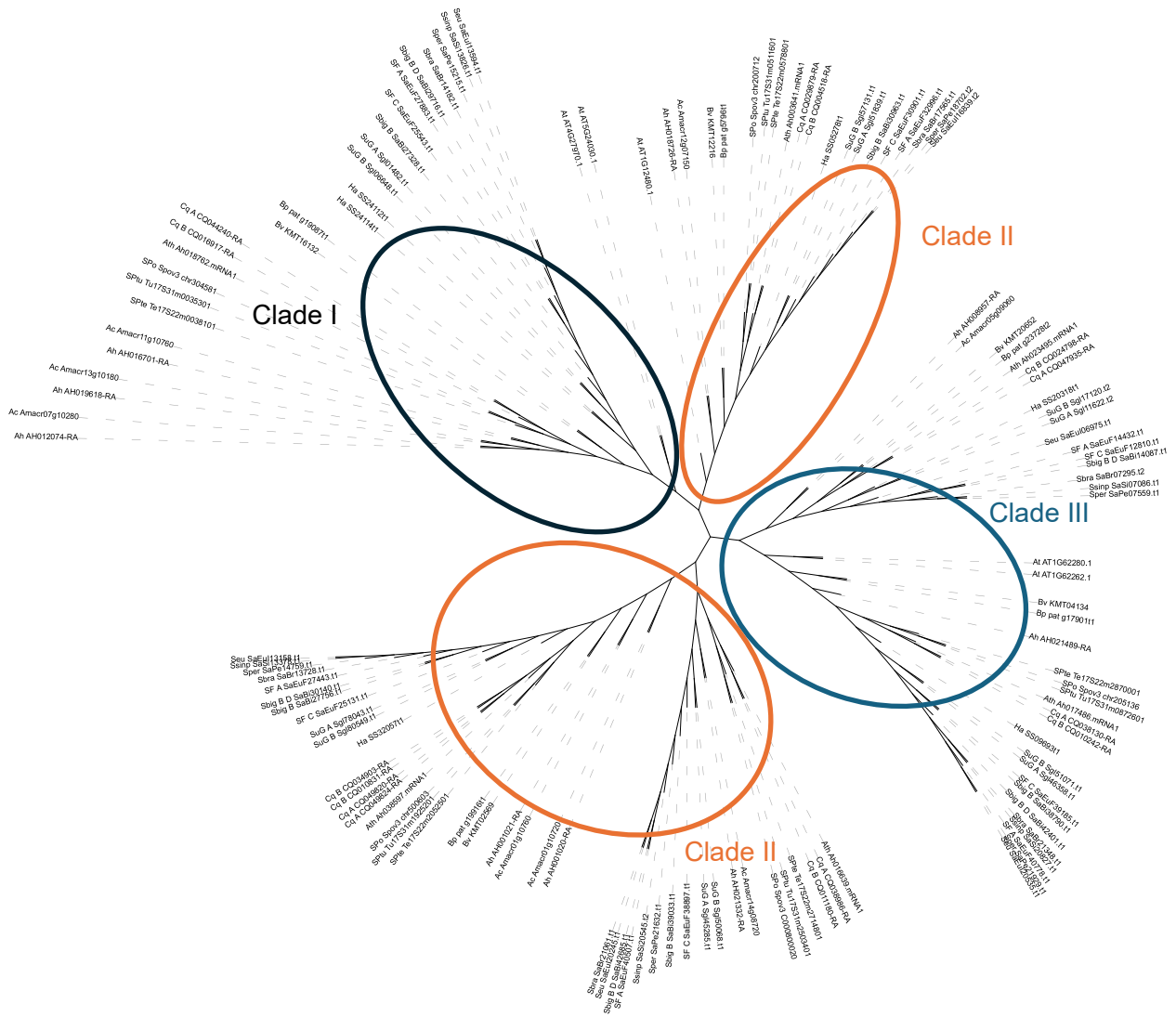
Supplementary Fig. 17. Maximum-likelihood phylogeny of HKT proteins from 18 plant species. Protein sequences were aligned with MUSCLE; the phylogeny was inferred with IQ-TREE using the best-fit substitution model selected by ModelFinder and node support estimated by 1,000 ultrafast bootstrap replicates. The tree is presented unrooted in iTOL with branch lengths omitted for clarity. Species abbreviations used in the tree are as follows: At, *Arabidopsis thaliana*; Ac, *Amaranthus cruentus*; Ah, *Amaranthus hypochondriacus*; Ath, *Atriplex hortensis*; Bp, *Beta patula*; Bv, *Beta vulgaris*; Cq_A and Cq_B, *Chenopodium quinoa* A and B subgenomes; Ha, *Haloxylon ammodendron*; SPo, *Spinacia oleracea*; SPte, *Spinacia tetrandra*; SPtu, *Spinacia turkestanica*; SuG_A and SuG_B, *Suaeda glauca* A and B subgenomes; Sbig_B and Sbig_D, *Salicornia bigelovii* B and D subgenomes; Seu, *Salicornia europaea* (Israel); Sper, *Salicornia persica* ssp. *iranica*; SF_A and SF_C, *Salicornia* sp. France A and C subgenomes; Sbra, *Salicornia brachiata*; Ssinp, *Salicornia sinus-persica*.



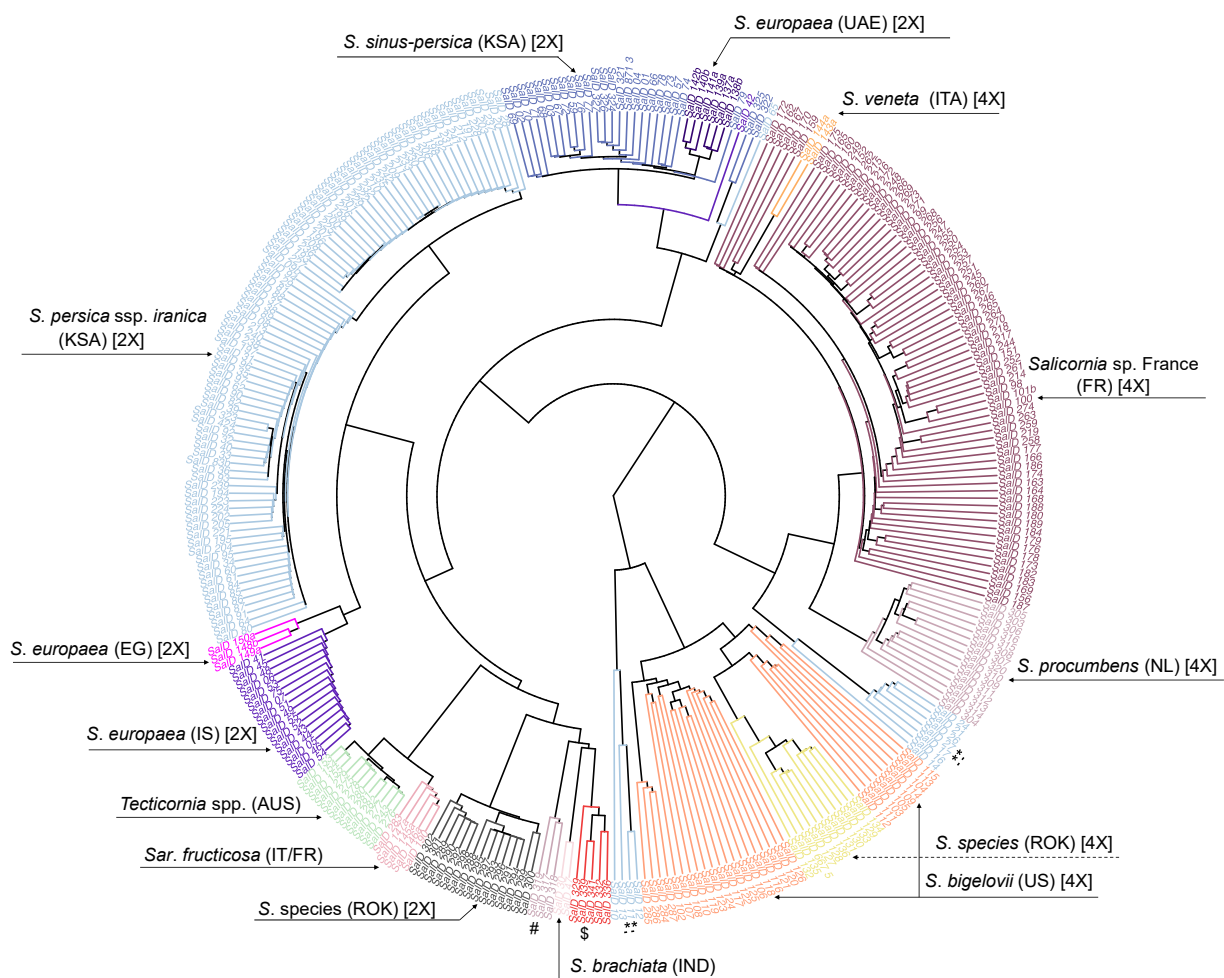
Supplementary Fig. 18. Maximum-likelihood phylogeny of CBL proteins from 18 plant species. Protein sequences were aligned with MUSCLE; the phylogeny was inferred with IQ-TREE using the best-fit substitution model selected by ModelFinder and node support estimated by 1,000 ultrafast bootstrap replicates. The tree is presented unrooted in iTOL with branch lengths omitted for clarity. Species abbreviations used in the tree are as follows: At, *Arabidopsis thaliana*; Ac, *Amaranthus cruentus*; Ah, *Amaranthus hypochondriacus*; Ath, *Atriplex hortensis*; Bp, *Beta patula*; Bv, *Beta vulgaris*; Cq_A and Cq_B, *Chenopodium quinoa* A and B subgenomes; Ha, *Haloxylon ammodendron*; SPo, *Spinacia oleracea*; SPte, *Spinacia tetrandra*; SPtu, *Spinacia turkestanica*; SuG_A and SuG_B, *Suaeda glauca* A and B subgenomes; Sbig_B and Sbig_D, *Salicornia bigelovii* B and D subgenomes; Seu, *Salicornia europaea* (Israel); Sper, *Salicornia persica* ssp. *iranica*; SF_A and SF_C, *Salicornia* sp. France A and C subgenomes; Sbra, *Salicornia brachiata*; Ssinp, *Salicornia sinu-persica*.



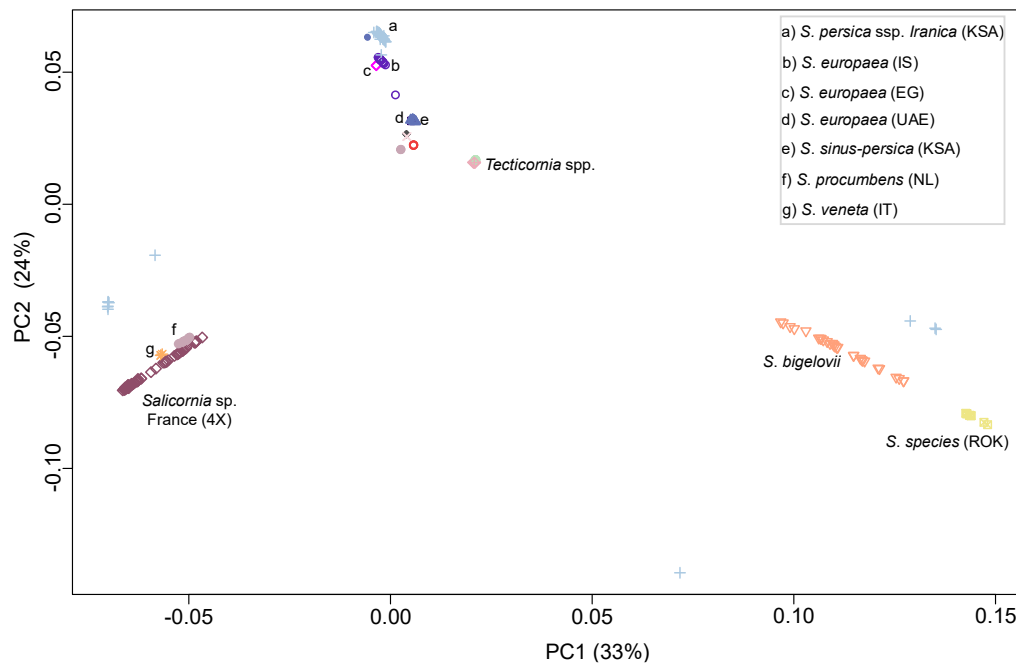
Supplementary Fig. 19. Maximum-likelihood phylogeny of CLC proteins from 18 plant species. Protein sequences were aligned with MUSCLE; the phylogeny was inferred with IQ-TREE using the best-fit substitution model selected by ModelFinder and node support estimated by 1,000 ultrafast bootstrap replicates. The tree is presented unrooted in iTOL with branch lengths omitted for clarity. Species abbreviations used in the tree are as follows: At, *Arabidopsis thaliana*; Ac, *Amaranthus cruentus*; Ah, *Amaranthus hypochondriacus*; Ath, *Atriplex hortensis*; Bp, *Beta patula*; Bv, *Beta vulgaris*; Cq_A and Cq_B, *Chenopodium quinoa* A and B subgenomes; Ha, *Haloxylon ammodendron*; SPo, *Spinacia oleracea*; SPte, *Spinacia tetrandra*; SPtu, *Spinacia turkestanica*; SuG_A and SuG_B, *Suaeda glauca* A and B subgenomes; Sbig_B and Sbig_D, *Salicornia bigelovii* B and D subgenomes; Seu, *Salicornia europaea* (Israel); Sper, *Salicornia persica* ssp. *iranica*; SF_A and SF_C, *Salicornia* sp. France A and C subgenomes; Sbra, *Salicornia brachiata*; Ssinp, *Salicornia sinuspersici*.



Supplementary Fig. 20. Maximum-likelihood phylogeny of SLAC proteins from 18 plant species. Protein sequences were aligned with MUSCLE; the phylogeny was inferred with IQ-TREE using the best-fit substitution model selected by ModelFinder and node support estimated by 1,000 ultrafast bootstrap replicates. The tree is presented unrooted in iTOL with branch lengths omitted for clarity. Species abbreviations used in the tree are as follows: At, *Arabidopsis thaliana*; Ac, *Amaranthus cruentus*; Ah, *Amaranthus hypochondriacus*; Ath, *Atriplex hortensis*; Bp, *Beta patula*; Bv, *Beta vulgaris*; Cq_A and Cq_B, *Chenopodium quinoa* A and B subgenomes; Ha, *Haloxylon ammodendron*; SPo, *Spinacia oleracea*; SPte, *Spinacia tetrandra*; SPtu, *Spinacia turkestanica*; SuG_A and SuG_B, *Suaeda glauca* A and B subgenomes; Sbig_B and Sbig_D, *Salicornia bigelovii* B and D subgenomes; Seu, *Salicornia europaea* (Israel); Sper, *Salicornia persica* ssp. *iranica*; SF_A and SF_C, *Salicornia* sp. France A and C subgenomes; Sbra, *Salicornia brachiata*; Ssinp, *Salicornia sinuipersica*.

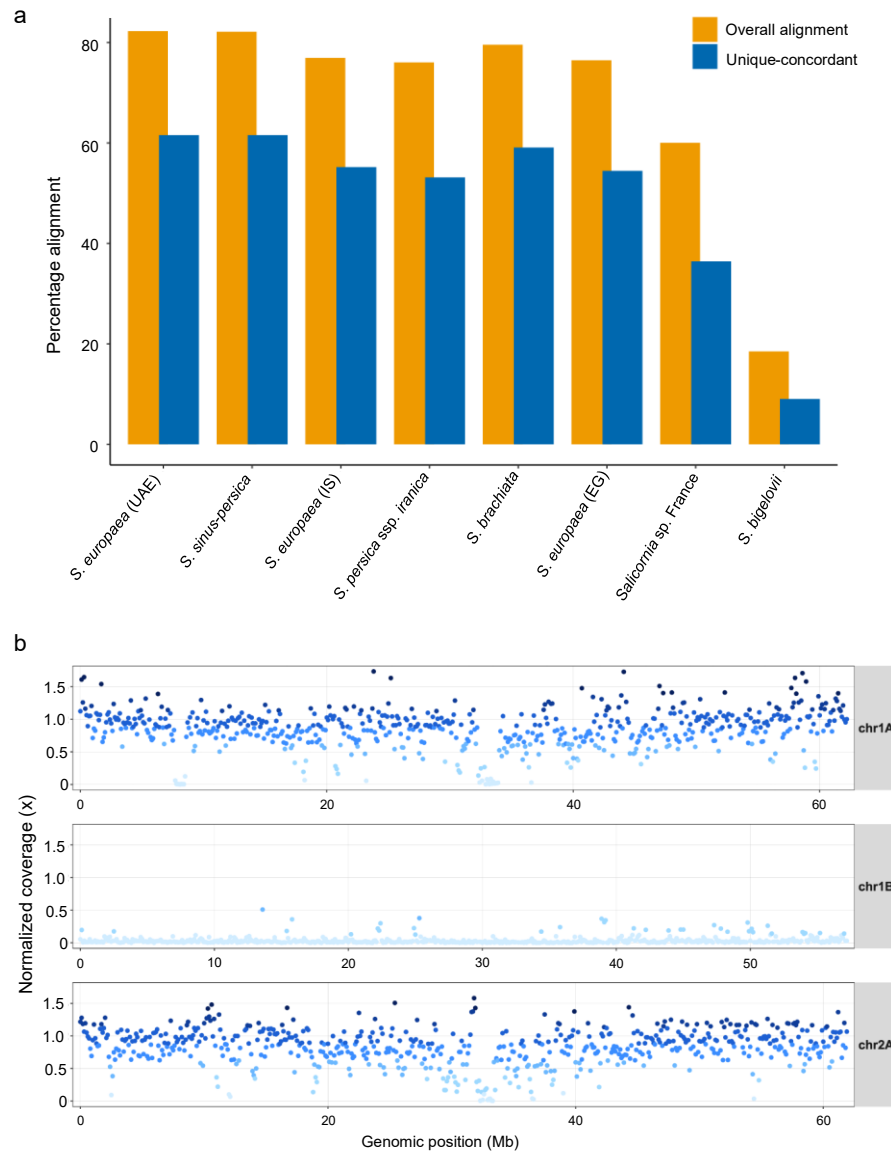


Supplementary Fig. 21. Unrooted neighbour-joining tree of *Salicornia* and related species based on k-mer analysis. Branch tips are labeled with the corresponding SalD_# accession identifiers. Distinct clades and subclades are highlighted and annotated according to species. Accessions marked with () represent putative *S. persica iranica* that were genetically distinct from the rest of the group. Accessions labeled with # represent putative diploid *S. europaea* or *S. procumbens* from the Netherlands, while those marked with (\$) represent putative *S. europaea* from China. ‘Sar.’ refers to the genus *Sarcocornia*. Country codes are AUS, Australia; EG, Egypt; FR, France; IND, India; IS, Israel; IT, Italy; KSA, Kingdom of Saudi Arabia; NL, the Netherlands; ROK, Republic of Korea; UAE, United Arab Emirates; US, United States of America.

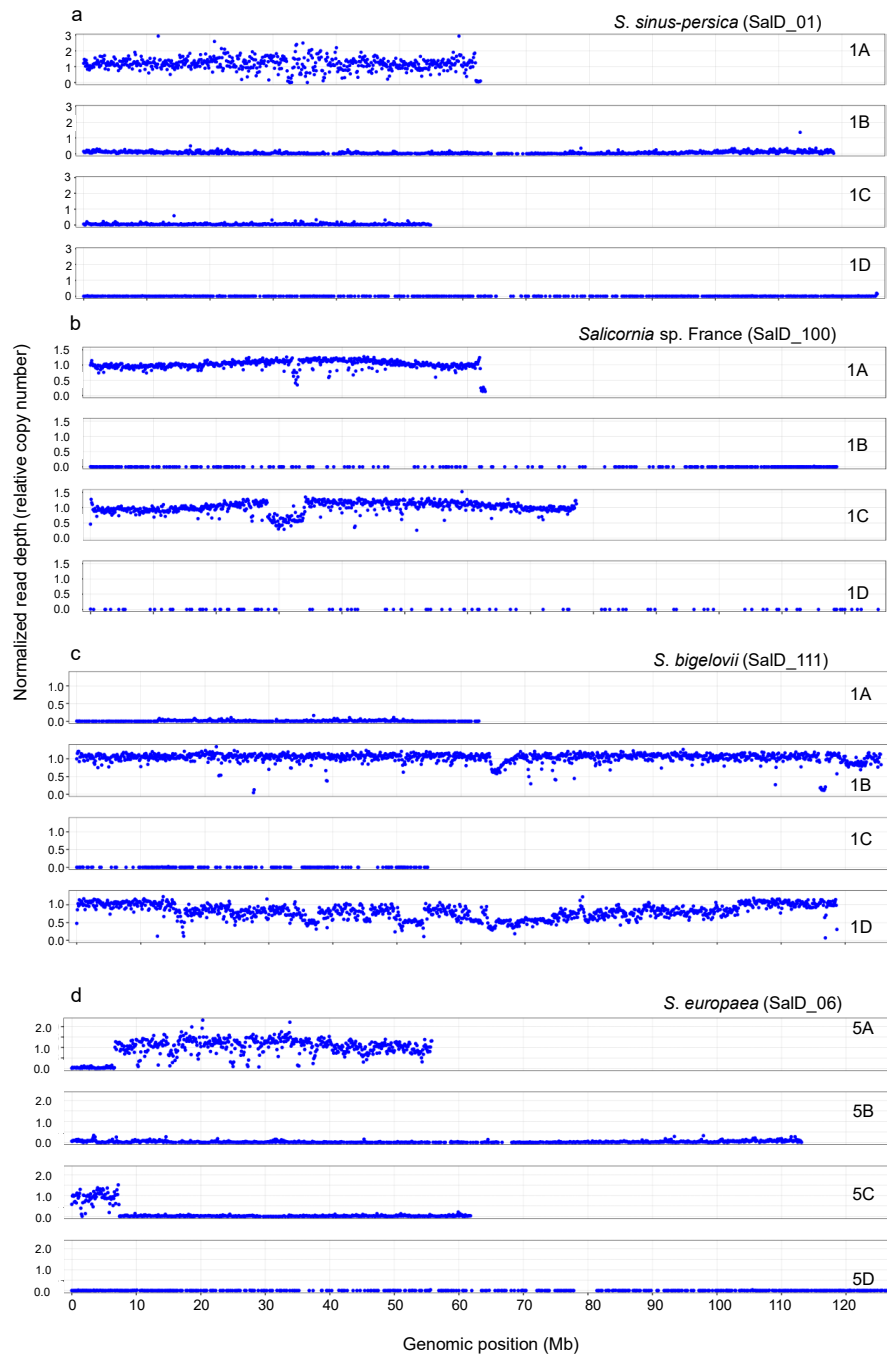


Supplementary Fig. 22. Principal component analysis (PCA) of *Salicornia* species based on k-mer data.

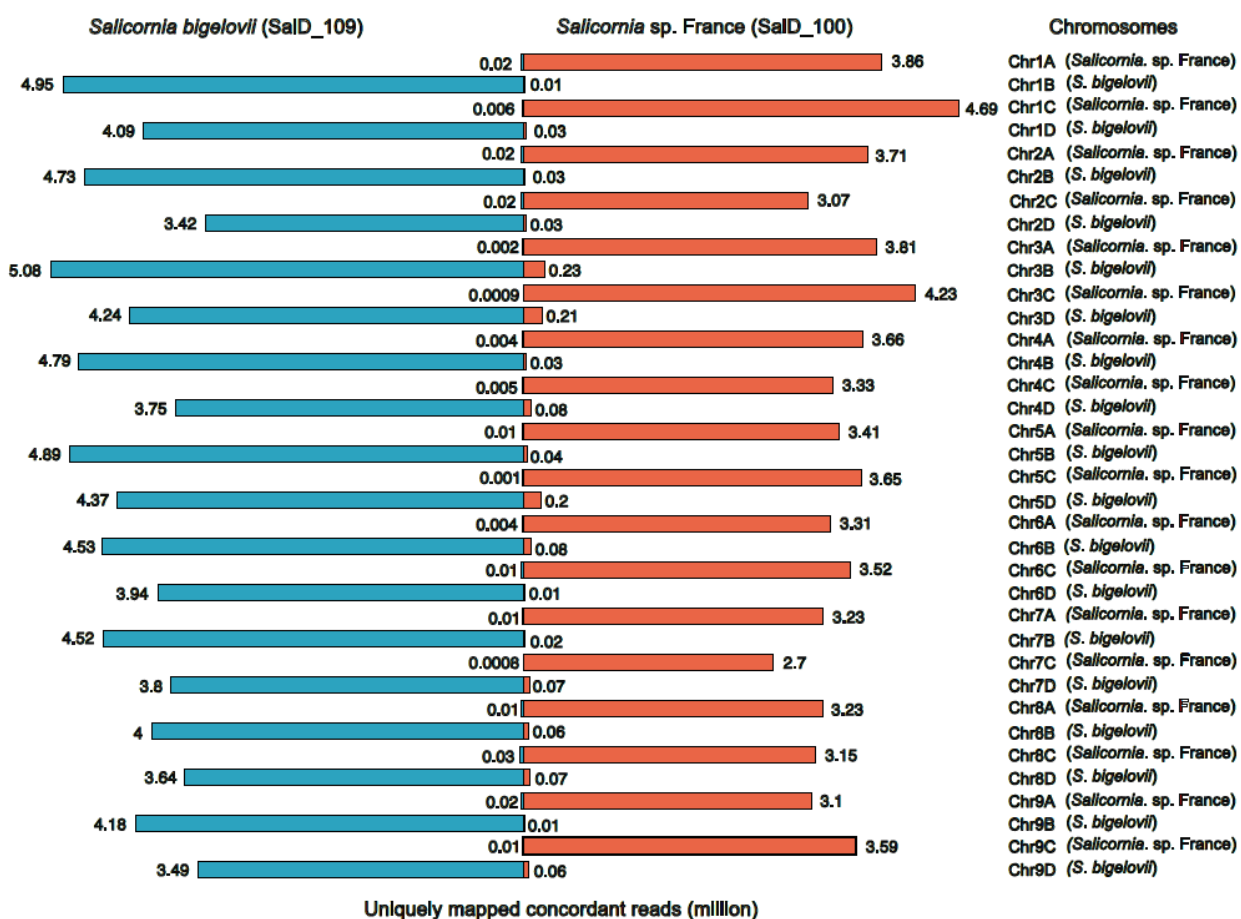
The first two principal components (PC1 and PC2) are shown. Individual accessions are represented by coloured shapes grouped according to species designation and genome composition. Light blue accessions marked with a plus sign represent putative *Salicornia persica* ssp. *iranica* that were genetically distinct from the rest of the group.



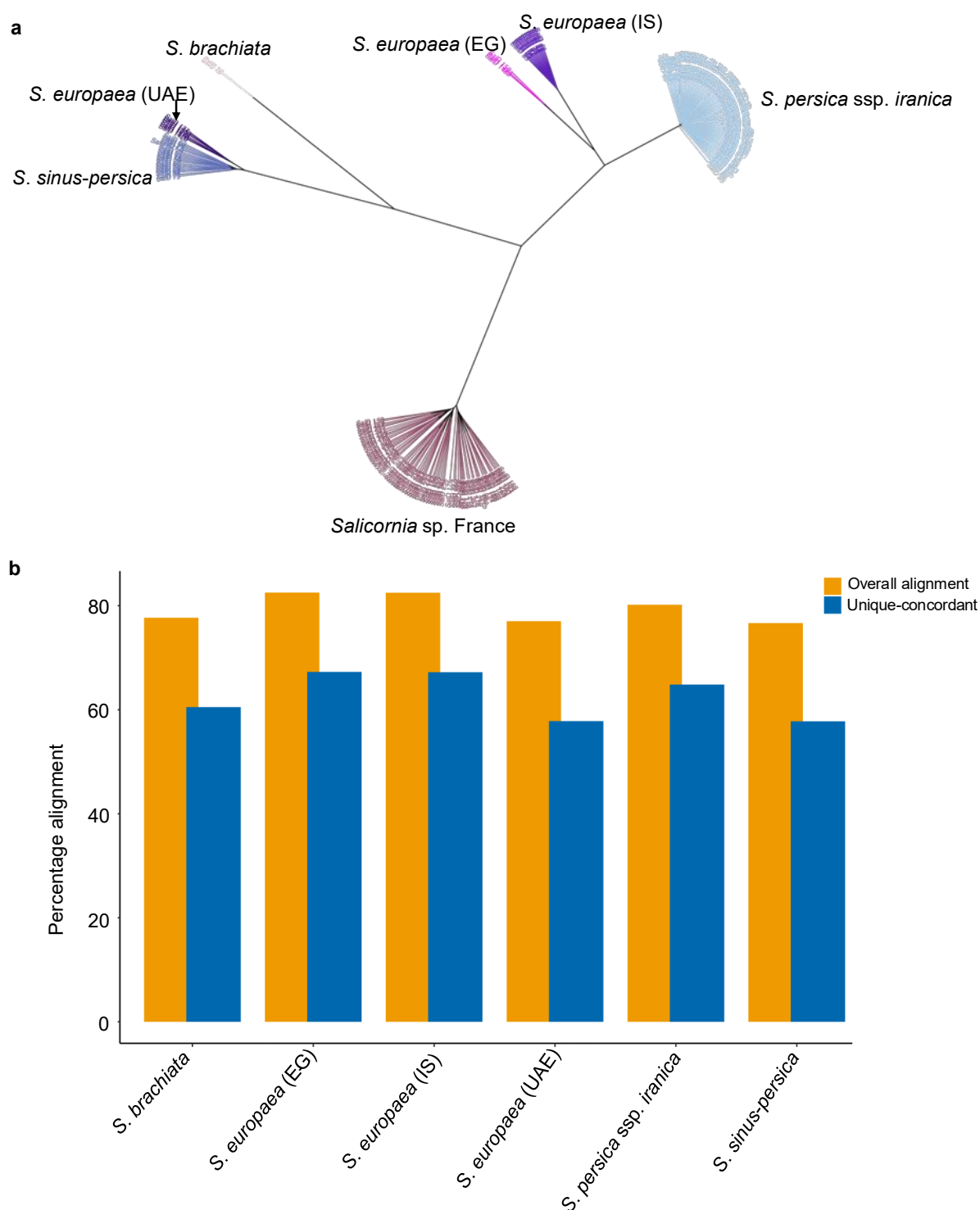
Supplementary Fig. 23. Genomic relationships between *Salicornia* accessions and the tetraploid *S. europaea*¹. **a** Percentage of read alignment (overall and unique-concordant) of different diploid *Salicornia* species and a tetraploid *Salicornia* sp. France to the published tetraploid *S. europaea* T2T reference genome. Values represent average read-mapping statistics across all accessions within each species. **b** Normalized read-depth coverage across three representative chromosomes of the published tetraploid *S. europaea* genome following mapping of the SalD_100 accession (*Salicornia* sp. France), which was also assembled in this study. A normalized coverage of 1 corresponds to two copies of each chromosome. These results indicate that *Salicornia* sp. France shares only one subgenome with the published *S. europaea* reference, whereas *S. bigelovii* shares none. In addition, the B subgenomes of *S. bigelovii* and *S. europaea* are distinct.



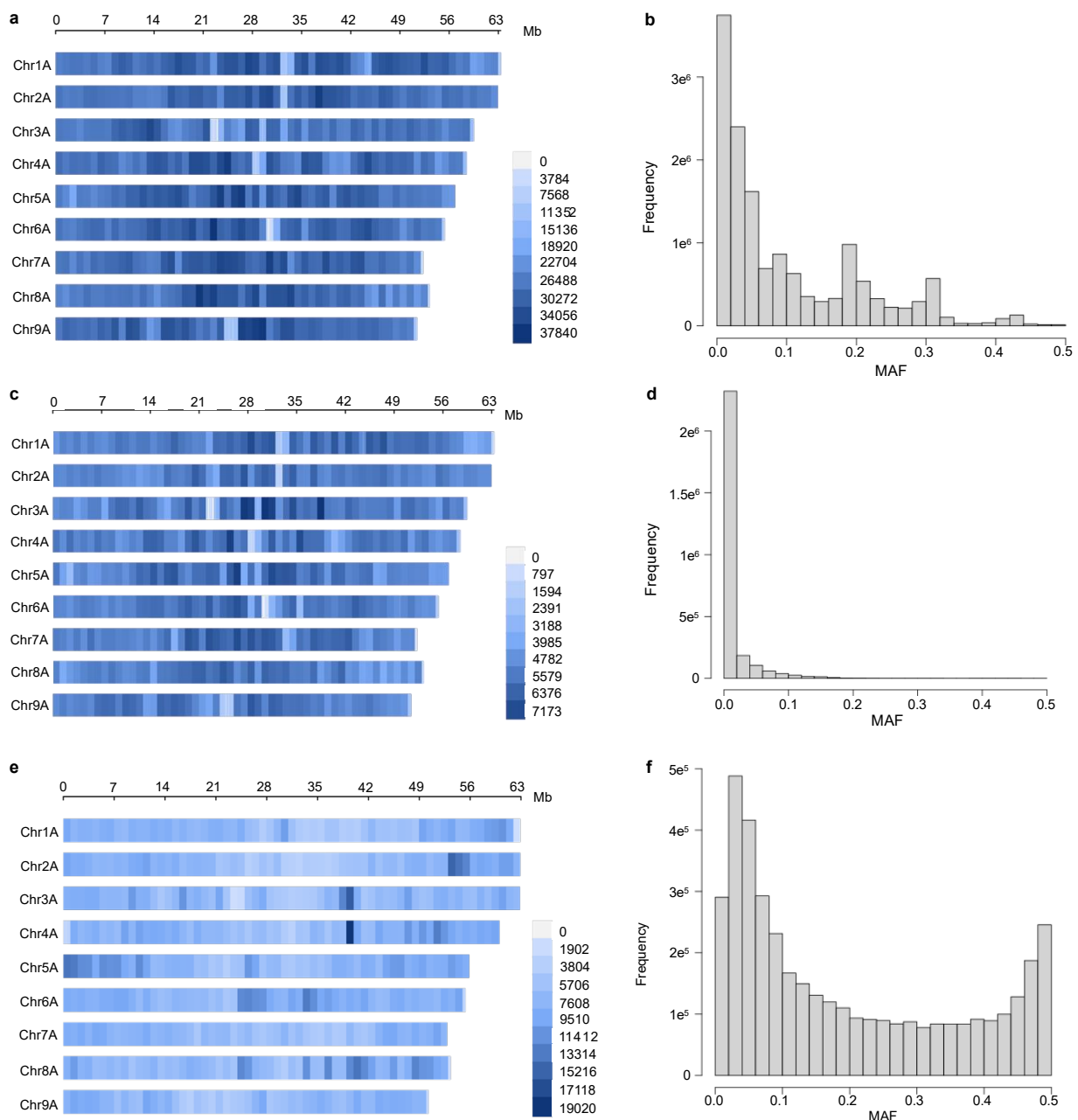
Supplementary Fig. 24. Read-depth coverage of *Salicornia* accessions mapped to a combined tetraploid reference. Chromosome 1 of the different subgenomes and the 5A:5C translocation in diploid *S. europaea* are shown. A normalized read count of 1 indicates two normal copies of the chromosome; 0.5 indicates a single copy; 0 indicates a nullisomic condition; and 2 indicates four normal copies. **a** SalD_01, a diploid *S. sinus-persica* from the Kingdom of Saudi Arabia (KSA). **b** SalD_100, a tetraploid *Salicornia* sp. France. **c** SalD_111, a tetraploid *S. bigelovii* from the USA. **d** SalD_06, *S. europaea* from Israel with a translocated segment on the short arms of 5A and 5C.



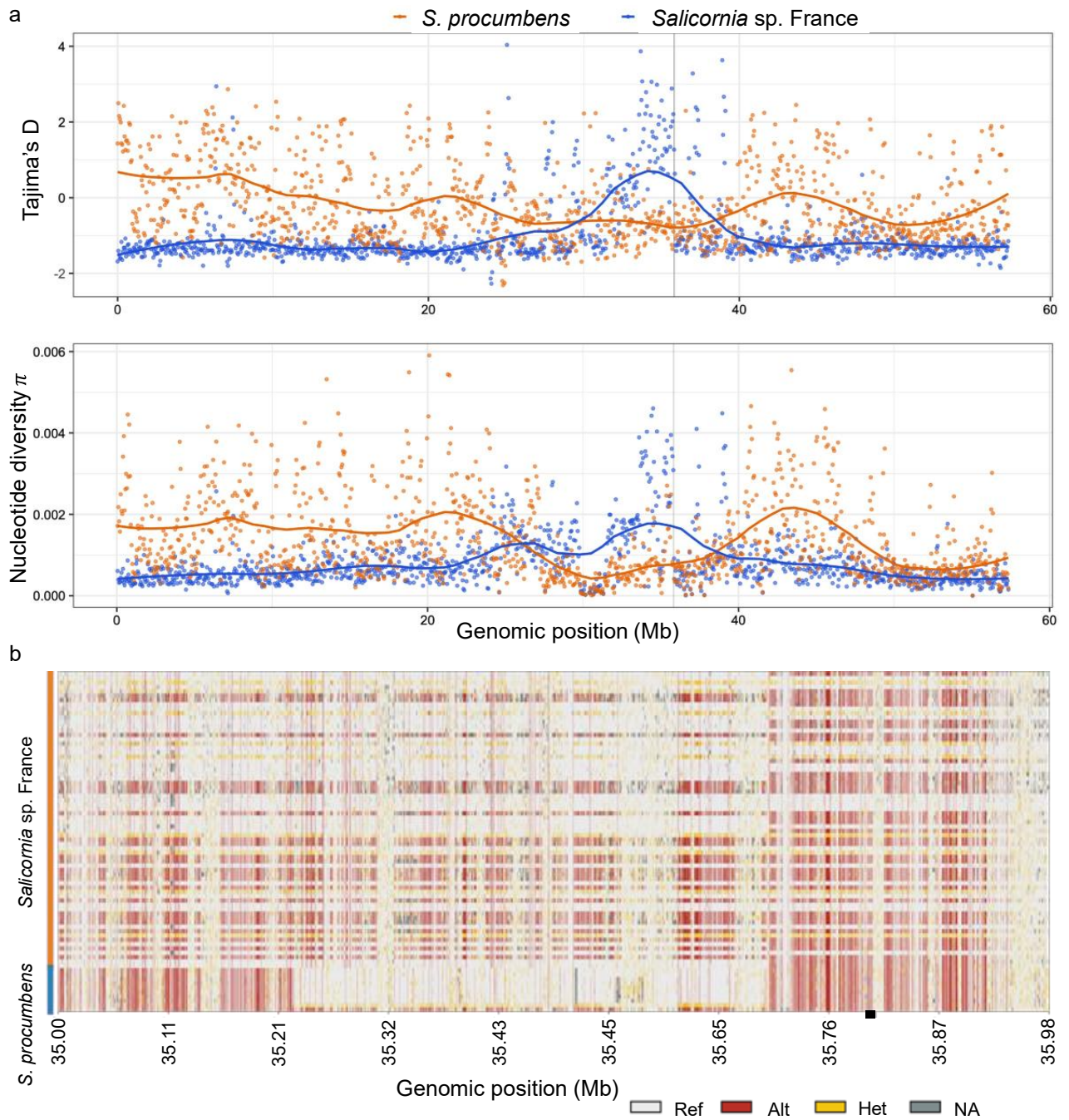
Supplementary Fig. 25. Subgenome-specific alignment patterns and mapping fidelity across *Salicornia* accessions. Read-mapping analysis of representative *Salicornia* accessions against an in silico octaploid reference assembly comprising combined subgenomes of *Salicornia* sp. France (A and C) and *Salicornia bigelovii* (B and D). Bar plots illustrate the distribution of uniquely mapped concordant reads (in millions) across chromosomes Chr1 to Chr9 for each of the four subgenomes (A, B, C and D). Blue bars show alignment metrics for *S. bigelovii* (SalD_109), demonstrating high specificity to the B and D target subgenomes. Orange bars show alignment metrics for *Salicornia* sp. France (SalD_100), showing high specificity to the A and C target subgenomes. Numerical labels indicate the precise read counts (in millions) mapping to each chromosomal locus. Cross-subgenome and homeologous mapping to non-target subgenomes remains negligible ($\leq 2\%$), validating the high stringency of subgenome assignment and the robustness of the downstream genomic analysis.



Supplementary Fig. 26. Relationships between diploid *Salicornia* species and the A subgenome of *Salicornia* sp. France. **a** Phylogenetic tree constructed using 1.4 million filtered single-nucleotide polymorphisms (SNPs) called on the A subgenome of tetraploid *Salicornia* sp. France. **b** Percentage of read alignment of different diploid *Salicornia* species to the A subgenome of tetraploid *Salicornia* sp. France.

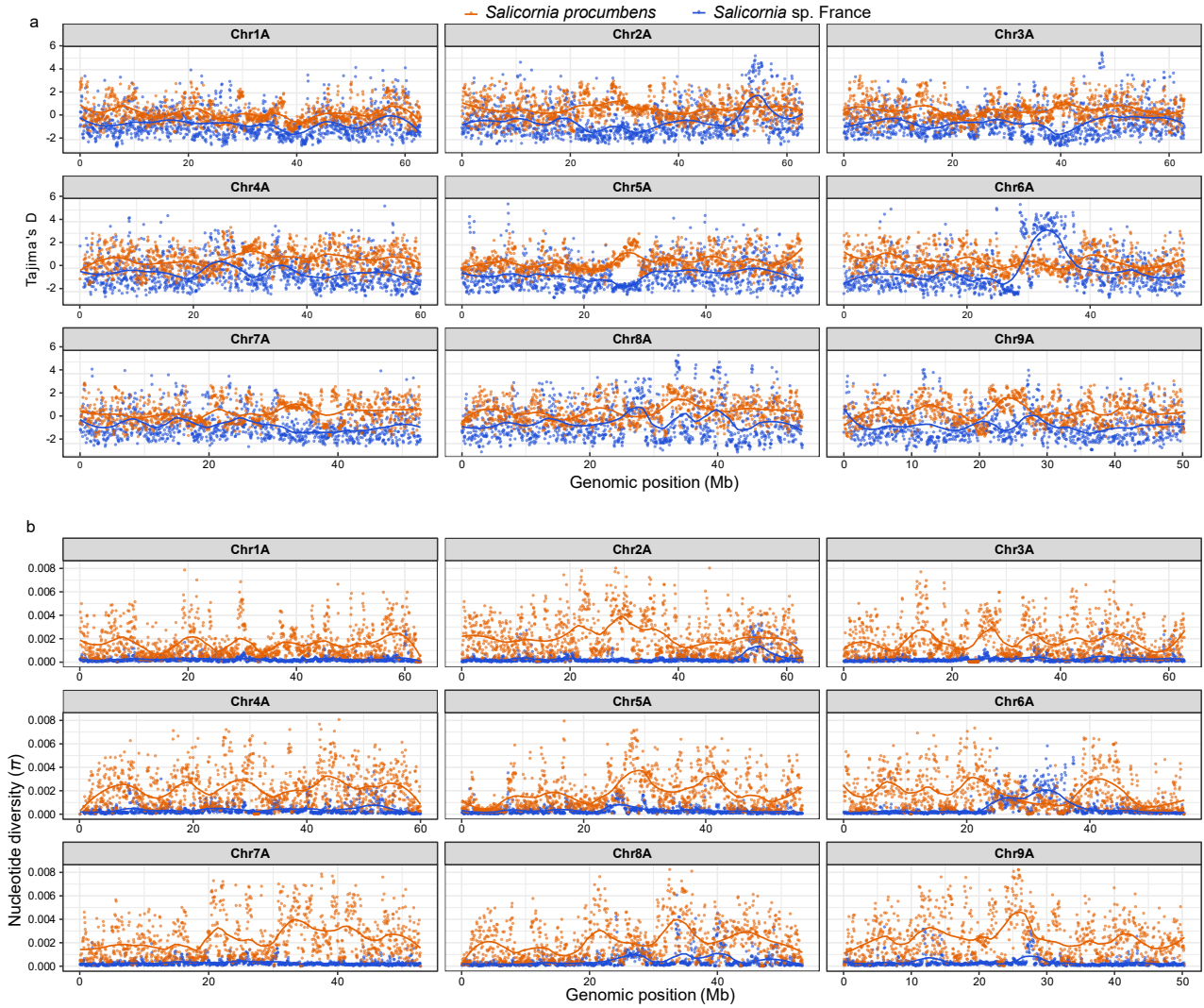


Supplementary Fig. 27. Genome-wide SNP density and minor allele frequency (MAF) in *Salicornia* species. **a** Genome-wide SNP density calculated from 14,535,034 biallelic SNPs identified in all diploid *Salicornia* species in the panel, using 1-Mb non-overlapping bins. **b** Distribution of MAF for the same set of diploid *Salicornia* species based on the 14.5 million SNPs. **c** Genome-wide SNP density calculated from 2,785,227 biallelic SNPs identified in diploid *S. persica* ssp. *iranica*, the largest diploid group in the resequencing panel, using 1-Mb bins. **d** Distribution of MAF for the 2.7 million SNPs identified in *S. persica* ssp. *iranica*. **e** Genome-wide SNP density calculated from 4,014,724 biallelic SNPs identified on the A-subgenome chromosomes of tetraploid *Salicornia* sp. France, using 1-Mb bins. **f** Distribution of MAF for the 4.0 million SNPs identified across the nine A-subgenome chromosomes.

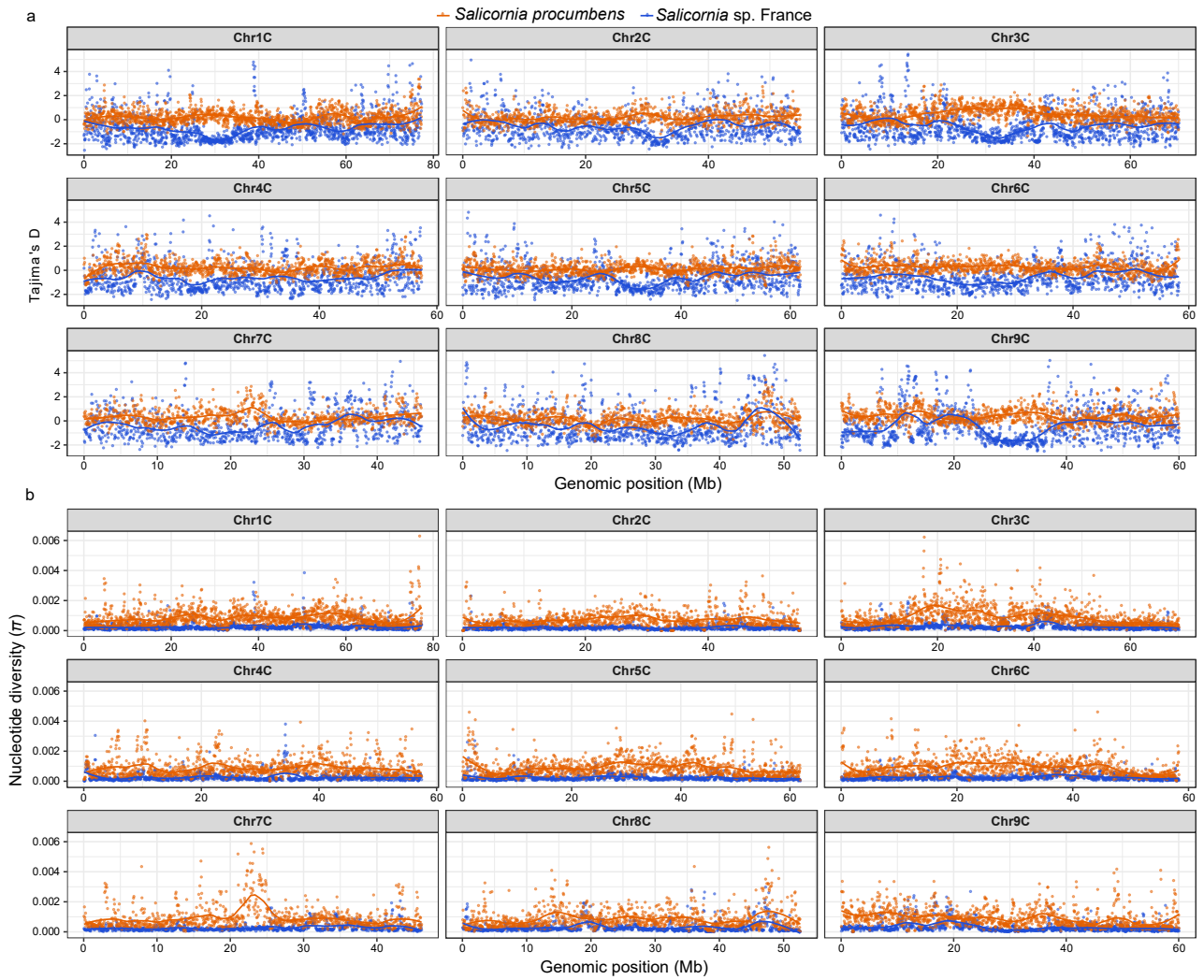


Supplementary Fig. 28. Evolutionary and structural divergence at an adaptive genomic hotspot on Chr6A. **a** Genome-wide comparison of Tajima's D and nucleotide diversity (π) profiles between *Salicornia* sp. France ($n = 71$) and *S. procumbens* ($n = 12$). Statistics were calculated in non-overlapping 100-kb windows along Chr6A using 850,414 SNPs called against the independent diploid *S. ramosissima* reference genome. Before variant calling, raw reads were quality-trimmed (Phred score ≥ 15) and aligned to the reference genome, and only uniquely mapped concordant read pairs were retained. Variants were filtered using bcftools (quality ≥ 30 , INFO/DP ≥ 20 , allele frequency $0.01 < AF < 0.99$), followed by filtering for MAF ≥ 0.01 , missing data $\leq 20\%$, and heterozygosity $\leq 20\%$. Blue and orange dots represent window-based estimates for *Salicornia* sp. France and *S. procumbens*, respectively, with solid lines indicating locally estimated scatterplot smoothing (LOESS) trends. The vertical grey guideline marks the candidate region (approximately 35 Mb) containing

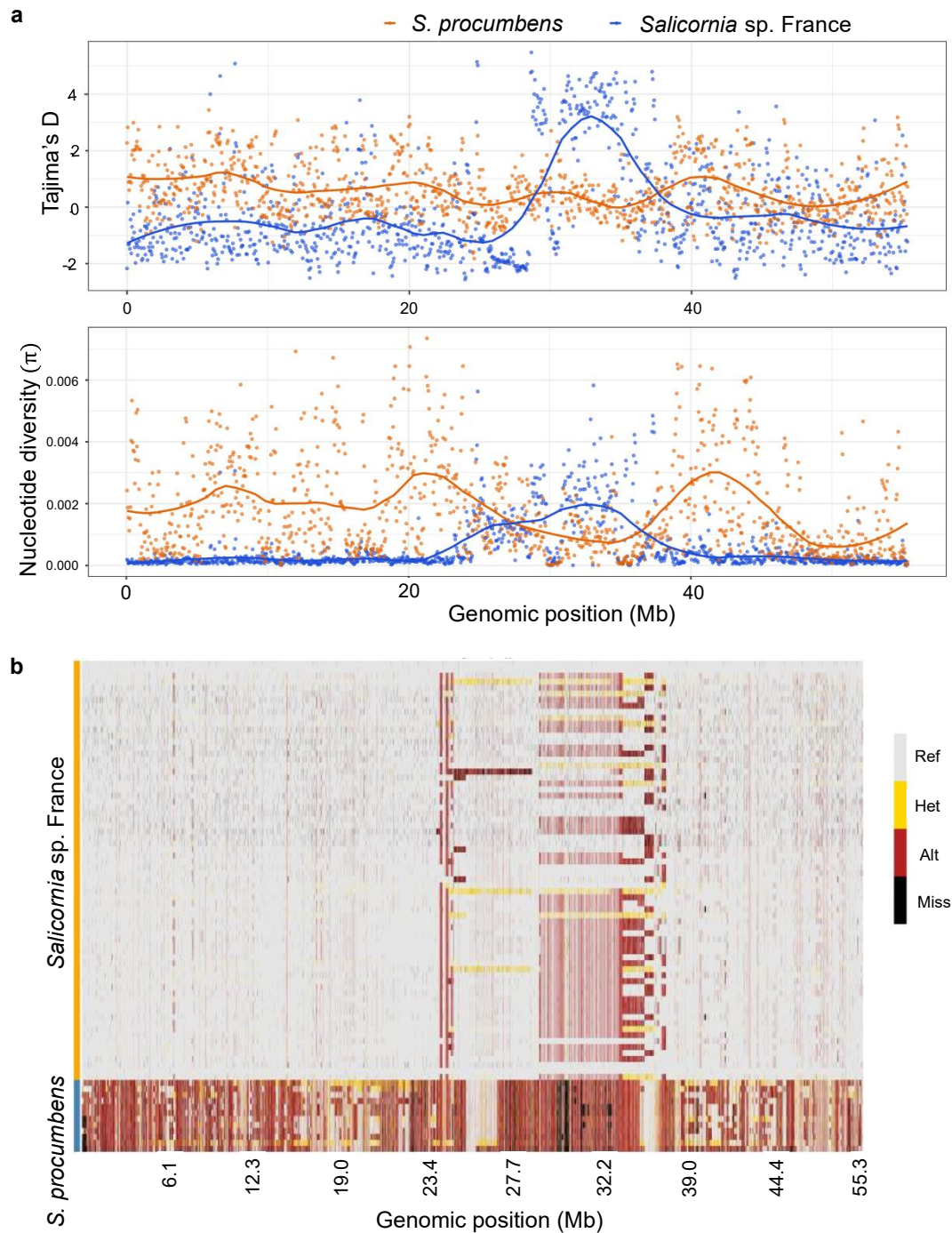
orthologues of the *Arabidopsis* hyperosmolality-gated calcium-permeable channel (OSCA) gene family locus (CSC1-like protein, *At4g02900*). *Salicornia* sp. France accessions originated from the highly fluctuating, hypersaline Étang de Thau, France, whereas *S. procumbens* accessions originated from the daily tidal mudflats of the Wadden Sea, the Netherlands. **b** High-resolution haplotype block structure of the candidate region spanning 35.0–35.9 Mb on chromosome Chr6A of the *S. ramosissima* reference genome. Rows represent individual accessions grouped by population (orange, *Salicornia* sp. France; blue, *S. procumbens*), and columns represent polymorphic sites called against the independent diploid *S. ramosissima* reference genome after applying the same filtering criteria described above. Within the OSCA region, the final filtered dataset exhibited mean per-sample missingness of 4.77% and 1.58%, and mean per-sample observed heterozygosity of 9.16% and 17.00%, for *Salicornia* sp. France and *S. procumbens*, respectively. Genotypes are colour-coded as homozygous reference (Ref, light grey), homozygous alternative (Alt, dark red), heterozygous (Het, gold) and missing data or deletion (NA, dark grey). The horizontal black bar indicates the physical positions of the target OSCA exons.



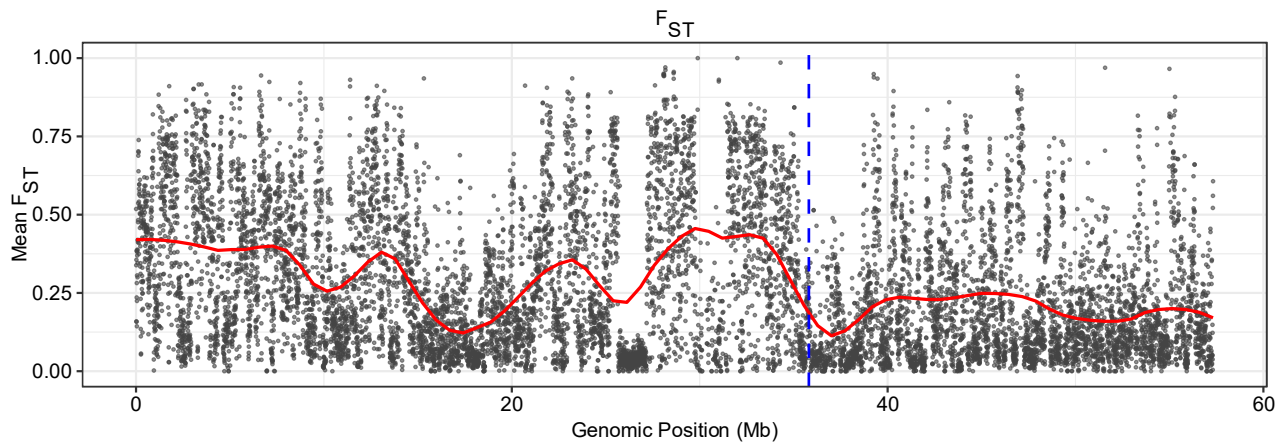
Supplementary Fig. 29. Tajima's D and nucleotide diversity in *S. procumbens* and *Salicornia* sp. France (A subgenome). **a** Chromosome-wide distribution of Tajima's D estimates for *Salicornia* sp. France (blue; $n = 71$) and *S. procumbens* (orange; $n = 12$). Variants were mapped to the chromosome-scale *Salicornia* sp. France reference genome and calculated in non-overlapping 50-kb windows across chromosomes Chr1A–Chr9A. Dots represent individual window estimates, and solid lines indicate locally estimated scatterplot smoothing (LOESS) trends highlighting genome-wide patterns. **b** Genome-wide nucleotide diversity (π) estimated across the same 50-kb windows. Individual window estimates and LOESS-smoothed trend lines are shown for *Salicornia* sp. France (blue) and *S. procumbens* (orange). For these analyses, variants were called against the *Salicornia* sp. France reference genome and filtered using bcftools (quality ≥ 30 , INFO/DP ≥ 20 , allele frequency $0.01 < AF < 0.99$), followed by filtering for minor allele frequency (MAF) ≥ 0.01 , missing data $\leq 20\%$, and heterozygosity $\leq 20\%$, resulting in 8,173,538 high-quality SNPs across all 18 chromosomes (Chr1A–Chr9A and Chr1C–Chr9C). Genome-wide statistics were calculated separately for the two populations using these filtered variants.



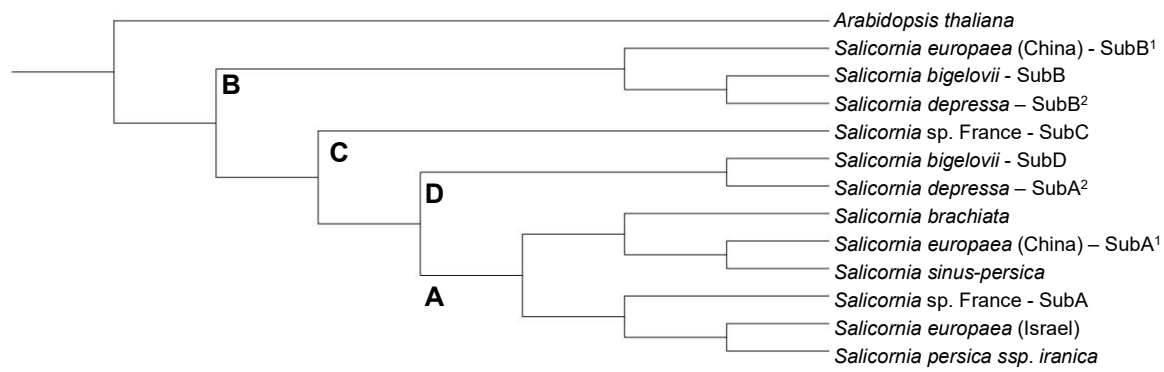
Supplementary Fig. 30. Tajima's D and nucleotide diversity in *S. procumbens* and *Salicornia* sp. France (C subgenome). **a** Chromosome-wide distribution of Tajima's D estimates for *Salicornia* sp. France (blue; $n = 71$) and *S. procumbens* (orange; $n = 12$). Variants were mapped to the chromosome-scale *Salicornia* sp. France reference genome and Tajima's D was calculated in non-overlapping 50-kb windows across chromosomes Chr1C–Chr9C. Dots represent individual window estimates, and solid lines indicate LOESS-smoothed trends highlighting genome-wide patterns. **b** Genome-wide nucleotide diversity (π) estimated across the same 50-kb windows along chromosomes Chr1C–Chr9C. Individual window estimates and LOESS-smoothed trend lines are shown for *Salicornia* sp. France (blue) and *S. procumbens* (orange). Variants were filtered using bcftools (quality ≥ 30 , INFO/DP ≥ 20 , allele frequency $0.01 < AF < 0.99$), followed by filtering for minor allele frequency (MAF) ≥ 0.01 , missing data $\leq 20\%$, and heterozygosity $\leq 20\%$, resulting in 8,173,538 high-quality SNPs across all 18 chromosomes (Chr1A–Chr9A and Chr1C–Chr9C). Genome-wide statistics were calculated separately for each population using the filtered variants mapped to the *Salicornia* sp. France reference genome.



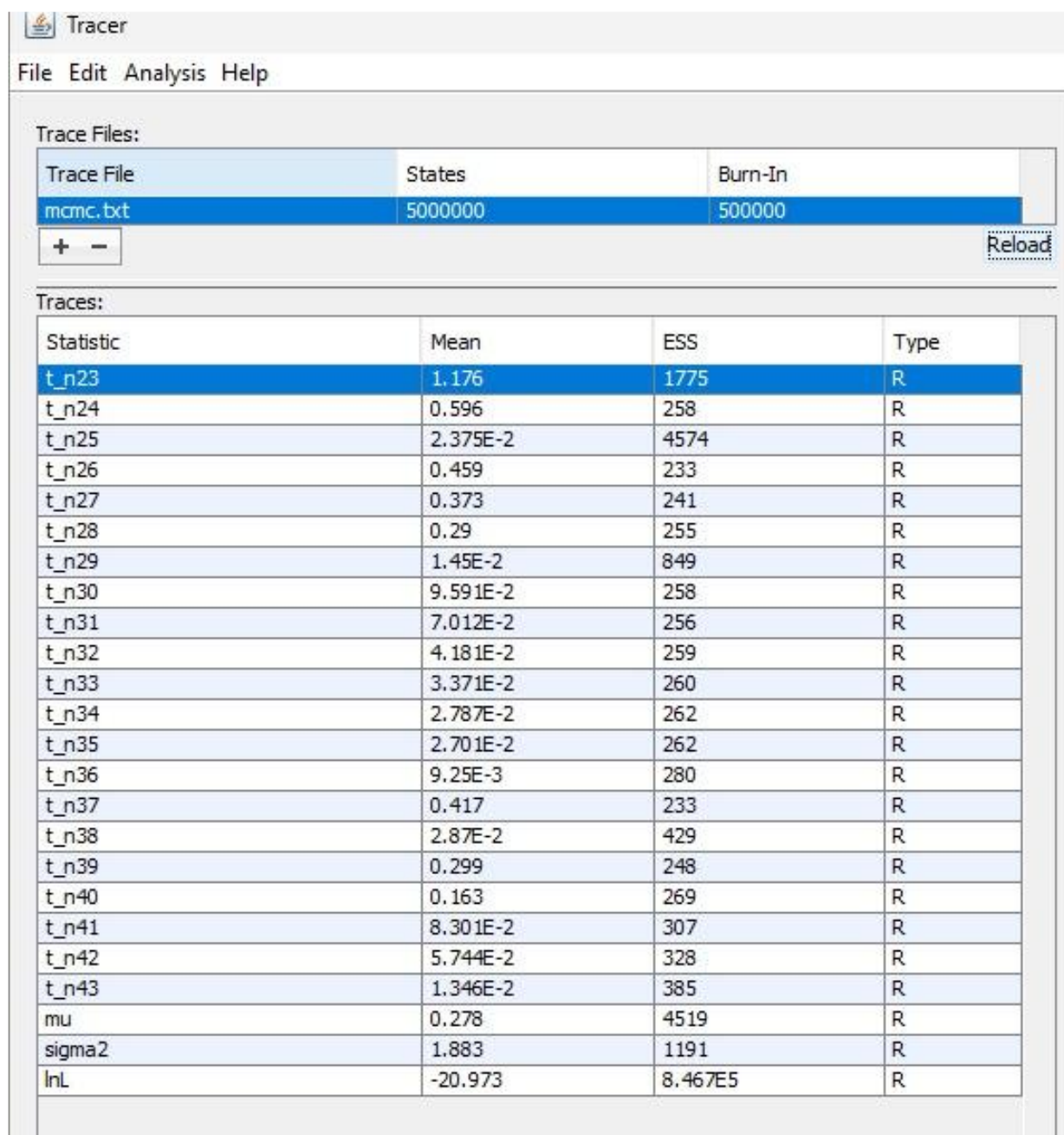
Supplementary Fig. 31. Population statistics and haplotype architecture across Chr6A in *Salicornia* sp. France. **a** Sliding-window comparison of Tajima's D and nucleotide diversity (π) between *Salicornia* sp. France ($n = 71$) and *Salicornia procumbens* ($n = 12$). Statistics were calculated in non-overlapping 50-kb windows along Chr6A using variants mapped to the chromosome-scale *Salicornia* sp. France reference genome, with solid lines indicating LOESS-smoothed trends. **b** Chromosome-wide haplotype visualisation across Chr6A (0–55.3 Mb). Rows represent individual accessions grouped by species, and columns represent polymorphic sites. Genotypes are colour-coded as homozygous reference (light grey), heterozygous (yellow), homozygous alternative (red) and missing data (dark grey). The analysis was performed using 547,872 high-quality SNPs extracted from Chr6A after applying the same filtering criteria described in Supplementary Figs. 17 and 18 (quality ≥ 30 , INFO/DP ≥ 20 , allele frequency $0.01 < AF < 0.99$, MAF ≥ 0.01 , missing data $\leq 20\%$, and heterozygosity $\leq 20\%$).



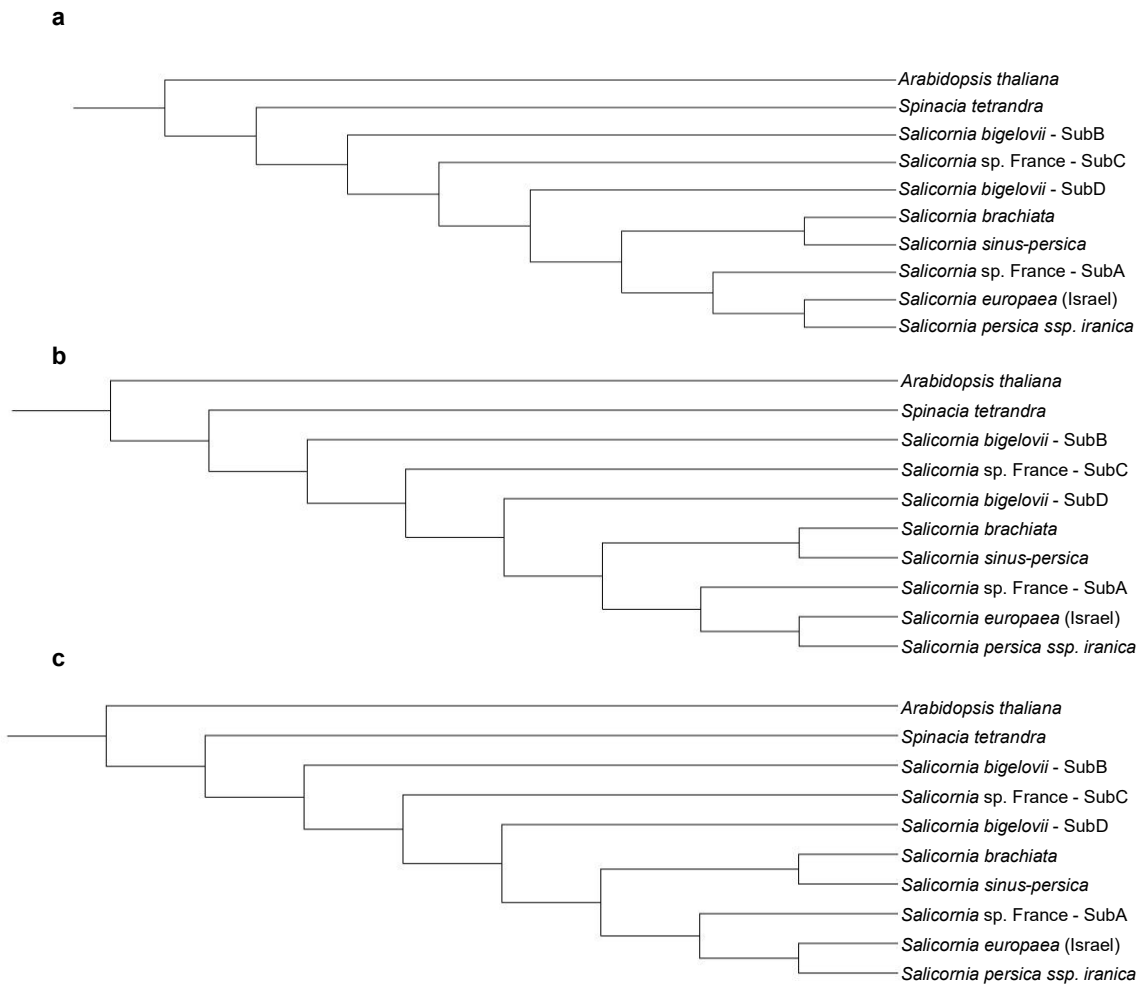
Supplementary Fig. 32. Genetic differentiation (F_{ST}) across Chr6A between *S. procumbens* and *Salicornia* sp. France. Mean Weir and Cockerham's F_{ST} was calculated in 10-kb sliding windows with a 5-kb step size using variants called against the *S. ramosissima* reference genome. To minimize bias resulting from unequal sample sizes, analyses were performed using balanced populations of 12 individuals per species. A total of 850,414 filtered SNPs from chromosome Chr6A were retained following variant filtering (MAF ≥ 0.01 , missing data $\leq 20\%$, and heterozygosity $\leq 20\%$). Each gray point represents the mean F_{ST} value for a single sliding window, and the red curve shows the LOESS-smoothed trend. The blue dashed vertical line indicates the genomic position of the OSCA/CSC1-like gene (35.8 Mb) on chromosome Chr6A of the *S. ramosissima* reference genome.



Supplementary Fig. 33. Single-copy gene phylogeny of the *Salicornia* genus. All allotetraploid *Salicornia* genomes were separated into their respective subgenomes for this analysis. The tree was reconstructed from multi-species single-copy gene protein sequences using the maximum likelihood method implemented in RAxML. Branch lengths have been ignored. Superscript numbers indicate the source publications from which the subgenome nomenclature was adopted: 1, Duan *et al.* ¹; 2, Kimball-Rhines *et al.* ².



Supplementary Fig. 35. Markov chain Monte Carlo (MCMC) convergence diagnostics for node age estimates. Analysis in Tracer shows sufficiently large effective sample sizes (ESS) for all nodes.



Supplementary Fig. 36. Topological congruence under three alignment strategies. Maximum-likelihood phylogenies of 4,977 single-copy orthologues reconstructed under three strategies. **a** Direct coding sequence (CDS) alignment. **b** Codon-aware back-translated alignment (PAL2NAL and trimAl -backtrans; GTR+F+R10). **c** Protein supermatrix (Q.plant+F+R6). Cladograms are rooted on *Arabidopsis thaliana* with branch lengths not drawn to scale. Numbers at internal nodes are ultrafast bootstrap values (1,000 replicates); Shimodaira–Hasegawa-like approximate likelihood ratio test (SH-aLRT) support was 100 at all internal nodes in **b** and **c**. All three reconstructions recover an identical topology, supporting the robustness of the inferred relationships among the *Salicornia* genomes and subgenomes.

Supplementary Table 1. Information of Omni-C, HiFi sequencing and 2C Nuclear DNA content (pg).

Species	Total number of Omni-C bases	Depth of Omni-C bases	Total number of HiFi bases	Depth of HiFi bases	2C Nuclear DNA content (pg)
<i>S. bigelovii</i>	58,619,560,800	28X	59,083,167,995	28X	4.372 ± 0.035
<i>Salicornia</i> . sp. France	91,860,302,400	78X	51,926,572,810	44X	2.309 ± 0.008
<i>S. europaea</i> (Israel)	60,172,820,100	116X	17,950,298,972	34X	1.187 ± 0.007
<i>S. persica</i> ssp. <i>iranica</i>	40,667,157,300	72X	80,163,298,693	143X	1.186 ± 0.011
<i>S. sinus-persica</i>	59,667,989,400	104X	80,013,892,812	140X	1.180 ± 0.017
<i>S. brachiata</i>	78,850,727,100	145X	26,837,871,621	49X	1.130 ± 0.008

Supplementary Table 2. Assembly and evaluation of six *Salicornia* genomes.

Species	Mapping rate of short reads (%)	Scaffold N50 (bp)	Scaffold N90 (bp)	Unplaced contigs length (bp)
<i>S. bigelovii</i>	99.88	113,138,572	98,786,103	11,292,989
<i>S. sp. France</i>	98.85	58,682,116	52,569,409	116,814,903
<i>S. europaea (Israel)</i>	97.32	56,823,638	51,428,091	5,028,763
<i>S. persica</i> ssp. <i>iranica</i>	99.88	56,944,142	43,511,563	43,511,563
<i>S. sinus-persica</i>	99.86	57,506,789	51,937,722	52,524,220
<i>S. brachiata</i>	99.72	55,846,030	49,949,492	48,130,604

Supplementary Table 3. Corresponding chromosome numbers between *S. bigelovii* and *S. depressa*.

<i>S. bigelovii</i>	<i>S. depressa</i>
1B	1B
1D	1A
2B	8B
2D	8A
3B	5B*
3D	7A*
4B	4B*
4D	4A*
5B	9B
5D	9A*
6B	7B*
6D	5A*
7B	3B
7D	3A
8B	2B
8D	2A*
9B	6B
9D	6A*

Chromosomes of *S. depressa* marked with an asterisk (*) have inverted orientation.

Supplementary Table 4. BUSCO-based correspondence between *Chenopodium quinoa* and *S. depressa* chromosomes.

<i>Chenopodium quinoa</i> Chr.	<i>Salicornia depressa</i> Chr.	BUSCO count	Other quinoa hits
CP135301.1_chr6A	1A	210	CP135292.1_chr1B (117), CP135291.1_chr1A (118), CP135302.1_chr6B (210)
CP135302.1_chr6B	1B	209	CP135301.1_chr6A (205), CP135291.1_chr1A (123), CP135292.1_chr1B (123)
CP135293.1_chr2A	2A	349	CP135292.1_chr1B (108), CP135294.1_chr2B (257)
CP135293.1_chr2A	2B	333	CP135294.1_chr2B (249)
CP135299.1_chr5A	3A	178	CP135300.1_chr5B (177), CP135296.1_chr3B (101)
CP135299.1_chr5A	3B	172	CP135300.1_chr5B (166)
CP135303.1_chr7A	4A	140	CP135297.1_chr4A (117), CP135298.1_chr4B (132)
CP135298.1_chr4B	4B	134	CP135297.1_chr4A (115), CP135303.1_chr7A (130)
CP135299.1_chr5A	5A	228	CP135300.1_chr5B (228), CP135296.1_chr3B (120), CP135295.1_chr3A (123)
CP135300.1_chr5B	5B	205	CP135299.1_chr5A (202), CP135303.1_chr7A (137)
CP135301.1_chr6A	6A	131	CP135305.1_chr8A (116), CP135306.1_chr8B (116), CP135302.1_chr6B (121)
CP135301.1_chr6A	6B	129	CP135306.1_chr8B (105), CP135302.1_chr6B (125), CP135305.1_chr8A (106)
CP135291.1_chr1A	7A	207	CP135294.1_chr2B (153), CP135303.1_chr7A (143)
CP135298.1_chr4B	7B	189	CP135295.1_chr3A (116), CP135296.1_chr3B (115), CP135297.1_chr4A (188)
CP135306.1_chr8B	8A	202	CP135307.1_chr9A (188), CP135304.1_chr7B (141), CP135305.1_chr8A (194)
CP135306.1_chr8B	8B	196	CP135305.1_chr8A (191), CP135307.1_chr9A (192), CP135304.1_chr7B (141)
CP135297.1_chr4A	9A	196	CP135298.1_chr4B (195), CP135296.1_chr3B (108), CP135295.1_chr3A (108)
CP135291.1_chr1A	9B	198	CP135296.1_chr3B (109), CP135295.1_chr3A (107), CP135294.1_chr2B (144)

Supplementary Table 5. LTR content from diploid genomes and tetraploid subgenomes analysed using EDTA.

	Copia	Gypsy	LTR unknown
<i>S. sinus-persica</i>	7.72%	17.81%	8.03%
<i>S. brachiata</i>	7.53%	20.22%	6.44%
<i>S. persica</i> ssp. <i>iranica</i>	7.06%	20.10%	7.51%
<i>S. europaea</i> (Israel)	7.81%	20.54%	8.36%
<i>S.</i> sp. France	7.25%	39.47%	10.22%
<i>S. bigelovii</i>	8.36%	41.59%	12.33%
<i>S.</i> sp. France_C	8.26%	25.58%	9.03%
<i>S.</i> sp. France_A	8.55%	20.64%	10.85%
<i>S. bigelovii</i> _D	8.45%	39.72%	10.17%
<i>S. bigelovii</i> _B	7.34%	43.75%	10.56%

Values are the total length of each LTR retrotransposon family, annotated with EDTA, expressed as a percentage of the assembly size of the corresponding genome or subgenome. Copia and Gypsy are LTR superfamilies; LTR unknown refers to LTR elements that could not be assigned to a superfamily. For the allotetraploids *Salicornia* sp. France and *S. bigelovii*, values are given both for the whole assembly and separately for each subgenome (SubA and SubC; SubB and SubD).

Supplementary Table 6. Summary of structural variants (SVs) identified from pairwise whole-genome comparisons among *Salicornia* species and subgenomes.

Query genome	<i>S. persica</i> ssp. <i>iranica</i>	<i>S. sinus-persica</i>	<i>S. brachiata</i>	<i>S. bigelovii</i> - SubB	<i>S. sp. France</i> - SubC
Reference genome	<i>S. europaea</i> (Israel)	<i>S. europaea</i> (Israel)	<i>S. europaea</i> (Israel)	<i>S. bigelovii</i> - SubD	<i>S. sp. France</i> - SubA
Deletion	3,526	6,289	6,391	4,957	3,028
Inversion (INV & related)	154	735	620	628	7,861
Translocation	38	261	238	213	2,097
Duplication	85	276	129	326	6,102
Total SV	3,803	7,561	7,378	6,124	19,088

Structural variants were detected using SyRI, and only SVs with a length ≥ 50 bp were retained for analysis. SVs were classified as deletions (DEL), inversions (INV, including Inverted translocated region and Inverted duplication), translocations (TRANS), and duplications (DUP). SubA, SubB, SubC, and SubD denote different subgenomes of polyploid *Salicornia* species. Each column represents one query–reference genome comparison as indicated in the table header.

Supplementary Table 7. Distribution of structural variants (SVs) across genomic features.

Query genome	<i>S. persica</i> ssp. <i>iranica</i>	<i>S. sinus-persica</i>	<i>S. brachiata</i>
Reference genome	<i>S. europaea</i> (Israel)	<i>S. europaea</i> (Israel)	<i>S. europaea</i> (Israel)
Exonic	216 (5.12%)	303 (3.34%)	418 (4.78%)
Intronic	457 (10.82%)	1,357 (14.96%)	1,414 (16.19%)
Intergenic	3,549 (84.06%)	7,413 (81.70%)	6,904 (79.03%)
Total	4,222	9,073	8,736

Structural variants (≥ 50 bp) identified by SyRI were mapped to the reference genome annotation. SVs overlapping gene models were categorized as exonic or intronic, whereas SVs outside annotated gene regions were classified as intergenic. Each SV was assigned to a single category based on a hierarchical priority order (exonic > intronic > intergenic) to avoid multiple counting. Percentages indicate the proportion of SVs in each category.

Supplementary Table 8. Number of NHX proteins per class across 18 plant species.

Species	Class I	Class II	Class III
<i>Arabidopsis thaliana</i>	4	2	2
<i>Amaranthus cruentus</i>	6	1	1
<i>Amaranthus hypochondriacus</i>	5	1	1
<i>Beta patula</i>	4	1	1
<i>Beta vulgaris</i>	3	1	1
<i>Spinacia oleracea</i>	5	1	1
<i>Spinacia tetrandra</i>	4	1	1
<i>Spinacia turkestanica</i>	5	1	1
<i>Chenopodium quinoa</i> (A B)	1 4	1 1	1 1
<i>Atriplex hortensis</i>	3	1	1
<i>Haloxylon ammodendron</i>	5	1	2
<i>Suaeda glauca</i> (A B)	3 3	1 1	1 1
<i>S. bigelovii</i> (D B)	5 2	1 1	1 1
<i>S. europaea</i> (France) (A C)	4 4	1 1	1 1
<i>S. europaea</i> (Israel)	4	1	1
<i>S. persica</i> ssp. <i>iranica</i>	4	1	1
<i>S. sinus-persica</i>	4	1	1
<i>S. brachiata</i>	4	1	1

Counts indicate the number of NHX proteins assigned to Class I–III based on the maximum-likelihood phylogeny (Supplementary Fig. 16). For polyploid species, values shown as “x | y” denote counts in the two subgenomes (A | B, B | D, or A | C as indicated in the table). Species are listed in the table in the following order: *Arabidopsis thaliana*, *Amaranthus cruentus*, *Amaranthus hypochondriacus*, *Beta patula*, *Beta vulgaris*, *Spinacia oleracea*, *Spinacia tetrandra*, *Spinacia turkestanica*, *Chenopodium quinoa* (A | B), *Atriplex hortensis*, *Haloxylon ammodendron*, *Suaeda glauca* (A | B), *Salicornia bigelovii* (B | D), *Salicornia* sp. (France; A | C), *Salicornia europaea* (Israel), *Salicornia persica* ssp. *iranica*, *Salicornia sinus-persica*, *Salicornia brachiata*.

Supplementary Table 9. Copy number of ion transporters and stress-related protein families across 18 plant species.

Species	HKT*	CBL	CLC	SLAC
<i>Arabidopsis thaliana</i>	1	10	7	5
<i>Amaranthus cruentus</i>	2	9	9	8
<i>Amaranthus hypochondriacus</i>	2	10	8	9
<i>Beta patula</i>	3	7	8	5
<i>Beta vulgaris</i>	3	7	10	5
<i>Spinacia oleracea</i>	5	6	7	5
<i>Spinacia tetrandra</i>	4	7	9	5
<i>Spinacia turkestanica</i>	2	7	8	5
<i>Chenopodium quinoa</i> (A B)	3 3	8 7	8 6	7 7
<i>Atriplex hortensis</i>	5	7	5	6
<i>Haloxylon ammodendron</i>	2	5	8	6
<i>Suaeda glauca</i> (A B)	1 1	7 7	8 7	6 6
<i>S. bigelovii</i> (D B)	2 1	6 6	6 6	5 5
<i>S. sp. France</i> (A C)	1 1	6 6	7 7	6 6
<i>S. europaea</i> (Israel)	2	6	7	6
<i>S. persica</i> ssp. <i>iranica</i>	2	6	7	6
<i>S. sinus-persica</i>	1	6	7	5
<i>S. brachiata</i>	3	6	7	6

Shown are the numbers of genes identified for each protein family, including HKT (High-affinity K⁺ Transporter), CLC (Chloride Channel), CBL (Calcineurin B-Like protein) and SLAC (Slow Anion Channel-Associated protein), based on combined BLAST and HMM searches. For polyploid species, values shown as “x | y” indicate gene copy numbers in the corresponding subgenomes (A | B, B | D, or A | C). *Arabidopsis has only one copy of HKT1.

Supplementary Table 10. Number of SLAC proteins per class across 18 plant species.

Species	Clade I	Clade II	Clade III
<i>Arabidopsis thaliana</i>	2	1	2
<i>Amaranthus cruentus</i>	3	4	1
<i>Amaranthus hypochondriacus</i>	3	4	2
<i>Beta patula</i>	1	2	2
<i>Beta vulgaris</i>	1	2	2
<i>Spinacia oleracea</i>	1	3	1
<i>Spinacia tetrandra</i>	1	3	1
<i>Spinacia turkestanica</i>	1	3	1
<i>Chenopodium quinoa</i> (A B)	1 1	4 4	2 2
<i>Atriplex hortensis</i>	1	3	2
<i>Haloxylon ammodendron</i>	2	2	2
<i>Suaeda glauca</i> (A B)	1 1	3 3	2 2
<i>S. bigelovii</i> (D B)	1 1	2 3	2 1
<i>S. sp.</i> France (A C)	1 1	3 3	2 2
<i>S. europaea</i> (Israel)	1	3	2
<i>S. persica</i> ssp. <i>iranica</i>	1	3	2
<i>S. sinus-persica</i>	1	2	1
<i>S. brachiata</i>	1	3	2

Counts indicate the number of SALC proteins assigned to Clade I–III. For polyploid species, values shown as “x | y” denote counts in the two subgenomes (A | B, B | D, or A | C as indicated in the table). Species are listed in the table in the following order: *Arabidopsis thaliana*, *Amaranthus cruentus*, *Amaranthus hypochondriacus*, *Beta patula*, *Beta vulgaris*, *Spinacia oleracea*, *Spinacia tetrandra*, *Spinacia turkestanica*, *Chenopodium quinoa* (A | B), *Atriplex hortensis*, *Haloxylon ammodendron*, *Suaeda glauca* (A | B), *Salicornia bigelovii* (B|D), *Salicornia* sp. France (A | C), *Salicornia europaea* (Israel), *Salicornia persica* ssp. *iranica*, *Salicornia sinus-persica*, *Salicornia brachiata*.

Supplementary Table 11. *D* statistics among diploid *Salicornia* sourced from the Middle East.

P1	P2	P3	outgroup	<i>D</i> statistic	Z- score	<i>P</i> value	f ₄ -ratio	BBAA	ABBA	BABA
<i>S. ssp.</i> <i>iranica</i>	<i>S.</i> <i>europaea</i> (Israel)	<i>S.</i> <i>europaea</i> (Egypt)	<i>S. sinus-</i> <i>persica</i>	0.244686	3.81197	0.000138	0.2979	180216	178196	108135
<i>S.</i> <i>europaea</i> (UAE)	<i>S. sinus-</i> <i>persica</i>	<i>S.</i> <i>europaea</i> (Israel)	<i>S. ssp.</i> <i>persica</i> <i>iranica</i>	0.075624	4.40402	1.06E-05	0.003053	3.85E+06	6602.58	5674.17

Supplementary Table 12. *Salicornia* diploid species, population size, segregating loci, Nei's index, and diversity.

Species	# Accession	Segregating loci	Nei's index
<i>Salicornia europaea</i> (UAE)	6	1,187,581	0.337
<i>S. sinus-persica</i>	20	2,098,273	0.267
<i>S. europaea</i> (Israel)	17	17,435	0.068
<i>S. persica</i> ssp. <i>iranica</i>	96	2,787,029	0.027
All diploid together	139	14,535,034	0.212

Supplementary Table 13. Genome-wide mean pairwise F_{ST} (Weir and Cockerham's F_{ST}).

Species	<i>S. europaea</i> , 2X (UAE)	<i>S. europaea</i> , 2X (Israel)	<i>S. persica</i> ssp. <i>iranica</i>
<i>S. sinus-persica</i>	0.2	0.78	0.8
<i>S. europaea</i> , 2X (UAE)	-	0.84	0.89
<i>S. europaea</i> , 2X (Israel)	0.84	-	0.59
<i>S. persica</i> ssp. <i>iranica</i>	0.89	0.59	-

Supplementary Table 14. Summary of reference genomes, genotyped species, and number of SNP after filtration.

Reference genome	Species genotypes	# of Filtered SNPs	Analysis
<i>Salicornia europaea</i> (Israel)	All A genome diploids (filtered for biallelic only)	14,542,597	Assembly error rate using short read mapping statistics
<i>Salicornia europaea</i> (Israel)	All A genome diploids (filtered for quality, maf, het and missingness)	10,515,087	Among diploids: population structure, nucleotide diversity (π), pairwise F_{ST} , Nei's index, D-statistics
<i>S. sinus-persica</i>	All A genome diploids (filtered for biallelic only)	12,363,351	Assembly error rate using short reads mapping statistics
Tetraploid <i>Salicornia</i> sp. France (A subgenome)	<i>Salicornia</i> sp. France	4,014,724	SNP density plot
Tetraploid <i>Salicornia</i> sp. France (A subgenome)	All A genome diploids + <i>Salicornia</i> sp. France (filtered for quality, maf, het and missingness)	1,432,526	Genomic relationships between A-subgenome diploid <i>Salicornia</i> species and tetraploid <i>Salicornia</i> sp. France assessed through phylogenetic tree and read alignment statistics.
Tetraploid <i>Salicornia</i> sp. France (A + C) subgenome	Tetraploid <i>Salicornia</i> sp. France and <i>S. procumbens</i>	8,173,539	Tajima's D , nucleotide diversity and haplotype blocks
<i>S. ramosissima</i> (Chr6A only)	Tetraploid <i>Salicornia</i> sp. France and <i>S. procumbens</i>	850,414	Tajima's D , nucleotide diversity and haplotype blocks

Supplementary Note 1. Chromosome assignment, numbering and subgenome relationships

The published *Salicornia* genomes (including *S. depressa*, the Chinese *S. europaea*, and *S. ramosissima*) were examined for subgenome assignment and chromosome numbering. The diploid *S. ramosissima* and tetraploid *S. europaea* followed consistent chromosome numbering and orientation, whereas *S. depressa* showed discrepancies. Therefore, we numbered our chromosomes and adjusted their orientations to maintain consistency with *S. ramosissima* and the Chinese *S. europaea* (Supplementary Fig. 5a, b).

Since *S. ramosissima* diploid shares only one subgenome (A) with the tetraploid species, we observed structural changes and relatively poor homology with the C subgenome of *Salicornia* sp. France (Supplementary Fig. 5a). We communicated with the authors of *S. depressa* regarding the discrepancies in chromosome numbering and orientation compared with other published genomes (Supplementary Fig. 6). Nevertheless, the published *S. depressa* genome showed closer genomic relationships with another *Salicornia* accession from the United States, *S. bigelovii*, sharing substantial homology.

The B subgenome of *S. bigelovii* showed closer genomic relationships with the B subgenome of *S. depressa*, whereas the D subgenome of *S. bigelovii* showed greater similarity to the A subgenome of *S. depressa* (Supplementary Fig. 7 and Supplementary Table 3).

For each *Chenopodium quinoa* chromosome, the *Salicornia depressa* chromosome with the highest number of shared BUSCOs is listed as the best match (Supplementary Table 4), along with the BUSCO count. Other *S. depressa* chromosomes mapping to the same *C. quinoa* chromosome with ≥ 100 shared BUSCOs are reported in the “Other quinoa hits” column. Kimball-Rhines *et al.*² reported that *S. depressa* chromosome assignment was performed using BUSCO ortholog mapping against *C. quinoa*; however, our analysis showed that discrepancies in chromosome numbering still remain (Supplementary Table 4). We also contacted the authors to make them aware of these issues and encourage harmonization of chromosome numbering and ordering for future research.

Supplementary Note 2. k-mer–based neighbour-joining phylogeny

We generated a *k*-mer matrix based on the presence and absence of *k*-mers across the entire population (Supplementary Data 9). Initially, there were 8 million *k*-mer based loci; after filtering out loci with less than 5% minor alleles in the population, 1,020,066 loci were retained for genotyping. The *k*-mer and SNP-based phylogenetic tree effectively differentiated any misclassified *Salicornia* spp. accessions. Five accessions (SalD_42, SalD_75, SalD_80, SalD_132, and SalD_323) clustered with clades different from their initially assigned taxonomic classifications and were corrected for the phylogenetic tree (Fig. 4 and Supplementary Fig. 21).

Supplementary Note 3. Subgenome composition

Subgenomic composition [<https://doi.org/10.5061/dryad.wh70rxwzn>].

Supplementary Note 4. Coverage graphs and mapping

Coverage graphs confirmed accurate mapping of all accessions to specific subgenomes (‘A’, ‘C’, ‘B’, or ‘D’) [<https://doi.org/10.5061/dryad.wh70rxwzn>]. In A-genome diploids, we identified a homoeologous translocation involving the telomeric regions of chromosomes 5A and 5C (Supplementary Fig. 24d). All genome coverage graphs are publicly available [<https://doi.org/10.5061/dryad.wh70rxwzn>]. Accessions SalD_153 and SalD_154, representing outlier *Tecticornia* spp. and *Sarcocornia* spp., genera closely related to *Salicornia*, showed poor read alignment (<10% overall) and sparse genome coverage (available at Dryad under [<https://doi.org/10.5061/dryad.wh70rxwzn>]).

Supplementary Note 5. SNP genotyping and filtering

After filtering (keeping minor allele frequency > 0.01, missing < 20% and het < 20%) from 10.4 million biallelic SNPs, approximately 1.4 million high-quality SNPs were retained and used to construct a phylogeny (Supplementary Table 14). The lowest average overall and unique-concordant mapping rates were observed for *S. sinus-persica* (76.6% and 57.7%, respectively). In contrast, the highest overall and unique-concordant mapping rates were obtained for *S. europaea* accessions from Egypt and Israel (82.5% and 67.2%, respectively) (Supplementary Data 9). Together, these results suggest that the A subgenome of tetraploid *Salicornia* sp.

France likely originated from a diploid lineage not represented in our current panel, as a direct progenitor would typically be expected to exhibit >90% overall alignment.

For the diploid species set, we identified 14,535,034 biallelic SNPs after initial filtering and variant calling against the *S. europaea* (Israel) assembly. After final filtering, 10,515,087 high-quality SNPs were retained. SNP density plots showed no strong genome-wide patterns of SNP distribution among diploids, although slightly higher densities were observed in pericentromeric regions compared with telomeric regions. The minor allele frequency (MAF) spectrum revealed an excess of rare variants, with the highest number of sites occurring at low MAF values (Supplementary Fig. 27b). In addition, small regions of reduced SNP density were detected in the pericentromeric regions of several chromosomes, including Chr1–Chr6 and Chr9 (Supplementary Fig. 27a). The SNP dataset for *S. persica* ssp. *iranica*—the largest diploid group in the panel—comprised 2,785,227 biallelic SNPs and exhibited similar genome-wide patterns of variant distribution and MAF (Supplementary Fig. 27c, d). However, the SNP density and MAF distribution patterns across the nine A-subgenome chromosomes of the tetraploid *Salicornia* sp. France differed from those of the diploid. SNPs were more uniformly distributed, with only a few regions showing higher density (a total of 4,014,724 biallelic SNPs across nine chromosomes), and the MAF distribution exhibited a parabolic pattern, indicating a substantial proportion of sites with relatively high MAF (Supplementary Fig. 27e, f).

The nine C subgenome chromosomes of *Salicornia* sp. France has similar SNP density pattern (Supplementary Fig. 26a), but MAF plot shows these chromosomes have higher number of rare variants (Supplementary Fig. 26b). Together, these patterns highlight pronounced genome-wide variation in SNP density and rare-variant distributions among *Salicornia* species, reflecting differences in evolutionary divergence across lineages.

Supplementary Note 6. Population structure and gene flow analysis

Population structure analysis within the diploid group of *Salicornia* species (Fig. 5a) revealed clear genetic differentiation and admixture patterns, consistent with the groupings observed in the NJ tree and PCA (Fig. 4 and Supplementary Fig. 22). At $K = 2$, *S. sinus-persica* and *S. europaea* accessions from the UAE formed a distinct cluster, separate from *S. persica* ssp. *iranica* and the remaining *S. europaea* accessions from Israel and Egypt (Fig. 5a). Notably, the UAE *S. europaea* exhibited a highly similar genetic structure to *S. sinus-persica*, while the *S. sinus-persica* accessions shared predominantly UAE *S. europaea* ancestry with minor ancestry from *S. persica* ssp. *iranica* (Fig. 5a). In contrast, *S. europaea* accessions from Israel and Egypt primarily carried *S. persica* ssp. *iranica* ancestry, with a small proportion derived from UAE *S. europaea* (Fig. 5a).

At $K = 3$, the population structure remained largely unchanged for most subgroups, except for the *europaea* accessions from Egypt, which displayed an admixed composition – approximately 50% *S. persica iranica* and 50% *S. europaea* from Israel (Fig. 5a). At $K = 4$, the Egyptian *S. europaea* accessions emerged as a distinct subpopulation. Further differentiation occurred at $K = 5$, where UAE *S. europaea* and *S. sinus-persica* separated into distinct subgroups. Finally, at $K = 6$, *S. persica iranica* split into two subgroups corresponding to their geographic origins: accessions from locations QL-1 and QL-2 (which are proximally situated) clustered together, while those from QL-3 formed a separate group (Fig. 5c). Three accessions from QL-1 shared most of their ancestry with the accessions from QL-3.

Using Dsuite on ~10.5 million filtered SNPs, we tested gene flow or introgression among diploid *Salicornia* species with reasonable population size. A significant introgression was detected between *S. europaea* populations from Israel and Egypt ($D \approx 0.25$, $Z > 3.8$; admixture ~30%), indicating substantial gene flow between *Salicornia* sourced from Israel and Egypt (Supplementary Table 11). In contrast, introgression between *S. sinus-persica* and *S. europaea* from Israel was weaker but significant ($D \approx 0.076$, $Z > 4.4$; f_4 -ratio ~3%), suggesting limited introgression (Supplementary Table 11). We did not observe significant introgression between other pairs such as *S. sinus-persica* and *S. persica* ssp. *iranica*.

Supplementary Note 7. Nucleotide diversity and population differentiation

The non-uniform genome-wide distribution of SNP variants was observed in all nine chromosomes (Supplementary Fig. 27a). As expected in a genetically diverse panel, we observed a high proportion of SNPs with low minor allele frequency (MAF) (Supplementary Fig. 27b). Genome-wide nucleotide diversity (π) across the four diploid *Salicornia* groups, estimated from 10.5 million SNPs (Supplementary Table 14), revealed substantial variation. *S. europaea* from the UAE exhibited the highest nucleotide diversity ($\pi =$

0.00119), followed by *S. sinus-persica* ($\pi = 0.00098$), while *S. europaea* from Israel showed reduced diversity ($\pi = 0.00023$). *S. persica* ssp. *iranica*, the only inland population sampled, displayed the lowest diversity ($\pi = 0.0000026$), indicating a pronounced reduction in genetic variation. When all accessions were analyzed collectively, nucleotide diversity was considerably higher ($\pi = 0.00272$), reflecting substantial genetic divergence among the sampled species and subpopulations. Consistent with the patterns of nucleotide diversity, Nei's genetic diversity index varied substantially among the four groups. Nei's genetic diversity was highest in *S. europaea* from the UAE (0.337) and *S. sinus-persica* (0.267), but was markedly reduced in *S. europaea* from Israel (0.068) and particularly in the inland *S. persica* ssp. *iranica* population (0.027) (Supplementary Table 12). Notably, *S. persica* ssp. *iranica* showed the lowest diversity despite having the largest sample size (96 accessions), indicating substantially reduced genetic variation within this inland population.

Population differentiation among diploid *Salicornia* subgroups, measured using Weir and Cockerham's F_{ST} , revealed pairwise genetic relationships between the subgroups. The lowest differentiation ($F_{ST} = 0.2$) was observed between *S. europaea* (UAE) and *S. sinus-persica*, corresponding to their genetic relationship observed in the phylogeny (Fig. 4) and population structure analysis (Supplementary Table 13, Fig. 5a). The *S. europaea* (Israel) and *S. persica* ssp. *iranica* showed moderate divergence ($F_{ST} = 0.59$), while *S. sinus-persica* and *S. europaea* (Israel) exhibited stronger differentiation ($F_{ST} = 0.78$). All other pairwise comparisons showed even stronger genetic divergence ($F_{ST} \geq 0.8$) (Supplementary Table 13).

Supplementary Note 8. Flow cytometry and genome size estimation

Relative fluorescence intensity of propidium iodide-stained nuclei was estimated using CytoFLEX flow cytometer (Beckman Coulter, Krefeld, Germany). *Solanum lycopersicum* cv. Stupicke polni tyckove rane (2C = 1.96 pg), *Glycine max* cv. Polanka (2C = 2.5 pg) and *Pisum sativum* cv. Ctirad (2C = 9.09 pg) were used as internal reference standards^{3,4}. The mean nuclear DNA content was calculated according to Doležel *et al.*⁵. The estimated genome sizes are summarized in Supplementary Table 1.

Supplementary Note 9. PacBio HiFi library preparation and sequencing

A modified CTAB protocol was employed: 5 g of tissue was mixed with 10 ml of CTAB buffer (2% w/v CTAB, 1.4 M NaCl, 100 mM Tris-HCl, pH 8, 20 mM EDTA, pH 8, 2% w/v PVP [MW 10,000], 1% v/v Triton X-100, 0.1% v/v β -mercaptoethanol, 1.9 mg mL⁻¹ RNase A, and 10 mg mL⁻¹ proteinase K)⁸. The mixture was incubated at 65°C for 45 min with periodic mixing, followed by three rounds of extraction with chloroform: isoamyl alcohol (24:1) to recover the aqueous phase. DNA was precipitated using 0.6 volumes of isopropanol, pelleted, washed with 70% (w/w) ethanol, and finally resuspended in TE buffer (pH 8.0). DNA quality was assessed by fragment size using an Advanced Analytical Fragment Analyzer (automated CE Systems) or BluePippin (Sage Analyzer) according to manufactures' instructions. Subsequently, HiFi libraries were constructed according to the "Procedure & Checklist - Preparing HiFi SMRTbell® Libraries using the SMRTbell Express Template Prep Kit 2.0" (PN 101-853-100, Pacific Biosciences). Ten micrograms of DNA was sheared using the Megaruptor 2 system (Diagenode) to achieve an average fragment size of 15–20 kb. Size-selected libraries were then sequenced on a PacBio Sequel II system in circular consensus sequencing (CCS) mode for 30 hours.

Supplementary Note 10. Genome assembly, scaffolding and quality assessment

Each contig of contig level assemblies were aligned with the NCBI Nucleotide database by blast (v2.14.1)⁶. Contamination contigs and low-quality contigs were removed by the result of BlobTools (v1.1.1)⁷. Contaminated contigs are those identified based on alignment results as originating from contigs that do not belong to Streptophyta, while low-quality contigs are those with a relatively low average sequencing depth. For *S. brachiata*, *S. sinus-persica*, *S. europaea* (Israel), *S. persica* ssp. *iranica*, and *S. bigelovii*, contigs were anchored using Omni-C data with Juicer (v2.0)⁸.

Using Illumina short-read data from the assembled accession and additional accessions within the species, we performed variant calling and estimated genome-wide polymorphism. The proportion of alternate alleles detected in the assembled accession was interpreted as a measure of sequencing or assembly error rate. The quality of all genome assemblies generated in this study was assessed using a k-mer-based approach implemented in Merquy (v1.3)⁹. For each genome, high-fidelity (HiFi) sequencing reads from the corresponding sample were used to construct a k-mer database, with a k-mer size of 19.

The k-mer database derived from the HiFi reads was compared with the corresponding genome assembly to estimate the consensus quality value (QV), which reflects the base-level accuracy of the assembled sequences. The QV values reported by Merqury were used to evaluate and compare the accuracy of all genome assemblies.

Supplementary Note 11. Identification of specific oligomers and labeling of oligo probes

Sets of 27,000 oligomers (45-mers) per chromosome arm were synthesized by Arbor Biosciences (Ann Arbor, MI, USA). The painting probes were labeled directly by Cy5 or indirectly by biotin or digoxigenin according to Beránková and Hříbová¹⁰.

45S rDNA probe was prepared from a clone pTA71¹¹ by nick translation (Roche Applied Science, Penzberg, Germany) using aminoallyl-dUTP-Cy5 (Jena Biosciences, Jena, Germany) according to the instructions of the manufacturer.

Chromosome spreads were prepared according to Salazar *et al.*¹² Hybridization mix (30 µl) containing 50% (v/v) formamide, 10% (w/v) dextran sulfate in 2x SSC and 10 ng/µl labeled probe was added onto slide and denatured for 40 s at 80°C. Hybridization was carried out overnight at 37°C. The sites of hybridization of digoxigenin- and biotin-labeled probes were detected using anti-digoxigenin-FITC (Roche Applied Sciences) and streptavidin-Cy3 (ThermoFisher Scientific), respectively. Chromosomes were counterstained with DAPI and mounted in Vectashield Antifade Mounting Medium (Vector Laboratories, Burlingame, CA, USA). The slides were examined with Axio Imager Z.2 Zeiss microscope (Zeiss, Oberkochen, Germany) equipped with Cool Cube 1 camera (Methasystems, Altlussheim, Germany) and appropriate optical filters. The signal capturing and picture processing were performed using ISIS software 5.4.7 (Metasystems) and the final image adjustment was completed in Adobe Photoshop CS5.

Supplementary Note 12. Subgenome differentiation and homoeologous chromosome assignment

The identified homologous chromosomes were subsequently classified into subgenomes via a repetitive *k*-mer based method implemented in SubPhaser (v1.2) with default parameters¹³. Based on the longest representative proteins from diploid *Salicornia* species and these subgenomes, orthologous groups were constructed using OrthoFinder (v2.4.0) with default parameters¹⁴. In parallel, whole-genome sequencing reads from *Salicornia* sp. France were mapped to both the *Salicornia* sp. France genome and the *S. bigelovii* subgenome using minimap2 (v2.1) under default settings¹⁵. Coverage information was calculated using sambamba (v1.0) with the parameters (-t 32 -w 1000000)¹⁶.

Supplementary Note 13. Phylogenetic analysis

Arabidopsis thaliana was included as an outgroup to root the tree due to its well-annotated reference genome and its position within core eudicots. The same outgroup was retained in the genus-level analysis (see "Subgenomes phylogenomic analysis") to maintain consistency across all phylogenetic reconstructions in this study. Within Amaranthaceae, weakly halophytic taxa, including *Amaranthus cruentus*, *A. hypochondriacus*, *Beta vulgaris*, *B. patula*, and *Spinacia* species (*Spinacia oleracea*, *Spinacia tetrandra*, *Spinacia turkestanica*), were included to provide phylogenetic context and close relatives of *Salicornia*. The three *Spinacia* species were specifically selected because high-quality genome assemblies and annotations are available for all of them, and because they represent both a cultivated species (*Spinacia oleracea*) and wild relatives (*Spinacia tetrandra* and *Spinacia turkestanica*), allowing robust comparative analyses while minimizing annotation-related bias. To represent independent halophytic lineages within the family, we further included *Atriplex hortensis*, *Haloxylon ammodendron*, *Chenopodium quinoa*, and *Suaeda glauca*. These taxa allow comparison between *Salicornia* and other salt-adapted lineages and account for the effects of polyploidy, lineage-specific gene family expansions, and copy-number variation, which can influence inferred gene-family dynamics and functional interpretations.

Supplementary Note 14. Gene Ontology enrichment analysis

A curated GO term description library was used to associate enriched GO terms with functional descriptions and ontology categories. The enrichment results were corrected for multiple testing, and significantly enriched GO terms were summarized and ranked based on statistical significance.

Supplementary Note 15. Ancestor karyotype construction

For ancestral genome reconstruction, we used six species (*S. brachiata*, *S. sinus-persica*, *S. europaea* (Israel) and *S. persica* ssp. *iranica*, *S. bigelovii*, *Salicornia* sp. France and *Haloxylon ammodendron*). For each species, genes located on chromosomes were first processed by retaining only the longest protein isoform per gene. These protein sequences were then used to infer orthologous relationships across species using OrthoFinder (v2.4.0) with default parameters, generating orthogroups based on homologous proteins.

Syntenic blocks were subsequently constructed following the pipeline described in processDrimm (<https://github.com/xjtu-omics/processDrimm>). Briefly, conserved blocks were defined as collinear segments containing at least 20 consecutive homologous proteins shared among species. These blocks represent conserved ancestral adjacencies at the protein-coding gene level.

The resulting syntenic blocks were used as input for IAGS (v 22.02.15)¹⁷ to reconstruct the ancestral karyotype under a parsimonious rearrangement framework.

Supplementary Note 16. Fastq and k-mer processing

Raw sequencing fastq files were trimmed using fastp (v0.23.2) with a Phred quality score threshold of 20 and a minimum read length of 150 bp¹⁸.

Briefly, *k*-mers for each accession are extracted and binned deterministically. Singleton *k*-mers, which are likely erroneous, are filtered out at this stage. Subsequently, an independent matrix is constructed for each bin, facilitating its use in downstream population analysis¹⁹. *k*-mers occurring less than 5% frequency in the population were removed from the *k*-mer matrix.

Supplementary Note 17. Neighbor-joining (NJ) phylogenetic trees

Briefly, phylogenetic clustering was performed using the ‘dist’ function to calculate genetic distances. The R (R programming language) packages of ‘ape’ and ‘phyclust’ were used to convert the genetic distance matrix to a phylo object and for plotting the tree, respectively²⁰. Additionally, a principal component analysis (PCA) plot was also generated based on *k*-mer data following Adhikari *et al.*^{21–23} to visualize the plot.

To further dissect the *Salicornia* germplasm collection and identify the subgenome carried by each accession, 318 accessions were mapped to an in silico octoploid reference (subgenomes A, C, B, D) assembly. This reference was constructed by merging two tetraploid assemblies *S. europaea* (4X) and *S. bigelovii* (4X) reported in this study. The resulting read-depth coverage plots enabled the assignment of accessions to specific subgenomes (A, B, C, and D). The approach described in Adhikari *et al.*²⁴ was adopted here to generate read-depth coverage plots and confirm subgenomes and ploidy. Using the octoploid genome, we mapped reads to several chromosomes absent in the diploid or tetraploid species, so a different normalization approach was applied than described in Adhikari *et al.*²¹ to plot coverage in these euploids. The top 10% of reads mapped in 100 kb bins were extracted to calculate the average of these bins (ten_avg). Then, 10% of this ten_avg value was used as a cutoff to select bins (i.e., bins with mapped read counts $\geq$ 10% of ten_avg) for normalization. The average of these selected bins in the truly mapped chromosomes (selected_bins_avg) was used as the normalization factor.

Therefore, average coverage 1 represents two copies of a chromosome, 0.5 represents a single copy, and values close to 0 suggest the chromosome is absent.

Supplementary Note 18. SNP filtration and downstream analyses

Further, the confirmed diploid accessions were mapped to the diploid *S. europaea* (Israel) genome assembly generated in this study to identify single nucleotide polymorphisms (SNPs) and perform genotyping. Trimmed sequences were mapped using the HISAT2 (v2.2.1) aligner, followed by variant calling with BCFtools^{25,26}. Raw variants were stringently filtered for quality (QUAL $\geq$ 30), read depth (DP $\geq$ 20), allele frequency ($0.01 < AF < 0.99$), and missing data ($\leq$ 20%), with additional filtering to exclude loci with MAF < 0.02 and heterozygosity (> 0.2) using in house script. Subsequently, the selected variants were merged into a single VCF file and then converted to a SNP matrix HapMap file using an inhouse script, which was then used for subsequent population genetic analyses. The population structure analysis was done in R (R programming) using the Landscape and Ecological Association Studies (LEA) package (v3.14.0)²⁷. Population structure was performed for all diploids *Salicornia* species together and also within the *S. persica* ssp. *iranica*. The codes

used in the *Salicornia* population analysis are available at Github [<https://github.com/laxmangene7/halophyte-salicornia-genomics>].

Supplementary Note 19. Nei's diversity indices, nucleotide diversity (π), segregating loci in diploids and F_{ST}

The genetic diversity of naturally collected diploid *Salicornia* populations was evaluated using Nei's diversity indices, nucleotide diversity (π), and the total number of segregating loci within genetically differentiated subgroups. These analyses were performed for *S. sinus-persica*, *S. europaea* from Israel, and *S. persica* ssp. *iranica*. For both diversity indices, genomic subsets were extracted from the VCF file using VCFtools (v0.1.17)²⁸. Nei's index was calculated following established methods²², while π was computed in 50 Kb sliding windows across the genome for each subgroup²⁹, with mean π values derived accordingly. To determine the total number of segregating loci within each subgroup, we excluded sites with entirely missing genotypes or those fixed (MAF = 0) and then tallied the remaining polymorphic loci.

Pairwise F_{ST} was estimated between the *S. sinus-persica* clade, *S. europaea* (Israel), and *S. persica* ssp. *iranica* populations using VCFtools (v0.1.17)^{28,30}. Mean pairwise F_{ST} was calculated in 100 Kb genomic windows for all three diploid population pairs, followed by computation of the genome-wide mean F_{ST} for each comparison.

Supplementary references

1. Duan, H. *et al.* A T2T gap-free genome provides insights into tetraploid *Salicornia europaea* evolution and aquaporin function in stress tolerance. *Plant Physiology* **199**, kiaf565 (2025).
2. Kimball-Rhines, C. *et al.* The annotated, chromosome-scale *Salicornia depressa* (American pickleweed) genome. *Journal of Heredity* **117**, 357–365 (2025).
3. Doležel, J., Sgorbati, S. & Lucretti, S. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum* **85**, 625–631 (1992).
4. Doležel, J. *et al.* Plant Genome Size Estimation by Flow Cytometry: Inter-laboratory Comparison. *Annals of Botany* **82**, 17–26 (1998).
5. Doležel, J., Bartoš, J., Voglmayr, H. & Greilhuber, J. Letter to the editor. *Cytometry Part A* **51A**, 127–128 (2003).
6. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410 (1990).
7. Laetsch, D. R. & Blaxter, M. L. BlobTools: Interrogation of genome assemblies. Preprint at <https://doi.org/10.12688/f1000research.12232.1> (2017).
8. Durand, N. C. *et al.* Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Systems* **3**, 95–98 (2016).
9. Rhie, A., Walenz, B. P., Koren, S. & Phillippy, A. M. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol* **21**, 245 (2020).
10. Beránková, D. & Hříbová, E. Bulkied Oligo-FISH for Chromosome Painting and Chromosome Barcoding. *Methods Mol Biol* **2672**, 445–463 (2023).
11. Gerlach, W. L. & Bedbrook, J. R. Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucl Acids Res* **7**, 1869–1885 (1979).
12. Salazar, O. R. *et al.* SOS1 tonoplast neo-localization and the RGG protein SALT1 are important in the extreme salinity tolerance of *Salicornia bigelovii*. *Nat Commun* **15**, 4279 (2024).
13. Jia, K.-H. *et al.* SubPhaser: a robust allopolyploid subgenome phasing method based on subgenome-specific k-mers. *New Phytologist* **235**, 801–809 (2022).
14. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* **20**, 238 (2019).
15. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
16. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
17. Gao, S. *et al.* IAGS: Inferring ancestor genome structure under a wide range of evolutionary scenarios. *Mol Biol Evol* **39**, msac041 (2022).
18. Chen, S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* **2**, e107 (2023).
19. Cavalet-Giorsa, E. *et al.* Origin and evolution of the bread wheat D genome. *Nature* **633**, 848–855 (2024).
20. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
21. Adhikari, L. *et al.* Genetic characterization and curation of diploid A-genome wheat species. *Plant Physiology* **188**, 2101–2114 (2022).
22. Adhikari, L. *et al.* Genomic characterization and gene bank curation of *Aegilops*: the wild relatives of wheat. *Front. Plant Sci.* **14**, 1268370 (2023).
23. Adhikari, L. *et al.* Dissecting the population structure, diversity and genetic architecture of rust resistance in wild emmer wheat (*Triticum turgidum* subsp. *dicoccoides*). *BMC Genomics* **27**, 284 (2026).
24. Adhikari, L. *et al.* A high-throughput skim-sequencing approach for genotyping, dosage estimation and identifying translocations. *Sci Rep* **12**, 17583 (2022).
25. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).

26. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907–915 (2019).
27. Frichot, E. & François, O. LEA: An R package for landscape and ecological association studies. *Methods Ecol Evol* **6**, 925–929 (2015).
28. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
29. Ahmed, H. I. *et al.* Einkorn genomics sheds light on history of the oldest domesticated wheat. *Nature* **620**, 830–838 (2023).
30. Weir, B. S. & Cockerham, C. C. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**, 1358–1370 (1984).