

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	PacBio HiFi sequencing was performed on a PacBio Sequel II system in circular consensus sequencing (CCS) mode. Omni-C libraries were sequenced on the Illumina HiSeq X platform by Dovetail Genomics. Whole-genome resequencing of 318 accessions was performed on an Illumina NovaSeq 6000 platform by Novogene (Beijing, China). Flow cytometry data were collected using CytExpert software (CytoFLEX platform, Beckman Coulter).
Data analysis	Genome assembly & QChiasm (v2.28) · BLAST (NCBI Nucleotide alignment, no version stated) · minimap2 (v2.28, HiFi read alignment to assemblies) · samtools (v1.16.1) · blobtools (v1.1.1) · Juicer (v2.0) · 3d-dna (v180922) · Juicebox (v2.17.00) · HapHic (v1.0.5) · BUSCO (v5.7.1, Viridiplantae) · QUAST (v5.2.0) · LTR_Finder (v1.0.7) · quarTeT (v1.2.0) · Chorus (v1.1)Subgenome/homoeolog assignmentD-GENIES (v1.5.0) · SubPhaser (v1.2) · OrthoFinder (v2.4.0) · minimap2 (v2.1, WGS coverage mapping) · sambamba (v1.0)Transcriptome & gene annotationHISAT2 (v2.2.1) · samtools (v1.16.1) · Trinity (v2.15.1) · CCS (v6.4.0) · Iso-Seq tool (v4.0.0) · EDTA (v2.0.0) · RepeatModeler (v2.0.1) · RepeatMasker (v4.1.2) · BRAKER3 (no version stated) · eggNOG-mapper (v2)SyntenyGenomeSyn (v1.41) · minimap2 (v2.21)PhylogenomicsOrthoFinder (v2.4.0) · RAXML (v8.0.0) · MCMCTREE (v1.0) · Tracer (v1.7.2) · MAFFT (v7.505) · trimAl (v1.5) · PAL2NAL (v14) · IQ-TREE (v3.0.1) · ModelFinder (no version stated) · WGD (v0.4.5) · HMMER (v3.1b2) · BLASTP (v2.14.1) · MUSCLE (v5.3) · CAFE5 (v1.1)Differential expression & enrichmentTrimmomatic (v0.39) · HISAT2 (2.2.1) · featureCounts (Subread v2.0.4) · DESeq2 (v1.42.0) · Cufflinks (v2.2.1) · SyRI (v1.5.5) · eggNOG-mapper (v2) · clusterProfiler (v4.10.1) · Mercator4 (no version stated)Population genomicsfastp (no version stated) · KGWASMatrix (no version stated) · HISAT2 (v2.2.1) · BCFtools (no version stated) · GATK (v4.3.0.0) · VCFtools (v0.1.17) · ape (v5.8.1) · phyclust (v0.1-34) · LEA (no version stated) · Pophelper (v2.3.1) · Dsuite (no version stated) · R (v4.3.0) · ggplot2.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data availability: The raw FASTQ files of the whole-genome sequencing panel of *Salicornia* accessions used in this study, along with PacBio HiFi long reads, Iso-Seq, and Omni-C data, have been deposited at the National Center for Biotechnology Information (NCBI) under BioProject accession code PRJNA1247873 [<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1247873>]. Six chromosome-scale genome assemblies generated in this study have been deposited at the Genome Warehouse in the National Genomics Data Center under project accession code PRJCA036664 [<https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA036664>]. The genome assemblies and gene annotation files are available at Dryad under DOI 10.5061/dryad.wh70rxwn [<https://doi.org/10.5061/dryad.wh70rxwn>]. Publicly available RNA-seq data for *S. europaea* (Israel) and *S. bigelovii* used in this study are available under BioProject accession [PRJNA733892; <https://www.ncbi.nlm.nih.gov/bioproject/607385>].

Code availability: All custom scripts used for genome assembly, whole-genome triploidization and divergence analyses, and intra-family transcriptomic comparisons are available at [<https://github.com/supermanwasd/Salicornia-genome>]. Scripts for the phylogenetic tree, genome coverage, and population structure analyses are available at [<https://github.com/laxmangene7/halophyte-salicornia-genomics>].

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	does not apply
Reporting on race, ethnicity, or other socially relevant groupings	does not apply
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	"We did not run a power analysis or use any statistical method to set sample sizes. The numbers here were set by what material we could get and by what is standard for this kind of comparative genomics. We assembled six genomes, one accession per species, selected to span the diploid and tetraploid lineages of <i>Salicornia</i> and their geographic range. We did not assemble replicate accessions because chromosome-scale assembly is expensive and the bottleneck is coverage of the genus, not replication within it. The resequencing panel of 318 accessions was everything we could obtain from germplasm collections and our own cultivation, so the group size reflects the collection size. Wild-collected seeds for whole-genome sequencing were sampled from individual plants, targeting 20–100 plants per 100 m² population to ensure broad representation of within-population genetic diversity. This target could not be met in all regions, as some populations were too small or seed viability was insufficient from individual plants. A statistical sampling method was not applied, as the genetic diversity of these populations was not known at the time of collection. Three individual plants representing each analyzed species were used for flow cytometry. Selection scans and detailed population-level comparisons focused on <i>Salicornia</i> sp. France (n = 71) and <i>S. procumbens</i> (n = 12), two closely related tetraploid populations that share the same subgenomes and cluster within the same phylogenetic clade. Population genomic analyses, including nucleotide diversity (π), Tajima's D, and FST, were performed across these two populations using sliding genomic windows, with each comparison based on thousands of window estimates rather than accession counts alone. For Nei's genetic diversity and pairwise FST
-------------	---

analyses, we focused on well-defined diploid species/populations. The diploid panel comprised 139 accessions, including *S. europaea* (UAE; $n = 6$), *S. sinus-persica* ($n = 20$), *S. europaea* (Israel; $n = 17$), and *S. persica* ssp. *iranica* ($n = 96$).

For the cytogenetics we captured at least five complete metaphase spreads per probe combination and genotype, which is what we normally use for oligo-FISH karyotyping and was enough to score chromosome number and translocation position without ambiguity. *S. procumbens* is the exception where we could only recover one complete spread. The differential expression comparisons used three biological replicates per condition, following the original studies the datasets came from.

Data exclusions

Data exclusions were applied at several analytical stages, using pre-established filtering criteria: In genome assembly, contigs identified by blobtools as contaminated (not originating from Streptophyta) or of low quality (low average sequencing depth) were excluded. In population genomic analyses, raw variants were filtered using bcftools ($QUAL \geq 30$, $INFO/DP \geq 20$, $0.01 < AF < 0.99$), followed by additional filtering for minor allele frequency ($MAF > 0.02$), missing data ($< 20\%$), and heterozygosity ($< 20\%$); only high-confidence, biallelic SNPs passing all criteria were retained. Structural variants shorter than 50 bp were excluded from structural-variant analysis. In differential expression analysis, genes with total raw counts < 10 across all samples were filtered out prior to testing. In phylogenomic analysis, seven orthogroups whose coding sequences could not be unambiguously back-translated to protein were excluded from the codon-aware alignment set (4,970 of 4,977 orthogroups retained). Accessions showing discordance between genetic clustering and passport information were re-evaluated and taxonomically curated rather than excluded outright.

Replication

Three independent rounds of genome size estimation using flow cytometry were performed for every analyzed plant. We did not repeat the sequencing experiments, since each assembly was validated against independent data rather than by repetition. For the two reference assemblies we re-sequenced the same plants with Illumina and mapped the reads back; polymorphic-site rates were 0.0024% (*S. sinus-persica*, SalD_871_3) and 0.00029% (*S. europaea*, SalD_43). Assembly completeness and continuity were also checked with BUSCO, LAI and Merquy and the chromosome-scale scaffolds were checked against the Omni-C contact maps.

The scaffolding was then confirmed cytogenetically. Oligo-FISH on metaphase chromosomes recovered the same chromosome numbers as the assemblies and independently detected the reciprocal translocations we had called from the sequence data.

For the computational analyses we repeated key inferences under different assumptions. The species tree was reconstructed from a direct CDS alignment, a codon-aware back-translated alignment and a protein supermatrix, and all three gave the same topology. Divergence times were run to convergence and checked in Tracer. Subgenome assignment was supported independently by k-mer analysis, SNP-based phylogeny and read-depth coverage, and the population structure was consistent across k-mer PCA, a SNP-based neighbour-joining tree and ADMIXTURE from $K = 2$ to $K = 6$.

All cytogenetic replication attempts were successful except one; we only recovered one complete metaphase spread for *S. procumbens* across all the slides prepared, so that observation rests on one cell. This is stated in the relevant figure legend.

Randomization

This is an observational comparative and population genomics study of wild-collected and germplasm-bank *Salicornia* accessions rather than a controlled experiment with treatment groups; therefore, random allocation into experimental groups was not applicable. Accessions were grouped by taxonomic identity, subgenome composition, and geographic origin. Initial taxonomic assignments (based on passport/collection records) were verified and, where necessary, corrected using genotype-based clustering: both k-mer-based and SNP-based neighbour-joining phylogenies were used to identify accessions that clustered discordantly with their assigned taxonomic group, and five accessions (SalD_42, SalD_75, SalD_80, SalD_132, and SalD_323) were reassigned accordingly prior to downstream analysis. Subgenome composition (A, B, C, or D) for all 318 resequenced accessions was independently confirmed by mapping reads to an in silico octoploid reference and evaluating normalized coverage across subgenome-specific chromosomes, rather than by geographic or passport-based assumption. Population subgroupings used for diversity, differentiation, and introgression analyses (e.g., *S. sinus-persica*, *S. europaea* from Israel/Egypt/UAE, *S. persica* ssp. *iranica*) were similarly defined by this genetically confirmed identity rather than arbitrary or investigator-selected allocation.

Blinding

Blinding was not relevant to this study, as it involved objective molecular and computational measurements (genome sequencing, assembly, and population-genetic analysis) rather than subjective scoring or assessment where investigator bias could influence data collection or interpretation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | | |
|-------------------------------------|--|
| n/a | Involved in the study |
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology and archaeology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |
| ☐ | ☒ Plants |

Methods

- | | |
|-------------------------------------|--|
| n/a | Involved in the study |
| ☒ | ☐ ChIP-seq |
| ☐ | ☒ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

- | No | Yes |
|-------------------------------------|---|
| ☒ | ☐ Public health |
| ☒ | ☐ National security |
| ☒ | ☐ Crops and/or livestock |
| ☒ | ☐ Ecosystems |
| ☒ | ☐ Any other significant area |

Experiments of concern

Does the work involve any of these experiments of concern:

- | No | Yes |
|-------------------------------------|--|
| ☒ | ☐ Demonstrate how to render a vaccine ineffective |
| ☒ | ☐ Confer resistance to therapeutically useful antibiotics or antiviral agents |
| ☒ | ☐ Enhance the virulence of a pathogen or render a nonpathogen virulent |
| ☒ | ☐ Increase transmissibility of a pathogen |
| ☒ | ☐ Alter the host range of a pathogen |
| ☒ | ☐ Enable evasion of diagnostic/detection modalities |
| ☒ | ☐ Enable the weaponization of a biological agent or toxin |
| ☒ | ☐ Any other potentially harmful combination of experiments and agents |

Plants

Seed stocks	Salicornia sinus-persica (KSA) seeds collected from wild grown mature plants collected in December 2021 from Tarout Island (Kingdom of Saudi Arabia) by Dr Melino and Ms Fiene. Seeds from individual plants were collected and stored in envelopes.
Novel plant genotypes	Salicornia europaea (IS) seeds were provided by Dr. Moshe Sagi from the Jacob Blaustein Institute for Desert Research (BIDR) and originate from the Dead Sea region (Israel). Salicornia persica ssp. iranica seeds collected from wild grown mature plants in December 2021 from Aushazia lake in Al-Qassim (Kingdom of Saudi Arabia) by Drs Melino and Al-Turki. Salicornia sp. France and Sarcocornia fructicosa seeds were collected from wild grown mature plants in October 2022 from Etang de Thau in the Marseille region (France) by Dr Melino and Ms Fiene. Sarcocornia unknown (FR) seeds were collected from wild grown mature stems of plants from the saline de reprise of Gruissan region (France) by Prof. Tester and Ms. Fiene. Salicornia bigelovii seeds were provided by Dr Glenn of the Environmental Research Laboratory, University of Arizona, Tucson, USA. Salicornia europaea (RAK and UAQ) seeds were provided by Mohammed Shahid (International Centre for Biosaline Agriculture, UAE) from Ras al-Khaimah (RAK) or Umm al-Quwain (United Arab Emirates) with an unknown collection date but received by authors in December 2022. Salicornia veneta (ITA) and Salicornia fructicosa (ITA) seeds were collected in November 2022 from the Venetian Lagoon of Chioggia region in Italy by Dr De Battisti. Tecticornia pergranulata (AU), Tecticornia pruinosa (AU) and Tecticornia pergranulata (AU) seeds were collected in December 2020 from the delta of the Gawler river in South Australia (Australia) by Ms Peri Coleman (Delta Environmental Consulting). In all examples seeds were collected into an envelope and if not sufficiently dried, seeds were dried in an oven at 50°C for 5 hours before storage in aluminium foil packs inside a secondary container at 4°C.
Authentication	Not applicable

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Nuclear genome size was estimated by flow cytometry as described previously (Salazar et al. 2024). Briefly, intact leaf tissues of a Salicornia species and 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) that served as internal reference standards, were chopped together in a glass Petri dish containing 500 µl Otto I solution (0.1M citric acid, 0.5% Tween 20; Otto, 1990). The crude suspension was

filtered through a 50 µm nylon mesh. Nuclei were then pelleted (300 x g, 2 min) and resuspended in 300 µl of Otto I solution. After 15 min of incubation on ice, 600 µl of Otto II solution supplemented with 50 µg/ml RNase and 50 µg/ml propidium iodide were added.

Instrument

Samples were analyzed using a CytoFLEX flow cytometer (Beckman Coulter, Krefeld, Germany).

Software

Analysis was performed using CytExpert Software for the CytoFLEX platform (Beckman Coulter), and 2 C nuclear DNA contents were calculated from means of G1 peak positions by applying the formula: 2 C nuclear DNA content = (sample G1 peak mean) × (standard 2 C DNA content) / (standard G1 peak mean). The mean nuclear DNA content in pg was converted to genome size in bp using the conversion factor 1 pg DNA = 0.978 Gbp (Doležel et al., 2003).

Cell population abundance

Three individual plants per genotype were sampled, and each sample was analyzed three times, each time on a different day. A minimum of 5000 nuclei per sample were analyzed.

Gating strategy

The low level threshold was set to channel 20 to eliminate particles with the lowest fluorescent intensity from the histogram, all remaining fluorescent events were recorded with no further gating used.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.