Supplementary Material

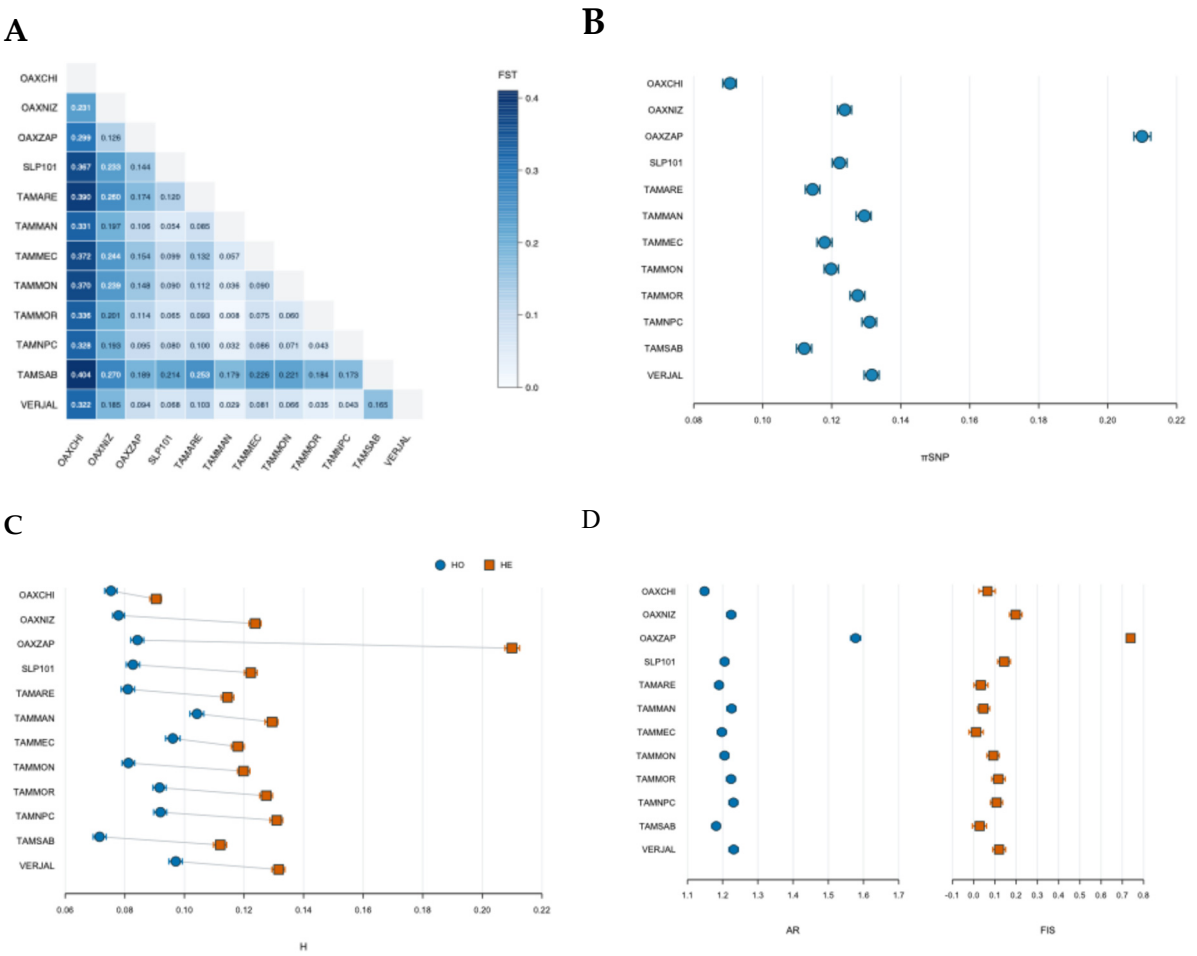
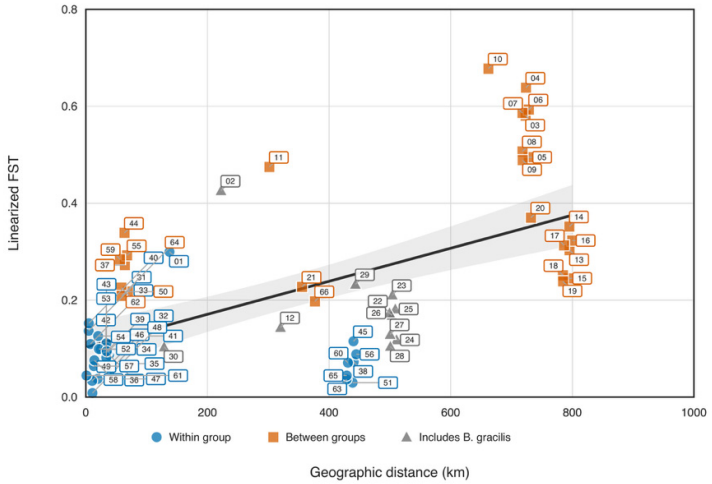


Figure S1. Genetic differentiation and diversity among Beaucarnea populations. A. Pairwise genetic differentiation among the 12 populations estimated using Hudson's F_{ST}; color intensity represents the magnitude of F_{ST}. **B**. Nucleotide diversity conditioned on the SNP panel (πSNP) by population; points represent estimates and bars indicate 95% confidence intervals. **C.** Observed (H_o) and expected (H_e) heterozygosity by population; lines connect the two estimates and bars indicate 95% confidence intervals. **D**. Rarefied allelic richness (AR) and inbreeding coefficient F_{IS} by population; bars indicate 95% confidence intervals.

A



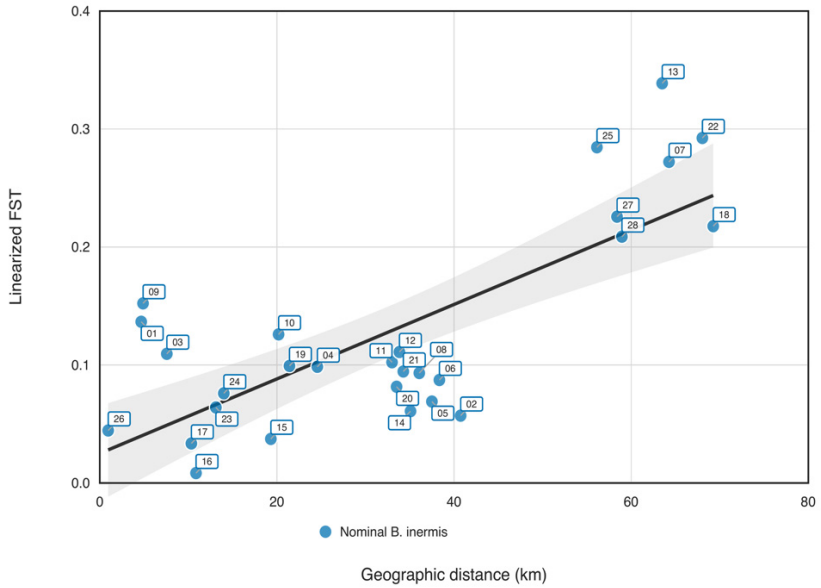
Population codes and localities

OAXCHI = Los Chimalapas
OAXNIZ = Nizanda
OAXZAP = Zapotitlan de las Salinas
SLP101 = Ciudad del Maíz
TAMARE = El Naranjo
TAMMAN = Ciudad Mante
TAMMEC = El Meco
TAMMON = Emiliano Zapata
TAMMOR = Antiguo Morelos
TAMNPC = Antiguo Morelos
TAMSAB = Ciudad Valles
VERJAL = Jalcomulco

Point ID = population pair

- 01 OAXCHI-OAXNIZ
- 02 OAXCHI-OAXZAP
- 03 OAXCHI-SLP101
- 04 OAXCHI-TAMARE
- 05 OAXCHI-TAMMAN
- 06 OAXCHI-TAMMEC
- 07 OAXCHI-TAMMON
- 08 OAXCHI-TAMMOR
- 09 OAXCHI-TAMNPC
- 10 OAXCHI-TAMSAB
- 11 OAXCHI-VERJAL
- 12 OAXNIZ-OAXZAP
- 13 OAXNIZ-SLP101
- 14 OAXNIZ-TAMARE
- 15 OAXNIZ-TAMMAN
- 16 OAXNIZ-TAMMEC
- 17 OAXNIZ-TAMMON
- 18 OAXNIZ-TAMMOR
- 19 OAXNIZ-TAMNPC
- 20 OAXNIZ-TAMSAB
- 21 OAXNIZ-VERJAL
- 22 OAXZAP-SLP101
- 23 OAXZAP-TAMARE
- 24 OAXZAP-TAMMAN
- 25 OAXZAP-TAMMEC
- 26 OAXZAP-TAMMON
- 27 OAXZAP-TAMMOR
- 28 OAXZAP-TAMNPC
- 29 OAXZAP-TAMSAB
- 30 OAXZAP-VERJAL
- 31 SLP101-TAMARE
- 32 SLP101-TAMMAN
- 33 SLP101-TAMMEC
- 34 SLP101-TAMMON
- 35 SLP101-TAMMOR
- 36 SLP101-TAMNPC
- 37 SLP101-TAMSAB
- 38 SLP101-VERJAL
- 39 TAMARE-TAMMAN
- 40 TAMARE-TAMMEC
- 41 TAMARE-TAMMON
- 42 TAMARE-TAMMOR
- 43 TAMARE-TAMNPC
- 44 TAMARE-TAMSAB
- 45 TAMARE-VERJAL
- 46 TAMMAN-TAMMEC
- 47 TAMMAN-TAMMON
- 48 TAMMAN-TAMMOR
- 49 TAMMAN-TAMNPC
- 50 TAMMAN-TAMSAB
- 51 TAMMAN-VERJAL
- 52 TAMMEC-TAMMON
- 53 TAMMEC-TAMMOR
- 54 TAMMEC-TAMNPC
- 55 TAMMEC-TAMSAB
- 56 TAMMEC-VERJAL
- 57 TAMMON-TAMMOR
- 58 TAMMON-TAMNPC
- 59 TAMMON-TAMSAB
- 60 TAMMON-VERJAL
- 61 TAMMOR-TAMNPC
- 62 TAMMOR-TAMSAB
- 63 TAMMOR-VERJAL
- 64 TAMNPC-TAMSAB
- 65 TAMNPC-VERJAL
- 66 TAMSAB-VERJAL

B



Population codes and localities

SLP101 = Ciudad del Maíz
TAMARE = El Naranjo
TAMMAN = Ciudad Mante
TAMMEC = El Meco
TAMMON = Emiliano Zapata
TAMMOR = Antiguo Morelos
TAMNPC = Antiguo Morelos
TAMSAB = Ciudad Valles

Point ID = population pair

- 01 SLP101-TAMARE
- 02 SLP101-TAMMAN
- 03 SLP101-TAMMEC
- 04 SLP101-TAMMON
- 05 SLP101-TAMMOR
- 06 SLP101-TAMNPC
- 07 SLP101-TAMSAB
- 08 TAMARE-TAMMAN
- 09 TAMARE-TAMMEC
- 10 TAMARE-TAMMON
- 11 TAMARE-TAMMOR
- 12 TAMARE-TAMNPC
- 13 TAMARE-TAMSAB
- 14 TAMMAN-TAMMEC
- 15 TAMMAN-TAMMON
- 16 TAMMAN-TAMMOR
- 17 TAMMAN-TAMNPC
- 18 TAMMAN-TAMSAB
- 19 TAMMEC-TAMMON
- 20 TAMMEC-TAMMOR
- 21 TAMMEC-TAMNPC
- 22 TAMMEC-TAMSAB
- 23 TAMMON-TAMMOR
- 24 TAMMON-TAMNPC
- 25 TAMMON-TAMSAB
- 26 TAMMOR-TAMNPC
- 27 TAMMOR-TAMSAB
- 28 TAMNPC-TAMSAB

Figure S2 (continued)

C

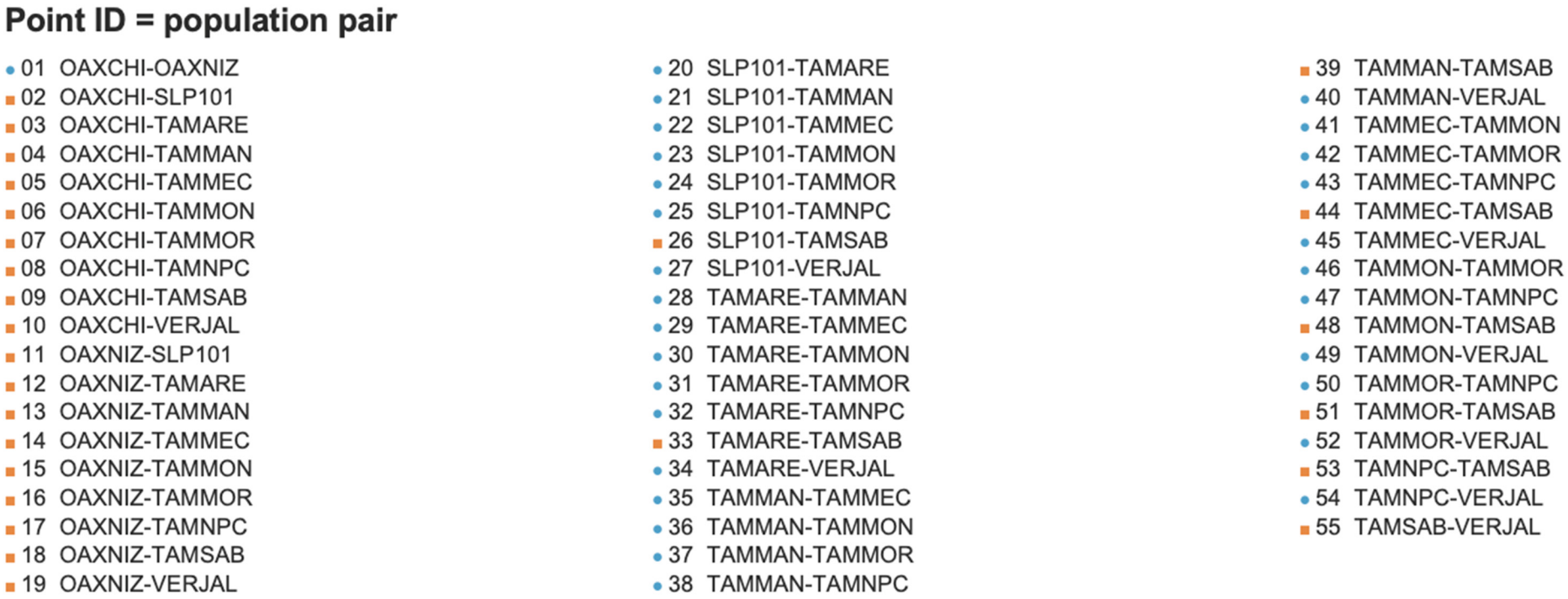


Figure S2. Sensitivity analyses of the association between geographic distance and genetic differentiation. **A.** Global analysis of the 12 populations, including *B. gracilis* as an external taxonomic comparison; symbols distinguish within-component and between-component comparisons and those involving *B. gracilis*. **B.** Analysis restricted to the eight populations nominally assigned to B. inermis. Points represent pairwise comparisons, and lines and shaded areas show the descriptive linear trends and their 95% confidence intervals. **C.** Identification key for the 55 population pairs shown in Figure 3B, comprising the 11 focal populations and excluding OAXZAP (*B. gracilis*). Point identifiers match the numerical labels in Figure 3B. Blue circles indicate within-component comparisons, and orange squares indicate between-component comparisons. Components are defined by the predominant population-level ancestry at K = 3 in the expanded dataset and do not represent taxonomic assignments.

Supplementary Methods S1 Genotypic similarity and sensitivity

We screened the 91 quality-controlled individuals for elevated pairwise similarity using exact genotype concordance and KING-robust kinship estimators (Manichaikul et al., 2010), implemented in Python. Comparisons used only jointly called genotypes, without imputation. Conservative kinship estimates ≥0.0884 were treated as exploratory flags, not diagnoses of clonality or specific familial relationships; this threshold was not calibrated for Beaucarnea.

Seven marker configurations were constructed separately for all individuals, the 85 focal-complex individuals, and the six OAXZAP individuals. Six combined minimum SNP call rates of 0.50, 0.80, or 0.90 with minor allele frequencies of 0.01 or 0.05, selecting one SNP per consensus contig by decreasing call rate and minor allele frequency. The seventh randomly selected one eligible SNP per contig at call rate ≥0.50 and minor allele frequency ≥0.05. The existing 22,510-SNP PCA panel was also examined. These 22 configurations overlap and are not independent replicates. Candidate pairs flagged in either initial full-sample panel (call rate ≥0.50 or ≥0.80; minor allele frequency ≥0.05) were assessed across configurations. Conditional percentile intervals used 500 bootstrap resamples of jointly called contigs per candidate pair and initial panel. Random operations used seed 20261029.

Eight pairs maintained elevated estimates across the seven full-sample configurations (Table S1); five involved only focal individuals. All 32 combinations retaining one individual per focal pair were evaluated, each with 80 individuals in 11 populations. Two additional scenarios considered all eight pairs, retaining either the higher-original-call-rate individual or the opposite member. A common 9,676-SNP subset of the original 13,232-SNP priority panel retained at least six called gene copies per population in every scenario. The complete focal sample was recalculated on the same subset. The two additional scenarios were also independently refiltered under the original call-rate, frequency, and within-contig LD criteria, yielding 11,997 and 12,225 SNPs.

Using the estimators defined in the main Methods, we recalculated allelic diversity and private richness rarefied to six gene copies, observed and expected heterozygosity, FIS, pairwise Hudson FST, nearest-neighbor isolation, the equal-weight four-metric priority index, and leave-one-population-out contributions to total diversity. The 32 scenarios retained the negative diversity–uniqueness association and positive contributions of OAXCHI, OAXNIZ, and TAMSAB. Priority ranks below the first two were sensitive to individual and marker selection (Table S2). Scenario ranges are not confidence intervals, and no new significance tests were performed for conservation metrics. These FIS estimates concern the priority panel, not the separate unfiltered diversity panel. The assessment did not establish genetic independence or resolve biological versus technical causes of similarity. PCA, ancestry, network, SNAPP, and spatial analyses were not repeated under these exclusions.

Reference: Manichaikul et al. (2010), Bioinformatics 26:2867–2873. https://doi.org/10.1093/bioinformatics/btq559

Table S1 Pairs with persistent elevated similarity

Conservative KING estimates for the one-SNP-per-contig panels with minor allele frequency ≥0.05. Persistence denotes exceeding 0.0884 across the full-sample configurations; estimates varied with filtering and do not confirm clones or pedigree categories.

Individual 1	Individual 2	φc at call ≥0.50	φc at call ≥0.80	φc at call ≥0.90
OAXCHI03	OAXCHI07	0.2563	0.3794	0.4258
OAXNIZ05	OAXZAP04	0.2529	0.3820	0.4154
OAXZAP03	OAXZAP06	0.3224	0.3592	0.4485
OAXZAP05	TAMNPC03	0.2656	0.4099	0.4697
TAMMAN02	TAMMOR01	0.3374	0.4442	0.4802
TAMMAN04	TAMMON07	0.3120	0.4322	0.4682
TAMMAN06	TAMNPC08	0.2845	0.4319	0.4643
TAMMAN07	TAMMOR04	0.3699	0.4542	0.4803

Table S2 Summary of the 32 focal-pair exclusion scenarios

Reference and exclusions use the same 9,676 SNPs. Maximum absolute changes are relative to the complete focal sample on that panel; they are not significance tests. Scenario counts are not probabilities.

Metric or result	Sensitivity outcome
Individuals / focal sampling sites	80 / 11 in each scenario; reference 85 / 11
Diversity–uniqueness Pearson r	−0.919 to −0.906; reference −0.914
Rarefied diversity maximum absolute change	0.0276 alleles per SNP (TAMMOR)
Private richness maximum absolute change	0.5199 per 1,000 SNPs (TAMSAB)
FIS maximum absolute change	0.0432 (TAMMAN)
FST maximum absolute change	0.0296 (TAMMAN–TAMMOR)
Priority ranks 1 and 2	OAXNIZ and TAMSAB in all 32 scenarios
Priority rank 3	SLP101 in 19 scenarios; TAMNPC in 13
Positive total-diversity contributors	OAXCHI, OAXNIZ, and TAMSAB in all scenarios

The two additional eight-pair scenarios also retained positive contributions of OAXCHI, OAXNIZ, and TAMSAB.

Supplementary Methods S2 Sensitivity to the inclusion of OAXZAP

OAXZAP (Beaucarnea gracilis) was treated as an external taxonomic comparison. To distinguish its influence from structure within the focal complex, PCA, ancestry inference, and the genetic-affinity network were recalculated after excluding its nine individuals from the 99-individual input matrix. Subsequent sample call-rate filtering (≥0.50) retained 85 individuals across 11 focal populations. No additional individuals were excluded on the basis of the separate pairwise-similarity assessment.

SNP filtering was repeated within the focal sample using the original criteria: call rate ≥0.50, minor allele frequency ≥0.05, and greedy within-contig LD pruning at r^2 >0.20 with at least 20 paired calls, prioritizing higher call rate and then minor allele frequency. The resulting panel contained 22,021 SNPs, compared with 22,510 in the expanded dataset. Thus, the comparison evaluates exclusion together with the accompanying refiltering, rather than holding the marker panel fixed.

PCA used mean-dosage imputation (2p), frequency standardization, and eigendecomposition as in the main analysis. Only the same 85 focal individuals were used to compare their positions in the expanded and focal ordinations. We calculated the correlation between pairwise Euclidean distances in the first two axes. A descriptive Procrustes comparison centered each configuration, normalized it to unit Frobenius norm, and aligned configurations by singular-value decomposition, allowing rotation and reflection. These comparisons were not significance tests.

Ancestry inference reused the supplied expectation-maximization implementation of the binomial admixture model, not the STRUCTURE software. K = 1–8 was evaluated on 6,000 evenly spaced retained SNPs, masking 5% of observed genotypes. Three fits per K (one for K = 1) were compared by validation cross-entropy and the one-standard-error rule. Five final fits used the full focal panel; replicate component labels were aligned before averaging. The original settings were retained: 200 iterations for cross-validation, 300 for final fitting, tolerance 10^{-5}, patience five, and seed 20260826. Because two K = 3 cross-validation fits reached the iteration limit, all three K = 3 fits were repeated with the same seeds and masks and a 600-iteration limit. This was a targeted convergence check, not an extended search across all K values.

Pairwise Hudson FST and 1,000 consensus-contig bootstrap replicates were recalculated for the 55 focal pairs. The network was reconstructed using the lowest FST quartile plus minimum-spanning-tree links; weights remained exp[−d/median(d)], where d = FST/(1−FST). Edge selection and node properties were recomputed rather than simply deleting OAXZAP from the existing graph. Original ancestry proportions retained in network node files are contextual annotations and do not enter edge selection or centrality calculations.

Full results and scripts are supplied in OAXZAP_supporting_data.zip. Python with NumPy and pandas was used. The conclusions are conditional on the supplied model, filtering and network rules. SNAPP, environmental models, and geographic-distance analyses were not repeated in this comparison. No fixed number of species, migration rate, or functional connectivity is inferred from K or network centrality.

Table S3 Comparison of expanded and focal analyses

Analysis or metric	Expanded dataset including OAXZAP	Focal complex excluding OAXZAP
Retained individuals / sampling sites	91 / 12	85 / 11
SNPs after call rate, frequency and LD filtering	22,510	22,021
PCA variance explained by PC1	8.79%	9.30%
PCA variance explained by PC2	3.90%	3.90%
Selected ancestry K	3	2
Network edges / density	22 / 0.3333	18 / 0.3273
Lowest-quartile FST threshold	0.08218	0.07258
Communities / weighted modularity	2 / 0.07343	3 / 0.06147
Highest weighted strength	TAMMAN	TAMMAN
Highest betweenness centrality	VERJAL	TAMMAN
Articulation populations	OAXNIZ, OAXZAP, TAMMAN, VERJAL	OAXNIZ, TAMMAN, VERJAL

For matched focal individuals, the correlation of first-two-axis PCA distances was 0.98990 and normalized Procrustes disparity was 0.01673. Across the same 55 focal population pairs, F_{ST} correlation was 0.999979; mean and maximum absolute changes were 0.000615 and 0.002012. The expanded graph contained 20 edges between focal populations: 17 were retained, three were removed, and one new focal edge was added after reconstruction.

Table S4 Targeted K = 3 convergence check without OAXZAP

Replicate	Iterations to convergence	Converged	Validation cross-entropy
1	184	True	0.422833
2	292	True	0.423755
3	410	True	0.424759

All extended K = 3 fits converged. Their mean validation cross-entropy was 0.423782, compared with 0.419410 for K = 2 (lower is better). This targeted check did not change the preference for K = 2. All five final K = 2 fits converged. Component labels at different K values are not directly equivalent. The broad PCA arrangement was preserved, while the selected K and some network roles depended on the sampled taxonomic scope.

Scores were calculated using the focal conservation panel of 13,232 SNPs from 85 individuals, excluding OAXZAP. Allelic diversity and private allelic richness were standardized to six gene copies. Each metric was ranked in ascending order and normalized as (rank − 1)/(N − 1), with N = 11 and average ranks for ties. Higher values of all four metrics, including FIS, received higher scores. Genetic isolation was defined as the minimum pairwise Hudson's FST to another focal population. The composite index is the arithmetic mean of the four scores, each weighted 0.25. Equal composite indices were ordered by descending private allelic richness score, then by the descending mean of the FIS and isolation scores, and finally by population code. Tiers correspond to ranks 1–3 (I), 4–6 (II), 7–9 (III), and 10–11 (IV). Scores represent relative ranks, not raw metric values; zero does not indicate absence of an attribute. FIS describes heterozygote deficiency and is not interpreted as direct evidence of inbreeding or a validated measure of extinction risk.

Supplementary Methods S3. Ancestry model selection and sensitivity to initialization

Five independent fits per K were completed using all 91 individuals and 22,510 SNPs, without imputation. The EM algorithm from the original deposit was retained, with a tolerance of 10^{-5} per observed allele for five consecutive iterations and a maximum of 5,000 iterations. The validation data were not used as a substitute for the full panel.

Table S5 Numerical convergence and variability among ancestry fits for K = 2 and K = 4.

K	Converged fits	Maximum iterations	Maximum Q RMSE relative to the best fit
2	5/5	147	0.0380
4	5/5	262	0.3405

Numerical convergence does not guarantee a global maximum or identical results across replicates. The individual fits and the best fit are retained alongside the consensus. The figures preserve the individual order used in Figure 1B. Matching colors do not imply correspondence between components at different K values.

Table S6. Normalized component scores and composite genetic priority for the 11 focal populations.

Population	Allelic diversity score	Private allelic richness score	*FIS* score	Genetic isolation score	Composite index	Rank	Tier
OAXNIZ	0.20	0.80	1.00	0.90	0.7250	1	I
TAMSAB	0.10	0.90	0.80	0.80	0.6500	2	I
SLP101	0.60	0.40	0.90	0.50	0.6000	3	I
TAMARE	0.30	0.70	0.50	0.70	0.5500	4	II
TAMNPC	0.80	0.50	0.60	0.30	0.5500	5	II
OAXCHI	0.00	1.00	0.00	1.00	0.5000	6	II
TAMMEC	0.40	0.60	0.10	0.60	0.4250	7	III
VERJAL	0.90	0.30	0.30	0.20	0.4250	8	III
TAMMON	0.50	0.10	0.70	0.40	0.4250	9	III
TAMMOR	0.70	0.20	0.40	0.05	0.3375	10	IV
TAMMAN	1.00	0.00	0.20	0.05	0.3125	11	IV

Sensitivity to initialization

At K = 4, the maximum RMSE after label alignment was 0.341, compared with 0.038 at K = 2. The K = 4 consensus is not interpreted as a stable four-component structure because it averages different solutions. The highest-likelihood fits are included in the accompanying supporting-data package for comparison; they are not claimed to represent global maxima.

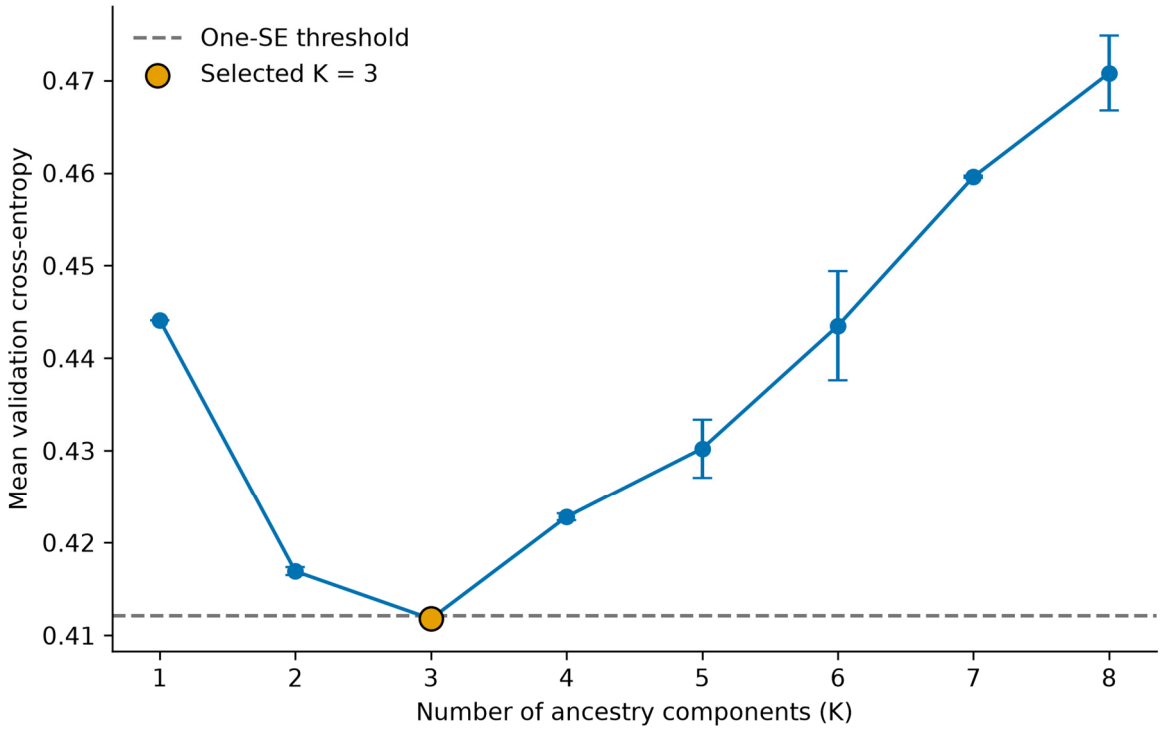


Figure S3. Selection of the number of ancestry components in the expanded interspecific dataset, including OAXZAP (*B. gracilis*). Points show mean validation cross-entropy for K = 1–8 using 6,000 SNPs and a shared mask withholding approximately 5% of observed genotypes. Error bars represent one standard error across three initialization replicates, except K = 1, which used one deterministic fit. They do not represent variation among independent validation folds. All 22 fits met the numerical stopping criterion. The dashed line indicates the one-standard-error threshold; K = 3 was selected.

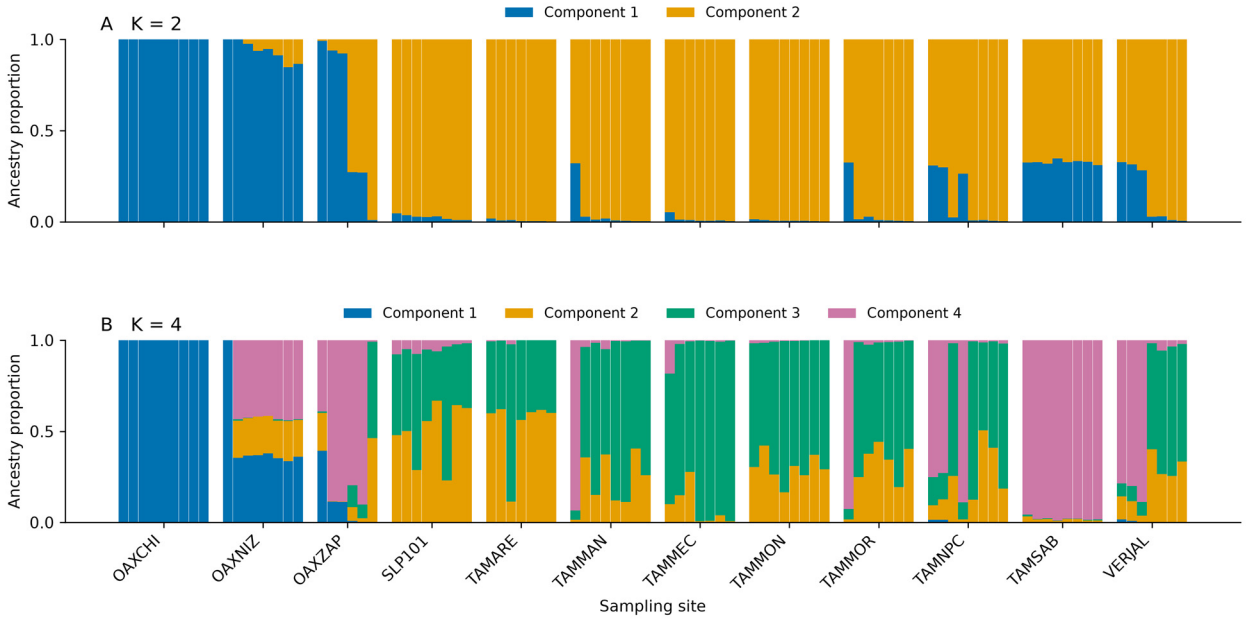


Figure S4. Individual ancestry proportions for K = 2 (A) and K = 4 (B) in the expanded interspecific dataset of 91 individuals from 12 sampling sites, including OAXZAP (*B. gracilis*), using 22,510 SNPs. Each bar represents one individual. Proportions are averages across five independently initialized full-data fits after label alignment to the highest-likelihood fit. All fits met the numerical stopping criterion. Individuals retain the same order as Figure 1B. Components are numbered separately within each K; matching colors across panels do not establish biological equivalence. Components are not taxonomic assignments. K = 4 showed substantial variation among initialized fits (maximum aligned Q RMSE relative to the highest-likelihood fit = 0.341); its averaged proportions therefore also reflect optimization variability. The highest-likelihood fits are included in the accompanying supporting-data package for comparison.

Supplementary Methods S4. Assessment of heterozygote deficiency in OAXZAP

We examined the distribution of missing genotypes and heterozygosity among the six retained OAXZAP individuals and across loci. Diversity statistics were recalculated using the broad (60,061 SNPs) and strict (6,444 SNPs) diversity panels, without genotype imputation. Observed heterozygosity was calculated as the proportion of heterozygous individuals at each locus, using only called genotypes. Expected heterozygosity was calculated from allele frequencies with a correction for the number of sampled gene copies. Population estimates were averaged across loci, and FIS was calculated as $1 - mean\ H_O/mean\ H_E$. Genotype patterns were also inspected to identify loci containing opposite homozygotes without heterozygotes.

The recalculated estimates reproduced πSNP = 0.2099 and $H_O - H_E = -0.1257$ in the broad panel, whereas $F_{IS} = 0.739$ was reproduced in the strict panel. These values therefore describe different marker panels. All six individuals had complete genotypes in the strict panel, where individual observed heterozygosity ranged from 0.050 to 0.077. Consequently, missing genotypes alone do not explain the heterozygote deficit.

Of the 3,720 SNPs polymorphic within OAXZAP in the strict panel, 2,232 contained both homozygous genotypes but no heterozygotes. At 2,017 SNPs distributed across 1,128 consensus contigs, OAXZAP03 and OAXZAP06 shared one homozygous genotype while the remaining four individuals shared the opposite genotype. Thus, both alleles were represented in the pooled sample, increasing expected heterozygosity, although neither allele combination occurred as a heterozygote at those loci.

This pattern is compatible with genetic heterogeneity within the pooled sample and a possible Wahlund effect. However, the two sets of individuals were identified through exploratory inspection of the same genotypes and do not constitute independently established biological subgroups. Complete genotype calls also do not rule out systematic genotyping errors or other technical effects. These results therefore document the pattern underlying the deficit but do not establish its biological or technical cause, or demonstrate interspecific hybridization.

Supplementary Methods S5. SNAPP sensitivity analyses

We compared marker selection, marker number and the number of diploid representatives per population. Several broad groupings recur across the 1,000-SNP panels, but incomplete MCMC mixing and changes in internal relationships preclude a general claim of phylogenetic stability. All panels include the same 12 population labels, including OAXZAP.

Panel construction and diagnostics

The original panel selected the highest-call-rate eligible individual per population (call rate ≥0.50; ties resolved by sample identifier). Among SNPs completely genotyped in these 12 individuals with minor allele frequency ≥0.05, one SNP per contig was retained by highest minor allele frequency, then lowest position and variant identifier. Candidates were ordered by a SHA256 key using seed 20260826 and variant identifier; the first 100 were retained. This preferential retention of common alleles changes the sampled frequency spectrum. One SNP per contig reduces within-contig linkage but does not establish independence among contigs.

The original XML used `non-polymorphic="true"` despite containing only variable sites. A separate comparison retained its individuals and 100 SNPs but used `non-polymorphic="false".` Additional panels selected one or two highest-call-rate individuals per population and required complete, polymorphic genotypes in that subset. SNPs within contigs and contigs between candidates were ordered by reproducible hash keys with seed 20260826, without an additional within-contig preference for high minor allele frequency. The 100-SNP panel is nested in the 1,000-SNP panel with one individual. Eligible contigs numbered 3,843 with one individual and 1,908 with two. All additional panels used `non-polymorphic="false";` the underlying source-panel filtering remains an ascertainment limitation.

Four independently initialized chains were run per panel, recording every 10 iterations and discarding the first 20% of stored samples per chain. The 1,000-SNP panel with one individual was extended from 170,000 to 340,000 total iterations per chain; other panels reached 170,000. Diagnostics comprised autocorrelation-based ESS per chain and their sum, classical split-$\hat{R}$, trace inspection and clade frequencies by chain. Four consecutive, approximately equal windows of retained samples assessed temporal stability. Internal node-indexed parameters were excluded from cross-chain $\hat{R}$ because their identities change with topology. ESS sum ≥100 and split-$\hat{R}$ ≤1.05 were screening criteria, not sufficient evidence of topological convergence. No pilot runs were included. Trees were not constrained or rerooted to place OAXZAP outside the focal complex.

Table SN1. Scalar and topological diagnostics. Min ESS is the minimum summed ESS across posterior, likelihood, prior, lambda, tree height and tree length. Δclade is the maximum between-chain clade-frequency difference. O = original 100 SNPs; C = corrected original panel; R100 = random 100 SNPs with one individual; R1000 = random 1,000 SNPs with one individual; R2 = random 1,000 SNPs with two individuals.

Panel	Iterations per chain	Min ESS sum	Max split-$\hat{R}$	Max Δclade
O	170,000	101.42	1.033	0.151
C	170,000	477.45	1.011	0.156
R100	170,000	491.72	1.009	0.155
R1000 initial	170,000	54.73	1.226	0.198
R1000	340,000	102.54	1.025	0.241
R2	170,000	83.55	1.140	0.284

Convergence and clade sensitivity

The two-individual panel retained 54,404 posterior trees and failed the declared scalar screening criteria. Individual-chain ESS values remained low, and visual trace inspection showed prolonged fluctuations and shifts. The extended one-individual panel passed the scalar criteria (minimum summed ESS 102.54; maximum split-$\hat{R}$ 1.025), but clade differences reached 0.241 between chains and 0.358 between windows. Passing scalar criteria therefore did not establish general topological stability.

Table SN2. Scalar diagnostics for the 1,000-SNP panel with two individuals. ESS sums are not a substitute for individual-chain mixing.

Parameter	ESS 1	ESS 2	ESS 3	ESS 4	ESS sum	Split-$\hat{R}$
posterior	64.85	80.34	70.51	30.44	246.14	1.018
likelihood	26.37	14.27	26.03	19.58	86.26	1.118
prior	28.74	15.01	25.22	14.57	83.55	1.098
lambda	170.25	63.60	195.66	143.77	573.28	1.025
treeHeight	36.52	23.08	33.70	24.23	117.54	1.095
tree.length	28.84	14.63	33.45	19.79	96.71	1.140

Table SN3. Selected clade frequencies across completed panels. Values are descriptive frequencies of sampled trees, not stabilized support where convergence is inadequate. R1000 denotes the extension to 340,000 iterations; other columns use 170,000.

Clade	O	C	R100	R1000	R2
Nine-population group[1]	0.854	0.881	0.573	1.000	1.000
OAXCHI + OAXNIZ + OAXZAP	0.234	0.227	0.447	0.999	1.000
Six-population group[2] + VERJAL	0.132	0.107	0.464	1.000	0.154
Six-population group[2]	0.116	0.069	0.039	0.005	1.000
TAMMOR + VERJAL	0.088	0.102	0.038	0.000	0.683
OAXNIZ + OAXZAP	0.272	0.297	0.972	0.535	0.586
TAMMAN + TAMMON	0.129	0.107	0.020	0.447	0.152

[1] SLP101, TAMARE, TAMMAN, TAMMEC, TAMMON, TAMMOR, TAMNPC, TAMSAB and VERJAL. [2] SLP101, TAMARE, TAMMAN, TAMMEC, TAMMON and TAMNPC.

The nine-population group recurred in all retained samples of both 1,000-SNP panels, and the OAXCHI–OAXNIZ–OAXZAP grouping also showed high sampled frequencies. Nevertheless, several internal relationships changed substantially between panels. A constant clade indicator does not rule out topological trapping. Differences cannot be attributed exclusively to sample size: SNP eligibility and composition also changed. The small original marker set is a major limitation and a plausible contributor to diffuse topological support; its contribution cannot be separated quantitatively from ascertainment, sampling, model assumptions and MCMC mixing. The analyses do not establish incomplete lineage sorting, historical gene flow or recent divergence as the causes of uncertainty.

Posterior trace assessment

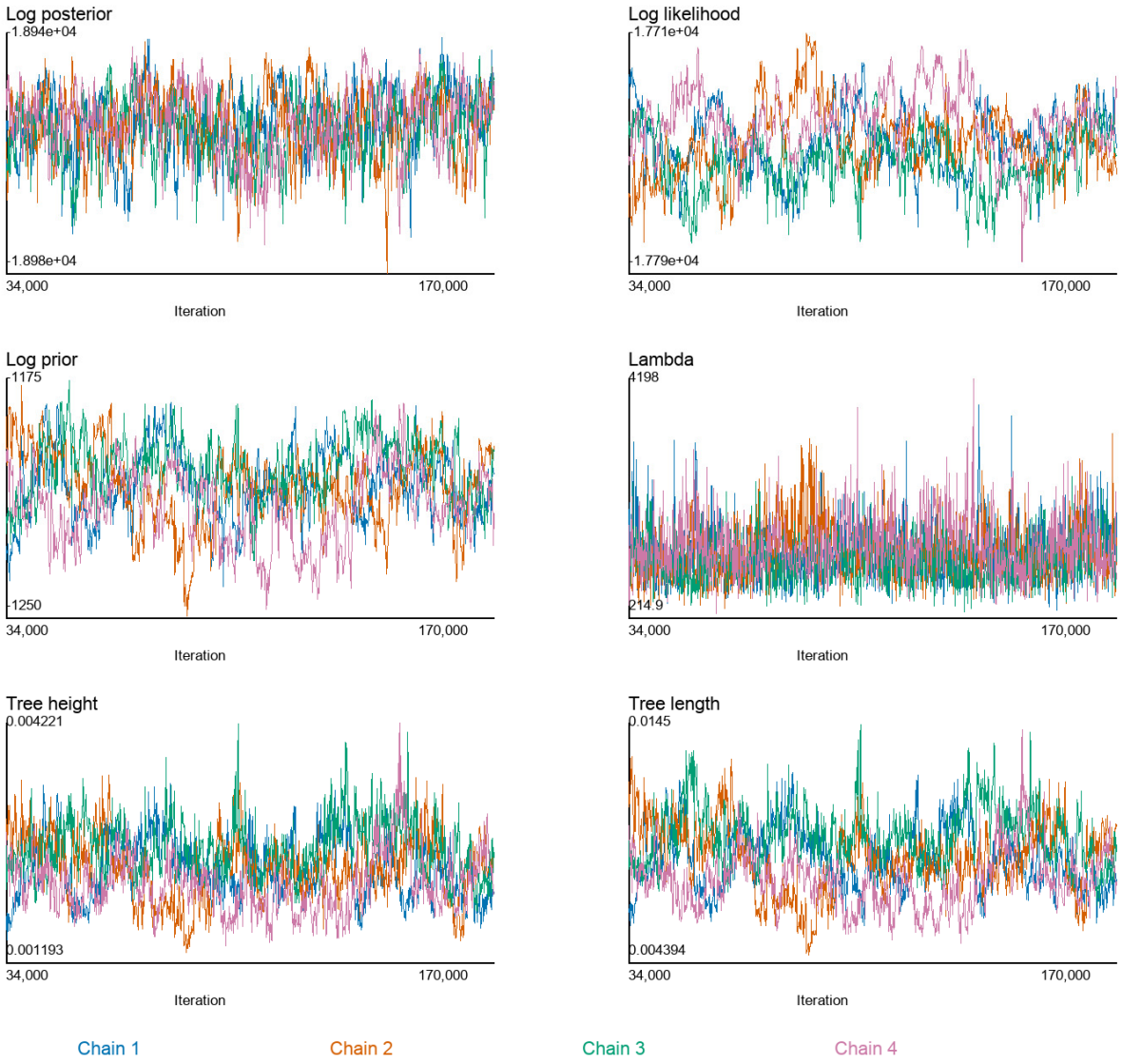


Figure SN1. Traces of six scalar parameters for the four completed chains of the 1,000-SNP panel with two diploid individuals per population. The plotted interval begins after the 20% discard (34,000 iterations). Colors distinguish chains. Every tenth retained logging point is displayed for readability; diagnostics used all retained samples. Similar posterior values alone do not establish adequate mixing in likelihood, prior or tree dimensions.

Temporal variation in clade frequencies

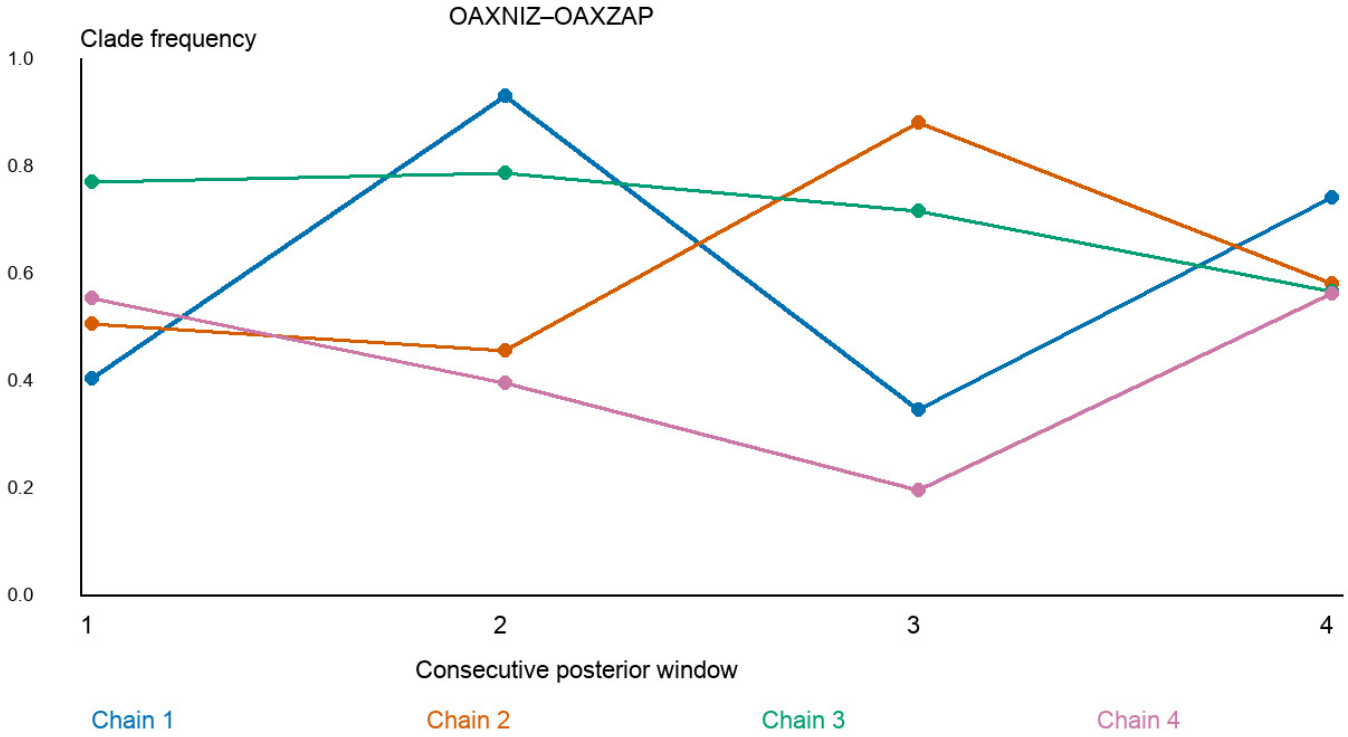


Figure SN2. Frequency of OAXNIZ–OAXZAP in four consecutive windows of the retained posterior samples for each chain of the two-individual panel. Windows contain 3,400 or 3,401 trees. The maximum within-chain window difference was 0.585; the maximum difference between full retained-chain frequencies was 0.284. This relationship remains unresolved despite a combined sampled frequency of 0.586.

The two-individual panel contained 4,291 distinct sampled topologies, with the most frequent at 0.318%. These counts describe the sampled posterior and do not by themselves demonstrate convergence or identify a biological process. Longer chains with renewed scalar and topological diagnostics are needed before stronger support claims. No further extensions are represented here.

Supporting data

The accompanying data archive provides panel summaries, scalar diagnostics, clade frequencies by chain and window, trace-window means, integrity checks, and sample and SNP selection tables for the comparison panels. Figures and tables use the SN prefix to distinguish the SNAPP sensitivity analyses from the other supplementary materials.

Supplementary Methods S6. Sensitivity of environmental results to occurrence assignments

Methods. We repeated the environmental PCA, directional niche-overlap estimation and distribution modeling under three scenarios: current labels (S0); historical labels for 21 records currently assigned to B. inermis but previously labeled B. recurvata (S1); and exclusion of those records (S2). Historical labels were alternatives, not verified taxonomic corrections. These scenarios did not directly exclude OAXNIZ because occurrence records could not be unambiguously linked to that genetically sampled population.

The procedures in Sections 2.7–2.9 were retained, with raster-cell deduplication and fitting repeated per scenario. PCA and niche analyses used the six bioclimatic variables from the Mexico 1910–2009 grids at 30 arc-seconds. Directional overlap in PC1–PC2 space used 10,000 posterior draws and 10,000 simulations per draw (seeds 20260826 and 20260827). Distribution models used WorldClim 2.1 at 2.5 arc-minutes, 20,000 background cells, five polynomial predictor terms and three geographic validation blocks. Regularization was reselected for each dataset. Suitable area used the 10th percentile of training presences; intervals used 100 bootstrap resamples. Cell-area-weighted Schoener's D compared normalized suitability projections over the shared valid Mexican grid.

Table ES1. Comparison of scenarios. I and R denote nominal B. inermis and B. recurvata datasets. Overlap is the probability of a source-group draw falling within the target's 95% niche region; brackets give 95% credible intervals. Area brackets are 95% bootstrap intervals.

Metric	S0	S1	S2
Input rows (I / R)	40 / 9	19 / 30	19 / 9
PCA and niche cells (I / R)	40 / 7	19 / 28	19 / 7
SDM cells (I / R)	36 / 6	18 / 24	18 / 6
PC1 + PC2 variance (%)	82.53	82.53	84.01
I within R (%)	0.015 [0–0.100]	85.54 [64.79–98.43]	0.050 [0–0.310]
R within I (%)	0.569 [0–5.621]	63.77 [42.55–84.79]	5.02 [0–57.54]
Schoener's D	0.474	0.623	0.459
Suitable area I (km²)	99,531 [71,303–227,550]	174,381 [57,209–419,900]	174,381 [57,209–419,900]
Suitable area R (km²)	34,705 [22,505–39,461]	379,542 [312,296–443,495]	34,705 [22,505–39,461]

Results. Historical assignments substantially increased overlap and altered projections. Exclusion retained low mean overlap but widened uncertainty. S0 and S1 contained identical observations with different group labels, explaining their equal PCA variance; S2 recalculated the PCA on a subset. Identical *B. recurvata* models in S0 and S2 reflect unchanged inputs, not independent validation. Environmental separation therefore depends on record composition and assignment and cannot establish differentiation between independently verified species or local adaptation.

Reproducibility. The accompanying archive retains record-level assignments, maps, full outputs, inputs and scripts, including analysis/reproducir_escenarios.py. Small samples and uncertain labels remain limitations. Reported intervals are conditional on each scenario. Numerical values are from the sensitivity runs; no new significance tests were performed.

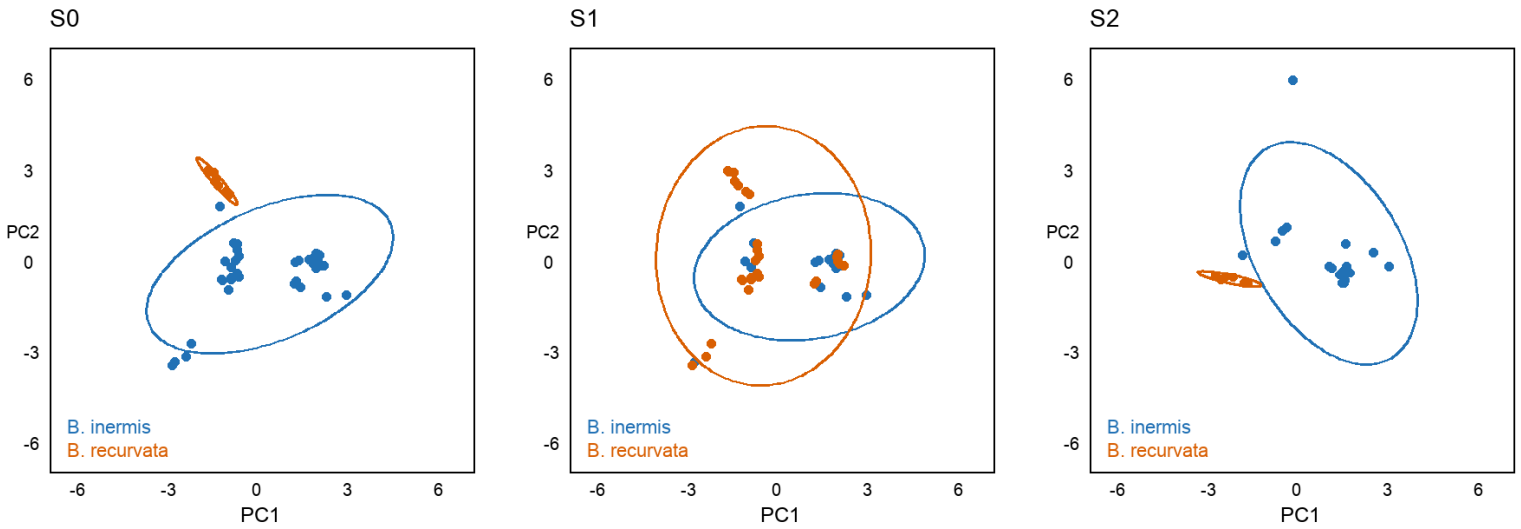


Figure ES1. Environmental PCA under current labels (S0), historical labels for 21 discordant records (S1), and exclusion of those records (S2). Points represent retained occurrence cells; colors indicate nominal species labels. Ellipses represent the estimated 95% distribution in PC1–PC2 space, not posterior niche-overlap estimates. S0 and S1 share the same observations and PCA axes with different group assignments; S2 recalculates the PCA after exclusion.

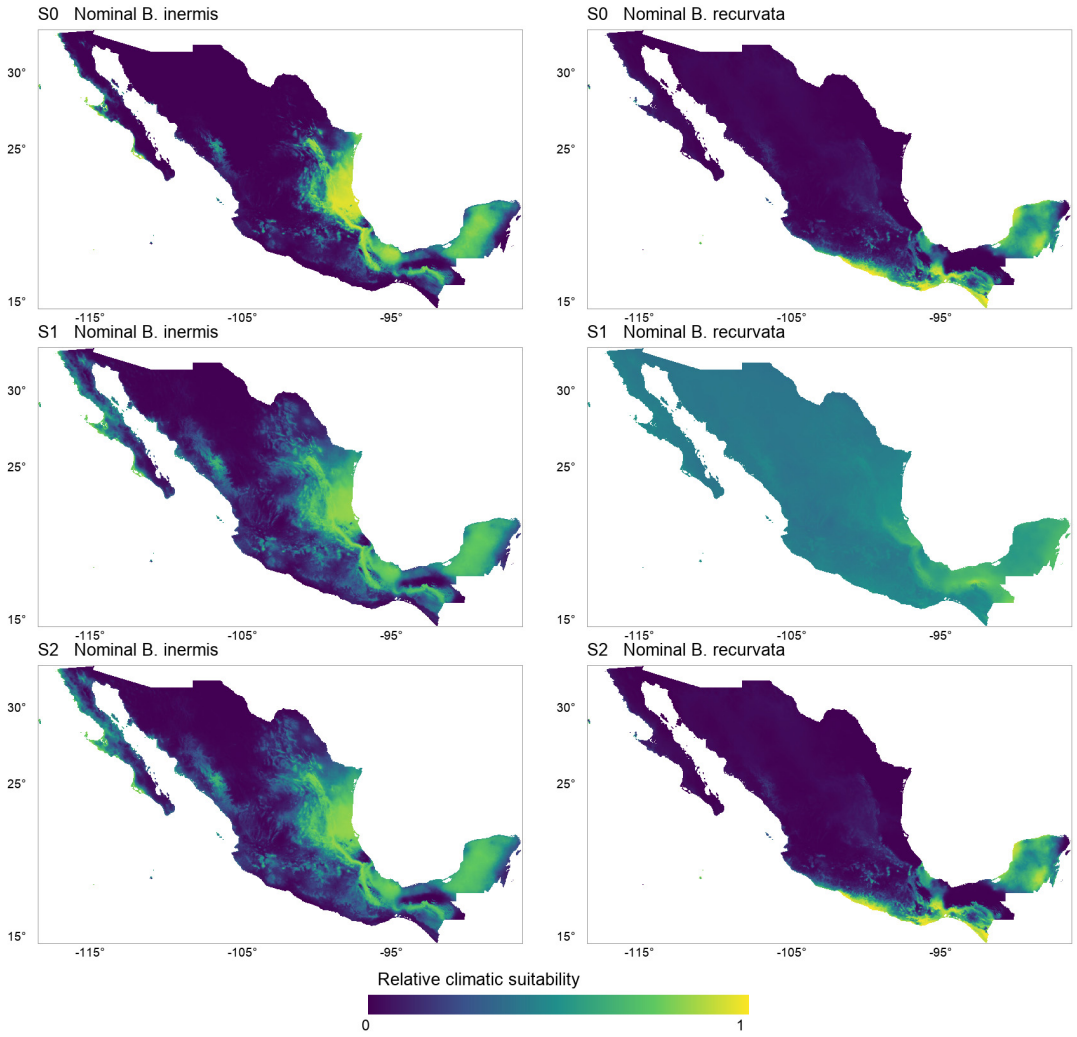


Figure ES2. Relative climatic suitability for the nominal *B. inermis* and *B. recurvata* datasets under S0–S2. All panels use the same 0–1 color scale and Mexican climatic domain; white cells are outside the valid domain. Projections are not occurrence probabilities or confirmed species distributions. Thresholded areas and uncertainty intervals are reported in Table ES1.