PNAS
www.pnas.org

Supplementary Information for

Individualistic evolutionary responses of central African rain forest plants to Pleistocene climatic fluctuations

Andrew J. Helmstetter, Kevin Béthune, Narcisse G. Kamdem, Bonaventure Sonké and Thomas L.P. Couvreur

Andrew J. Helmstetter.
E-mail: andrew.j.helmstetter@gmail.com

This PDF file includes:

Contents

Sampling and sequencing

Family	Genus	Epithet	Samples	75/75	Total length (bp)	SNPs	Loci
Arecaceae	*Podococcus*	*acaulis*	37	119	277199	2565	137
Arecaceae	*Podococcus*	*barteri*	88	120	216143	1918	144
Arecaceae	*Sclerosperma*	*mannii*	129	160	316075	3378	155
Annonaceae	*Annickia*	*affinis*	112	351	756665	5963	262
Annonaceae	*Anonidium*	*mannii*	109	365	766765	5311	293
Annonaceae	*Monanthotaxis*	*enghiana*	105	356	781943	3469	243
Annonaceae	*Greenwayodendron*	*suaveolens*	145	338	650322	5960	312

Table S1. Information on sampling of individuals used for genetic data. 75/75 indicates the number of loci (after paralogs were removed) reconstructed using the hybpiper pipeline for phylogenetic inference in which $\geq$75% of the exon length was recovered in $\geq$75% of individuals. Total length is the sum of the length of all 75/75 supercontigs (exon + introns). Loci indicates the number of loci and SNPs recovered using the SECAPR pipeline. We note that three individuals of *P. acaulis* did not have location data, and so were excluded from analyses with a geographic component.

Species information

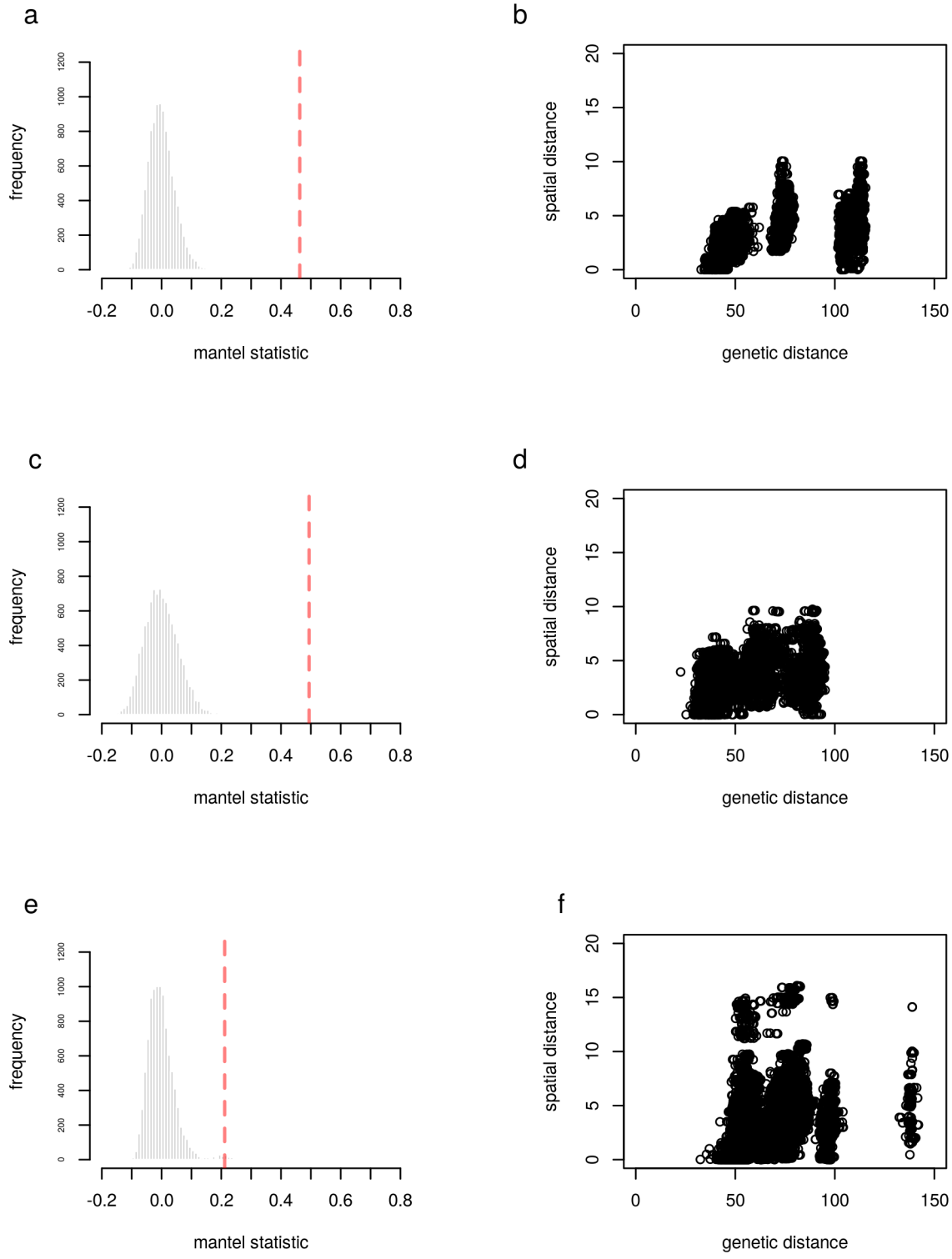
genus	Annickia.affinis	Anonidium.mannii	Greenwayodendron.suaveolens
Habit	Tree	Tree	Tree
Max size (m)	up to 30	8–30	8–45
Forest stratum	canopy	canopy	canopy
Habitat	lowland evergreen rain forest	lowland evergreen rain forest	lowland to premontane evergreen/semi-decidious forest
Soil ecology	terra firma/periodically flooded soils	terra firma	terra firma
Sex	Hermaphrodite	Androdioecious	Androdioecious
Pollination	entomophilous	entomophilous	entomophilous
Number of fruits	3 to 34	1	2 to 8
Number of seeds	1	50–100	1–4
Fruit dimension (cm)	2–3.5 x 1–1.5	25–50 x 10–30	0.7–1.6 x 0.7–1.6
Fruit structure	stipitate, fleshy	stipitate, syncarpous, fleshy	stipitate, mericarp medium hard, fleshy inside
color at maturity	dark purple - black	yellow	green reddish
Dispersal syndrome (primary / secondary)	bird-monkey	monkey, elephants	bird-monkey / elephant
Inferred dispersal	potentially medium to widespread	potentially limited	potentially medium to widespread

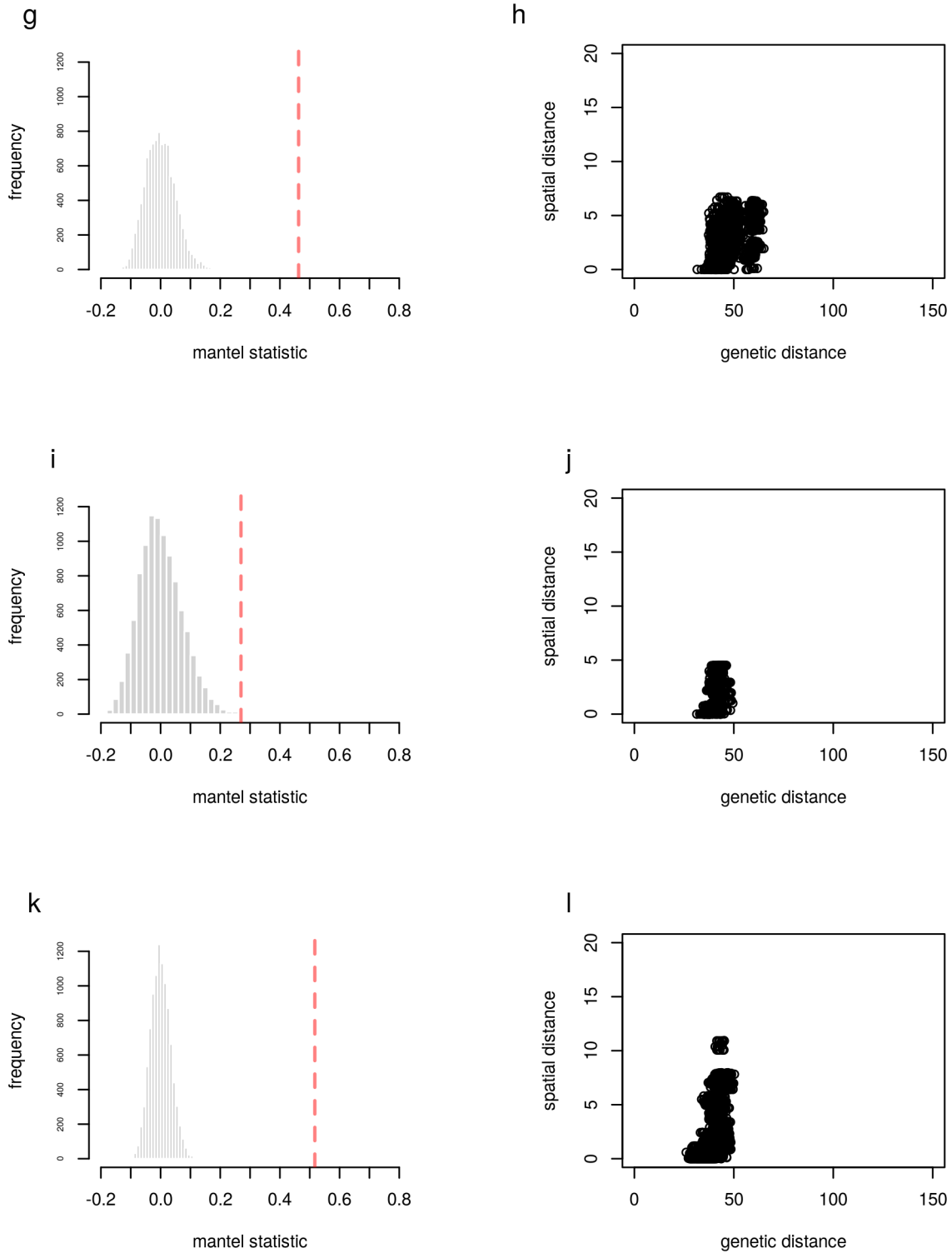
genus	Monanthotaxis.enghiana	Podococcus.acaulis	Podococcus.barteri	Sclerosperma.mannii
Habit	Liana	Shrub	Shrub	Shrub
Max size (m)	up to 15	3–4	2–3	2–5
Forest stratum	understory/canopy	understory	understory	understory
Habitat	lowland to premontane evergreen/semi-decidious forest	lowland evergreen rain forest	lowland evergreen rain forest	lowland evergreen rain forest
Soil ecology	terra firma/periodically flooded soils	terra firma soil	terra firma/periodically flooded soils	periodically flooded soil
Sex	Hermaphrodite	Monoecious	Monoecious	Monoecious
Pollination	entomophilous	entomophilous	entomophilous	entomophilous
Number of fruits	5 to 15	20 to 30	20 to 30	up to 17
Number of seeds	1–4	1– 3	1– 3	1
Fruit dimension (cm)	1.4–3.4 x 8–9	2–3 × 0.5–1	3 x 0.6	2.5–3 x 2–3
Fruit structure	moniliformes, fleshy	sessil, fleshy	sessil, fleshy	sessile, mericarp hard, fleshy inside
color at maturity	dull purple/brown	dark purple	dark purple	brown
Dispersal syndrome (primary / secondary)	bird-monkey	Ruminant-rodent	Ruminant-rodent	Ruminant-rodent / elephants
Inferred dispersal	potentially medium to widespread	potentially limited	potentially limited	potentially limited

Table S2. Information on the taxonomy, morphology, life history and ecology of the seven chosen species in our study.

Mantel tests

We tested the amount of spatial structure in each of our species SNP datasets using Mantel tests. This approaches assesses the correlation between geographic distance and genetic distance, i.e. the presence of isolation-by-distance (IBD). The Mantel statistic is the correlation between two dissimilarity matrices (pairwise geographic and pairwise genetic distance). The genetic matrix consisted of pairwise distances among all individuals for each species. We used the *mantel* function ('pearson' method) in the R package 'vegan' (1) and calculated significance by performing 9999 random permutations . We then compared these to the observed value of the Mantel statistic. We found that all species exhibited significant isolation-by-distance patterns.


a
frequency
mantel statistic
b
spatial distance
genetic distance
c
frequency
mantel statistic
d
spatial distance
genetic distance
e
frequency
mantel statistic
f
spatial distance
genetic distance


g
frequency
0
200
400
600
800
1000
1200
-0.2
0.0
0.2
0.4
0.6
0.8
mantel statistic
h
spatial distance
0
5
10
15
20
0
50
100
150
genetic distance
i
frequency
0
200
400
600
800
1000
1200
-0.2
0.0
0.2
0.4
0.6
0.8
mantel statistic
j
spatial distance
0
5
10
15
20
0
50
100
150
genetic distance
k
frequency
0
200
400
600
800
1000
1200
-0.2
0.0
0.2
0.4
0.6
0.8
mantel statistic
l
spatial distance
0
5
10
15
20
0
50
100
150
genetic distance

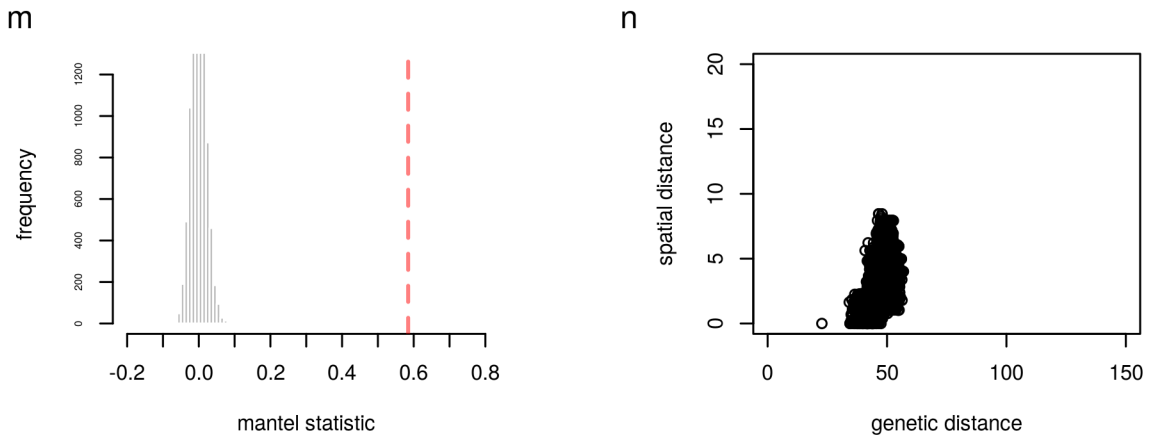


Fig. S1. Mantel tests results for (a-b) *A. affinis*, (c-d) *A. mannii*, (e-f) *G. suaveolens*, (g-h) *M. enghiana*, (i-j) *P. acaulis*, (k-l) *P. barteri* and (m-n) *S. mannii*. Panels on the left show histograms of 9999 random permutations and the observed value as a dashed red line. Panels on the right show plots of genetic distance vs spatial distance.

Genetic clustering including geographic information with TESS

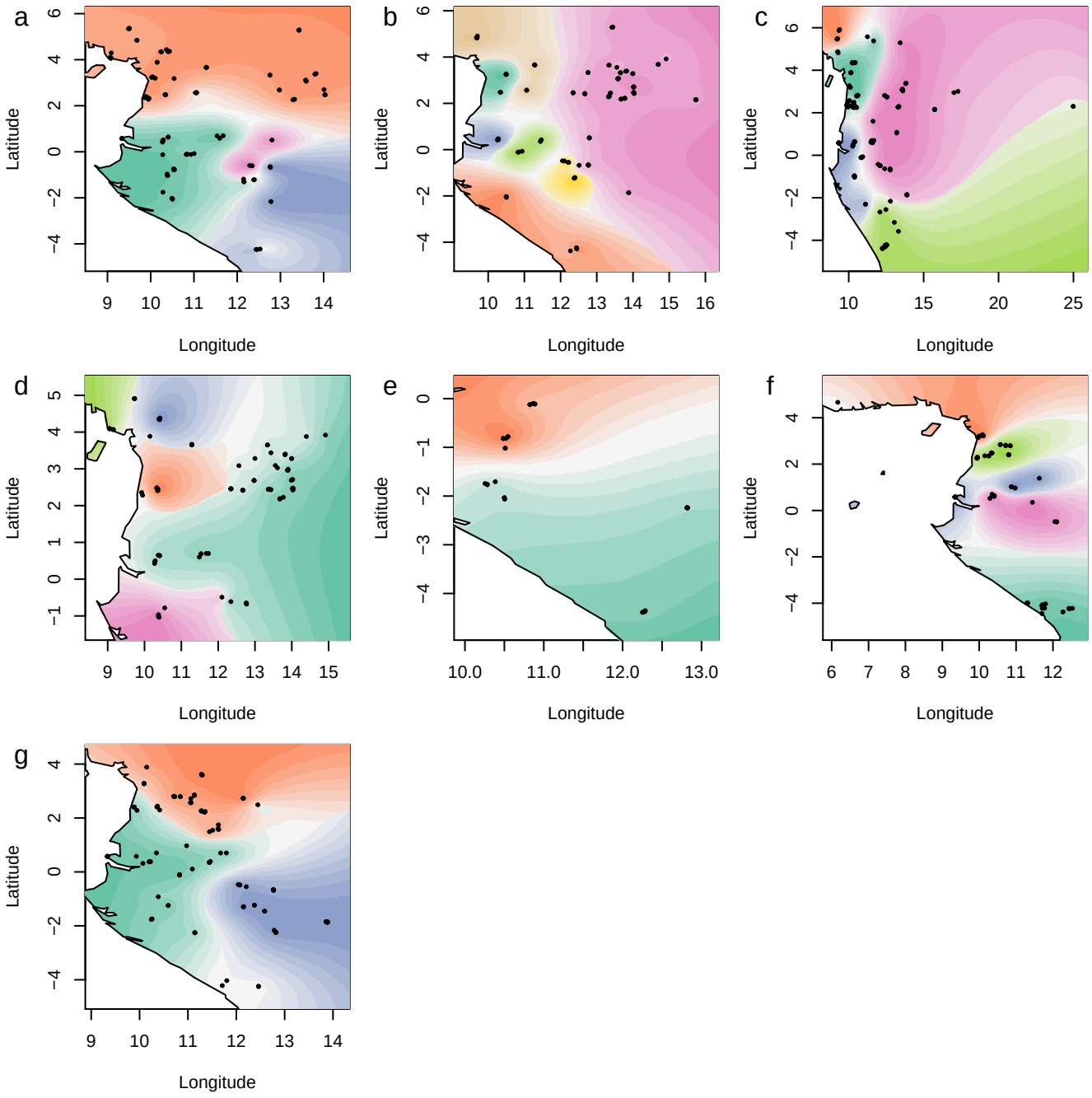
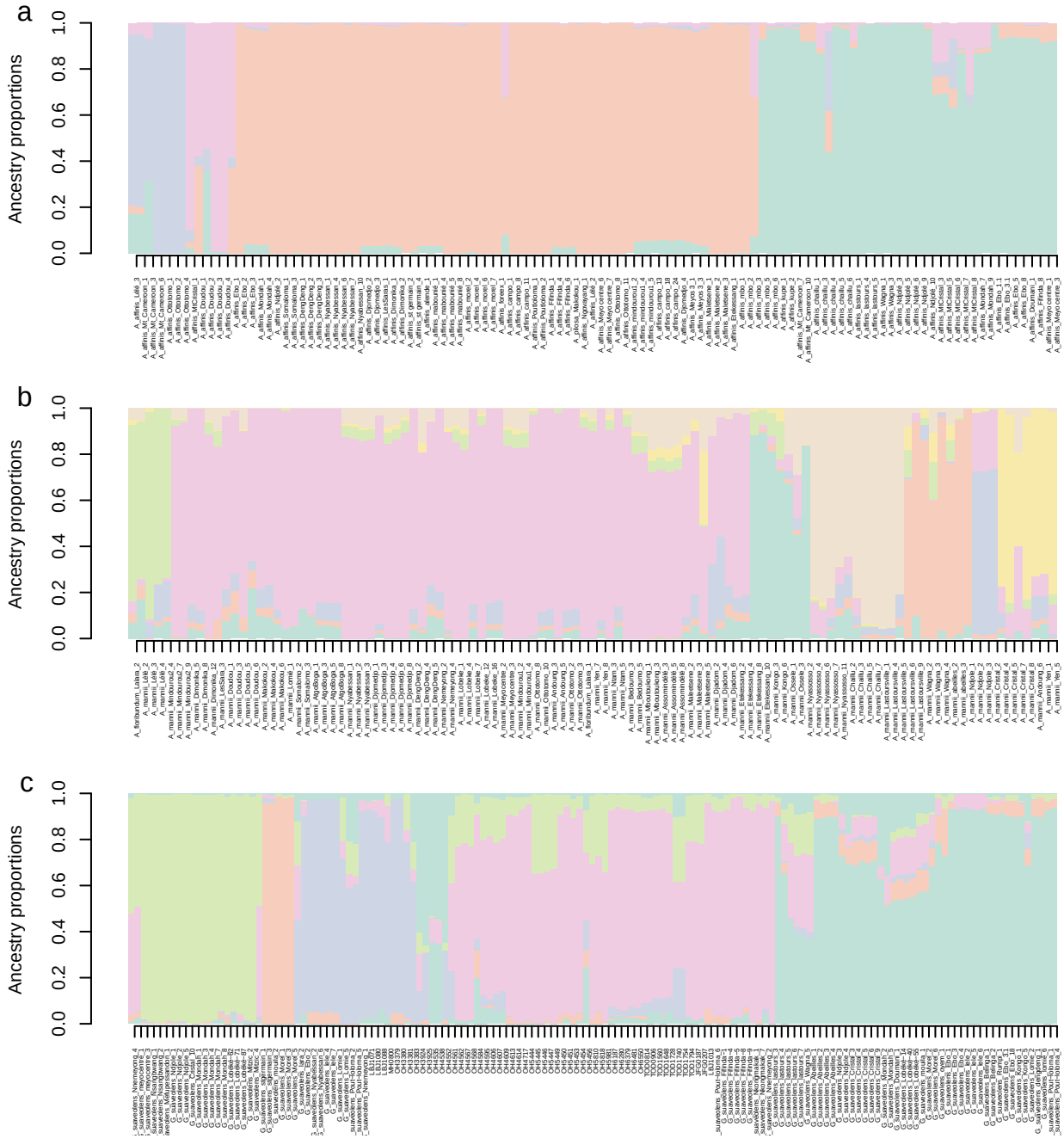
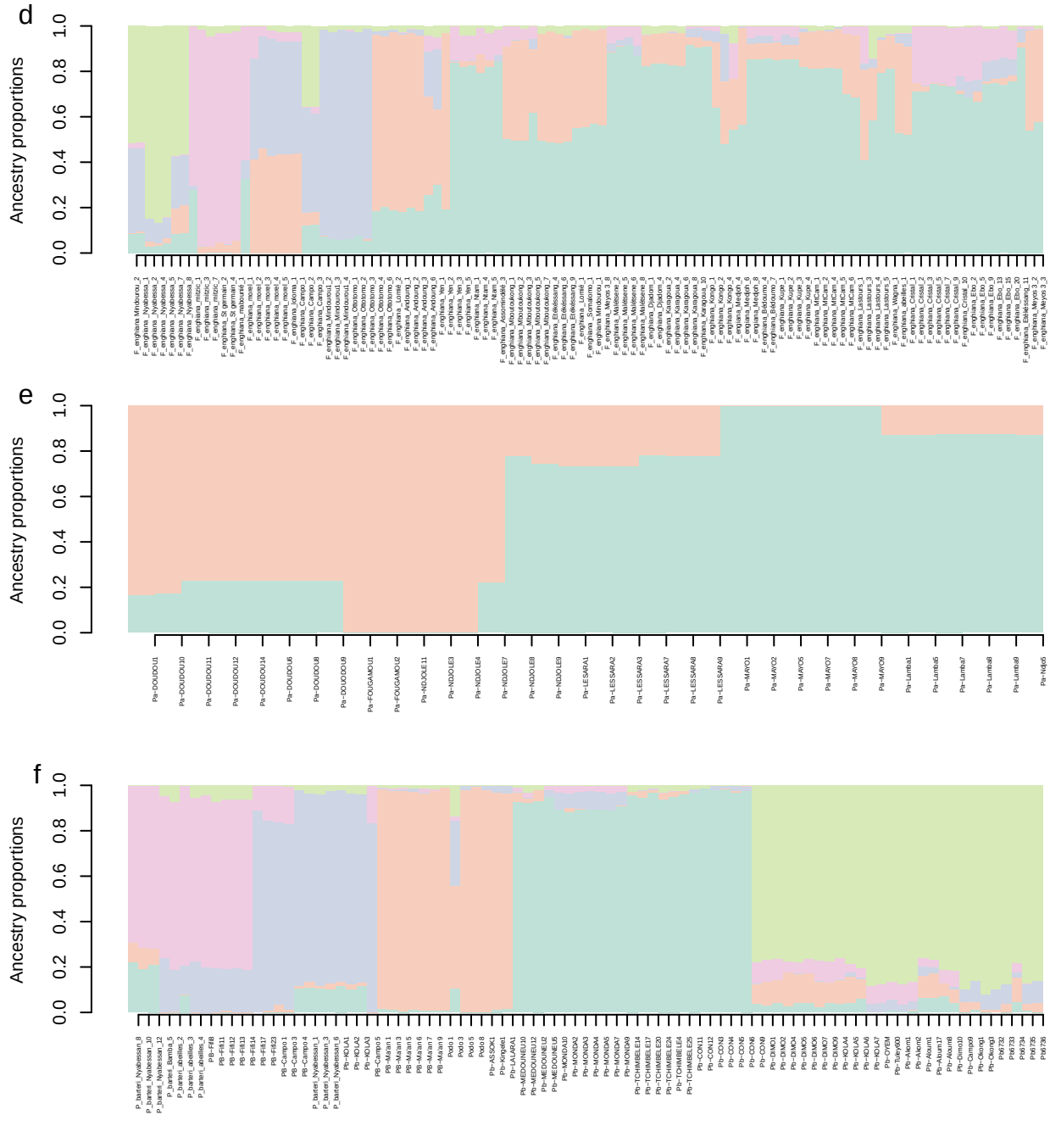


Fig. S2. Maps depicting genetic clusters inferred from the TESS3 analysis for (a) *Annickia affinis*, (b) *Anonidium mannii*, (c) *Greenwayodendron suaveolens*, (d) *Monanthotaxis enghiana*, (e) *Podococcus acaulis*, (f) *Podococcus barteri* and (g) *Sclerosperma mannii*. This approach takes into account geographic information as well as genetic data. Black points represent locations of individuals. Interpolated values of ancestry coefficients are displayed for clusters inferred and the color gradient corresponds to the level of ancestry.


a
b
c
Ancestry proportions
0.0
0.2
0.4
0.6
0.8
1.0


d
e
f
Ancestry proportions
0.0
0.2
0.4
0.6
0.8
1.0

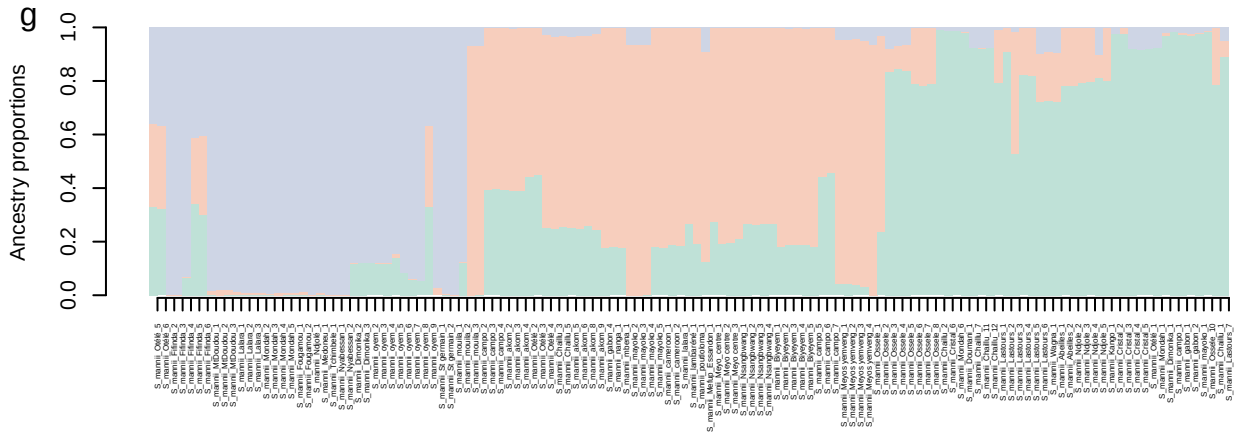
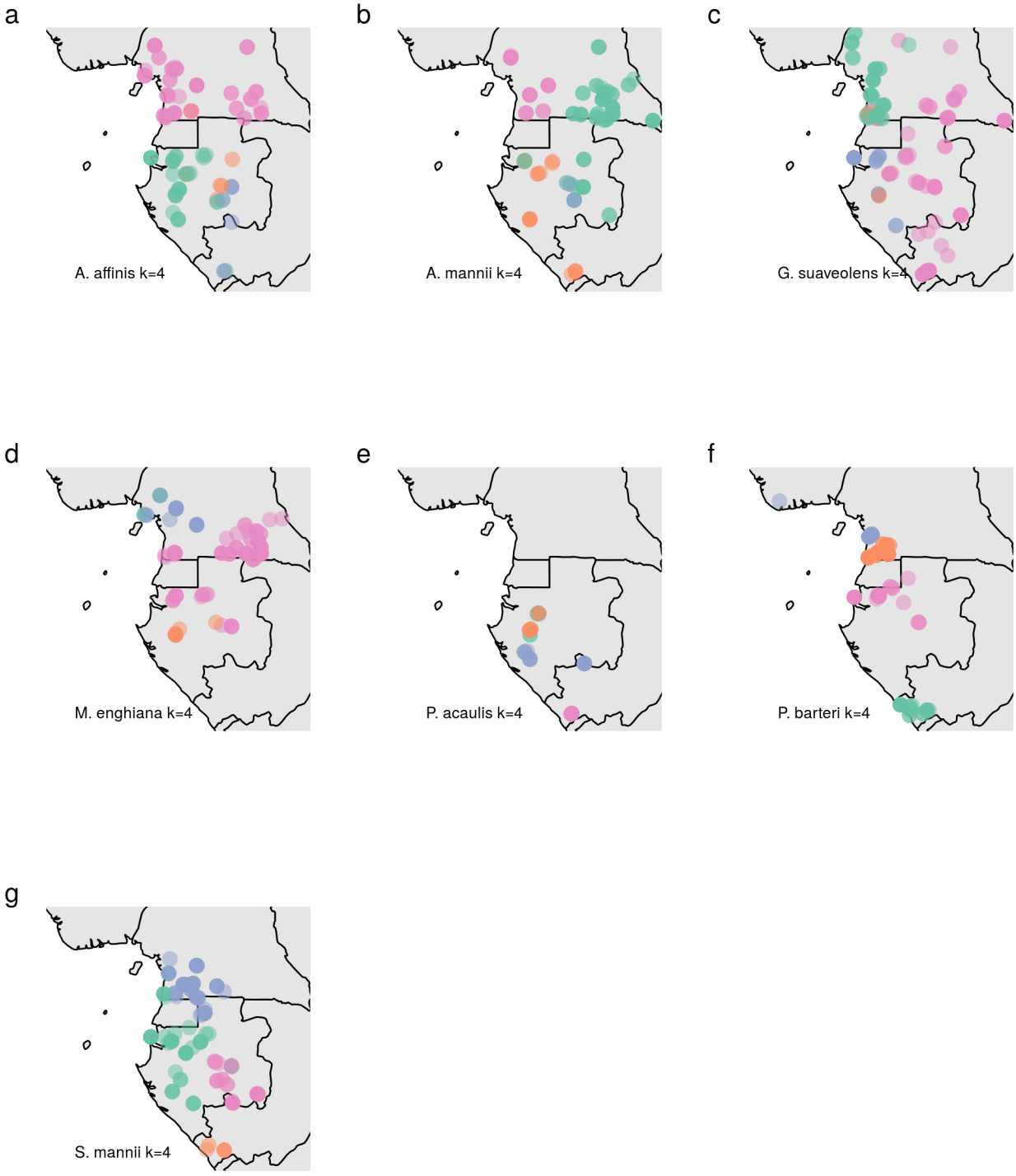


Fig. S3. Barplot of ancestry proportions inferred using TESS3 for (a) *Annickia affinis*, (b) *Anonidium mannii*, (c) *Greenwayodendron suaveolens*, (d) *Monanthotaxis enghiana*, (e) *Podococcus acaulis*, (f) *Podococcus barteri* and (g) *Sclerosperma mannii*. Number of genetic clusters was chosen for each species using the cross-validation store.

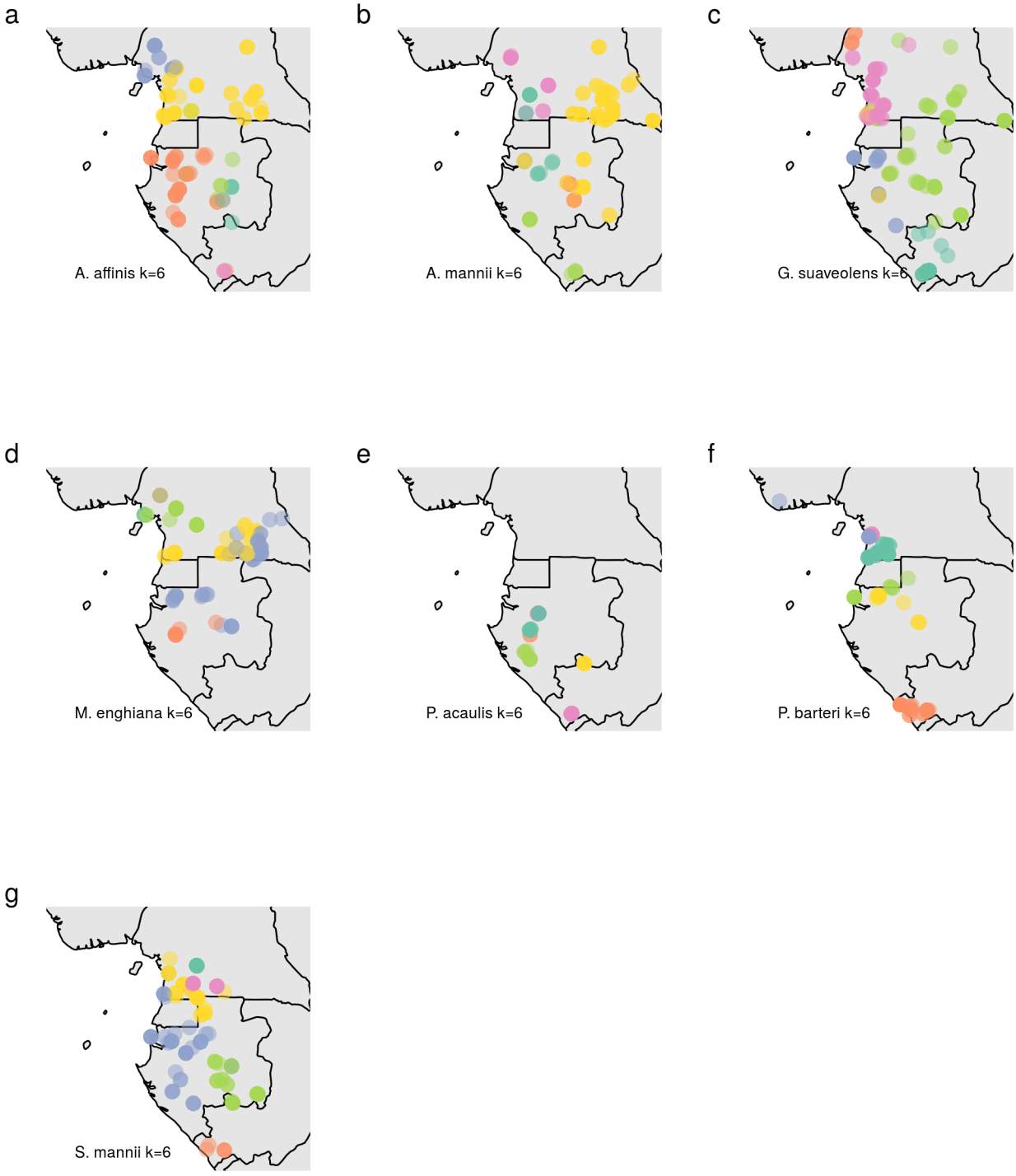
Clustering sensitivity

DAPC analyses were presented based on values of K selected based on BIC. Here we present DAPC results for other values of K (from 2-8) for each species. This was to determine whether changing values of K might lead to similar patterns of population genteic structure across species. Instead we find that structure is generally idiosyncratic, and this does not change as the number of clusters increases.


a
A. affinis k=3
b
A. mannii k=3
c
G. suaveolens k=3
d
M. enghiana k=3
e
P. acaulis k=3
f
P. barteri k=3
g
S. mannii k=3


a
A. affinis k=4
b
A. mannii k=4
c
G. suaveolens k=4
d
M. enghiana k=4
e
P. acaulis k=4
f
P. barteri k=4
g
S. mannii k=4


a
A. affinis k=5
b
A. mannii k=5
c
G. suaveolens k=5
d
M. enghiana k=5
e
P. acaulis k=5
f
P. barteri k=5
g
S. mannii k=5


a
A. affinis k=6
b
A. mannii k=6
c
G. suaveolens k=6
d
M. enghiana k=6
e
P. acaulis k=6
f
P. barteri k=6
g
S. mannii k=6

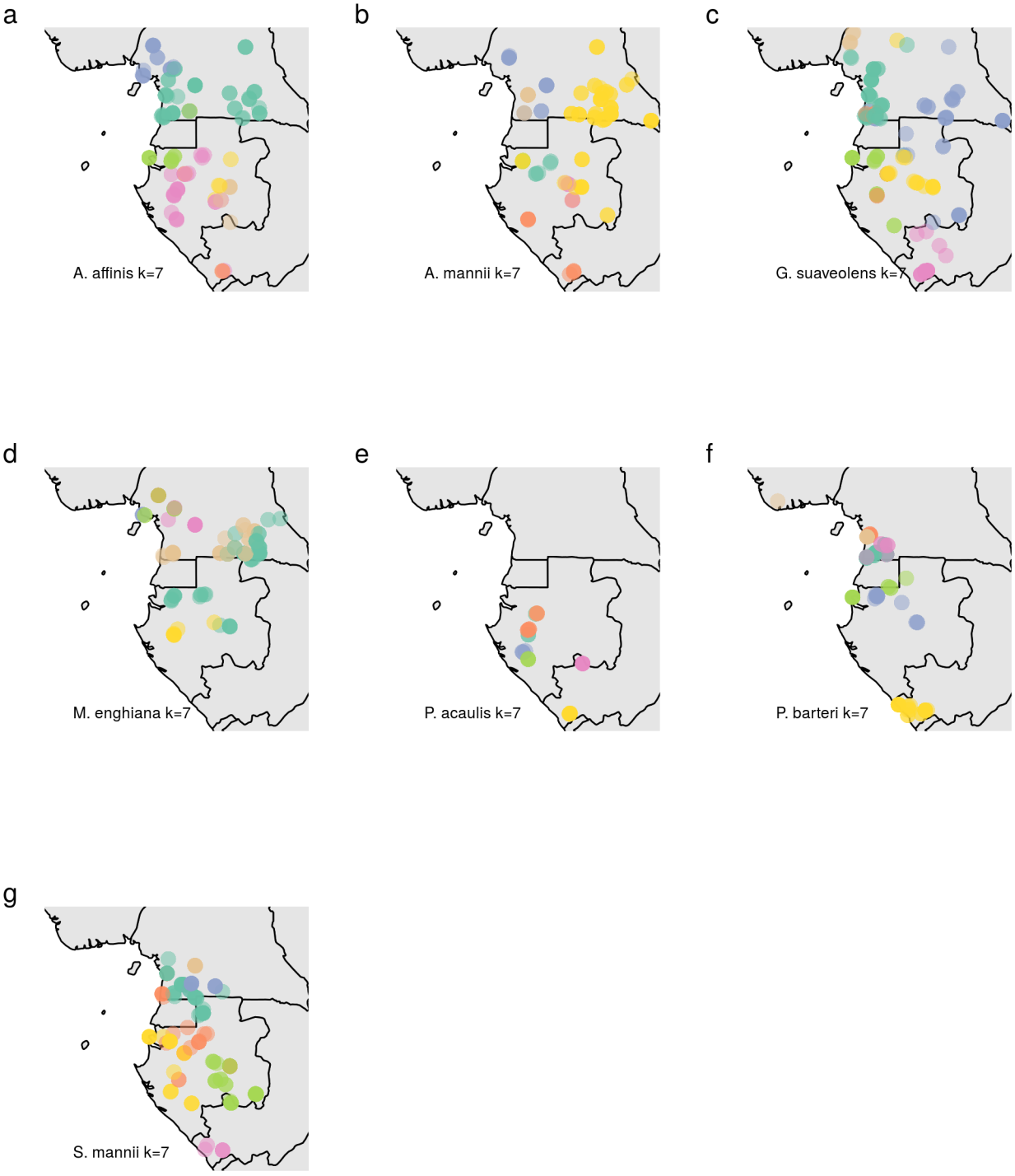

a
A. affinis k=7
b
A. mannii k=7
c
G. suaveolens k=7
d
M. enghiana k=7
e
P. acaulis k=7
f
P. barteri k=7
g
S. mannii k=7

Fig. S4. Assessment of sensitivity of DAPC analyses. We ran DAPC analyses for values of k from 2 (minimum K inferred) to 8 (maximum K inferred) for each species. Analyses were run as detailed in the materials and methods section. Sets of maps are shown for each value of K, and individuals are colored by cluster membership. The name of the species and the value of K are shown in the bottom left of each plot.

Individual-level maximum-likelihood trees

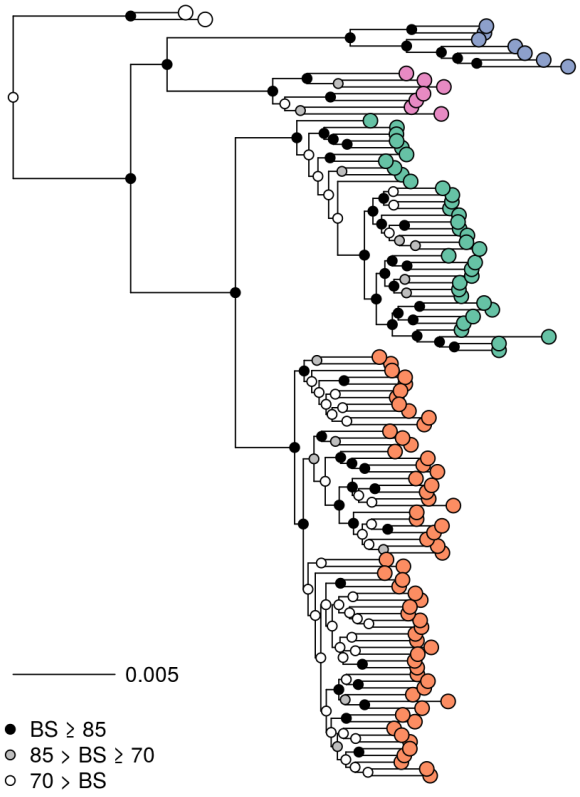


Fig. S5. Maximum likelihood phylogenetic tree of *A. affinis*. Colours on tips correspond to inferred DAPC clusters. Tips with white circles are outgroup taxa. Circles at nodes represent confidence in the preceding branch and show the results of 100 bootstraps inferred using RAxML.

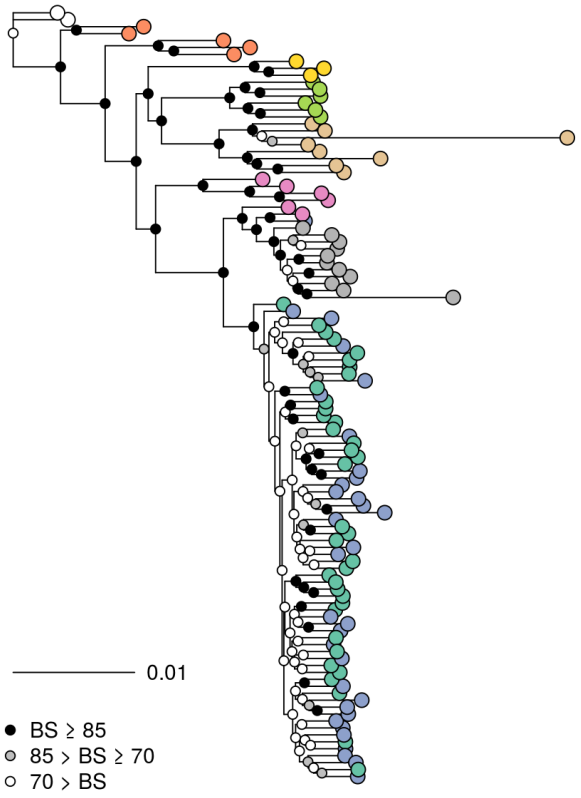


Fig. S6. Maximum likelihood phylogenetic tree of *A. mannii*. Colours on tips correspond to inferred DAPC clusters. Tips with white circles are outgroup taxa. Circles at nodes represent confidence in the preceding branch and show the results of 100 bootstraps inferred using RAxML

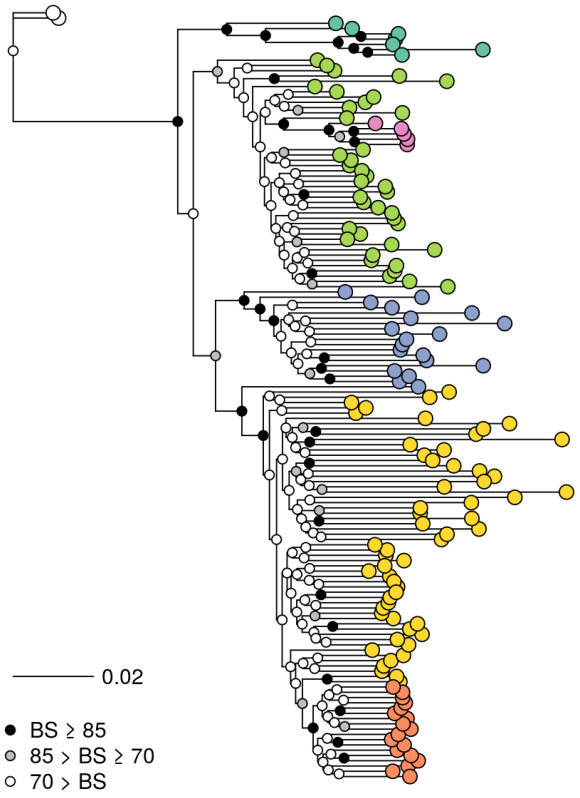


Fig. S7. Maximum likelihood phylogenetic tree of *G. suaveolens*. Colours on tips correspond to inferred DAPC clusters. Tips with white circles are outgroup taxa. Circles at nodes represent confidence in the preceding branch and show the results of 100 bootstraps inferred using RAxML.

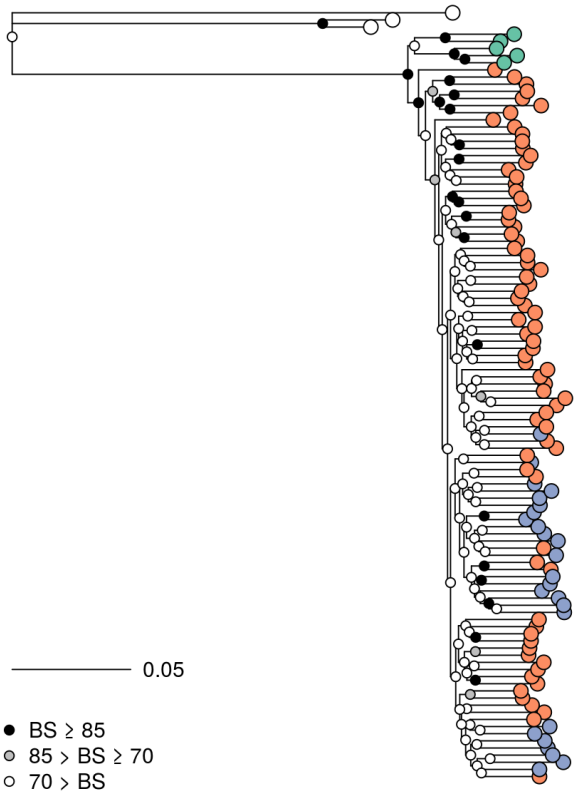


Fig. S8. Maximum likelihood phylogenetic tree of *M. enghiana*. Colours on tips correspond to inferred DAPC clusters. Tips with white circles are outgroup taxa. Circles at nodes represent confidence in the preceding branch and show the results of 100 bootstraps inferred using RAxML.

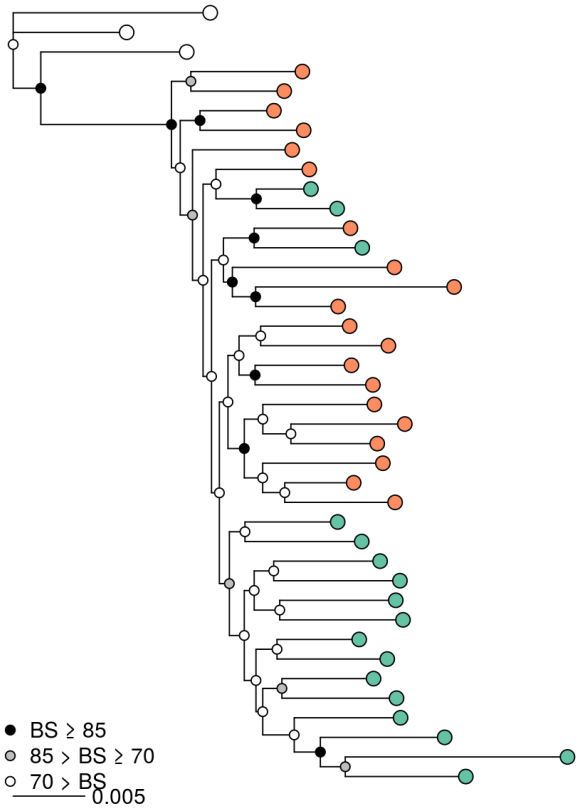


Fig. S9. Maximum likelihood phylogenetic tree of *P. acaulis*. Colours on tips correspond to inferred DAPC clusters. Tips with white circles are outgroup taxa. Circles at nodes represent confidence in the preceding branch and show the results of 100 bootstraps inferred using RAxML.

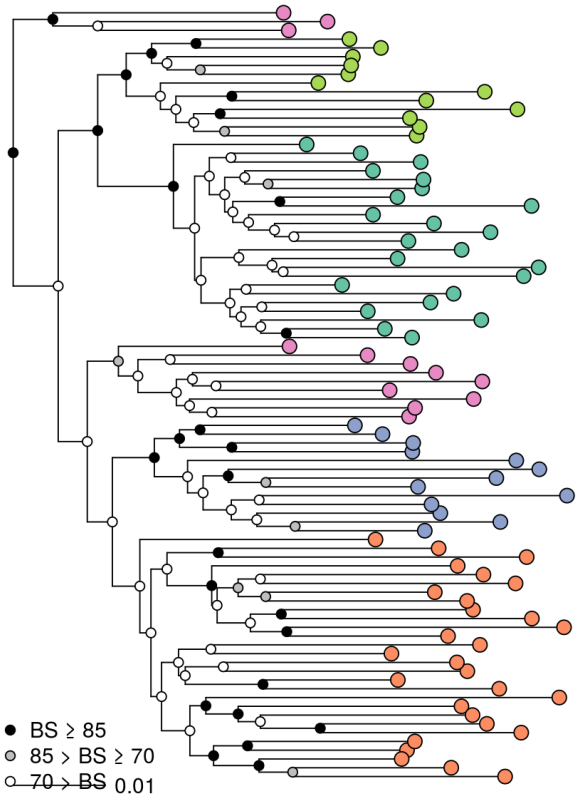


Fig. S10. Maximum likelihood phylogenetic tree of *P. barteri*. Colours on tips correspond to inferred DAPC clusters. Tips with white circles are outgroup taxa. Circles at nodes represent confidence in the preceding branch and show the results of 100 bootstraps inferred using RAxML.

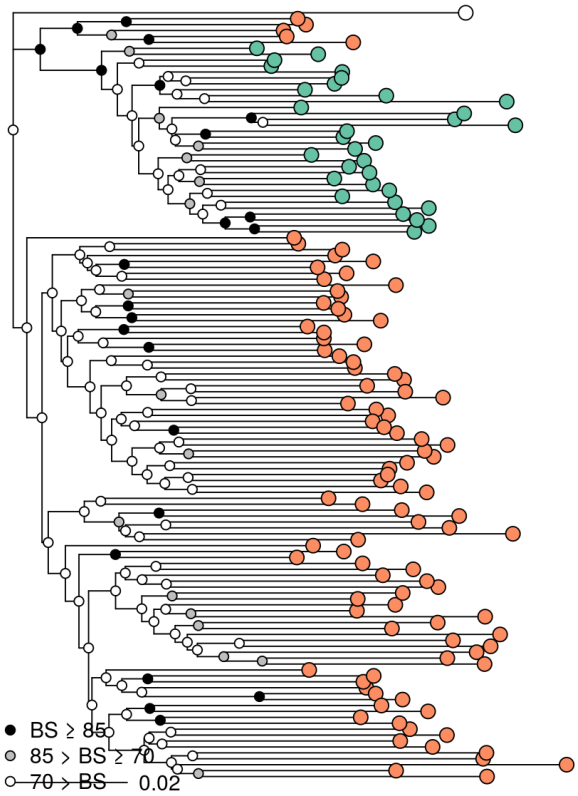


Fig. S11. Maximum likelihood phylogenetic tree of *S. mannii*. Colours on tips correspond to inferred DAPC clusters. Tips with white circles are outgroup taxa. Circles at nodes represent confidence in the preceding branch and show the results of 100 bootstraps inferred using RAxML.

Genetic diversity and putative refugia

One of our predictions (Table 1) is that the stability of refugial areas may lead to the accumulation of genetic diversity over time when compared to areas outside of refugia. Rapid expansion during interglacial periods would likely lead to low levels of diversity in newly (re)colonised areas, due to founder effects. We examined the feasibility of this hypothesis in central African rain forests by performing regressions on three different measures of individual-level expected heterozygosity (H_e), observed heterozygosity (H_o) and allelic richness. We compared these metrics of diversity to distance from either Maley's (2) or Anhuf's (3) refugia (see Fig. 1 in the main text). In the first instance we calculated the distance from each individual to the approximate central point of the nearest of Maley's refugia. In the second instance we calculated the distance from the coast (representing Anhuf's hypothesis of coastal refugia) for each individual and plotted this against genetic diversity. Two species (*G. sauveolens* and *P. barteri*) had conspicuous outliers that we identified using the *boxplot* function in R and removed from the dataset. We performed regression analyses for each species and diversity statistic, fitting linear models for each variable independently. We also log-transformed variables to account for any non-normality in our data and found extremely similar patterns (see Dryad repository for plots).

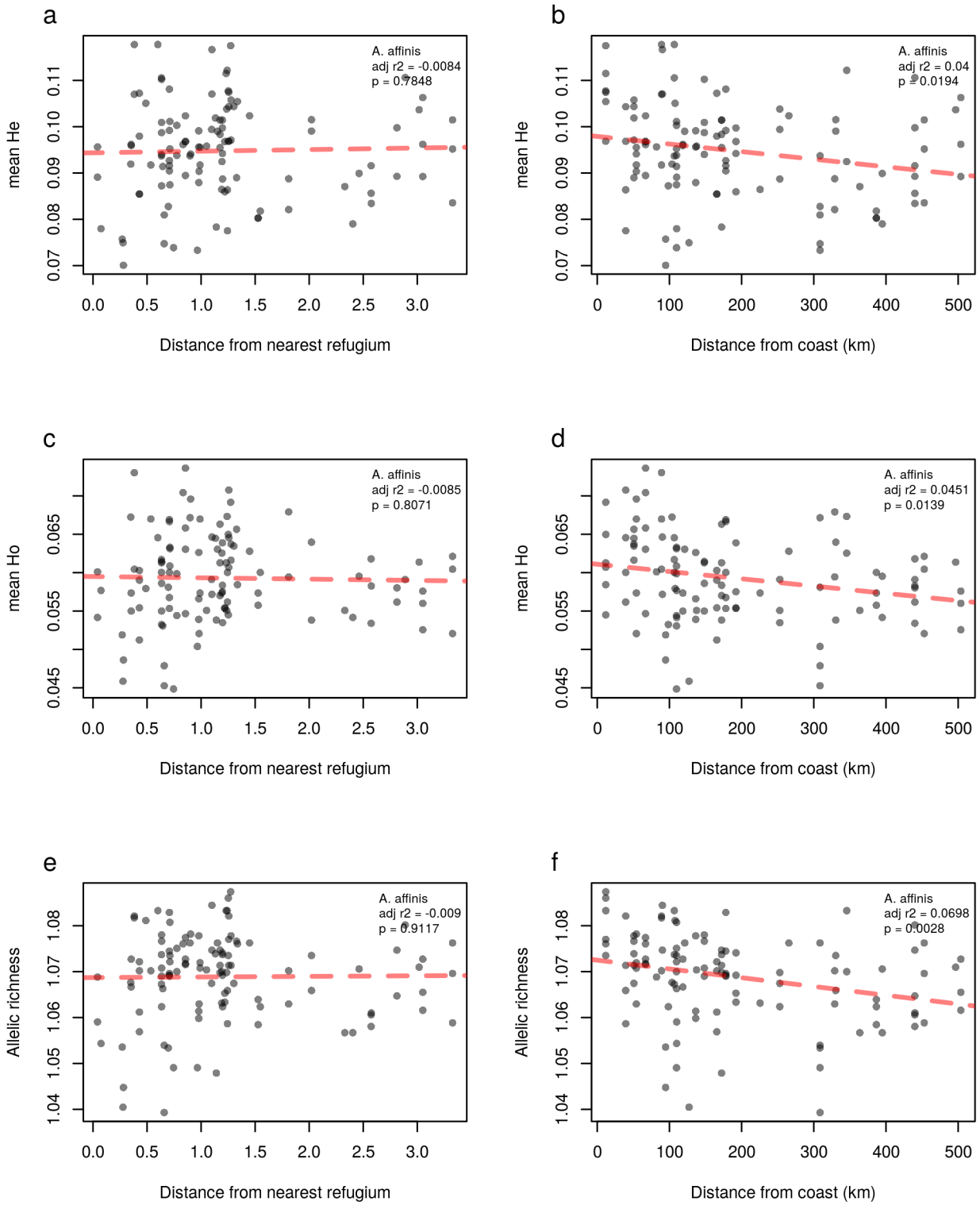


Fig. S12. Regressions of three different per-individual genetic diversity statistics (H_o, H_e and allelic richness) against distance to nearest refugium (a,c,e) and distance from coast (b,d,f) for *A. affinis*. Species names, R-squared values and p values are shown on each plot. The dotted red line indicates the fitted line of the regression model.

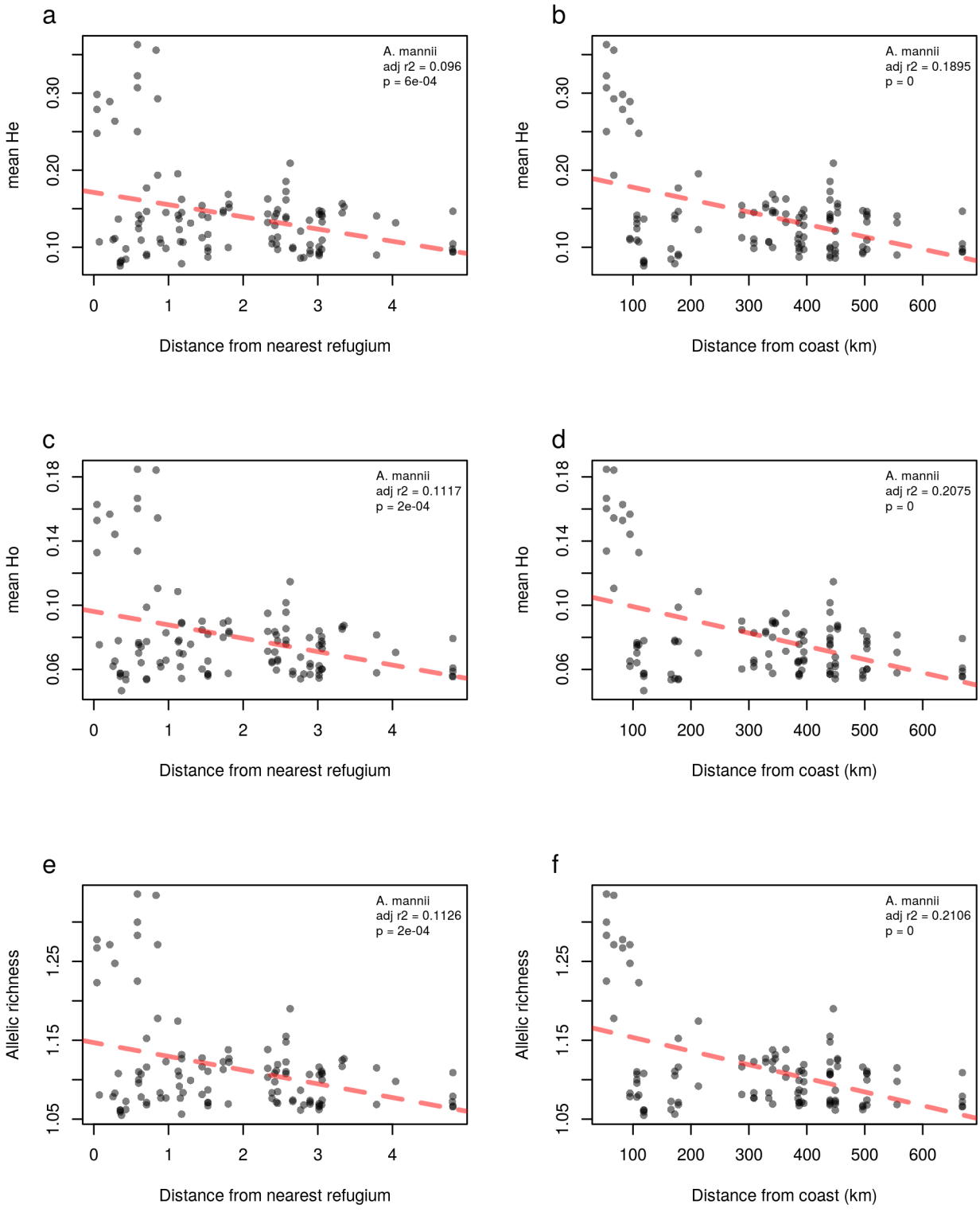


Fig. S13. Regressions of three different per-individual genetic diversity statistics (Ho, He and allelic richness) against distance to nearest refugium (a,c,e) and distance from coast (b,d,f) for *A. mannii.* Species names, R-squared values and p values are shown on each plot. The dotted red line indicates the fitted line of the regression model.

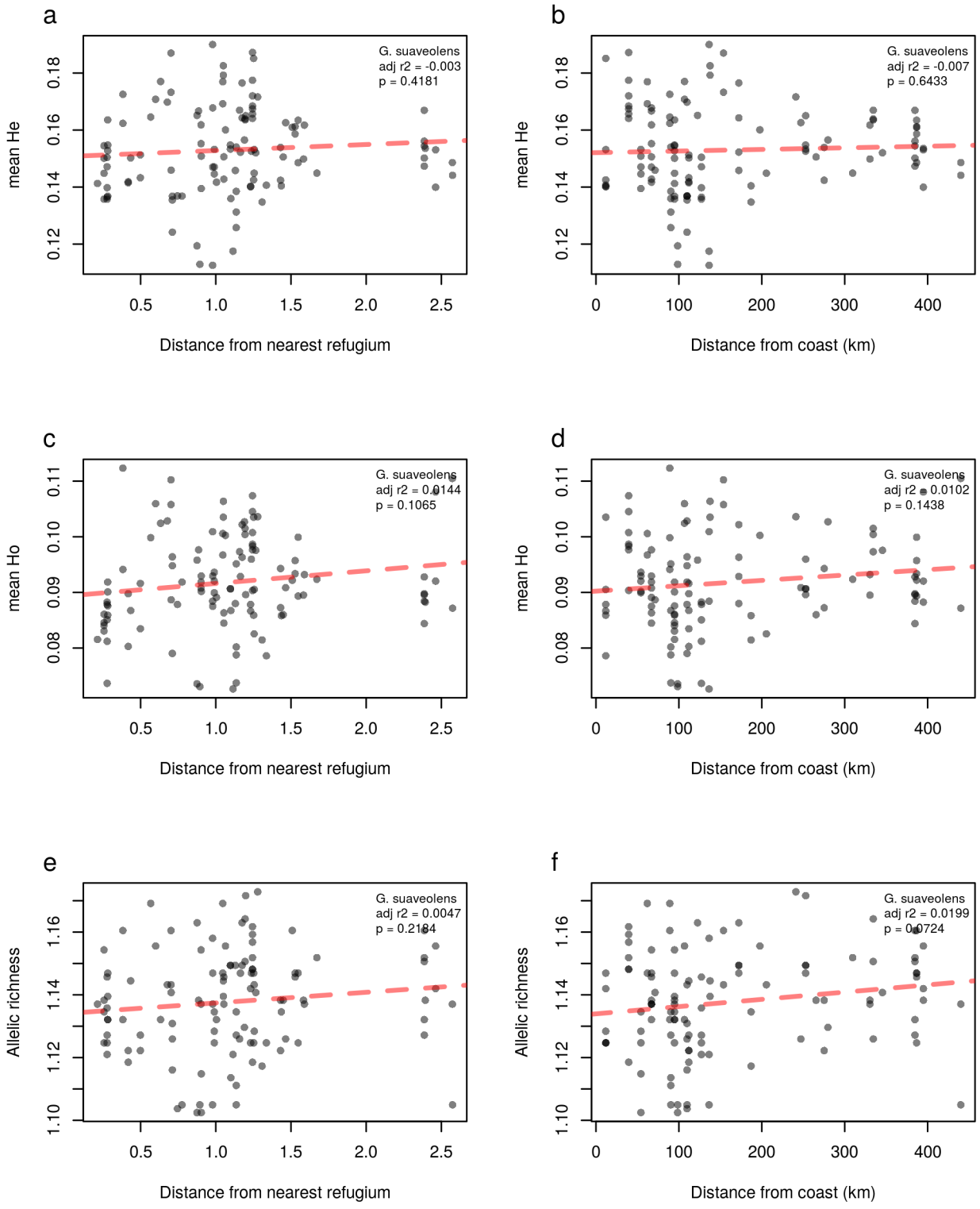


Fig. S14. Regressions of three different per-individual genetic diversity statistics (Ho, He and allelic richness) against distance to nearest refugium (a,c,e) and distance from coast (b,d,f) for *G. suaveolens*. Species names, R-squared values and p values are shown on each plot. The dotted red line indicates the fitted line of the regression model.

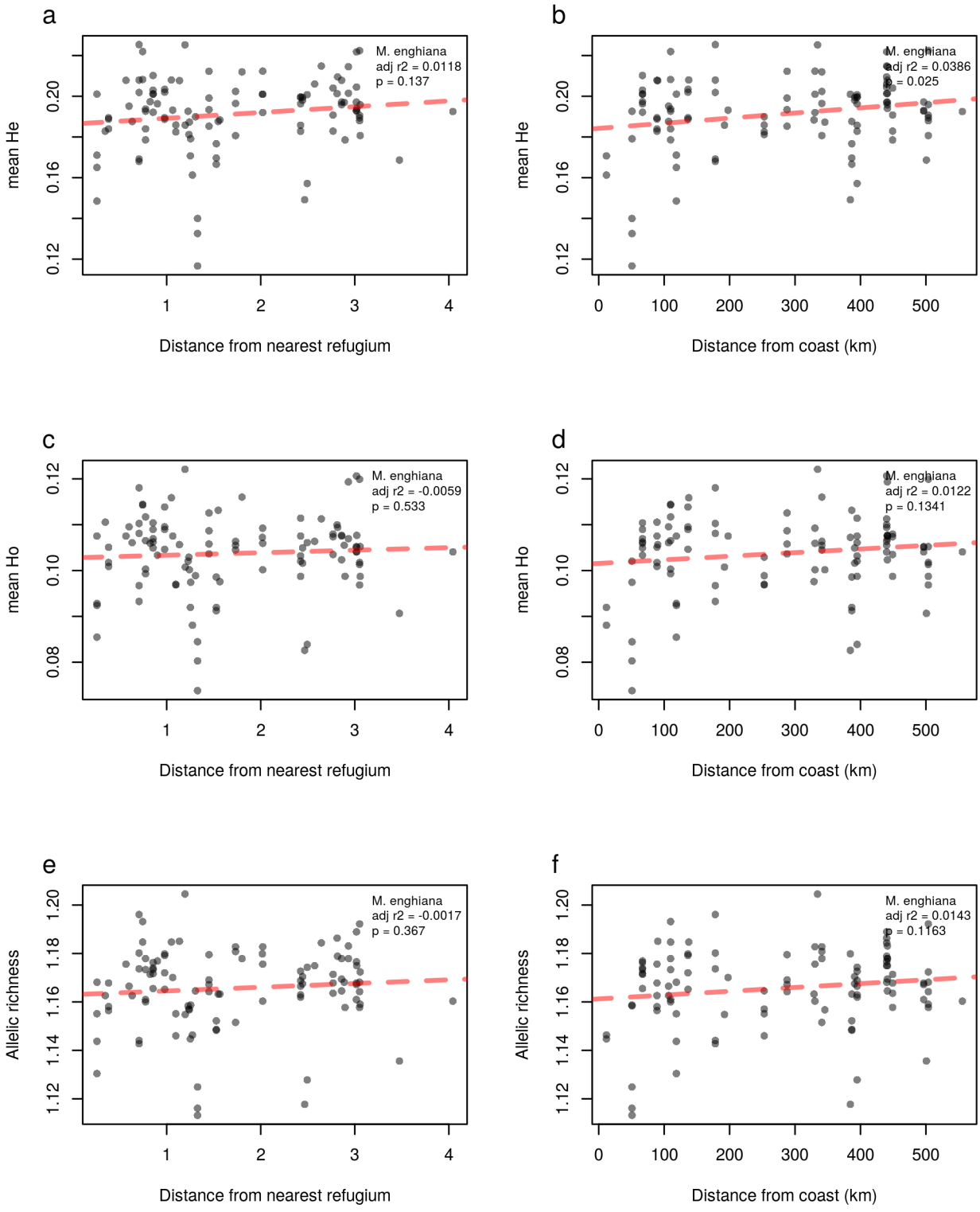


Fig. S15. Regressions of three different per-individual genetic diversity statistics (Ho, He and allelic richness) against distance to nearest refugium (a,c,e) and distance from coast (b,d,f) for *M. enghiana*. Species names, R-squared values and p values are shown on each plot. The dotted red line indicates the fitted line of the regression model.

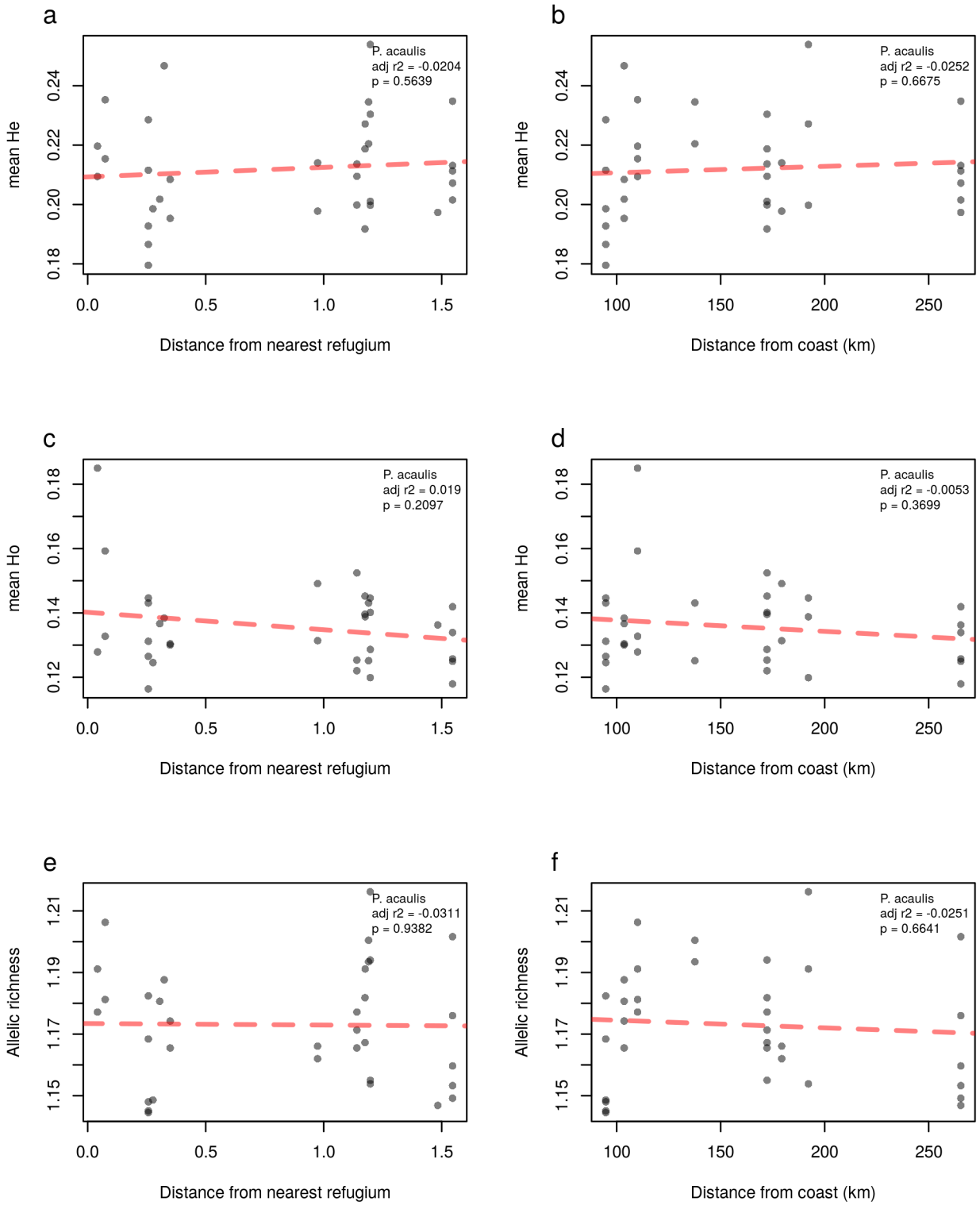


Fig. S16. Regressions of three different per-individual genetic diversity statistics (Ho, He and allelic richness) against distance to nearest refugium (a-c) and distance from coast (d-f) for *P. acaulis*. Species names, R-squared values and p values are shown on each plot. The dotted red line indicates the fitted line of the regression model.

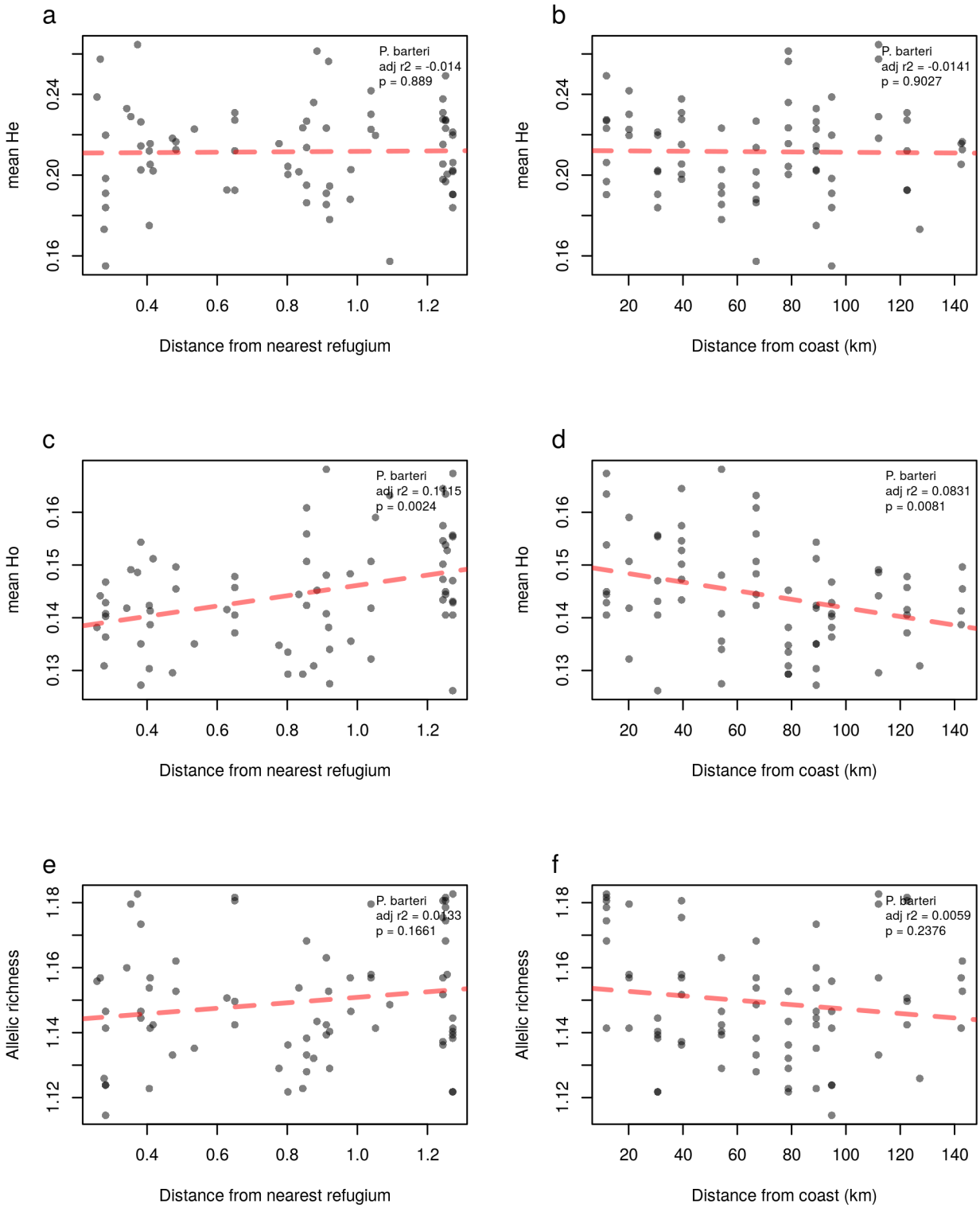


Fig. S17. Regressions of three different per-individual genetic diversity statistics (Ho, He and allelic richness) against distance to nearest refugium (a-c) and distance from coast (d-f) for *P. barteri*. Species names, R-squared values and p values are shown on each plot. The dotted red line indicates the fitted line of the regression model.

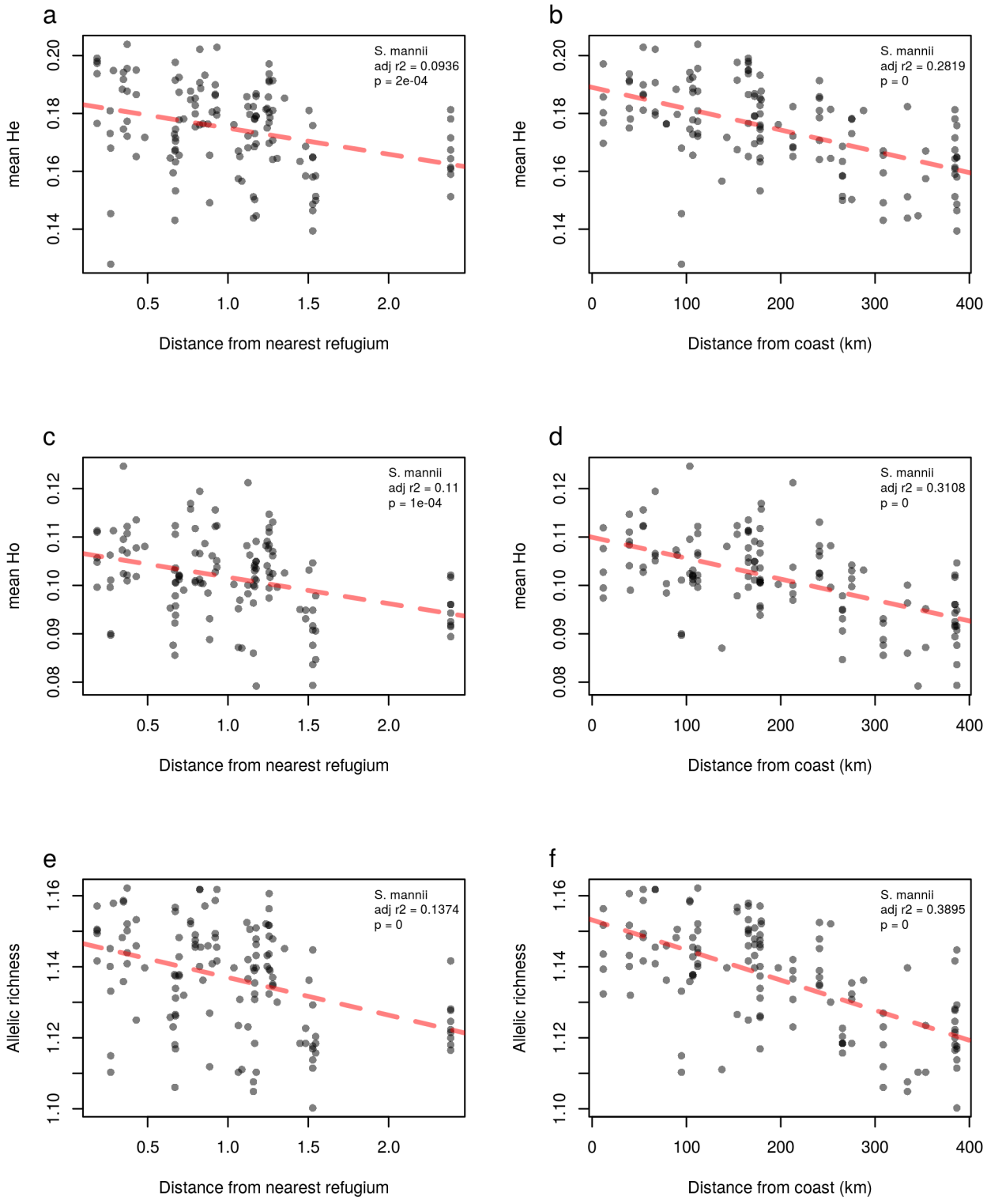


Fig. S18. Regressions of three different per-individual genetic diversity statistics (Ho, He and allelic richness) against distance to nearest refugium (a-c) and distance from coast (d-f) for *S. mannii*. Species names, R-squared values and p values are shown on each plot. The dotted red line indicates the fitted line of the regression model.

Analysis of MOlecular VAriance (AMOVA)

The AMOVA approach is used to assess population differentation using genetic markers (4). A pairwise genetic-distance matrix is used to test how much genetic variation can be explained by a given stratification of the dataset. We used AMOVA to test how much genetic variance could be explained by the North-South climatic inversion. Genetic clusters inferred from DAPC were grouped based on whether they were north or south of the climatic inversion discontinuity at the climate inversion (where the majority of its range was found). Then, each genetic cluster was considered as a second layer of structure, or subpopulation. AMOVAs were performed using the *poppr.amova* function in the R package 'poppr' (5) and assessed significance by performing 999 randomisations, shuffling individual labels. We did not include within-individual variance in our models. We then repeated our analysis, randomly shuffling population assignments in our dataset to calculate the significance of the population structure.

We also tested how much variance was explained by previously proposed glacial refugia or estimated areas of stability. We split samples into two groups, those considered inside stable areas and those outside. We defined presence in Maley's refugia using 1st quantile of distances from refugia in each dataset. Individuals closer than the 1st quantile distance were considered as inside a stable area, and individuals further than this were considered outside stable areas. We took a similar approach for the climate-based stable areas, using the difference between present and past climate at an individuals location instead of the geographic distance from a point. We subtracted rasters of present-day climate values from past climate values (LGM; MIROC) for annual mean temperature (bio1) and annual mean rainfall (bio12). We rescaled both variables to have a mean of 1, then added them together to produce a raster where smaller values indicated stability through time. This approach gives equal weight to rainfall and temperature variability. Finally, we used the binary presence-absence maps output in our ecological niche modelling (ENM) analyses to split individuals into groups for the ENM-based stable areas. We followed the same approach as detailed above, but did not include genetic cluster as a subpopulation layer of structure.

test	sp	sigma	percentage	p
Variation between North/South	anni	-506.71	-59.74	0.00
Variation among clusters within North/South	anni	1123.22	132.42	0.00
Variation within clusters	anni	231.70	27.32	0.75
Variation between North/South	anon	303.49	43.45	0.00
Variation among clusters within North/South	anon	226.36	32.41	0.00
Variation within clusters	anon	168.67	24.15	0.00
Variation between North/South	green	12.34	2.26	0.01
Variation among clusters within North/South	green	9.23	1.69	0.08
Variation within clusters	green	524.50	96.05	0.21
Variation between North/South	mona	-112.90	-41.63	0.00
Variation among clusters within North/South	mona	169.69	62.58	0.00
Variation within clusters	mona	214.38	79.06	0.66
Variation between North/South	podo_b	14.82	11.00	0.00
Variation among clusters within North/South	podo_b	37.98	28.17	0.00
Variation within clusters	podo_b	82.01	60.83	0.11

Table S3. Components of covariance from AMOVA analyses.The components of covariance table shows how much variance is detected at each stratification. Sigma represents the variance for each hierarchical level and to the right is the percent of the total.

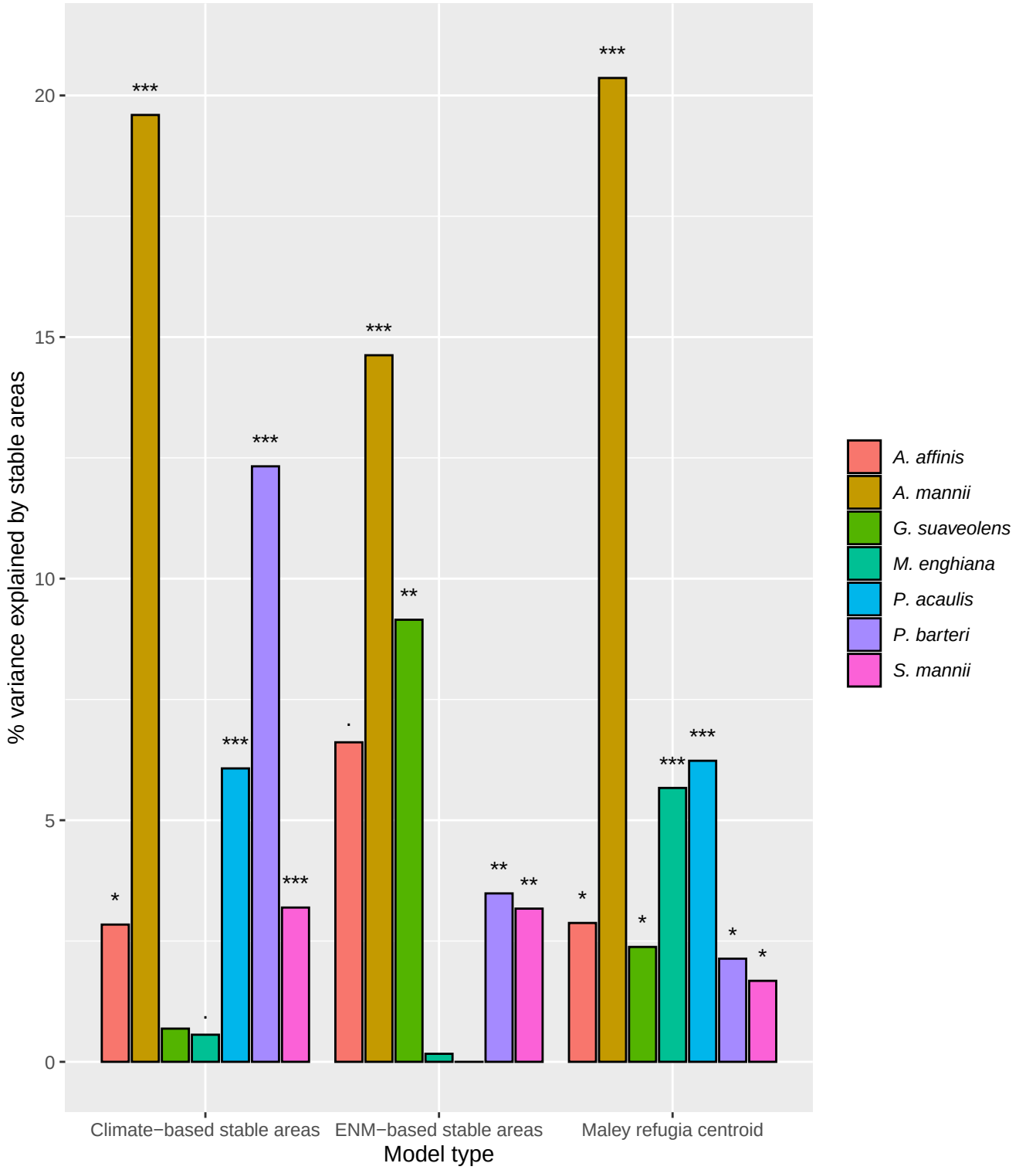


Fig. S19. Barplot showing the amount of genetic variance explained by different concepts of refugia (1st quantile); climate-based, ENM-based and Maley's (1996) refugia. Bars are coloured by species and signficance is denoted above each bar (. = p < 0.1, * = p <0.05, ** = p <0.01, *** = p<0.001). No *P. acaulis* individuals occurred within estimated ENM stable areas so the corresponding AMOVA was not conducted.

Fossil calibrations

Fossil name	Hard lower bound	Soft upper bound (95%)	Exponential mean
Sabalites carolinensis	85.80	88.80	1.00
Attaleinae	54.80	60.79	2.00
Hyphaeneocarpon indicum	64.00	67.00	0.80

Table S4. Fossil calibrations used for the palm backbone tree inference. Parameters are for priors with an exponential distribution.

78 ## Backbone trees

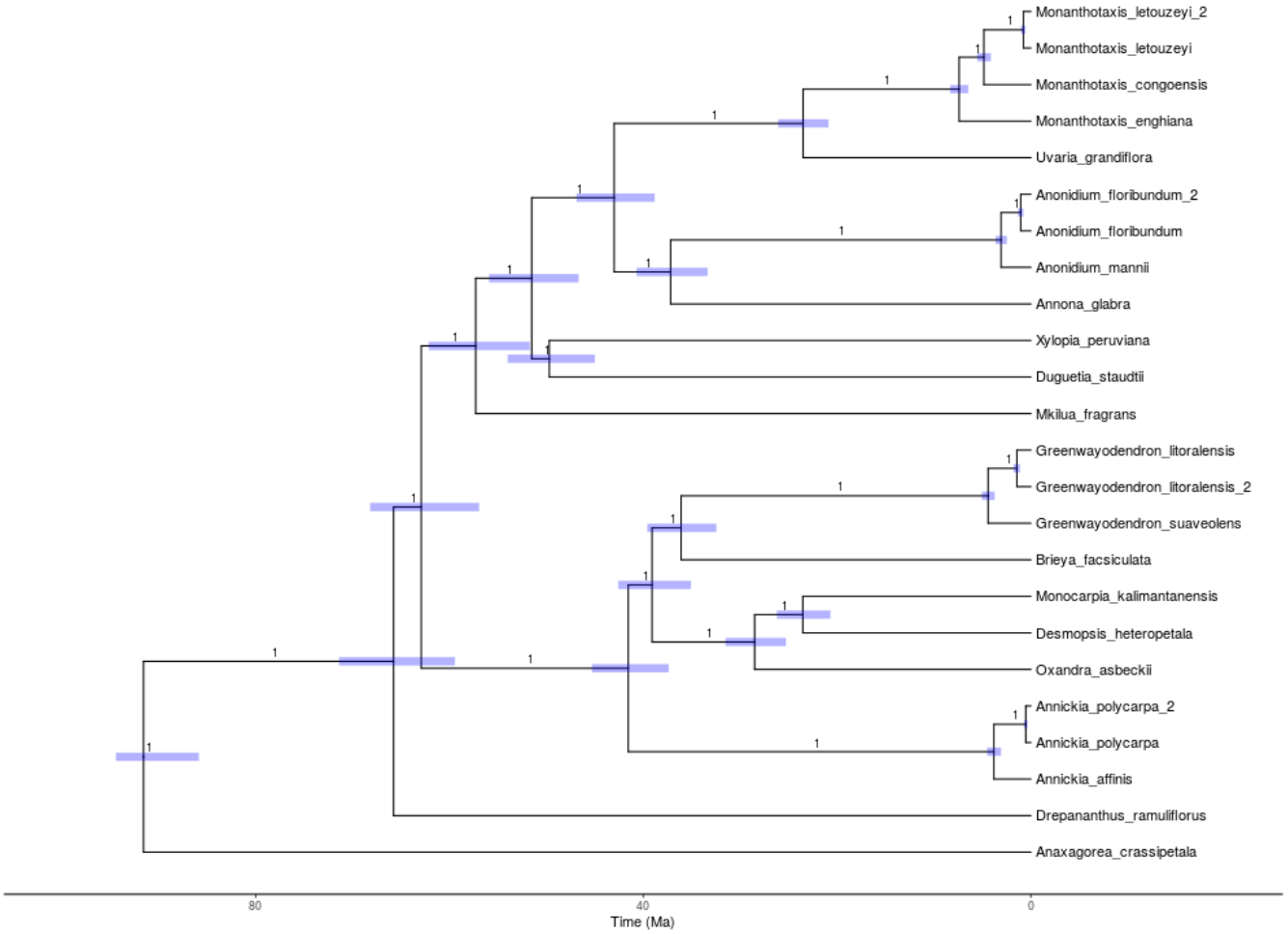


Fig. S20. Backbone tree of Annonaceae constructed using BEAST2. The tree was calibrated with a single fossil calibration at the root node. Posterior probabilities are indicated above the corresponding branch. 95% highest posterior densities (HPD) of node age are shown with blue bars at nodes.

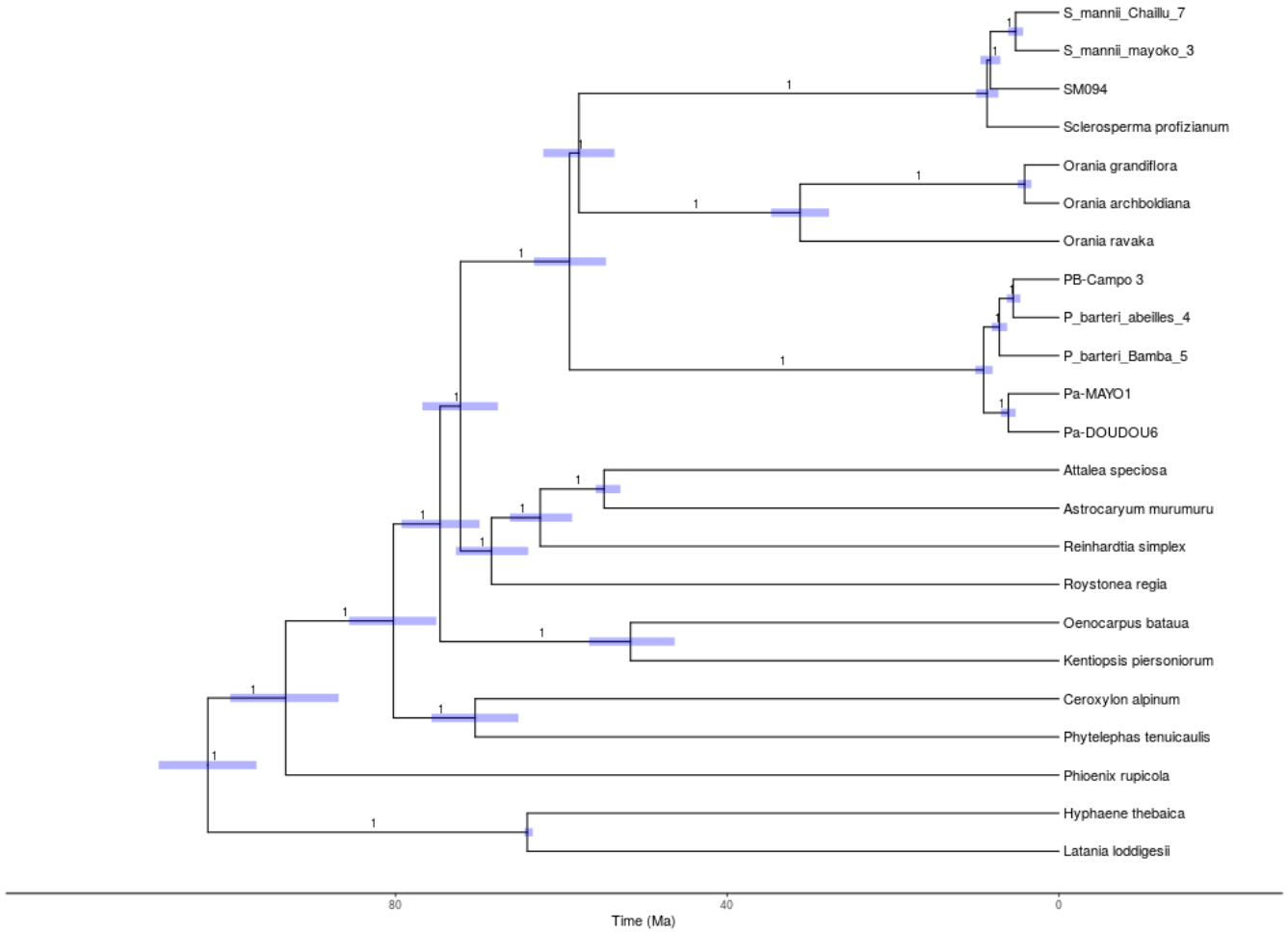


Fig. S21. Backbone tree of Arecaceae constructed using BEAST2. The tree was calibrated with a single fossil calibration at the root node. Posterior probabilities are indicated above the corresponding branch. 95% highest posterior densities (HPD) of node age are shown with blue bars at nodes.

Backbone-and-patch trees

To achieve a tree with all species and backbone taxa we performed a backbone-and-patch approach as outlined in (6). For each species we first rescaled the StarBEAST tree, setting the root age as the time of divergence between the focal species and its closest outgroup (e.g. *Annickia affinis* and *A. polycarpa*). We then shortened the tip leading to its representative in the backbone tree using the stem/crown ratio of our StarBEAST tree. The StarBEAST tree was then grafted to the shortened tip while the tree remained ultrametric. We repeated this for each plant family, Annonaceae (Fig. S22) and Arecaceae (Fig. S23), separately.

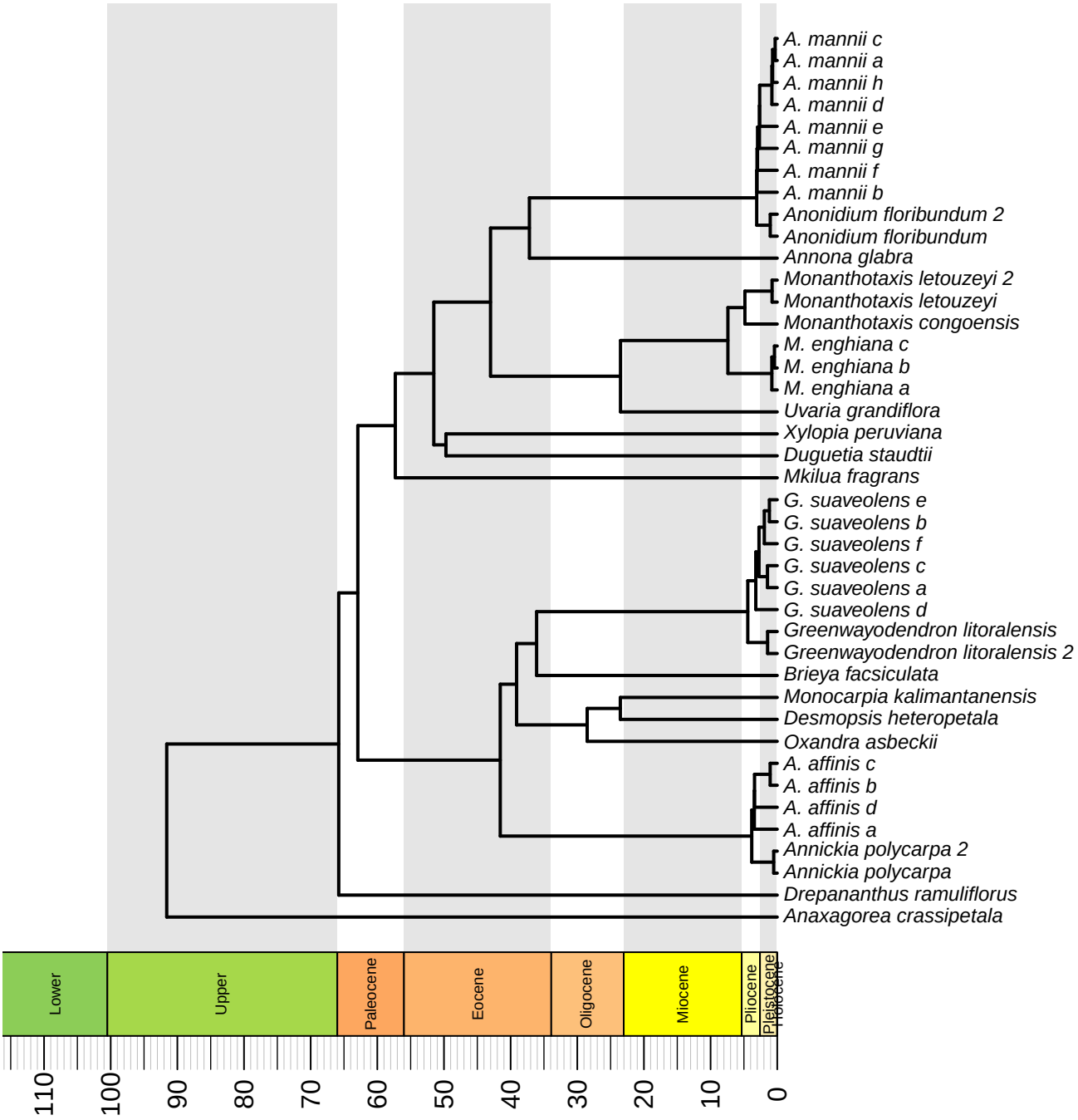


Fig. S22. Phylogenetic trees showing the timing of population divergence events in four Annonaceae species. The tree was built using the backbone-and-patch approach whereby a backbone tree was first constructed for Annonaceae (Fig. S20) and then StarBEAST trees were rescaled and grafted into place on the backbone trees (see methods for further details). Timescale and epochs are shown on the x axis.

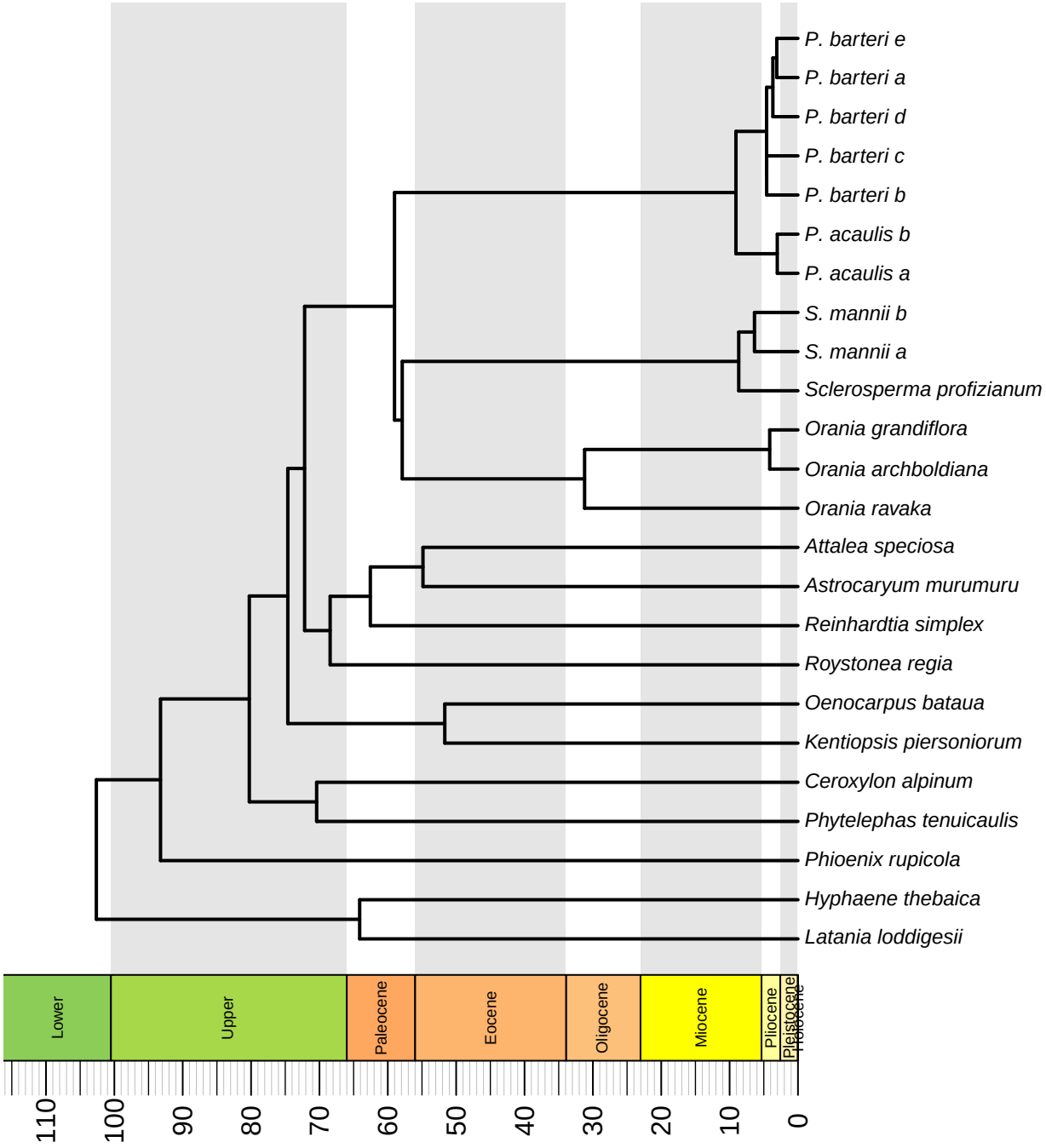


Fig. S23. Phylogenetic trees showing the timing of population divergence events in three Palm species. The tree was built using the backbone-and-patch approach whereby a backbone tree was first constructed for Arecaceae (Fig. S21) and then StarBEAST trees were rescaled and grafted into place on the backbone trees (see methods for further details). Timescale and epochs are shown on the x axis.

Phylogenetic trees reconstructed using StarBEAST and DENIM

Summary of the three molecular dating methods used.

1. StarBEAST is a typical coalescent Bayesian inference method, allowing us to jointly date and infer topologies for phylogenetic trees.

2. DENIM extends this by accounting for migration in the data, providing a useful comparison to understand the potential impact of low-level migration.

3. The backbone-and-patch approach produces a tree with both inter and intraspecific relationships, and avoids the additional noise produced when using secondary calibrations by rescaling the tree based on fossil calibrations on backbone tree.

In preliminary analyses we examined the effect of taking phased alleles into account (7) as output by SECAPR (8). In addition, we examined estimated ages in *A. affinis* and found that they extensively overlapped with those inferred using a consensus sequence (Fig. S30b). However, this approach doubles the number of tips, resulting in significantly more computational time, so for these two reasons we used a single sequence per individual. We also ran our StarBEAST analysis with a relaxed uncorrelated lognormal clock model, in comparison with a strict clock to compare divergence times (Fig. S30c). We found that times were generally similar but convergence was much harder to achieve and required computation time was significantly longer. Monophyly constraints were used on the focal Annonaceae species as preliminary runs yielded non-monophyletic species in some cases. Preliminary runs indicated that this was not necessary for the palm species.

Divergence estimation notwithstanding ILS and migration (DENIM). In order to account for migration amongst different genetic clusters we also used the Divergence estimation notwithstanding ILS and migration (DENIM) approach (9). The model used is similar to IMa2 (10), though the phylogenetic tree is estimated under a birth-death model rather than being assumed. Migration is calculated between pairs of branches using an n-island model, providing information about the number of migrations within each gene tree. DENIM can perform poorly when levels of migration are high. Nevertheless, taking migration into account when it may be present can help to improve the estimation of topology and divergence times and acts as a useful comparison to our StarBEAST analyses. We ran DENIM analyses using the same loci and substitution models as our StarBEAST analyses. We assumed a strict clock and followed the instructions in the DENIM manual in regard to prior choice. In some cases (notably *A. mannii* and *G. suaveolens*) we had to set upper limits on clock rate to avoid the analysis finding stationarity at implausably high rates. We conducted multiple (2-3) runs for each species, and summarised the combined the posterior distribution of trees from each run to produce a maximum-clade credibility (MCC) tree.

It is important to note the coalescent phylogenetic analyses results could be influenced by gene flow among genetic clusters leading to underestimation of the timing of divergence events (11). However, our DENIM approach takes into account low levels of migration, and yielded broadly similar topologies and divergence times when compared to analyses without migration (Fig. S24-S29). Furthermore, our clustering analyses revealed limited evidence for admixture between genetic clusters (Fig. S3). Species in which admixture was more prevalent (e.g. *M. enghiana*) may have actually diverged earlier than estimated in our phylogenetic trees (see Fig. 3 in the main text). However, this admixture is not necessarily indicative of ongoing gene flow. Introgression events may cause loci to deviate from the molecular clock. Given that we selected markers by this metric, this may have helped to reduce the effect of any potential gene flow.

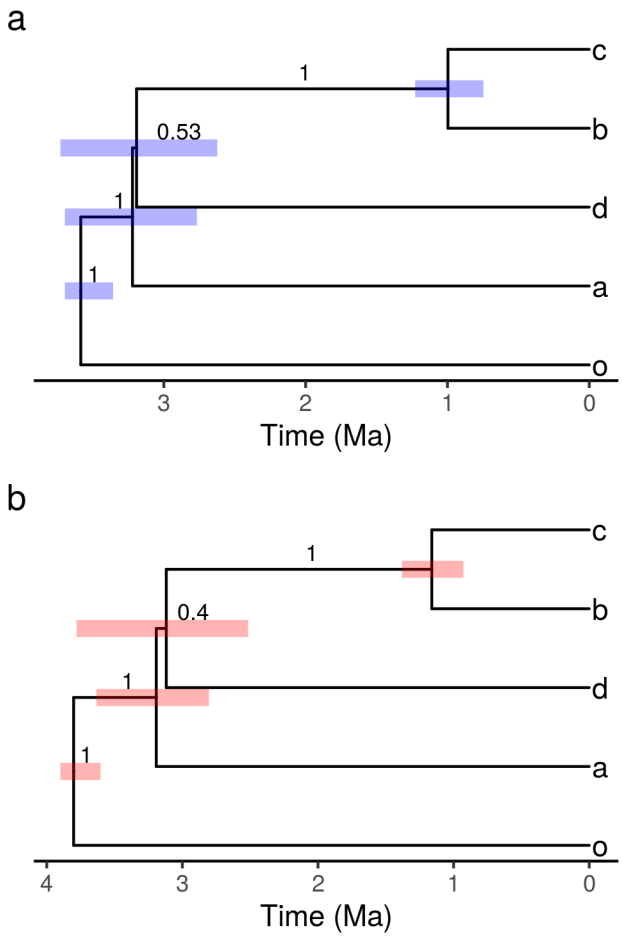


Fig. S24. Phylogenetic trees inferred using (a) StarBEAST and (b) DENIM for *A. affinis*. Genetic clusters are labelled alphabetically from 'a' to 'd' and outgroups are labelled 'o'. Node ages represent target heights and node bars represent 95% highest posterior densities of divergence times. Posterior probabilities are annotated on trees for each divergence event (0-1).

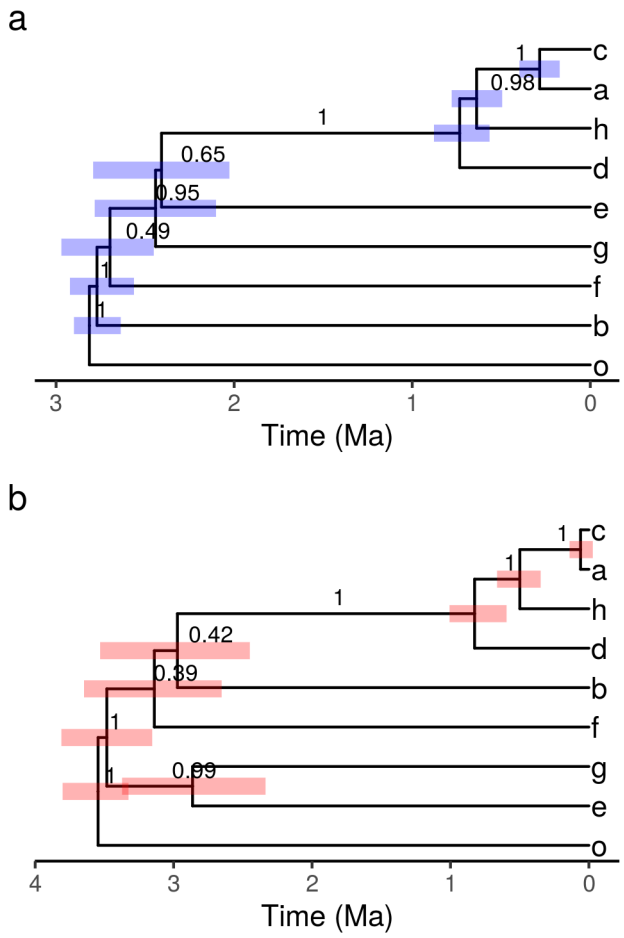


Fig. S25. Phylogenetic trees inferred using (a) StarBEAST and (b) DENIM for *A. mannii*. Genetic clusters are labelled alphabetically from 'a' to 'h' and outgroups are labelled 'o'. Node ages represent target heights and node bars represent 95% highest posterior densities of divergence times. Posterior probabilities are annotated on trees for each divergence event (0-1).

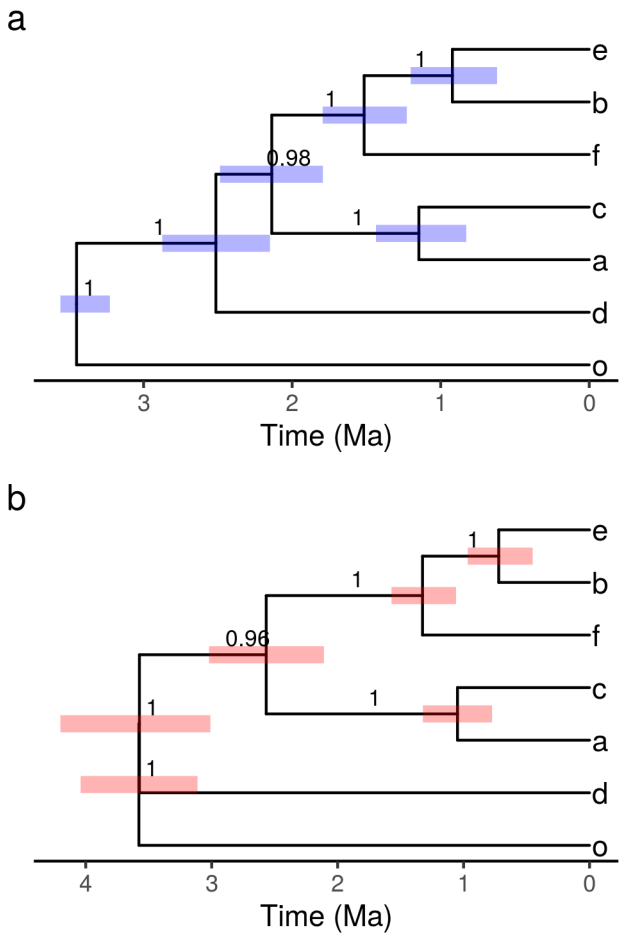


Fig. S26. Phylogenetic trees inferred using (a) StarBEAST and (b) DENIM for *G. suaveolens*. Genetic clusters are labelled alphabetically from 'a' to 'f' and outgroups are labelled 'o'. Node ages represent target heights and node bars represent 95% highest posterior densities of divergence times. Posterior probabilities are annotated on trees for each divergence event (0-1).

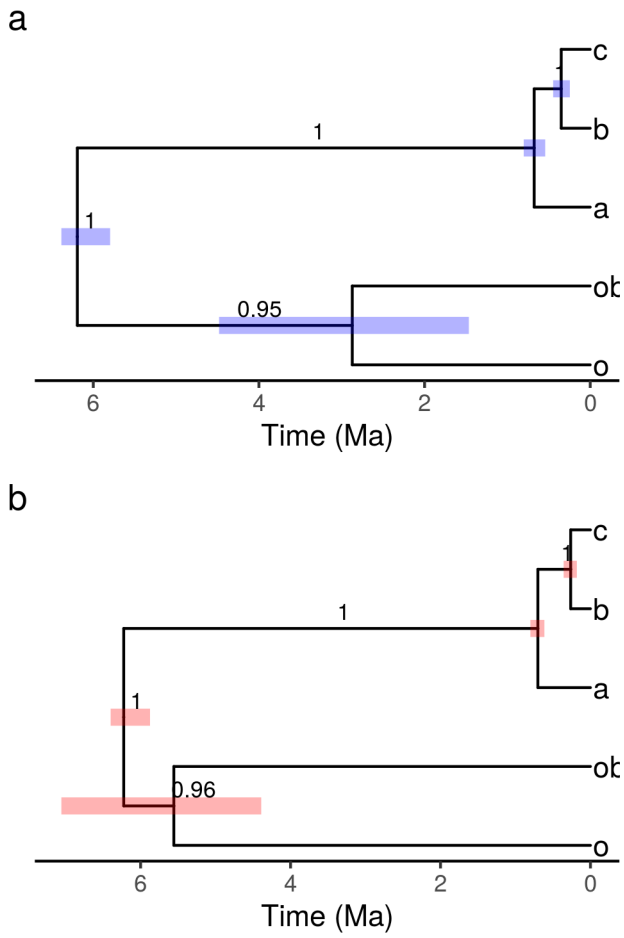


Fig. S27. Phylogenetic trees inferred using (a) StarBEAST and (b) DENIM for *M. enghiana*. Genetic clusters are labelled alphabetically from 'a' to 'c' and outgroups are labelled 'o' and 'ob'. Node ages represent target heights and node bars represent 95% highest posterior densities of divergence times. Posterior probabilities are annotated on trees for each divergence event (0-1).

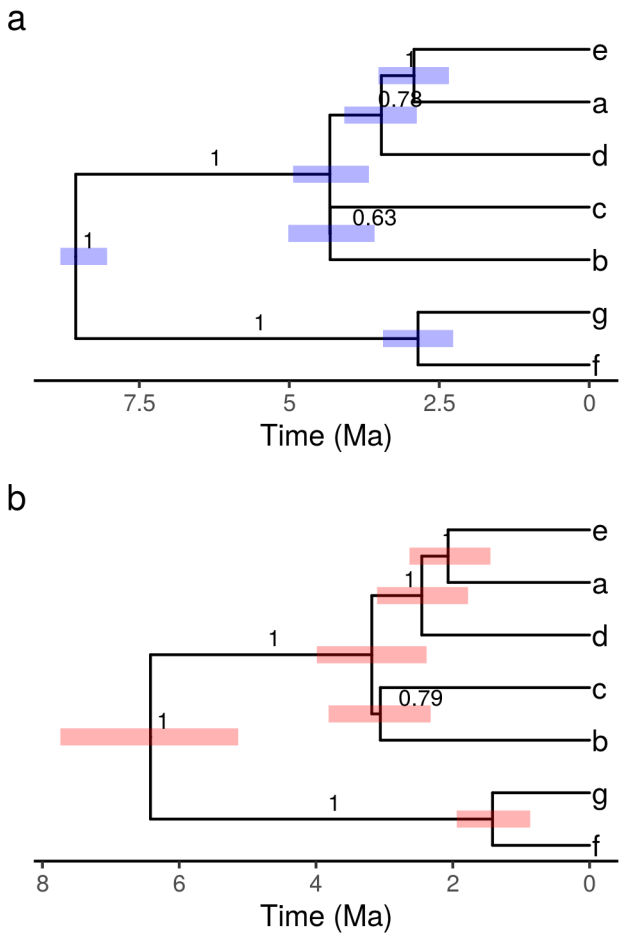


Fig. S28. Phylogenetic trees inferred using (a) StarBEAST and (b) DENIM for *P. acaulis* and *P. barteri*. Genetic clusters are labelled alphabetically from 'a' to 'e' for *P. barteri* and 'f' to 'g' for *P. acaulis*. Node ages represent target heights and node bars represent 95% highest posterior densities of divergence times. Posterior probabilities are annotated on trees for each divergence event (0-1).

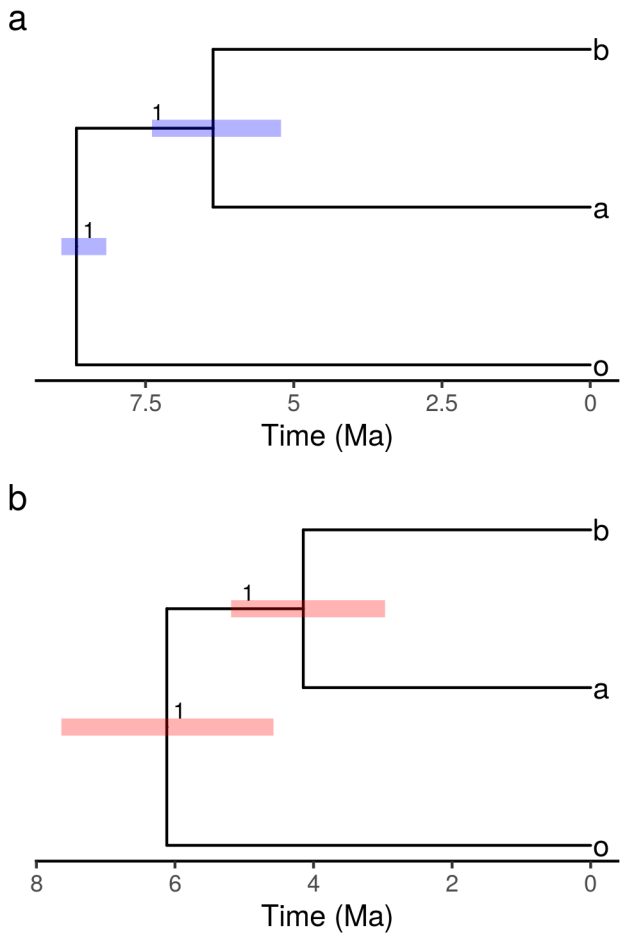


Fig. S29. Phylogenetic trees inferred using (a) StarBEAST and (b) DENIM for *S. mannii*. Genetic clusters are labelled alphabetically from 'a' to 'b' and outgroups are labelled 'o'. Node ages represent target heights and node bars represent 95% highest posterior densities of divergence times. Posterior probabilities are annotated on trees for each divergence event (0-1).

Testing the effect of using alleles and lognormal clock models

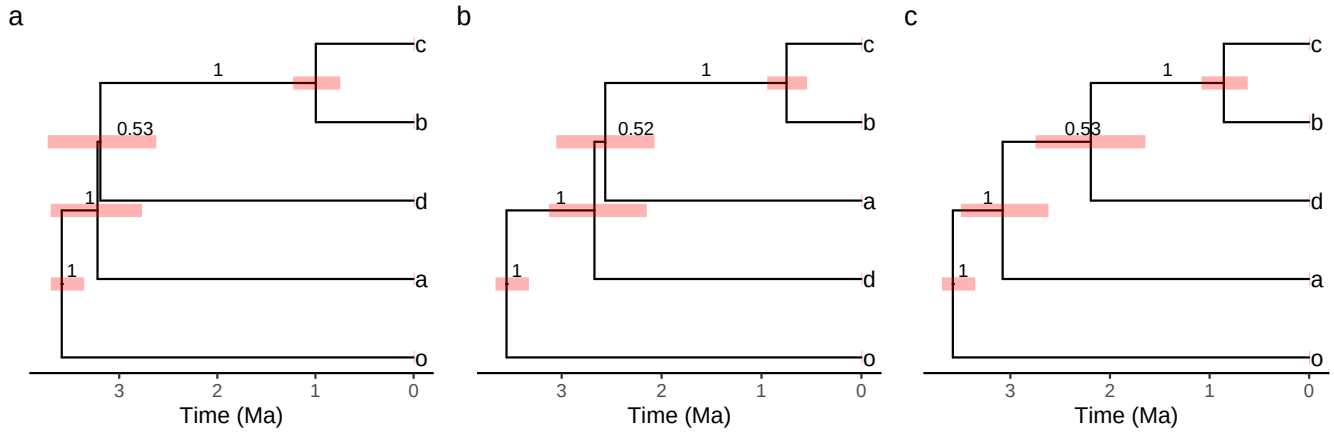


Fig. S30. Phylogenetic reconstruction in *A. affinis* using (a) single consensus sequence, (b) two allele sequences per individual for assessing divergence times and (c) as in (a) but with uncorrelated lognormal relaxed clock models instead of strict clock. Blue bars around nodes indicate 95% confidence intervals. Pie charts at nodes show posterior probability percentages. We note that the random set of five individuals per cluster used in (b) is different to those in (a) and (c).

Variable.name	Description
BIO1	Annual Mean Temperature
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO3	Isothermality (BIO2/BIO7)
BIO4	Temperature Seasonality (standard deviation)
BIO5	Max Temperature of Warmest Month
BIO6	Min Temperature of Coldest Month
BIO7	Temperature Annual Range (BIO5-BIO6)
BIO8	Mean Temperature of Wettest Quarter
BIO9	Mean Temperature of Driest Quarter
BIO10	Mean Temperature of Warmest Quarter
BIO11	Mean Temperature of Coldest Quarter
BIO12	Annual Precipitation
BIO13	Precipitation of Wettest Month
BIO14	Precipitation of Driest Month
BIO15	Precipitation Seasonality (Coefficient of Variation)
BIO16	Precipitation of Wettest Quarter
BIO17	Precipitation of Driest Quarter
BIO18	Precipitation of Warmest Quarter
BIO19	Precipitation of Coldest Quarter
altitude	elevation (m)
soil order	USDA soil taxonomy (ST) developed by United States Department of Agriculture and the National Cooperative
Distance from closest Maley refugia centroid	Distance (km) from central point of Maley's (1996) proposed refugia
Distance from coast	Distance from coast as calculated with distance from distance_to_coastline_lowres function in the 'distance_to_coast' R package
Habitat stability as extracted from ENMs	Areas predicted as suitable (TSS threshold) in both past (MPI_ESM_P) and present ENMs.
BIO1 stability	Absolute values of differences between past (MIROC_ESM) and present annual mean temperature
BIO12 stability	Absolute values of differences between past (MIROC_ESM) and present annualrainfall
Combined climatic stability as calculated from climate data	Sum of BIO1 stability and BIO12 stability
MEM	Moran Eigenvector's Maps - multivariate summary of geographic structure
BDRICM_M_250m	Depth to bedrock (R horizon))
BDTICM_M_250m	Absolute depth to bedrock (cm)
BLDFIE_M_sl2_250m	Bulk density (kg m_3) of the fine earth fraction ($<$ 2 mm)
CECSOL_M_sl2_250m	Cation exchange capacity of soil
CLYPPT_M_sl2_250m	Proportion of clay particles ($<$ 0.002 mm) in the fine earth fraction
CRFVOL_M_sl2_250m	Volumetric fraction of coarse fragments ($>$ 2 mm)
OCSTHA_M_sd1_250m	Organic carbon stocks
ORCDRC_M_sl2_250m	Organic carbon density
PHIKCL_M_sl2_250m	Soil pH in KCl solution
SLTPPT_M_sl2_250m	Proportion of silt particles ($>$ 0.002 mm and $<$ 0.05 mm) in the fine earth fraction
SNDPPT_M_sl2_250m	Proportion of sand particles ($>$ 0.05 mm) in the fine earth fraction
TAXNWRB_250m	observed taxonomic class in the WRB system (target variable)

Table S5. Variables used in the ENM and db-RDA analyses. Note that variables related to soil order, TAXNWRB, distances from refugia, MEMs and measures of stability were only used in db-RDA analyses. Further information on soil variables can be found here: https://www.isric.org/explore/soilgrids/faq-soilgrids

Ecological Niche Models (ENMs)

We obtained data for the 19 bioclimatic variables from the worldclim database (http://www.worldclim.org) including present-day data ('wc2.1_2.5m_bio') and data for three different past-climate models ('CCSM4' (CCL), 'MIROC_ESM' (MRL), 'MPI_ESM_P' (MEL)). We also included data from 11 different soil variables, downloaded from http://www.soilgrids.org (accessed March 2020). Elevation data was obtained using the *get_elev_raster* function in the R package 'elevatr'. All data were resampled to 2.5 arcmin resolution to match present-day climate data. Soil and altitude variables were kept constant in the past and present. We identified and removed correlated environmental variables (>0.9) by running the *raster.cor.matrix* function in the R package 'ENMTools' (12). The full set of input variables can be found in S5. We collated a total of 1691 individual locations across the seven focal species (see supplementary data for details).

We constructed ENMs for each of the seven species using the R package 'biomod2' (13) and climatic, altitude and soil variables. We created a convex hull around sample points (5 degrees) and used this as a mask to create a background area for modelling (see R package 'ConR' for additional details (14)). We performed 10 replicates of 10,000 pseudoasbences with the surface range envelop model (SRE) with a threshold of 0.05 (5%). An SRE is consructed (using the specified quantile) for the species and pseudoabsences are extracted outside of this envelope. Note that pseudoabsences are randomly generated and will cause repeated runs to be different We inferred ENMs using five different model algorithms: Generalized linear model (GLM), Boosted regression trees (GBM), Artificial neural networks (ANN), Random forest (RF), MAXENT.Phillips (maxent). We evaluated model fit using Cohen's Kappa (KAPPA) and true skill statistic (TSS). Once all algorithms were run we performed ensemble modeling using all models. We then projected models (of all algorithms) into past (and present) climate, performing a further ensemble modelling step. Finally we converted ensemble projections to binary presence-absence maps using KAPPA and TSS. These are plotted below, showing areas suitable for each species under a given statistic and present or past climate model. We highlight areas that were suitable in the past only, present only, or both. Details of model adequacy can be found in Table S6. We used results from the 'CCSM4' (AMOVA, db-RDA) and 'MIROC_ESM' (calculation of climatic stable areas) past-climate models for downstream analyses.

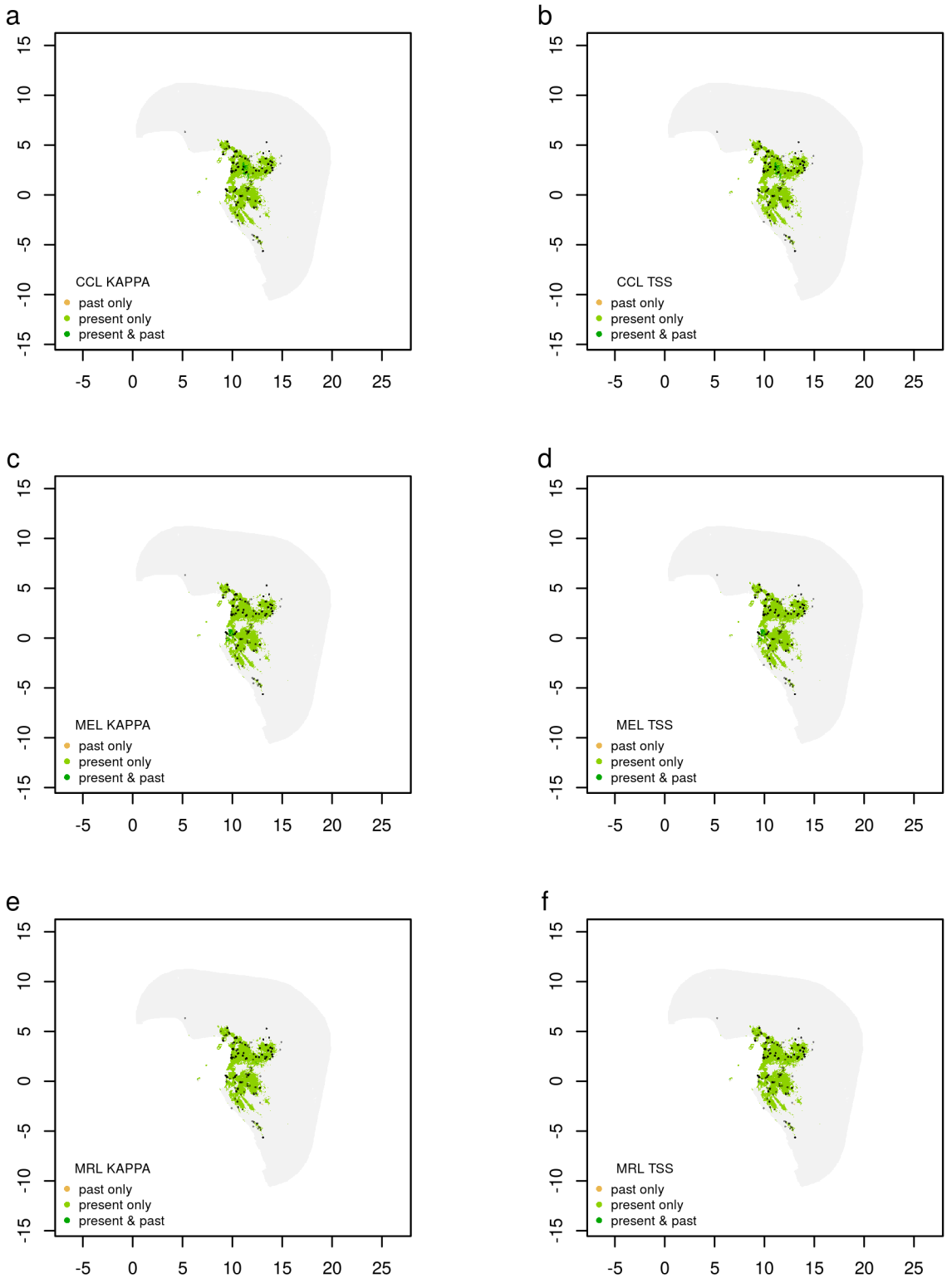


Fig. S31. Habitat suitability and stable areas inferred from ecological niche modelling for *A. affinis*. Suitability was summarised as binary maps for present and past climate, these were added together so that values of 0.5 indicate habitat only suitable in the past, values of 1 indicate areas only suitable in the present and values of 1.5 indicate areas suitable during both periods (stable areas). Results are from three different models of past climate: CCL (a-b), MEL (c-d), and MRL (e-f). Two statistics were used to infer binary maps: KAPPA (a, c, e) and the true skill statistic (TSS; b, d, f). Individual locations used in the modelling are shown as black points on each plot.

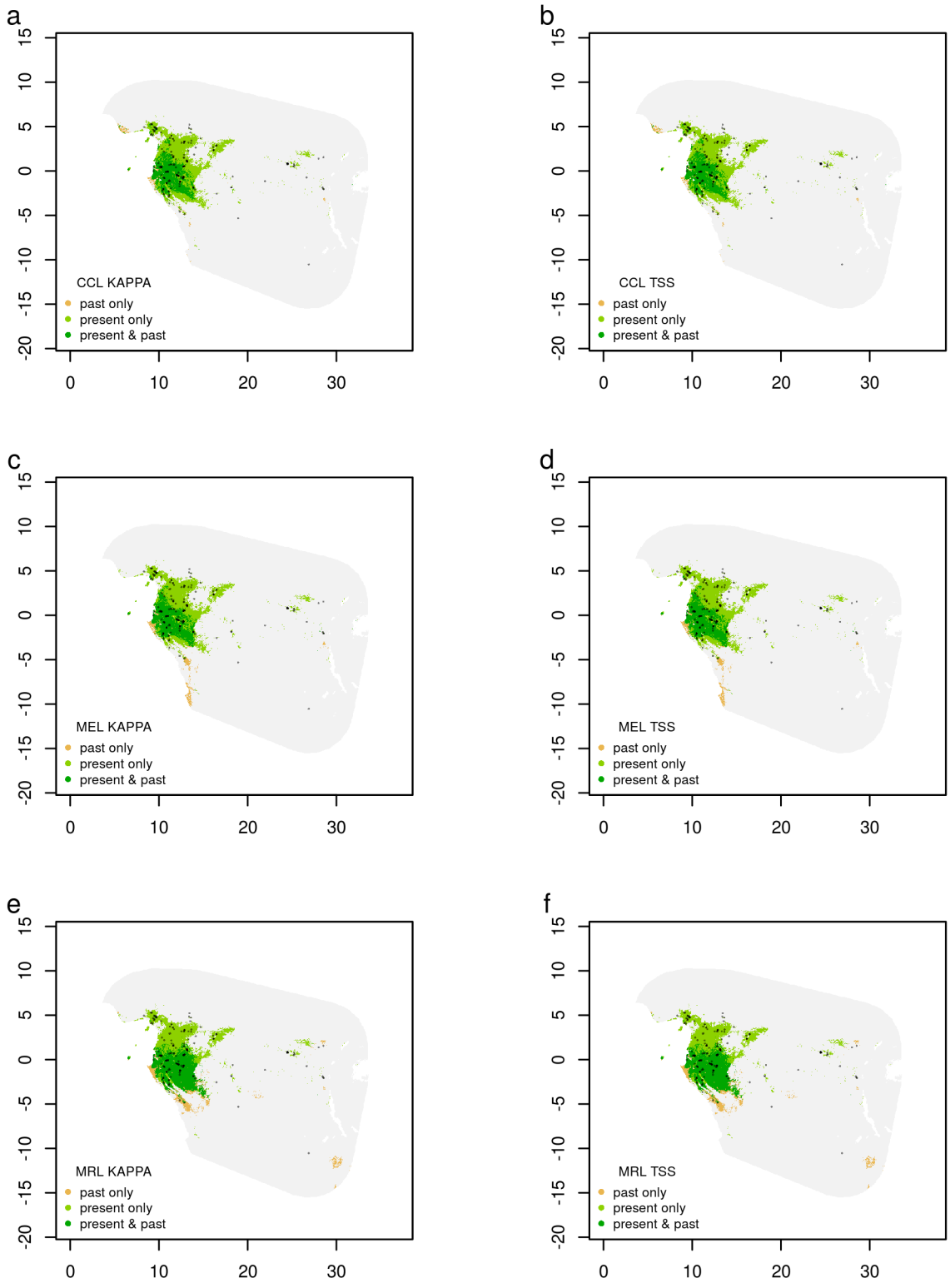


Fig. S32. Habitat suitability and stable areas inferred from ecological niche modelling for *A. mannii*. Suitability was summarised as binary maps for present and past climate, these were added together so that values of 0.5 indicate habitat only suitable in the past, values of 1 indicate areas only suitable in the present and values of 1.5 indicate areas suitable during both periods (stable areas). Results are from three different models of past climate: CCL (a-b), MEL (c-d), and MRL (e-f). Two statistics were used to infer binary maps: KAPPA (a, c, e) and the true skill statistic (TSS; b, d, f). Individual locations used in the modelling are shown as black points on each plot.

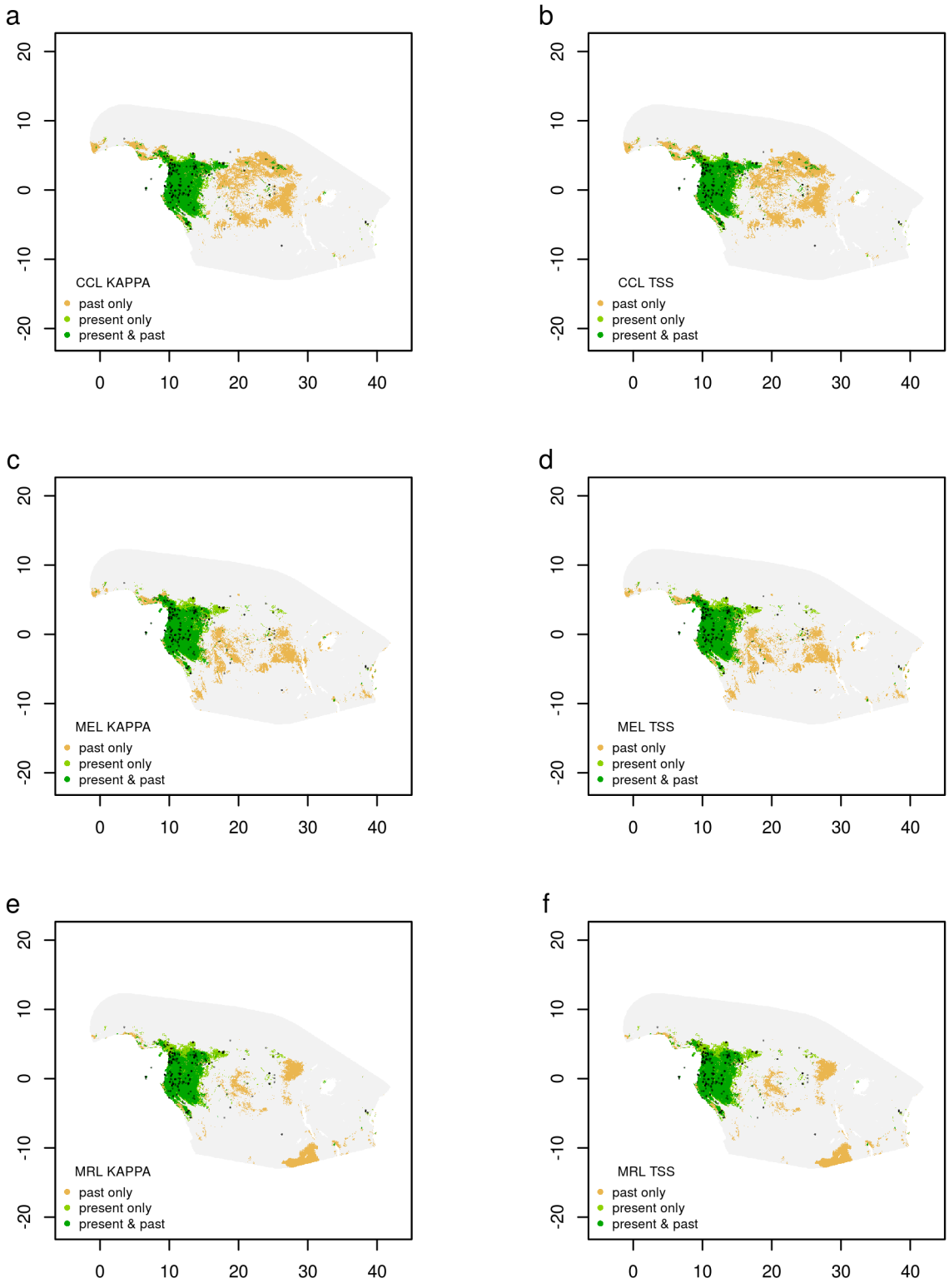


Fig. S33. Habitat suitability and stable areas inferred from ecological niche modelling for *G. suaveolens*. Suitability was summarised as binary maps for present and past climate, these were added together so that values of 0.5 indicate habitat only suitable in the past, values of 1 indicate areas only suitable in the present and values of 1.5 indicate areas suitable during both periods (stable areas). Results are from three different models of past climate: CCL (a-b), MEL (c-d), and MRL (e-f). Two statistics were used to infer binary maps: KAPPA (a, c, e) and the true skill statistic (TSS; b, d, f). Individual locations used in the modelling are shown as black points on each plot.

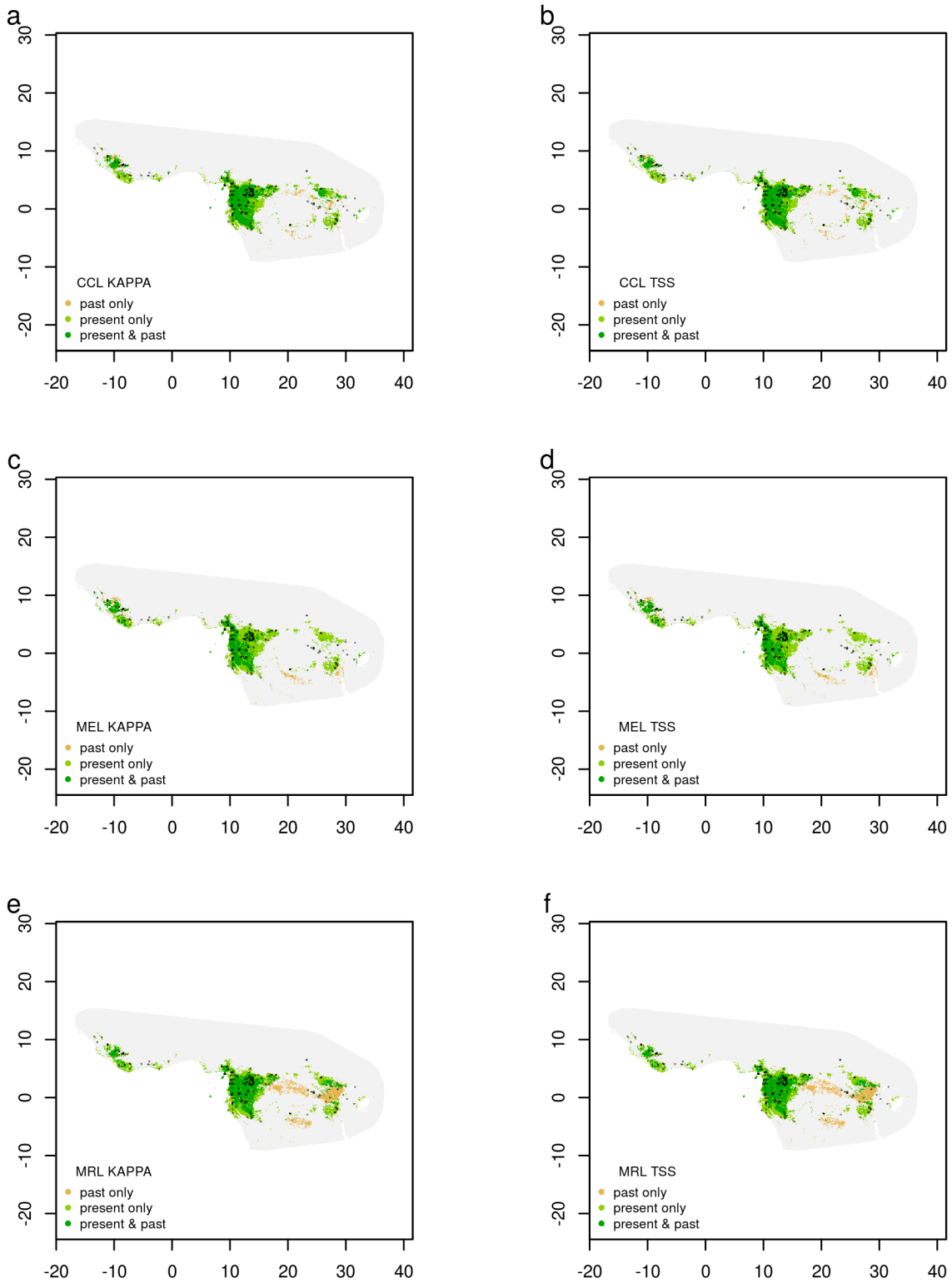


Fig. S34. Habitat suitability and stable areas inferred from ecological niche modelling for *M. enghiana*. Suitability was summarised as binary maps for present and past climate, these were added together so that values of 0.5 indicate habitat only suitable in the past, values of 1 indicate areas only suitable in the present and values of 1.5 indicate areas suitable during both periods (stable areas). Results are from three different models of past climate: CCL (a-b), MEL (c-d), and MRL (e-f). Two statistics were used to infer binary maps: KAPPA (a, c, e) and the true skill statistic (TSS; b, d, f). Individual locations used in the modelling are shown as black points on each plot.

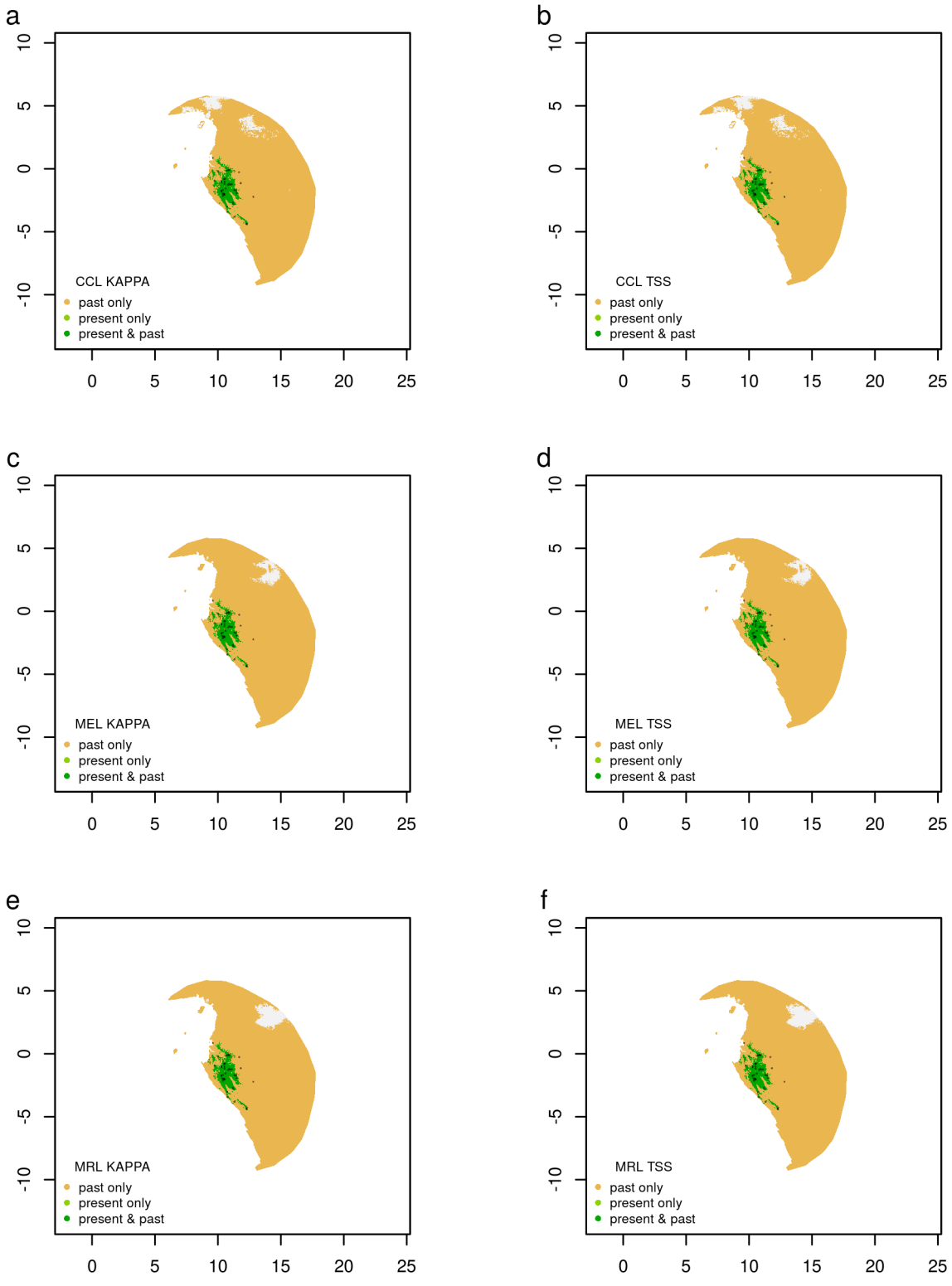


Fig. S35. Habitat suitability and stable areas inferred from ecological niche modelling for *P. acaulis*. Suitability was summarised as binary maps for present and past climate, these were added together so that values of 0.5 indicate habitat only suitable in the past, values of 1 indicate areas only suitable in the present and values of 1.5 indicate areas suitable during both periods (stable areas). Results are from three different models of past climate: CCL (a-b), MEL (c-d), and MRL (e-f). Two statistics were used to infer binary maps: KAPPA (a, c, e) and the true skill statistic (TSS; b, d, f). Individual locations used in the modelling are shown as black points on each plot.

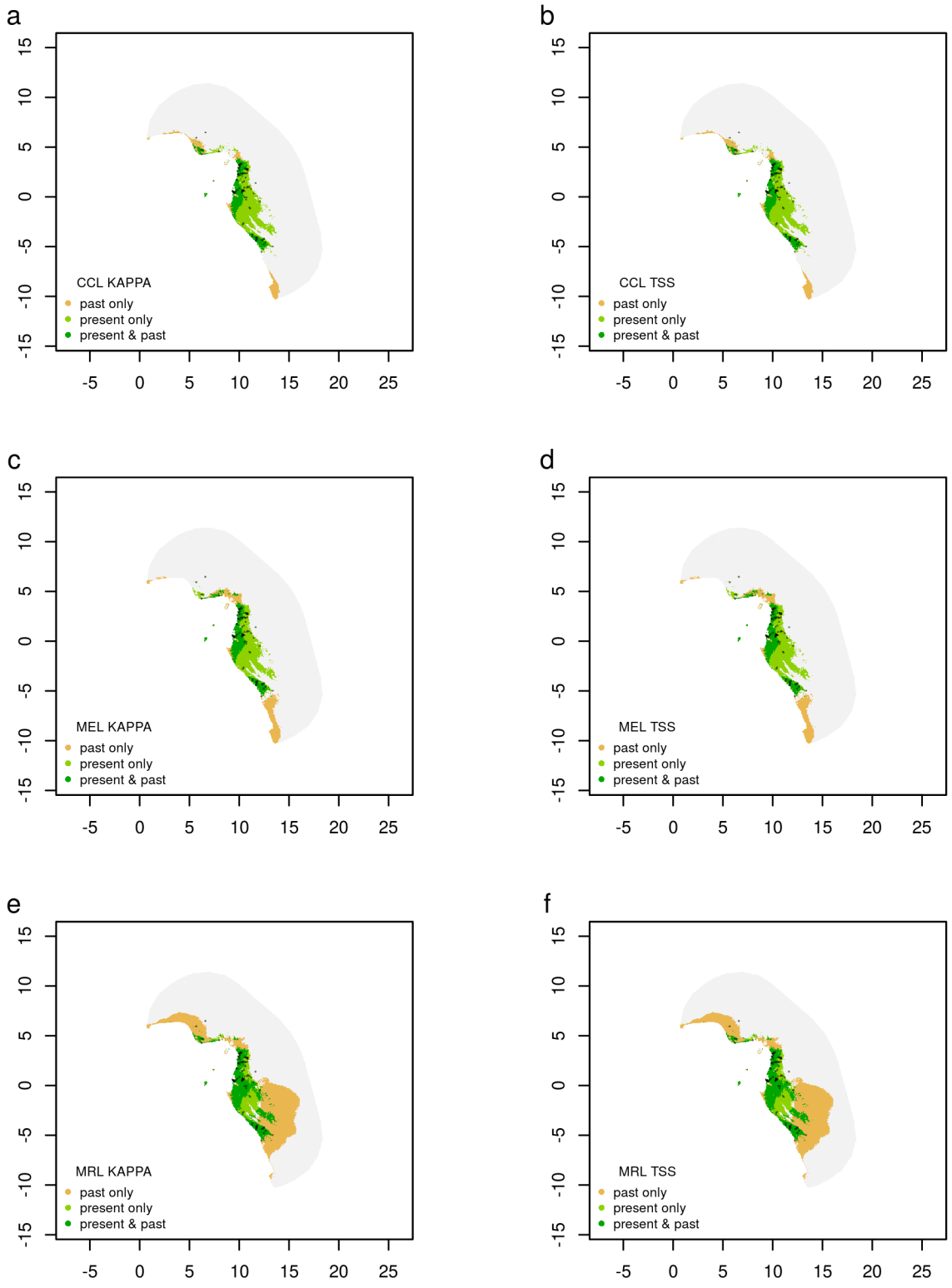


Fig. S36. Habitat suitability and stable areas inferred from ecological niche modelling for *P. barteri*. Suitability was summarised as binary maps for present and past climate, these were added together so that values of 0.5 indicate habitat only suitable in the past, values of 1 indicate areas only suitable in the present and values of 1.5 indicate areas suitable during both periods (stable areas). Results are from three different models of past climate: CCL (a-b), MEL (c-d), and MRL (e-f). Two statistics were used to infer binary maps: KAPPA (a, c, e) and the true skill statistic (TSS; b, d, f). Individual locations used in the modelling are shown as black points on each plot.

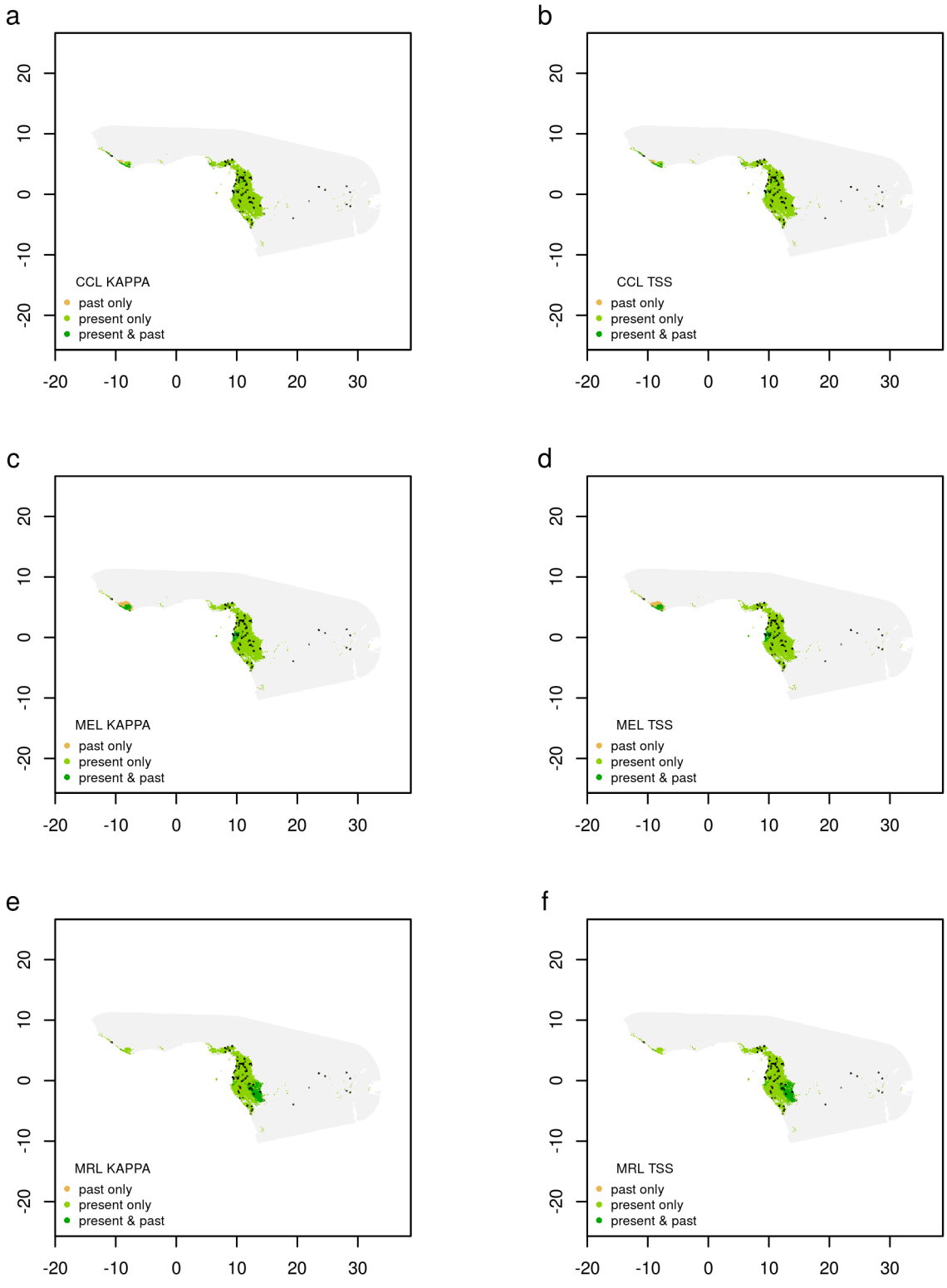


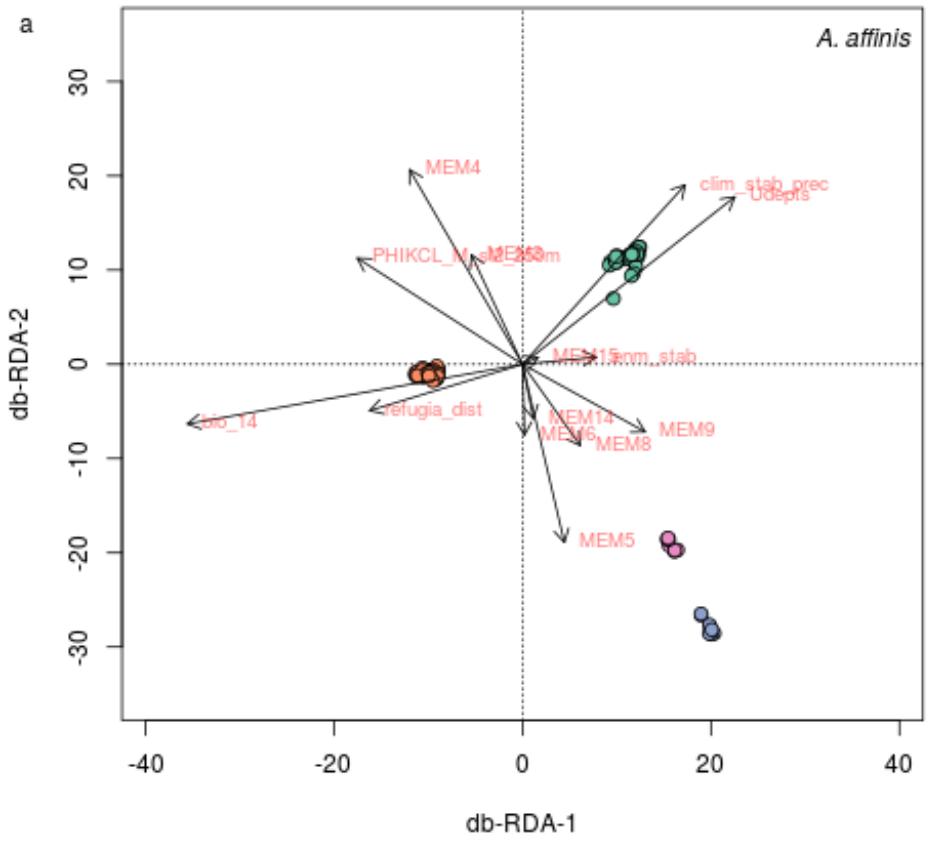
Fig. S37. Habitat suitability and stable areas inferred from ecological niche modelling for *S. mannii*. Suitability was summarised as binary maps for present and past climate, these were added together so that values of 0.5 indicate habitat only suitable in the past, values of 1 indicate areas only suitable in the present and values of 1.5 indicate areas suitable during both periods (stable areas). Results are from three different models of past climate: CCL (a-b), MEL (c-d), and MRL (e-f). Two statistics were used to infer binary maps: KAPPA (a, c, e) and the true skill statistic (TSS; b, d, f). Individual locations used in the modelling are shown as black points on each plot.

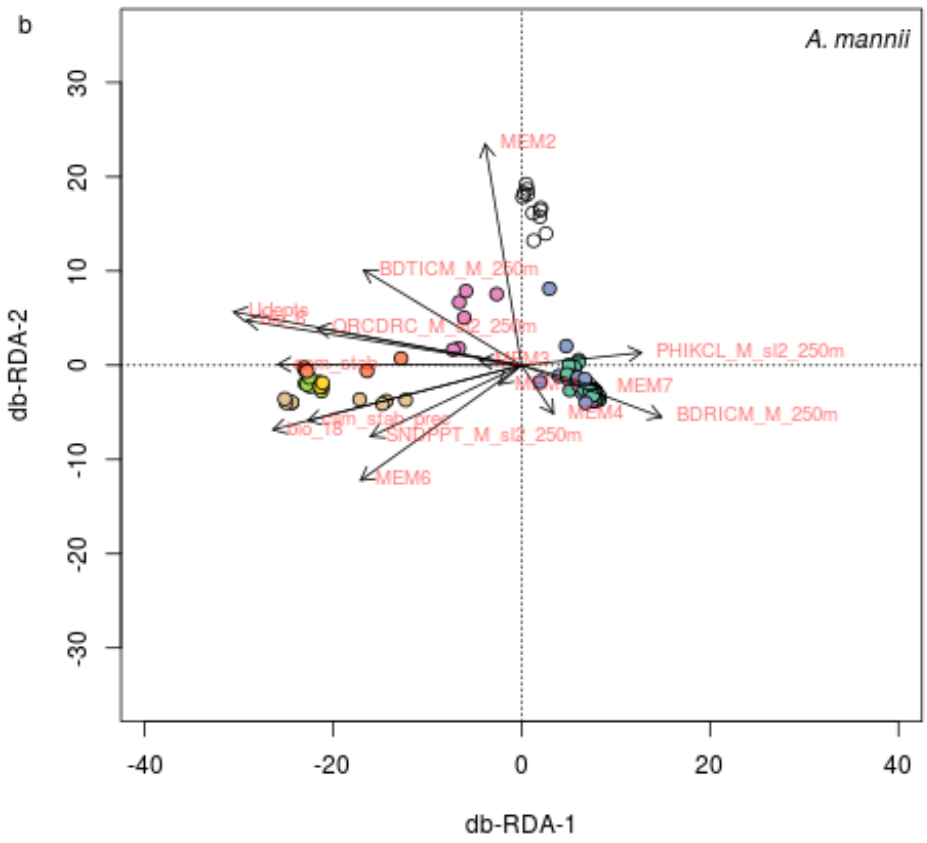
variable	model	statistic	A_affinis	A_mannii	G_suaveolens	M_enghiana	P_acaulis	P_barteri	S_mannii
KAPPA	ANN	Cutoff	819.10	768.30	810.90	841.25	842.50	885.05	795.80
KAPPA	ANN	Sensitivity	60.33	62.97	66.57	63.50	61.15	63.36	62.74
KAPPA	ANN	Specificity	97.55	97.59	97.53	97.34	98.00	98.47	97.59
KAPPA	ANN	Testing.data	0.45	0.45	0.51	0.51	0.52	0.55	0.45
KAPPA	GBM	Cutoff	617.75	592.10	598.20	603.50	613.90	616.60	608.35
KAPPA	GBM	Sensitivity	62.14	62.27	69.94	64.48	77.47	61.52	69.53
KAPPA	GBM	Specificity	96.66	94.56	94.67	96.98	94.40	97.63	95.28
KAPPA	GBM	Testing.data	0.44	0.39	0.45	0.51	0.50	0.51	0.43
KAPPA	GLM	Cutoff	824.70	824.90	860.95	828.15	861.40	895.60	801.40
KAPPA	GLM	Sensitivity	58.92	59.38	66.39	60.05	66.49	62.02	54.59
KAPPA	GLM	Specificity	96.75	97.97	97.78	98.26	97.56	98.55	97.13
KAPPA	GLM	Testing.data	0.41	0.45	0.51	0.51	0.52	0.55	0.39
KAPPA	MAXENT	Cutoff	572.35	554.35	575.40	576.70	532.80	616.00	469.50
KAPPA	MAXENT	Sensitivity	56.12	55.18	63.02	62.38	64.03	65.58	57.12
KAPPA	MAXENT	Specificity	99.47	99.34	99.40	99.38	99.24	99.45	99.37
KAPPA	MAXENT	Testing.data	0.59	0.54	0.64	0.66	0.64	0.68	0.54
KAPPA	RF	Cutoff	849.94	824.05	761.40	841.17	845.55	865.05	847.44
KAPPA	RF	Sensitivity	58.91	65.61	65.46	52.69	67.03	67.15	61.84
KAPPA	RF	Specificity	96.95	95.80	95.97	96.69	94.62	96.76	97.47
KAPPA	RF	Testing.data	0.40	0.38	0.43	0.44	0.41	0.48	0.40
TSS	ANN	Cutoff	288.00	278.70	282.50	241.20	209.90	257.30	262.30
TSS	ANN	Sensitivity	89.67	91.52	89.15	88.16	94.87	93.95	85.65
TSS	ANN	Specificity	82.98	72.77	81.62	82.30	71.97	72.90	91.89
TSS	ANN	Testing.data	0.80	0.80	0.78	0.79	0.81	0.83	0.78
TSS	GBM	Cutoff	340.70	201.10	189.85	289.50	370.95	291.70	407.35
TSS	GBM	Sensitivity	89.66	88.85	89.87	85.14	91.80	89.21	87.46
TSS	GBM	Specificity	90.75	83.02	89.71	83.90	82.68	80.82	82.55
TSS	GBM	Testing.data	0.81	0.77	0.80	0.76	0.82	0.78	0.77
TSS	GLM	Cutoff	366.60	425.60	442.30	422.10	393.90	392.10	327.00
TSS	GLM	Sensitivity	91.01	90.29	92.88	92.26	94.02	94.11	86.65
TSS	GLM	Specificity	90.19	91.62	90.77	82.87	90.93	90.82	81.71
TSS	GLM	Testing.data	0.81	0.82	0.84	0.82	0.85	0.85	0.76
TSS	MAXENT	Cutoff	62.70	44.10	45.90	39.10	29.10	66.30	68.10
TSS	MAXENT	Sensitivity	90.97	90.65	93.91	91.88	93.36	92.43	93.73
TSS	MAXENT	Specificity	83.21	92.06	74.63	74.00	73.94	65.82	74.42
TSS	MAXENT	Testing.data	0.81	0.83	0.84	0.80	0.82	0.82	0.79
TSS	RF	Cutoff	467.56	414.75	336.00	434.22	504.45	450.70	370.61
TSS	RF	Sensitivity	91.04	91.39	93.92	89.87	88.84	89.74	90.93
TSS	RF	Specificity	88.81	89.51	76.63	86.54	89.67	89.52	89.99
TSS	RF	Testing.data	0.80	0.81	0.78	0.76	0.79	0.79	0.81

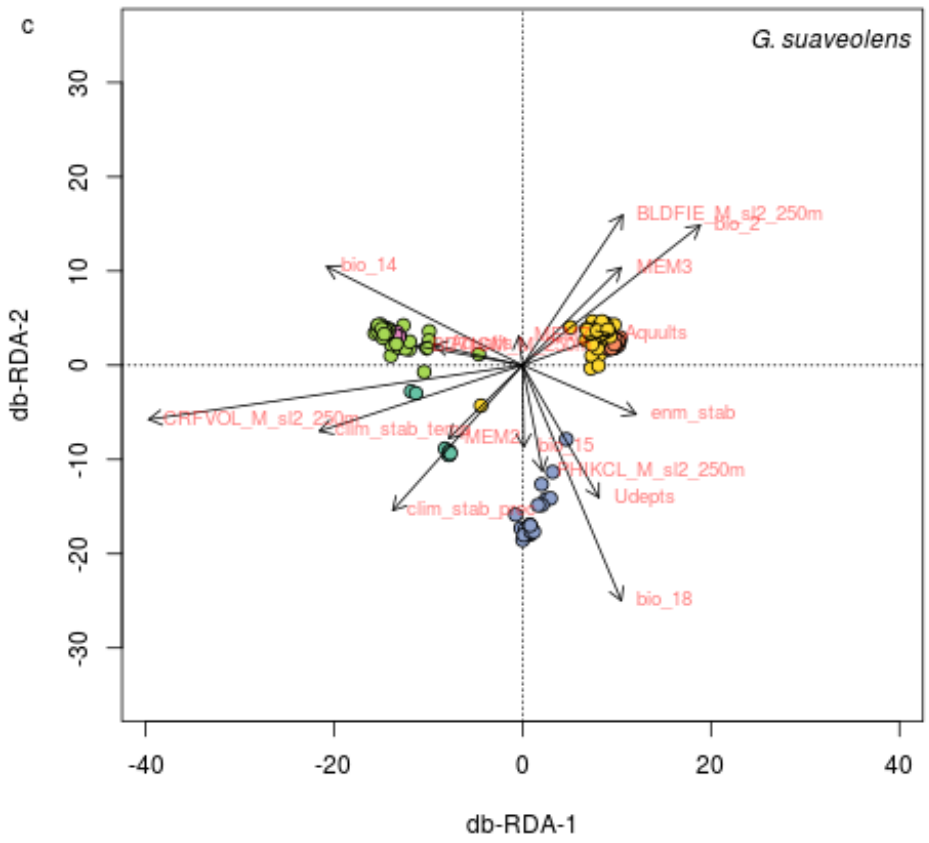
Table S6. Evaluation of ecological niche models extracted frim 'models.out' files from biomod2 runs. Values for each pseudoabsence replicate were averaged across species.

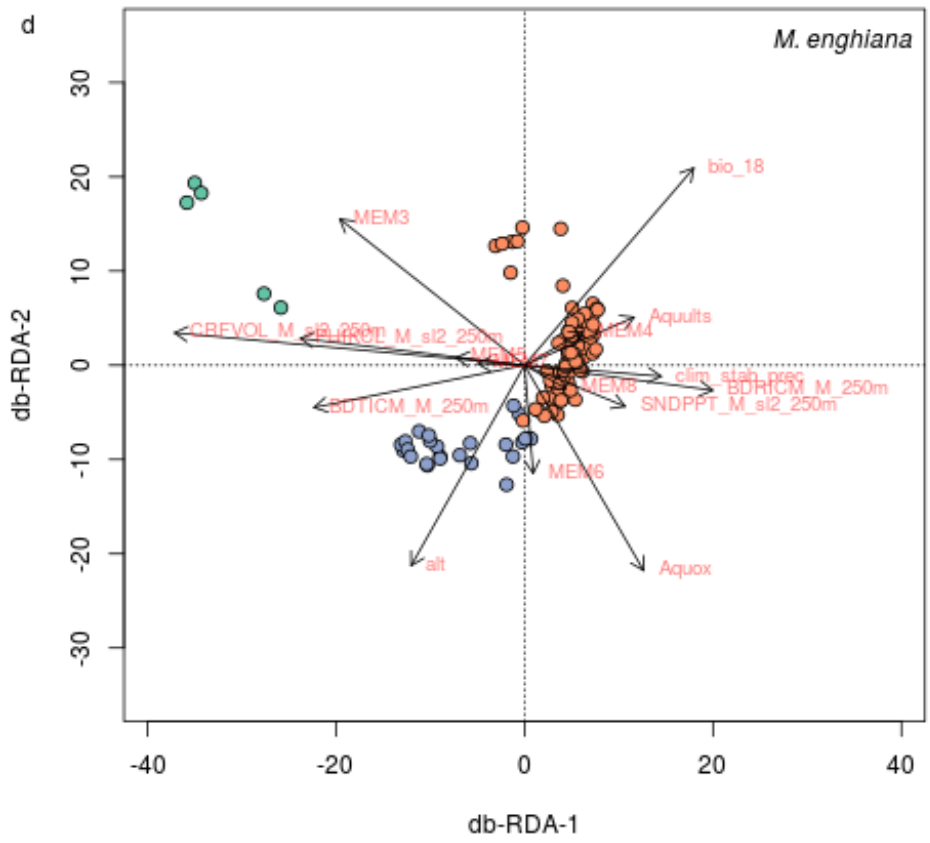
distance-based RedunDancy Analysis (db-RDA)

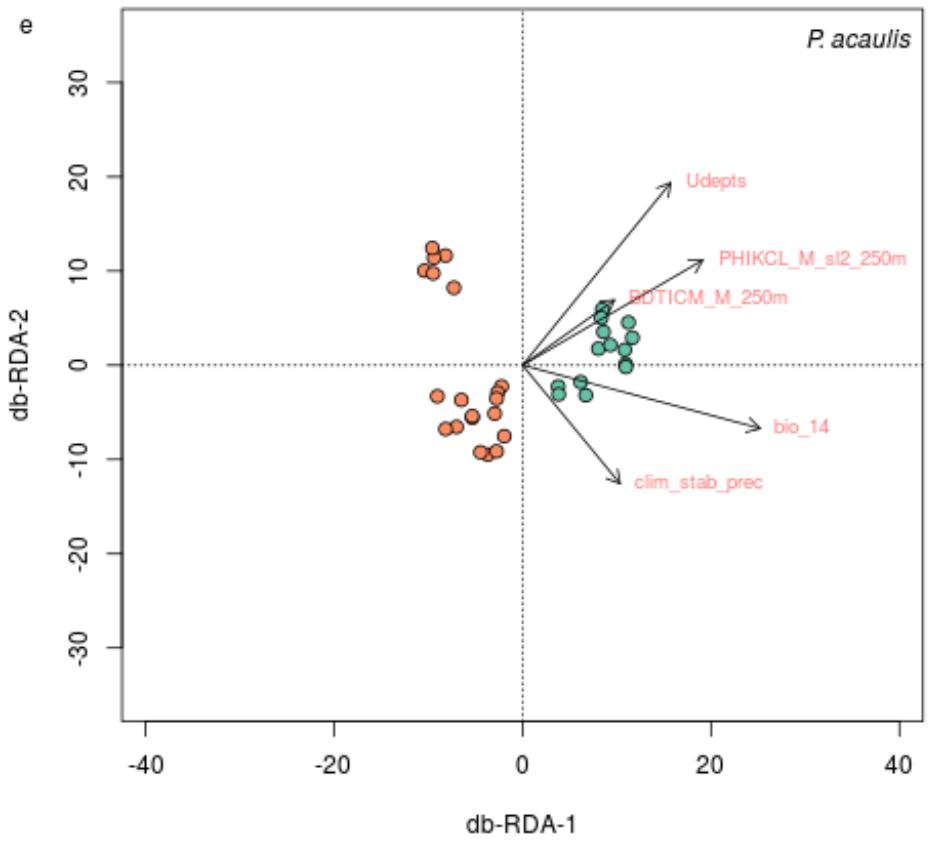
We used db-RDA to uncover how genetic variance could be attributed to different geographic, historic and biological variables. Our approach was based on the following tutorial (https://github.com/laurabenestan/db-RDA-and-db-MEM/). Genetic distances were calculated using Principal Coordinates Analysis (PCoA) on a genetic distance matrix, which was used as a response variable. We used Moran Eigenvector's Maps (MEMs) to summarise spatial structure in the data. This multivariate approach outputs db-MEMs where the first explain patterns at large spatial scales, with spatial scale decreasing with each db-MEM. Explanatory variables included those variables used for ENMs, MEMs, several variables related to stable areas and USDA soil taxonomy. A list of the explanatory variables used can be found in Table S5. Unlike the AMOVA analyses, climate-based measures of stability for rainfall and temperature were kept separate for db-RDA analyses. Correlation among explanatory variables was examined and variables were pruned based on a threshold of 0.9 (Pearson). We then used Variance Inflation Factor (VIF) to look for multicolinearity within the model. Variables with a VIF of >10 were iteratively removed the model, starting with the variable with the largest VIF until all VIF values were <10. We built our db-RDA model by adding and removing variables in order to maximise the explained variance, but ensuring that the amount explained did not exceed the that of the full model, to avoid overfitting. Once this process finished, the final model was obtained and we visualized the results using biplots, shown below. The relative importance of different categories of explanatory variables for each species is shown in Figure 5 in the main text.

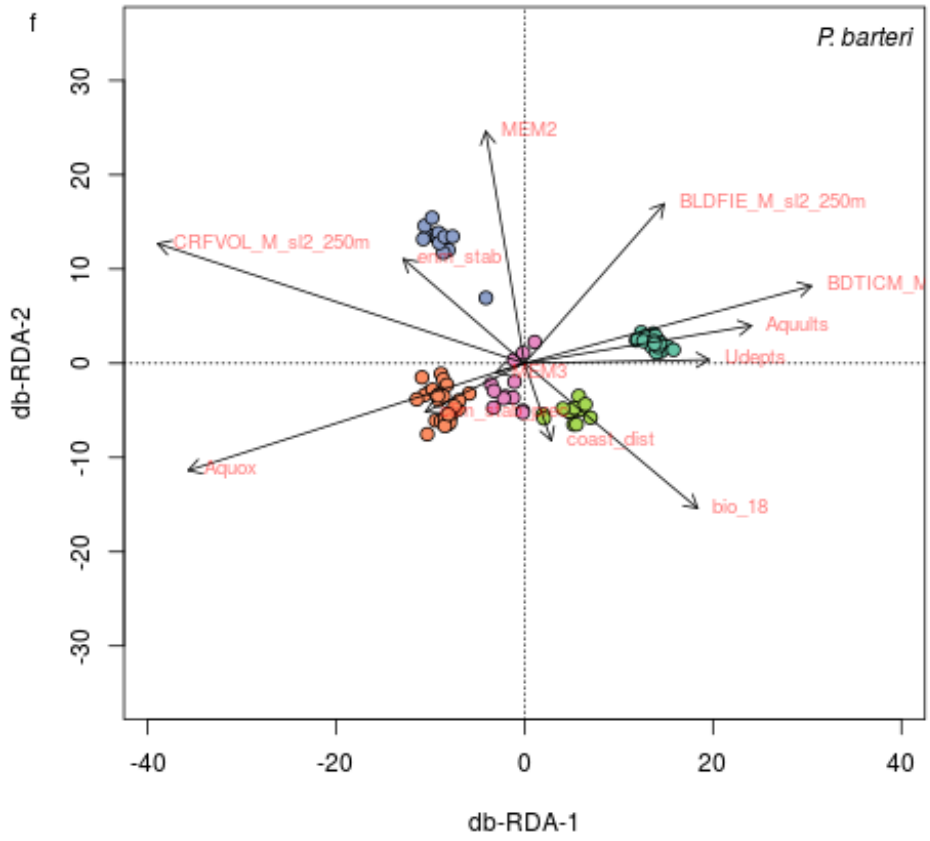



b
A. mannii
MEM2
BDTICM_M_250m
ORCDRC_M_sl2_250m
PHIKCL_M_sl2_250m
MEM7
MEM4
BDRICM_M_250m
SNDPPT_M_sl2_250m
MEM6
db-RDA-2
db-RDA-1
30
20
10
0
-10
-20
-30
-40
-20
0
20
40


c
G. suaveolens
BLDFIE_M_sl2_250m
bio_2
bio_14
MEM3
Aquults
enm_stab
CRFVOL_M_sl2_250m
clim_stab_temp
MEM2
bio_15
PHIKCL_M_sl2_250m
Udepts
clim_stab_prec
bio_18
db-RDA-2
db-RDA-1
30
20
10
0
-10
-20
-30
-40
-20
0
20
40


d
M. enghiana
db-RDA-2
db-RDA-1
30
20
10
0
-10
-20
-30
-40
-20
0
20
40
bio_18
MEM3
Aquults
MEM4
clim_stab_prec
BDTICM_M_250m
SNDPPT_M_sl2_250m
MEM6
alt
Aquox


e
P. acaulis
Udepts
PHIKCL_M_sl2_250m
BDTICM_M_250m
bio_14
clim_stab_prec
db-RDA-2
db-RDA-1
30
20
10
0
-10
-20
-30
-40
-20
0
20
40


f
P. barteri
MEM2
BLDFIE_M_sl2_250m
CRFVOL_M_sl2_250m
BDTICM_M
Aquults
Udepts
MEM3
coast_dist
Aquox
bio_18
db-RDA-2
db-RDA-1
30
20
10
0
-10
-20
-30
-40
-20
0
20
40

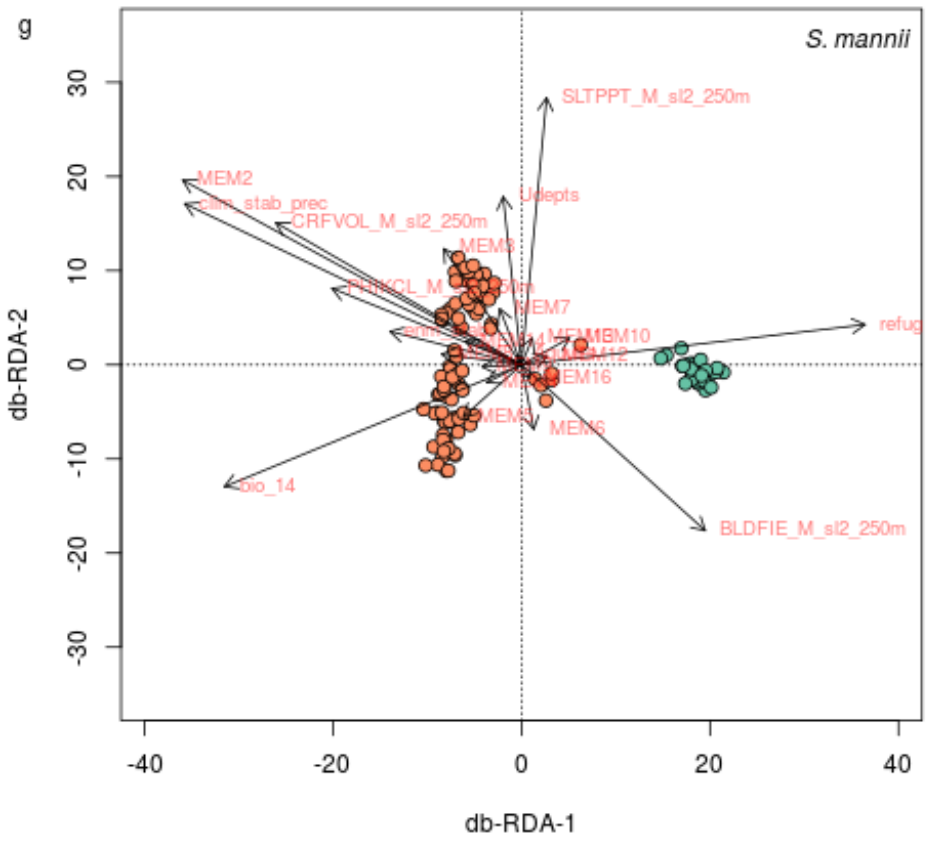


Fig. S38. Biplots of the distance-based redundancy analyses (db-RDA) for (a) *A. affinis*, (b) *A. mannii*, (c) *G. suaveolens*, (d) *M. enghiana*, (e) *P. acaulis*, (f) *P. barteri* and (g) *S. mannii*. The first two RDA axes are shown on the x and y axes. Arrows indicate the direction in which increasing values for variables, shown adjacent in red, explain genetic variance. The length of the arrow shows the level of variation explained by the associated variable. Individuals are represented as circles (point locations scaled by a factor of 10), coloured by their associated DAPC cluster.

	Species	Variable	Category	R-squared
[redacted]	anni	bio_14	Climate	0.12911
2	anni	PHIKCL_M_sl2_250m	Soil	0.07142
3	anni	clim_stab_prec	Stability	0.06003
4	anni	MEM5	Geography	0.04111
5	anni	MEM3	Geography	0.02306
6	anni	MEM4	Geography	0.02179
7	anni	MEM9	Geography	0.01858
8	anni	MEM6	Geography	0.01693
9	anni	enm_stab	Stability	0.01524
10	anni	Inceptisols	Soil	0.01939
11	anni	refugia_dist	Stability	0.00854
12	anni	MEM8	Geography	0.00851
13	anni	MEM15	Geography	0.00695
14	anni	MEM14	Geography	0.00712
16	anon	Inceptisols	Soil	0.09663
17	anon	MEM6	Geography	0.05654
18	anon	MEM2	Geography	0.05662
19	anon	bio_6	Climate	0.05389
20	anon	ORCDRC_M_sl2_250m	Soil	0.02507
21	anon	MEM7	Geography	0.02345
22	anon	SNDPPT_M_sl2_250m	Soil	0.01217
23	anon	enm_stab	Stability	0.01054
24	anon	clim_stab_prec	Stability	0.00866
25	anon	MEM3	Geography	0.00890
26	anon	PHIKCL_M_sl2_250m	Soil	0.01461
27	anon	bio_18	Climate	0.00901
28	anon	BDRICM_M_250m	Soil	0.00692
29	anon	MEM4	Geography	0.00667
30	anon	BDTICM_M_250m	Soil	0.00825
31	anon	MEM11	Geography	0.00468
33	green	CRFVOL_M_sl2_250m	Soil	0.15570
34	green	bio_18	Climate	0.06829
35	green	bio_2	Climate	0.02023
36	green	bio_14	Climate	0.01987
37	green	PHIKCL_M_sl2_250m	Soil	0.00989
38	green	Alfisols	Soil	0.00967
39	green	BLDFIE_M_sl2_250m	Soil	0.00532
40	green	Inceptisols	Soil	0.00477
41	green	bio_15	Climate	0.00411
42	green	BDTICM_M_250m	Soil	0.00487
43	green	Ultisols	Soil	0.00391
44	green	MEM2	Geography	0.00340
45	green	MEM3	Geography	0.00356
46	green	enm_stab	Stability	0.00359
47	green	clim_stab_prec	Stability	0.00362
48	green	MEM7	Geography	0.00265
49	green	clim_stab_temp	Stability	0.00245
51	mona	CRFVOL_M_sl2_250m	Soil	0.04345
52	mona	bio_18	Climate	0.02349
53	mona	MEM3	Geography	0.01622
54	mona	SNDPPT_M_sl2_250m	Soil	0.01124
55	mona	alt	Geography	0.01269
56	mona	BDRICM_M_250m	Soil	0.00756
57	mona	MEM7	Geography	0.00747
58	mona	clim_stab_prec	Stability	0.00431
59	mona	MEM4	Geography	0.00455
60	mona	BDTICM_M_250m	Soil	0.00352
61	mona	MEM6	Geography	0.00349
62	mona	PHIKCL_M_sl2_250m	Soil	0.00321
63	mona	MEM8	Geography	0.00207

64	mona	Oxisols	Soil	0.00207
65	mona	Ultisols	Soil	0.00272
66	mona	MEM5	Geography	0.00300
68	podoa	bio_14	Climate	0.03020
69	podoa	Inceptisols	Soil	0.02291
70	podoa	clim_stab_prec	Stability	0.02255
71	podoa	BDTICM_M_250m	Soil	0.01590
72	podoa	PHIKCL_M_sl2_250m	Soil	0.01027
74	podob	CRFVOL_M_sl2_250m	Soil	0.06950
75	podob	Oxisols	Soil	0.03819
76	podob	bio_18	Climate	0.03055
77	podob	BLDFIE_M_sl2_250m	Soil	0.01924
78	podob	coast_dist	Stability	0.01501
79	podob	MEM2	Geography	0.01285
80	podob	Ultisols	Soil	0.01176
81	podob	Inceptisols	Soil	0.00772
82	podob	BDTICM_M_250m	Soil	0.00900
83	podob	MEM3	Geography	0.00522
84	podob	clim_stab_prec	Stability	0.00408
85	podob	enm_stab	Stability	0.00315
87	sclero	MEM2	Geography	0.05109
88	sclero	bio_14	Climate	0.05146
89	sclero	SLTPPT_M_sl2_250m	Soil	0.01623
90	sclero	MEM3	Geography	0.01969
91	sclero	Inceptisols	Soil	0.00988
92	sclero	MEM6	Geography	0.00877
93	sclero	MEM7	Geography	0.00595
94	sclero	MEM5	Geography	0.00509
95	sclero	MEM8	Geography	0.00518
96	sclero	MEM10	Geography	0.00403
97	sclero	PHIKCL_M_sl2_250m	Soil	0.00379
98	sclero	CRFVOL_M_sl2_250m	Soil	0.00313
99	sclero	Ultisols	Soil	0.00279
100	sclero	BLDFIE_M_sl2_250m	Soil	0.00266
101	sclero	refugia_dist	Stability	0.00193
102	sclero	clim_stab_prec	Stability	0.00293
103	sclero	MEM14	Geography	0.00278
104	sclero	MEM9	Geography	0.00276
105	sclero	MEM4	Geography	0.00234
106	sclero	MEM12	Geography	0.00191
107	sclero	MEM13	Geography	0.00162
108	sclero	enm_stab	Stability	0.00156
109	sclero	MEM16	Geography	0.00110

Table S7. Variable importance for final db-RDA models. For each species the variables present in the final db-RDA model are shown and the corresponding R-squared value. The category each model falls into is also noted.

References

1. Oksanen J, et al. (2019) vegan: Community Ecology Package. R package version 2.5-2. *Cran R*.
2. Maley J (1996) The African rain forest - Main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh Section B: Biological Sciences* 104:31–73.
3. Anhuf D, et al. (2006) Paleo-environmental change in Amazonian and African rainforest during the LGM. *Palaeogeography, Palaeoclimatology, Palaeoecology* 239(3-4):510–527.
4. Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131(2):479–491.
5. Kamvar ZN, Tabima JF, Grunwald NJ (2014) Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2014(1):1–14.
6. Upham NNS, Esselstyn JA, Jetz W (2019) Inferring the mammal tree: species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biology* 17(12):e3000494.
7. Andermann T, et al. (2019) Allele Phasing Greatly Improves the Phylogenetic Utility of Ultraconserved Elements. *Systematic Biology* 68(1):32–46.
8. Andermann T, Cano Á, Zizka A, Bacon C, Antonelli A (2018) SECAPR-A bioinformatics pipeline for the rapid and user-friendly processing of targeted enriched Illumina sequences, from raw reads to alignments. *PeerJ* 2018(7):e5175.
9. Jones GR (2019) Divergence Estimation in the Presence of Incomplete Lineage Sorting and Migration. *Systematic Biology* 68(1):19–31.
10. Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution* 27(4):905–920.
11. Leaché AD, Harris RB, Rannala B, Yang Z (2014) The influence of gene flow on species tree estimation: A simulation study. *Systematic Biology* 63(1):17–30.
12. Warren DL, Glor RE, Turelli M (2010) ENMTools: A toolbox for comparative studies of environmental niche models. *Ecography* 33(3):607–611.
13. Thuiller W, Georges D, Engler R (2013) biomod2: Ensemble platform for species distribution modeling. *R package version* 2(7):r560.
14. Dauby G, et al. (2017) ConR: An R package to assist large-scale multispecies preliminary conservation assessments using distribution data. *Ecology and Evolution* 7(24):11292–11303.