Supplemental Information for:

Genomic evidence of introgression and adaptation in a model subtropical tree species, *Eucalyptus grandis*

Marja M. Mostert-O'Neill, S. Melissa Reynolds, Juan J. Acosta, David J. Lee, Justin O. Borevitz, Alexander A. Myburg

Table of Contents:

Appendix S1 *Supplementary methods*

Cryptic relatedness assessment through identity by decent analysis
In order to assess the relatedness among individuals and subset a balanced population dataset, identity by decent analysis was conducted in SNP & Variation Suite v8.x (SVS8, Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com). Pairwise identity by decent estimates were calculated (Purcell *et al.* 2007) and used to select two half-siblings (the two half-sib seedlings with the lowest pair-wise identity by decent estimate in each family) from families with more than two seedlings represented. Identity by decent analysis of the balanced sampling is given in **Figure S1**.

Efficient Inference of Local Ancestry (*EILA*) testing and optimization of breakpoint penalty variable, λ, using samples of known ancestry
To test and optimise the *EILA* method (Yang *et al.* 2013) for inferring *E. grandis* or non-*E. grandis* ancestry in all samples from the native range, a set of control samples with known ancestry was used. These included two *E. grandis* from the native range as positive controls - 001_1 (Windsor 1 provenance); 220_1 (Orara West 72 provenance), two *E. nitens*, two *E. dunnii*, two *E. urophylla*, one *E. nitens* x *E. grandis* (NG) F1 hybrid, one *E. grandis* x *E. nitens* (GN) F1 hybrid, one *E. grandis* x *E. nitens* backcross to *E. grandis* (GNxG), one *E. grandis* x *E. urophylla* (GU) F1 hybrid, one *E. grandis* x *E. urophylla* backcross to *E. urophylla* (GUxU) hybrid, and two three-way hybrids of crosses between a *E. grandis* x *E. nitens* F1 hybrid and a *E. grandis* x *E. urophylla* F1 hybrid (GNxGU). These samples were provided by two forestry companies, Sappi and Mondi, and previously genotyped using the EUChip60K SNP chip (unpublished).

The values of λ tested were 15 (default), 20, 25, 30, 35, and 40. Breakpoint penalties below 30 had numerous improbable segment assignments. At $\lambda = 40$, improbable assignments were resolved in the control samples, however, this improvement was only marginal over the assignments at $\lambda = 30$ (**Figure S3**). Since natural introgression events will likely involve smaller genomic sections than those resolved in the F1 and backcross hybrid controls, we report $\lambda = 15$ and $\lambda = 30$ for suspected introgressed provenances (**Figure S8-S10**), and $\lambda = 30$ for all other provenances (**Table S6**).

Principal component analysis of environmental variables
To reduce redundant Environmental Association Analyses (EAA) for variables that are highly correlated, Principal Component Analysis (PCA) was applied to the 11 environmental variables, excluding latitude and longitude, obtained from Atlas of Living Australia (ALA; occurrence data download at http://www.ala.org.au. Accessed 28 November 2019) using the *prcomp* function in R with scaled data to account for magnitude differences in units of measure. The analysis was first performed on all provenances across the species range sampled, then on the Northern and Southern subpopulations, separately. Environmental variable contributions the each of the first four principal components, also referred to as the variable loadings (Holland 2019), are reported in **Table S3**, along with the percentage variance explained by each of the first four principal components.

Table S1 Seed stocks and study population

Location/provenance	Region	Subpop.†	Latitude	Longitude	Elevation (m)	Köppen classification	Families/ seedlots	Supplier‡
Windsor Tableland 1	Windsor	North	-16.183	144.983	1161	Subtropical: no dry season	8	CSIRO
Windsor Tableland 2	Windsor	North	-16.200	145.017	1134	Subtropical: no dry season	6	CSIRO
Copperlode	Copperlode	North	-16.967	145.667	368	Subtropical: no dry season	20	CSIRO§ & DAF
Tinaroo Creek Road	Atherton	North	-17.083	145.567	925	Subtropical: no dry season	8	CSIRO§
Tinaroo Dam	Atherton	North	-17.133	145.567	893	Subtropical: no dry season	1	DAF
Wild River, QLD	Atherton	North	-17.367	145.400	927	Subtropical: moderately dry winter	1	DAF
Mt. Spec, Paluma	Paluma	North	-18.933	146.117	822	Subtropical: no dry season	15	CSIRO§
Paluma	Paluma	North	-19.000	146.150	796	Subtropical: no dry season	3	DAF
Eungella	Eungella	Mackay	-21.075	148.632	294	Subtropical: no dry season	7	CSIRO§
Finch Hatton Gorge	Finch Hatton Gorge	Mackay	-21.197	148.504	753	Subtropical: no dry season	10	CSIRO§
Brooweena	Brooweena	South	-25.550	152.267	210	Subtropical: no dry season	3	CSIRO
Kin Kin¶	Sunshine Coast	South	-26.263	152.876	142	Subtropical: no dry season	4	CSIRO
Woondum	Sunshine Coast	South	-26.300	152.817	306	Subtropical: no dry season	6	CSIRO§
Yabba	Sunshine Coast	South	-26.533	152.400	475	Temperate: no dry season (hot summer)	9	CSIRO
Mapleton	Sunshine Coast	South	-26.600	152.867	325	Subtropical: no dry season	9	CSIRO§
Conondale	Sunshine Coast	South	-26.667	152.600	498	Temperate: no dry season (warm summer)	9	CSIRO§‡
Beerwah	Sunshine Coast	South	-26.850	153.000	20	Subtropical: no dry season	1	DAF
Mt. Mee	Sunshine Coast	South	-27.133	152.717	368	Subtropical: no dry season	7	CSIRO
Mt. Tamborine	Mt. Tamborine & Mt Lindsay	South	-27.917	153.183	396	Subtropical: no dry season	6	CSIRO
Mt. Lindsay	Mt. Tamborine & Mt Lindsay	South	-28.350	152.750	352	Temperate: no dry season (warm summer)	11	CSIRO§
Cherry Tree	Cherry Tree	South	-28.900	152.817	236	Subtropical: no dry season	3	CSIRO§
Newfoundland	Newfoundland	South	-29.946	153.162	52	Subtropical: no dry season	1	DAF

† Subpopulation assignment based on geographic and genetically distinct groupings.

‡ CSIRO - Australian Tree Seed Centre, Commonwealth Scientific and Industrial Research Organisation; DAF – Department of Agriculture and Fisheries.

§ A selection of 108 seedlots will be established in large common garden trials in various exotic locations by Camcore (NC State University, Raleigh, NC).

¶ GPS coordinates provided for Kin Kin (-26.200;153.167) appeared improbable, as such the coordinates for Kin Kin the town were used as proxy environment.

Table S1 (continued) *Seed stocks and study population*

Location/provenance	Region	Subpop.†	Latitude	Longitude	Elevation (m)	Köppen classification	Families/ seedlots	Supplier‡
Bagawa	Coffs Harbour	South	-30.117	152.900	299	Temperate: no dry season (warm summer)	10	CSIRO
Wedding Bells	Coffs Harbour	South	-30.167	153.117	95	Subtropical: no dry season	8	CSIRO
Lower Bucca	Coffs Harbour	South	-30.178	153.074	108	Temperate: no dry season (hot summer)	1	DAF
Orara East	Coffs Harbour	South	-30.225	153.068	195	Temperate: no dry season (hot summer)	1	DAF
Orara West (71)	Coffs Harbour	South	-30.332	152.973	189	Temperate: no dry season (warm summer)	1	DAF
Orara West (72)	Coffs Harbour	South	-30.341	153.011	353	Temperate: no dry season (hot summer)	10	CSIRO
Tuckers Nob	Coffs Harbour	South	-30.372	152.951	259	Temperate: no dry season (warm summer)	2	DAF
Pine Creek	Coffs Harbour	South	-30.414	153.009	18	Subtropical: no dry season	2	DAF
Gladstone	Coffs Harbour	South	-30.509	152.838	242	Temperate: no dry season (hot summer)	2	DAF
Newry	Coffs Harbour	South	-30.536	152.956	21	Subtropical: no dry season	2	DAF
Bulahdelah	Bulahdelah	South	-32.338	152.149	119	Temperate: no dry season (warm summer)	9	CSIRO§

† Subpopulation assignment based on geographic and genetically distinct groupings.

‡ CSIRO - Australian Tree Seed Centre, Commonwealth Scientific and Industrial Research Organisation; DAF – Department of Agriculture and Fisheries.

§ A selection of 108 seedlots will be established in large common garden trials in various exotic locations by Camcore (NC State University, Raleigh, NC).

¶ GPS coordinates provided for Kin Kin (-26.200;153.167) appeared improbable, as such the coordinates for Kin Kin the town were used as proxy environment.

Table S2 Atlas of Living Australia (ALA) data layers (occurrence data downloaded at http://www.ala.org.au. Accessed 28 November 2019)

Environmental data layer	Description on ALA	ALA layer metadata ID
Aridity index	Mean annual aridity index	715
Distance to coast (km)	Distance - to coast	819
Elevation (m)	Elevation (metres above mean sea level)	674
Nitrogen pre-European (kg/ha)	Pre-European estimate of mean annual store of total plant-available soil nitrogen (NTot0.Base)	831
Phosphorus pre-European (kg/ha)	Pre-European estimate of mean annual store of plant-available mineral phosphorus (PMnl0.Base)	811
Precipitation driest period (Bio14, mm)	Precipitation - driest period (Bio14)	872
Precipitation wettest period (Bio13, mm)	Precipitation - wettest period (Bio13)	866
Soil depth (m)	Solum depth (surface and subsoil layers) (metres)	816
Soil pedality	Grade of soil pedal structure (grade)	841
Temperature coldest period min (Bio06, °C)	Temperature - coldest period min (Bio06)	867
Temperature warmest period min (Bio05, °C)	Temperature - warmest period max (Bio05)	879

Table S3 Environmental variable contributions (loadings[†]) to principal components based on PCA of the species range, Northern and Southern subpopulations, respectively.

Environmental variable	PC1	PC2	PC3	PC4
A) Species range				
Variance explained	36.19%	26.67%	12.49%	11.18%
Aridity index	-0.7358	-0.34818	0.463489	0.098947
Distance to coast	-0.01064	0.571959	0.054696	-0.75133
Elevation	-0.82003	0.232567	0.302009	-0.36352
Nitrogen pre-European	0.354505	-0.79569	-0.27007	-0.16857
Phosphorus pre-European	-0.25321	-0.58885	-0.5594	-0.38201
Precipitation wettest period (Bio13)	-0.96854	0.022773	-0.03245	0.076564
Precipitation driest period (Bio14)	0.473295	-0.61862	0.397747	0.253254
Soil depth	-0.67319	-0.56881	-0.22896	-0.07007
Soil pedality	-0.62337	-0.41749	-0.21003	0.127841
Temperature warmest period (Bio05)	0.053738	0.650658	-0.64302	0.315834
Temperature coldest period (Bio06)	-0.75522	0.383608	-0.09691	0.397787
B) Northern subpopulation				
Variance explained	41.69%	30.88%	17.31%	6.88%
Aridity index	-0.31402	0.892449	0.192953	0.173317
Distance to coast	-0.96381	-0.11648	0.052044	0.076061
Elevation	-0.82622	0.529001	-0.15539	0.038265
Nitrogen pre-European	0.851314	-0.34622	-0.35694	0.138626
Phosphorus pre-European	0.727169	0.186128	-0.63936	-0.15223
Precipitation wettest period (Bio13)	0.410503	0.808183	0.108733	0.348354
Precipitation driest period (Bio14)	0.245466	0.484514	0.437436	-0.71185
Soil depth	0.618202	0.620613	-0.37468	0.040586
Soil pedality	0.659165	0.623779	0.342341	0.037658
Temperature warmest period (Bio05)	0.61323	-0.68781	0.375622	0.069938
Temperature coldest period (Bio06)	0.447471	-0.13252	0.837531	0.203278
C) Southern subpopulation				
Variance explained	46.50%	21.58%	8.40%	6.92%
Aridity index	0.855132	-0.38123	-0.08103	0.037575
Distance to coast	-0.57329	0.758541	0.089771	0.056064
Elevation	-0.04536	0.717134	0.466161	0.413577
Nitrogen pre-European	0.852057	0.311406	-0.00772	0.098637
Phosphorus pre-European	0.704893	0.434399	-0.34798	0.029393
Precipitation wettest period (Bio13)	0.72233	-0.46618	0.143999	0.041349
Precipitation driest period (Bio14)	0.732229	-0.20926	0.063184	0.4426
Soil depth	0.79463	0.204985	0.153285	-0.1377
Soil pedality	0.580163	0.14063	0.574057	-0.53091
Temperature warmest period (Bio05)	-0.83035	-0.29053	0.078366	-0.08087
Temperature coldest period (Bio06)	-0.33845	-0.66807	0.43222	0.264928

† Indicative of how the variable contributes to each principal component. Positive loadings indicates positive correlations and negative loading indicate negative correlations (Holland 2019).

Table S4 Environmental Association Analysis (EAA) and Generalised Dissimilarity Modelling (GDM)

Data set	Number of SNPs detected in EAA (< 1% missing data)	Number of sites incl. in GDM	Null deviance	% deviance explained by GDM	% deviance explained by GDM excl. geographic distance	Geographic distance Imp.	Sig.	Aridity index Imp.	Sig.	Distance to coast Imp.	Sig.	Elevation Imp.	Sig.	Nitrogen pre-European Imp.	Sig.	Phosphorus pre-European Imp.	Sig.	Precipitation wettest period (Bio13) Imp.	Sig.	Precipitation driest period (Bio14) Imp.	Sig.	Soil depth Imp.	Sig.	Soil pedality Imp.	Sig.	Temperature warmest period (Bio05) Imp.	Sig.	Temperature coldest period (Bio06) Imp.	Sig.
Model A - species range (m = 25 099)	NA	26	9.4	71.4	68.1	4.54	**0.00**	NA	NA	0.00	1.00	0.21	0.48	0.65	0.32	0.86	0.26	0.88	0.36	4.66	0.14	0.26	0.30	NA	NA	5.12	0.10	NA	NA
Model B - species range excl. Eungella, Finch Hatton Gorge and Brooweena (m = 25 099)	NA	23	4.0	87.6	84.5	3.55	**0.00**	NA	NA	0.02	0.72	0.36	0.46	NA	NA	0.50	0.26	0.13	0.56	2.79	0.06	0.94	0.24	0.02	0.60	1.55	0.16	0.14	0.44
Northern subpopulation																													
North all SNPs (m = 25 099)	NA	6	0.2	100.0	93.8	6.14	**0.02**	NA	NA	0.00	0.18	0.00	0.22	NA	NA	0.01	0.10	NA	NA	2.00	**0.04**	0.00	0.14	0.01	0.08	NA	NA	0.01	**0.04**
North PC1-associated	186	6	1.2	98.3	98.3	NA	NA	NA	NA	1.16	**0.00**	0.06	**0.00**	NA	NA	NA	NA	NA	NA	NA	NA	0.06	**0.04**	0.15	0.08	NA	NA	NA	NA
North PC2-associated	158	6	1.3	98.7	98.7	NA	NA	NA	NA	1.93	**0.00**	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.65	**0.04**	0.01	0.16	0.01	0.08
North PC3-associated	207	6	1.8	92.0	87.6	4.80	**0.04**	NA	NA	NA	NA	NA	NA	NA	NA	0.02	**0.04**	NA	NA	0.15	0.20	0.20	0.12	2.55	0.10	NA	NA	‡	‡
North Bio14-associated	135	6	1.4	97.5	93.4	4.20	**0.02**	NA	NA	0.02	0.08	NA	NA	NA	NA	NA	NA	NA	NA	33.17	**0.02**	NA	NA	4.92	0.08	NA	NA	NA	NA
North Soil depth-associated	98	6	1.2	95.7	95.6	0.13	**0.00**	NA	NA	NA	NA	NA	NA	NA	NA	2.44	**0.00**	NA	NA	0.08	0.20	17.20	**0.00**	NA	NA	NA	NA	0.60	0.06
North Soil pedality-associated	149	6	1.5	97.2	96.3	0.93	**0.00**	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.02	0.20	0.42	0.12	12.02	**0.00**	NA	NA	0.09	0.10
North Bio05-associated	217	6	1.8	99.2	99.1	0.15	**0.00**	NA	NA	0.10	**0.00**	0.07	**0.02**	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.08	**0.00**	0.04	**0.02**	0.02	0.22
Southern subpopulation																													
South excl. Brooweena all SNPs (m = 25 099)	NA	17	§	§	§	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
South PC1-associated	34	17	13.8	64.6	62.0	4.04	**0.00**	0.78	0.36	5.57	0.10	NA	NA	0.70	0.40	34.86	0.10	NA	NA	1.66	0.46	NA	NA	NA	NA	0.07	0.58	NA	NA
South PC2-associated	35	17	12.6	61.3	61.3	NA	NA	NA	NA	3.04	0.32	1.34	0.34	0.50	0.42	19.83	**0.04**	0.00	0.82	0.05	0.62	NA	NA	0.01	0.68	0.52	0.33	0.26	0.60
South Soil pedality-associated	27	17	10.9	49.2	47.1	4.31	0.06	0.24	0.70	1.05	0.56	4.00	0.30	NA	NA	9.40	0.16	NA	NA	11.47	0.18	NA	NA	6.64	0.18	NA	NA	0.26	0.52
South Bio06-associated	54	17	14.2	68.8	66.3	3.64	**0.00**	0.18	0.62	0.91	0.44	0.31	0.48	NA	NA	15.02	**0.00**	NA	NA	1.07	0.46	NA	NA	0.00	1.00	0.60	0.58	5.55	0.04

Red text indicate significance values where p < 0.05.

† Percent change in deviance explained by the full model and the model with the variable permuted.

‡ GDM failed to fit when this variable is permuted, indicative that this variable is important in the model.

§ No model could be fitted.

Table S5 Gene Ontology (GO) biological process enrichment data sets

Data set	Number of SNP loci	Number of unique genes overlapping 4 kb window of loci	GO biological process enrichment detected
Total SNP-captured space (reference/background)	25 099	13014	n/a
North all environment-associated	n/a	3050	yes
- North PC1-associated	1601	820	yes
- North PC2-associated	1517	797	yes
- North PC3-associated	1390	728	yes
- North Bio14-associated	1591	790	yes
- North Soil depth-associated	1684	779	yes
- North Soil pedality-associated	2582	1610	no
- North Bio05-associated	1536	823	yes
South all environment-associated	n/a	1277	yes
- South PC1-associated	1000	441	yes
- South PC2-associated	1168	551	yes
- South Soil pedality-associated	1155	524	no
- South Bio06-associated	1138	475	yes
North non-associated	6340	6036	no
Random	513	374	no

Table S6 Gene Ontology (GO) biological process enrichment results for 4 kb window around each environment-associated SNP

| Cluster	GO term ID	GO term description	Genes in category in total SNP captured set	Genes in category in cluster	Adjusted p-value
North PC1-associated	GO:0006952	defense response	355	52	0.000024
North PC2-associated	GO:0006952	defense response	355	51	0.000025
North PC3-associated	GO:0006952	defense response	355	45	0.000556
North Bio14-associated	GO:0006952	defense response	355	55	0.000000
North Soil depth-associated	GO:0006952	defense response	355	60	0.000000
North Soil pedality-associated	NA	NA	NA	NA	NA
North Bio05-associated	GO:0006952	defense response	355	50	0.000186
North all environment-associated	GO:0006952	defense response	355	143	0.000000
South PC1-associated	GO:0006952	defense response	355	35	0.000036
South PC2-associated	GO:0006952	defense response	355	37	0.001005
South Soil pedality-associated	NA	NA	NA	NA	NA
South Bio06-associated	GO:0006468	protein phosphorylation	822	56	0.010926
South Bio06-associated	GO:0006952	defense response	355	35	0.000230
South all environment-associated	GO:0006468	protein phosphorylation	822	126	0.000433
South all environment-associated	GO:0007165	signal transduction	365	62	0.029692
South all environment-associated	GO:0006952	defense response	355	72	0.000004
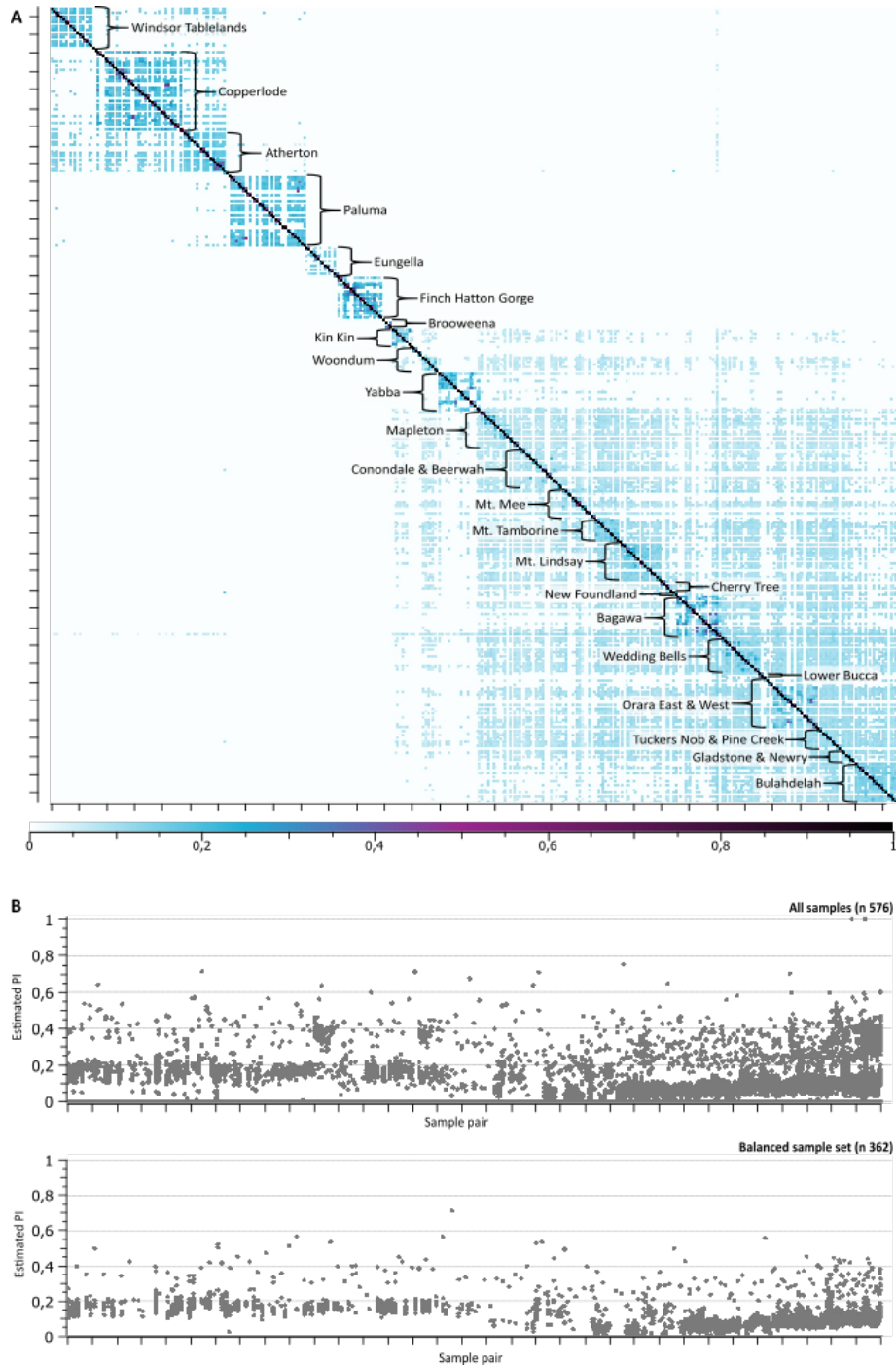
Figure S1 Pairwise identity by decent estimates of relatedness. A) Heatmap of relatedness among the 362 individuals representing a balanced sampling of the population. First-degree relatedness as expected for parent-offspring or full-siblings is ~0.5 (purple) and second-degree relatedness values, as for half-siblings, are expected to be ~0.25 (cyan). **B**) Estimated Pairwise Identity (PI) among all sample pairs of the full sample set (top) and balanced sample set (bottom).
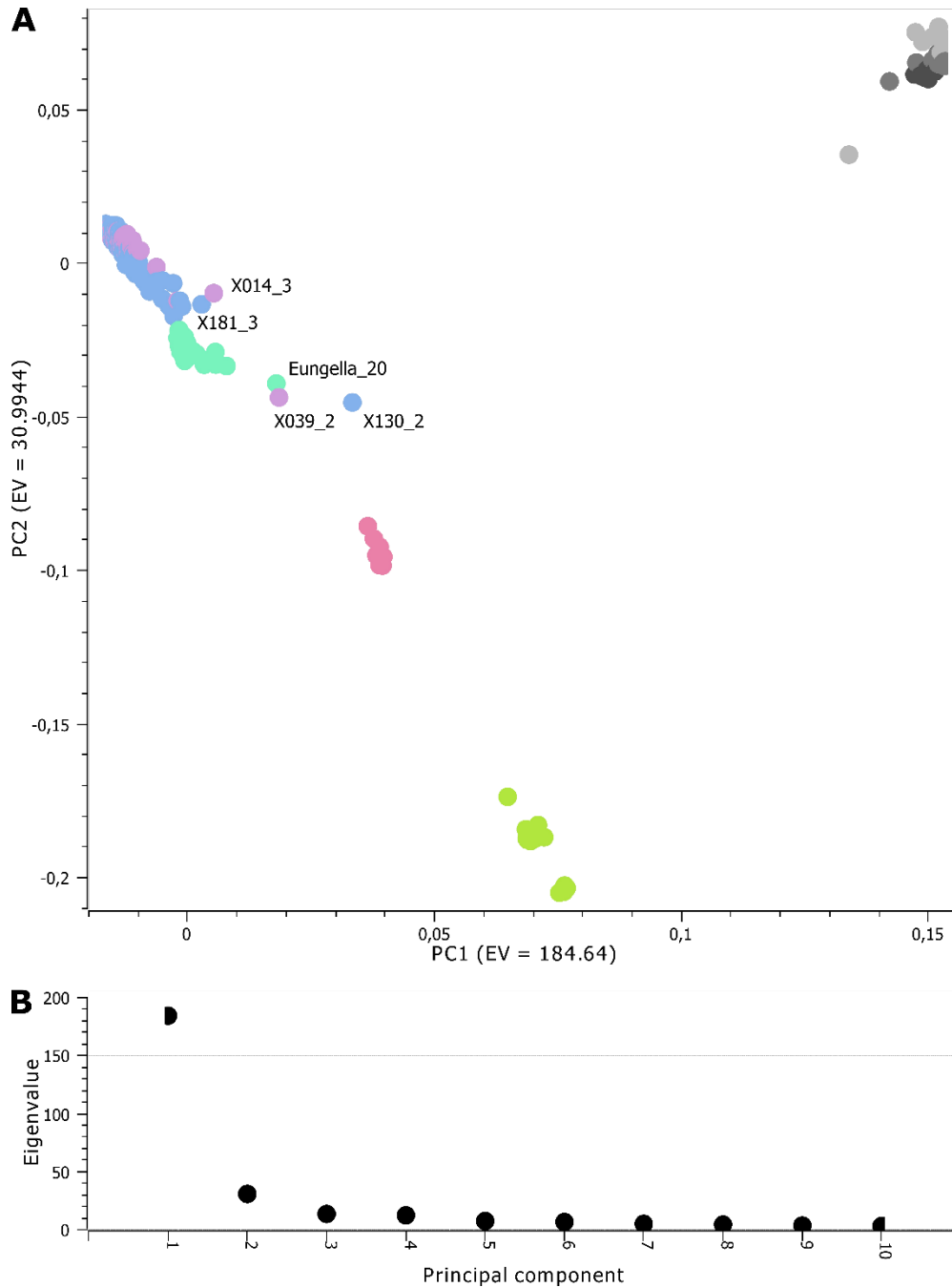
Figure S2 Species delineation by principal component analysis for *E. grandis* and other species used as reference populations in ancestry mapping. A) PCA plot of the first two principal components of all *E. grandis* individuals (coloured according to subpopulation assignment; North as purple, Mackay as turquoise and South as blue) and reference samples from other species as obtained from Silva-Junior *et al.,* (2015). Samples from the *Maidenaria* section are in shades of grey (*E. nitens* in light grey, *E. viminalis* in medium grey and *E. dunnii* in dark grey). *E. urophylla* and *E. saligna* from the *Latoangulatae* section are coloured lime green and pink, respectively. Five suspected hybrids are individually labelled by sample ID. B) Screeplot of the first ten principal components' eigenvalues.
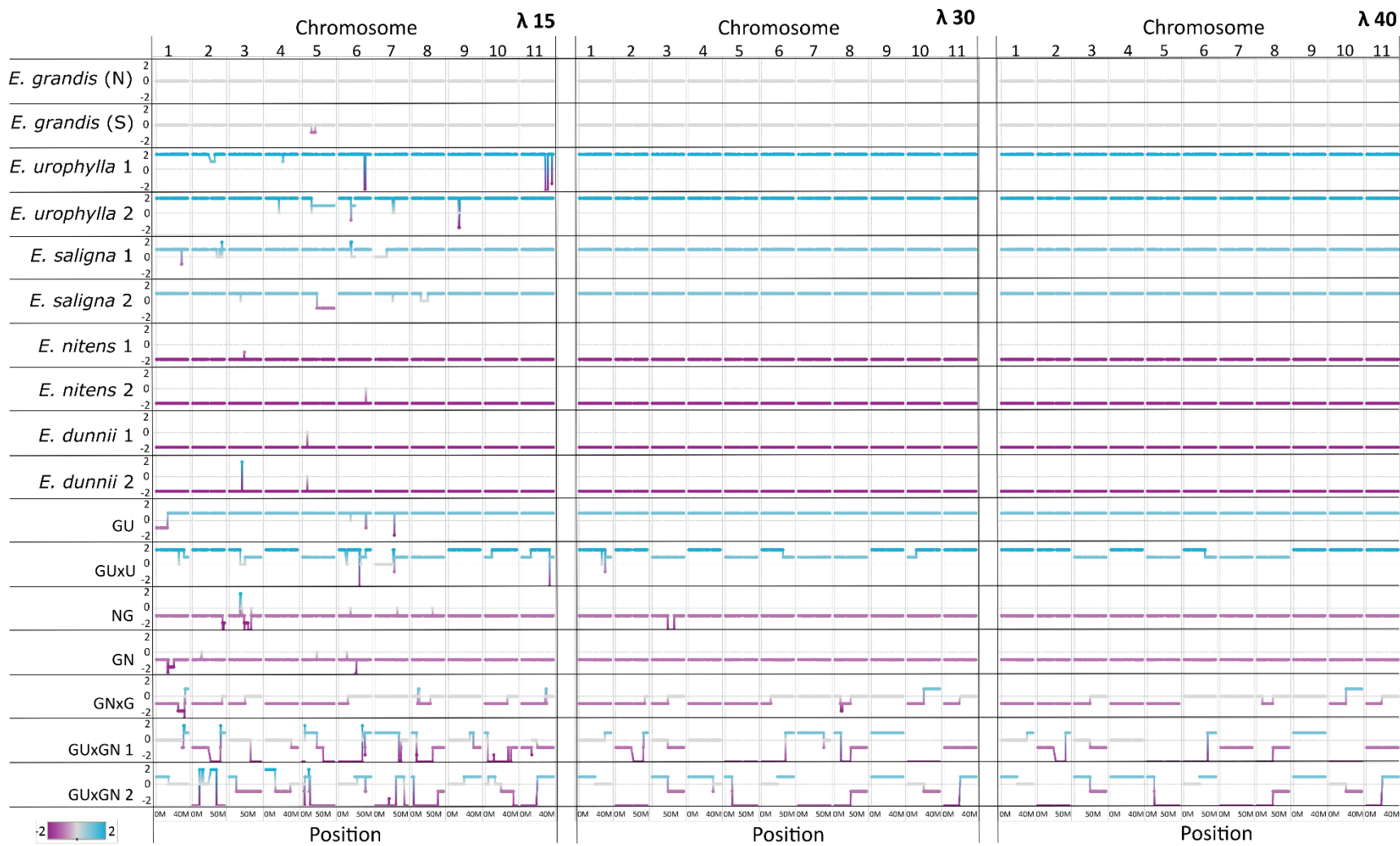
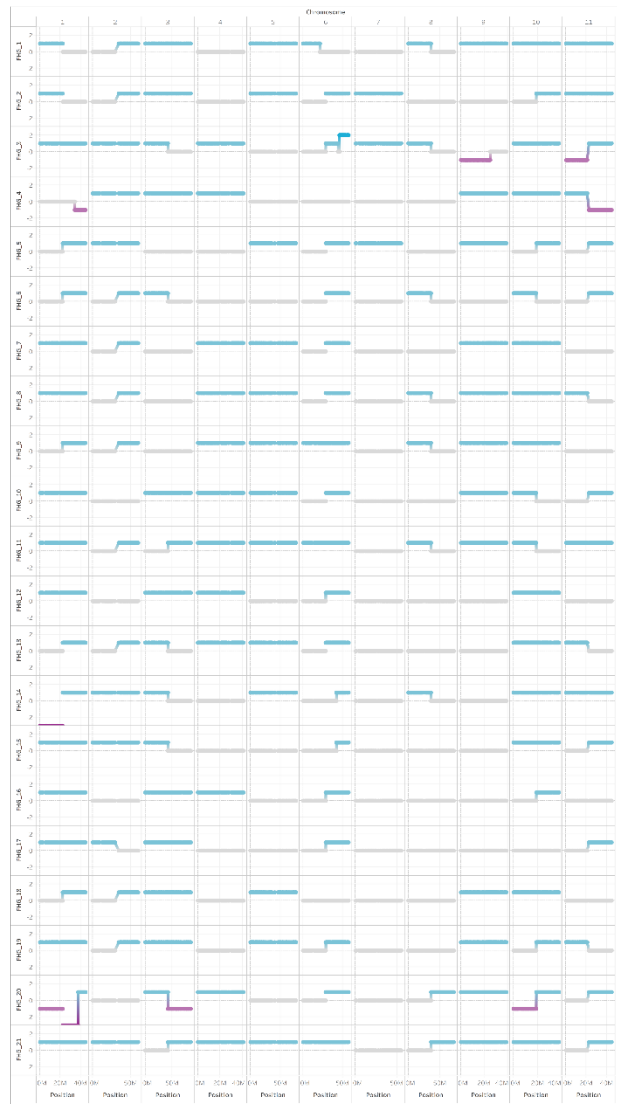**Figure S3 Optimization of ancestry assignment of chromosomal segments to *E. grandis* (anc1, grey), non-*E. grandis Latoangulatae* (anc3, cyan) or *Maidenaria*-like (anc2, purple) using *EILA* (Yang et al. 2013) for control samples.** Control sample ancestry mapping for breakpoint (where ancestry switched from one ancestral population to another) penalty values of $\lambda = 15$, $\lambda = 30$, $\lambda = 40$. Each colour dot represents a SNP (7736 SNPs shown along the 11 chromosome with Mbp positions on the x-axis). Both alleles assigned to *E. grandis* (anc1) = 0; one allele assigned to *E. grandis* the other to anc2 = -1, both alleles assigned to anc2 = -2, one allele assigned to *E. grandis* and the other to anc3 = 1, both alleles assigned to anc3 = 2, one allele assigned to anc2 and the other to anc3 = -2.5.
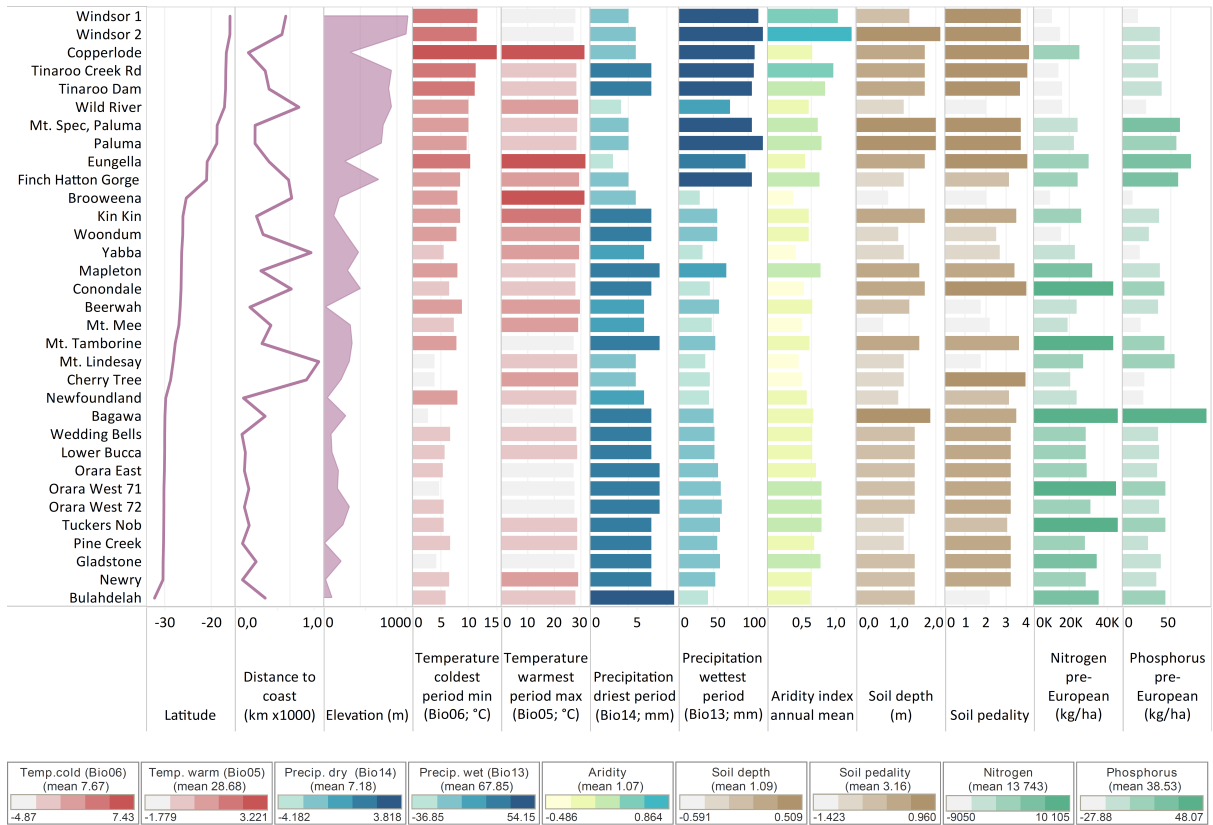
Figure S4 Environmental data for realised niche represented in sample collection. The trends of environmental data layers as obtained from Atlas of Living Australia (occurrence data downloaded at http://www.ala.org.au. Accessed 28 November 2019). Colour shading correlates with the deviation from the mean. For each variable, the mean of the observed values are indicated with the colour scale legend.
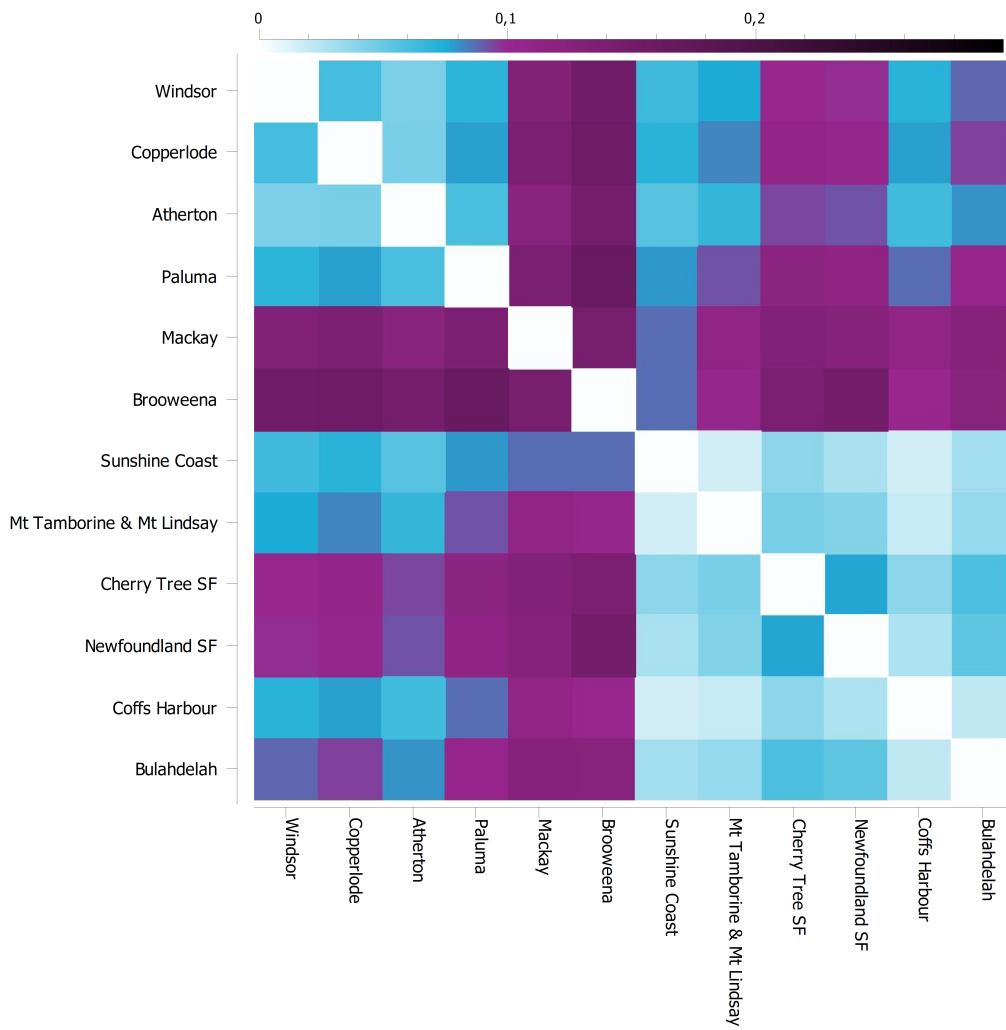
Figure S5 Heatmap of F_{ST} estimates by region (see Table S1) across the species range as ordered by decreasing latitude
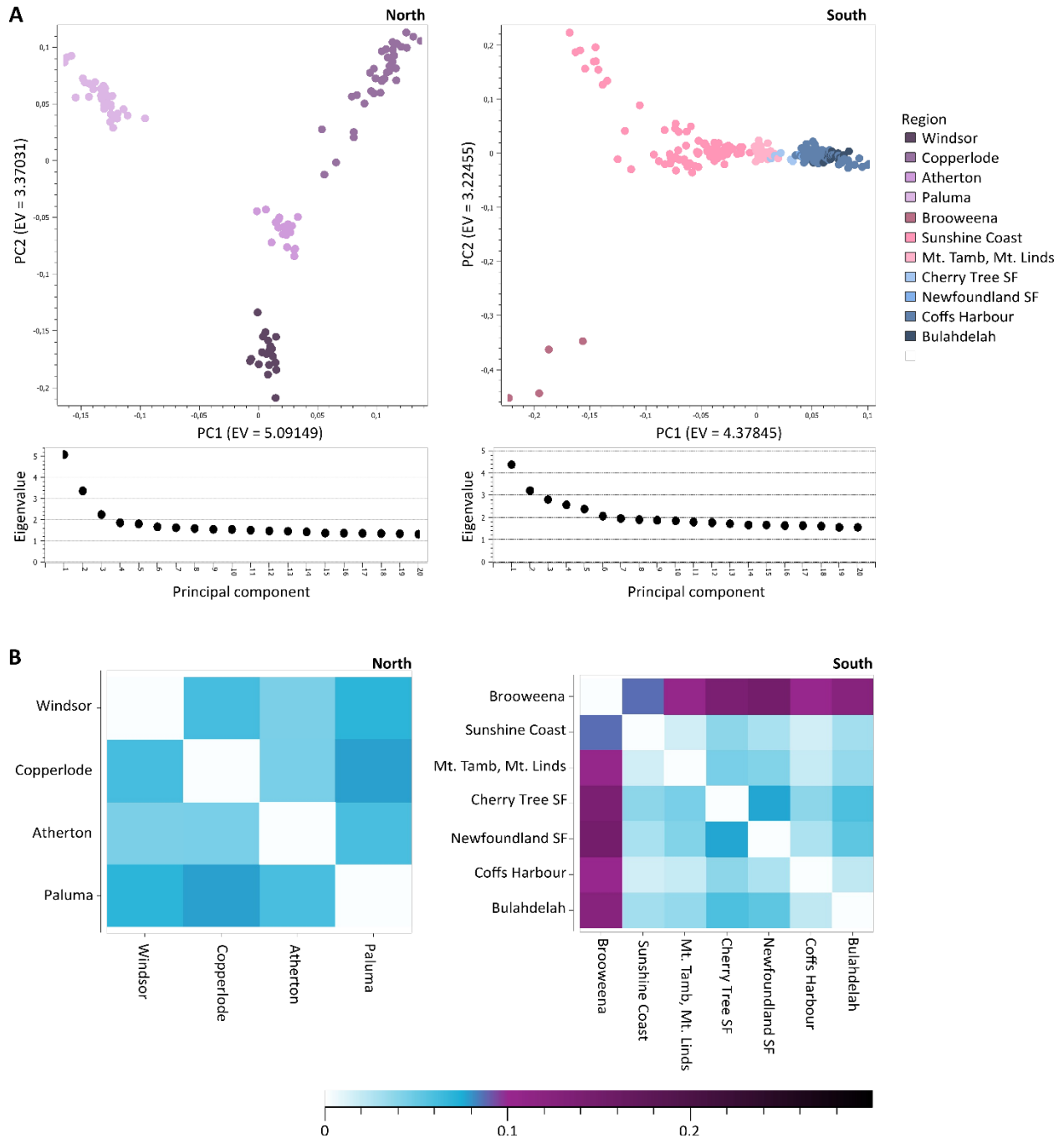
Figure S6 Population structure of the Northern and Southern subpopulations. A) Population structure is given based on the first two components of the PCA with individuals coloured by region (top), and screeplot of PCA (bottom) for Northern (left) and Southern (right) subpopulations. **B**) Heatmaps of F_{ST} estimates by region for Northern subpopulation (left) and Southern subpopulation (right).

A

Cross-entropy

Number of ancestral populations

B

Cross-entropy

Number of ancestral populations

C

Cross-entropy

Number of ancestral populations
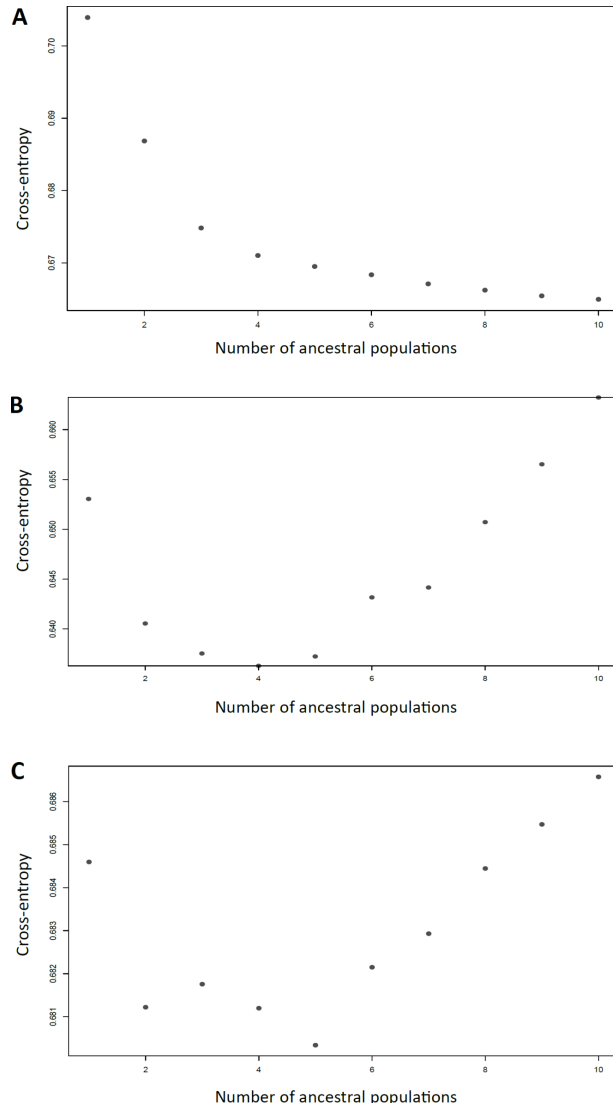
Figure S7 Cross-entropy of sNMF for 362 samples representing the species range (A), Northern subpopulation (B) and Southern subpopulation (C). The number of subpopulations, *K*, assessed were one to ten with 20 repetitions each using the *LEA* R package (Frichot & François 2015). Most plausible number of subpopulations populations revealed as either a "knee" in the plot as at *K* = 3 in **A**, or as the minimum cross-entropy value as *K* = 4 in **B** and *K* = 5 in **C**. These values of *K* were used for correction of population structure in the LFMM analyses.
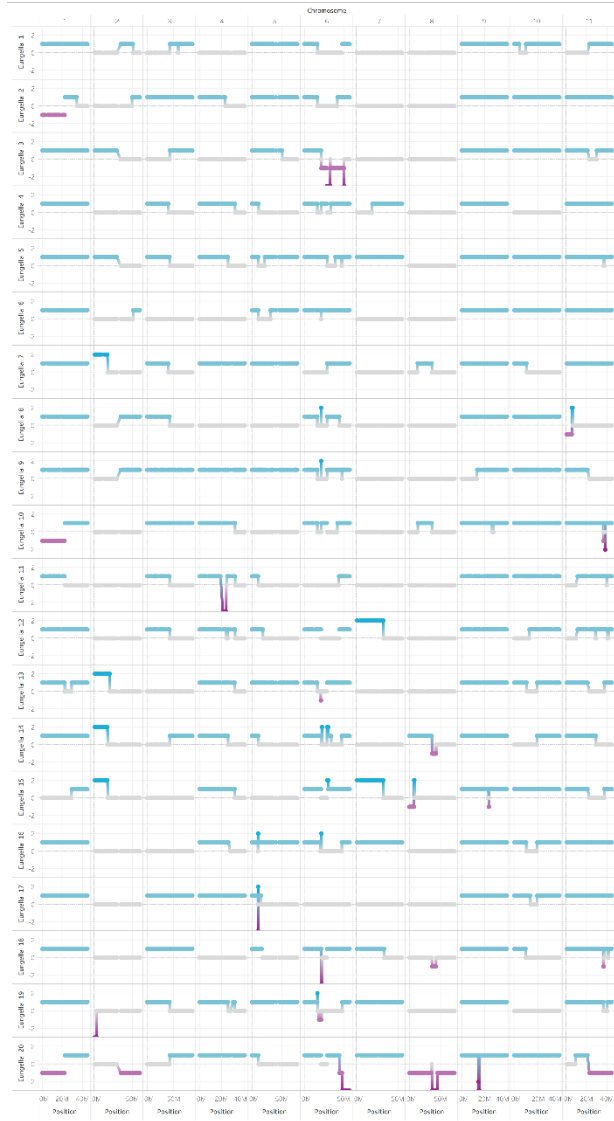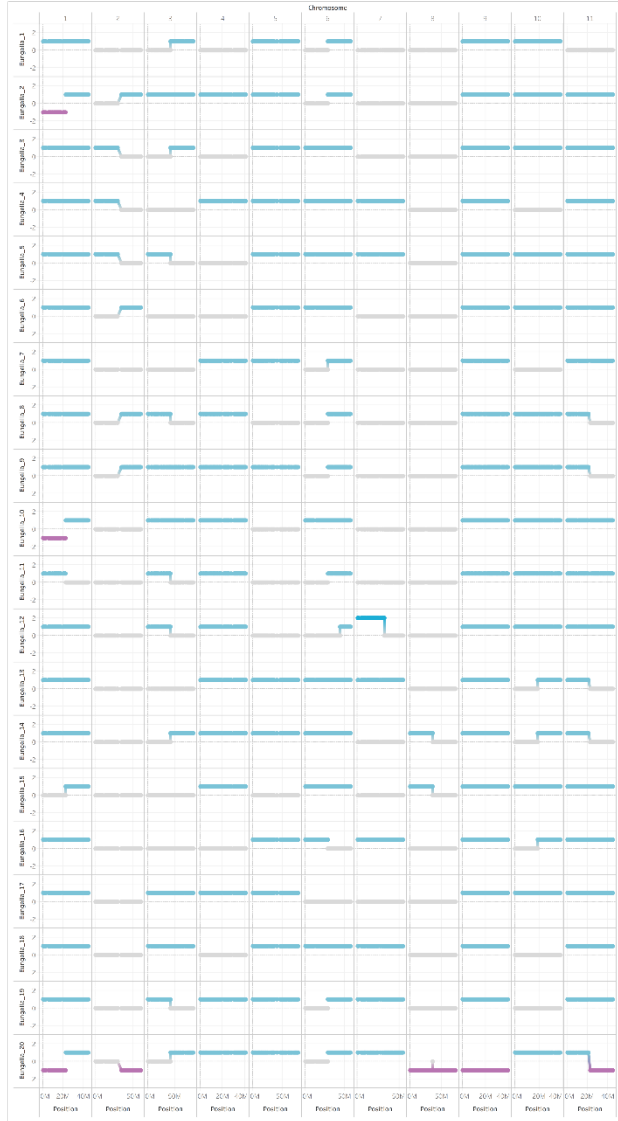
Figure S8 Ancestry assignment of chromosomal segments to *E. grandis* (anc1, grey), non-*E. grandis* *Latoangulatae* (anc3, cyan) or *Maidenaria*-like (anc2, purple) using *EILA* for the Eungella provenance. Each colour dot represents a SNP (7736 SNPs shown along the 11 chromosome with Mbp positions on the x-axis). Both alleles assigned to *E. grandis* (anc1) = 0; one allele assigned to *E. grandis* the other to anc2 = -1, both alleles assigned to anc2 = -2, one allele assigned to *E. grandis* and the other to anc3 = 1, both alleles assigned to anc3 = 2, one allele assigned to anc2 and the other to anc3 = -2.5. Breakpoint penalty $\lambda = 30$.
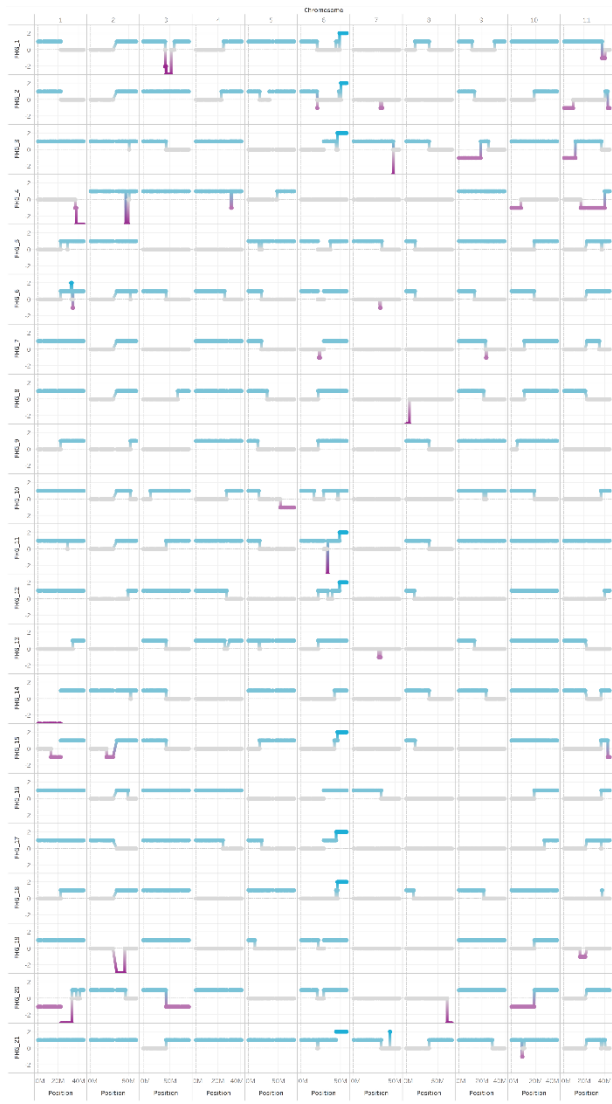
Figure S9 Ancestry assignment of chromosomal segments to *E. grandis* (anc1, grey), non-*E. grandis Latoangulatae* (anc3, cyan) or *Maidenaria*-like (anc2, purple) using *EILA* for the Finch Hatton Gorge provenance. Each colour dot represents a SNP (7736 SNPs shown along the 11 chromosome with Mbp positions on the x-axis). Both alleles assigned to *E. grandis* (anc1) = 0; one allele assigned to *E. grandis* the other to anc2 = -1, both alleles assigned to anc2 = -2, one allele assigned to *E. grandis* and the other to anc3 = 1, both alleles assigned to anc3 = 2, one allele assigned to anc2 and the other to anc3 = -2.5. Breakpoint penalty λ = 30.

18
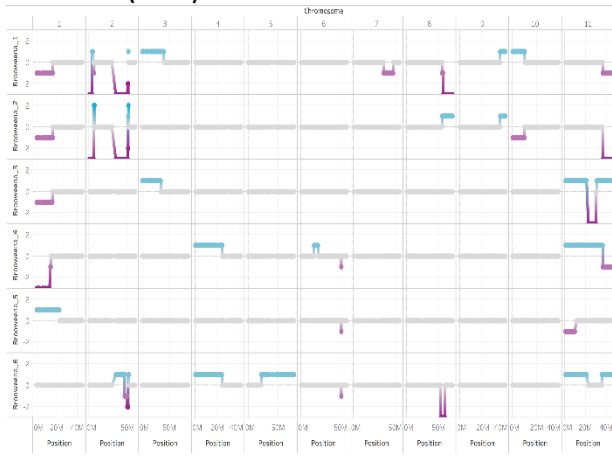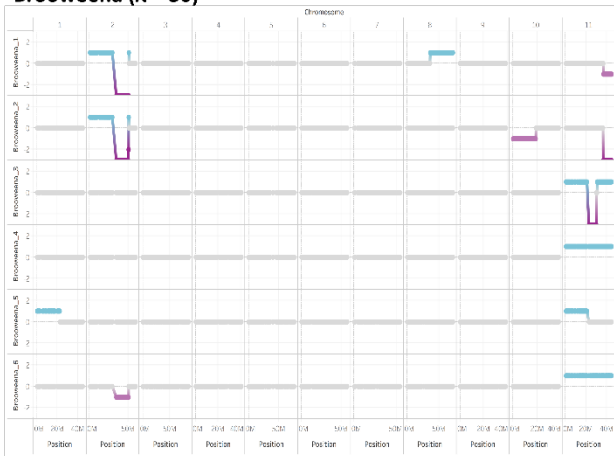
Figure S10 Ancestry assignment of chromosomal segments to *E. grandis* (anc1, grey), non-*E. grandis Latoangulatae* (anc3, cyan) or *Maidenaria*-like (anc2, purple) using *EILA* for the Brooweena provenance. Each colour dot represents a SNP (7736 SNPs shown along the 11 chromosome with Mbp positions on the x-axis). Both alleles assigned to *E. grandis* (anc1) = 0; one allele assigned to *E. grandis* the other to anc2 = -1, both alleles assigned to anc2 = -2, one allele assigned to *E. grandis* and the other to anc3 = 1, both alleles assigned to anc3 = 2, one allele assigned to anc2 and the other to anc3 = -2.5. Breakpoint penalty λ = 30.
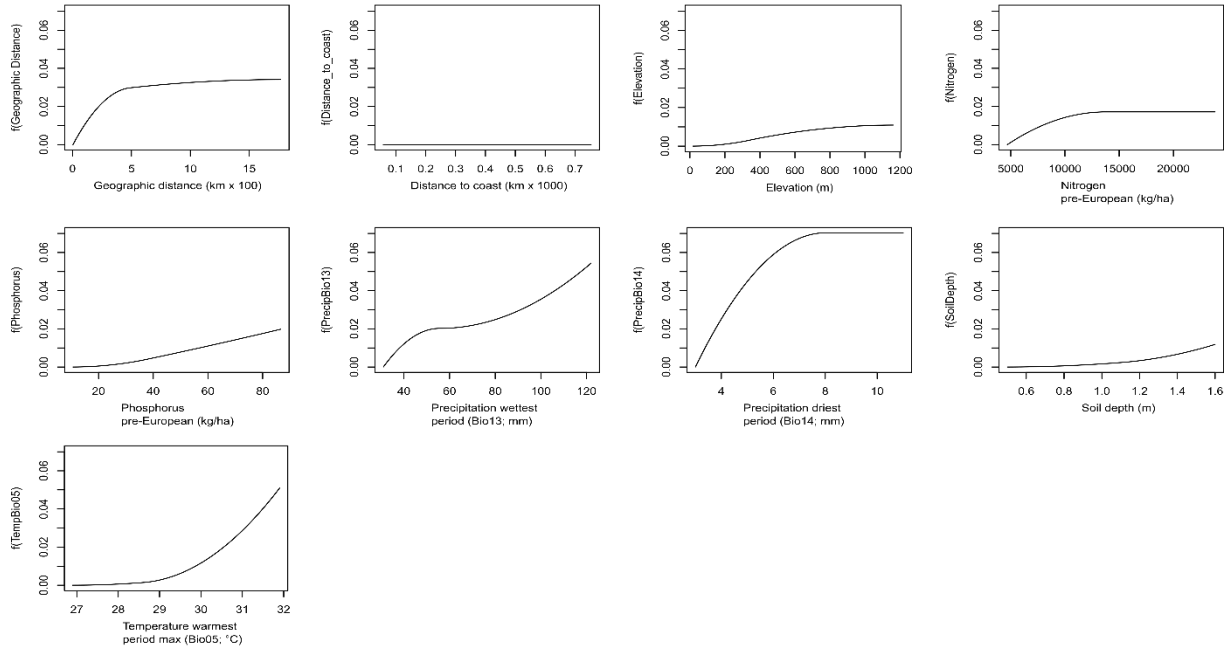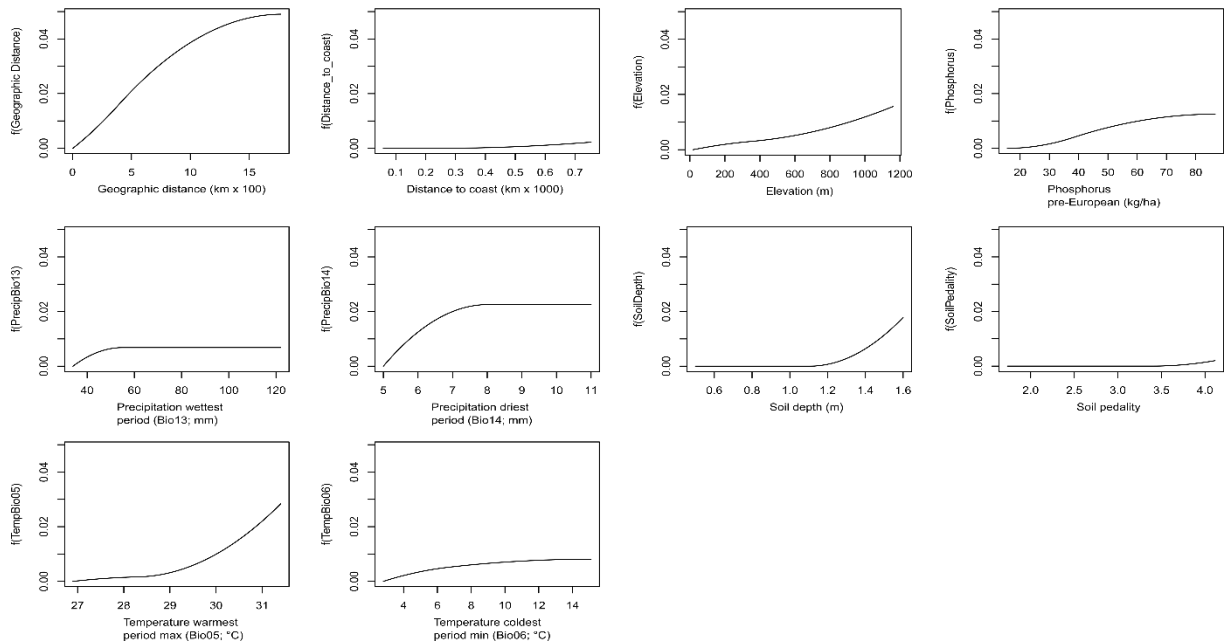
Model A

Model B

Figure S11 Generalised Dissimilarity Model (GDM) *I*-spline plots for predictor variables with coefficients > 0 for model A – entire species range, and model B – species range excluding introgressed provenances. The change in coefficient (*f*) given on the y-axis across the environmental variable range given on the x-axis. The importance and significance of predictor variables with coefficients > 0 are given in **Table S4**.
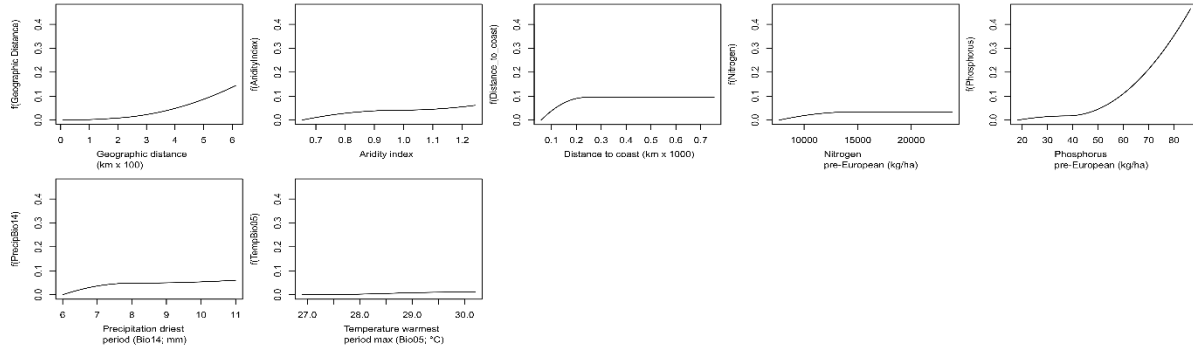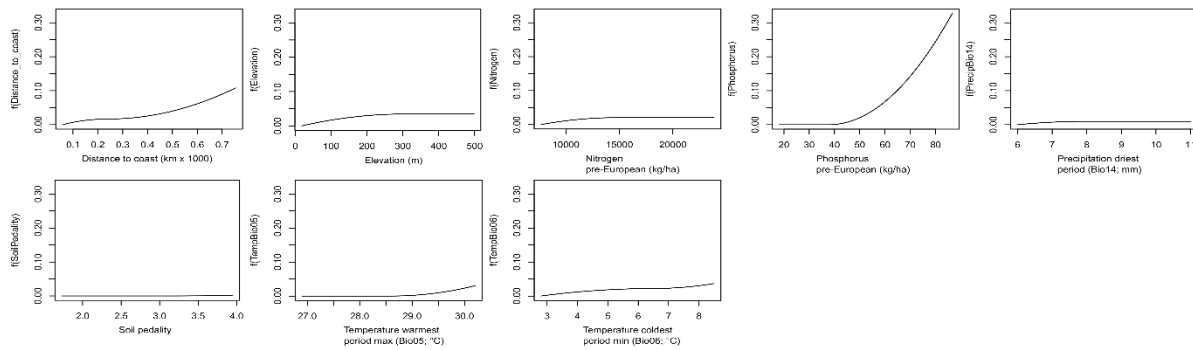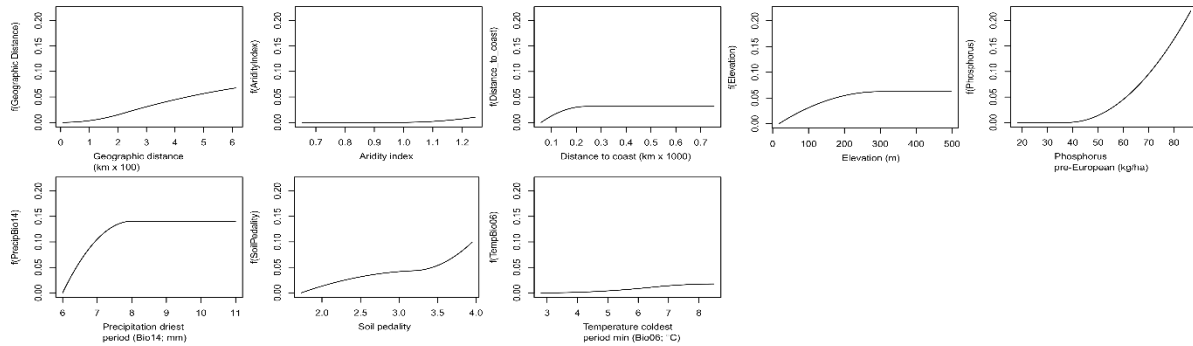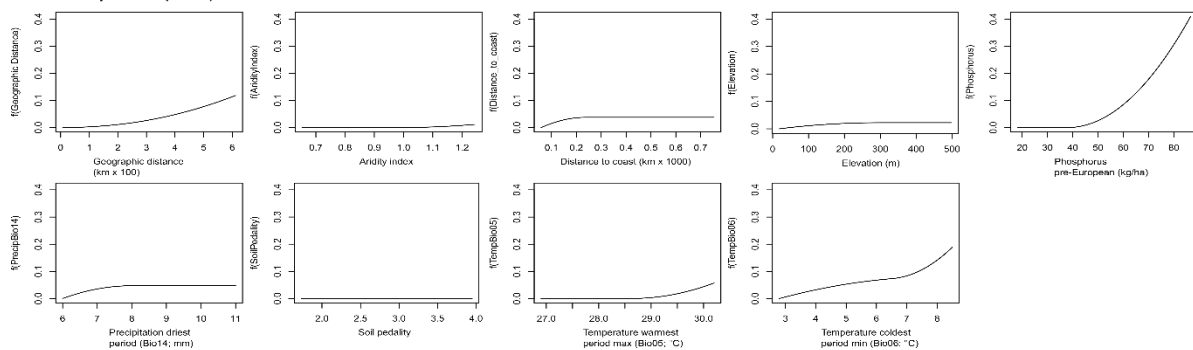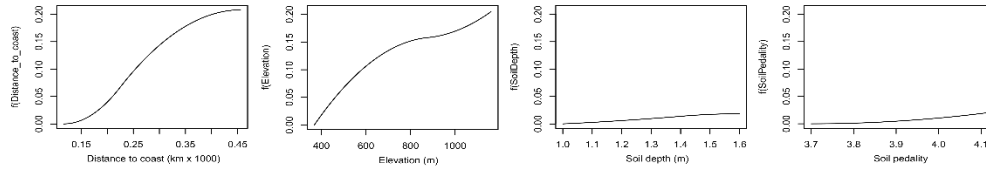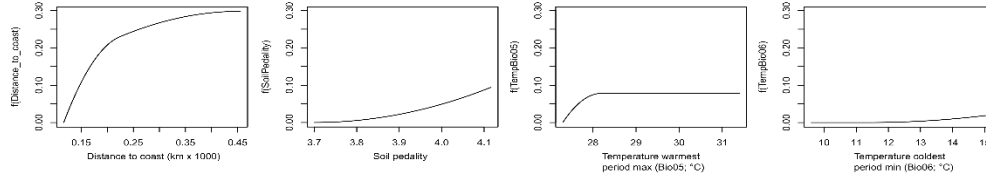
Figure S32 Generalised Dissimilarity Model (GDM) *I*-spline plots for predictor variables with coefficients > 0 for the Southern subpopulation environment associated SNP sets. The change in coefficient (*f*) given on the y-axis across the environmental variable range on the x-axis. The importance and significance of predictor variables with coefficients > 0 are given in **Table S4**.
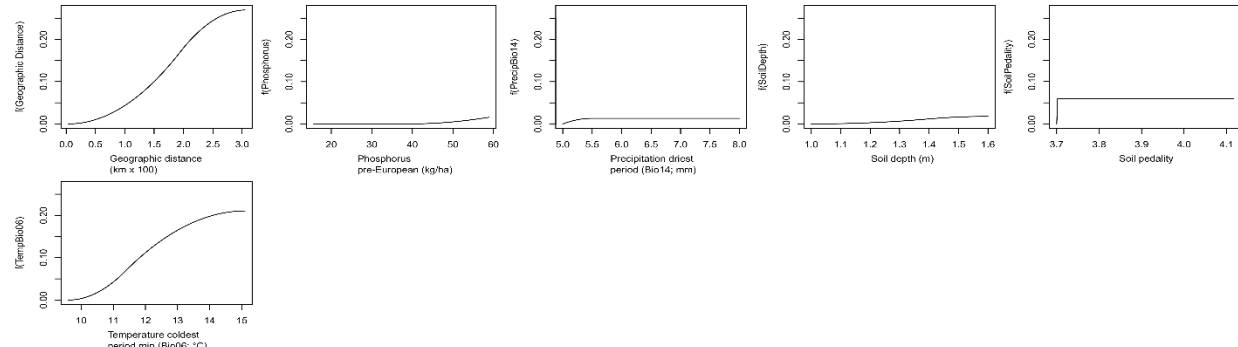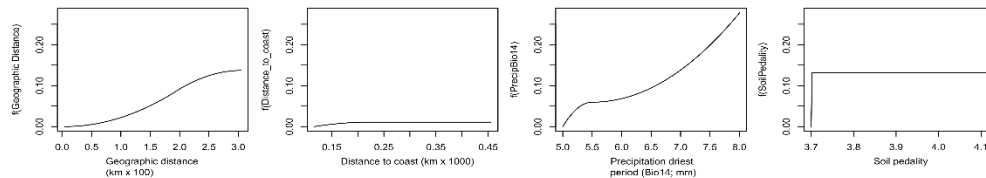
North-PC1

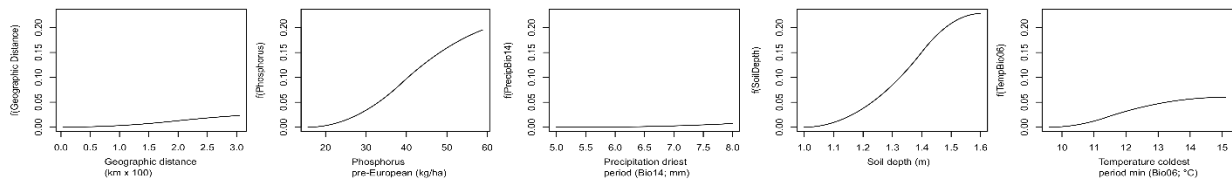**North-PC2**

North-PC3

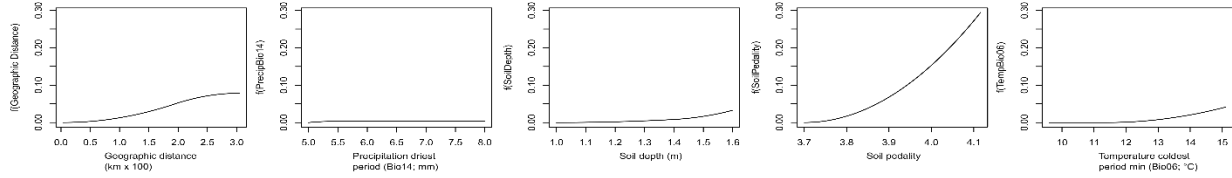**North-Precipitation (Bio14)**

Figure S43 (continued on next page) Generalised Dissimilarity Model (GDM) *I*-spline plots for predictor variables with coefficients > 0 for the Northern subpopulation environment associated SNP sets. The change in coefficient (*f*) given on the y-axis across the environmental variable range on the x-axis. The importance and significance of predictor variables with coefficients > 0 are given in **Table S4**.

22

Figure S53 (continued from previous page) Generalised Dissimilarity Model (GDM) *I*-spline plots for predictor variables with coefficients > 0 for the Northern subpopulation environment associated SNP sets. The change in coefficient (*f*) given on the y-axis across the environmental variable range on the x-axis. The importance and significance of predictor variables with coefficients > 0 are given in **Table S4**.
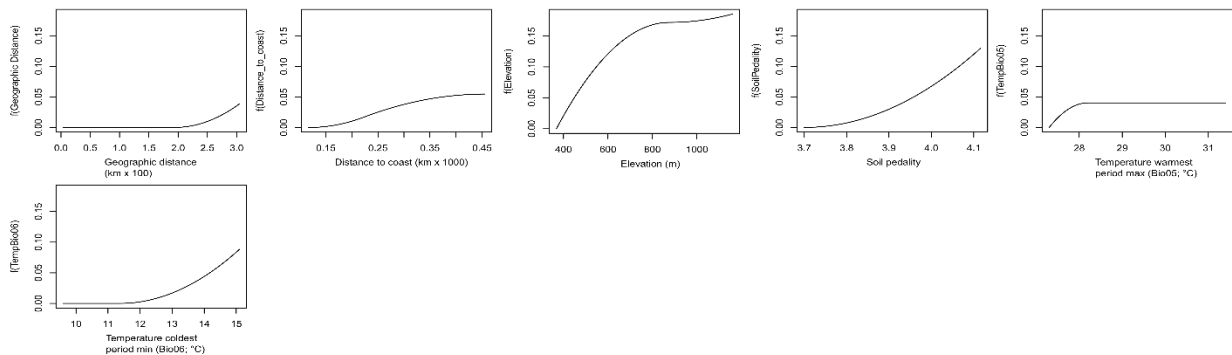
References

Brondani R, Brondani C, Grattapaglia D (2002) Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. *Molecular Genetics and Genomics* **267**, 338-347.

Brondani RP, Williams ER, Brondani C, Grattapaglia D (2006) A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. *BMC Plant Biology* **6**, 20.

Brondani RPV, Brondani C, Tarchini R, Grattapaglia D (1998) Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *TAG Theoretical and Applied Genetics* **97**, 816-827.

Byrne M, Marquezgarcia M, Uren T, Smith D, Moran G (1996) Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. *Australian Journal of Botany* **44**, 331-341.

Frichot E, François O (2015) LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution* **6**, 925-929.

Holland SM (2019) Principal components analysis (PCA). *Department of Geology, University of Georgia, Athens, GA*, 30602-32501.

Jones OR, Wang J (2010) COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**, 551-555.

Purcell S, Neale B, Todd-Brown K*, et al.* (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* **81**, 559-575.

Thamarus K, Groom K, Murrell J, Byrne M, Moran G (2002) A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre, and floral traits. *Theoretical and Applied Genetics* **104**, 379-387.

Yang JJ, Li J, Buu A, Williams LK (2013) Efficient inference of local ancestry. *Bioinformatics* **29**, 2750-2756.