Landscape genomic prediction for restoration of a *Eucalyptus* foundation species under climate change

Megan A. Supple[a,b,*], Jason G. Bragg[a,c], Linda M. Broadhurst[d], Adrienne B. Nicotra[a], Margaret Byrne[e], Rose L. Andrew[f], Abigail Widdup[a], Nicola C. Aitken[a], Justin O. Borevitz[a,g]

[a]Research School of Biology, The Australian National University, 134 Linnaeus Way, Canberra, ACT 2601, Australia; [b]Department of Ecology and Evolutionary Biology, University of California at Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA; [c]National Herbarium of New South Wales, The Royal Botanic Gardens and Domain Trust, Mrs Macquaries Road, Sydney, NSW 2000, Australia; [d]Centre for Australian National Biodiversity Research, Commonwealth Scientific and Industrial Research Organisation (CSIRO) National Research Collections and Facilities, PO Box1600, Canberra, ACT 2601, Australia; [e]Science and Conservation Division, Department of Parks and Wildlife Western Australia, Locked Bag 104, Bentley Delivery Centre, WA, 6983, Australia; [f]School of Environmental and Rural Science, University of New England, W55 Geology Rd, Armidale, NSW 2351, Australia; [g]Centre of Excellence in Plant Energy Biology, The Australian National University, 134 Linnaeus Way, Canberra, ACT 2601, Australia

[*]corresponding author
Megan Supple
1156 High Street
Mail Stop EEB
Santa Cruz, CA 95064, USA
megan.a.supple@gmail.com

key words: conservation genomics, phenotypic plasticity, landscape restoration

Abstract

As species face rapid environmental change, we can build resilient populations through restoration projects that incorporate predicted future climates into seed sourcing decisions. *Eucalyptus melliodora* is a foundation species of a critically endangered community in Australia that is a target for restoration. We examined patterns of genomic and phenotypic variation to make empirical based recommendations for seed sourcing. We examined isolation by distance and isolation by environment, determining gene flow up to 500 km and associations with environmental variables. Climate chamber studies revealed extensive phenotypic variation both within and among sampling sites, but no site-specific differentiation in phenotypic plasticity. Overall our results suggest that seed can be sourced broadly across the landscape, providing ample diversity for adaptation to environmental change. Application of our landscape genomic model to *E. melliodora* restoration projects can identify genomic variation suitable for predicted future climates, thereby increasing the long term probability of successful restoration.

Introduction

Species around the globe face rapidly changing environments, often in combination with habitat degradation and fragmentation. These factors are expected to have a negative impact on biodiversity (Lindenmayer et al., 2010). Three processes enable species survival in the face of altered conditions--migration, adaptation, and phenotypic plasticity (Aitken & Whitlock, 2013; Aitken et al., 2008; Hoffmann et al., 2015; Nicotra et al., 2010). An important conservation strategy is to assist these natural processes to help build more resilient communities. We can help species shift to regions with their preferred environmental conditions by assisting migration of gene pools across the landscape (Aitken & Whitlock, 2013; Aitken et al., 2008). We can aid populations to survive in situ by ensuring that sufficient genomic variation exists for adaptation to changing environments (Hoffmann et al., 2015). We can enable individuals to respond

to a greater range of environments by conserving existing phenotypic plasticity (Nicotra et al., 2010).

Seed sourcing during landscape restoration provides an ideal opportunity to apply scientific knowledge to enable these key processes and improve conservation outcomes (Broadhurst et al., 2008; Prober et al., 2015). For example, seed sources can be selected to restore historical patterns of gene flow across a fragmented landscape, incorporate high genomic diversity, and/or increase phenotypic plasticity. Seed sources can also be matched with current or projected future climates, enabling assisted migration to favorable environments (Aitken & Whitlock, 2013; Williams et al., 2014).

Historically, restoration often focused on geographically restricted local sources of seed under the premise that this would improve restoration outcomes by reducing the risk of maladaptation to local conditions and preventing outbreeding depression (Broadhurst et al., 2008). However, there are several potential drawbacks to this narrow local focus. In a fragmented system, narrow local seed sourcing reduces the number of potential source populations, thereby reducing the pool of available genetic material. This reduced gene pool may result in inbreeding depression in future generations, especially if combined with small population size (Broadhurst et al., 2008). Obtaining potential seed sources from a wider geographical area can increase genomic and phenotypic diversity, thereby increasing the ability of the species to survive in situ (Broadhurst et al., 2008). Additionally, the focus on maintaining local adaptation in situ assumes a static environment, not the rapidly changing environment that occurs today. As local conditions change, traits and genes that may have conferred an advantage in the past might not be suitable in the future environment. In recent year, climate adjusted provenancing has been proposed, which is a seed sourcing strategy that incorporates climate variability and focuses on sourcing seed that is predicted to be adapted to future climates (Byrne et al., 2013; Prober et al., 2015). This strategic assisted migration of variation across the landscape can aid in the establishment of populations that are more adaptable to future environments (Prober et al., 2015).

To determine the appropriate seed sourcing strategy and to identify optimal seed sources for a reforestation project, empirical knowledge of genomic variation for

the target species can provide valuable information. The technology now exists to assess genomic variation in any target species, enabling determination of patterns of Isolation By Distance (IBD) and Isolation By Environment (IBE). IBD is the association between genomic distance and geographic distance resulting from patterns of dispersal. IBE is the association between genomic distance and environmental distance, while controlling for geographic distance (Wang & Bradburd, 2014). Landscape genomic models can be generated by fitting geographic and environmental variables to the observed genomic diversity. These predictive models can optimize the genetic material selected for restoration and should improve long term outcomes (Hoffmann et al., 2015; Williams et al., 2014).

The extent of phenotypic plasticity in potential seed sources can be measured in growth assays of seedling traits across contrasting environmental conditions. The magnitude of the environmental response can be compared among maternal lines or populations and may identify populations that differ in their response to the environment. Such differing responses have been seen in some species of *Eucalyptus* (Andrew et al., 2010; Byrne et al., 2013; McLean et al., 2014).

Eucalyptus melliodora (A.Cunn. ex Schauer), commonly called yellow box, is an iconic Australian species that is the subject of extensive restoration efforts across its distribution. It is a foundation species of a critically endangered ecological community: the White Box–Yellow Box–Blakely's Red Gum Grassy Woodland and Derived Native Grassland (Department of Environment and Climate Change and Water, 2011; Department of the Environment and Heritage, 2006; Threatened Species Scientific Committee, 2006). This woodland community exists in a fragmented landscape, with less than 5% of its original distribution remaining, mostly in small remnant patches (Department of Environment and Climate Change and Water, 2011; Department of the Environment and Heritage, 2006; Threatened Species Scientific Committee, 2006). Efforts to restore this endangered woodland community are ongoing and restoration practitioners are seeking scientific recommendations to improve seed sourcing. Climate change is an important consideration in seed sourcing decisions because species distribution modelling predicts that most eucalypts will need to shift their distributions

considerably in response (González-Orozco et al., 2016). In particular, for *E. melliodora* ecological niche modelling predicts that by 2090 the species distribution will shift toward the southeast and suitable areas will decrease by 77% as a result of environmental changes (Broadhurst et al., *in review*).

Here we survey genomic variation in 275 individuals from 37 sites across the present range of *E. melliodora*. We fit the genotypic data to geographic distance and key environmental variables at the site of origin. We find that effects of genomic isolation by distance begin at approximately 500km. This empirical estimate of "local" is much farther than what is often practiced for local provenancing. We also find that features of the abiotic environment can further explain genomic differentiation after accounting for geographic distance. We also examine seedling growth characteristics under simulated climate conditions and find significant variation in growth traits both within and among sites, but no significant variation in phenotypic plasticity across sites. Our landscape genomic model, which can empirically define local provenances and identify variation suitable for predicted future climates, can help build resilient populations through scientifically based restoration.

Results

Genotyping by Sequencing

We selected leaf material from 39 sites, sampling 3-10 trees per site (Supplemental Table S1). For each sample we Illumina sequenced a Genotyping by Sequencing (GBS) library (Elshire et al., 2011) and used a reference alignment-based approach to call genotypes. We conducted a preliminary analysis, based on 123,227 SNPs and removed 69 samples due to greater than 60% missing data. Visual examination of a cluster dendrogram of genomic distance between samples showed that technical replicates cluster closely together (Supplemental Fig. S1). A preliminary principal coordinate analysis (PCA) identified 19 samples that were strong genomic outliers (Supplemental Fig. S2), likely misidentified samples or recent hybrids. This result is

consistent with minor morphological differences noted in these samples, as well as previous microsatellite work (Broadhurst et al., *in review*). After removal of poor quality and geographic and genomic outlier samples, we reran the genotyping with the remaining 280 samples, resulting in 9,781 SNPs after filtering. A second preliminary PCA identified an additional 5 outlier samples that we considered sufficiently differentiated from the main *E. melliodora* cluster to merit removal for downstream analyses (Supplemental Fig. S3). We removed these samples and reran the missing data filter. The final data set included 275 samples from 37 sites (Fig. 1A), genotyped at 9,378 physically distinct SNPs (>300 bp apart).

Genomic Analyses

A PCA of genomic distance among samples showed continuous variation with little suggestion of discrete population structure (Fig. 1B). The samples largely formed a single cluster, with the first PCA axis corresponding roughly to latitude. Outside of the main cluster, samples from the northernmost site separated out along the first PCA axis (y-axis) and a few samples from two other sites separated out along the second PCA axis (x-axis). Together, the first two PCA axes explained 3.0% of the genomic variation among individuals. The Mantel test estimated that the natural log of the geographic distance between samples explained 2.3% of the variation in individual genomic distance, indicating weak, but statistically significant, isolation by distance ($p=0.0001$). We summarized genomic diversity between sampling sites using pairwise F_{st}. For all comparisons F_{st} was low (mean $F_{st}=0.04$, sd=0.02) (Supplemental Table S2). The maximum $F_{st}=0.10$ occurs between sites 3 and 13, which are separated by over 1200 km. Similar to the PCA of genomic distance among samples, the PCA of F_{st} between sampling sites corresponded roughly to latitude (Fig. 1C). In contrast, the first two axes of the PCA of F_{st} between sampling sites explained a higher percentage of variation (37.1%). These results highlight the tremendous amount of genomic variation within sampling sites, as well as the ability of thousands of independent genomic markers to distinguish between more distant sampling sites.

The site by site pairwise F_{st} matrix was used to test for geographic and environmental associations using generalized dissimilarity modelling (GDM) (Ferrier et al., 2007; Fitzpatrick & Keller, 2015; Thomassen et al., 2011). Of the 28 environmental variables considered for the model, we removed 12 variables because the single variable model explained less than 5% of the deviance (bioclimatic variables 2, 5, 6, 9, 10, 14, 17, 19; elevation; water at depth; Prescott Index; and MrVBF). We removed an additional 9 variables due to high correlation and lower explanatory power than another remaining variable (bioclimatic variables 1, 4, 7, 12, 13, 15, 18; surface nitrogen; and surface phosphorus) (Supplemental Table S3). We ran permutation testing on a model with the remaining 7 variables and geography. This highlighted an additional 2 variables with low statistical significance and low explanatory power that we removed from the final model (surface water and bioclimatic variable 8). We also removed phosphorus at depth because, although it explained a substantial amount of genomic variation, the sampled sites were not well distributed across the range of values.

As a result, we included four environmental variables in the final model: isothermality (bioclim 3), mean temperature of the coldest quarter (bioclim 11), precipitation of the wettest quarter (bioclim 16), and total soil nitrogen at 100-200 cm (nitrogen at depth) (Supplemental Fig. S4). The GDM model with these four variables plus geographic distance explained 40% of the variation in sampling site genomic differentiation (F_{st}). The GDM model showed a positive non-linear relationship between environmental distance and genomic distance (Supplemental Fig. S4A). To test the predictive power of the GDM model, we used a cross validation approach by generating 1000 models with a random 30% of sampling sites removed. GDM proved satisfactory at predicting genomic differences between removed sites (cross validation correlation mean=0.73, standard deviation=0.12) (Supplemental Fig. S4B).

Of the four environmental variables, nitrogen at depth showed the strongest relationship with genomic distance, with changes in genomic distance predicted across the range of nitrogen values (Supplemental Fig. S4C). Mean temperature of the coldest quarter was the second strongest predictor, showing a similar pattern as nitrogen (Supplemental Fig. S4D). Precipitation of the wettest quarter was the third strongest

environmental predictor, predicting the largest change in genomic distance between 250 and 400 mm (Supplemental Fig. S4E). Isothermality (mean diurnal range divided by annual temperature range) was the final predictor, predicting the most change in genomic distance at higher values (Supplemental Fig. S4F).

Geographic distance showed a non-linear relationship with genomic distance. The geographic spline predicted no genomic differentiation until close to 500 km, at which point an increase in geographic distance predicted an increase in genomic distance (Fig. 2). Randomly subsampling sites showed that the predicted genomic distance for large geographic distances was quite variable, but for sites less than 500 km apart, all iterations consistently predicted little genomic differentiation between sites (Supplemental Fig. S4H).

To project the final GDM model onto the current environmental landscape, we first delineated the geographic extent of the analysis by defining an *E. melliodora* distribution polygon. We then projected the GDM model onto this region using the current values of the environmental variables across the landscape. This analysis partitioned the landscape into a number of regions with different predicted genomic compositions, including northern coastal, northern inland, and southern regions (Fig. 3A). While the biggest differences occurred in regions with few sampling sites, there is a distinction between the northern and southern sites, as well as between sites on opposite sides of the Great Dividing Range in the southern region (e.g. site G versus site A , Fig. 3A). These projections highlight where environmental filtering of genotypes may have occurred due to different selective pressures.

We compared the GDM model projected onto current conditions to the GDM model projected onto 2070 climate predictions. This analysis scored each position across the landscape based on how much genomic change was predicted to occur in response to changing environmental conditions (Fig. 3B). For the middle north region (around sites K2 and 14) and the southern areas towards the coast, the models predicted more intense natural selection in response to climate change. Thus, these areas could be prioritized for assisted migration.

We also used the GDM model to compare the genomic composition under future environmental conditions at a single location to the genomic composition under current climate conditions across the landscape. This comparison is useful for identifying optimal seed sources for restoration sites given climate change scenarios. We demonstrated this utility by selecting two hypothetical reforestation sites and identifying distinct regions that would provide favorable seed sources for each site (Fig. 4). The analysis for the southern reforestation site identified a large portion of the southern distribution, centered at the reforestation site. For this site it appears that the selected areas are largely driven by the pattern of isolation by distance, in particular the lack of genetic differentiation for long geographical distances. The analysis for the northern reforestation site identified a more limited range of areas across the landscape, possibly driven in part by a decreased power due to lower sampling intensity in the north. In addition to identifying a narrow region in the north that is centered on the reforestation site, a number of more distant areas along the coast were also identified, indicating these selected areas are driven more by patterns of isolation by environment than isolation by distance.

These genomic analyses suggest that for woodland restoration a geographically wider and environmental model based approach to seed sourcing would allow incorporation of more genetic diversity and enable better matching of the selected genotypes to current and predicted future environmental conditions at the reforestation site.

Growth Experiments

We conducted a climate controlled growth experiment to measure variation in seedling growth traits among sampling sites and assay phenotypic plasticity. We grew seedlings from six sites, with six maternal lines per site, at two different climate regimes (average summer conditions and 5^{o}C hotter summer conditions). For analysis of seedling height and total leaf length, we analyzed a total of 291 seedlings (from 32 maternal lines representing six sampling sites) that were determined to be well established at the five week measurement. For analysis of the relative height increment, we analyzed a total

of 560 seedlings (from all 36 maternal lines) for which were able to calculate this metric. There were four seedlings that were outliers for the relative growth increment (>0.035). These outliers had little effect on the results of the linear models, so we included them in the final analysis.

The models for all three response variables (seedling height, total leaf length, and relative height increment) showed that all fixed effects (sampling site, maternal line nested within sampling site, and experimental condition) were statistically significant at the p=0.05 level (Supplemental Table S4). Experimental condition explained a small percentage of the variation (1.2-8.1%), as did sampling site (1.8-17.7%) (Supplemental Fig. S5). Maternal line tended to explain a larger amount of variation (10.6-27.6%). However, most of the variation remained unexplained (56.6-71.5%). None of the three response variables showed significant variation in phenotypic plasticity across sites (all maternal line/sampling site by experimental condition interactions p>0.50) (Supplemental Fig. S6 and Table S5).

We then conducted an outdoor drought experiment using a subset of seedlings from the temperature experiment. We analyzed 146 seedlings representing 20 maternal lines from five sampling sites. These seedlings were grouped into 73 pairs, with one of each pair assigned to each treatment—well watered versus drought. We analyzed variation in four response variables: stomatal conductance, leaf length to width ratio, relative chlorophyll content (SPAD index), and specific leaf area (SLA, leaf area divided by dry mass).

The droughted seedlings had significantly lower stomatal conductance rates than the well watered ones, indicating that the seedlings were stressed (p<0.00001) (Supplemental Table S6). Treatment explained most of the variation in stomatal conductance (62.3%), while maternal line and sampling site explained only a small amount of variation (5.8% and 0.9% respectively) (Supplemental Fig. S7). For the remaining three response variables (leaf length to width ratio, SPAD, and SLA), much of the variation was unexplained (40.5%-70%). Experimental condition was not statistically significant and explained little to no variation (0.0-4.4%) (Supplemental Fig. S7 and Table S6). Sampling site and maternal line were statistically significant in the

linear models at the p=0.05 level and explained some variation (6.7-21.2%) (Supplemental Fig. S7 and Table S6). Smaller, thicker leaves, and thus lower SLA values, were expected for droughted seedlings and for seedlings grown from seed collected from drier areas. Consistent with this expectation, the seedlings subjected to drought conditions showed lower SLA values. However, seedlings from drier sampling sites (D and T3) showed higher SLA values than more mesic sites (B, G, and 11), contrary to expectation (Supplemental Fig. S7). None of the four response variables showed significant variation in phenotypic plasticity across sites (all maternal line/sampling site by experimental condition interactions p>0.13) (Supplemental Fig. S8 and Table S7).

In addition to measuring seedling growth traits, we also examined the shape of the leaves from the drought experiment seedlings. We noted substantial variation in leaf shape, both among sites and within (Supplemental Fig. S9). The remarkable amount of phenotypic variation in the seedlings is consistent with the high levels of genomic variation measured both among sites and within sites.

Discussion

Eucalyptus melliodora is a foundation species in a critically endangered woodland community that now occupies a fraction of its former distribution and is the subject of restoration projects across its native range. Our examination of the distribution of genomic and phenotypic variation across the range of this species provides valuable information for sourcing seed for restoration, including empirically defining local provenances and matching genotypes to predicted future environmental conditions.

Examining the relationship between genomic and geographic distance, we are able to empirically define "local" in this species to be on the order of 500 km, which is substantially farther than the current practice. This new definition encourages restoration projects to source seed more broadly across the landscape. In a highly fragmented landscape this will increase the number of potential source sites, potentially

313 enabling the collection of higher quality seed with increased genetic diversity
314 (Broadhurst et al., 2008). Incorporating more naturally occurring genomic variation can
315 increase the adaptive potential of the restored population by providing the substrate for
316 adaptation to rapidly changing environmental conditions.
317 By modelling genomic variation across the landscape, we can understand the
318 environmental factors that shape patterns of genomic variation and identify variation
319 suitable for predicted future climates. We found several environmental variables that
320 have played a role in the structure of genomic variation across the landscape. Of these,
321 the climate variables are predicted to change rapidly over time. Change in soil nitrogen
322 content might occur over longer time scales, but it is difficult to forecast due to complex
323 biotic feedbacks (Brevik, 2013). This suggests that optimal seed sourcing will need to
324 balance the tracking of rapidly changing climate variables with the need to account for
325 variables that are more stable due to their dependence on stable features of geology,
326 topography, or hydrology. This also highlights an important concern that key
327 environmental variables may become uncoupled, resulting in less than ideal conditions
328 for this species across the landscape.
329 Our analyses of phenotypic variation found no site-specific variation in
330 phenotypic plasticity that would enable us to identify provenances better able to cope
331 with rapid environmental change. However, plasticity is trait specific and traits that are
332 hypothesized to be important for establishment and survival should continue to be
333 investigated because they may provide valuable information for restoration projects.
334 Importantly, our growth experiments support the results of the genomic analyses,
335 showing the remarkable extent of variation both among sites and within sites, further
336 supporting our recommendation to source seed to incorporate the high level of variation
337 that occur naturally in *E. melliodora*.
338 The results of this study are promising for the future of *E. melliodora* across its
339 native distribution. We found high genomic and phenotypic diversity within sites, as well
340 as shared across the range. This naturally occurring variation can provide a basis for
341 adaptation to a rapidly changing environment and it should be incorporated into
342 restoration projects through strategic seed sourcing. It is important to note that our

343 genomic analyses were based on mature trees that predate extensive land clearing for
344 agriculture. The same analyses in seedlings or saplings at these sites may show
345 different results, although our phenotypic studies using seedlings produced concordant
346 results. It remains to be determined if human modifications of the landscape have
347 disrupted historical patterns of gene flow, resulting in more fragmented and inbred
348 populations.

349 Our landscape genomic model can guide seed selection by empirically
350 defining local provenances and identifying variation suitable for predicted future
351 climates. This understanding of the relationship between environmental and genomic
352 variation can be combined with other types of information, such as basic biological
353 knowledge of the ecological community and best agronomic practices in restoration, to
354 establish foundation species and ecosystems with the highest probability of success in
355 a rapidly changing environment.

356 Methods

357 Sample Collection

358 We obtained *E. melliodora* leaf samples from mature trees at 38 sites across the
359 species' range through a community science project described in Broadhurst et al. (*in*
360 *review*) (Supplemental Table S1). From each site, a citizen scientist collected leaf
361 samples from up to 30 trees, put the samples in silica gel for drying, and shipped them
362 to CSIRO for processing. In addition to leaf material, they also collected seeds from the
363 sampled trees when available. We sampled an additional seven trees planted at a
364 single site in Western Australia, well outside the species' natural distribution.

365 Genotyping by Sequencing

366 We selected 3 to 10 trees per sampling site for sequencing and we processed each of
367 the seven trees from Western Australia twice, using different leaves from the same tree,
368 to serve as technical replicates. No power analysis was used to determine sample size

during the design of the study. Sample size was determined based on our experience and judgment, with consideration of the availability of samples. We sequenced these 379 samples using a modified Genotyping-By-Sequencing (GBS) protocol (Elshire et al., 2011). Briefly, we extracted genomic DNA from approximately 50 mg of leaf tissue using the Qiagen DNeasy Plant 96 Kit, digested with PstI for genome complexity reduction, and ligated with a uniquely barcoded sequencing adapter pair. We then individually PCR amplified each sample to avoid sample bias. We pooled samples in equimolar concentrations and extracted library amplicons between 350 and 600 bp from an agarose gel. We sequenced the library pool on an Illumina HiSeq2500 using a 101-bp paired-end protocol at the Biomolecular Resource Facility at the Australian National University, generating almost 260 million read pairs.

We checked the quality of the raw sequencing reads with FastQC (v0.10.1, (Andrews, 2012)). We used AXE (v0.2.6, (Murray & Borevitz, 2017a)) to demultiplex the sequencing reads according to each sample's unique combinatorial barcode and were unable to assign 11% of read pairs to a sample. We used trimit from libqcpp (v0.2.5, (Murray & Borevitz, 2017b)) to clean the reads for each sample, using default parameters, except q=20. This involved removing adapter contamination due to read-through of small fragments (20% of read pairs) and merging overlapping pairs (49% of read pairs), both steps using algorithms based on a global alignment of read pairs. We also used trimit for sliding window quality trimming (11% of reads). We then used a custom script to remove sequencing reads that did not begin with the expected restriction site sequence (16% of reads). We aligned sequencing reads to the *E. grandis* reference genome (v2.0, (Bartholomé et al., 2015; JGI, 2015; Myburg et al., 2014)), including all nuclear, chloroplast, mitochondrial, and ribosomal scaffolds, but used only nuclear scaffolds for downstream analyses. We aligned reads using bwa-mem (v0.7.5a-r405, (Li, 2013)), as paired reads (-p) and treating shorter split hits as secondary alignments (-M), with 88% of reads successfully mapped. We used GATK's HaplotypeCaller in GVCF mode (v3.6-0-g89b7209, (McKenna et al., 2010)) to call variants for each sample with heterozygosity (-hets) increased to 0.005, indel

heterozygosity (-indelHeterozygosity) increased to 0.0005, and the minimum number of reads sharing the same alignment start (-minReadsPerAlignStart) decreased to 4.

We used GATK's GenotypeGVCFs (v3.6-0-g89b7209, (McKenna et al., 2010)) for a preliminary round of joint genotyping across all samples based on the individual variant calls and again increasing the heterozygosity (-hets) to 0.005 and the indel heterozygosity (-indelHeterozygosity) to 0.0005. For basic filtering, we used GATK to remove variants that were indels, had no variation relative to the reference, were non-biallelic SNPs, had QD<2.0 ("variant call confidence normalized by depth of sample reads supporting a variant"), MQ>40.0 ("Root Mean Square of the mapping quality of reads across all samples"), or MQRankSum<-12.5 ("Rank Sum Test for mapping qualities of REF versus ALT reads"). We removed samples with more than 60% missing data and SNPs with more than 80% missing data. We examined the genomic distance between samples to verify that technical replicates clustered closely together. We used a preliminary PCA, based on genomic distance between samples, to identify outlier samples. We removed outlier samples and poorly sequenced samples from the dataset for final genotyping and all downstream analyses.

We reran GATK's joint genotyping on the final sample set. We again used GATK to remove variants that were indels, SNPs with no variation relative to the reference, and non-biallelic SNPs. We determined final filtering thresholds by examining parameter distributions. A locus was retained for subsequent analysis if ExcessHet<13.0 ("phred-scaled p-value for exact test of excess heterozygosity"), -0.3<InbreedingCoeff<0.3 ("likelihood-based test for the inbreeding among samples"), MQ>15.0 ("Root Mean Square of the mapping quality of reads across all samples"), -10.0<MQRankSum<10.0 ("Rank Sum Test for mapping qualities of REF versus ALT reads"), and QD>8.0 ("variant call confidence normalized by depth of sample reads supporting a variant"). We ran a second preliminary PCA analysis to identify additional outlier samples. Finally, we used VCFtools (v0.1.12b, (Danecek et al., 2011)) to remove SNPs with greater than 60% missing data and thin the SNPs so that none were closer than 300 bp.

Genomic Analyses

To examine the genomic structure of *E. melliodora* and how it is influenced by geography, we conducted individual-based analyses. For these analyses, we converted the final genotypic data (a vcf file) to a sample-by-SNP matrix and imported it into a genind object (R adegenet v2.0.1, (Jombart, 2008)). We calculated the pairwise genomic distances between individuals using a euclidean distance in *dist* (R stats v3.1.2, (R Core Team, 2015)). To visualize the genomic distance among samples, we ran a PCA using *dudi.pco* (R ade4 v1.7-4, (Dray & Dufour, 2007)). We plotted the first two PCA components, with samples colored in a rainbow gradient based on sample latitude. We calculated the geographic distance between samples based on their GPS coordinates using *earth.dist* (R fossil v0.3.7, (Vavrek, 2011)). We used a *mantel* test (R vegan v2.4-0, (Oksanen et al., 2016)) to quantify the linear relationship between the genomic distance between individuals and the natural log of geographic distance.

To examine the role that environmental factors play in driving the genomic structure across the landscape, we used Generalized Dissimilarity Modelling (GDM), which uses matrix regression to estimate the non-linear relationship between genomic distance and environmental distance (Ferrier et al., 2007; Fitzpatrick & Keller, 2015; Thomassen et al., 2011). We then used this model to predict the distribution of genomic variation across the landscape under current environmental conditions, as well as predicted future conditions. We obtained environmental data for the GDM from climate, elevation, soil, and landscape raster layers. Climate variables included 19 bioclimatic variables for the current time period (1960-1990), at 30 arc second resolution (WorldClim, 2016b). Elevation was from a digital elevation model aggregated from 90 m resolution (CGIAR-CSI, 2016). Soil data included available water capacity, total nitrogen, and total phosphorus sampled at the surface (0-5 cm) and at depth (100-200 cm) (CSIRO, 2016). Landscape data included the Prescott Index (a measure of water balance) and MrVBF (a topographic index) (CSIRO, 2016). For future prediction, we used the 19 bioclimatic variables predicted for 2070 at 30 arc second resolution based on GCM MIROC5 for representative concentration pathway 8.5 (WorldClim, 2016a), which is a greenhouse gas concentration trajectory showing continual increase in

457 emissions over time. We determined the values for each variable at each sampling site
458 based on GPS coordinates and used those values to calculate the environmental
459 distances between sites.
460 To determine the genomic distances between sampling sites, we used the
461 sample by SNP matrix to calculate pairwise F_{st} (Weir & Cockerham, 1984) using
462 *pairwise.WCfst* (R hierfstat v0.04-22, (Goudet & Jombart, 2015)). We ran a sampling
463 site level PCA on the pairwise F_{st} matrix using *dudi.pco* (R ade4 v1.7-4, (Dray & Dufour,
464 2007)) and calculated the percent of variation explained for each PCA axis. For the
465 GDM, we scaled the F_{st} matrix to between 0 and 1 by subtracting the minimum value
466 and then dividing by the maximum value. We generated the GDM model using *gdm* (R
467 gdm v1.2.3, (Manion et al., 2016)) with the scaled F_{st} matrix, geographic distance
468 between sites, and environmental distances for the 28 variables for the current time
469 period. Initially, we generated a GDM model for each environmental variable separately
470 and excluded variables from further analysis if the deviance explained by the model was
471 less than 5%. For the remaining variables, we calculated Pearson's correlation for site
472 values between pairwise sets of variables. If a pair of variables had a correlation
473 greater than 60%, we excluded the variable with the lowest explanatory power from
474 subsequent analysis. We conducted permutation testing using gdm.varImp (R gdm
475 v1.2.3, (Manion et al., 2016)) with 1000 permutations to determine the explanatory
476 power and statistical significance of the remaining variables and excluded additional
477 inconsequential variables. We generated a final GDM model with the remaining
478 environmental variables.
479 We cross validated the GDM model using a random 70% of the spatial
480 sampling sites as training data and the remaining 30% of sites as test data and ran
481 1000 resampled iterations. We used the GDM models from the training sites to predict
482 the genomic dissimilarity between the test sites and used Pearson's correlation to
483 compare the predicted values to the observed values. To test the robustness of the
484 geographic prediction from the GDM model, we visualized the geographic splines from
485 100 of these GDM models.

To project the final GDM model onto the current environmental landscape, we first delineated the geographic extent of the analysis by defining an *E. melliodora* distribution polygon. We downloaded 14,977 *E. melliodora* occurrence records from the Atlas of Living Australia (ALA, 2016), of which we removed 189 because they were well outside the expected distribution or were sparse records on the distribution margin. We generated the polygon using ahull (R alphahull v2.1, (Pateiro-López & Rodríguez-Casal, 2010)), with alpha=15 and gBuffer (R rgeos v0.3-21, (Bivand & Rundel, 2016)), with a 20 km buffer. We then transformed the environmental rasters based on the model splines (*gdm.transform*), took a PCA of the transformed layers (*prcomp* R stats v3.1.2, (R Core Team, 2015)), and predicted across space (*predict*). We visualized the result by graphing the first three components of a PCA using a red-green-blue plot (Fitzpatrick & Keller, 2015). We also projected the model onto a predicted future environmental landscape with the same procedure, except we substituted the current bioclimatic rasters with the future ones for 2070 that were predicted under a high CO2 emission scenario. We calculated the expected change in the distribution of genomic variation over time using the *predict* function with time=T.

We examined the implications of the GDM model for seed sourcing decisions by selecting two hypothetical reforestation sites. We compared predicted future GDM values at these two hypothetical reforestation sites to current climate GDM values across the landscape of potential seed sources. This enabled us to generate a map of predicted genomic similarity of potential seed sources to the hypothetical reforestation sites under climate change.

Growth Experiments

To examine the effect of provenance and environment on phenotype, we conducted experiments in climate controlled growth chambers under two different climate regimes. No power analysis was used to determine sample size during the design of the experiment. Sample size was determined based on our experience and judgment, with consideration of the availability of seed and space in the growth chambers. We selected six sites (11, B, D, G, T1, T3) and six maternal trees per site that had sufficient

seed (asterisks in Fig. 1A). For each chamber, we grew eight or nine replicate seedlings from each maternal tree. To ensure we had a seedling for each intended replicate, four seeds were planted per pot (6.5 cm x 6.5 cm x 20 cm pots with soil that was 80% Martin's mix and 20% sand). We germinated seeds in climate controlled chambers with 12 hours of light at 25°C and 12 hours of dark at 15°C. We set lights to mimic summer morning light (photosynthetic photon flux 370 nm=82, 400 nm=83, 420 nm=78, 450 nm=37, 530 nm=31, 620 nm=72, 660 nm=28, 735 nm=34, 850 nm=89, 6500 K=94 μmol/m^2/s). We watered all seeds twice daily to keep the soil moist. We culled to one seedling per pot 12-14 days after planting.

Three weeks after germination, we sorted seedlings into treatment chambers based on a randomized block design. Climate conditions were determined with SolarCalc (Spokas & Forcella, 2006) to mimic average summer conditions (sampling site 11) and hotter conditions (5°C temperature increase; sampling site T3). We ran the experimental conditions for 12-14 weeks and took phenotypic measurements at five time points (1, 2, 3, 5, and 11 weeks after the experimental treatment began). Measurements included seedling height, number of leaves, and total leaf length.

For the analysis of seedling height and total leaf length, we used the measurements at five weeks after the experimental treatment began and used only seedlings that were determined to be well established at that time. We also calculated a relative height increment for each seedling by determining the last measurement when the seedling had two or fewer leaves and the first measurement with eight or more leaves. The relative height increment is the difference between the natural log of the two height measures, divided by the difference in time.

We investigated phenotypic plasticity by examining interaction plots between maternal line and experimental conditions for three response variables: seedling height, total leaf length, and relative height increment. We statistically tested for an interaction between sampling site/maternal line and experimental condition with linear mixed-effect models using *lmer (*R lme4 v1.1-10, (Bates et al., 2015)) for each of the three response variables. Due to a lack of power to consider maternal line nested within sampling site, we ran two models for each response variable—one with maternal line and one with

sampling site. These models included the experimental condition, sampling site or maternal line, and and their interactions as fixed effects. We included germination chamber and block as random effects. We identified outliers visually and ran the models with and without outliers to determine if they affected the results.

We visualized the distribution of values for the three response variables across the six sampling sites using box plots. We quantified the distribution of phenotypic variation with linear mixed-effect models using *lmer (*R lme4 v1.1-10, (Bates et al., 2015)). For each of the three response variables, the model included maternal line nested within sampling site and experimental condition as main effects, with no interaction term, and germination chamber and block as random effects.

After completion of the chamber experiment, we conducted an outdoor covered drought experiment on the 16 week old seedlings. No power analysis was used to determine sample size during the design of the experiment. Sample size was determined based on our experience and judgment, with consideration of the availability of space in the covered growth facility. We selected 160 seedlings from five sampling sites, with four maternal lines per site. We paired each seedling with a seedling of similar size from the same maternal line and treatment chamber. We randomly assigned each seedling of the pair to a different drought treatment group. We transplanted the seedlings to PVC tubes (9 cm diameter x 50 cm height with sand, perlite, and slow release osmocote) and kept them well watered for seven weeks, allowing them to acclimate to the outdoor conditions. Then we imposed two treatments: well watered and drought. For the well watered treatment, we watered the seedlings to saturation as needed (between three times per week and twice per day, depending on the weather). For the drought treatment, we watered as necessary to reach (but not exceed) 50% saturation.

We measured leaf traits on each seedling three weeks after the initiation of treatment. We measured stomatal conductance with a porometer (SC-1 Leaf Porometer by Decagon Devices) and determined that water stress was induced in the droughted seedlings. We determined the leaf length to width ratio from a scan of the most recent fully expanded leaf from each seedling using image analysis software

(WD3 WinDIAS Leaf Image Analysis System by Delta-T Devices). This leaf was initiated prior to the start of treatment, but expanded while under treatment conditions. We took additional measurements two months after the initiation of treatment. We used a chlorophyll meter (SPAD – 502 by Konica Minolta) to determine the SPAD index, which measures relative chlorophyll content; reduction in SPAD index would indicate detrimental effects of water limitation. We calculated specific leaf area (SLA, leaf area divided by dry mass) by scanning a single leaf from each seedling to determine the leaf area (WD3 WinDIAS Leaf Image Analysis System by Delta-T Devices) and weighed oven dried leaves. For analysis, we excluded data for seedlings that died during the experiment. We also excluded the experimental treatment pair of any dead seedlings.

We visualized phenotypic plasticity by examining interaction plots between maternal line and experimental conditions for four response variables: stomatal conductance, leaf length to width ratio, SPAD index, and SLA. We statistically tested for an interaction between sampling site/maternal line and experimental condition with linear mixed-effect models using *lmer (*R lme4 v1.1-10, (Bates et al., 2015)) for each of the four response variables. Due to a lack of power to consider maternal line nested within sampling site, we ran two models for each response variable--one with maternal line and one with sampling site. These models included the experimental condition, sampling site or maternal line, and their interactions as fixed effects. We included block and sample pairings as random effects.

We visualized the distribution of values for the four response variables across the five sampling sites using box plots. We quantified the distribution of phenotypic variation with linear mixed-effect models using *lmer (*R lme4 v1.1-10, (Bates et al., 2015)). For each of the four response variables, the model included maternal line nested within sampling site and experimental condition as main effects, with no interaction term, and block and sample pairings as random effects. Due to a lack of power, the p-value for the sampling site term was determined from a model without maternal line.

Acknowledgments

We thank the ANU Bioinformatics Consultancy for computational support and advice, the ANU Statistical Consulting Unit for statistical advice, the Centre for Biodiversity Analysis for advice on GDM modelling, and the ANU NCRIS Plant Growth Facility and RSB Plant Services for assistance with growth experiments.

Data Access

GBS sequencing reads are available at the NCBI Sequence Read Archive (SRA) (http://www.ncbi.nlm.nih.gov/sra) under BioProject PRJNA413429. Growth experiment data and scripts for genomic and phenotypic analyses are available at https://github.com/LaMariposa/emelliodora.

References

Aitken, S. N., & Whitlock, M. C. (2013). Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 367–388. https://doi.org/10.1146/annurev-ecolsys-110512-135747

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, *1*(1), 95–111. https://doi.org/10.1111/j.1752-4571.2007.00013.x

ALA. (2016). Atlas of Living Australia. Retrieved November 24, 2016, from http://www.ala.org.au/

Andrew, R. L., Wallis, I. R., Harwood, C. E., & Foley, W. J. (2010). Genetic and environmental contributions to variation and population divergence in a broad-spectrum foliar defence of Eucalyptus tricarpa. *Annals of Botany*, *105*(5), 707–717. https://doi.org/10.1093/aob/mcq034

628 Andrews, S. (2012). FastQC. Retrieved from
629 https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

630 Bartholomé, J., Mandrou, E., Jenkins, J., Nabihoudine, I., Klopp, C., Schmutz, J.,
631 Plomion, C., & Gion, J.-M. (2015). High-resolution genetic maps of Eucalyptus
632 improve Eucalyptus grandis genome assembly. *New Phytologist*, *206*, 1283–1296.

633 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects
634 Models Using lme4. *Journal of Statistical Software*, *67*(1).
635 https://doi.org/10.18637/jss.v067.i01

636 Bivand, R., & Rundel, C. (2016). rgeos: Interface to Geometry Engine - Open Source
637 (GEOS).

638 Brevik, E. (2013). The Potential Impact of Climate Change on Soil Properties and
639 Processes and Corresponding Influence on Food Security. *Agriculture*, *3*(3), 398–
640 417. https://doi.org/10.3390/agriculture3030398

641 Broadhurst, L. M., Lowe, A., Coates, D. J., Cunningham, S. A., McDonald, M., Vesk, P.
642 A., & Yates, C. (2008). Seed supply for broadscale restoration: Maximizing
643 evolutionary potential. *Evolutionary Applications*, *1*(4), 587–597.
644 https://doi.org/10.1111/j.1752-4571.2008.00045.x

645 Broadhurst, L. M., Mellick, R., Knerr, N., Li, L., & Supple, M. A. (n.d.). Land availability
646 may be more important than genetic diversity in the range shift response of the
647 widely distributed eucalypt Eucalyptus melliodora (Yellow Box, Myrtaceae).

648 Byrne, M., Prober, S. M., McLean, E. H., Steane, D. A., Stock, W. D., Potts, B. M., &
649 Vaillancourt, R. E. (2013). *Adaptation to climate in widespread eucalypt species*.
650 Gold Coast: National Climate Change Adaptation Research Facility.

651 CGIAR-CSI. (2016). SRTM 90m Digital Elevation Data. Retrieved September 5, 2016,
652 from http://srtm.csi.cgiar.org/

653 CSIRO. (2016). Soil and Landscape Grid of Australia. Retrieved November 29, 2016,
654 from http://www.clw.csiro.au/aclep/soilandlandscapegrid/index.html

655 Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A.,
656 Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R.
657 (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158.
658 https://doi.org/10.1093/bioinformatics/btr330

659 Department of Environment and Climate Change and Water. (2011). National Recovery
660 Plan for White Box - Yellow Box - Blakely's Red Gum Grassy Woodland and
661 Derived Native Grassland.

662 Department of the Environment and Heritage. (2006). *EPBC Act Policy Statements:*
663 *White Box - Yellow Box - Blakely's Red Gum grassy woodlands and derived native*
664 *grasslands*.

665 Dray, S., & Dufour, A.-B. (2007). The ade4 Package: Implementing the Duality Diagram
666 for Ecologists. *Journal of Statistical Software*, *22*(4).
667 https://doi.org/10.18637/jss.v022.i04

668 Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., &
669 Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach
670 for high diversity species. *PloS One*, *6*(5), e19379.
671 https://doi.org/10.1371/journal.pone.0019379

672 Ferrier, S., Manion, G., Elith, J., & Richardson, K. (2007). Using generalized dissimilarity
673 modelling to analyse and predict patterns of beta diversity in regional biodiversity
674 assessment. *Diversity and Distributions*, *13*(3), 252–264.
675 https://doi.org/10.1111/j.1472-4642.2007.00341.x

676 Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level
677 modelling of biodiversity: Mapping the genomic landscape of current and future
678 environmental adaptation. *Ecology Letters*, *18*(1), 1–16.
679 https://doi.org/10.1111/ele.12376

680 González-Orozco, C. E., Pollock, L. J., Thornhill, A. H., Mishler, B. D., Knerr, N., Laffan,
681 S. W., Miller, J. T., Rosauer, D. F., Faith, D. P., Nipperess, D. A., Kujala, H., Linke,
682 S., Butt, N., Külheim, C., Crisp, M. D., & Gruber, B. (2016). Phylogenetic
683 approaches reveal biodiversity threats under climate change. *Nature Climate*
684 *Change*, *1*(12), 1110–1114. https://doi.org/10.1038/nclimate3126

685 Goudet, J., & Jombart, T. (2015). hierfstat: Estimation and Tests of Hierarchical F-
686 Statistics.

687 Hoffmann, A., Griffin, P., Dillon, S., Catullo, R., Rane, R., Byrne, M., Jordan, R.,
688 Oakeshott, J., Weeks, A., Joseph, L., Lockhart, P., Borevitz, J., & Sgrò, C. (2015). A
689 framework for incorporating evolutionary genomics into biodiversity conservation
690 and management. *Climate Change Responses*, *2*(1).
691 https://doi.org/10.1186/s40665-014-0009-x

692 JGI. (2015). Eucalyptus grandis v2.0 (Rose gum). Retrieved August 27, 2015, from
693 https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Egrandis_er

694 Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic
695 markers. *Bioinformatics*, *24*(11), 1403–1405.
696 https://doi.org/10.1093/bioinformatics/btn129

697 Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with
698 BWA-MEM. *arXiv*, 1303.3997v1.

699 Lindenmayer, D. B., Steffen, W., Burbidge, A. A., Hughes, L., Kitching, R. L., Musgrave,
700 W., Stafford Smith, M., & Werner, P. A. (2010). Conservation strategies in response
701 to rapid climate change: Australia as a case study. *Biological Conservation*, *143*(7),
702 1587–1593. https://doi.org/10.1016/j.biocon.2010.04.014

703 Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., & Fitzpatrick, M. C. (2016). gdm:
704 Functions for Generalized Dissimilarity Modeling.

705 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A.,
706 Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The
707 Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation
708 DNA sequencing data. *Genome Research*, *20*(9), 1297–1303.
709 https://doi.org/10.1101/gr.107524.110

710 McLean, E. H., Prober, S. M., Stock, W. D., Steane, D. A., Potts, B. M., Vaillancourt, R.
711 E., & Byrne, M. (2014). Plasticity of functional traits varies clinally along a rainfall
712 gradient in Eucalyptus tricarpa. *Plant, Cell and Environment*, *37*(6), 1440–1451.
713 https://doi.org/10.1111/pce.12251

714 Murray, K. D., & Borevitz, J. O. (2017a). Axe: rapid, competitive sequence read
715 demultiplexing using a trie. *bioRxiv*. https://doi.org/10.1101/160606

716 Murray, K. D., & Borevitz, J. O. (2017b). libqcpp : A C++14 sequence quality control
717 library. *The Journal of Open Source Software*, *2*(13).
718 https://doi.org/10.21105/joss.00232

719 Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J.,
720 … Schmutz, J. (2014). The genome of Eucalyptus grandis. *Nature*, *509*(7505),
721 356–362. https://doi.org/10.1038/nature13308

722 Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius,
723 U., Poot, P., Purugganan, M. D., Richards, C. L., Valladares, F., & van Kleunen, M.

(2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science*, *15*(12), 684–692. https://doi.org/10.1016/j.tplants.2010.09.008

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2016). vegan: Community Ecology Package.

Pateiro-López, B., & Rodríguez-Casal, A. (2010). Generalizing the Convex Hull of a Sample : The R Package Alphahull. *Journal of Statistical Software*, *34*(5). https://doi.org/http://dx.doi.org/10.18637/jss.v034.i05

Prober, S. M., Byrne, M., McLean, E. H., Steane, D. A., Potts, B. M., Vaillancourt, R. E., & Stock, W. D. (2015). Climate-adjusted provenancing: a strategy for climate-resilient ecological restoration. *Frontiers in Ecology and Evolution*, *3*(65). https://doi.org/10.3389/fevo.2015.00065

R Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Spokas, K., & Forcella, F. (2006). Estimating hourly incoming solar radiation from limited meteorological data. *Weed Science*, *54*(1), 182–189. https://doi.org/10.1614/WS-05-098R.1

Thomassen, H. A., Fuller, T., Buermann, W., Milá, B., Kieswetter, C. M., Jarrín-V., P., … Smith, T. B. (2011). Mapping evolutionary process: A multi-taxa approach to conservation prioritization. *Evolutionary Applications*, *4*(2), 397–413. https://doi.org/10.1111/j.1752-4571.2010.00172.x

Threatened Species Scientific Committee. (2006). *Advice to the Minister for the Environment and Heritage*.

Vavrek, M. J. (2011). Fossil: Palaeoecological and Paelaeogeogrpahical Analysis Tools. *Palaeontologia Electronica*, *14*(1), 1–16.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*(23), 5649–5662. https://doi.org/10.1111/mec.12938

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, *38*(6), 1358–1370. https://doi.org/10.2307/2409936

Williams, A. V, Nevill, P. G., & Krauss, S. L. (2014). Next generation restoration genetics: Applications and opportunities. *Trends in Plant Science*, *19*(8), 529–537. https://doi.org/10.1016/j.tplants.2014.03.011

WorldClim. (2016a). CMIP5 30-seconds. Retrieved September 30, 2016, from http://www.worldclim.org/cmip5_30s

WorldClim. (2016b). WorldClim 1.4. Retrieved November 28, 2016, from http://www.worldclim.org/current

Figure Legends

Figure 1: Map of sampling sites and PCA of genomic distance between samples

(A) A map of the geographic locations of the 37 sampling sites in southeastern Australia. Sampling locations are labeled with the site name and color coded in a rainbow gradient based on latitude. Black asterisks indicate 6 sites used for growth chamber experiments. The gray background shading indicates the species distribution polygon. (B) Principal coordinate analysis of the genomic distance between individual samples. Samples are labeled with a sample name that indicates the site name and tree number. Samples are color coded by site to match the map. The percentage on each axis indicates how much of the genomic variation between individuals was explained by the axis. (C) Principal coordinate analysis of F_{st} between sampling sites. Sampling sites are labeled by name and color coded to match the map. The percentage on each axis indicates how much of the variation in F_{st} between sampling sites was explained by the axis.

FIgure 2: Estimated genomic variation as a function of geographic distance

The geographic spline estimated from the GDM model showing little predicted genomic change between sites less than 500 km apart and increasing genomic variation as geographic distance increases beyond 500 km.

Figure 3: Predicted spatial and temporal variation in genomic composition

(A) The spatial distribution of predicted genomic variation based on projecting the GDM model onto geography and current environmental conditions. Regions with similar colors are predicted to have similar genomic composition. (B) The predicted temporal genomic variation based on comparing the GDM model projected onto current environmental conditions and predicted environmental conditions for 2070. The higher the difference (green colors), the more genomic

783 change predicted between current and 2070 conditions. Sampling sites are labeled in black
784 text.

785 **Figure 4: Optimal seed sourcing locations for hypothetical reforestation sites**
786 The predicted genomic similarity of hypothetical reforestation sites (indicated by black circles) to
787 potential seed sourcing locations under a climate change scenario for 2070. Dark green areas
788 indicate seed sourcing areas predicted to best match future conditions at the hypothetical
789 reforestation sites; white and brown areas indicate areas of potential genomic mismatch.

790 **Supplementary Figures**
791 **Figure S1: Technical replicate dendrogram**
792 Dendrogram based on genomic distance between samples showing the strong clustering of
793 technical replicates (denoted with an "R" after the sample name and highlighted in yellow).
794 Note that three of the technical replicates failed to pass quality control and are not included in
795 the dendrogram. Additional sample pairs show strong clustering. In each case, the individuals
796 of the pair are from the same sampling site, indicating samples that are closely related.

797 **Figure S2: Species identification PCA**
798 PCA of genomic distance between samples showing strong outliers that are likely misidentified
799 samples or hybrids. The vertical line at 50 on PCA axis 1 indicates the cutoff, with all samples
800 to the right removed from further analyses.

801 **Figure S3: Outlier PCA**
802 PCA of genomic distance between samples for the confirmed *E. melliodora* samples. The five
803 samples on the left were deemed outliers and removed from further analyses.

804 **Figure S4: Generalized dissimilarity modelling (GDM) results**
805 (A) Non-linear relationship between environmental distance and genomic distance. Points are
806 site pairs; the line is the predicted relationship. (B) Relationship between predicted genomic
807 distance and observed genomic distance. Points are site pairs; the line indicates where
808 observation and prediction match. (C-G) Predicted splines showing the estimated relationship
809 between the environmental variable and genomic distance for (C) total nitrogen content at 100-
810 200 cm of soil depth, (D) mean temperature of the coldest quarter, (E) precipitation of the

wettest quarter, (F) isothermality, and (G) geography. (H) Geographic splines from 100 iterations of sampling 70% of sites. Each solid black line is an iteration; dashed grey line is the full model prediction.

Figure S5: Variation in seedling growth in chamber experiment

Box plots showing variation between chamber treatments (left) and sampling sites (right) for three seedling growth traits.

Figure S6: Interaction plots for chamber experiment

Plots showing interactions between three seedling growth traits and the experimental conditions. Each line represents a maternal line, with color indicating the sampling site.

Figure S7: Variation in leaf traits in drought experiment

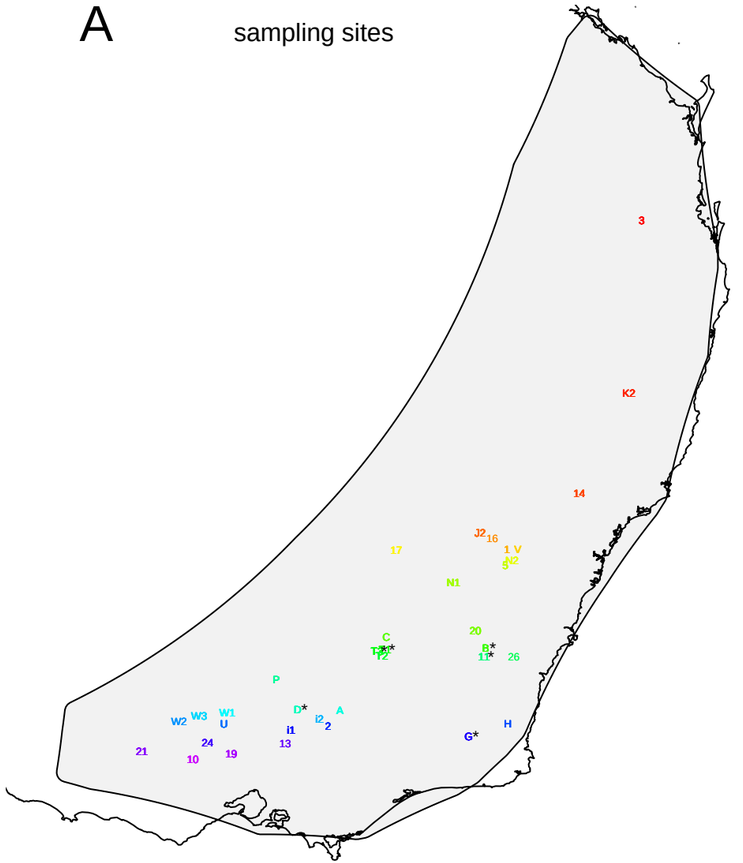
Box plots showing variation between water treatments (left) and sampling sites (right) for four leaf traits.

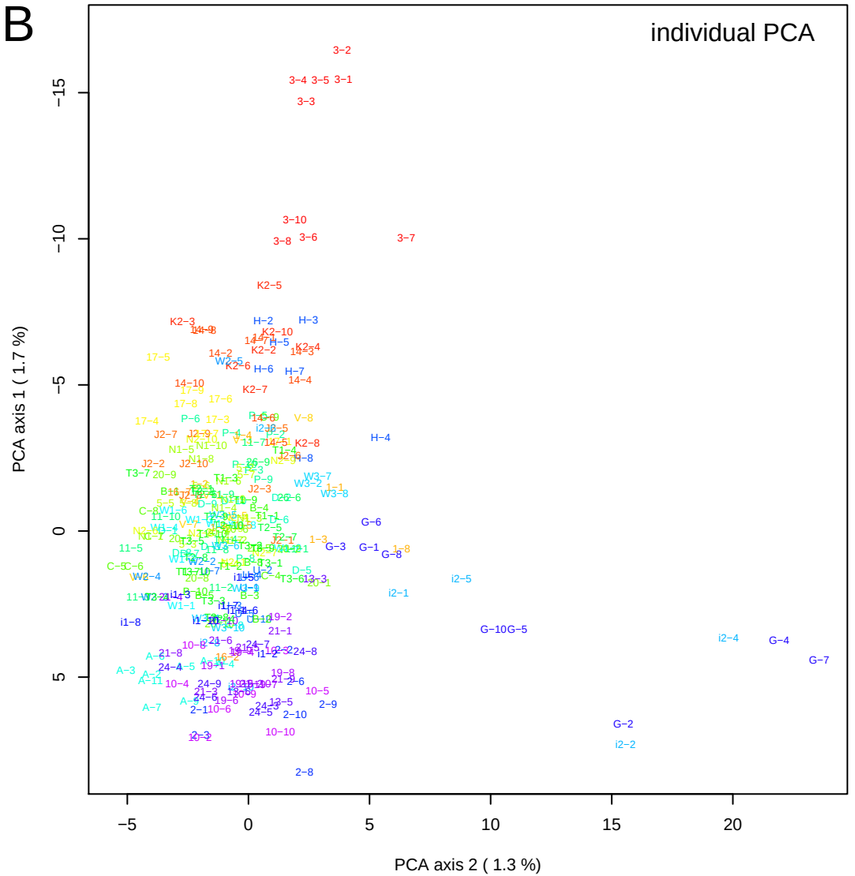
Figure S8: Interaction plots for drought experiment

Plots showing interactions between the four leaf traits and the water treatment. Each line represents a maternal line, with color indicating the sampling site.

Figure S9: Variation in leaf shape

One representative leaf from each maternal line in the drought experiment. Each row shows a single sampling site, identified by site ID and state location (ACT=Australian Capital Territory, VIC=Victoria, NSW=New South Wales). Each leaf is identified by its sampling site, maternal line, and replicate number).


A
sampling sites


B
individual PCA
PCA axis 1 (1.7 %)
PCA axis 2 (1.3 %)

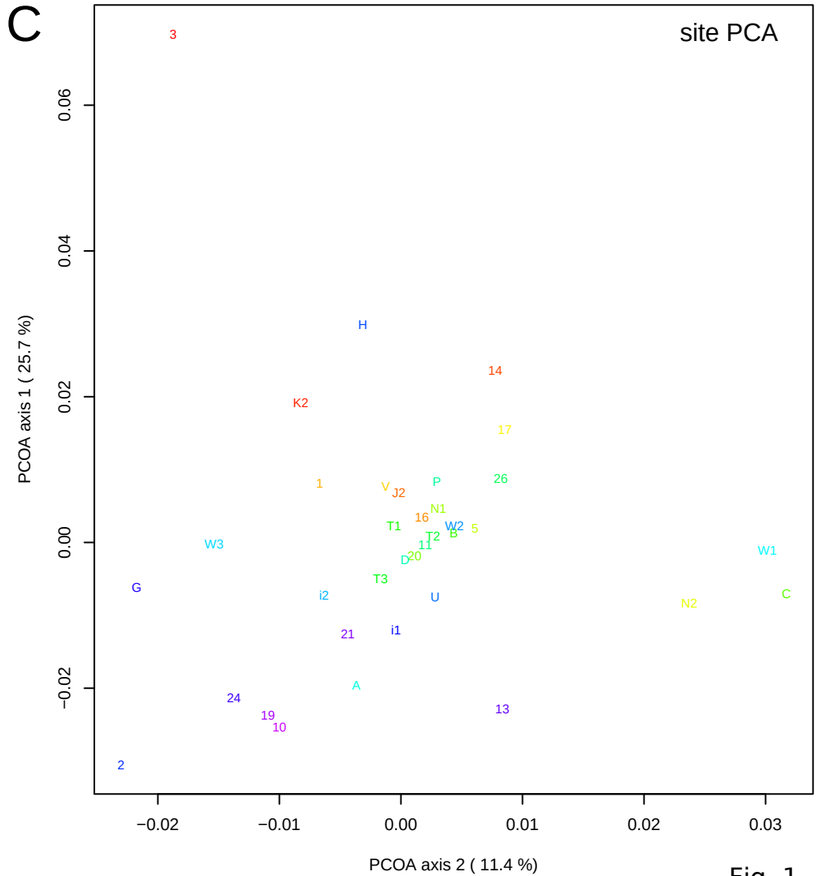

C
site PCA
PCOA axis 1 (25.7 %)
PCOA axis 2 (11.4 %)

Fig. 1

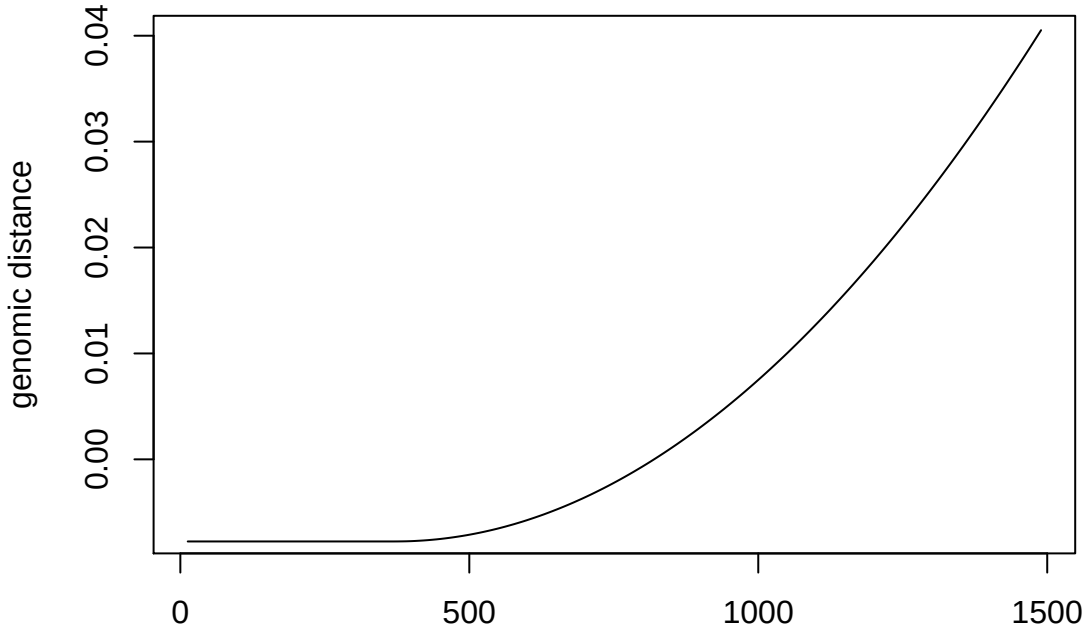

genomic distance
0.00
0.01
0.02
0.03
0.04
0
500
1000
1500
geographic distance (km)

Fig. 2

(A) Predicted Genomic Composition Across Space

(B) Predicted Change in Genomic Composition Over Time

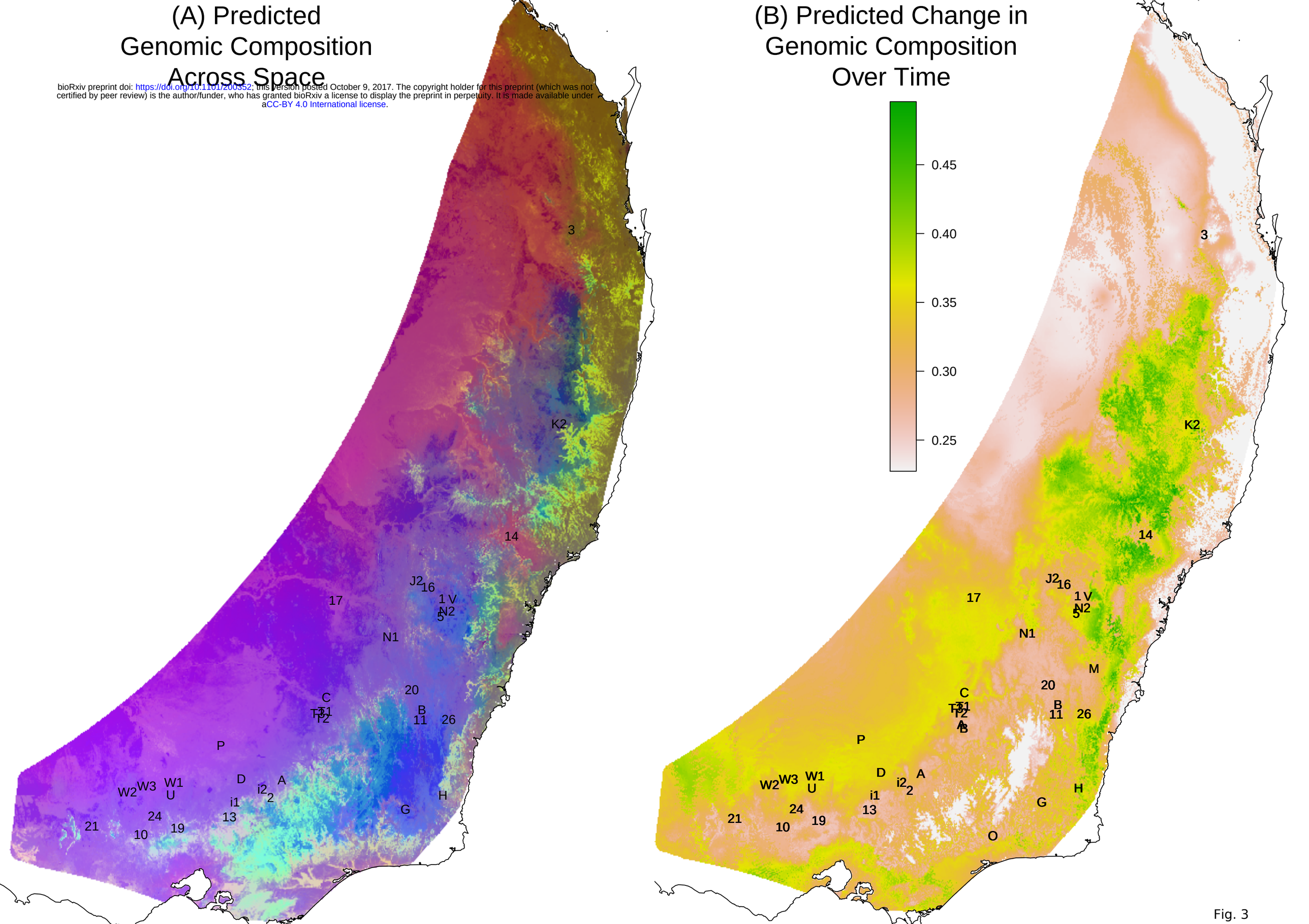


Fig. 3

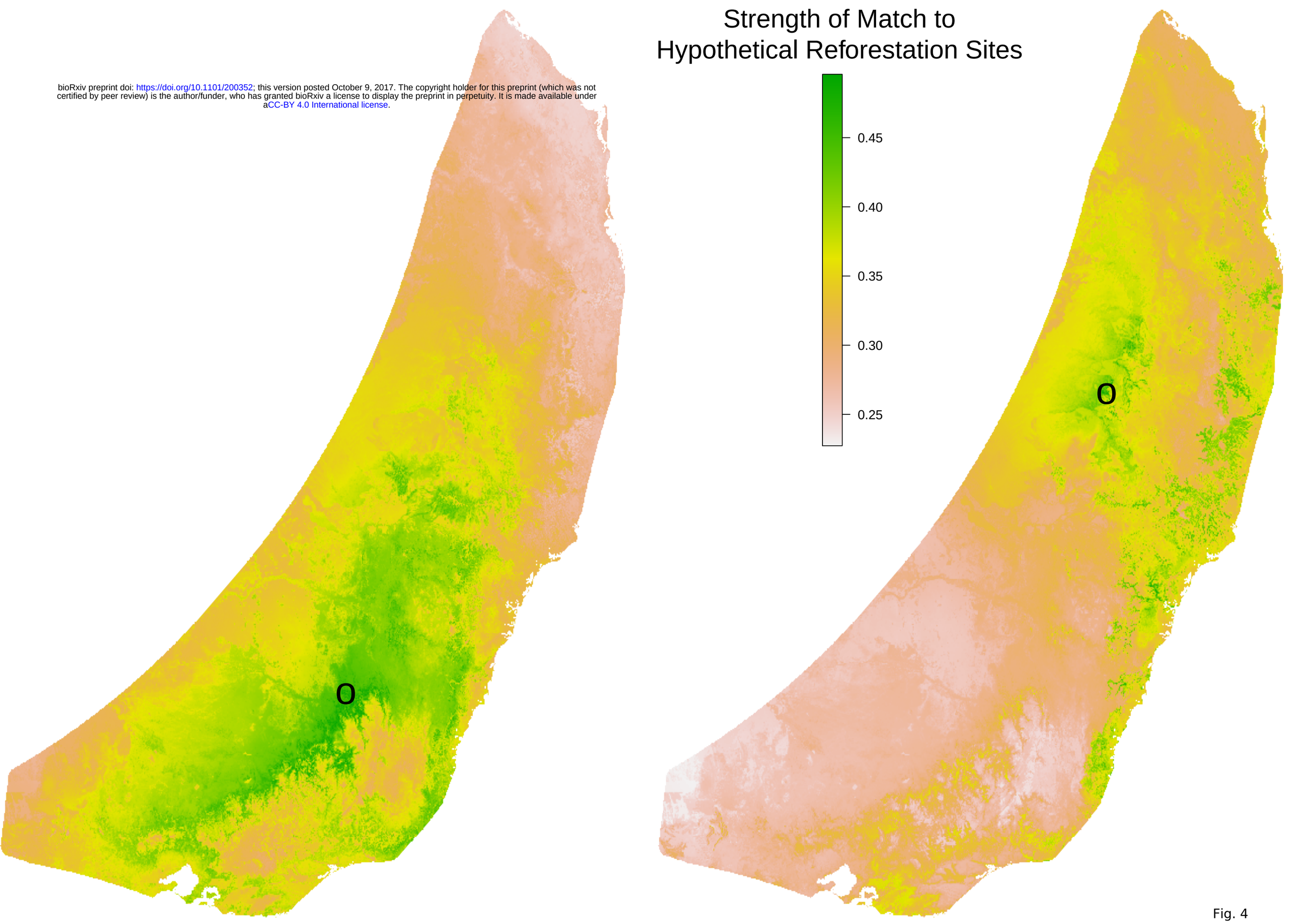


Fig. 4

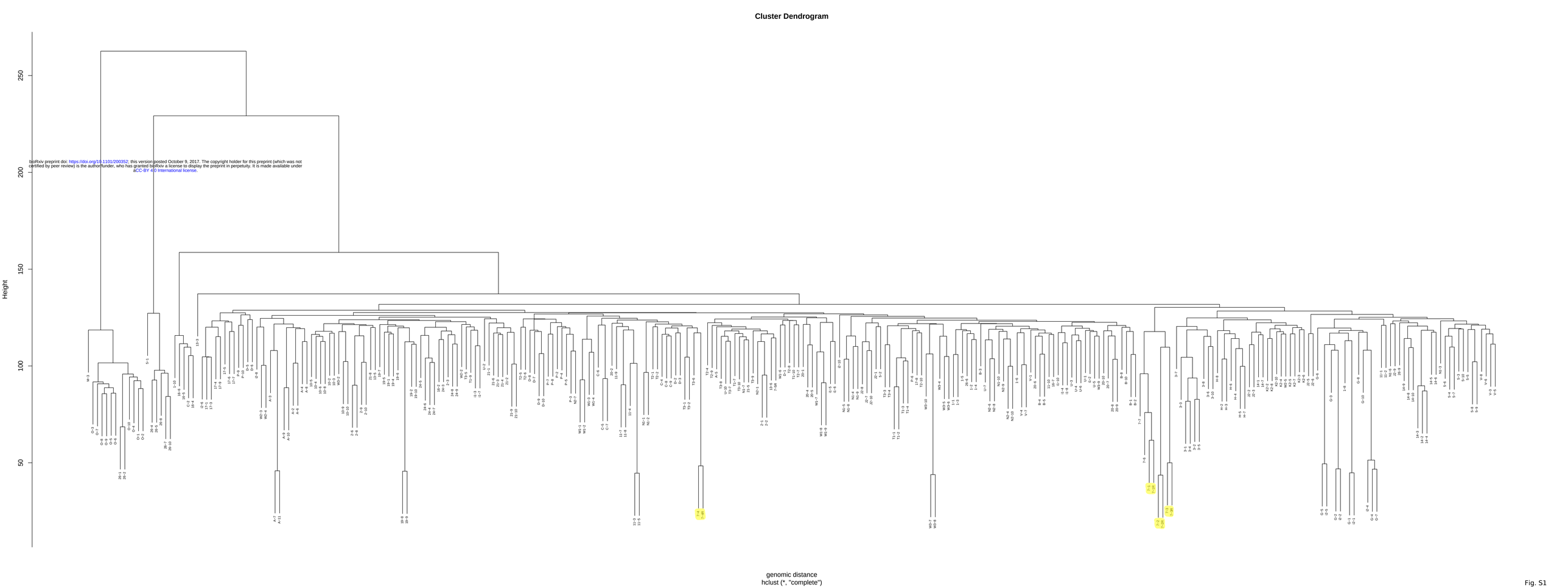


Fig. S1

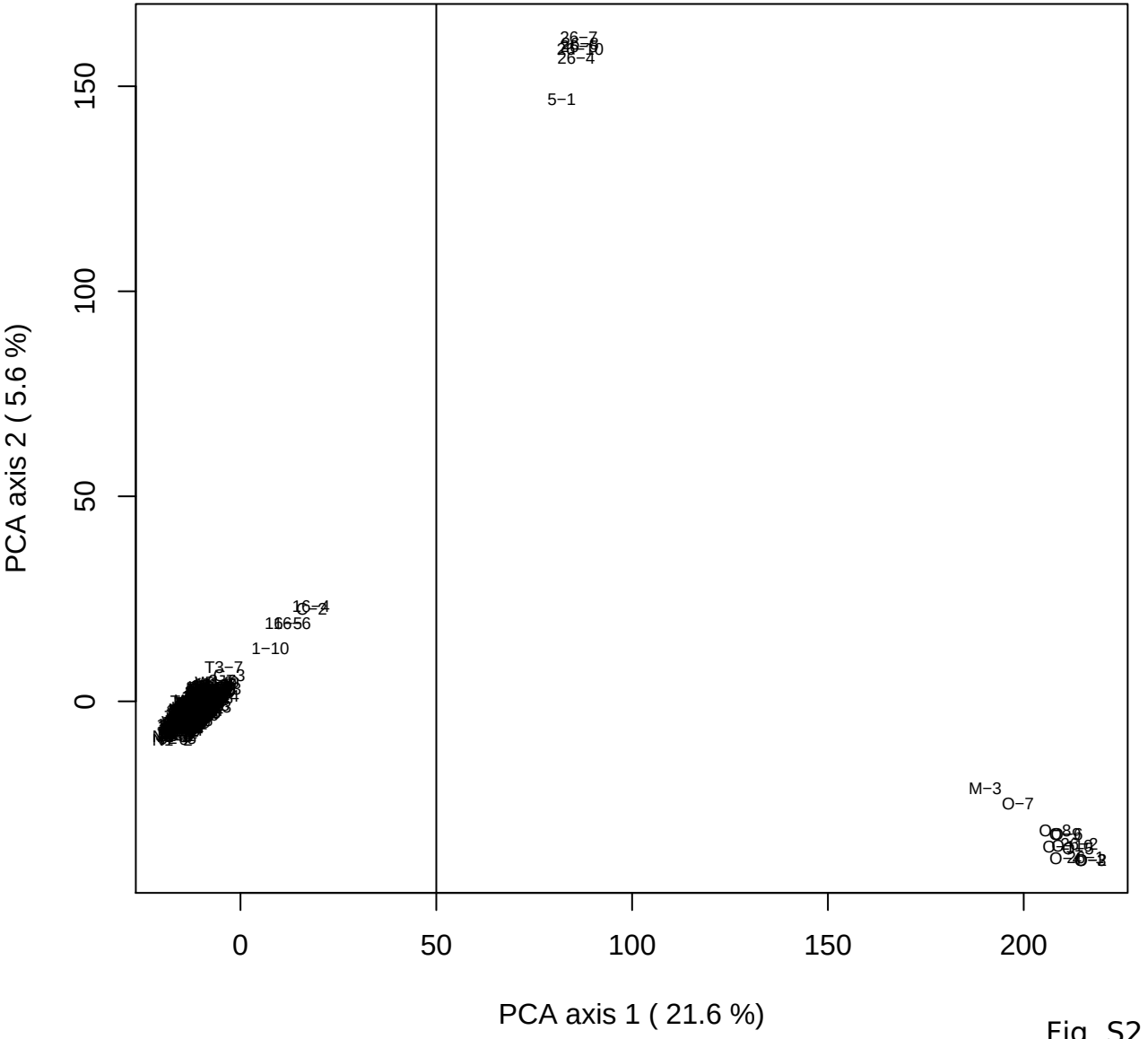

PCA axis 2 (5.6 %)
PCA axis 1 (21.6 %)
0
50
100
150
200
26−7
26−4
5−1
16−4
1−10
T3−7
M−3
O−7

Fig. S2

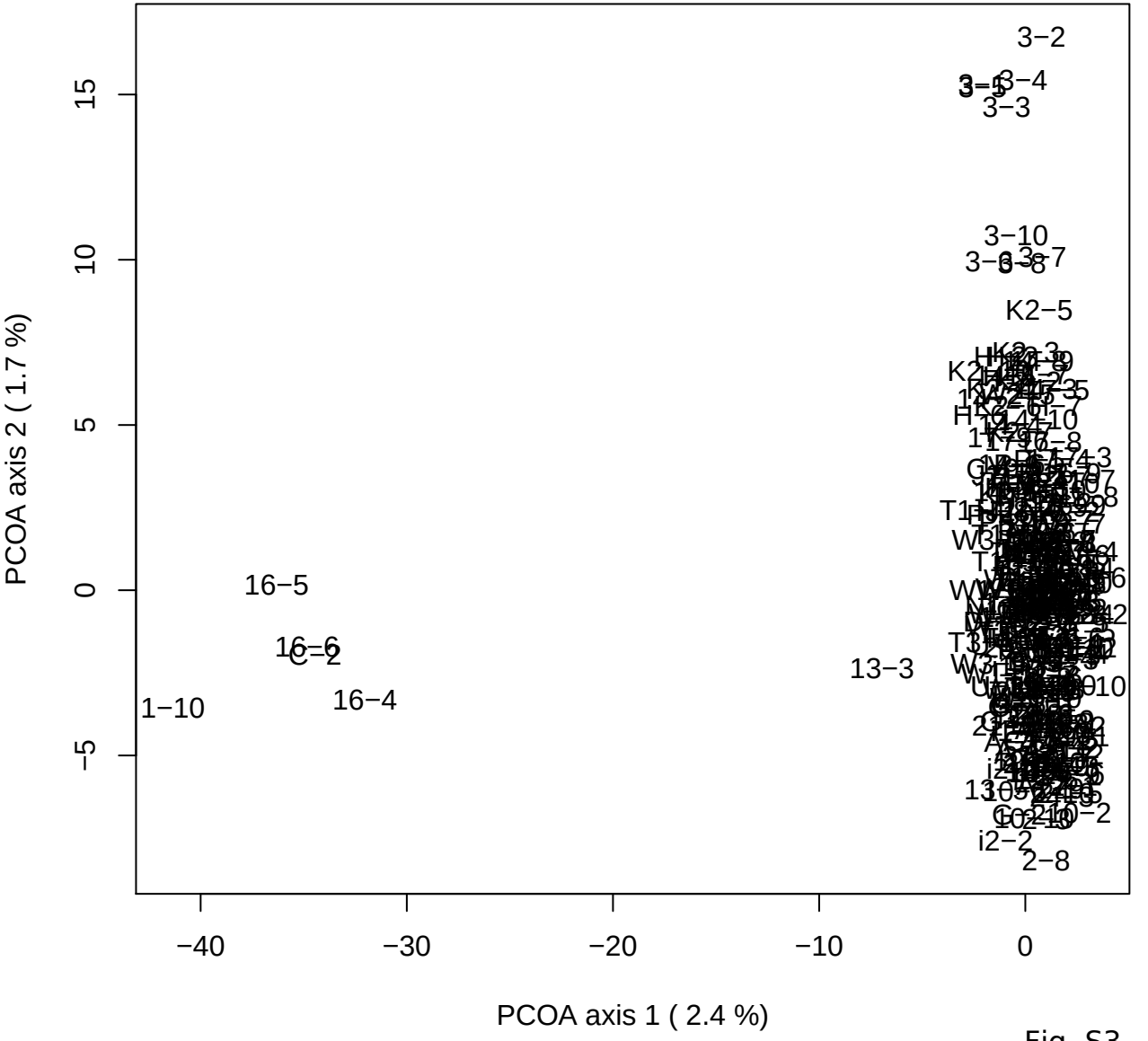

PCOA axis 2 (1.7 %)
PCOA axis 1 (2.4 %)
15
10
5
0
−5
−40
−30
−20
−10
0
3−2
3−4
3−3
3−10
3−7
K2−5
16−5
16−6
16−4
1−10
13−3
i2−2
2−8

Fig. S3

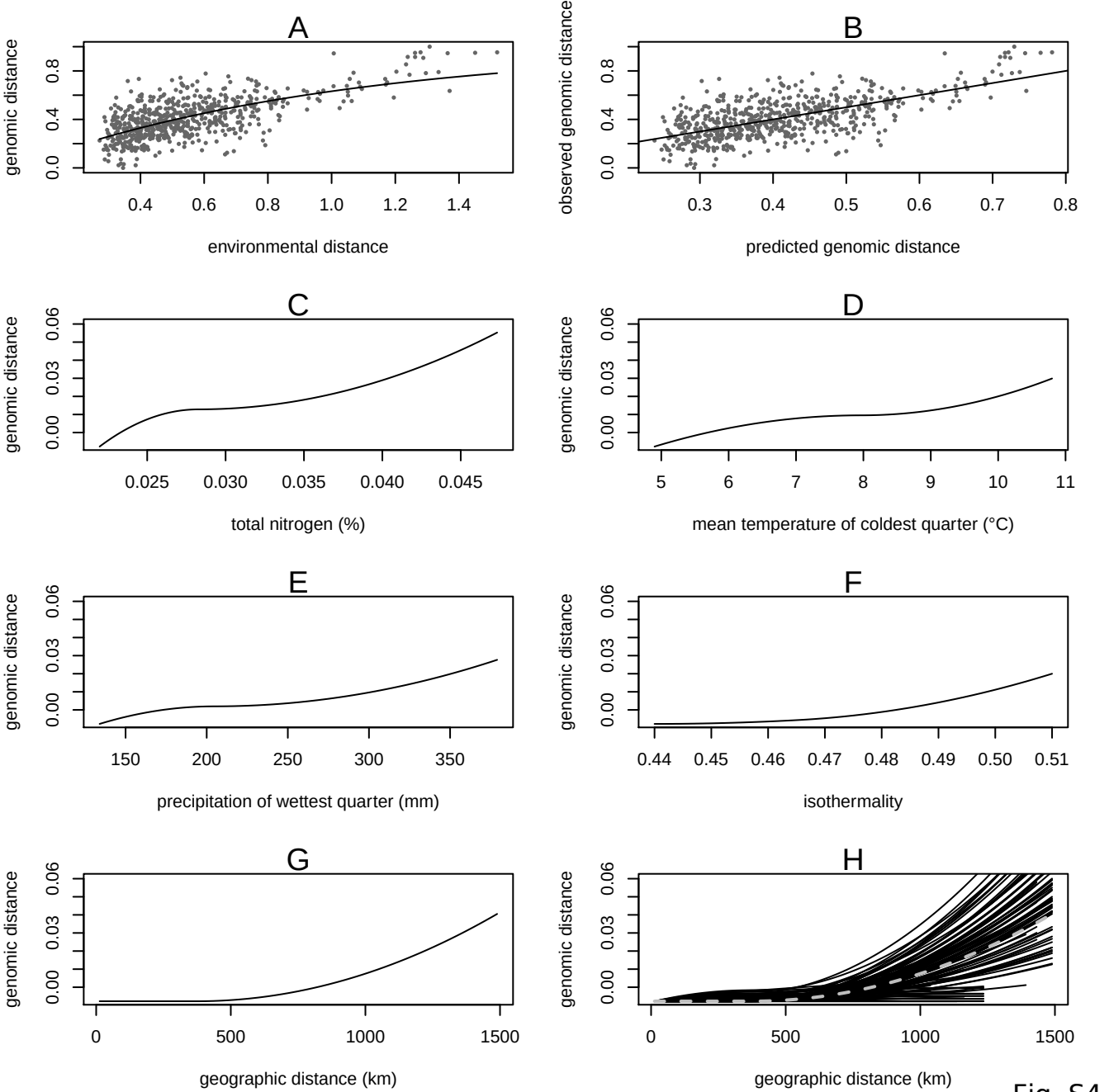
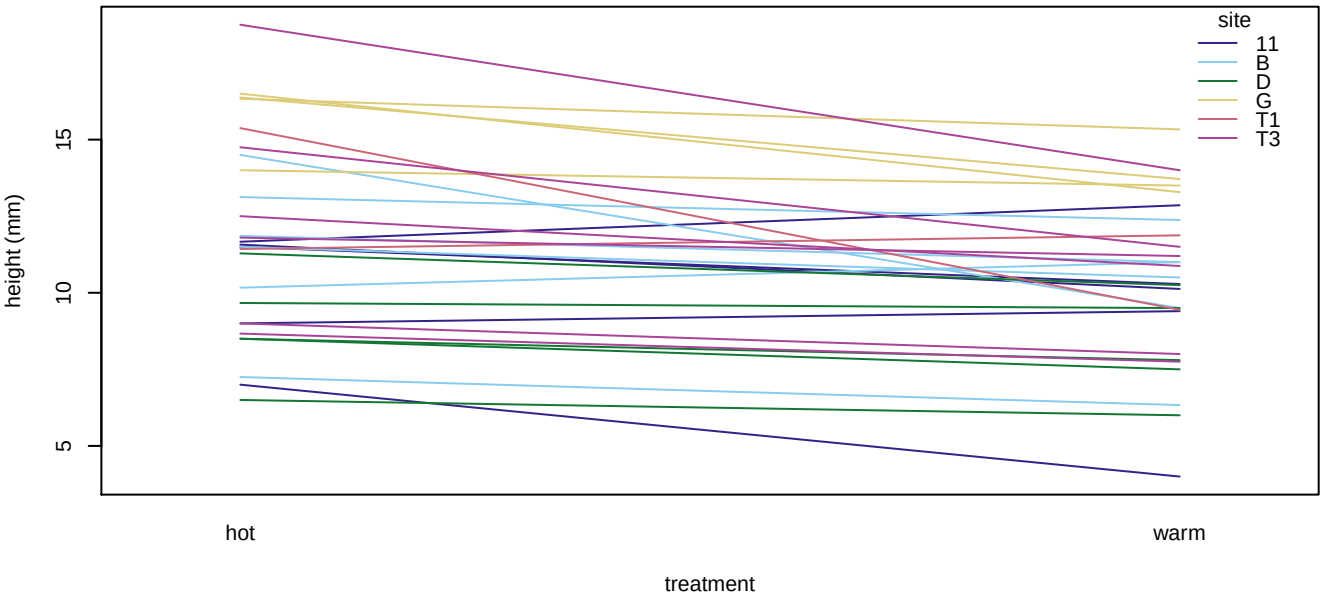
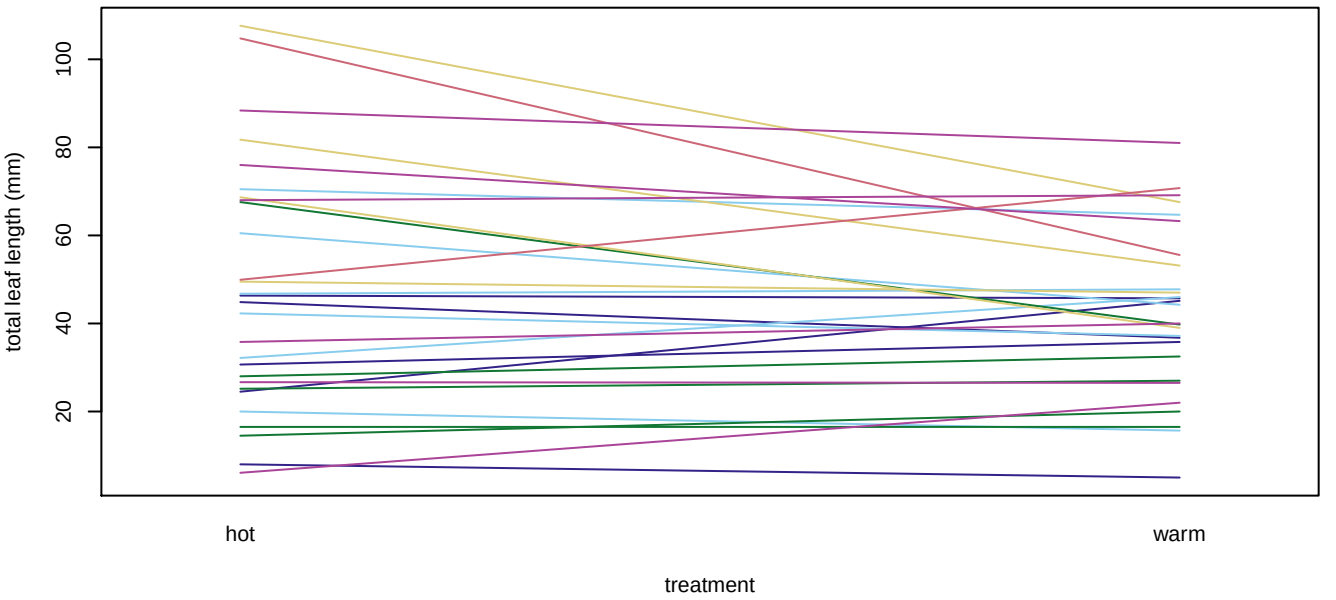

A
genomic distance
0.0
0.4
0.8
0.4
0.6
0.8
1.0
1.2
1.4
environmental distance
B
observed genomic distance
0.0
0.4
0.8
0.3
0.4
0.5
0.6
0.7
0.8
predicted genomic distance
C
genomic distance
0.00
0.03
0.06
0.025
0.030
0.035
0.040
0.045
total nitrogen (%)
D
genomic distance
0.00
0.03
0.06
5
6
7
8
9
10
11
mean temperature of coldest quarter (°C)
E
genomic distance
0.00
0.03
0.06
150
200
250
300
350
precipitation of wettest quarter (mm)
F
genomic distance
0.00
0.03
0.06
0.44
0.45
0.46
0.47
0.48
0.49
0.50
0.51
isothermality
G
genomic distance
0.00
0.03
0.06
0
500
1000
1500
geographic distance (km)
H
genomic distance
0.00
0.03
0.06
0
500
1000
1500
geographic distance (km)

Fig. S4

height (mm)
5
10
20
30
hot
warm
11
B
D
G
T1
total leaf length (mm)
0
50
150
250
relative height increment (per day)
0.00
0.04
0.08
treatment
site

Fig. S5


site
11
B
D
G
T1
T3
height (mm)
5
10
15
hot
warm
treatment


total leaf length (mm)
20
40
60
80
100
hot
warm
treatment

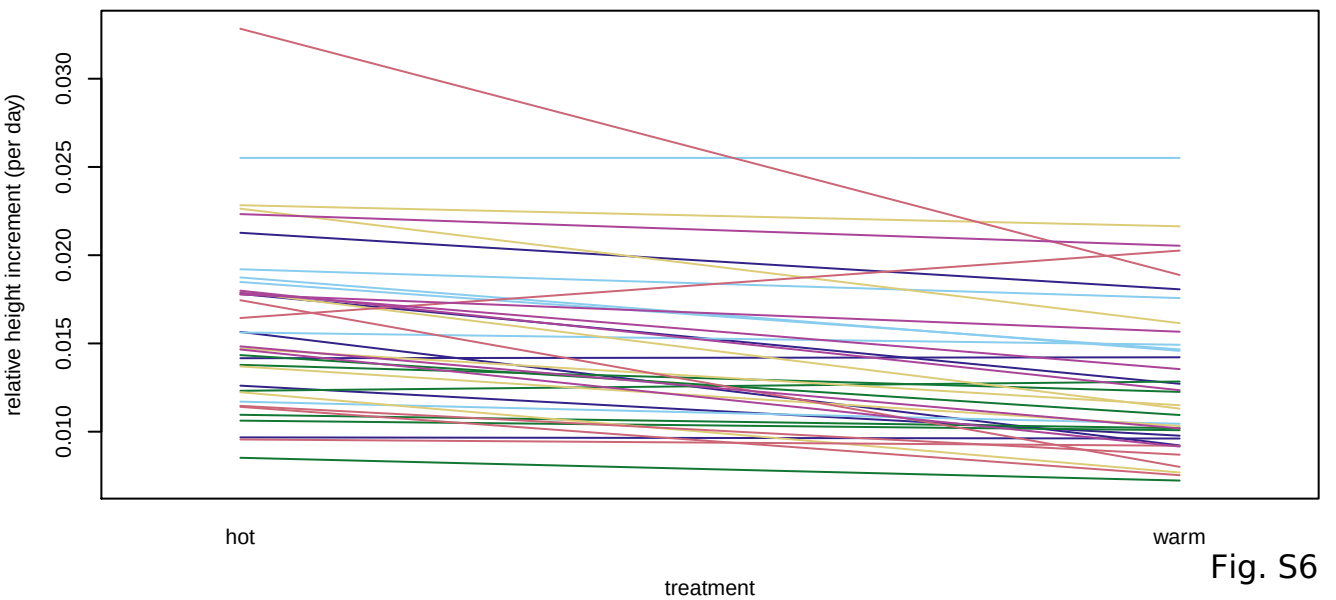
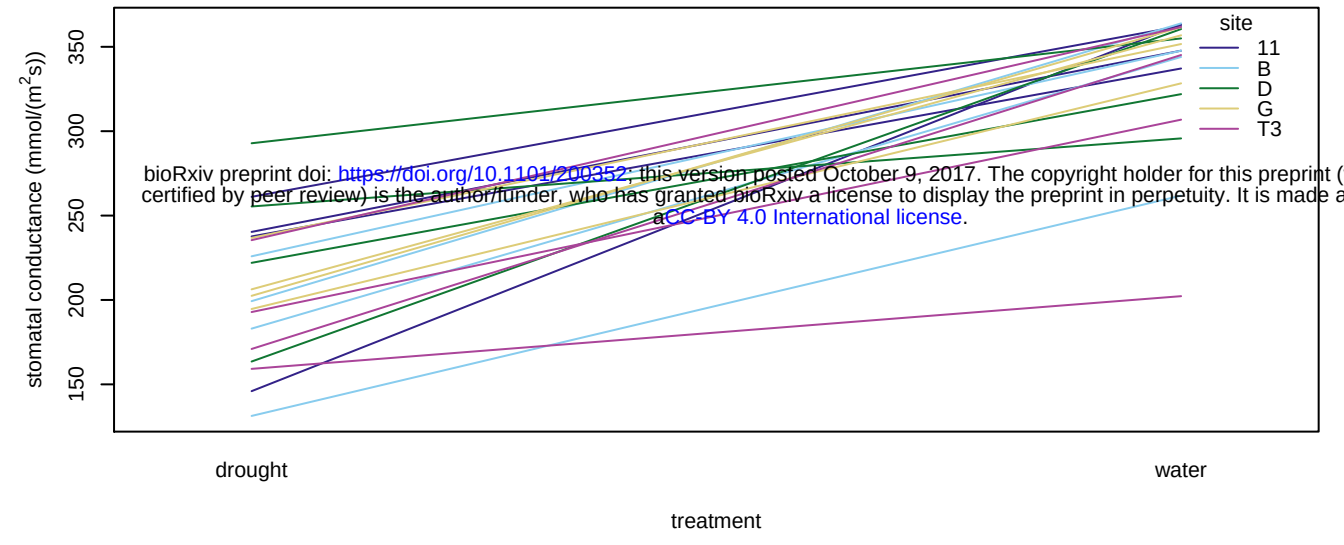
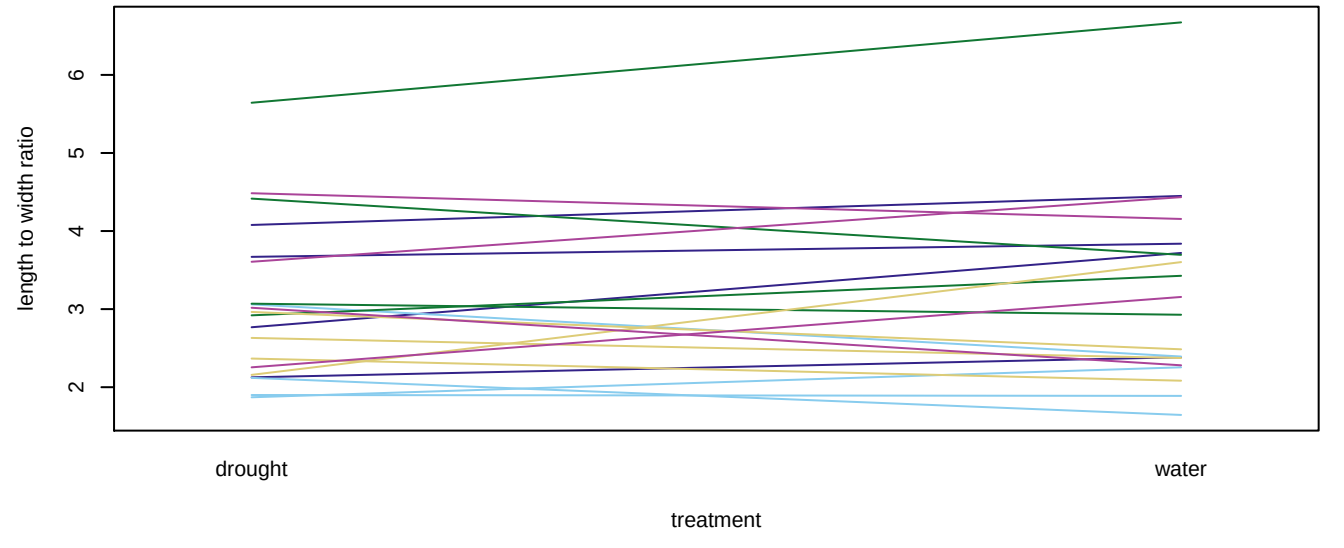

relative height increment (per day)
0.010
0.015
0.020
0.025
0.030
hot
warm
treatment

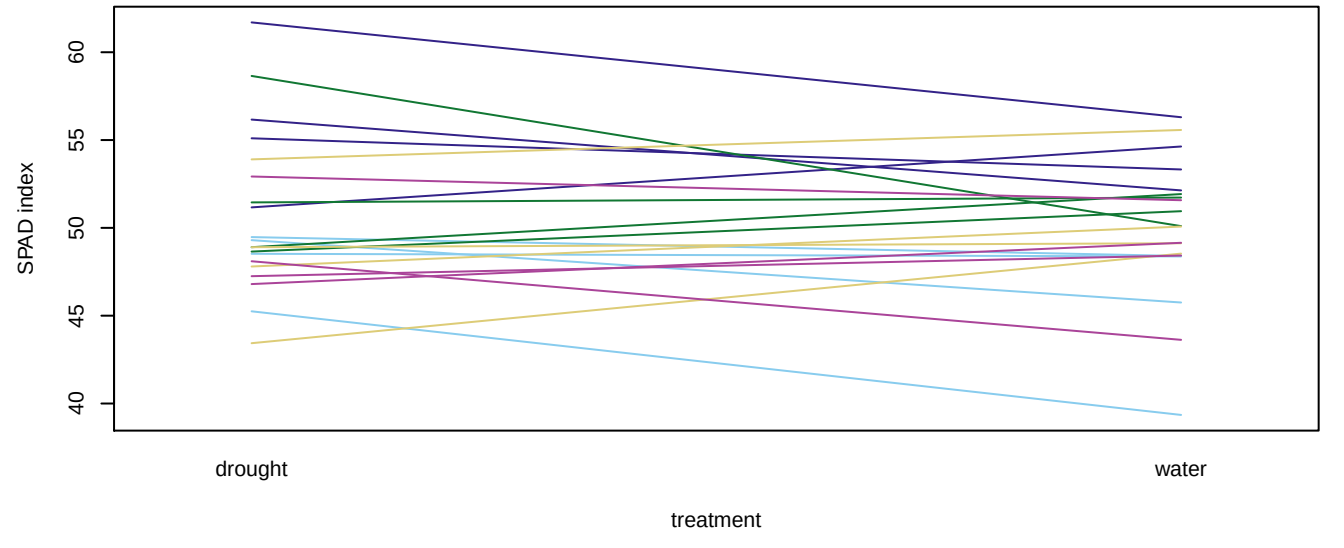
Fig. S6

stomatal conductance (mmol/(m2s))
100
200
300
400
drought
water
11
B
D
G
T3
length to width ratio
2
4
6
8
SPAD index
30
40
50
60
SLA (cm2/g)
20
60
100
140
treatment
site

Fig. S7







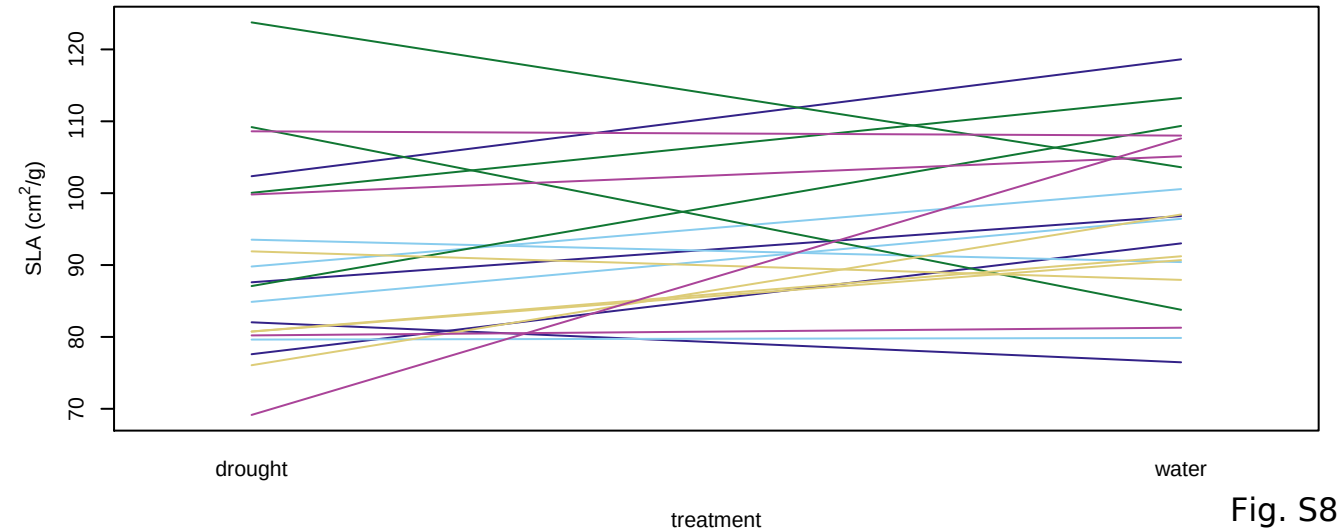


Fig. S8

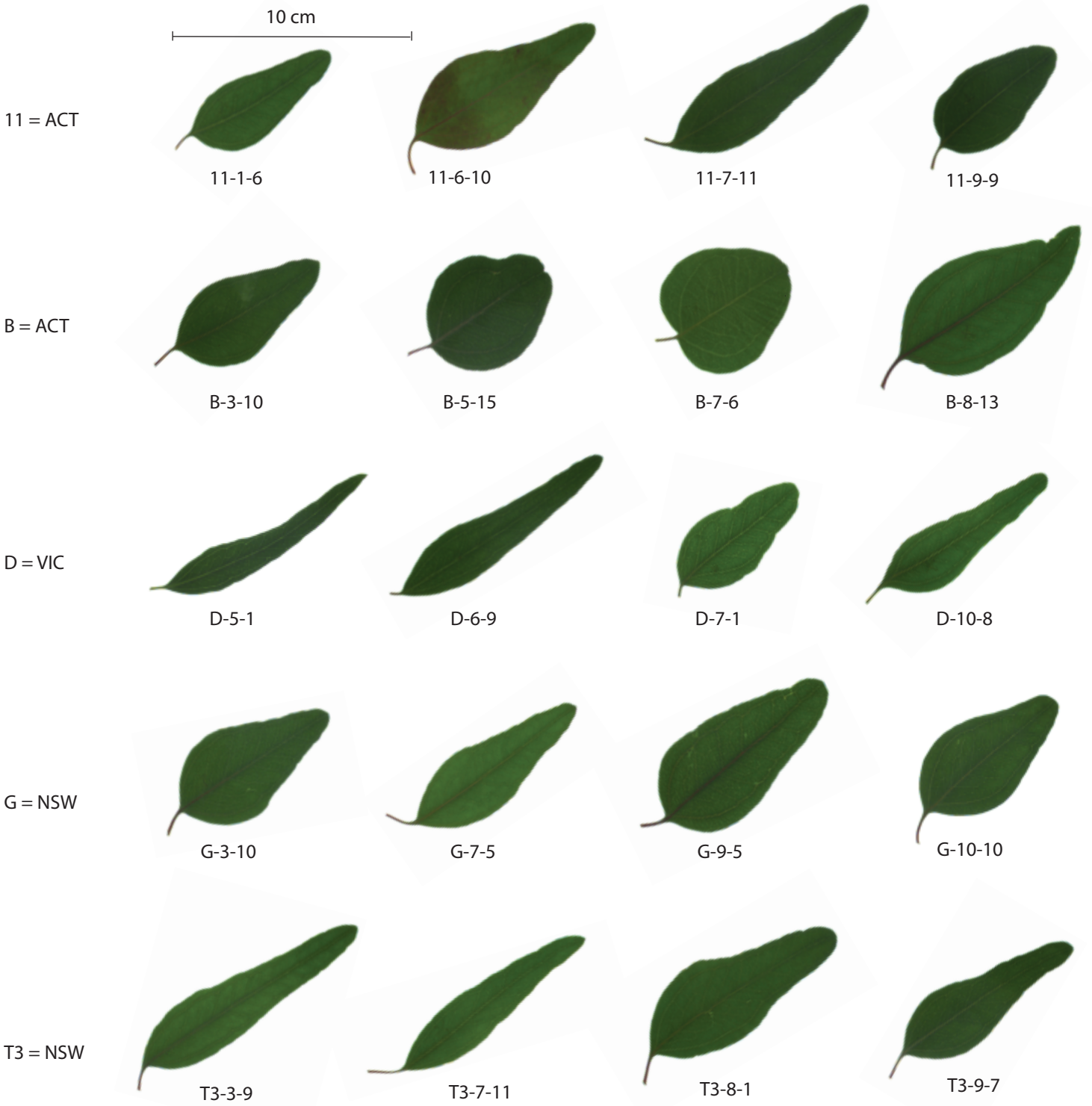

10 cm
11 = ACT
11-1-6
11-6-10
11-7-11
11-9-9
B = ACT
B-3-10
B-5-15
B-7-6
B-8-13
D = VIC
D-5-1
D-6-9
D-7-1
D-10-8
G = NSW
G-3-10
G-7-5
G-9-5
G-10-10
T3 = NSW
T3-3-9
T3-7-11
T3-8-1
T3-9-7

Fig. S9

Table S1: *E. melliodora* sampling information

site name	latitude	longitude	# sampled	# outliers*	# in final analysis	chamber experiment	drought experiment
1	-33.416668	149.55055	10	1 secondary	7		
2	-36.6755	146.26477	10		7		
3	-27.368696	152.03023	10		9		
5	-33.710145	149.52355	10	1 primary	9		
7	-33.936944	116.963333	7	7 geographic	0		
10	-37.285278	143.77778	10		8		
11	-35.40026	149.1318	10		9	yes	yes
13	-36.997778	145.474944	6		3		
14	-32.389167	150.881667	10		10		
16	-33.224724	149.28056	7	3 secondary	2		
17	-33.44218	147.51628	10		8		
19	-37.179444	144.483333	10		9		
20	-34.911785	148.971179	10		9		
21	-37.144361	142.839428	10		9		
24	-36.980278	144.047611	10		7		
26	-35.39474	149.677129	10	7 primary	2		
A	-36.388	146.481	10		10		
B	-35.235664	149.16422	10		7	yes	yes
C	-35.03333	147.33333	9	1 secondary	6		
D	-36.3675	145.70232	10		9	yes	yes
G	-36.86476	148.85016	10		9	yes	yes
H	-36.63518	149.56728	10		7		
i1	-36.75	145.583333	10		8		
i2	-36.54041	146.10739	10		7		
J2	-33.121143	149.05785	10		9		
K2	-30.54493	151.79367	10		8		
M	-34.630472	149.870892	3	1 primary	0		
N1	-34.036884	148.5675	10		10		
N2	-33.6166	149.64452	10		7		
O	-37.4279	147.8924	10	10 primary	10		
P	-35.819633	145.31187	10		9		
T1	-35.26495	147.310133	10		7	yes	
T2	-35.37515	147.25261	10		9		
T3	-35.299683	147.165733	10		10	yes	yes
U	-36.632668	144.35019	10		8		
V	-33.423	149.752	10		7		
W1	-36.42797	144.41092	10		8		
W2	-36.580471	143.521639	10		4		
W3	-36.980278	144.047611	10		8		

* primary outliers refers to outlier samples identified with the first PCA analysis
secondary outliers refers to outlier samples identified with the second PCA analysis
geographic refers to samples outside the natural distribution

Table S2: Pairwise population Fst

	10	11	13	14	16	17	19	2	20	21	24	26	3	5	A	B	C	D	G	H	i1	i2	J2	K2	N1	N2	P	T1	T2	T3	U	V	W1	W2	W3
1	0.043	0.027	0.037	0.034	-0.005	0.036	0.034	0.046	0.010	0.023	0.035	0.028	0.053	0.019	0.037	0.029	0.047	0.011	0.032	0.038	0.018	0.012	0.016	0.021	0.008	0.044	0.014	0.026	0.011	0.027	0.028	0.013	0.031	0.027	0.022
10		0.047	0.047	0.065	0.019	0.056	0.030	0.034	0.028	0.018	0.047	0.030	0.097	0.043	0.038	0.049	0.058	0.028	0.056	0.072	0.029	0.022	0.039	0.039	0.041	0.053	0.045	0.042	0.028	0.034	0.028	0.046	0.063	0.027	0.042
11			0.049	0.054	0.028	0.040	0.057	0.051	0.028	0.034	0.050	0.041	0.080	0.035	0.045	0.030	0.051	0.029	0.064	0.055	0.037	0.027	0.031	0.026	0.027	0.053	0.035	0.041	0.021	0.039	0.040	0.034	0.056	0.027	0.046
13				0.056	0.011	0.050	0.053	0.045	0.031	0.022	0.051	0.015	0.103	0.047	0.046	0.035	0.043	0.026	0.057	0.067	0.035	0.025	0.045	0.056	0.037	0.065	0.039	0.069	0.033	0.036	0.022	0.048	0.063	0.006	0.054
14					0.016	0.028	0.055	0.073	0.033	0.046	0.062	0.044	0.057	0.047	0.055	0.049	0.053	0.031	0.061	0.043	0.041	0.039	0.025	0.017	0.035	0.053	0.029	0.049	0.022	0.039	0.042	0.047	0.051	0.036	0.048
16						0.029	0.028	0.051	-0.002	0.009	0.016	0.047	0.053	0.008	0.041	0.010	0.025	-0.008	0.040	0.044	0.034	-0.004	0.000	0.030	0.004	0.021	0.003	0.034	0.000	0.009	-0.005	-0.004	0.042	0.020	0.034
17							0.055	0.063	0.038	0.040	0.056	0.037	0.062	0.043	0.048	0.043	0.041	0.020	0.068	0.050	0.031	0.035	0.023	0.019	0.033	0.050	0.022	0.031	0.018	0.040	0.029	0.031	0.057	0.031	0.042
19								0.036	0.034	0.028	0.043	0.063	0.097	0.047	0.042	0.043	0.062	0.021	0.065	0.071	0.026	0.023	0.040	0.045	0.043	0.057	0.036	0.050	0.029	0.036	0.027	0.042	0.063	0.043	0.055
2									0.024	0.025	0.048	0.071	0.097	0.040	0.032	0.049	0.061	0.024	0.053	0.073	0.033	0.031	0.034	0.047	0.044	0.067	0.052	0.051	0.031	0.026	0.027	0.054	0.068	0.051	0.048
20										0.022	0.037	0.007	0.067	0.021	0.033	0.018	0.035	0.014	0.050	0.045	0.020	0.021	0.013	0.015	0.021	0.029	0.019	0.033	0.012	0.023	0.021	0.022	0.039	0.020	0.034
21											0.032	0.028	0.079	0.024	0.028	0.030	0.041	0.018	0.045	0.055	0.013	0.026	0.022	0.032	0.019	0.044	0.021	0.029	0.016	0.025	0.013	0.026	0.044	0.023	0.024
24												0.059	0.097	0.054	0.045	0.045	0.066	0.036	0.065	0.078	0.046	0.038	0.040	0.044	0.043	0.058	0.045	0.056	0.038	0.045	0.046	0.039	0.073	0.043	0.053
26													0.067	0.027	0.059	0.027	0.047	0.010	0.044	0.057	0.036	0.008	0.009	0.022	0.022	0.049	0.008	0.054	0.013	0.034	0.026	0.026	0.053	0.029	0.048
3														0.070	0.093	0.066	0.097	0.068	0.092	0.061	0.086	0.073	0.060	0.039	0.063	0.097	0.056	0.079	0.064	0.070	0.077	0.058	0.093	0.062	0.079
5															0.049	0.029	0.045	0.019	0.057	0.053	0.029	0.036	0.022	0.027	0.030	0.026	0.028	0.032	0.016	0.034	0.028	0.035	0.047	0.028	0.042
A																0.046	0.053	0.026	0.063	0.068	0.031	0.040	0.032	0.039	0.036	0.054	0.030	0.039	0.027	0.025	0.032	0.042	0.057	0.034	0.050
B																	0.040	0.021	0.058	0.052	0.027	0.029	0.028	0.045	0.024	0.039	0.036	0.039	0.021	0.027	0.019	0.020	0.057	0.022	0.046
C																		0.034	0.074	0.055	0.041	0.036	0.040	0.048	0.040	0.047	0.036	0.052	0.026	0.037	0.039	0.042	0.052	0.037	0.057
D																			0.048	0.038	0.014	0.011	0.012	0.007	0.023	0.036	0.011	0.017	0.008	0.013	0.009	0.019	0.033	0.012	0.022
G																				0.062	0.057	0.005	0.042	0.049	0.051	0.068	0.045	0.064	0.038	0.049	0.044	0.056	0.072	0.056	0.059
H																					0.051	0.052	0.031	0.025	0.048	0.076	0.049	0.052	0.037	0.056	0.034	0.052	0.070	0.050	0.058
i1																						0.019	0.017	0.032	0.026	0.049	0.024	0.032	0.011	0.030	0.008	0.029	0.054	0.033	0.035
i2																							0.029	0.027	0.032	0.037	0.023	0.024	0.013	0.029	0.016	0.019	0.048	0.011	0.034
J2																								0.015	0.018	0.030	0.006	0.027	0.016	0.021	0.022	0.022	0.035	0.021	0.029
K2																									0.020	0.033	0.013	0.025	0.008	0.032	0.037	0.022	0.052	0.017	0.024
N1																										0.032	0.016	0.027	0.010	0.026	0.024	0.023	0.043	0.019	0.031
N2																											0.039	0.062	0.027	0.040	0.033	0.040	0.058	0.042	0.061
P																												0.025	0.010	0.021	0.011	0.023	0.035	0.022	0.017
T1																													0.009	0.033	0.021	0.037	0.048	0.039	0.053
T2																														0.014	0.016	0.021	0.029	0.000	0.019
T3																															0.019	0.033	0.047	0.016	0.033
U																																0.022	0.038	0.009	0.034
V																																	0.056	0.022	0.043
W1																																		0.044	0.071
W2																																			0.036

Table S3: Pearson's correlation between environmental variables across sampling sites

	% explained	bioclim1	bioclim2	bioclim3	bioclim4	bioclim5	bioclim6	bioclim7	bioclim8	bioclim9	bioclim10	bioclim11	bioclim12	bioclim13	bioclim14	bioclim15	bioclim16	bioclim17	bioclim18	bioclim19	elevation	surface water	depth water	surface nitrogen	depth nitrogen	surface phosphorus	depth phosphorus	Prescott Index	MrVBF
bioclim1	9.86	1.00	0.32	-0.10	0.34	0.84	0.79	0.32	0.29	0.45	0.95	0.95	-0.46	-0.25	-0.59	0.21	-0.25	-0.53	-0.15	-0.45	-0.71	-0.33	-0.47	-0.24	-0.15	0.50	0.47	-0.50	0.64
bioclim2	2.87		1.00	0.00	0.61	0.43	-0.26	0.80	0.37	-0.17	0.44	0.10	-0.22	-0.24	0.14	-0.44	-0.31	0.20	0.17	-0.56	0.10	-0.15	-0.22	-0.49	-0.53	0.10	0.01	-0.29	0.22
bioclim3	22.42			1.00	-0.76	-0.55	-0.12	-0.59	0.52	-0.56	-0.36	0.11	0.39	0.60	0.22	0.47	0.53	0.11	0.69	-0.26	0.35	0.54	0.57	0.06	-0.01	0.14	0.29	0.20	-0.39
bioclim4	12.17				1.00	0.72	-0.03	0.95	-0.10	0.31	0.62	0.03	-0.38	-0.55	-0.07	-0.61	-0.54	0.08	-0.34	-0.16	-0.19	-0.47	-0.58	-0.35	-0.35	0.03	-0.13	-0.31	0.48
bioclim5	3.49					1.00	0.64	0.67	-0.05	0.66	0.95	0.67	-0.65	-0.60	-0.56	-0.16	-0.58	-0.47	-0.52	-0.29	-0.74	-0.60	-0.68	-0.31	-0.19	0.26	0.15	-0.57	0.71
bioclim6	2.45						1.00	-0.14	-0.08	0.69	0.67	0.88	-0.47	-0.23	-0.77	0.47	-0.20	-0.76	-0.43	-0.07	-0.87	-0.34	-0.37	0.03	0.19	0.34	0.35	-0.41	0.52
bioclim7	10.82							1.00	0.01	0.18	0.58	0.02	-0.39	-0.54	0.01	-0.65	-0.55	0.13	-0.25	-0.30	-0.12	-0.44	-0.52	-0.42	-0.43	0.02	-0.14	-0.34	0.42
bioclim8	10.61								1.00	-0.64	0.17	0.29	0.08	0.27	0.02	0.11	0.16	0.05	0.60	-0.65	0.23	0.38	0.26	-0.33	-0.38	0.20	0.28	-0.14	0.13
bioclim9	1.51									1.00	0.52	0.43	-0.46	-0.44	-0.52	0.10	-0.35	-0.51	-0.70	0.26	-0.74	-0.61	-0.58	0.12	0.24	0.11	0.05	-0.29	0.37
bioclim10	3.89										1.00	0.80	-0.54	-0.43	-0.54	-0.03	-0.42	-0.44	-0.29	-0.41	-0.69	-0.47	-0.60	-0.31	-0.23	0.42	0.33	-0.54	0.71
bioclim11	17.30											1.00	-0.40	-0.12	-0.64	0.43	-0.12	-0.63	-0.11	-0.39	-0.74	-0.24	-0.33	-0.12	-0.01	0.50	0.51	-0.45	0.54
bioclim12	13.04												1.00	0.88	0.76	0.24	0.90	0.76	0.75	0.56	0.64	0.51	0.43	0.61	0.44	0.22	0.29	0.91	-0.64
bioclim13	18.37													1.00	0.46	0.61	0.98	0.44	0.80	0.37	0.54	0.56	0.47	0.55	0.41	0.29	0.42	0.79	-0.59
bioclim14	0.00														1.00	-0.36	0.45	0.97	0.60	0.34	0.72	0.35	0.35	0.29	0.13	-0.05	-0.07	0.65	-0.52
bioclim15	11.46															1.00	0.61	-0.42	0.24	0.16	-0.11	0.22	0.19	0.34	0.36	0.25	0.40	0.24	-0.19
bioclim16	22.31																1.00	0.44	0.76	0.45	0.49	0.52	0.42	0.60	0.45	0.36	0.47	0.82	-0.58
bioclim17	0.00																	1.00	0.59	0.35	0.71	0.34	0.30	0.29	0.12	0.04	0.01	0.66	-0.46
bioclim18	18.87																		1.00	-0.10	0.68	0.62	0.50	0.13	-0.05	0.31	0.42	0.51	-0.43
bioclim19	2.97																			1.00	0.09	0.02	0.06	0.76	0.75	-0.04	-0.04	0.72	-0.41
elevation	0.24																				1.00	0.63	0.61	-0.01	-0.17	-0.24	-0.12	0.48	-0.57
surface water	17.03																					1.00	0.91	0.10	0.06	0.12	0.28	0.31	-0.24
depth water	3.56																						1.00	0.10	0.10	-0.13	0.06	0.27	-0.37
surface nitrogren	6.05																							1.00	0.94	0.32	0.30	0.78	-0.38
depth nitrogen	8.56																								1.00	0.25	0.22	0.63	-0.23
surface phosphorus	24.23																									1.00	0.93	0.19	0.22
depth phosphorus	24.27																										1.00	0.18	0.14
Prescott Index	4.88																											1.00	-0.65
MrVBF	2.72																												1.00

Table S4: Percent of variation explained and p-values for non-interaction linear models for chamber experiment

	seedling height		total leaf length		relative height increment	
	% explained	p-value	% explained	p-value	% explained	p-value
sampling site	17.7	<0.00001	8.2	0.00063	1.8	<0.00001
maternal line:sampling site	10.6	0.00031	17.2	0.00001	27.6	<0.00001
experimental condition	5.4	0.00032	1.2	0.05084	8.1	<0.00001
germination chamber	0	--	0.3	--	0.8	--
block	1.1	--	1.6	--	5.1	--
residual	65.2	--	71.5	--	56.6	--

Table S5: P-values of interaction term in linear model for chamber experiment

	response variable		
	seedling height	total leaf length	relative height increment
maternal line	0.89	0.58	0.67
sampling site	0.63	0.51	0.53

Table S6: Percent of variation explained and p-values for non-interaction linear models for drought experiment

	stomatal conductance		leaf length to width ratio		SPAD		SLA	
	% explained	p-value	% explained	p-value	% explained	p-value	% explained	p-value
sampling site	0.9	0.02859	20	0.00007	19.5	0.0001	6.7	0.00864
maternal line:sampling site	5.8	0.00088	21.2	0.00131	10.2	0.03344	7.9	0.07407
experimental condition	62.3	<0.00001	0	0.49232	0	0.54236	4.4	0.02135
sample pairing	0	--	15.8	--	8.7	--	4.9	--
block	7.6	--	2.5	--	2.2	--	6.1	--
residual	23.4	--	40.5	--	59.4	--	70	--

Table S7: P-values of interaction term in linear model for drought experiment

	response variable			
	stomatal conductance	leaf length to width ratio	SPAD	SLA
maternal line	0.13	0.47	0.56	0.27
sampling site	0.51	0.78	0.31	0.82