



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Orel, HK;McLay, TGB;Murphy, DJ;Cantrill, DJ;Udovicic, F;Fahey, PS;Franklin, DC;Lewis, D;Docherty, PG;White, A;Bayly, MJ;Fowler, RM

Title:

Genetic Differentiation of Geographically Overlapping Sister Species of Eucalyptus in Northern Australia

Date:

2025-06-01

Citation:

Orel, H. K., McLay, T. G. B., Murphy, D. J., Cantrill, D. J., Udovicic, F., Fahey, P. S., Franklin, D. C., Lewis, D., Docherty, P. G., White, A., Bayly, M. J. & Fowler, R. M. (2025). Genetic Differentiation of Geographically Overlapping Sister Species of Eucalyptus in Northern Australia. *Ecology and Evolution*, 15 (6), pp.e71454-. <https://doi.org/10.1002/ece3.71454>.

Persistent Link:


<https://hdl.handle.net/11343/362385>

License:

[CC BY](#)

RESEARCH ARTICLE OPEN ACCESS

Genetic Differentiation of Geographically Overlapping Sister Species of *Eucalyptus* in Northern Australia

Harvey K. Orel¹ | Todd G. B. McLay^{1,2,3} | Daniel J. Murphy³ | David J. Cantrill³ | Frank Udovicic³ | Patrick S. Fahey^{4,5} | Donald C. Franklin⁶ | Donna Lewis^{6,7} | Philip G. Docherty⁸ | Adam White⁹ | Michael J. Bayly¹ | Rachael M. Fowler^{1,3} 

¹School of BioSciences, the University of Melbourne, Parkville, Victoria, Australia | ²National Biodiversity DNA Library, Environomics, NCMI, CSIRO, Parkville, Victoria, Australia | ³Royal Botanic Gardens Victoria, Melbourne, Victoria, Australia | ⁴Department of the Environment, Tourism, Science and Innovation, Queensland Herbarium and Biodiversity Science, Toowong, Queensland, Australia | ⁵Queensland Alliance for Agriculture and Food Innovation, the University of Queensland, St Lucia, Queensland, Australia | ⁶Research Institute for Environment & Livelihoods, Charles Darwin University, Darwin, Northwest Territories, Australia | ⁷Terrestrial Ecosystem Research Network (TERN), University of Adelaide, Adelaide, South Australia, Australia | ⁸Independent Scholar, Cable Beach, Western Australia, Australia | ⁹Australian National Herbarium, Centre for Australian National Biodiversity Research, Canberra, Australian Capital Territory, Australia

Correspondence: Rachael M. Fowler (rachael.fowler@unimelb.edu.au)

Received: 29 December 2024 | **Revised:** 21 April 2025 | **Accepted:** 7 May 2025

Funding: This work was supported by Eucalypt Australia, 59 735 884 485.

Keywords: Australia | *Eucalyptus* | *Eudessia* | introgression | phylogeography | sister speciation

ABSTRACT

In the large genus *Eucalyptus*, which dominates most of Australia's open forests and woodlands, genetic studies commonly show signs of introgression between closely related, co-occurring species. Here we assessed genetic variation in two sister species with geographically overlapping distributions in northern Australia. One species, *Eucalyptus tetradonta*, is a dominant species in many lowland savannas. It is found on gravelly red-lateritic to sandy soils, with a large distribution spanning over 2000 km east-to-west and 1000 km north-to-south, from the Kimberley region of Western Australia to northern Queensland. The other, *E. megasapala*, was taxonomically separated from *E. tetradonta* in 2006 on the basis of its larger sepals, prominently ribbed buds and operculum, and fruit and peduncle shape. It typically occurs on rocky substrates in northeast Queensland, within the range of *E. tetradonta*, where the two species can form stands within a few hundred metres of one another. Contrary to expectations, DArTseq genotyping showed strong differentiation between the two species (F_{ST} : 0.28) and little evidence of genetic admixture. Analyses of data from the Australasian Virtual Herbarium showed significant differences in flowering times between the two species. SNP outlier analyses identified multiple loci potentially under selection that are associated with differences between the two species, including the gene FT-interacting protein 7 (FTIP7), which is known to modulate flowering time in plants. Based on current data, it is unclear whether differentiation of these species is the product of parapatric/sympatric speciation or if it was allopatric with secondary geographic overlap. Within *E. tetradonta*, genetic variation showed a strong signal of isolation-by-distance, with an east–west trend in the pattern of genetic relatedness. The most substantial genetic break in *E. tetradonta* was associated with the Carpentarian Gap, a region of seasonally arid, alluvial plains on the southern margin of the Gulf of Carpentaria known as a biogeographic barrier for other Australian biota.

Harvey Orel and Todd McLay should be considered joint first author.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Ecology and Evolution* published by British Ecological Society and John Wiley & Sons Ltd.

1 | Introduction

An evolutionary pattern is emerging in some diverse, temperate tree genera, including *Quercus* (oaks), *Salix* (willows), *Pinus* (pines), and *Eucalyptus* (gum trees) (Keim et al. 1989; Hardig et al. 2000; Thorsson et al. 2007; Willyard et al. 2009; Hipp et al. 2019; Li et al. 2021; Zhang et al. 2024). These genera are characterised by a high level of species diversity and a wide range of environmental and genetic variation within species (Cannon and Petit 2020; Crowl et al. 2020; López de Heredia et al. 2020). They can have high levels of hybridisation between related species, suggesting leaky species boundaries and incomplete reproductive isolation (Cannon and Petit 2020; Buck and Flores-Renteria 2022; Ma et al. 2024). Groups of species with high levels of within-species variation and high levels of hybridisation between species have been referred to as ‘syngameons’ (Cannon and Petit 2020; Sanderson et al. 2023). Such a strategy may be an advantage in long-lived trees, as high levels of intraspecific diversity and low or weak reproductive barriers between species may provide adaptive advantages to changing climatic and environmental conditions, as the genetic entities can produce offspring that are extremely plastic (Wang et al. 2020; Cannon and Petit 2020). However, the question remains of how such a high level of species diversity can be generated and maintained when species are not reproductively isolated.

In Australia, there are over 700 species of *Eucalyptus*, which dominate many vegetation types (APC 2024; Nicolle 2024). *Eucalyptus* species are known to hybridise extensively (Pryor 1953; Griffin et al. 1988; McKinnon et al. 2004), especially those within the same subgenera, sections, or series (Larcombe et al. 2015). Inferring relationships between eucalypt species can be complex due to a combination of recent/contemporary hybridisation and plastid capture (e.g., McKinnon et al. 2004, 2010; Pollock et al. 2013, 2015; Nevill et al. 2014), and earlier phylogenetic studies were limited by the molecular tools available at the time (Bayly 2016). High-throughput sequencing technologies have allowed greater insight into the extent of hybridisation, including patterns of both plastid (Schuster et al. 2018; Fahey et al. 2021) and nuclear DNA introgression (Rutherford et al. 2018, 2019, 2021, 2023; Collins et al. 2019; Murray et al. 2019; von Takach Dukai et al. 2019; Fahey et al. 2022a, 2022b), among species that co-occur or almost so.

Speciation in syngameons and plants in general is considered to occur primarily via allopatry, where populations of a species become separated by edaphic, climatic, or physical barriers (Binks et al. 2021; Hernández-Hernández et al. 2021). This separation prevents gene flow between populations, resulting in reproductive isolation and eventually speciation and associated fixed genetic and morphological differences between species. However, if speciation does not lead to strong reproductive barriers, secondary contact between populations of those species can provide opportunities for genetic exchange and hybridisation, resulting in co-occurring species that show levels of genetic admixture (Buck et al. 2023; Cannon and Petit 2020).

Sympatric speciation (occurring without geographic barriers) is considered to be less common and more difficult to identify and

demonstrate than allopatric speciation (Fitzpatrick et al. 2008; Papadopoulos et al. 2011; Bird et al. 2012; Richards et al. 2019; Sun et al. 2022). Nonetheless, sympatric speciation has been postulated in several different organisms, predominantly in animals (Kautt et al. 2020; Sutra et al. 2019), but more evidence is emerging for sympatric speciation in plants (Osborne et al. 2019, 2020; Tavares et al. 2022). Under allopatry, divergence between populations can be promoted through factors such as genetic drift (which can be impacted by generation times and population sizes) and environmental selection, without the mitigating impact of gene flow to reconnect the lineages (Sobel et al. 2010). However, in sympatry, the mechanisms to enable speciation must promote divergence in the face of possible gene flow and therefore have to be significantly impactful over relatively shorter time periods (Wellborn and Langerhans 2015). A number of isolating mechanisms have been implicated as leading to sympatric speciation in plants, including both pre-zygotic mechanisms (e.g., phenology (Osborne et al. 2019, 2020), pollinator preference (Chen 2013), pollen-tube inhibition (Pellegrino 2016)) and post-zygotic mechanisms (e.g., chromosomal rearrangements (Wang et al. 2020), chromosome duplications and polyploidy (Tate et al. 2005), reduced germination success (de Holanda et al. 2015)). These mechanisms may act in concert to enhance and maintain reproductive isolation. In eucalypts, allopatric speciation has generally been assumed or inferred (Ladiges 1997; McGowen et al. 2001), and although parapatric differentiation of ecotypes has been inferred (Foster et al. 2007), examples of sympatric speciation have not been substantiated.

Eucalyptus subg. *Eudesmia* is a clade of 22 species that occurs across large parts of western, central, and northern Australia. A high level of introgression and hybridisation has been detected using phylogenomic analyses (McLay et al. 2023), suggesting that gene flow between species is common within this subgenus. The centre of phylogenetic diversity for *E.* subg. *Eudesmia* is in northern Queensland, although species diversity is highest in Western Australia (McLay et al. 2023). The most widespread species in the subgenus is *E. tetradonta*, which occurs in Western Australia, the Northern Territory, and Queensland in the monsoonal tropics of northern Australia (Figure 1). It occurs on a range of soil types including red bauxitic laterites, red earths, infertile sand plains, and sand on sandstone, in pure stands or mixed with a wide variety of other species (Fox et al. 2001; Bean 2006; Franklin 2022). In 2006, some Queensland populations previously included in *E. tetradonta* were described as a new species, *E. megasepala* (Bean 2006). Morphologically, the two species are similar, with *E. megasepala* being differentiated from *E. tetradonta* by larger sepals, prominently ribbed buds, conical opercula, strongly flattened peduncles, and fruits that are square in cross-section (Bean 2006). Additionally, in the region in which their ranges overlap, the two species display distinct fine-scale ecological preferences, with *E. megasepala* often occurring on rocky rises, and *E. tetradonta* tending to occur on lower slopes. Despite these differences, the range of *E. megasepala* is entirely nested within the range of *E. tetradonta*, and plants of the two species can be found within metres of one another (RMF, personal observation) in discrete, adjacent stands (Franklin 2022). Our recent phylogenomic analysis (McLay et al. 2023) confirmed that *E. megasepala* and *E. tetradonta* formed a distinct clade within *E.* subg. *Eudesmia* (taxonomically

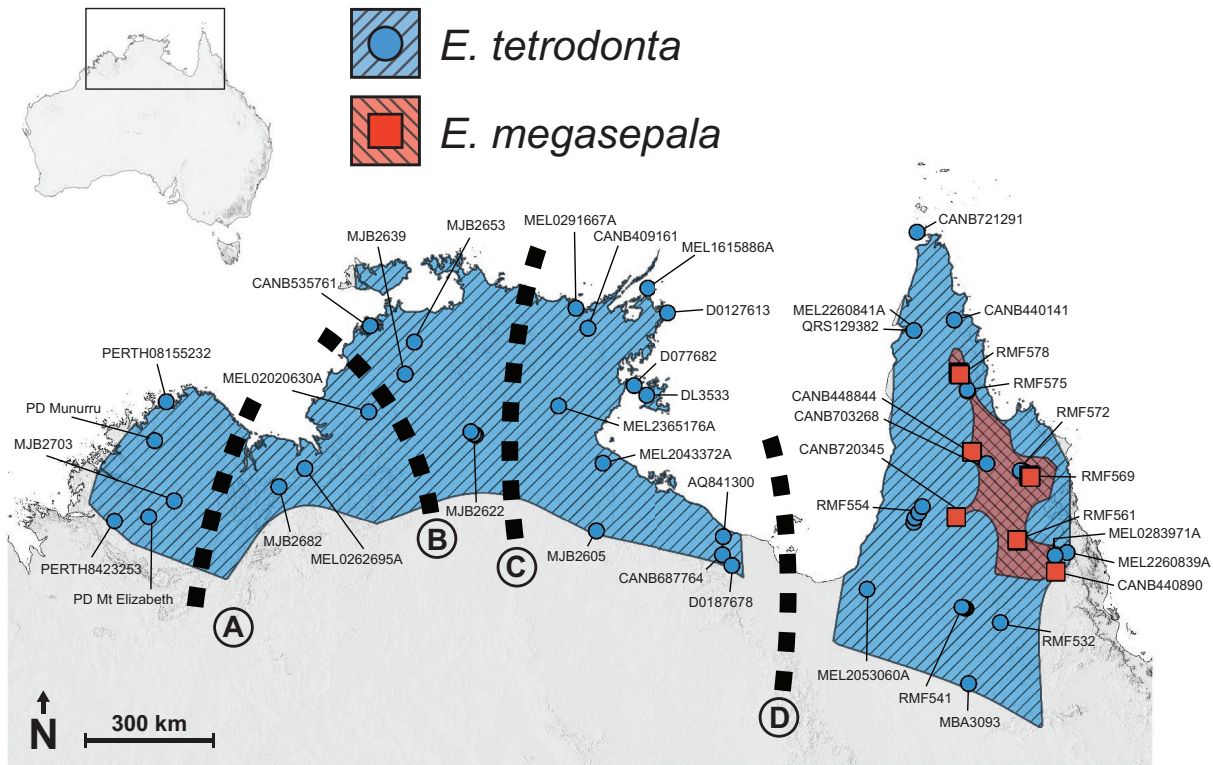


FIGURE 1 | Geographic ranges of *Eucalyptus tetrodonta* and *E. megasepala* (diagonally-lined regions), including collection sites (circles or squares) for samples included in this study. Thick dashed lines indicate notable biogeographic barriers identified by Edwards et al. (2017): (A) Ord Basin; (B) Daly River Plains; (C) Arnhem Land Plateau; (D) Carpentarian Gap. Photographs compare habit and fruit morphology of each species. Voucher specimen numbers or collector numbers are indicated for all collecting sites (details in Table S1). Photos: F. Udovicic and M. Bayly (*E. tetrodonta*); D. Nicolle (*E. megasepala*; CC BY-NC 4.0, <https://www.inaturalist.org/observations/96957531>).

recognised at the level of series; Bayly et al. 2023), but resolved *E. megasepala* nested among Queensland samples of *E. tetrodonta*, questioning the distinctiveness of the more recently described species.

To investigate the relationship between the geographically overlapping sister species, *E. megasepala* and *E. tetrodonta*, and the genetic structure of *E. tetrodonta* across the entire species' range, we performed Diversity Arrays Technology

sequencing (DARtseq) on 94 samples. DARtseq is a genome reduction method that allows for the detection of a high number of informative SNPs from across the genome (Melville et al. 2017). DARtseq has been widely applied across many plant systems for studies of genetic diversity, population structure, hybridization, speciation, phylogeography, etc. (Steane et al. 2015; Edet et al. 2018; Rutherford et al. 2018; Rutherford 2020; Fahey et al. 2022a; Rosser et al. 2023; Palssson et al. 2024; Vašut et al. 2024). Contrary to our expectations, tests of diversity and differentiation identified strong genetic distinction between the two species and low evidence of gene flow. This result was investigated more thoroughly using genetic outlier analysis to identify SNPs that strongly differentiated the two species (Ahrens et al. 2018; Kidner et al. 2021) and ecological data in the form of broad-scale environmental variables and flowering time data collected from herbarium specimen metadata. We also detected very low genetic structure in *E. tetradonta*, despite this species occurring in a wide geographic range across several potential biogeographic barriers (Edwards et al. 2017). Here we identify high genetic differentiation between co-occurring *Eucalyptus* species that are strongly supported as sister taxa.

2 | Methods

2.1 | Study Species and Sampling

Eucalyptus tetradonta and *E. megasepala* are two sister species (McLay et al. 2023) with geographic distributions in northern Australia. *Eucalyptus tetradonta* is common and widespread, spanning over 1000km north-to-south and over 2000km east-to-west from the Kimberley region of Western Australia to the Cape York Peninsula in Queensland (Figure 1). It occurs across a number of biogeographic barriers in northern Australia, including the Carpentarian Gap, around which there is a disjunction in its distribution (see Figure 1). *Eucalyptus megasepala* is restricted in distribution to far northeastern Queensland, where it overlaps in distribution with *E. tetradonta* on Cape York Peninsula. Both species are upright, fibrous-barked trees with lanceolate to falcate, opposite to sub-opposite adult leaves, and the two species are hard to distinguish if buds, flowers, or fruits (capsules) are lacking (Slee et al. 2020). However, there are several key differences between the two species. *Eucalyptus tetradonta* has capsules that are round in cross-section with four small teeth (remnant sepals), flower buds with a dome-shaped cap and at most faint ribbing, and sepals that are 1–3 mm long; it flowers from July to October. In contrast, *E. megasepala* has capsules that are square in cross-section with very prominent ribs, flower buds with a cap that tapers to a blunt tip and very prominent ribbing, and sepals that are 5–9 mm long; it flowers from April to July. In addition, *E. tetradonta* and *E. megasepala* tend to occupy different habitats in their geographic area of overlap; *E. tetradonta* is more common in more fertile downslopes environments, while *E. megasepala* is more common in less fertile soils on upslopes environments. However, the two species occur within metres of one another in some parts of their range, especially on the extensive Laura sandstone (Franklin 2022).

For this study, leaf material was collected from 94 plant specimens, including 73 *E. tetradonta* samples, 18 *E. megasepala*

samples, and one each of three outgroup species (*E. chartaboma*, *E. eudesmioides*, *E. similis*—see Table S1), representing distinct clades of *E.* subg. *Eudesmia* (McLay et al. 2023). Sampling included both field and herbarium collections. Where possible, population-level sampling was completed across the range of each species (10 populations of *E. tetradonta* and three populations of *E. megasepala*). Most population samples consisted of five individuals, collected approximately 1 km apart to reduce the likelihood of sibling relatedness between plants, or of stems being genetically identical root suckers; *E. tetradonta* being one of only two eucalypt species known to produce root suckers (Lacey and Whelan 1976; Franklin et al. 2010). Leaf tissue was provided to Diversity Arrays Technology Pty Ltd. (Canberra, Australia) for DNA extraction and sequencing using the DARtseq platform. Genotype data were generated from total genomic DNA using proprietary analytical pipelines (DARt Pty Ltd).

2.2 | Data Filtering

Following sequencing, two sets of genotype data were requested from DARt; one dataset including all samples ('total data'), and one dataset with outgroup samples excluded ('no outgroup data'). This format was specified to maximise the number of ingroup-specific SNPs retained during DARt pipeline filtering for analyses that did not require outgroups. Of the 94 and 91 samples initially included in each of these respective datasets (total data/no outgroup data), five herbarium samples failed sequencing (*E. tetradonta*—MBA3093.2, MEL0283971A, MEL1615886A, CANB 721291; *E. megasepala*—CANB 440890) and were excluded from the rest of the study.

The datasets received from DARt consisted of 44,270 SNPs (total data) and 31,371 SNPs (no outgroup data). Both datasets were subject to SNP filtering in the R package *dartR* v.2.9.7 (Mijangos et al. 2022) (see Figure 2 for a flowchart detailing the filtering process). Both datasets were initially screened for samples with high levels (> 70%) of missing data and eight ingroup samples (CANB720345, D0077682, D0187678, MEL2020630A, MEL2043372A, MEL2260840A, MEL2260841A, RMF569.5) were removed from both datasets resulting in a 'total data' set of 81 samples (Figure 2; Dataset 1) and a 'no outgroup data' set of 78 samples (Figure 2; Dataset 2). Both datasets were then filtered to remove instances of more than one SNP identified from a single locus (i.e., potentially linked SNPs) using the function 'gl.filter.secondarys', and potentially duplicated or paralogous loci using 'gl.filter.hamming' (both with default parameters).

For further analyses of relationships between species and populations, the 'no outgroup data' (Figure 2; Dataset 2) set was subjected to further filtering steps, retaining SNPs with a read depth between 5 and 50, and reproducibility (RepAvg) value ≥ 0.99 . We then filtered the data using two different call rate (CR: the proportion of samples with no missing data for a SNP) thresholds for conducting separate population genetic analyses capable of handling different levels of missing data. For this, we used thresholds of CR ≥ 0.80 and CR ≥ 0.95 .

The data produced from the CR ≥ 0.80 filtering were used to form two datasets, Dataset 3.1 and Dataset 3.2 (Figure 2), for input into PCA, admixture, and isolation-by-distance analyses.

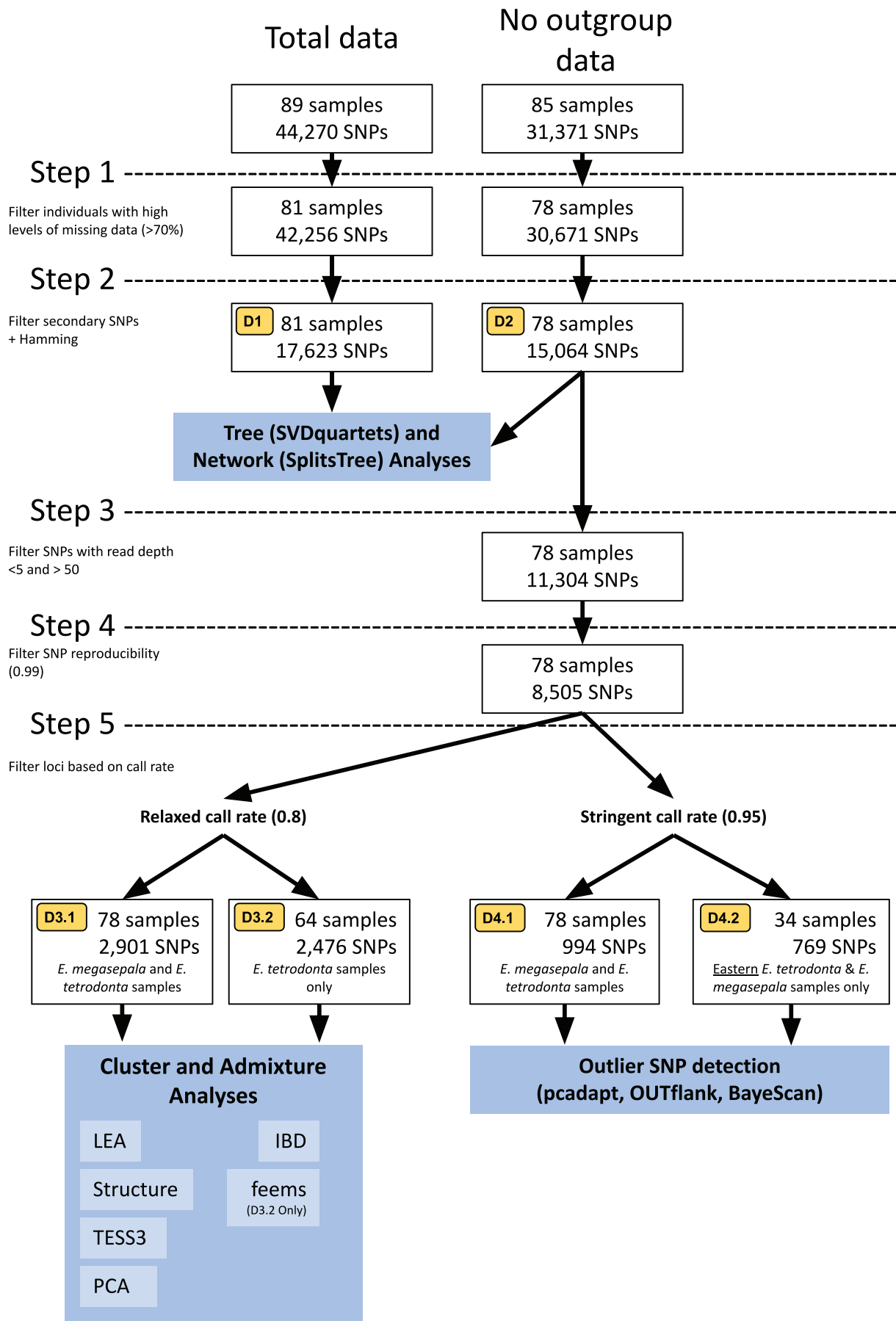


FIGURE 2 | Flowchart of data processing steps with resultant datasets (yellow boxes, e.g., D3.2=Dataset 3.2) and methods of analysis (blue boxes).

Dataset 3.1 included all samples of *E. megasepala* and *E. tetradonta*; we also removed all *E. megasepala* samples from Dataset 3.1 (also recalculating locus statistics and removing monomorphic loci) to produce Dataset 3.2, which included only samples of *E. tetradonta*.

The data produced from the $CR \geq 0.95$ filtering was used to form two datasets, Dataset 4.1 and Dataset 4.2 (Figure 2), for input to tests aimed at detecting SNPs under divergent selection between *E. megasepala* and *E. tetradonta*. Dataset 4.1 included all samples of *E. megasepala* and *E. tetradonta*, and Dataset 4.2 included only samples of *E. tetradonta* collected from a geographic area that could be considered overlapping with that of *E. megasepala* (those at most 222 km from any *E. megasepala* collections recorded on AVH 2023, and cleaned per the *Analysis of flowering time* section).

2.3 | Phylogenetic Analyses

We undertook phylogenetic analyses of the ‘total data’ set (Figure 2; Dataset 1) and the ‘no outgroup data’ set (Figure 2; Dataset 2) using tree and network approaches. For both datasets, we ran SVDquartets (Chifman and Kubatko 2014) in PAUP* v.4.0a (Swofford 2002) to produce phylogenetic trees using a site-based coalescent approach. For this, SNP matrices were formatted to a single sequence per sample, with ambiguity codes used for heterozygous sites, and samples were treated as individuals (i.e., not grouped by species or population). We exhaustively sampled all quartets, using Quartet FM (Reaz et al. 2014) for quartet assembly, and inferred node support with 100 non-parametric bootstrap replicates. To generate phylogenetic networks, we imported the SNP matrices into SplitsTree4 (Huson and Bryant 2006) and viewed the network under default import settings.

2.4 | Ordinations, Diversity Measures, Admixture and Isolation-By-Distance

To investigate patterns of genetic variation and structure in *E. megasepala* and *E. tetradonta*, we ran several analyses using Dataset 3.1 and Dataset 3.2 (Figure 2). For each dataset, we carried out successive k -means clustering using the ‘find.clusters’ function of *adeget* v.2.1.10 (Jombart 2008), testing values of k from 1 to 40, retaining 100 principal components and evaluating goodness-of-fit for each k -value using the Bayesian Information Criterion (BIC). We also ran Pearson principal component analyses (PCAs) on each dataset using the ‘gl.pcoa’ function in *dartR* with default settings. To further examine genetic diversity and differentiation between species and populations of *E. megasepala* and *E. tetradonta*, we calculated unbiased expected heterozygosity (uH_e), a metric that accounts for variation in sample size between populations, from Dataset 3.1 using the ‘gl.report.heterozygosity’ function in *dartR*. We also calculated F_{ST} estimates between the species and populations (excluding singleton specimens from herbarium collections) from Dataset 3.1 using the ‘gl.fst.pop’ function and 100 bootstrap replicates in *dartR*.

Genetic structure in Dataset 3.1 and Dataset 3.2 was first examined using the R package LEA v.3.8.0 (Frichot and

François 2015) to estimate admixture coefficients for each sample using sparse nonnegative matrix factorisation (snmf) with the function ‘snmf’. We ran snmf for values of k from 1 to 10, with 40 replicate runs per k -value, and estimated cross-entropy. The optimal value of k for each dataset was identified by the interpretation of cross-entropy scores, with the value of k with the lowest cross-entropy score considered optimal. To supplement our assessment of admixture with LEA and test for consistent results across different methods, we performed additional analyses aimed at detecting population structure and admixture using the program Structure v.2.3.4 (Pritchard et al. 2000). In Structure, we conducted 20 replicate runs for each value of k from 1 to 7 using 300,000 MCMC cycles with a burn-in of 100,000 cycles. We used Structure Harvester v.0.6.94 (Earl and vonHoldt 2012) to investigate optimal values of k , following the method of Evanno et al. (2005), and runs were combined in CLUMPAK (Kopelman et al. 2015) to obtain ancestry proportions for major clusters of each k -value. Ancestry proportions for $k = 2-7$ were then visualised as bar plots. The non-spatially explicit models employed by LEA snmf and Structure are known to be susceptible to producing inaccurate results under IBD, usually in the form of detecting additional artificial genetic clusters (Frantz et al. 2009; Meirmans 2012; Perez et al. 2018). To determine whether IBD was present in our Datasets 3.1 and 3.2, we conducted Mantel tests and generated IBD plots for each dataset using the ‘gl.ibd’ function of *dartR* with 9999 permutations to calculate pairwise genetic and geographic distances between individual samples using Euclidean distance.

Following the detection of IBD in *E. tetradonta* (Dataset 3.2), we ran *FEEMS* (Marcus et al. 2021) to estimate effective migration surfaces and visualise spatially heterogeneous IBD across the range of *E. tetradonta*. For this, we converted Dataset 3.2 to the PLINK format with PLINK v.1.90b7 (Purcell et al. 2007; Chang et al. 2015) and the ‘gl2plink’ function in *dartR*. We used the R package *dggridR* v.3.0.0 (Barnes et al. 2017) to generate two triangular grids covering the species’ distribution with differing cell sizes of 389 km² and 1557 km². We conducted *FEEMS* runs using both grids, performing leave-one-out cross-validation over a grid of tuning parameter values (λ ; which controls the strength of the penalisation placed on the output of the migration surface) to select the optimal λ value for each run.

Finally, we used the R package *triangulaR* v.0.0.0.9000 (Wiens and Colella 2024) to calculate hybrid indices and interclass heterozygosity between populations of *E. megasepala* and *E. tetradonta* from Dataset 4.1, testing five allele frequency difference thresholds (0.5, 0.6, 0.7, 0.8, 0.9). Samples from herbarium vouchers for which a single specimen was sampled at a locality were treated as separate, singleton populations.

2.5 | Detection of Outlier Loci

To identify loci potentially under selection that may be associated with biological differences between *E. megasepala* and *E. tetradonta*, we used three different programs for detecting outlier SNPs in Dataset 4.1 and Dataset 4.2. *Pcadapt* v.4.3.3 identifies F_{ST} outliers by ascertaining population structure using PCA and

then determining associations between principal components and locus variation for each SNP, with the underlying assumption that loci experiencing divergent selection exhibit non-typical relationships to principal components (Luu et al. 2017; Privé et al. 2020). We ran *pcadapt* using Mahalanobis distance to compute *p*-values for SNPs with minor allele frequencies > 5% for 20 principal components (PCs). The optimal number of PCs was identified from an interpretation of the scree plot using Cattell's Rule (Cattell 1966). The R package *qvalue* v.2.30.0 (Storey et al. 2019) was then used to transform *p*-values to *q*-values, and SNPs with a false discovery rate (FDR) *q* < 0.05 were identified as outliers.

OutFLANK v.0.2 (Whitlock and Lotterhos 2015) calculates an F_{ST} measure of genetic differentiation for each locus and estimates a null χ^2 distribution based on the neutral loci, then uses this as the null distribution to identify loci under selection (i.e., in the right tail) of the full χ^2 distribution. We ran *OutFLANK* through *dartR* using the 'gl.outflank' function with default settings.

BayeScan v.2.1 (Foll and Gaggiotti 2008) estimates F_{ST} coefficients for two components, a population-specific component (β) and a locus-specific component (α), and considers loci to have departed from neutrality when the locus-specific component is necessary to explain the observed pattern of diversity (i.e., when α is significantly different from zero). Positive values of α suggest divergent selection, while negative values suggest balancing selection. A reversible-jump MCMC algorithm is then used to estimate the posterior probability of a neutral selection model (that does not include the α component) and a selection model (that includes the α component) for each locus. To run *BayeScan*, genlight objects for each dataset were converted to *BayeScan* format in *dartR* using the 'gl2bayescan' function, with samples assigned to populations based on their species. We then ran *BayeScan* twice using different values for prior odds; this parameter sets the prior odds for the neutral model, where a value of 1 is equivalent to assuming that for each locus the neutral model is equally as likely as the model including selection, and increasing this value increases the likelihood of the neutral model. Higher prior odds values tend to eliminate false positives at the expense of reducing the power to detect markers under selection, and the prior odds value has a decreasing influence with increasing dataset size. For the first run, we used the default prior odds value of 10 (recommended for hundreds of markers) with 10,000 iterations after a burn-in of 200,000 iterations. For the second run, we used a more relaxed prior odds value of 2, with 10,000 iterations after a burn-in of 200,000 iterations. For both runs, all other settings were left at their defaults, and any SNPs with *q*-values < 0.2 (corresponding to allowing 10% of false positives) were treated as outliers.

2.6 | Annotation of Outlier Loci

To investigate the identities of outlier loci and their potential relation to gene function, outlier SNP information (including SNP position, minor allele frequency, and other metadata) was exported to a vcf file from *dartR*. The outlier vcf files were imported into Geneious Prime 2023.2.1 (<http://www.geneious.com/>) and mapped to the annotated *E. grandis* genome (Myburg et al. 2014); loci that did not map to this reference genome during the initial

DarT mapping procedure were excluded from further analyses. In Geneious, we extracted 750bp of flanking sequence from both sides of the SNP to create contigs 1501 bp long. From the annotations in each contig, we recorded the SNP position (including whether the SNP was inside or outside a coding sequence (CDS) with the number of bases from the nearest CDS, and if inside a CDS, whether or not the SNP occurred inside an exon or intron), and *E. grandis* Gene ID, Protein ID, and Product. To associate gene ontology (GO) annotation terms to outlier SNPs using the Gene Ontology database (Carbon et al. 2009; Gene Ontology Consortium 2004), *E. grandis* Gene IDs were searched in Batch Entrez (<https://www.ncbi.nlm.nih.gov/sites/batchentrez>) to obtain gene names searchable on the UniProt ID mapping portal (<https://www.uniprot.org/id-mapping>), whereafter we used the results returned by this portal to download GO IDs. REVIGO (<http://revigo.irb.hr>) was then used to summarise and visualise the effects of GO terms across all outlier SNPs.

2.7 | Environmental Variables

To compare the overlap in environmental conditions between the two species, we used all 19 bioclimatic variables from WorldClim 2 (Fick and Hijmans 2017) across the entire range of *E. tetradonta* and eight soil attributes (AWC, available water capacity; BDW, bulk density; CLY, clay; ECE, effective cation exchange capacity; NTO, total nitrogen; pHc, pH CaCl₂; PTO, total phosphorus; and SND, sand) from the Soil and Landscape Grid of Australia (Rossel et al. 2015). The soil attributes were limited to the range of *E. megasepala* and were downloaded using package *slga* (<https://github.com/obrl-soil/slga>, accessed October 2023). Specimen data were downloaded for both species from the Australasian Virtual Herbarium (AVH; <https://avh.chah.org.au/>; accessed October 2023), and data were cleaned manually to remove records with erroneous data. Environmental variables were appended to the specimen data based on spatial coordinates using the packages *raster* (<https://cran.r-project.org/web/packages/raster/index.html>, accessed October 2023) and *geodata* (<https://cran.r-project.org/web/packages/geodata/index.html>, accessed October 2023). Pairwise regressions were used to identify autocorrelated relationships between variables, and the dataset was filtered so that variables were independent. The final set of bioclimatic and soil variables was compared across the range of *E. tetradonta* with a PCA using the *factoextra* package (<https://cran.r-project.org/web/packages/factoextra/index.html>, accessed October 2023).

2.8 | Analysis of Flowering Time

Following the method of Osborne et al. (2020), we used herbarium data to estimate and compare flowering time in *E. megasepala* and *E. tetradonta*. For this, all herbarium records for both species were downloaded from the AVH (accessed 2023). To subset these data to records where the individual tree represented by the voucher specimen was in flower, we retained only those records returning a match for the word "flower" under the Darwin Core (Wieczorek et al. 2012) categories 'reproductiveCondition' and 'occurrenceRemarks'. Following this, records with missing geographic coordinate data (i.e., decimal latitude and longitude) and missing or incomplete

collection date information (i.e., year, month, and day all required) were removed. To remove duplicate specimens held across different herbaria, we identified records with duplicate 'locality' information and duplicate 'recordNumber' and 'recordedBy' information, and retained only a single specimen from the duplicate set. The resulting data were then checked by eye, and any duplicates that had passed previous filtering steps were manually removed. The phenology and identity of specimens held at CANB (10/20 *E. megasepala* and 37/85 total *E. tetradonta* records) were verified in-person by TGBM; the identity of specimens at BRI (5/20 *E. megasepala* and 15/85 *E. tetradonta*) had previously been verified by A.R. Bean. We classed each record as a single-point observation of flowering time and converted the flowering date to days after 1 January of the collection year. To investigate patterns of flowering time with respect to geography, the *E. tetradonta* dataset was divided into two additional datasets, based on locality, for comparison with *E. megasepala*. The first subset included only specimens collected from areas that could be considered to overlap with the range of *E. megasepala* (any *E. tetradonta* specimens <222 km from any *E. megasepala*); the second included only specimens collected from areas that could be considered to be non-overlapping with the range of *E. megasepala* (any *E. tetradonta* specimens >222 km from any *E. megasepala*). To generate the 'overlapping' and 'non-overlapping' subsets, the geographic range of *E. megasepala* was converted from point data to a polygon in QGIS v.3.16 using the Concave Hull v.2.0 plugin (<https://plugins.qgis.org/plugins/concavehull/>) and expanded outwards by ~222 km (2°) using the 'Buffer' tool. Records of *E. tetradonta* falling within the buffered polygon were then extracted for the 'overlapping' subset using the 'Clip' tool, and the inverse of these was extracted for the 'non-overlapping' subset. The distribution of dates from each *E. tetradonta* dataset was compared with those of the *E. megasepala* dataset using Mann–Whitney U tests.

3 | Results

3.1 | Phylogenetic Inferences

The phylogenetic trees produced by SVDquartets analyses of Dataset 1 and Dataset 2 were topologically congruent and resolved *E. megasepala* and *E. tetradonta* both on well-supported branches (Bootstrap Support (BS): >99% for the stem branches of each species across both datasets). The Dataset 1 tree (Figure 3A), which includes outgroup samples, resolved the two species as sister to each other with high support (BS: 99%). Overall, relationships between populations within each of these species were poorly resolved, although SVDquartets did recover a weakly supported clade (BS: 50%) of populations of *E. tetradonta* from west of the Carpentarian Gap (Figure 3A,B).

The phylogenetic networks produced by SplitsTree showed *E. megasepala* and *E. tetradonta* to be distinct from each other (Figure 3C), with each species forming well-defined lineages with little reticulation between them. The network generated from Dataset 1 (not shown) showed evidence of separate reticulation between the outgroups and each of *E. megasepala* and Queensland populations of *E. tetradonta*. Populations within

each species were not clearly defined, although samples of *E. tetradonta* tended to group based on geographic location (Figure 3C).

3.2 | Patterns of Genetic Variation, Diversity and Structure

BIC scores for each k -value tested by k -means clustering varied across Dataset 3.1 and Dataset 3.2 (Fig. Figure 4C,D). For Dataset 3.1, the optimum value of k was identified as 2 (BIC: 364.2), with the clusters identified under this value corresponding to the species *E. megasepala* and *E. tetradonta*. The PCA for Dataset 3.1 also retrieved two strongly distinct groups that corresponded with the two species (Figure 4A). The proportion of genetic variation explained by the first four and 10 principal components was 22.4% and 32.5%, respectively, with the first principal component accounting for 14.5% of the variation in the dataset and strongly separating *E. megasepala* and *E. tetradonta* (Figure 4A). For Dataset 3.2, k -means clustering returned the lowest BIC score for $k=1$ (BIC: 292.4), with BIC scores steadily increasing for each value of k after this (Figure 4D). Under $k=2$, the second-best supported value by BIC (BIC: 292.9), the two resultant groups corresponded with *E. tetradonta* samples from either side of the Carpentarian Gap (Figure S1). The PCA for Dataset 3.2 (Figure 4B) did weakly distinguish these groups, albeit with a lower proportion of variation explained by the first four (13.4%) and 10 (26.3%) principal components. In particular, the first and second principal components only accounted for 6.1% and 2.5% of variation, respectively.

We found similar levels of genetic diversity in *E. megasepala* (uH_e : 0.14) and *E. tetradonta* (uH_e : 0.10). Likewise, genetic diversity between populations of *E. tetradonta* was similar across its range (QLD: four populations, mean uH_e : 0.094; NT: four populations, mean uH_e : 0.082; WA: two populations, mean uH_e : 0.087).

Both methods used for identifying the optimal k -value from results of LEA snmf and structure analyses suggested $k=2$ as optimal for Dataset 3.1 and Dataset 3.2. For Dataset 3.1, the groups delineated under $k=2$ for both analyses corresponded with *E. megasepala* and *E. tetradonta* (Figure 5A; Figure S2); and genetic differentiation between the two species and clusters measured using F_{ST} was 0.28. Estimated ancestry coefficients showed low levels of admixture between the species, particularly for the LEA results, although one individual of *E. megasepala* (herbarium specimen CANB448844) showed a higher proportion of mixed ancestry (36% *E. tetradonta*, 64% *E. megasepala*) than all other samples (Figure 5A). This sample was also placed closer to *E. tetradonta* in the PCA for Dataset 3.1 (Figure 4A) but did not show greater levels of mixed ancestry in the structure analysis (Figure S2); it is likely that the PCA and LEA results for the sample were affected by the high percentage of missing data for this sample (38% missing) compared with other samples of *E. megasepala*. For Dataset 3.2, the groups delineated under $k=2$ in LEA and structure analyses weakly corresponded to samples collected from either side of the Carpentarian Gap; however, we found large amounts of admixture between these groups, most notably

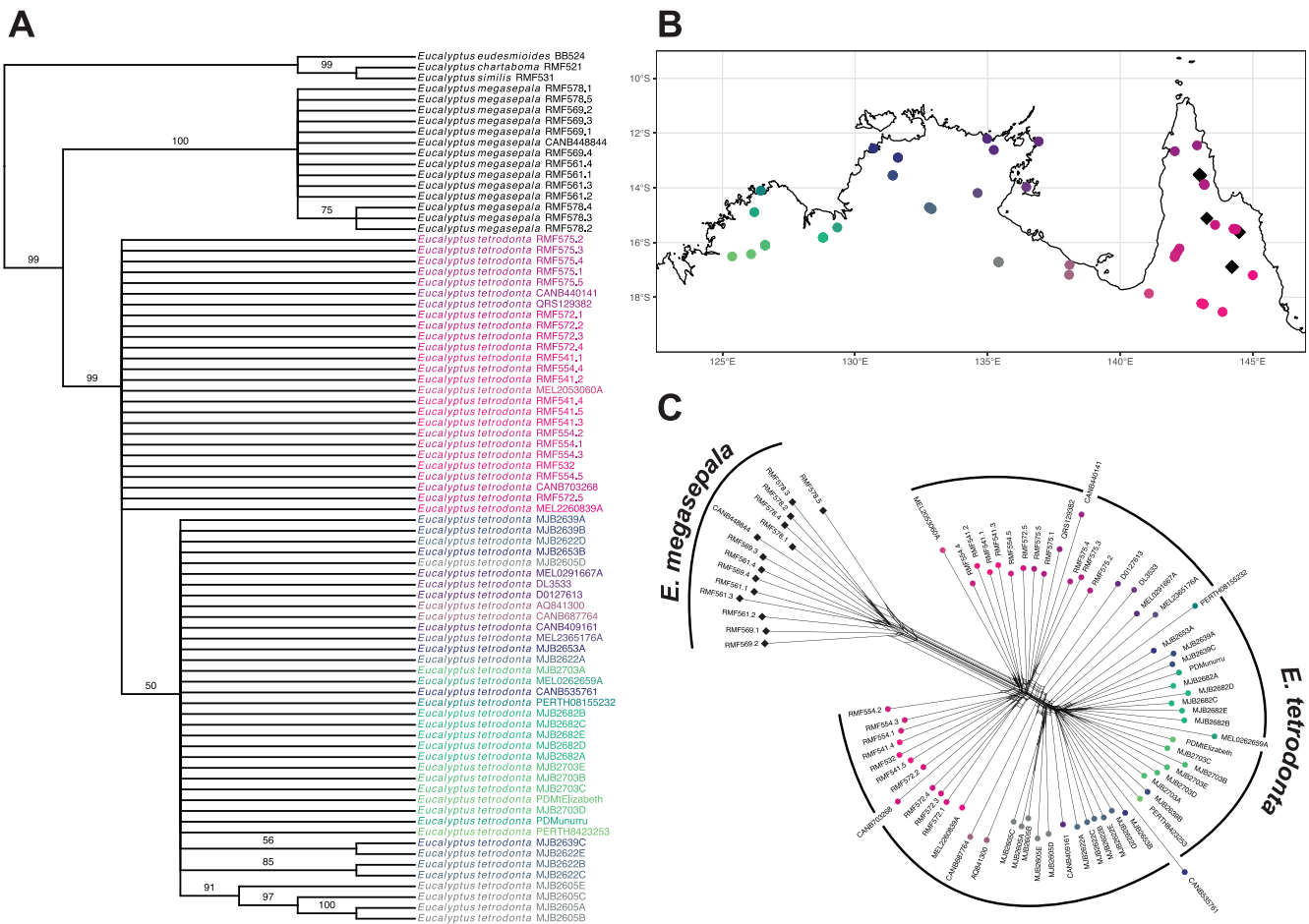


FIGURE 3 | Results of phylogenetic analyses. (A) SVDquartets phylogeny generated from analysis of Dataset 1 (81 samples [3 outgroup], 17,623 unlinked SNPs), with branch support denoted by bootstrap values above branches and branches < 50% supported collapsed. (B) Map showing sample localities for *Eucalyptus tetradonta* (circles) and *E. megasepala* (diamonds), with *E. tetradonta* samples coloured by latitude and longitude and corresponding with tip colours in (A) and (C). (C) Neighbour-Net network generated from Dataset 2 (78 samples [no outgroups], 15,064 unlinked SNPs).

among samples collected from the middle of the longitudinal range (Figure 5D,E; Figure S3). Mantel tests for IBD detected a weak signal of IBD in the dataset with both species (Dataset 3.1; Mantel $r=0.12$, $p=0.0098$); however, IBD plots showed discrete patches of genetic differentiation irrespective of geographic differentiation corresponding to pairwise comparisons of *E. megasepala* and *E. tetradonta* (Figure 5C). The detection of a comparatively much stronger signal of IBD in *E. tetradonta* (Dataset 3.2; Mantel $r=0.32$, $p=0.0001$) suggests that genetic variation is strongly linked to geographic proximity in that species (Figure 5F). Given the discrete patches recovered in the IBD plot of Dataset 3.1 (Figure 5C), it is likely that the weak signal of IBD picked up for that dataset was simply an artefact of the very strong signal in *E. tetradonta*. The results of *FEEMS* (Figure S4) were consistent across both of the tested grid cell sizes and found gene flow to be continuously reduced from north-central Queensland across the Carpentarian Gap and into the Northern Territory, with isolated patches of low gene flow also occurring in some populations of *E. tetradonta* in the Kimberley region, Western Australia. Aside from these isolated patches in Western Australia, a generally continuous region of higher gene flow was suggested to occur from the Kimberley, Western Australia through Arnhem Land, Northern Territory, and reaching the

tip of Cape York Peninsula, Queensland. Hybrid indices and interclass heterozygosity statistics estimated in *triangulaR* did not provide clear evidence for hybridisation (Figure S5).

3.3 | Outlier Locus Detection and Characterisation with GO Terms

From the 994 SNPs present in Dataset 4.1, *pcadapt* identified 12 outliers among all samples of *E. megasepala* and *E. tetradonta* (Table 1). From the 769 SNPs present in Dataset 4.2, which included only geographically overlapping samples of *E. megasepala* and *E. tetradonta*, *pcadapt* identified 17 outliers (Table 2). For both datasets, *OutFLANK* and *BayeScan* did not detect any outliers. Of the 12 and 17 SNP outliers identified by *pcadapt* for each dataset, three SNP outliers were unique to Dataset 4.1, and eight SNP outliers were unique to Dataset 4.2 (Figure 6A,C). Outlier SNPs were distributed fairly evenly across the genome, being located on all but two chromosomes (3 and 10) of the *E. grandis* genome (Figure 6A,C).

Outlier genes from Dataset 4.1 were involved in various biological processes related to cell wall organisation, response to heat, transmembrane transport, protein modification, and

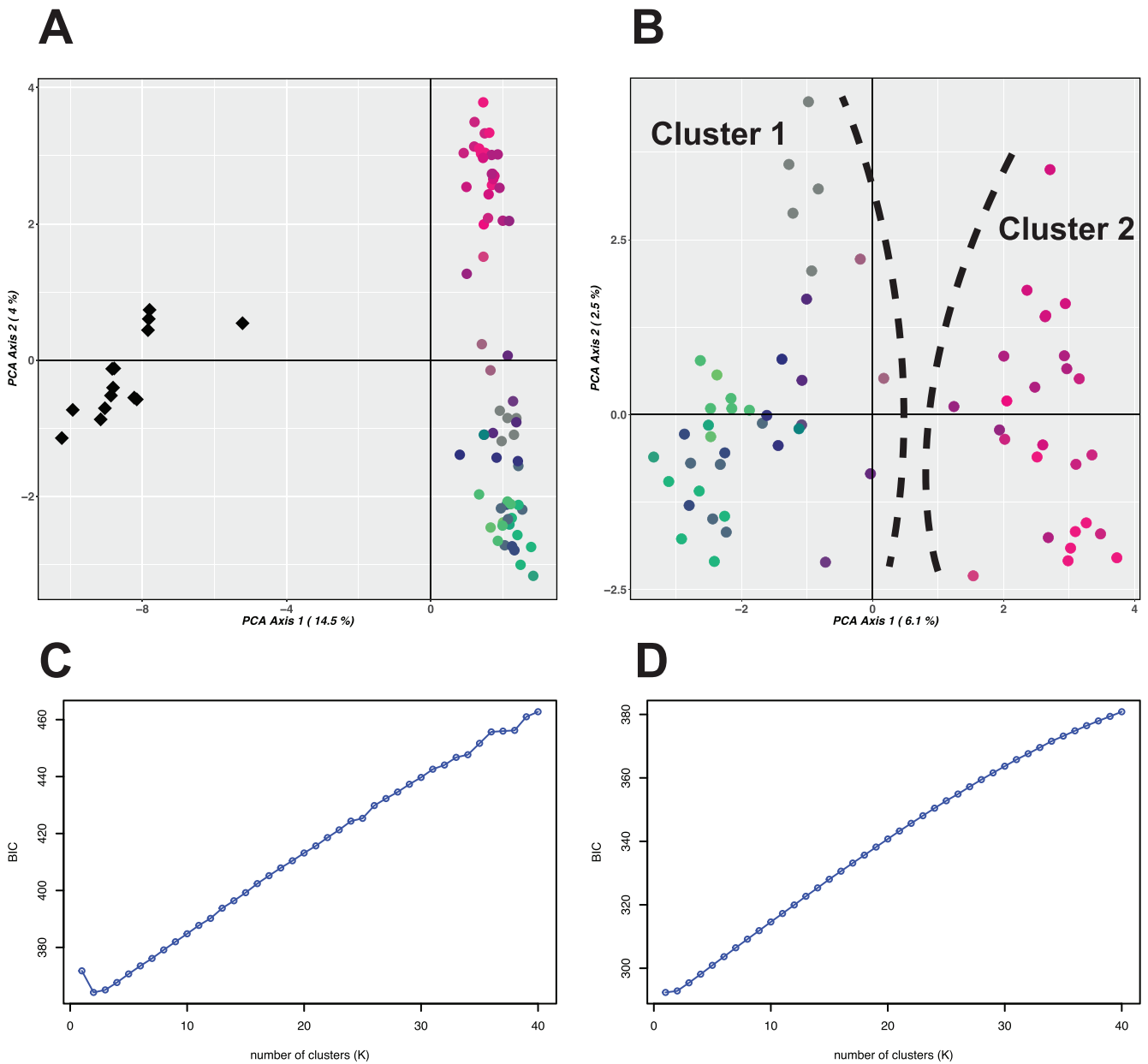


FIGURE 4 | Results of k -means clustering and PCA for Dataset 3.1 (*Eucalyptus megasepala* and *E. tetradonta*; A, C) and Dataset 3.2 (*E. tetradonta* only; B, D). (A) Plot of first two principal components, with two distinct genomic clusters detected by k -means clustering (lowest BIC in C) corresponding to species (diamonds=*E. megasepala*, circles=*E. tetradonta*); *E. tetradonta* samples are coloured by latitude and longitude to correspond with Figure 3. (B) Plot of first two principal components for samples of *E. tetradonta*, coloured by latitude and longitude to match Figure 3, with clusters for $k=2$ separated by dashed lines, and plot showing support for $k=1$ using k -means clustering (D).

phosphatidylinositol dephosphorylation (Figure 6B). In particular, allele frequencies for genes linked to heat response, transport, and protein modification by small protein removal (such as ubiquitin or a ubiquitin-like protein) were higher in *E. megasepala*; conversely, allele frequencies for genes linked to protein ubiquitination and phosphorylation were higher in *E. tetradonta* (Figure 6B). Outlier genes from Dataset 4.2 were, for the most part, related to similar biological processes as those identified in Dataset 4.1, but were also involved in ammonium homeostasis (Figure 6D). Allele frequencies for genes linked to cell wall organisation (and xyloglucan metabolic process) were also higher for *E. megasepala*; this was opposite to allele frequencies

being higher in *E. tetradonta* for the same process in Dataset 4.1. Notably, the gene encoding FT-interacting protein 7 was identified as an outlier in Dataset 4.2; however, the specific function of this gene in relation to flowering was lost by the broadness of the GO term associated with the *E. grandis* gene in UniProt (= GO:0016020: “membrane”). This limitation, which stems from uncertainty surrounding gene function and annotation in eucalypts, presumably also affected the level of GO information retrieved for other outliers, leading to the return of broad GO terms that lacked sufficient specificity to be of major use in assigning fine-scale functions to outlier genes (Huntley et al. 2014; The Gene Ontology Consortium 2019).

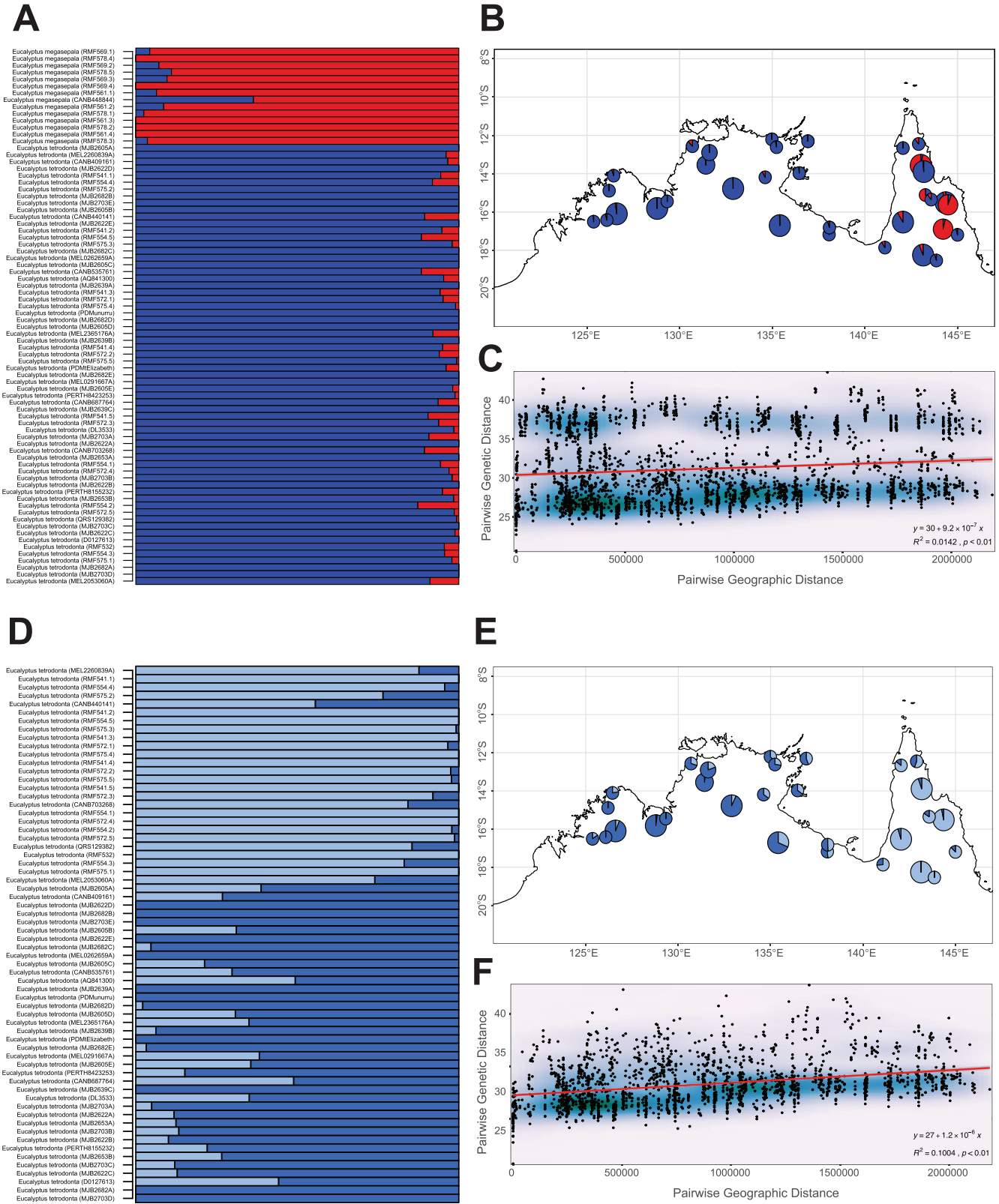


FIGURE 5 | LEA admixture coefficients and isolation-by-distance plots for Dataset 3.1 (*Eucalyptus megasepala* and *E. tetradonta*; A–C) and Dataset 3.2 (*E. tetradonta* only; D–F). (A, D) Structure-like barplots of ancestry coefficients for all individuals. (B, E) Mean population ancestry coefficients mapped as pie-charts according to collection locality, with pies scaled according to individuals in each population (smallest $n = 1$, largest $n = 5$). (C, F) Isolation-by-distance plots, showing individual pairwise genetic distance against geographic distance.

TABLE 1 | Details of SNPs identified by outlier analyses of Dataset 4.1. Numbers in the FigureIndex column relate to those in Figure 6A.

| FigureIndex | LocusName | MinorAlleleFreq | AlleleFreq | AlleleFreq/Meg | AlleleFreqTet | AlleleFreqRatio | DetectedBy | Chromosome | SnpPosition | SnpDistanceFromGene | SnpPositionInGene | GeneID | Product | UniProtSymbol | UniProtProteinName | UniProtGOIDs | |
|-------------|-----------------|-----------------|------------|----------------|---------------|-----------------|------------|------------|-------------|---------------------|-------------------|--------------|---|----------------|--|--|--|
| 1 | 83585763-38-T/C | 0.1365636364 | 71.43 | 0.79 | 0.79 | 0.71 | peadap1 | 1 | 30431004 | 696 | - | LOC104438613 | auxin-responsive protein SAUR50 | EUGRSUZ_F03969 | Magnesium transporter | GO:0015095; GO:0015693; GO:0016020 | |
| 2 | 83590581-14-A/G | 0.1365636364 | 75 | 0 | 0 | 0.75 | peadap1 | 1 | 33162595 | 41 | - | LOC104437376 | small heat shock protein, chloroplastic | EUGRSUZ_A01416 | SHSP domain-containing protein | GO:0009408 GO:0004439; GO:0004445; GO:0034485; GO:0046856 | |
| 3 | 83587306-21-C/A | 0.1495506494 | 82.14 | 0 | 0 | 0.82 | peadap1 | 1 | 39470440 | Inside | Intron | LOC104411748 | probable xyloglucan endotransglucosylase/hydrolase protein 10 | EUGRSUZ_A01968 | Inositol polyphosphate-related phosphatase domain-containing protein | GO:0003886; GO:0016020; GO:0022857; GO:0055085 | |
| 4 | 4839540-46-A/C | 0.1428571429 | 78.57 | 0 | 0 | 0.79 | peadap1 | 2 | 14098765 | Inside | Exon | LOC104427397 | protein PNS1 | EUGRSUZ_K01535 | Choline transporter-like protein | GO:0015095; GO:0016020 | |
| 5 | 83591896-44-C/T | 0.1883116883 | 96.43 | 1.59 | 0.95 | 0.95 | peadap1 | 5 | 2968432 | Inside | Exon | LOC104443574 | type 1 inositol polyphosphate 5-phosphatase 2 | EUGRSUZ_E00306 | Magnesium transporter | GO:0005794; GO:0016020 | |
| 6 | 8747499-19-G/C | 0.1466666667 | 75 | 0.82 | 0.74 | 0.74 | peadap1 | 7 | 47977231 | Inside | Exon | LOC104453901 | protein CANDIDATE G-PROTEIN COUPLED RECEPTOR 7 | EUGRSUZ_G02451 | Protein GPR107 | GO:0005794; GO:0016020 | |
| 7 | 4837849-50-G/A | 0.1755246753 | 85.71 | 2.38 | 0.83 | 0.83 | peadap1 | 8 | 57212870 | Inside | Exon | LOC104417548 | putative ubiquitin-conjugating enzyme E2 38 | EUGRSUZ_H04049 | UBC core domain-containing protein | GO:0061631 GO:0004672; GO:0004842; GO:0005524; GO:000004648; GO:0016567 | |
| 8 | 45459831-18-G/A | 0.1133333333 | 39.29 | 100 | -0.61 | -0.61 | peadap1 | 9 | 13133652 | Inside | Exon | LOC104418337 | putative U-box domain-containing protein 50 | EUGRSUZ_I00562 | U-box domain-containing protein 50 | | |
| 9 | 83593474-16-G/C | 0.3289473684 | 7.14 | 38.71 | -0.32 | -0.32 | peadap1 | 9 | 24494485 | >750 | - | - | - | - | - | - | |
| 10 | 83591433-36-G/A | 0.2105263158 | 89.29 | 5.65 | 0.84 | 0.84 | peadap1 | 9 | 35737964 | Inside | Exon | LOC104419563 | LOW QUALITY PROTEIN: RING-H2 finger protein ATL8 | EUGRSUZ_I02211 | RING-type domain-containing protein | GO:0016020 GO:0008233; GO:0070646 | |
| 11 | 83594040-51-A/G | 0.1346153846 | 75 | 0 | 0.75 | 0.75 | peadap1 | 11 | 30228334 | Inside | Exon | LOC104425740 | hypothetical protein | EUGRSUZ_K02178 | PPPDE domain-containing protein | | |
| 12 | 83587884-52-C/T | 0.1474358974 | 78.57 | 0.78 | 0.78 | 0.78 | peadap1 | 11 | 34028667 | Inside | Exon | LOC104426197 | adenine/guanine permease AZG2 | EUGRSUZ_K02759 | Adenine/guanine permease AZG2 | GO:0005345; GO:0003886 | |

TABLE 2 | Details of SNPs identified by outlier analyses of Dataset 4.2. Numbers in the FigureIndex column relate to those in Figure 6C.

| FigureIndex | LocusName | MinorAlleleFreq | AlleleFreqMeg | AlleleFreqTet | AlleleFreqRatio | DetectedBy | Chromosome | SnpPosition | SnpDistanceFromGene | SnpPositionInGene | GeneID | Product | UniProtSymbol | UniProtProteinName | UniProtGOIDs |
|-------------|-----------------|-----------------|---------------|---------------|-----------------|------------|------------|-------------|---------------------|-------------------|--------------|---|----------------|--|---|
| 1 | 83590881-14-A/G | 0.3088235294 | 75 | 0 | 0.75 | pendapt 1 | 1 | 33162595 | 41 | - | LOC104437376 | small heat shock protein, chloroplatic | EUGRSUZ_A01416 | SHSP domain-containing protein | GO:0009408 GO:0004553; GO:0010411; GO:0016762; GO:0042546; GO:0048046; GO:0071555 |
| 2 | 83587306-21-G/A | 0.3382352941 | 82.14 | 0 | 0.82 | pendapt 1 | 1 | 39470440 | Inside | Intron | LOC104441748 | probable xyloglucan endotransglucosylase/hydrolase protein 16 | EUGRSUZ_A01968 | Xyloglucan endotransglucosylase/hydrolase (EC 2.4.1.207) | GO:0005886; GO:0016020; GO:0022857; GO:0055085 |
| 3 | 4839540-46-A/C | 0.3235294118 | 78.57 | 0 | 0.79 | pendapt 2 | 2 | 14098765 | Inside | Exon | LOC104427397 | protein PNS1 | EUGRSUZ_K01535 | Choline transporter-like protein | GO:0016020 |
| 4 | 83594832-59-G/C | 0.3484848485 | 78.57 | 2.63 | 0.76 | pendapt 2 | 2 | 14109488 | Inside | Exon | LOC104425302 | FT-interacting protein 7 | EUGRSUZ_K01536 | C2 domain-containing protein | GO:0008757; GO:0032529; GO:0046872 |
| 5 | 8746136-38-C/A | 0.2575757576 | 60.71 | 0 | 0.61 | pendapt 2 | 2 | 44310969 | Inside | Intron | LOC104433048 | loganic acid O-methyltransferase | EUGRSUZ_B02682 | S-adenosylmethionine-dependent methyltransferase | GO:0003680; GO:0003700; GO:0005634 |
| 6 | 83594747-25-G/T | 0.4545454545 | 96.43 | 7.89 | 0.89 | pendapt 4 | 4 | 33911151 | 21 | - | LOC10442013 | AT-hook motif nuclear-localized protein 23 | EUGRSUZ_D02188 | PPC domain-containing protein | GO:0004439; GO:0004445; GO:0034485; GO:0046856 |
| 7 | 83591896-44-C/T | 0.4117647059 | 96.43 | 2.5 | 0.94 | pendapt 5 | 5 | 2968432 | Inside | Exon | LOC104443574 | type I inositol polyphosphate 5-phosphatase 2 | EUGRSUZ_E00306 | Inositol polyphosphate-related phosphatase domain-containing protein | GO:0004672; GO:0005324; GO:0006468; GO:0016020 |
| 8 | 83592756-34-C/T | 0.3676470588 | 85.71 | 2.5 | 0.83 | pendapt 6 | 6 | 2233479 | Inside | Exon | LOC104447867 | probable inactive receptor kinase A.5g67200 | EUGRSUZ_F00189 | Protein kinase domain-containing protein | GO:0005886; GO:0008519; GO:0072488; GO:0097272 |
| 9 | 83587691-50-C/G | 0.3529411765 | 85.71 | 0 | 0.86 | pendapt 6 | 6 | 22233137 | Inside | Exon | LOC104432553 | ammonium transporter 1 member 1 | EUGRSUZ_B02160 | Ammonium transporter | GO:0005737; GO:0007049; GO:0008289; GO:0016020; GO:0051301 |
| 10 | 83590416-40-G/C | 0.2352941176 | 57.14 | 0 | 0.57 | pendapt 6 | 6 | 53265901 | Inside | Exon | LOC104450763 | patellin-3 protein CANDIDATE G-PROTEIN COUPLED RECEPTOR 7 | EUGRSUZ_F04114 | CRAL-TRIO domain-containing protein | GO:0005794; GO:0016020 |
| 11 | 8747499-19-G/C | 0.3235294118 | 75 | 2.5 | 0.73 | pendapt 7 | 7 | 47977231 | Inside | Exon | LOC104453901 | protein TIC 20-IV, chloroplatic-like | EUGRSUZ_G02451 | Protein GPR107 | - |
| 12 | 11944478-18-G/T | 0.2575757576 | 50 | 92.11 | -0.42 | pendapt 7 | 7 | 5133505 | 7 | - | LOC104455794 | probable indole-3-pyruvate monooxygenase YUCCA7 | EUGRSUZ_G02451 | Flavin-containing monooxygenase (EC 1.-.-.-) | GO:0004497; GO:0004499; GO:0050660; GO:0050661 |
| 13 | 4838063-11-C/T | 0.1911764706 | 25 | 1.5 | 0.10 | pendapt 8 | 8 | 39965166 | Inside | Exon | LOC104414392 | putative ubiquitin-conjugating enzyme E2 38 | EUGRSUZ_H02874 | UBC core domain-containing protein | GO:0001601 |
| 14 | 4837849-50-G/A | 0.3676470588 | 85.71 | 2.5 | 0.83 | pendapt 8 | 8 | 57212870 | Inside | Exon | LOC104417548 | LOW QUALITY PROTEIN: RING-H2 finger protein ATL8 | EUGRSUZ_H04049 | RING-type domain-containing protein | GO:00016020 |
| 15 | 83591433-36-G/A | 0.3787878788 | 89.29 | 0 | 0.89 | pendapt 9 | 9 | 35737964 | Inside | Exon | LOC104419563 | adenine/guanine permease AZG2 | EUGRSUZ_I02211 | Adenine/guanine permease | GO:0005345; GO:0005886 |
| 16 | 83594040-51-A/G | 0.3088235294 | 75 | 0 | 0.75 | pendapt 11 | 11 | 30228534 | >750 | - | - | - | - | - | - |
| 17 | 83587884-52-G/T | 0.3235294118 | 78.57 | 0 | 0.79 | pendapt 11 | 11 | 34028667 | Inside | Exon | LOC104426197 | adenine/guanine permease AZG2 | EUGRSUZ_K02759 | Adenine/guanine permease | GO:0005345; GO:0005886 |

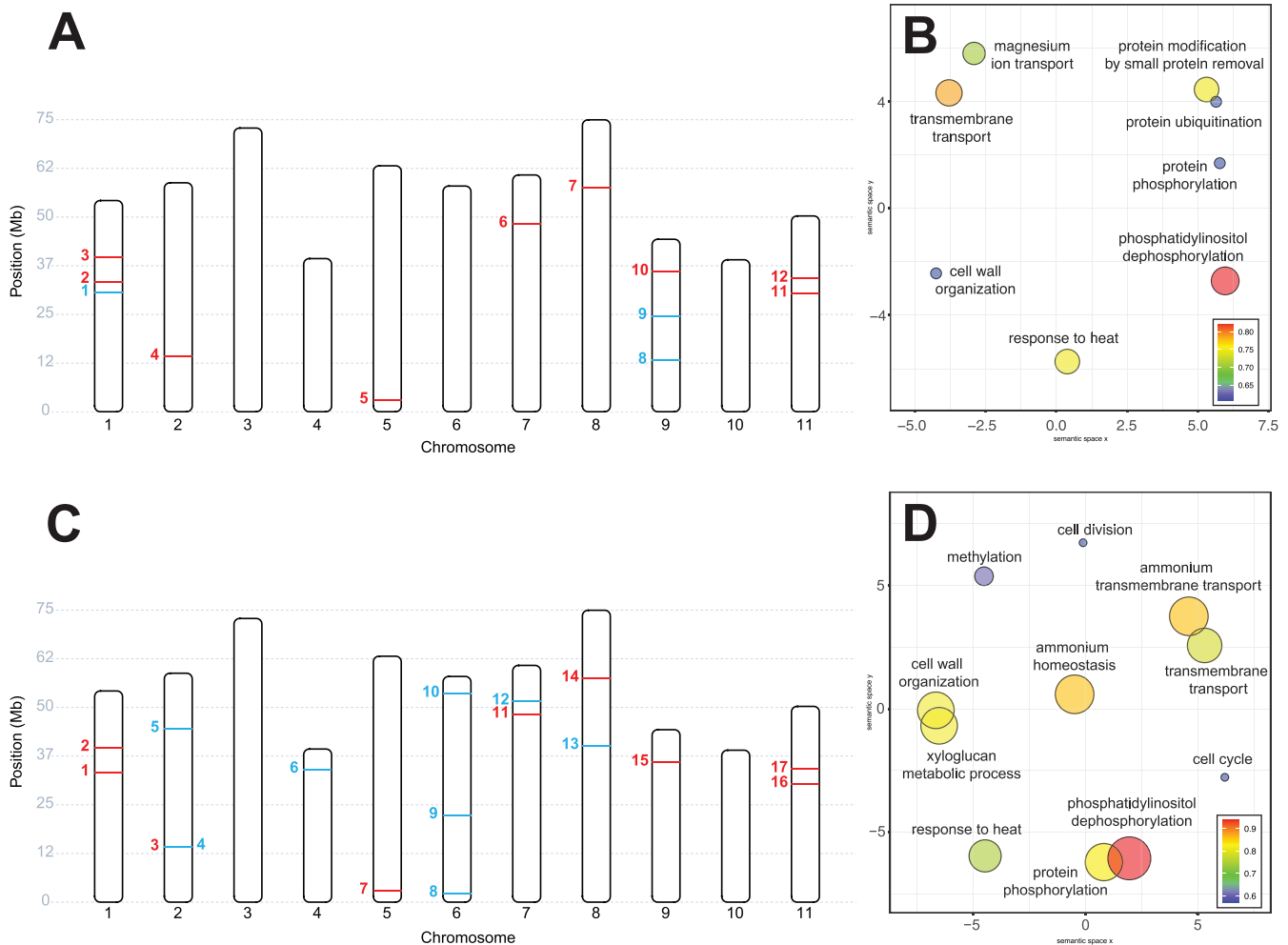


FIGURE 6 | Outlier SNP locations and REVIGO gene ontology visualisations for Dataset 4.1 (*Eucalyptus megasepala* and all *E. tetradonta*; A, B) and Dataset 4.2 (*E. megasepala* and geographically proximal *E. tetradonta* only; C, D). (A, C) Positions of outlier SNPs identified by *pcadapt* on the *E. grandis* genome (Myburg et al. 2014). SNPs in blue were identified uniquely for their respective dataset, while those in red were identified for both datasets 4.1 and 4.2. Numbers correspond to SNP information presented in Tables 1 and 2. (B, D) REVIGO scatterplots showing Gene Ontology (GO) biological processes associated with outlier SNPs in semantic space (where axes have no intrinsic meaning and simply aid visualisation). For each plot, circles denoting GO terms are scaled and coloured according to the amount of allele fixation (i.e., the absolute difference in allele frequencies: $|(AF_{E.meg} - AF_{E.tet})/100|$) of outlier SNPs associated with each term between *E. megasepala* and *E. tetradonta* from 0 to 1, with higher values corresponding to higher fixation between species.

3.4 | Comparisons of Environmental Conditions and Flowering Times

Flowering time was significantly different between collection records of the two species (Mann–Whitney U test; $p = 3.75 \times 10^{-6}$). The median flowering time of *E. megasepala* was 9 June, 48 days before the median for *E. tetradonta* on 27 July (Figure 7). Flowering times for the geographically overlapping and non-overlapping groups of *E. tetradonta* were both significantly different from *E. megasepala* (overlapping Mann–Whitney U test, $p = 2.75 \times 10^{-4}$; non-overlapping Mann–Whitney U test, $p = 3.00 \times 10^{-6}$), with the median flowering time for overlapping *E. tetradonta* falling 2 days before that of the non-overlapping group (25 July vs. 27 July). The difference between the flowering time of the two groups of *E. tetradonta* was non-significant (Mann–Whitney U test; $p = 0.47$).

The final set of bioclimatic variables included in the PCA were BIO3 (isothermality), BIO10 (mean temperature of the warmest

quarter), BIO16 (precipitation of wettest quarter), and BIO17 (precipitation of driest quarter); these variables failed to discriminate any difference between *E. tetradonta* and *E. megasepala* (Figure S6). As with the BIOCLIM variables, the soil variables were not able to distinguish the two species in northern Queensland.

4 | Discussion

4.1 | Strong Genetic Differentiation between *E. tetradonta* and *E. megasepala* Provides Novel Insights Into the Evolutionary Dynamics of Eucalypt Speciation

Studies of eucalypt phylogenetics or population genetics routinely show signals of gene flow and/or introgression between closely related species, especially using plastid markers (Jackson et al. 1999; McKinnon et al. 1999, 2001, 2004, 2010; Pollock

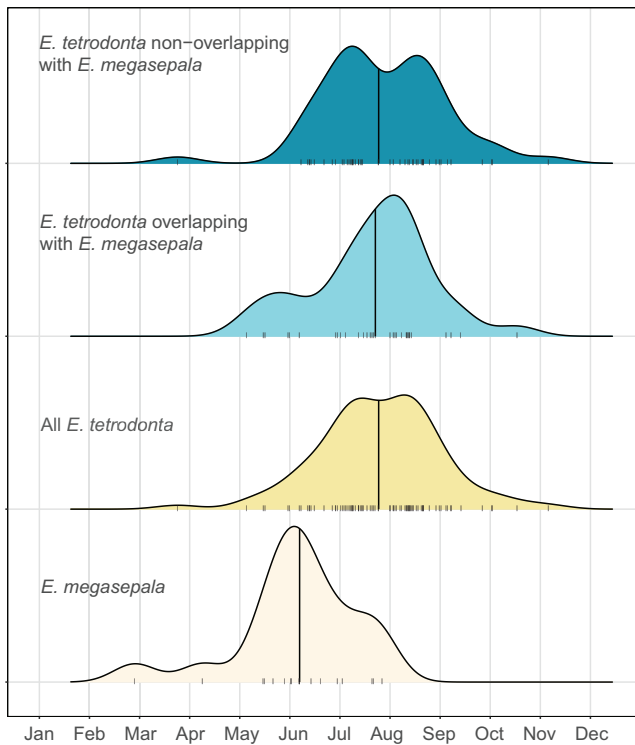


FIGURE 7 | Flowering time differences between *Eucalyptus megasepala* and *E. tetradonta* from cleaned AVH (2023) records represented by density curves. The bottom two curves include observations for all collections of *E. tetradonta* and *E. megasepala*, while the top two curves include collections of *E. tetradonta* from geographic ranges that are non-overlapping and overlapping with *E. megasepala*, respectively. Vertical black lines show the median flowering time for each group. Ticks under density curves represent individual observations.

et al. 2013, 2015; Nevill et al. 2014; Fahey et al. 2021) but also using multilocus nuclear DNA datasets (Rutherford et al. 2018, 2019, 2021, 2023; Collins et al. 2019; Murray et al. 2019; von Takach Dukai et al. 2019; Fahey et al. 2022b). Our recent phylogenomic analyses of *Eucalyptus* subg. *Eudesmia*, to which our target species belong, also showed a legacy of gene flow throughout the evolution of the lineage (McLay et al. 2023). Such patterns are also commonly found in other large temperate tree genera, such as the oak genus, *Quercus*, which has been the subject of extensive study (McVay et al. 2017; Cavender-Bares et al. 2015; Cannon and Petit 2020; Hipp et al. 2020). Given this background, the level of genetic differentiation observed between *E. tetradonta* and *E. megasepala* is surprising. The fact that the two species are sisters to one another and overlap geographically in northern Queensland led us to hypothesise that we would detect some level of genetic admixture; in reality, there is almost none.

Although the ancestry plots from LEA (Figure 5A) and STRUCTURE (Figure S2) show some admixture between samples of *E. tetradonta* and *E. megasepala*, this likely reflects the shared ancestry (sister relationship) of the two species, rather than recent genetic introgression between them (see Figure S5). For *E. tetradonta*, the LEA plot shows mixed ancestry for a number of samples (up to ~13% “*E. megasepala*” ancestry), but there is no geographic pattern that could be considered consistent with recent introgression. For instance, one *E. tetradonta* sample with mixed ancestry is from the zone of

overlap between the two species (CANB 703268), but others are from outside the range of *E. megasepala* in Queensland (RMF 554.5, RMF 541.2, RMF 541.5, CANB 440141), the Northern Territory (CANB 535761) and Western Australia (MJB 2703) (see Figure 1 for locations of samples), making introgression an unlikely explanation for the admixture. The sample of *E. megasepala* (CANB 448844) showing more substantial admixture in LEA, but not in STRUCTURE, is explained as an artefact of the very high proportion of missing data (38% missing) for that sample. For *E. megasepala*, most samples show a proportion of mixed ancestry in both LEA and STRUCTURE plots, but the level of admixture is reasonably similar among all samples (0%–11% “*E. tetradonta* ancestry” in LEA (excluding CANB 448844) and ~5%–27% in STRUCTURE). Given the close phylogenetic and phenotypic relatedness of *E. megasepala* and *E. tetradonta*, in concert with high levels of heterozygosity and ILS in *Eucalyptus* species, this low but consistent pattern of mixed ancestry represents a background signal of the ancestral genetic pool of both species, rather than recent introgression.

The finding of strong genetic differentiation and limited evidence for an intergradation zone between sister species growing in adjacent habitats contrasts with the findings of many other studies on *Eucalyptus*. Extensive examples of intergradation zones where gene flow between closely related species occurs exist for *Eucalyptus* taxa growing both in adjacent geographic regions (e.g., *E. melanophloia* and *E. whitei* (Holman et al. 2011), *E. saligna* and *E. botryoides* (Passioura and Ash 1993) and Scribbly Gum spp. (Rutherford et al. 2023)), and abutting but contrasting habitat (e.g., *E. microcarpa* and *E. wimmerensis* (Fahey et al. 2022b), *E. rubida* and *E. aggregata* (Field et al. 2011), and the multiple examples of Robins et al. (2021)). The differentiation between the two species examined here and the maintenance of strong species boundaries is likely the outcome of a complex interplay of environmental factors and biological speciation mechanisms (Levin 2006; Mallet et al. 2009; Givnish 2010), in addition to the lineage being far less speciose than the *Eucalyptus* lineages where large intergradation zones are typically observed (e.g., section *Eucalyptus*, section *Adnataria* and section *Glandulosae* which each hold > 50 species (Nicolle 2024)). Evidence for a lack of gene flow has also been observed between other sympatric and closely related species of *Eucalyptus*, including *E. salubris* and *E. ravidia* (Binks et al. 2021), *E. dalrympleana* and *E. rubida* (Fahey et al. 2024), and *E. polybractea* and *E. viridis* (Fahey et al. 2022b). In combination with previously characterised ecological differences between the two species studied here, results from this research offer some insight into the factors potentially involved in the speciation and maintenance of clear genetic distinction between the two focal species and *Eucalyptus* species more broadly.

Species phenology is often an important ecological marker in distinguishing closely related species (e.g., in palms (Osborne et al. 2019), *Metrosideros* (Osborne et al. 2020), and irises (Osmolovsky et al. 2022)), and in limiting gene flow between sympatric or parapatric species or populations, including in eucalypts (Potts and Reid 1985; McGowen et al. 2001; Foster et al. 2007). Phenology was noted as a distinguishing character in the original description of *E. megasepala* (Bean 2006), and our analysis of herbarium specimen metadata revealed the extent of

phenological differences between *E. tetradonta* and *E. megasepala*; in particular, there is a strong separation of flowering times in Queensland, where the two species co-occur (Figure 7). The presence of overlaps in the data may be due to seasonality shifts in the wet-dry seasons in northern Australia that impact the onset of flowering in both species in different years, rather than an indication of actual flowering overlap. Notably, our investigation of outlier SNPs between the two species identified an SNP exhibiting perfect segregation between the two species, and that SNP resides within the gene FT-interacting protein 7 (FTIP7), known to modulate flowering time in plants (Pin and Nilsson 2012; Yue 2012; Song et al. 2018). FTIP7 is a homologue of FTIP1, which mediates the movement of Flowering Locus T (FT) through the phloem. FT plays a central role in determining the onset of flowering (Pin and Nilsson 2012). Although we have not fully characterised differences in FTIP7 between the two species, our finding that one of the few segregating SNPs between the two species is associated with a gene that relates to a difference in phenology suggests this locus could be important in maintaining species boundaries.

Further ecological differentiation between the two species is found in subtle edaphic variation, even over extremely short distances where the two species can occur in the same vicinity on similar yet distinct substrate profiles. In Queensland, where the two species co-occur, *E. tetradonta* tends to occur on deeper, well-drained, sandy, colluvial, and laterite soils and on downslope sites relative to *E. megasepala* (Franklin 2022). In contrast, *E. megasepala* occurs on harsher, upslope sites (on ridges, plateaus or outcrops) of lower quality, larger particle soils (Bean 2006; Franklin 2022). Our analyses based on environmental and soil data were not able to identify any differentiation, but this is likely due to the lack of fine-scale granularity of the publicly available datasets used in the analyses. Geological mapping has identified Cape York Peninsula as a complex mosaic of ancient and more contemporary soils that are highly leached with low levels of nutrients (Wannan 2014). There is some evidence of heavy metals in the soils, such as manganese, gold, copper, and mineral ores such as bauxite, which may impact plant distributions in the area, but no current mapping of profiles is fine-scale enough to differentiate between the soil preferences of the two species across their overlapping ranges (Biggs and Philip 1994; Denaro and Ewers 1995; Pain et al. 1995). Our outlier analyses also point to some loci associated with soil chemistry and heavy metal tolerances (e.g., ammonium and magnesium transport, Table 2). Slight differences in soil tolerances have been shown to impact distributions of other eucalypts over short ranges (e.g., Parsons 1969). If such a difference in soil chemistry is inhibiting either species, then it is possible that any potential hybrids are not well-suited to survive on either of the substrates, leading to reinforcement of the differentiation between the two species (Rieseberg and Carney 1998; Givnish 2010).

Interactions between edaphic variables and climate are likely to have some bearing on variation in reproductive/flowering phenology in this system. The effects of temperature and rainfall have been shown to strongly influence the timing of flowering in eucalypts (Rawal et al. 2015a, 2015b), and many eucalypt species exhibit a flush of flowering following seasonal rainfall (Law et al. 2000). Northern Australia is defined by its strongly

monsoonal climate, with a distinct, almost rainless dry season spanning 7–8 months followed by a torrential wet season (Woinarski et al. 2007), and relative to southern eucalypts, the reproductive biology of savannah species in northern Australia is likely to be even more tightly linked to this seasonal variation. Flowering of *E. tetradonta* is observed to be seasonally linked, with bud formation and flowering occurring following the onset of the dry season (Setterfield and Williams 1996; Franklin 2022). Setterfield and Williams (1996) noted that peak seedfall of a Kakadu (Northern Territory) population of *E. tetradonta* between 1992 and 1994 coincided with the first rains of the wet season and hypothesised that seeds that fall following the onset of rains experience increased germination success and reduced loss due to ant harvesting. In *E. tetradonta* and *E. megasepala*, differences in flowering time may be related to differences in soil moisture and water retention as a consequence of soil or substrate variation, namely, the occurrence of *E. tetradonta* on more fertile, deep soils of downslope sites and *E. megasepala* on shallower and less fertile, rocky upslope sites. Despite experiencing equivalent day lengths, temperatures and rainfall across their overlapping geographic ranges, the moisture retaining capacity of the soil/substrate is likely to vary considerably between the two species (i.e., the rocky upslope sites of *E. megasepala* could have lower soil water availability compared to *E. tetradonta*). Some anecdotal evidence of this has been observed in the field; in November 2015 (i.e., late in the dry season), foliage on *E. megasepala* near Laura, Queensland was severely desiccated, whereas that of *E. tetradonta* immediately downslope was not (DCF, personal observation). It is conceivable that the soil moisture conditions of *E. megasepala* may trigger an earlier flowering time relative to *E. tetradonta*.

While we have some insight into the underlying mechanism(s) that likely maintain species boundaries and may have led to speciation, what is less clear is whether speciation of *E. tetradonta* and *E. megasepala* occurred in close proximity as the two species now occur, or whether geographic separation between two populations allowed for differentiation prior to secondary contact. Sympatric or parapatric speciation is conceivable via local adaptation, potentially driven by subtle genetic changes at fine scales, for instance, through selection of traits associated with edaphic factors and reinforced by shifts in flowering time. Alternatively, vicariance and secondary contact and thus allopatric differentiation is also plausible, with barriers to gene flow, such as discrete flowering times and different ecological tolerances, developing in allopatry and preventing admixture on secondary contact. Given debate about the relative importance of sympatric versus allopatric speciation in plants (Fitzpatrick et al. 2008; Mallet et al. 2009; Bock et al. 2023), further studies of *E. tetradonta* and *E. megasepala* could be worthwhile. These could better quantify ecological differentiation of the two species (given that environmental factors used here were largely overlapping), model their distributions under past climates and also study the extent of genetic differentiation, including in outlier loci, at a finer scale than done here. Inclusion of plastid DNA markers, which show widespread introgression in other eucalypt groups (McKinnon et al. 2010; Pollock et al. 2013, 2015; Nevill et al. 2014; Fahey et al. 2021) could also provide further insight into the extent of genetic differentiation between the two species.

The level of genetic differentiation that we detected at thousands of short SNP loci potentially points towards larger, more significant physical and genetic variation in the genomes of these two species that could be better detected with whole genome sequencing or longer DNA sequence reads. Genomic rearrangements and microsyntenic differences between the genomes of *Eucalyptus* species have recently been revealed to be a driver of genetic variation within the genus (Ferguson et al. 2023). Several papers by Ferguson et al. (2023, 2024b, 2024a) have shown how dynamic eucalypt genomes can be, with a high level of microsyntenic variation between species. Their studies focus on more distantly related species than *E. tetradonta* and *E. megasepala*, so it is unclear how generalisable their findings may be to this speciation scenario. Given the outlier SNP findings might point towards important speciation mechanisms in plants (flowering time, soil substrates), there is also the possibility that differentiation between the two species relies on a handful of ‘barrier loci’ rather than genomic-scale changes (Barton and Bengtsson 1986; Ravinet et al. 2017; Kautt et al. 2020). Further sampling of both specimens (populations more broadly, distinctly looking for possible hybrids, etc.) and genomic (whole genome and whole genome resequencing), together with better outgroup sampling, might lend itself to coalescent analyses and investigations of speciation with/without gene flow. This type of genetic sampling could be complemented with soil chemistry and moisture data for the edaphic profiles of both species, cytological data, common garden experiments and assessment of crossability to provide a more complete picture of differentiation and speciation between the two species.

4.2 | Minimal Genetic Variation and Geographic Structure Exists in *E. tetradonta* across Northern Australia

In contrast with the significant genetic differentiation observed between the two species in Cape York Peninsula, *E. tetradonta* exhibits notably low genetic differentiation across its range throughout northern Australia (encompassing some 800,000 km²). Genetic variation in the species follows a clear IBD pattern (Figure 5), with most genetic structure occurring between samples collected from either side of the Carpentarian Gap. However, the specimens we sampled immediately adjacent to the Gap were genetic intermediates of the two genetic clusters, indicating possible admixture in the past between the eastern and western ranges of the species or the interruption of an IBD continuum caused by population extinction and a lack of suitable habitats (Figure 5).

The plains across northern Australia that *E. tetradonta* occupies are old and geologically stable, but the climate and sea level have fluctuated significantly between periods of glacial maxima and warmer interglacial periods (Chivas et al. 2001; De Deckker 2001). During glacial maxima, sea levels dropped by up to 120–140 m, the Gulf of Carpentaria was reduced to a shallow lake, and land connections formed between Cape York and New Guinea across the Torres Strait (via the Torres Sill), and Arnhem Land and New Guinea across the Arafura Sea (via the Arafura Sill) (De Deckker 2001; Hope et al. 2004), providing vast tracts of newly revealed land and the opportunity for range expansions for a variety of organisms (Reeves et al. 2008; Edwards

et al. 2017). During warm interglacial periods, sea levels were 5–8 m higher than today (Hope et al. 2004), resulting in a marine environment in the Gulf with the sea extending further inland than at present (Reeves et al. 2008). Together with glacial cycles and sea-level changes, the region has also experienced short-term patterns of aridity cycling (~2.6Ka), presenting another environment for expansion and adaptation (De Deckker 2001).

These historical environmental dynamics have undoubtedly impacted the biota of the region, and the Carpentarian Gap has been shown to be a barrier to many species distributions and the cause of major vicariance events in many taxa (Bowman et al. 2010; Edwards et al. 2017). While there are numerous biogeographic barriers that have been inferred across northern Australia (Edwards et al. 2017), the Carpentarian Gap is the only one that appears to have been somewhat of a barrier to gene flow in *E. tetradonta*. Even so, our results suggest that it has not had a strong enough influence on the genetic structure in the species to produce discrete genetic groups from either side of the gap.

Bowman et al. (2010) suggested that the widespread distribution of *E. tetradonta* could be explained by an ability to grow on well-drained clay soils, such as on the plains of the Gulf of Carpentaria, as well as oligotrophic soils derived from sandstone and laterite that are present either side of the Carpentarian Gap on Cape York Peninsula and the Top End/Kimberley (Bowman et al. 2010). Certainly, the wide ecological tolerances of *E. tetradonta* do not preclude it from possible range expansions onto the Torres and Arafura Sills during glacial periods; support for higher gene flow across the northern populations in our *FEEMS* analysis may reflect this. Despite these tolerances, the present-day range of *E. tetradonta* has a small east–west disjunction (of ~300 km; Figure 1) near the southeast corner of the Gulf of Carpentaria. The fact that the species is persisting on the periphery of the disjunction suggests that it is either at the edge of its ecological tolerances and this zone is acting as a narrow geographic border restricting gene flow or that the species simply hasn't yet reoccupied the area and reconnected the two genetic populations.

An alternative explanation for the relatively low genetic differentiation of *E. tetradonta* across its large northern range (including across the Carpentarian Gap) is the long-distance pollen transfer ability of a number of likely pollinator species. The little red flying fox (*Pteropus scapulatus*) has a widespread distribution across northern and eastern Australia, and eDNA studies have identified *E. tetradonta* and *E. megasepala* flower nectar and pollen to be major components of their diet (Bradford et al. 2022). Observational and tracking data indicate that the little red flying fox is extremely mobile and is capable of long-range movements of >400 km per night or 3000–6000 km per year (Vardon and Tidemann 1999; Welbergen et al. 2020; Westcott et al. 2020) and is thus more than capable of the transfer of pollen across the Carpentarian Gap. Likewise, the nectarivorous varied lorikeet (*Psitteuteles versicolor*) is a documented visitor of *E. tetradonta* flowers, with a wide distribution across northern Australia (Franklin and Noske 2000; ALA 2024). Flight distances are not known for the varied lorikeet; however, other lorikeet species are known to take daily flights to feeding sites up to 35 km away (Southerton et al. 2004), and flock influxes associated with eucalypt flowering events indicate flight

distances of the varied lorikeet may be substantial (Woinarski and Tidemann 1991; D. Franklin pers. obs).

Within the eucalypts, patterns of distribution suggest that the Carpentarian Gap has played an important role in shaping extant diversity. Biogeographic analyses across the whole of *Eucalyptus* indicate two major phylogeographic regions across northern Australia, divided at the Carpentarian Gap; (1) monsoonal areas in Western Australia and Northern Territory and (2) more tropical areas of northern Queensland (González-Orozco et al. 2014). Several other widespread species or sister-species pairs in *E.* subg. *Eudesmia* (based on relationships inferred by McLay et al. 2023) are separated across the Carpentarian Gap including: *E. miniata* (west) and its sister *E. chartaboma* (east, on Cape York Peninsula); *E. lirata* (west) and its sister *E. similis* (east); and *E. phoenicea*, which has disjunct populations either side of the Carpentarian Gap, with those in the east considered by Nicolle (2024) to represent a distinct, undescribed species, referred to by the informal name “*Eucalyptus* sp. Battlecamp (Nicolle 1789)”. For all of these examples of eucalypt species that are closely related to *E. tetradonta*, the Carpentarian Gap has seemingly played a key role in creating vicariant populations that have led to subsequent differentiation and speciation. The lack of strong phylogeographic structure in *E. tetradonta* across the Gap could indicate either that *E. tetradonta* is in the early stages of this process, or that the Gap does not represent a substantial barrier to the species because of its ecological tolerances or pollinator dispersal abilities.

Comparing this finding to other plant groups is difficult because biogeographic patterns across northern Australia using genetic methods are understudied in plants. Presently, our biogeographic understanding of the region is more heavily informed by studies on animals (including reptiles (Moritz et al. 2016; Potter et al. 2016; Fenker et al. 2024), mammals (Cremona et al. 2021; Potter et al. 2012, 2024; Umbrello et al. 2024), frogs (Catullo et al. 2014), and flies (Manawaduge et al. 2023)), which have generally found more discrete genetic groupings in northern Australia, associated with various biogeographic barriers, than we observed in *E. tetradonta*. Further phylogeographic investigations of plant species with both wide and narrow distributions and diverse life history and reproductive traits are necessary in order to build a better understanding of plant biogeography in the northern Australian monsoon tropics using comparative phylogeography.

Author Contributions

Harvey K. Orel: data curation (equal), formal analysis (lead), methodology (supporting), software (lead), visualization (lead), writing – original draft (lead), writing – review and editing (lead). **Todd G. B. McLay:** formal analysis (supporting), investigation (supporting), methodology (supporting), project administration (equal), resources (supporting), writing – original draft (lead), writing – review and editing (lead). **Daniel J. Murphy:** funding acquisition (supporting), resources (supporting), writing – review and editing (supporting). **David J. Cantrill:** conceptualization (equal), funding acquisition (equal), resources (supporting), writing – review and editing (supporting). **Frank Udovicic:** resources (supporting), writing – review and editing (supporting). **Patrick S. Fahey:** formal analysis (supporting), writing – review and editing (supporting). **Donald C. Franklin:** resources (supporting), writing – review and editing (supporting). **Donna Lewis:** resources (supporting), writing – review and editing (supporting). **Philip G.**

Docherty: resources (supporting), writing – review and editing (supporting). **Adam White:** resources (supporting), writing – review and editing (supporting). **Michael J. Bayly:** conceptualization (equal), funding acquisition (equal), project administration (equal), resources (supporting), writing – original draft (supporting), writing – review and editing (supporting). **Rachael M. Fowler:** data curation (equal), formal analysis (supporting), investigation (lead), methodology (lead), project administration (equal), resources (lead), visualization (supporting), writing – review and editing (supporting).

Acknowledgements

This study was funded through Eucalypt Australia (grant 59 735 884 485). Plant collecting permits were provided by the Western Australian Department of Biodiversity Conservation and Attractions (FT61000813-1B), Main Roads Western Australia (D22#429977-21/1082), Northern Territory Parks and Wildlife Commission (68953), Northern Territory Department of Infrastructure Planning and Logistics (Road Operations; 22110095), and the Queensland Parks and Wildlife Service (PTU18-001474). We would like to thank the Australian Tropical Herbarium, Queensland Herbarium, Australian National Herbarium, University of Melbourne Herbarium, National Herbarium of Victoria, and Western Australian Herbarium for provision of leaf material for this study. Staff at ATH assisted with field work. The Peter Doherty Institute allowed access to their sonicator. R.M.F. and T.G.B.M. were successive holders of the Pauline Ladiges Postdoctoral Fellowship in Plant Systematics, supported by The University of Melbourne Botany Foundation and the Royal Botanic Gardens Victoria Foundation. Bruce Wannan, John Clarkson, John Neldner, and Eda Addicot are thanked for their thoughts on the relationships between the two species. Open access publishing facilitated by The University of Melbourne, as part of the Wiley - The University of Melbourne agreement via the Council of Australian University Librarians.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Supporting data and files for this manuscript are now publicly available here: <https://doi.org/10.26188/27997571.v1>.

References

- Ahrens, C. W., P. D. Rymer, A. Stow, et al. 2018. “The Search for Loci Under Selection: Trends, Biases and Progress.” *Molecular Ecology* 27: 1342–1356. <https://doi.org/10.1111/mec.14549>.
- ALA. 2024. *Atlas of Living Australia*. ALA. <https://bie.ala.org.au/species/https://biodiversity.org.au/afd/taxa/21e61335-d76e-454f-8329-4cd7dfd31ee0>.
- APC. 2024. *Australian Plant Census. The Australian National Species List*. APC. <https://biodiversity.org.au/nsl/services/search/taxonomy>.
- AVH. 2023. *The Australasian Virtual Herbarium, Council of Heads of Australasian Herbaria*. AVH. <https://avh.chah.org.au>.
- Barnes, R., K. Sahr, G. Evenden, A. Johnson, and F. Warmerdam. 2017. dggridR: discrete global grids for R. R Package Version 01 12, 963. <https://github.com/r-barnes/dggridR/>.
- Barton, N., and B. O. Bengtsson. 1986. “The Barrier to Genetic Exchange Between Hybridising Populations.” *Heredity* 57: 357–376.
- Bayly, M. J. 2016. “Phylogenetic Studies of Eucalypts: Fossils, Morphology and Genomes.” *Proceedings of the Royal Society of Victoria* 128, no. 1: 12–24.
- Bayly, M. J., T. G. McLay, D. J. Cantrill, and R. M. Fowler. 2023. “A New Classification of Sections and Series in *Eucalyptus* Subg. *Eudesmia* (Myrtaceae).” *Muelleria* 41: 50–54.

- Bean, A. R. 2006. “*Eucalyptus Megasepala* AR Bean (Myrtaceae), a New Species From Queensland Allied to *E. Tetrodonta* F. Muell.” *Austrobaileya* 7, no. 2: 305–310.
- Biggs, A. J. W., and S. R. Philip. 1994. *Soils of Cape York Peninsula*. Queensland Department of Primary Industries, Land Resources Bulletin No. QV95001.
- Binks, R. M., D. A. Steane, and M. Byrne. 2021. “Genomic Divergence in Sympatry Indicates Strong Reproductive Barriers and Cryptic Species Within *Eucalyptus salubris*.” *Ecology and Evolution* 11: 5096–5110. <https://doi.org/10.1002/ece3.7403>.
- Bird, C. E., I. Fernandez-Silva, D. J. Skillings, and R. J. Toonen. 2012. “Sympatric Speciation in the Post “Modern Synthesis” Era of Evolutionary Biology.” *Evolutionary Biology* 39: 158–180.
- Bock, D. G., Z. Cai, C. Elphinstone, et al. 2023. “Genomics of Plant Speciation.” *Plant Communications* 4: 100599. <https://doi.org/10.1016/j.xplc.2023.100599>.
- Bowman, D. M., G. K. Brown, M. F. Braby, et al. 2010. “Biogeography of the Australian Monsoon Tropics.” *Journal of Biogeography* 37, no. 2: 201–216.
- Bradford, M., M. Venz, K. L. Bell, et al. 2022. “The Diet of a Specialist Nectarivore in Australia: The Little Red Flying-Fox (*Pteropus scapulatus*, Pteropodidae).” *Austral Ecology* 47: 619–628. <https://doi.org/10.1111/aec.13143>.
- Buck, R., and L. Flores-Rentería. 2022. “The Syngameon Enigma.” *Plants* 11: 895. <https://doi.org/10.3390/plants11070895>.
- Buck, R., D. Ortega-Del Vecchyo, C. Gehring, et al. 2023. “Sequential Hybridization May Have Facilitated Ecological Transitions in the Southwestern Pinyon Pine Syngameon.” *New Phytologist* 237: 2435–2449.
- Cannon, C. H., and R. J. Petit. 2020. “The Oak Syngameon: More Than the Sum of Its Parts.” *New Phytologist* 226: 978–983.
- Carbon, S., A. Ireland, C. J. Mungall, S. Shu, B. Marshall, and S. Lewis. 2009. “AmiGO: Online Access to Ontology and Annotation Data.” *Bioinformatics* 25, no. 2: 288–289. <https://doi.org/10.1093/bioinformatics/btn615>.
- Cattell, R. B. 1966. “The Scree Test for the Number of Factors.” *Multivariate Behavioral Research* 1, no. 2: 245–276. https://doi.org/10.1207/s15327906mbr0102_10.
- Catullo, R. A., R. Lanfear, P. Doughty, and J. S. Keogh. 2014. “The Biogeographical Boundaries of Northern Australia: Evidence From Ecological Niche Models and a Multi-Locus Phylogeny of *Uperoleia* Toadlets (Anura: Myobatrachidae).” *Journal of Biogeography* 41, no. 4: 659–672.
- Cavender-Bares, J., A. González-Rodríguez, D. A. Eaton, A. A. Hipp, A. Beulke, and P. S. Manos. 2015. “Phylogeny and Biogeography of the American Live Oaks (*Quercus* Subsection *Virentes*): A Genomic and Population Genetics Approach.” *Molecular Ecology* 24: 3668–3687.
- Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. “Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets.” *GigaScience* 4, no. 1: s13742. <https://doi.org/10.1186/s13742-015-0047-8>.
- Chen, G. F. 2013. “Sexual Isolation in Two Bee-Pollinated *Costus* (Costaceae).” *Plant Reproduction* 26: 3–16. <https://doi.org/10.1007/s00497-012-0203-6>.
- Chifman, J., and L. Kubatko. 2014. “Quartet Inference From SNP Data Under the Coalescent Model.” *Bioinformatics* 30: 3317–3324.
- Chivas, A. R., A. García, S. Van Der Kaars, et al. 2001. “Sea-Level and Environmental Changes Since the Last Interglacial in the Gulf of Carpentaria, Australia: An Overview.” *Quaternary International* 83: 19–46.
- Collins, T. L., R. L. Andrew, and J. J. Bruhl. 2019. “Morphological, Phytochemical and Molecular Analyses Define Species Limits in *Eucalyptus magnificata* (Myrtaceae) and Lead to the Discovery of a New Rare Species.” *Australian Systematic Botany* 32: 12–28.
- Cremona, T., A. M. Baker, S. J. B. Cooper, R. Montague-Drake, A. M. Stobo-Wilson, and S. M. Carthew. 2021. “Integrative Taxonomic Investigation of *Petaurus breviceps* (Marsupialia: Petauridae) Reveals Three Distinct Species.” *Zoological Journal of the Linnean Society* 191, no. 2: 503–527. <https://doi.org/10.1093/zoolinnean/zlaa060>.
- Crowl, A. A., P. S. Manos, J. D. McVay, A. R. Lemmon, E. M. Lemmon, and A. L. Hipp. 2020. “Uncovering the Genomic Signature of Ancient Introgression Between White Oak Lineages (*Quercus*).” *New Phytologist* 226: 1158–1170. <https://doi.org/10.1111/nph.15842>.
- De Deckker, P. 2001. “Late Quaternary Cyclic Aridity in Tropical Australia.” *Palaeogeography, Palaeoclimatology, Palaeoecology* 170: 1–9. [https://doi.org/10.1016/S0031-0182\(01\)00233-4](https://doi.org/10.1016/S0031-0182(01)00233-4).
- Denaro, T. J., and G. R. Ewers. 1995. *Mineral Resource Assessment – Cape York Peninsula Land Use Strategy*. Queensland Minerals and Energy Review Series, Department of Minerals and Energy, Queensland.
- Earl, D. A., and B. M. VonHoldt. 2012. “STRUCTURE HARVESTER: A Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method.” *Conservation Genetics Resources* 4: 359–361.
- Edet, O. U., Y. S. A. Gorafi, S. Nasuda, and H. Tsujimoto. 2018. “DarTseq-Based Analysis of Genomic Relationships Among Species of Tribe Triticeae.” *Scientific Reports* 8: 1–11. <https://doi.org/10.1038/s41598-018-34811-y>.
- Edwards, R. D., M. D. Crisp, D. H. Cook, and L. G. Cook. 2017. “Congruent Biogeographical Disjunctions at a Continent-Wide Scale: Quantifying and Clarifying the Role of Biogeographic Barriers in the Australian Tropics.” *PLoS One* 12: e0174812.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. “Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study.” *Molecular Ecology* 14: 2611–2620.
- Fahey, P. S., R. M. Fowler, F. Udovicic, D. J. Cantrill, and M. J. Bayly. 2021. “Use of Plastid Genome Sequences in Phylogeographic Studies of Single Species Can Be Misleading Without Comprehensive Sampling of Co-Occurring, Related Species.” *Tree Genetics & Genomes* 17: 43.
- Fahey, P. S., M. Rossetto, J. J. Bruhl, R. Andrew, and D. Andrew. 2024. “A Genetic Investigation of Relationships and Species Boundaries Between *Eucalyptus Dalrympleana* and Allied Taxa in New South Wales.” *Telopea* 27: 177–196.
- Fahey, P. S., F. Udovicic, D. J. Cantrill, and M. J. Bayly. 2022a. “A Box on the River: The Phylogenetics and Phylogeography of *Eucalyptus Baueriana* (*Eucalyptus* Sect. *Adnataria* Ser. *Heterophloiae*).” *PLoS One* 17, no. 11: e0276117.
- Fahey, P. S., F. Udovicic, D. J. Cantrill, D. Nicolle, T. G. B. McLay, and M. J. Bayly. 2022b. “A Phylogenetic Investigation of the Taxonomically Problematic *Eucalyptus Odorata* Complex (*E.* Sect. *Adnataria* Ser. *Subbuxaeales*): Evidence for Extensive Interspecific Gene Flow and Reticulate Evolution.” *Australian Systematic Botany* 35: 403–435.
- Fenker, J., J. Melville, and C. Moritz. 2024. “Dragons in the Tropics – Phylogeography and Speciation in *Diporiphora* Lizards and Common Geographic Breaks in Co-Distributed Taxa.” *Molecular Phylogenetics and Evolution* 197: 108090.
- Ferguson, S., A. Jones, K. Murray, R. Andrew, B. Schwessinger, and J. Borevitz. 2024a. “Plant Genome Evolution in the Genus *Eucalyptus* Is Driven by Structural Rearrangements That Promote Sequence Divergence.” *Genome Research* 34: 606–619.

- Ferguson, S., A. Jones, K. Murray, et al. 2024b. "Exploring the Role of Polymorphic Interspecies Structural Variants in Reproductive Isolation and Adaptive Divergence in *Eucalyptus*." *GigaScience* 13: giae029.
- Ferguson, S., A. Jones, K. Murray, B. Schwessinger, and J. O. Borevitz. 2023. "Interspecies Genome Divergence Is Predominantly due to Frequent Small Scale Rearrangements in *Eucalyptus*." *Molecular Ecology* 32: 1271–1287.
- Fick, S. E., and R. J. Hijmans. 2017. "WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37: 4302–4315.
- Field, D. L., D. J. Ayre, R. J. Whelan, and A. G. Young. 2011. "Patterns of Hybridization and Asymmetrical Gene Flow in Hybrid Zones of the Rare *Eucalyptus aggregata* and Common *E. rubida*." *Heredity* 106: 841–853.
- Fitzpatrick, B. M., J. A. Fordyce, and S. Gavrillets. 2008. "What, if Anything, Is Sympatric Speciation?" *Journal of Evolutionary Biology* 21: 1452–1459.
- Foll, M., and O. Gaggiotti. 2008. "A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective." *Genetics* 180: 977–993. <https://doi.org/10.1534/genetics.108.092221>.
- Foster, S. A., G. E. McKinnon, D. A. Steane, B. M. Potts, and R. E. Vaillancourt. 2007. "Parallel Evolution of Dwarf Ecotypes in the Forest Tree *Eucalyptus globulus*." *New Phytologist* 175: 370–380.
- Fox, I. D., V. J. Neldner, G. W. Wilson, and P. J. Bannink. 2001. *The Vegetation of the Australian Tropical Savannas*. Environment Protection Agency.
- Franklin, D. 2022. *A Field Study of the Eucalypts of North-East Queensland*. DC Franklin. <https://www.nqnhg.org/eucalyptsneq>.
- Franklin, D. C., R. M. Gunton, J. O. N. Schatz, and M. J. Lawes. 2010. "Resprouting Responses of Trees in a Fire-Prone Tropical Savanna Following Severe Tornado Damage." *Austral Ecology* 35, no. 6: 685–694.
- Franklin, D. C., and R. A. Noske. 2000. "Nectar Sources Used by Birds in Monsoonal North-Western Australia: A Regional Survey." *Australian Journal of Botany* 48, no. 4: 461–474.
- Frantz, A. C., S. Cellina, A. Krier, L. Schley, and T. Burke. 2009. "Using Spatial Bayesian Methods to Determine the Genetic Structure of a Continuously Distributed Population: Clusters or Isolation by Distance?" *Journal of Applied Ecology* 46: 493–505.
- Frichot, E., and O. François. 2015. "LEA: An R Package for Landscape and Ecological Association Studies." *Methods in Ecology and Evolution* 6: 925–929.
- Gene Ontology Consortium. 2004. "The Gene Ontology (GO) Database and Informatics Resource." *Nucleic Acids Research* 32: D258–D261. <https://doi.org/10.1093/nar/gkh036>.
- Givnish, T. J. 2010. "Ecology of Plant Speciation." *Taxon* 59, no. 5: 1329–1366. <https://doi.org/10.1002/tax.595003>.
- González-Orozco, C. E., A. H. Thornhill, N. Knerr, S. Laffan, and J. T. Miller. 2014. "Biogeographical Regions and Phytogeography of the Eucalypts." *Diversity and Distributions* 20: 46–58.
- Griffin, A. R., I. P. Burgess, and L. Wolf. 1988. "Patterns of Natural and Manipulated Hybridisation in the Genus *Eucalyptus* L'herit.—1 A Review." *Australian Journal of Botany* 36: 41–66.
- Hardig, T. M., S. J. Brunsfeld, R. S. Fritz, M. Morgan, and C. M. Orians. 2000. "Morphological and Molecular Evidence for Hybridization and Introgression in a Willow (*Salix*) Hybrid Zone." *Molecular Ecology* 9: 9–24. <https://doi.org/10.1046/j.1365-294X.2000.00757.x>.
- Hernández-Hernández, T., E. C. Miller, C. Román-Palacios, and J. J. Wiens. 2021. "Speciation Across the Tree of Life." *Biological Reviews* 96: 1205–1242. <https://doi.org/10.1111/brv.12698>.
- Hipp, A. L., P. S. Manos, M. Hahn, et al. 2020. "Genomic Landscape of the Global Oak Phylogeny." *New Phytologist* 226: 1198–1212.
- Hipp, A. L., A. T. Whittemore, M. Garner, et al. 2019. "Genomic Identity of White Oak Species in an Eastern North American Syngameon1." *Annals of the Missouri Botanical Garden* 104: 455–477.
- de Holanda, A. S. S., A. Vicentini, M. J. G. Hopkins, and C. E. Zartman. 2015. "Phenotypic Differences Are Not Explained by Pre-Zygotic Reproductive Barriers in Sympatric Varieties of the Humiria Balsamifera Complex (Humiriaceae)." *Plant Systematics and Evolution* 301: 1767–1779. <https://doi.org/10.1007/s00606-014-1191-9>.
- Holman, J. E., J. M. Hughes, and R. J. Fensham. 2011. "Origins of a Morphological Cline Between *Eucalyptus melanophloia* and *Eucalyptus whitei*." *Australian Journal of Botany* 59: 244–252.
- Hope, G., A. P. Kershaw, S. van der Kaars, et al. 2004. "History of Vegetation and Habitat Change in the Austral-Asian Region." *Quaternary International* 118: 103–126.
- Huntley, R. P., M. A. Harris, Y. Alam-Faruque, et al. 2014. "A Method for Increasing Expressivity of Gene Ontology Annotations Using a Compositional Approach." *BMC Bioinformatics* 15: 155. <https://doi.org/10.1186/1471-2105-15-155>.
- Huson, D. H., and D. Bryant. 2006. "Application of Phylogenetic Networks in Evolutionary Studies." *Molecular Biology and Evolution* 23: 254–267.
- Jackson, H. D., D. A. Steane, B. M. Potts, and R. E. Vaillancourt. 1999. "Chloroplast DNA Evidence for Reticulate Evolution in *Eucalyptus* (Myrtaceae)." *Molecular Ecology* 8: 739–751.
- Jombart, T. 2008. "Adegenet: A R Package for the Multivariate Analysis of Genetic Markers." *Bioinformatics* 24: 1403–1405.
- Kautt, A. F., C. F. Kratochwil, A. Nater, et al. 2020. "Contrasting Signatures of Genomic Divergence During Sympatric Speciation." *Nature* 588: 106–111.
- Keim, P., K. N. Paige, T. G. Whitham, and K. G. Lark. 1989. "Genetic Analysis of an Interspecific Hybrid Swarm of *Populus*: Occurrence of Unidirectional Introgression." *Genetics* 123: 557–565. <https://doi.org/10.1093/genetics/123.3.557>.
- Kidner, J., P. Theodorou, J. O. Engler, M. Taubert, and M. Husemann. 2021. "A Brief History and Popularity of Methods and Tools Used to Estimate Micro-Evolutionary Forces." *Ecology and Evolution* 11: 13723–13743. <https://doi.org/10.1002/ece3.8076>.
- Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. "Clumpak: A Program for Identifying Clustering Modes and Packaging Population Structure Inferences Across K." *Molecular Ecology Resources* 15: 1179–1191.
- Lacey, C. J., and P. I. Whelan. 1976. "Observations on the Ecological Significance of Vegetative Reproduction in the Katherine-Darwin Region of the Northern Territory." *Australian Forestry* 39: 131–139.
- Ladiges, P. Y. 1997. "Phylogenetic History and Classification of Eucalypts." In *Eucalypt Ecology Individuals to Ecosystems*, edited by J. Williams and J. Woinarski, 16–29. Cambridge University Press.
- Larcombe, M. J., B. Holland, D. A. Steane, et al. 2015. "Patterns of Reproductive Isolation in *Eucalyptus*—A Phylogenetic Perspective." *Molecular Biology and Evolution* 32: 1833–1846.
- Law, B., C. Mackowski, L. Schoer, and T. Tweedie. 2000. "Flowering Phenology of Myrtaceous Trees and Their Relation to Climatic, Environmental and Disturbance Variables in Northern New South Wales." *Austral Ecology* 25: 160–178. <https://doi.org/10.1046/j.1442-9993.2000.01009.x>.
- Levin, D. A. 2006. "Flowering Phenology in Relation to Adaptive Radiation." *Systematic Botany* 31: 239–246. <https://doi.org/10.1600/03636440677585928>.

- Li, X., G. Wei, Y. A. El-Kassaby, and Y. Fang. 2021. "Hybridization and Introgression in Sympatric and Allopatric Populations of Four Oak Species." *BMC Plant Biology* 21: 266. <https://doi.org/10.1186/s12870-021-03007-4>.
- López de Heredia, U., F. Mora-Márquez, P. G. Goicoechea, L. Guillardín-Calvo, M. C. Simeone, and Á. Soto. 2020. "ddRAD Sequencing-Based Identification of Genomic Boundaries and Permeability in *Quercus Ilex* and *Q. suber* Hybrids." *Frontiers in Plant Science* 11: 564414. <https://doi.org/10.3389/fpls.2020.564414>.
- Luu, K., E. Bazin, and M. G. B. Blum. 2017. "Pcadapt: An R Package to Perform Genome Scans for Selection Based on Principal Component Analysis." *Molecular Ecology Resources* 17: 67–77. <https://doi.org/10.1111/1755-0998.12592>.
- Ma, X. G., Y. B. Ren, and H. Sun. 2024. "Introgression and Incomplete Lineage Sorting Blurred Phylogenetic Relationships Across the Genomes of Sclerophyllous Oaks From Southwest China." *Cladistics* 40: 357–373.
- Mallet, J., A. Meyer, P. Nosil, and J. L. Feder. 2009. "Space, Sympatry and Speciation." *Journal of Evolutionary Biology* 22: 2332–2341. <https://doi.org/10.1111/j.1420-9101.2009.01816.x>.
- Manawaduge, C. G., A. R. Clarke, and D. A. Hurwood. 2023. "Divergent East-West Lineages in an Australian Fruit Fly (*Bactrocera jarvisi*), associated With the Carpentaria Basin Divide." *PLoS One* 18: e0276247.
- Marcus, J., W. Ha, R. F. Barber, and J. Novembre. 2021. "Fast and Flexible Estimation of Effective Migration Surfaces." *eLife* 10: e61927.
- McGowen, M. H., R. J. E. Wiltshire, B. M. Potts, and R. E. Vaillancourt. 2001. "The Origin of *Eucalyptus vernicosa*, a Unique Shrub Eucalypt." *Biological Journal of the Linnean Society* 74: 397–405.
- McKinnon, G. E., G. J. Jordan, R. E. Vaillancourt, D. A. Steane, and B. M. Potts. 2004. "Glacial Refugia and Reticulate Evolution: The Case of the Tasmanian Eucalypts." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 359: 275–284.
- McKinnon, G. E., J. J. Smith, and B. M. Potts. 2010. "Recurrent Nuclear DNA Introgression Accompanies Chloroplast DNA Exchange Between Two Eucalypt Species." *Molecular Ecology* 19: 1367–1380.
- McKinnon, G. E., D. A. Steane, B. M. Potts, and R. E. Vaillancourt. 1999. "Incongruence Between Chloroplast and Species Phylogenies in *Eucalyptus* Subgenus *Monocalyptus* (Myrtaceae)." *American Journal of Botany* 86: 1038–1046.
- McKinnon, G. E., R. E. Vaillancourt, H. D. Jackson, and B. M. Potts. 2001. "Chloroplast Sharing in the Tasmanian Eucalypts." *Evolution* 55: 703–711.
- McLay, T. G. B., R. M. Fowler, P. S. Fahey, et al. 2023. "Phylogenomics Reveals Extreme Gene Tree Discordance in a Lineage of Dominant Trees: Hybridization, Introgression, and Incomplete Lineage Sorting Blur Deep Evolutionary Relationships Despite Clear Species Groupings in *Eucalyptus* Subgenus *Eudesmia*." *Molecular Phylogenetics and Evolution* 187: 107869.
- McVay, J. D., D. Hauser, A. L. Hipp, and P. S. Manos. 2017. "Phylogenomics Reveals a Complex Evolutionary History of Lobed-Leaf White Oaks in Western North America." *Genome* 60: 733–742.
- Meirmans, P. G. 2012. "The Trouble With Isolation by Distance." *Molecular Ecology* 21: 2839–2846.
- Melville, J., M. L. Haines, K. Boysen, et al. 2017. "Identifying Hybridization and Admixture Using SNPs: Application of the DArTseq Platform in Phylogeographic Research on Vertebrates." *Royal Society Open Science* 4: 161061. <https://doi.org/10.1098/rsos.161061>.
- Mijangos, J. L., B. Gruber, O. Berry, C. Pacioni, and A. Georges. 2022. "dartR v2: An Accessible Genetic Analysis Platform for Conservation, Ecology and Agriculture." *Methods in Ecology and Evolution* 13: 2150–2158.
- Moritz, C., M. K. Fujita, D. Rosauer, et al. 2016. "Multilocus Phylogeography Reveals Nested Endemism in a Gecko Across the Monsoonal Tropics of Australia." *Molecular Ecology* 25, no. 6: 1354–1366.
- Murray, K. D., J. K. Janes, A. Jones, H. M. Bothwell, R. L. Andrew, and J. O. Borevitz. 2019. "Landscape Drivers of Genomic Diversity and Divergence in Woodland *Eucalyptus*." *Molecular Ecology* 28: 5232–5247.
- Myburg, A. A., D. Grattapaglia, G. A. Tuskan, et al. 2014. "The Genome of *Eucalyptus grandis*." *Nature* 510: 356–362.
- Nevill, P. G., T. Despres, M. J. Bayly, G. Bossinger, and P. K. Ades. 2014. "Shared Phylogeographic Patterns and Widespread Chloroplast Haplotype Sharing in *Eucalyptus* Species With Different Ecological Tolerances." *Tree Genetics & Genomes* 10, no. 4: 1079–1092.
- Nicolle, D. 2024. Classification of the Eucalypts, Genus *Eucalyptus*. Version 7. <http://www.dn.com.au/Classification-Of-The-Eucalypts.pdf>.
- Osborne, O. G., A. Ciezarek, T. Wilson, et al. 2019. "Speciation in *Howea* Palms Occurred in Sympatry, Was Preceded by Ancestral Admixture, and Was Associated With Edaphic and Phenological Adaptation." *Molecular Biology and Evolution* 36: 2682–2697.
- Osborne, O. G., T. Kafle, T. Brewer, M. P. Dobрева, I. Hutton, and V. Savolainen. 2020. "Sympatric Speciation in Mountain Roses (*Metrosideros*) on an Oceanic Island." *Philosophical Transactions of the Royal Society B* 375: 20190542.
- Osmolovsky, I., M. Shifrin, I. Gamliel, J. Belmaker, and Y. Sapir. 2022. "Eco-Geography and Phenology Are the Major Drivers of Reproductive Isolation in the Royal Irises, a Species Complex in the Course of Speciation." *Plants* 11: e3306.
- Pain, C. F., J. R. Wilford, and J. C. Dohrenwend. 1995. *Regolith-Terrain Mapping of Cape York Peninsula*. Cape York Peninsula Land Use Strategy, Office of the Co-ordinator General of Queensland, Brisbane, Department of the Environment, Sport and Territories, Canberra, and Australian Geological Survey Organisation, Canberra.
- Palsson, R. L., I. R. H. Telford, J. J. Bruhl, and R. L. Andrew. 2024. "Population Genetic Structure and Range Limits of *Prostanthera Cineolifera* (Lamiaceae), a Vulnerable Shrub With a Patchy Distribution." *Conservation Genetics* 25: 1231–1251. <https://doi.org/10.1007/s10592-024-01637-3>.
- Papadopoulos, A. S., W. J. Baker, D. Crayn, et al. 2011. "Speciation With Gene Flow on Lord Howe Island." *Proceedings of the National Academy of Sciences of the United States of America* 108: 13188–13193.
- Parsons, R. F. 1969. "Physiological and Ecological Tolerances of *Eucalyptus incrassata* and *E. socialis* to Edaphic Factors." *Ecology* 50: 386–390.
- Passioura, J. A., and J. E. Ash. 1993. "Phenotypic, Genetic and Ecological Variation in the *Eucalyptus saligna*-*E. botryoides* Complex." *Australian Journal of Botany* 41: 393–412.
- Pellegrino, G. 2016. "Sympatric Reinforcement of Reproductive Barriers Between *Neotinea tridentata* and *N. ustulata* (Orchidaceae)." *Journal of Plant Research* 129: 1061–1068. <https://doi.org/10.1007/s10265-016-0855-7>.
- Perez, M. F., F. F. Franco, J. R. Bombonato, et al. 2018. "Assessing Population Structure in the Face of Isolation by Distance: Are We Neglecting the Problem?" *Diversity and Distributions* 24: 1883–1889.
- Pin, P. A., and O. Nilsson. 2012. "The Multifaceted Roles of FLOWERING LOCUS T in Plant Development." *Plant, Cell & Environment* 35: 1742–1755.
- Pollock, L. J., M. J. Bayly, P. G. Nevill, and P. A. Vesk. 2013. "Chloroplast DNA Diversity Associated With Protected Slopes and Valleys for Hybridizing *Eucalyptus* Species on Isolated Ranges in South-Eastern Australia." *Journal of Biogeography* 40, no. 1: 155–167.

- Pollock, L. J., M. J. Bayly, and P. A. Vesk. 2015. "The Roles of Ecological and Evolutionary Processes in Plant Community Assembly: The Environment, Hybridization and Introgression Influence Co-Occurrence of *Eucalyptus*." *American Naturalist* 185: 784–796.
- Potter, S., J. G. Bragg, B. M. Peter, K. Bi, and C. Moritz. 2016. "Phylogenomics at the Tips: Inferring Lineages and Their Demographic History in a Tropical Lizard, *Carlia amax*." *Molecular Ecology* 25: 1367–1380.
- Potter, S., M. D. Eldridge, D. A. Taggart, and S. J. Cooper. 2012. "Multiple Biogeographical Barriers Identified Across the Monsoon Tropics of Northern Australia: Phylogeographic Analysis of the *Brachyotis* Group of Rock-Wallabies." *Molecular Ecology* 21, no. 9: 2254–2269.
- Potter, S., C. Moritz, M. P. Piggott, et al. 2024. "Museum Skins Enable Identification of Introgression Associated With Cytonuclear Discordance." *Systematic Biology* 73: syae016.
- Potts, B. M., and J. B. Reid. 1985. "Variation in the *Eucalyptus Gunnii-Archeri* Complex. II. The Origin of Variation." *Australian Journal of Botany* 33, no. 5: 519–541.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155: 945–959.
- Privé, F., K. Luu, B. J. Vilhjálmsson, and M. G. B. Blum. 2020. "Performing Highly Efficient Genome Scans for Local Adaptation With R Package Pcadapt Version 4." *Molecular Biology and Evolution* 37: 2153–2154. <https://doi.org/10.1093/molbev/msaa053>.
- Pryor, L. D. 1953. "Genetic Control in *Eucalyptus* Distributions." *Proceedings of the Linnean Society of New South Wales* 78: 8–18.
- Purcell, S., B. Neale, K. Todd-Brown, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81: 559–575.
- Ravinet, M., R. Faria, R. K. Butlin, et al. 2017. "Interpreting the Genomic Landscape of Speciation: A Road Map for Finding Barriers to Gene Flow." *Journal of Evolutionary Biology* 30: 1450–1477.
- Rawal, D. S., S. Kasel, M. R. Keatley, and C. R. Nitschke. 2015a. "Climatic and Photoperiodic Effects on Flowering Phenology of Select Eucalypts From South-Eastern Australia." *Agricultural and Forest Meteorology* 214: 231–242. <https://doi.org/10.1016/j.agrformet.2015.08.253>.
- Rawal, D. S., S. Kasel, M. R. Keatley, and C. R. Nitschke. 2015b. "Herbarium Records Identify Sensitivity of Flowering Phenology of Eucalypts to Climate: Implications for Species Response to Climate Change." *Austral Ecology* 40: 117–125. <https://doi.org/10.1111/aec.12183>.
- Reaz, R., M. S. Bayzid, and M. S. Rahman. 2014. "Accurate Phylogenetic Tree Reconstruction From Quartets: A Heuristic Approach." *PLoS One* 9: e104008.
- Reeves, J. M., A. R. Chivas, A. García, et al. 2008. "The Sedimentary Record of Palaeoenvironments and Sea-Level Change in the Gulf of Carpentaria, Australia, Through the Last Glacial Cycle." *Quaternary International* 183, no. 1: 3–22.
- Richards, E. J., M. R. Servedio, and C. H. Martin. 2019. "Searching for Sympatric Speciation in the Genomic Era." *BioEssays* 41: 1900047.
- Rieseberg, L. H., and S. E. Carney. 1998. "Tansley Review No. 102: Plant Hybridization." *New Phytologist* 140: 599–624. <https://doi.org/10.1046/j.1469-8137.1998.00315.x>.
- Robins, T. P., R. M. Binks, M. Byrne, and S. D. Hopper. 2021. "Landscape and Taxon Age Are Associated With Differing Patterns of Hybridization in Two *Eucalyptus* (Myrtaceae) Subgenera." *Annals of Botany* 127: 49–62.
- Rossel, R. V., C. Chen, M. J. Grundy, R. Searle, D. Clifford, and P. H. Campbell. 2015. "The Australian Three-Dimensional Soil Grid: Australia's Contribution to the GlobalSoilMap Project." *Soil Research* 53: 845–864.
- Rosser, N. L., A. Quinton, H. Davey, D. J. Ayre, and A. J. Denham. 2023. "Genetic Assessment of the Value of Restoration Planting Within an Endangered Eucalypt Woodland." *Scientific Reports* 13: 1–11. <https://doi.org/10.1038/s41598-023-33720-z>.
- Rutherford, S. 2020. "Insights Into Speciation and Species Delimitation of Closely Related Eucalypts Using an Interdisciplinary Approach." *Australian Systematic Botany* 33, no. 1: 110–127. <https://doi.org/10.1071/SB18042>.
- Rutherford, S., M. Rossetto, J. G. Bragg, et al. 2018. "Speciation in the Presence of Gene Flow: Population Genomics of Closely Related and Diverging *Eucalyptus* Species." *Heredity* 121: 126–141.
- Rutherford, S., M. Rossetto, J. G. Bragg, and J. S. Wan. 2023. "Where to Draw the Boundaries? Using Landscape Genomics to Disentangle the Scribbly Gum Species Complex." *American Journal of Botany* 110, no. 11: e16245.
- Rutherford, S., M. van Der Merwe, P. G. Wilson, R. M. Kooyman, and M. Rossetto. 2019. "Managing the Risk of Genetic Swamping of a Rare and Restricted Tree." *Conservation Genetics* 20: 1113–1131.
- Rutherford, S., J. S. Wan, J. M. Cohen, D. Benson, and M. Rossetto. 2021. "Looks Can Be Deceiving: Speciation Dynamics of Co-Distributed *Angophora* (Myrtaceae) Species in a Varying Landscape." *Evolution* 75, no. 2: 310–329.
- Sanderson, B. J., D. Ghambir, G. Feng, et al. 2023. "Phylogenomics Reveals Patterns of Ancient Hybridization and Differential Divergence That Contribute to Phylogenetic Conflict in Willows, Poplars, and Close Relatives." *Systematic Biology* 72: 1220–1232. <https://doi.org/10.1093/sysbio/syad042>.
- Schuster, T. M., S. D. Setaro, J. F. Tibbits, et al. 2018. "Chloroplast Variation Is Incongruent With Classification of the Australian Bloodwood Eucalypts (Genus *Corymbia*, Family Myrtaceae)." *PLoS One* 13, no. 4: e0195034.
- Setterfield, S. A., and J. R. Williams. 1996. "Patterns of Flowering and Seed Production in *Eucalyptus Miniata* and *E. Tetrodonta* in a Tropical Savanna Woodland, Northern Australia." *Australian Journal of Botany* 44, no. 2: 107–122. <https://doi.org/10.1071/BT9960107>.
- Slee, A. V., M. I. H. Brooker, S. M. Duffy, and J. G. West. 2020. *EUCLID Eucalypts of Australia*. 4th ed. <https://apps.lucidcentral.org/euclid/text/intro/about.htm>. Centre for Australian National Biodiversity Research.
- Sobel, J. M., G. F. Chen, L. R. Watt, and D. W. Schemske. 2010. "The Biology of Speciation." *Evolution* 64: 295–315.
- Song, S., Y. Chen, L. Liu, et al. 2018. "OsFTIP7 Determines Auxin-Mediated Anther Dehiscence in Rice." *Nature Plants* 4: 495–504. <https://doi.org/10.1038/s41477-018-0175-0>.
- Southern, S. G., P. Birt, J. Porter, and H. A. Ford. 2004. "Review of Gene Movement by Bats and Birds and Its Potential Significance for Eucalypt Plantation Forestry." *Australian Forestry* 67, no. 1: 44–53. <https://doi.org/10.1080/00049158.2004.10676205>.
- Steane, D. A., B. M. Potts, E. McLean, et al. 2015. "Genome-Wide Scans Reveal Cryptic Population Structure in a Dry-Adapted Eucalypt." *Tree Genetics & Genomes* 11: 33. <https://doi.org/10.1007/s11295-015-0864-z>.
- Storey, J. D., A. J. Bass, A. Dabney, and D. Robinson. 2019. qvalue: Q-value estimation for false discovery rate control. R package version 2.30.0. <http://github.com/jdstorey/qvalue>.
- Sun, N., L. Yang, F. Tian, et al. 2022. "Sympatric or Micro-Allopatric Speciation in a Glacial Lake? Genomic Islands Support Neither." *National Science Review* 9: nwac291.
- Sutra, N., J. Kusumi, J. Montenegro, et al. 2019. "Evidence for Sympatric Speciation in a Wallacean Ancient Lake." *Evolution* 73, no. 9: 1898–1915. <https://doi.org/10.1111/evo.13821>.

- Swofford, D. L. 2002. *PAUP: Phylogenetic Analysis Using Parsimony (And Other Methods)*, version 4. Sinauer Associates.
- von Takach Dukai, B., C. Jack, J. Borevitz, D. B. Lindenmayer, and S. C. Banks. 2019. "Pervasive Admixture Between Eucalypt Species Has Consequences for Conservation and Assisted Migration." *Evolutionary Applications* 12: 845–860.
- Tate, J. A., J. Fuertes Aguilar, S. J. Wagstaff, J. C. La Duke, T. A. Bodo Slotta, and B. B. Simpson. 2005. "Phylogenetic Relationships Within the Tribe Malveae (Malvaceae, Subfamily Malvoideae) as Inferred From ITS Sequence Data." *American Journal of Botany* 92: 584–602.
- Tavares, M. M., M. Ferro, B. S. S. Leal, and C. Palma-Silva. 2022. "Speciation With Gene Flow Between Two Neotropical Sympatric Species (*Pitcairnia* spp.: Bromeliaceae)." *Ecology and Evolution* 12: e8834.
- The Gene Ontology Consortium. 2019. "The Gene Ontology Resource: 20 Years and Still GOing Strong." *Nucleic Acids Research* 47: D330–D338. <https://doi.org/10.1093/nar/gky1055>.
- Thorsson, A. T., S. Pálsson, A. Sigurgeirsson, and K. Anamthawat-Jonsson. 2007. "Morphological Variation Among *Betula Nana* (Diploid), *B. pubescens* (Tetraploid) and Their Triploid Hybrids in Iceland." *Annals of Botany* 99: 1183–1193. <https://doi.org/10.1093/aob/mcm060>.
- Umbrello, L. S., H. Newton, A. M. Baker, K. J. Travouillon, and M. Westerman. 2024. "Vicariant Speciation Resulting From Biogeographic Barriers in the Australian Tropics: The Case of the Red-Cheeked Dunnart (*Sminthopsis virginiae*)." *Ecology and Evolution* 14: e70215. <https://doi.org/10.1002/ece3.70215>.
- Vardon, M. J., and C. R. Tidemann. 1999. "Flying-Foxes (*Pteropus Alecto* and *P. Scapulatus*) in the Darwin Region, North Australia: Patterns in Camp Size and Structure." *Australian Journal of Zoology* 47, no. 4: 411–423. <https://doi.org/10.1071/ZO99022>.
- Vašut, R. J., M. Pospíšková, J. Lukavský, and J. Weger. 2024. "Detection of Hybrids in Willows (*Salix*, Salicaceae) Using Genome-Wide DArTseq Markers." *Plants* 13: 639. <https://doi.org/10.3390/plants13050639>.
- Wang, H., H. Yin, C. Jiao, et al. 2020. "Sympatric Speciation of Wild Emmer Wheat Driven by Ecology and Chromosomal Rearrangements." *Proceedings of the National Academy of Sciences* 117: 5955–5963. <https://doi.org/10.1073/pnas.1920415117>.
- Wannan, B. 2014. "Review of the Phytogeography of Cape York Peninsula: A Flora That Illustrates the Development of the Australian Sclerophyll Biota." *Australian Journal of Botany* 62, no. 2: 85–113. <https://doi.org/10.1071/BT13295>.
- Welbergen, J. A., J. Meade, H. E. Field, et al. 2020. "Extreme Mobility of the World's Largest Flying Mammals Creates Key Challenges for Management and Conservation." *BMC Biology* 18: 1–13. <https://doi.org/10.1186/s12915-020-00829-w>.
- Wellborn, G. A., and R. B. Langerhans. 2015. "Ecological Opportunity and the Adaptive Diversification of Lineages." *Ecology and Evolution* 5: 176–195.
- Westcott, D. A., A. McKeown, M. Bradford, et al. 2020. The Little Red Flying-fox: Ecology and Management of Australia's most abundant and enigmatic flying-fox, Report for the Queensland Department of Environment and Science.
- Whitlock, M. C., and K. E. Lotterhos. 2015. "Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model Through Trimming the Distribution of F_{ST} ." *American Naturalist* 186: S24–S36.
- Wieczorek, J., D. Bloom, R. Guralnick, et al. 2012. "Darwin Core: An Evolving Community-Developed Biodiversity Data Standard." *PLoS One* 7, no. 1: e29715. <https://doi.org/10.1371/journal.pone.0029715>.
- Wiens, B. J., and J. P. Colella. 2024. triangulaR: An R Package for Identifying AIMS and Building Triangle Plots Using SNP Data From Hybrid Zones. *bioRxiv* 2024–03.
- Willyard, A., R. Cronn, and A. Liston. 2009. "Reticulate Evolution and Incomplete Lineage Sorting Among the Ponderosa Pines." *Molecular Phylogenetics and Evolution* 52: 498–511. <https://doi.org/10.1016/j.ympev.2009.02.011>.
- Woinarski, J., B. Mackey, H. Nix, and B. Traill. 2007. *The Nature of Northern Australia: Natural Values, Ecological Processes and Future Prospects*. ANU Press.
- Woinarski, J. C. Z., and S. C. Tidemann. 1991. "The Bird Fauna of a Deciduous Woodland in the Wet-Dry Tropics of Northern Australia." *Wildlife Research* 18, no. 4: 479–500. <https://doi.org/10.1071/WR9910479>.
- Yue, W. 2012. Identification and characterization of Flowering Locus T Interaction Protein 7 in *Arabidopsis*. PhD Thesis, National University of Singapore.
- Zhang, W. P., Y. M. Ding, Y. Cao, et al. 2024. "Uncovering Ghost Introgression Through Genomic Analysis of a Distinct Eastern Asian Hickory Species." *Plant Journal* 119: 1386–1399. <https://doi.org/10.1111/tj.16859>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.