BOTANIC
GARDENS
OF SYDNEY

Research Centre for Ecosystem Resilience

Population genomics of *Acacia harpophylla* f.Muell. Ex benth. (Brigalow)

Final Report, June 2025

Date of Publication:
June 2025

Population genomics of *Acacia harpophylla* f.Muell. Ex benth. (Brigalow)

Research Centre for Ecosystem Resilience (ReCER), Botanic Gardens of Sydney

Executive summary

This study examines the genetic diversity, population structure, and conservation considerations for *Acacia harpophylla* (Brigalow), a tree species native to eastern Australia. A total of 225 individuals were sampled from 39 sites across Queensland and New South Wales.

Results showed low population structure and strong gene flow among sites, even those hundreds of kilometres apart, indicating that genetic material (seeds or pollen) is moving effectively across the landscape. Differences between trees mostly followed an isolation-by-distance pattern, meaning trees that are closer together tend to be more genetically similar. The species overall exhibits high genetic diversity and low levels of inbreeding, both indicators of good genetic health.

However, around 14% of individuals showed unusually high genetic variation within themselves, consistent with polyploidy (having extra sets of chromosomes), a condition often linked to clonality. These potentially clonal individuals were mainly found at a few sites, where a lack of seedlings suggests limited or no sexual reproduction.

These findings have direct implications for conservation and restoration. While Brigalow shows broad genetic connectivity and resilience, areas with high levels of clonality may face long-term reproductive challenges. Restoration efforts should prioritise the use of genetically diverse, sexually reproducing individuals from multiple populations, avoid sourcing from highly clonal sites, and aim to maintain or enhance gene flow across the landscape.

Acknowledgements

We acknowledge the Traditional Custodians of the land where the plants in this study were found and pay our respects to Elders past and present.

We thank all collectors who contributed to this study, including Joel Cohen, Nathan Emery, Justin Collette, Laura Watts, Lily Mickaill, Sam Padgett, Katherine Thomson, and Chad Beranek (BGoS); Amanda Elledge, Craig Thornton, Yan Lim, and Rory Wellham (DETSI); as well as Leif Black, Sean Nixon, and Pramod Shrestha (CQU).

We acknowledge the contributions of Craig Thornton and Amanda Elledge (DETSI), Nathan Brooks-English (CQU), and Nanjappa Ashwath (previously CQU) to the study design.

This study was funded by the Australian and Queensland Governments through the Paddock to Reef Integrated Monitoring, Modelling, and Reporting program.

Published by the Research Centre for Ecosystem Resilience (ReCER), Botanic Gardens of Sydney March 2025.

This report should be cited as:

McMaster ES, Rossetto M (2025). Population genomics of *Acacia harpophylla* f.Muell. Ex benth. (Brigalow). Research Centre for Ecosystem Resilience, Royal Botanic Garden Sydney.

Key terms

Adaptive potential: The ability of a population to survive and reproduce under changing environmental conditions.

Allele: Alternate versions of DNA sequence at a given location within the genome.

Fixed allele: Where all individuals in a population or deme have the same allele at a given locus. This can be driven by genetic bottleneck (see below) or natural selection.

Gene flow: The movement of genetic material such as differing alleles (see above) between individuals or a group of individuals by processes such as dispersal of seed and pollen.

Genet: Amongst clonal species, a genet is the genotype of an individual that develops from a single gamete (via sexual reproduction) and produces individuals with an identical genotype via asexual reproduction.

Genetic bottleneck: A sharp reduction in genetic diversity caused by a reduction in population size or by sampling of seed from too few individuals.

Genetic diversity: The totality of genetic variation present in a population and a determinant of adaptability and adaptive potential.

Genetic structure: The amount and distribution of genetic variation within and between populations of the same species across the landscape.

Heterozygosity: The average proportion of DNA sites in the genome where the two copies (or alleles) are different across a sample of individuals. Heterozygous refers to having two different copies of DNA at a particular location within the genome, as opposed to homozygous (see below) where the two copies at a location are the same.

Homozygous: Having two identical copies of DNA at a particular location within the genome, as opposed to heterozygous where the two copies at a location are different. See above.

Inbreeding: The mating of individuals that are genetically closely related within a population leading to the production of progeny.

Inbreeding depression: A decrease in fitness of progeny from closely related parents due to an increase in homozygosity with the genome.

Isolation by distance: The term used to describe the change of shared genetic material across geographic space due to dispersal ability limiting the mating between individuals. Importantly this often does not lead to discrete population structure and local adaptation if populations are continuous across the landscape.

Minor allele frequency: The most common allele at a given locus is referred to as the "major allele"; while the least common allele is the "minor allele". The minor allele frequency is therefore the frequency at which the minor allele occurs within a population.

Outbreeding: The mating of individuals that are not genetically closely related leading to the production of progeny.

Population: A group of individuals that grow in the same place at the same time and interbreed freely. A population may be composed of sub-populations, or demes.

Private alleles: Alleles observed only in a single population, thus being private to that population.

Progeny: Offspring of a plant, typically seed.

Ramet: Amongst clonal species, an individual produced asexually from a genet and thus with an identical genotype (there may be two or more ramets produced from a single genet).

Site: a specific locality where genetic samples or seeds have been collected.

Sustainability/Self-sustaining: The ability of a planting to survive and reproduce in the long-term with minimal further active investment.

Contents

1. Introduction

Acacia harpophylla F.Muell. ex Benth. (Brigalow) is a tree from the Fabaceae family, typically 5–20 m tall, known for its extensive suckering, especially after disturbance. It has furrowed, grey or brown bark and angled branchlets. The narrow, silvery-grey phyllodes are 10–20 cm long and leathery, with numerous veins and a single gland at the base. It produces bright yellow to whitish globose flower heads from July to October. The seed pods are 3–11 cm long, straight to slightly curved, with longitudinal veins and sparse to no hairs. The species name refers to the falcate (curved) shape of its phyllodes.

This species is found in western New South Wales and Queensland, it thrives in dense scrubs, low forests, and scattered clumps within open box and Myall communities. *Acacia harpophylla* plays a significant role in its ecosystem, providing habitat for wildlife and stabilizing the soil in its native regions. Historically, it has been extensively cleared for agricultural land, but efforts are being made to conserve the remaining stands in protected areas. As a hardy species, it is also a key component of the Brigalow Belt, a vast and ecologically significant region in Australia.

The Brigalow Belt stretches across approximately 408,242 km^2 of central Queensland and northern New South Wales, making up two major IBRA (Interim Biogeographic Regionalisation for Australia) bioregions: Brigalow Belt North (BBN) and Brigalow Belt South (BBS). This region serves as a transition zone between the tropical rainforests of the coast and the semi-arid interior of Australia. It is characterized by acacia-wooded grasslands, and the Brigalow tree (*Acacia harpophylla*) was once a dominant species in much of the area, particularly in the fertile lowlands. Today, however, the landscape has been heavily cleared for agriculture, with only small patches of the original habitat remaining.

Despite the region's agricultural transformation, the Brigalow Belt is recognized as one of Australia's biodiversity hotspots and still supports diverse ecosystems and provides critical habitat for numerous species. Conservation efforts are focused on preserving the remaining patches of Brigalow, as well as restoring degraded areas and protecting critical habitats, such as the Great Artesian Basin and the Darling Downs.

In this pilot study, we analysed a whole-genome SNP dataset of *Acacia harpophylla* sampled across its range in New South Wales and Queensland to examine genetic diversity patterns across the species. We evaluated genetic diversity at the individual, site, and population levels, assessed kinship and clonality, investigated connectivity and gene flow, and identified potential instances of ploidy variation. These findings contribute to the ecological and genetic knowledge of *A. harpophylla* and provide a baseline dataset in support of efforts to conserve its genetic diversity and manage its populations effectively.

Photos. Various aspects of *Acacia harpophylla* (Brigalow), showcasing its distinctive tree form, flower structures, and a stand of trees. Photos by Amanda Elledge and Craig Thornton (2024).

2. Methods

2.1 Sampling

Samples of *Acacia harpophylla* were collected across its entire range following the Restore & Renew sampling strategy, an approach that has already been applied across hundreds of species to date (Rossetto et al 2019). The Flora, Fauna, and Freshwater Research Cluster at CQUniversity and the Department of the Environment, Tourism, Science and Innovation gathered 146 samples from 26 sites in Queensland, while the Research Centre for Ecosystem Resilience at the Botanic Gardens of Sydney collected 79 samples from 13 sites in New South Wales. Geographic coordinates were recorded for each plant, and approximately 3 g of leaf tissue was collected from each individual. Sampling was designed to reflect the genetic diversity across populations of varying sizes, based on known distribution patterns and the number of individuals at each site. Collectors aimed to sample six individuals per site. After collection, samples were dried and stored in silica gel to preserve them until DNA extraction.

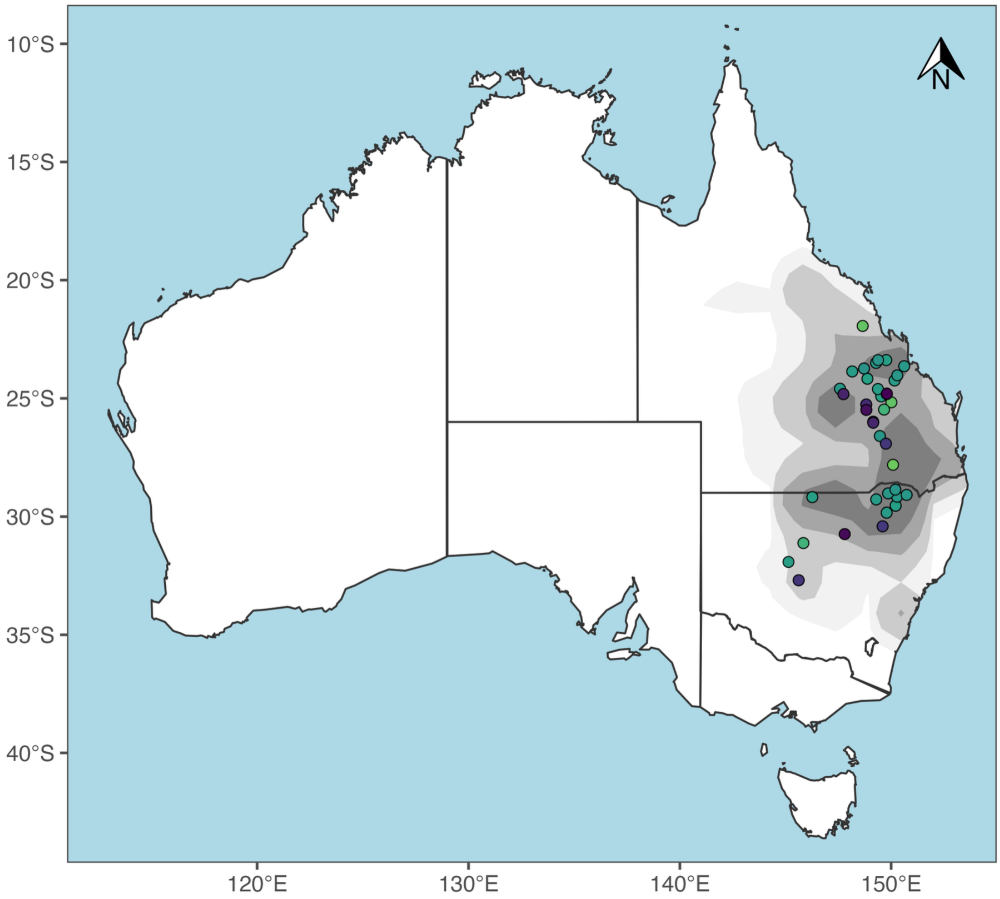


Figure 1. *Acacia harpophylla* distribution and sampling. Density distribution of *Acacia harpophylla*, with grey-shaded regions representing areas where 99%, 95%, 80%, and 50% (light to dark) of Atlas of Living Australia records are concentrated. Colored points indicate the samples used in this study, with individual heterozygosity represented on a gradient: yellow-green (H_O 0–0.25) for low to mid heterozygosity and blue ($H_O > 0.25$) for high heterozygosity individuals. Photo by Craig Thornton (2024).

2.2 DNA Extraction and DArTseq

Genotyping was performed using medium-density DArTseq, a reduced representation sequencing technique provided by Diversity Arrays Technology Australia Pty Ltd (Canberra, ACT, Australia). The DArTseq method involves a genome restriction digest followed by sequencing of the digested fragments using an Illumina platform. Single Nucleotide Polymorphisms (SNPs) were called using proprietary analytical pipelines developed by DArT Pty Ltd (Jaccoud et al., 2001; Kilian et al., 2012).

2.3 Genetic Analyses

Data filtering was performed using custom scripts implemented with RRtools (v0.1.0, see https://github.com/jasongbragg/RRtools) to prepare the dataset for analysis. Loci with more than 20% missing data were removed, as were fixed loci. To minimize linkage disequilibrium effects, the dataset was further reduced to include only one SNP per locus. Samples with more than 30% missing loci were excluded. Following these filtering steps, the dataset comprised 225 samples and 28,151 loci, which were used for subsequent analyses.

The distribution and sampling map for the populations with latitude/longitude information was made using species occurrence data for *Acacia harpophylla*, which was obtained from the Atlas of Living Australia (ALA) using the galah R package (v1.5.0, see https://rdrr.io/cran/galah/; (Westgate et al., 2023)). The spatial distribution of occurrences was visualized using the ozmaps package (v0.4.2, see https://rdrr.io/cran/ozmaps/) and density contours were generated with geom_hdr from the ggdensity package (v0.1.0, see https://rdrr.io/cran/ggdensity/).

Genetic diversity statistics, including observed heterozygosity (H_O), expected heterozygosity ($\mathrm{H_E}$), and the inbreeding coefficient (F_{IS}), were calculated for each population using the faststats function from the fastDiversity package (v0.1.0, see https://github.com/eilishmcmaster/fastDiversity/). A minimum minor allele frequency (MAF) filter of 5% and a maximum locus missingness threshold of 20% were applied. Additionally, individual-level H_O and F_{IS} values were calculated across the species dataset.

Pairwise kinship estimates were obtained using the identity-by-descent (IBD) method in PLINK (Purcell et al. 2007; Chang et al. 2015), implemented via the snpgdsIBDMoM function in SNPRelate (v1.36.0, see http://rdrr.io/bioc/SNPRelate/; (Zheng et al., 2012)). The analysis applied a 5% MAF filter and a maximum locus missingness of 20%. Individuals exceeding the theoretical kinship threshold for clonality (kin > $1/2^{3/2}$) (Manichaikul et al., 2010) were considered genetically identical, and the sample with lower data quality was excluded from principal component analysis (PCA).

PCA was conducted using the adegenet package (v1.7–22, see http://rdrr.io/cran/adegenet/; (Jombart, 2008)) to visualize genetic clustering within the multispecies dataset. Population differentiation was assessed by calculating fixation index (F_{ST}) between sites with four or more samples using the snpgdsFst function from the SNPRelate package. The Weir and Hill (Weir & Hill, 2002) estimator was used for F_{ST} calculations, applying a minor allele frequency (MAF) filter of 5% and a missingness filter of 20% to ensure robust estimates. For both PCA and F_{ST}, clones were excluded.

Additional R packages used for analysis and visualization included dplyr (v1.1.4, see https://rdrr.io/cran/dplyr/), ggplot2 (v3.5.1, see https://rdrr.io/cran/ggplot2/; (Wickham, 2016)), ggthemes (v5.0.0, see https://rdrr.io/cran/ggthemes/; (Arnold, 2024)), RColorBrewer (v1.1–3, see https://rdrr.io/cran/RColorBrewer/), ComplexHeatmap (v2.18.0, see https://rdrr.io/bioc/ComplexHeatmap/; (Gu, 2022)), and ggpubr (v0.6.0, see https://rdrr.io/cran/ggpubr/).

3. Results and Discussion

3.1 Population structure

Acacias can be quite variable in the way genetic diversity is distributed across the landscape and on how mating system and extent of clonality impact on distribution-wide genetic diversity and divergence (Rossetto et al. 2020). The analysis of population structure in *Acacia harpophylla* reveals no major genetic differentiation across its range. The variance between individuals is minimal, as shown by the PCA analysis, where the first principal component (PC1) explains less than 2% of the total variance (Figure 2). The limited variance observed is partly attributable to isolation by distance. Specifically, PC1 is moderately and significantly positively correlated with latitude (Pearson r = 0.46, t = 6.99, df = 178, p = 5.20e-11), with a 95% confidence interval of [0.34, 0.57], and moderately and significantly positively correlated with longitude (Pearson r = 0.61, t = 10.41, df = 178, p < 2.2e-16), with a 95% confidence interval of [0.52, 0.70] (Figure 2).There is some separation of the Careunga NR site from the rest of the population along PC2, however this is likely because these individuals are all closely related (Table 2), which can affect the PCA results.

Additionally, the genetic data indicate substantial gene flow between sites (either historical or current), with F_{ST} values remaining consistently low (Figure 3), even between sites separated by hundreds of kilometres, never exceeding 0.2 (Figure 4). These patterns are consistent with those observed in other large, widespread tree species with continuous distributions, such as *Eucalyptus* (particularly those west of the GDR), which are predominantly wind-pollinated and dispersed.

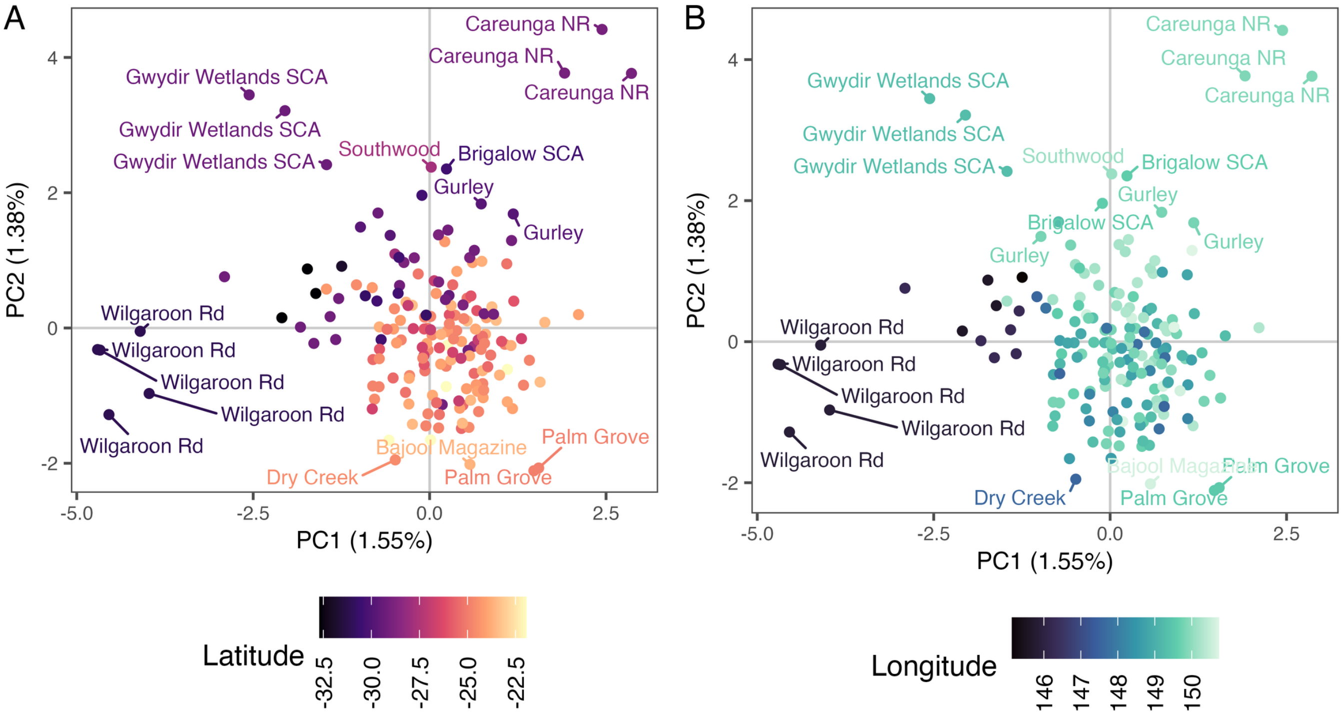
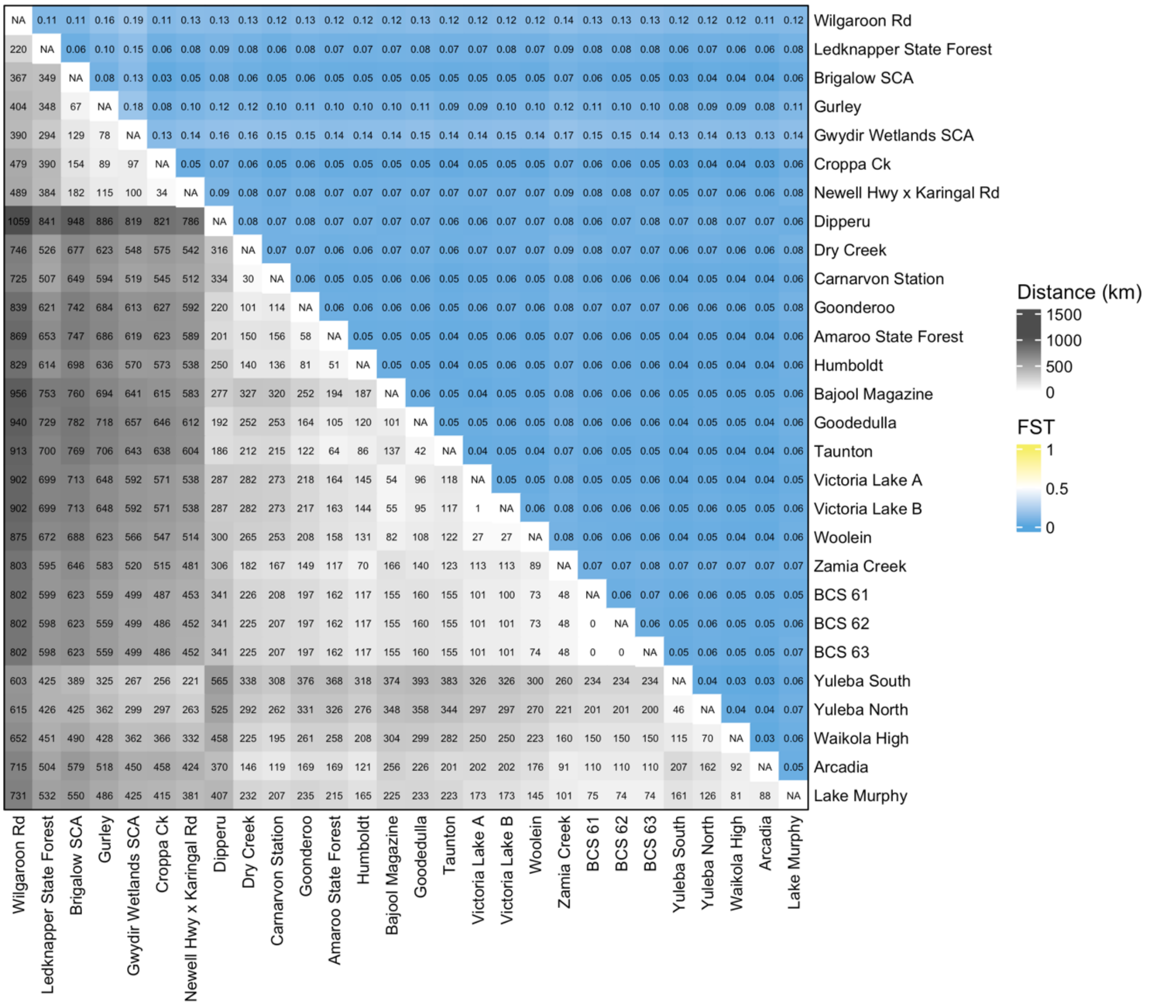


Figure 2. Principal components analysis colored by latitude or longitude. Principal component analysis (PCA) of SNP data, filtered for MAF ≥ 5% and a maximum locus missingness of 20%, and randomly subsampled to 1,000 loci (clones excluded). Each point represents a sample, with labels indicating the site names. The analysis shows that PC1 is moderately and significantly positively correlated with latitude (Pearson r = 0.46, t = 6.99, df = 178, p = 5.20e-11; 95% CI: [0.34, 0.57]) and longitude (Pearson r = 0.61, t = 10.41, df = 178, p < 2.2e-16; 95% CI: [0.52, 0.70]).

A PCA on genotype data simplifies complex genetic differences to reveal how individuals or populations relate to each other genetically. In this case, the first principal component (PC1) explains less than 2% of the variance, which means there is very little genetic differentiation or structure among the samples.



	Wilgaroon Rd	Ledknapper State Forest	Brigalow SCA	Gurley	Gwydir Wetlands SCA	Croppa Ck	Newell Hwy x Karingal Rd	Dipperu	Dry Creek	Carnarvon Station	Goonderoo	Amaroo State Forest	Humboldt	Bajool Magazine	Goodedulla	Taunton	Victoria Lake A	Victoria Lake B	Woolein	Zamia Creek	BCS 61	BCS 62	BCS 63	Yuleba South	Yuleba North	Waikola High	Arcadia	Lake Murphy
Wilgaroon Rd	NA	0.11	0.11	0.16	0.19	0.11	0.13	0.13	0.13	0.12	0.13	0.12	0.12	0.12	0.13	0.12	0.12	0.12	0.12	0.14	0.13	0.13	0.13	0.12	0.12	0.12	0.11	0.12
Ledknapper State Forest	220	NA	0.06	0.10	0.15	0.06	0.08	0.09	0.08	0.06	0.08	0.07	0.07	0.07	0.08	0.07	0.07	0.08	0.07	0.09	0.08	0.08	0.08	0.06	0.07	0.06	0.06	0.08
Brigalow SCA	367	349	NA	0.08	0.13	0.03	0.05	0.08	0.06	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.07	0.06	0.05	0.05	0.03	0.04	0.04	0.04	0.06
Gurley	404	348	67	NA	0.18	0.08	0.10	0.12	0.12	0.10	0.11	0.10	0.10	0.10	0.11	0.09	0.09	0.10	0.10	0.12	0.11	0.10	0.10	0.08	0.09	0.09	0.08	0.11
Gwydir Wetlands SCA	390	294	129	78	NA	0.13	0.14	0.16	0.16	0.15	0.15	0.14	0.14	0.14	0.15	0.14	0.14	0.15	0.14	0.17	0.15	0.15	0.14	0.13	0.14	0.13	0.13	0.14
Croppa Ck	479	390	154	89	97	NA	0.05	0.07	0.06	0.05	0.06	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.07	0.06	0.06	0.05	0.03	0.04	0.04	0.03	0.06
Newell Hwy x Karingal Rd	489	384	182	115	100	34	NA	0.09	0.08	0.07	0.08	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.09	0.08	0.08	0.07	0.05	0.07	0.06	0.06	0.08
Dipperu	1059	841	948	886	819	821	786	NA	0.08	0.07	0.08	0.07	0.07	0.07	0.07	0.06	0.06	0.07	0.07	0.08	0.06	0.07	0.08	0.07	0.08	0.07	0.07	0.06
Dry Creek	746	526	677	623	548	575	542	316	NA	0.07	0.07	0.06	0.06	0.07	0.07	0.06	0.07	0.07	0.07	0.09	0.08	0.07	0.07	0.06	0.07	0.06	0.06	0.08
Carnarvon Station	725	507	649	594	519	545	512	334	30	NA	0.06	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.05	0.08	0.06	0.06	0.06	0.04	0.05	0.04	0.04	0.06
Goonderoo	839	621	742	684	613	627	592	220	101	114	NA	0.06	0.06	0.06	0.06	0.05	0.06	0.07	0.06	0.08	0.07	0.07	0.07	0.06	0.06	0.06	0.05	0.08
Amaroo State Forest	869	653	747	686	619	623	589	201	150	156	58	NA	0.05	0.05	0.05	0.04	0.05	0.06	0.05	0.07	0.06	0.06	0.05	0.04	0.05	0.04	0.04	0.06
Humboldt	829	614	698	636	570	573	538	250	140	136	81	51	NA	0.05	0.05	0.04	0.05	0.06	0.05	0.07	0.06	0.06	0.06	0.04	0.05	0.04	0.04	0.06
Bajool Magazine	956	753	760	694	641	615	583	277	327	320	252	194	187	NA	0.06	0.05	0.04	0.05	0.05	0.08	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.06
Goodedulla	940	729	782	718	657	646	612	192	252	253	164	105	120	101	NA	0.05	0.05	0.06	0.05	0.08	0.06	0.06	0.06	0.04	0.05	0.05	0.04	0.06
Taunton	913	700	769	706	643	638	604	186	212	215	122	64	86	137	42	NA	0.04	0.05	0.04	0.07	0.06	0.05	0.05	0.04	0.05	0.04	0.04	0.06
Victoria Lake A	902	699	713	648	592	571	538	287	282	273	218	164	145	54	96	118	NA	0.05	0.05	0.08	0.05	0.05	0.06	0.04	0.05	0.04	0.04	0.05
Victoria Lake B	902	699	713	648	592	571	538	287	282	273	217	163	144	55	95	117	1	NA	0.06	0.08	0.06	0.06	0.06	0.05	0.06	0.05	0.05	0.06
Woolein	875	672	688	623	566	547	514	300	265	253	208	158	131	82	108	122	27	27	NA	0.08	0.06	0.06	0.06	0.04	0.05	0.04	0.04	0.06
Zamia Creek	803	595	646	583	520	515	481	306	182	167	149	117	70	166	140	123	113	113	89	NA	0.07	0.07	0.08	0.07	0.07	0.07	0.07	0.07
BCS 61	802	599	623	559	499	487	453	341	226	208	197	162	117	155	160	155	101	100	73	48	NA	0.06	0.07	0.06	0.06	0.05	0.05	0.05
BCS 62	802	598	623	559	499	486	452	341	225	207	197	162	117	155	160	155	101	101	73	48	0	NA	0.06	0.05	0.06	0.05	0.05	0.06
BCS 63	802	598	623	559	499	486	452	341	225	207	197	162	117	155	160	155	101	101	74	48	0	0	NA	0.05	0.06	0.05	0.05	0.07
Yuleba South	603	425	389	325	267	256	221	565	338	308	376	368	318	374	393	383	326	326	300	260	234	234	234	NA	0.04	0.03	0.03	0.06
Yuleba North	615	426	425	362	299	297	263	525	292	262	331	326	276	348	358	344	297	297	270	221	201	201	200	46	NA	0.04	0.04	0.07
Waikola High	652	451	490	428	362	366	332	458	225	195	261	258	208	304	299	282	250	250	223	160	150	150	150	115	70	NA	0.03	0.06
Arcadia	715	504	579	518	450	458	424	370	146	119	169	169	121	256	226	201	202	202	176	91	110	110	110	207	162	92	NA	0.05
Lake Murphy	731	532	550	486	425	415	381	407	232	207	235	215	165	225	233	223	173	173	145	101	75	74	74	161	126	81	88	NA

Figure 3. Pairwise distance and F_{ST} between sites as a heatmap. Lower triangle (greys) shows pairwise distance between sites in kilometers (km). upper triangle (blues) indicates pairwise F_{ST} calculated using Weir and Hill (2002) method. Clones are excluded. The fixation index (F_{ST}) indicates how different allele frequencies are between sites. When F_{ST} is low (approaching 0), sites have similar allele frequencies, indicating free gene flow/connectivity. When F_{ST} is high (approaching 1), sites have very divergent allele frequencies, indicating substantial isolation and divergence.

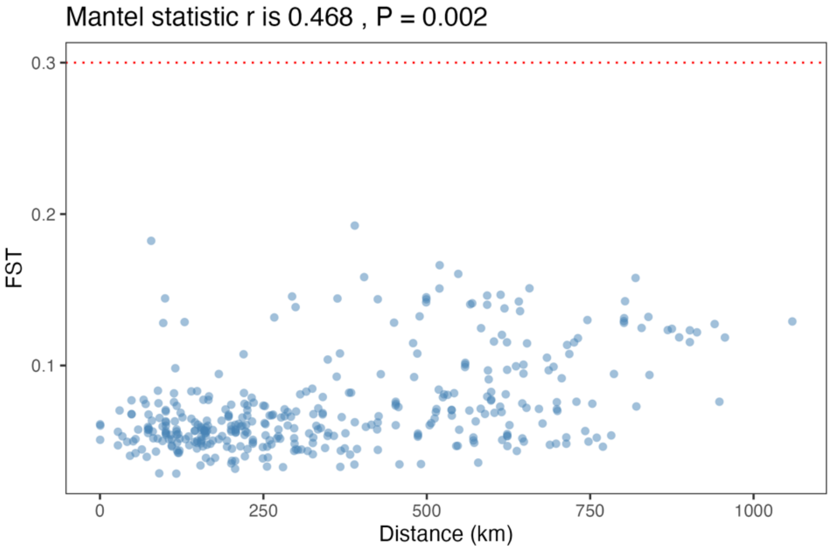


Figure 4. Pairwise distance vs F_{ST} between sites as scatter plot. Each point represents a pairwise measurement between two sites. Clones are excluded.

3.2 Diversity, Kinship and Clonality

Acacia harpophylla exhibits high levels of heterozygosity and low inbreeding, which are favourable for fitness and adaptive potential (Table 1). Although some degree of inbreeding is observed, with a few individuals displaying very low heterozygosity, no sites show high levels of inbreeding. These findings align with the species' preference for outbreeding, a pattern that also supports the absence of population structure.

A key finding is that 14% of the individuals sampled (31/225) exhibit notably higher heterozygosity ($H_O > 0.25$, $F_{IS} < 0$), which distinguishes them from the rest of the population (H_O 0–0.25, $F_{IS} > 0$) (Figure 5a). This increase in heterozygosity may suggest the presence of polyploid individuals, as indicated by read count ratios that deviate from diploid expectations (Figure 5). Other species of *Acacia*, such as *Acacia ulicifolia*, are known to exhibit ploidy variation.

Notably, most of these high-heterozygosity individuals were found to be clones (80% or 25/31 high-heterozygosity individuals were genetically identical to at least one other individual) (Figure 5, Table 2). If these individuals are polyploid, their high clonality could be a result of the fact that polyploids often rely on asexual reproduction (apomixis) when sexual reproduction is limited.

At sites predominantly inhabited by high-heterozygosity individuals, such as Quambone Rd ([-30.740, 147.803]) and Yathong NR ([-32.693, 145.623]), field observations reported a uniformity among individuals within stands, with no juveniles and evidence of senescence (Table 2). This supports the hypothesis that the putatively polyploid clones may not be capable of sexual reproduction, further supporting the idea of asexual persistence.

While most putative polyploid individuals (those with high heterozygosity) were clonal, not all clones were polyploid. In fact, the majority of clones had average heterozygosity, with 62% (41/66) of the detected clones showing this characteristic. Notably, Biniguy, Brigalow Catchment Site (BCS) W, and Warilda sites were made up predominantly of single genets of average-heterozygosity clones (Table 2). This implies that asexual reproduction, either through apomixis or resprouting, can be common especially where outcrossed sexual reproduction is no longer possible (inbreeding avoidance).

In addition to clonality, several instances of first- and second-degree relationships (parent-offspring or sibling relationships) were identified, but these were confined to the same site (Table 2), suggesting no evidence of first-generation dispersal between sites. Notably, some of these relationships were observed between average-heterozygosity individuals and high-heterozygosity, potential polyploid individuals, particularly in sites such as Gwydir Wetlands SCA, Palm Grove, and Careunga NR. This implies that the average-heterozygosity individuals may be related to the high-heterozygosity individuals, either as siblings or parents/offspring.

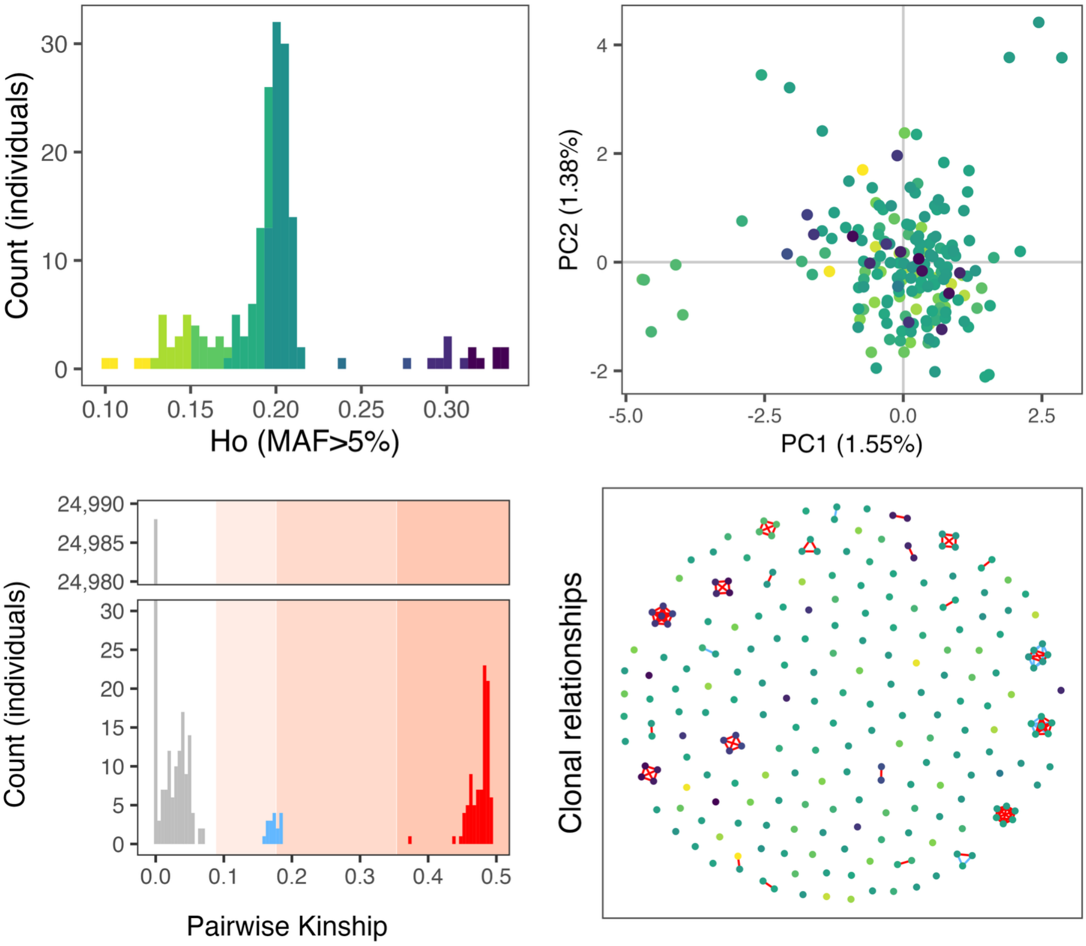


Figure 5. Summary of individual diversity analyses in *Acacia harpophylla*. (A) Histogram of observed heterozygosity for 225 samples (MAF ≥ 5%). (B) PCA of samples, excluding clones, colored by individual heterozygosity (as defined in A). (C) Histogram of pairwise kinship values, with background shading indicating theoretical kinship ranges: clones (dark pink), first-degree relatives (parent-offspring/full siblings; medium pink), and second-degree relatives (half-siblings; light pink). Clonal relationships are colored red, and first-/second-degree relatives are blue. (D) Network of genetic relationships, where each point represents an individual, and lines indicate close relationships: clones (red) and first-/second-degree relatives (blue). In plots A, B, and D, higher heterozygosity is shown in blue, and lower heterozygosity in yellow.

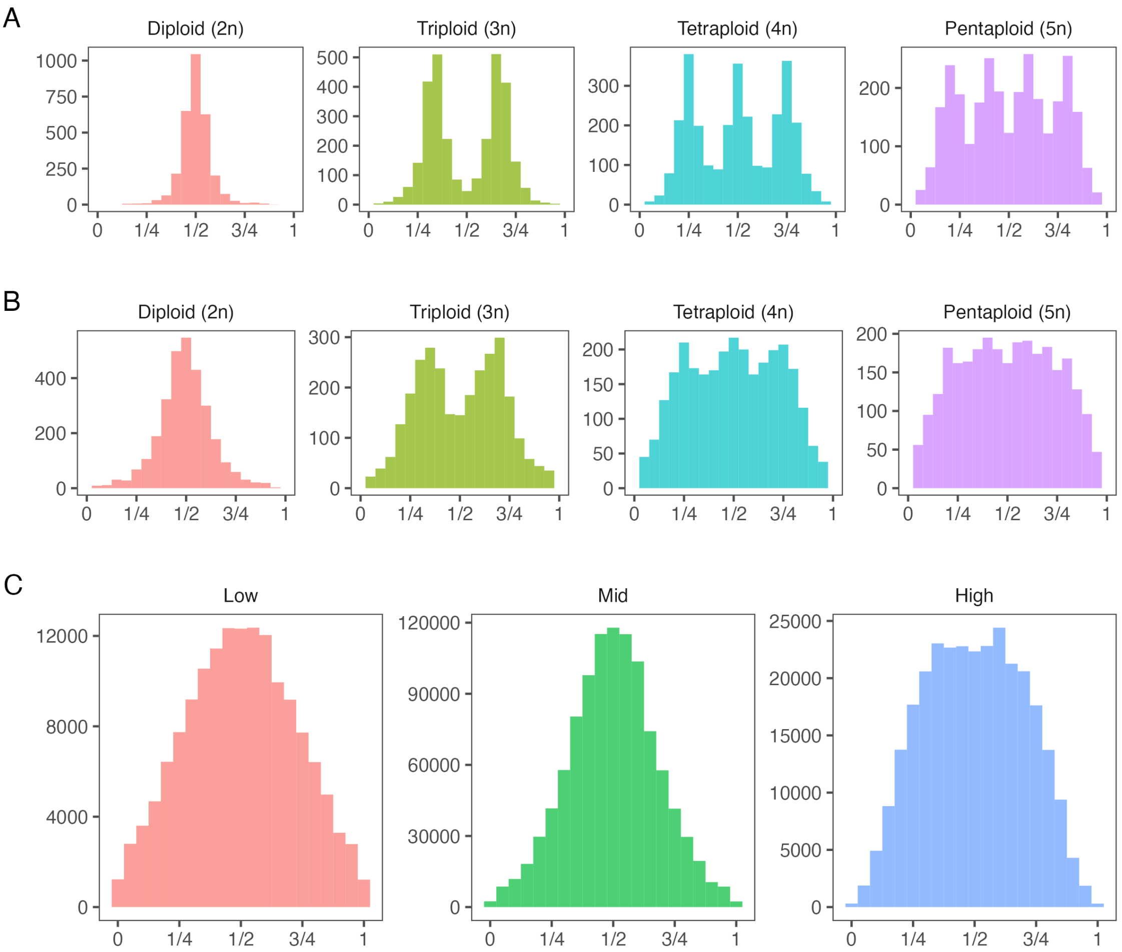


Figure 6. Analysis of read count ratios for ploidy variation in *A. harpophylla*. Rows (A) and (B) show the expected read count ratio patterns for diploid (2n), triploid, tetraploid, and pentaploid scenarios, with (B) representing a noisier dataset scenario. Row (C) presents observed read count ratios for *A. harpophylla* samples, categorized by diversity levels: low (H_O < 0.17, F_{IS} > 0.4), mid (H_O 0.17–0.25, F_{IS} 0.4–0), and high (H_O > 0.25, F_{IS} < 0). The mid-diversity pattern is consistent with diploidy, while the high-diversity pattern suggests possible polyploidy, potentially tetraploidy, though resolution is insufficient for a definitive conclusion. The low-diversity pattern appears typical of more inbred diploidy with reduced heterozygosity.

3.3 Preliminary implications and future directions

From a management perspective, the genetic structure of *A. harpophylla* does not pose major concerns. The high levels of gene flow and observed diversity associated with latitude suggests that climate-driven provenance (planting individuals from higher latitudes to prepare for future climate conditions) could be considered for restoration. However, the low variance across latitudes indicates that this may not meaningfully impact the species' adaptability. Additional associative studies looking into temporal environmental niche variations will further enhance our interpretative power.

To confirm the levels of ploidy variation, flow cytometry on fresh leaf samples could be used to assess the genome size of suspected polyploids. This would provide definitive evidence of polyploidy and help further refine restoration strategies.

As preliminary guidance to restoration practices, the data suggest that focus should be on planting genetically diverse individuals from multiple sites to ensure broad genetic representation. It is important to avoid collecting from sites known to harbor clonal or polyploid individuals, as these may have reduced fertility and could fail to contribute effectively to the overall genetic pool.

From a site conservation viewpoint, sites with high levels of clonality or close genetic relatedness among individuals may be at risk of extinction due to low reproductive output. Introducing unrelated individuals from other sites in these areas could enhance reproductive success and improve local long-term viability.

Genetic data can also be used to inform specific collection recommendations that maximize diversity, ensuring that restoration efforts are well-supported by a varied genetic foundation.

In conclusion, while *A. harpophylla* generally exhibits a homogenous genetic profile with high genetic diversity but low diversification, the presence of polyploidy and clonality in certain individuals requires careful consideration in management efforts. Ensuring genetic diversity in these programs is critical for the species' persistence in a rapidly changing environment.

3.4 Summary

Key findings

- *Acacia harpophylla* exhibits little genetic difference between locations, even those far apart
    - Indicates effective dispersal of pollen and seeds and limited local adaptation
- High overall genetic diversity and low inbreeding levels indicates strong potential for adaptation and resilience
- Approximately 14% of individuals show signs of polyploidy (increased number of chromosomes)
    - This can happen randomly in plants
    - Polyploids often struggle to reproduce sexually and rely more on asexual reproduction, which we see in *A. harpophylla*
- Not all clones are putative polyploids
    - Sexual reproduction via apomixis or resprouting also occurs occasionally in non-polyploid plants
    - In *A. harpophylla*, this may occur when there are limited mate options to avoid inbreeding

Conservation and restoration guidance

- Source planting material from genetically diverse individuals across multiple sites
- Avoid using clonal or polyploid individuals where possible, due to their reduced reproductive potential
- Consider genetic supplementation at highly clonal or genetically uniform sites to improve local reproductive success
- Use future ploidy testing (e.g., flow cytometry) to confirm polyploidy and inform site-specific strategies

4. Tables

Table 1. Summary diversity statistics per site. For each site, the table reports centroid latitude and longitude, observed heterozygosity (H_O), expected heterozygosity (H_E), inbreeding coefficient (F_{IS}), and the number of individuals (n). Analyses were conducted on 8,036 polymorphic loci with a minimum minor allele frequency of 5% and a maximum locus missingness of 20%. Clones are included, which can decrease F_{IS}.

Observed heterozygosity (H_O) shows how much genetic variation is present in a population, where higher values mean more diversity. F_{IS} measures inbreeding, where positive values mean more inbreeding, negative values mean more outcrossing. Both help assess how genetically healthy and resilient a population is.

Site	**Latitude**	**Longitude**	H_O	H_E	F_{IS}	**n**
Dipperu	-21.940	148.652	0.146	0.217	0.324	5
Goodedulla	-23.385	149.674	0.203	0.224	0.095	4
Taunton	-23.504	149.288	0.199	0.242	0.179	6
Bajool Magazine	-23.645	150.618	0.199	0.234	0.151	5
Amaroo State Forest	-23.742	148.716	0.203	0.242	0.158	6
Goonderoo	-23.865	148.158	0.199	0.237	0.161	6
Victoria Lake A	-24.031	150.287	0.202	0.237	0.15	5
Victoria Lake B	-24.032	150.293	0.195	0.232	0.161	6
Humboldt	-24.175	148.876	0.195	0.237	0.18	6
Woolein	-24.248	150.163	0.194	0.241	0.195	6
Dry Creek	-24.597	147.578	0.203	0.221	0.081	4
Zamia Creek	-24.610	149.374	0.153	0.203	0.243	4
BCS W	-24.804	149.796	0.204	0.189	-0.077	6
BCS 61	-24.814	149.796	0.156	0.232	0.329	6
BCS 62	-24.816	149.794	0.155	0.232	0.332	6
BCS 63	-24.817	149.793	0.224	0.242	0.074	6
Carnarvon Station	-24.821	147.742	0.242	0.234	-0.037	6
Palm Grove	-24.916	149.532	0.197	0.122	-0.613	6
Isla Gorge	-25.167	150.011	0.146	0.18	0.186	3
Arcadia	-25.258	148.821	0.22	0.252	0.127	6
Lake Murphy	-25.474	149.665	0.151	0.222	0.317	5
Lonesome Holding	-25.488	148.826	0.297	0.226	-0.315	6
Waikola High	-26.030	149.1506	0.221	0.249	0.113	6
Waikola Low	-25.988	149.1508	0.193	0.206	0.061	6
Yuleba North	-26.594	149.466	0.205	0.243	0.155	6
Yuleba South	-26.916	149.753	0.217	0.25	0.129	6
Southwood	-27.811	150.090	0.148	0.186	0.204	3
Newell Hwy x Karingal Rd	-28.864	150.201	0.203	0.237	0.141	6
Careunga NR	-29.022	149.857	0.207	0.151	-0.367	6
Warialda Rd	-29.085	150.733	0.205	0.176	-0.162	6
Croppa Ck	-29.164	150.281	0.197	0.245	0.196	6
Ledknapper State Forest	-29.180	146.270	0.179	0.234	0.237	6
Gwydir Wetlands SCA	-29.278	149.293	0.174	0.184	0.057	6
Biniguy	-29.538	150.214	0.199	0.102	-0.961	6
Gurley	-29.834	149.788	0.202	0.198	-0.021	6
Brigalow SCA	-30.411	149.596	0.21	0.249	0.157	6
Quambone Rd	-30.740	147.802	0.323	0.224	-0.442	6
Wilgaroon Rd	-31.120	145.847	0.176	0.208	0.153	5
Yathong NR	-32.694	145.624	0.291	0.228	-0.276	12

Table 2: Relationships detected in *A. harpophylla*. Each row represents an individual sample, including site location (latitude and longitude), family group, genet (clonal lineage identifier), observed heterozygosity (H_O), and inbreeding coefficient (F_{IS}). Families consist of individuals linked by zero-, first-, or second-degree relationships, while a genet represents a group of genetically identical individuals (clones). Families or genets with two or more individuals are highlighted in red.

H_O and F_{IS} were calculated using 8,036 polymorphic loci, with a minimum minor allele frequency of 5% and a maximum locus missingness of 20%. Individuals with high heterozygosity (H_O >0.25), which may indicate polyploidy, are highlighted in green.

Sample	Site	Latitude	Longitude	Family	Genet	H_O	F
CQUD3	Dipperu	-21.939	148.652	1	1	0.161	0.430
CQUD2	Dipperu	-21.940	148.652	2	2	0.136	0.520
CQUD4	Dipperu	-21.940	148.652	3	3	0.129	0.544
CQUD1	Dipperu	-21.940	148.652	4	4	0.143	0.495
CQUD5	Dipperu	-21.940	148.652	5	5	0.159	0.438
CQUGP3	Goodedulla	-23.384	149.770	6	6	0.200	0.293
CQUGP1	Goodedulla	-23.385	149.385	7	7	0.213	0.248
CQUGP4	Goodedulla	-23.385	149.770	8	8	0.195	0.309
CQUGP5	Goodedulla	-23.385	149.770	9	9	0.202	0.285
CQUT01	Taunton	-23.504	149.289	10	10	0.202	0.287
CQUT02	Taunton	-23.504	149.289	11	11	0.190	0.328
CQUT03	Taunton	-23.504	149.289	12	12	0.191	0.324
CQUT04	Taunton	-23.504	149.288	13	13	0.209	0.261
CQUT05	Taunton	-23.504	149.288	14	14	0.210	0.259
CQUT06	Taunton	-23.505	149.288	15	15	0.191	0.325
CQUB1	Bajool Magazine	-23.645	150.618	16	16	0.194	0.313
CQUB2	Bajool Magazine	-23.645	150.618	17	17	0.192	0.322
CQUB3	Bajool Magazine	-23.645	150.618	18	18	0.208	0.264
CQUB4	Bajool Magazine	-23.645	150.618	19	19	0.202	0.286
CQUB6	Bajool Magazine	-23.645	150.618	20	20	0.197	0.303
CQUASF4	Amaroo State Forest	-23.742	148.717	21	21	0.205	0.274
CQUASF5	Amaroo State Forest	-23.742	148.716	22	22	0.199	0.297
CQUASF6	Amaroo State Forest	-23.742	148.716	23	23	0.206	0.273
CQUASF1	Amaroo State Forest	-23.742	148.717	24	24	0.216	0.237
CQUASF2	Amaroo State Forest	-23.742	148.717	25	25	0.200	0.294
CQUASF3	Amaroo State Forest	-23.743	148.716	26	26	0.195	0.312
CQUG6	Goonderoo	-23.865	148.159	27	27	0.201	0.291
CQUG4	Goonderoo	-23.865	148.159	28	28	0.204	0.279
CQUG5	Goonderoo	-23.865	148.158	29	29	0.197	0.305
CQUG3	Goonderoo	-23.866	148.158	30	30	0.184	0.349
CQUG1	Goonderoo	-23.866	148.158	30	32	0.202	0.288
CQUG2	Goonderoo	-23.866	148.159	31	31	0.205	0.277
CQUVLA01	Victoria Lake A	-24.033	150.293	32	33	0.199	0.296
CQUVLA02	Victoria Lake A	-24.032	150.293	33	34	0.196	0.307
CQUVLA03	Victoria Lake A	-24.032	150.293	34	35	0.202	0.285
CQUVLA04	Victoria Lake A	-24.032	150.293	35	36	0.210	0.259
CQUVLA05	Victoria Lake A	-24.032	150.293	36	37	0.201	0.291
CQUVLA06	Victoria Lake B	-24.032	150.293	37	38	0.199	0.297
CQUVLB01	Victoria Lake B	-24.031	150.286	38	39	0.195	0.312
CQUVLB02	Victoria Lake B	-24.031	150.286	38	39	0.195	0.312
CQUVLB03	Victoria Lake B	-24.031	150.286	39	40	0.198	0.300
CQUVLB04	Victoria Lake B	-24.031	150.286	40	41	0.197	0.305
CQUVLB05	Victoria Lake B	-24.031	150.285	41	42	0.184	0.348
CQUH5	Humboldt	-24.174	148.877	42	43	0.194	0.315
CQUH6	Humboldt	-24.175	148.877	43	44	0.201	0.290
CQUH4	Humboldt	-24.175	148.876	44	45	0.196	0.308
CQUH3	Humboldt	-24.175	148.876	45	46	0.191	0.324
CQUH2	Humboldt	-24.175	148.876	46	47	0.196	0.308
CQUH1	Humboldt	-24.175	148.876	47	48	0.191	0.326
CQUW2	Woolein	-24.248	150.163	48	49	0.203	0.282
CQUW1	Woolein	-24.248	150.163	49	50	0.204	0.278
CQUW3	Woolein	-24.248	150.163	50	51	0.198	0.302

CQUW5	Woolein	-24.248	150.163	51	52	0.179	0.366
CQUW4	Woolein	-24.248	150.163	52	53	0.204	0.279
CQUW6	Woolein	-24.249	150.163	53	54	0.176	0.378
CQUDC6	Dry Creek	-24.596	147.577	54	55	0.209	0.261
CQUDC1	Dry Creek	-24.597	147.578	55	56	0.201	0.289
CQUDC2	Dry Creek	-24.597	147.578	56	57	0.202	0.285
CQUDC3	Dry Creek	-24.597	147.578	57	58	0.201	0.289
CQUZC6	Zamia Creek	-24.610	149.374	58	59	0.133	0.531
CQUZC5	Zamia Creek	-24.610	149.374	59	60	0.119	0.580
CQUZC2	Zamia Creek	-24.610	149.374	60	61	0.146	0.483
CQUZC1	Zamia Creek	-24.610	149.374	61	62	0.204	0.278
DNRW1	BCS W	-24.802	149.795	62	63	0.169	0.401
DNRW2	BCS W	-24.802	149.795	62	63	0.173	0.388
DNRW3	BCS W	-24.802	149.795	62	63	0.172	0.392
DNRW4	BCS W	-24.803	149.795	62	63	0.173	0.390
DNRW5	BCS W	-24.805	149.797	63	64	0.197	0.302
DNRW6	BCS W	-24.808	149.801	64	65	0.336	-0.188
DNR61-4	BCS 61	-24.814	149.796	65	66	0.133	0.531
DNR61-3	BCS 61	-24.814	149.796	66	67	0.146	0.485
DNR61-5	BCS 61	-24.814	149.796	67	68	0.168	0.406
DNR61-2	BCS 61	-24.814	149.796	68	69	0.195	0.311
DNR61-6	BCS 61	-24.814	149.796	69	70	0.147	0.482
DNR61-1	BCS 61	-24.814	149.796	70	71	0.139	0.507
DNR62-2	BCS 62	-24.815	149.794	71	72	0.154	0.455
DNR62-5	BCS 62	-24.815	149.795	72	73	0.153	0.459
DNR62-4	BCS 62	-24.816	149.795	73	74	0.137	0.515
DNR62-3	BCS 62	-24.816	149.795	74	75	0.175	0.380
DNR62-1	BCS 62	-24.816	149.794	75	76	0.158	0.440
DNR62-6	BCS 62	-24.816	149.794	76	77	0.147	0.479
DNR63-1	BCS 63	-24.817	149.793	77	78	0.330	-0.166
DNR63-2	BCS 63	-24.817	149.793	78	79	0.199	0.297
DNR63-6	BCS 63	-24.817	149.793	79	80	0.173	0.387
DNR63-5	BCS 63	-24.817	149.793	80	81	0.192	0.321
DNR63-4	BCS 63	-24.817	149.793	81	82	0.207	0.269
DNR63-3	BCS 63	-24.817	149.793	82	83	0.239	0.155
CQUCS2	Carnarvon Station	-24.821	147.742	84	85	0.203	0.283
CQUCS1	Carnarvon Station	-24.821	147.742	83	84	0.208	0.263
CQUCS4	Carnarvon Station	-24.821	147.742	83	84	0.209	0.260
CQUCS3	Carnarvon Station	-24.821	147.742	85	86	0.314	-0.111
CQUCS5	Carnarvon Station	-24.821	147.742	85	86	0.310	-0.096
CQUCS6	Carnarvon Station	-24.822	147.742	86	87	0.206	0.271
CQUPG3	Palm Grove	-24.916	149.532	87	89	0.195	0.310
CQUPG2	Palm Grove	-24.916	149.532	87	88	0.195	0.312
CQUPG1	Palm Grove	-24.916	149.532	87	88	0.197	0.305
CQUPG5	Palm Grove	-24.916	149.532	87	88	0.197	0.305
CQUPG6	Palm Grove	-24.916	149.532	87	88	0.199	0.298
CQUPG4	Palm Grove	-24.916	149.532	87	88	0.198	0.302
CQUIG1	Isla Gorge	-25.167	150.012	88	90	0.163	0.424
CQUIG2	Isla Gorge	-25.167	150.012	89	91	0.124	0.563
CQUIG4	Isla Gorge	-25.167	150.011	90	92	0.149	0.474
CQUA6	Arcadia	-25.258	148.821	91	93	0.204	0.278
CQUA5	Arcadia	-25.258	148.821	92	94	0.310	-0.094
CQUA4	Arcadia	-25.258	148.821	93	95	0.202	0.284
CQUA1	Arcadia	-25.258	148.821	94	96	0.198	0.299
CQUA2	Arcadia	-25.258	148.822	95	97	0.195	0.312
CQUA3	Arcadia	-25.258	148.822	96	98	0.206	0.272
CQULM6	Lake Murphy	-25.474	149.665	97	99	0.134	0.526
CQULM4	Lake Murphy	-25.474	149.665	98	100	0.143	0.494
CQULM2	Lake Murphy	-25.474	149.665	99	101	0.152	0.461
CQULM5	Lake Murphy	-25.474	149.665	100	102	0.180	0.362
CQULM3	Lake Murphy	-25.474	149.665	101	103	0.142	0.497
CQULH4	Lonesome Holding	-25.488	148.826	102	104	0.309	-0.092
CQULH5	Lonesome Holding	-25.488	148.825	102	104	0.329	-0.162
CQULH3	Lonesome Holding	-25.488	148.826	102	104	0.311	-0.099

CQULH2	Lonesome Holding	-25.488	148.826	102	104	0.329	-0.163
CQULH6	Lonesome Holding	-25.488	148.825	103	105	0.301	-0.063
CQULH1	Lonesome Holding	-25.488	148.826	104	106	0.198	0.299
DNRLHL4	Waikola Low	-25.988	149.151	105	107	0.199	0.297
DNRNHL3	Waikola Low	-25.988	149.151	105	107	0.196	0.308
DNRNHL1	Waikola Low	-25.988	149.151	106	108	0.194	0.314
DNRNHL2	Waikola Low	-25.988	149.151	106	108	0.195	0.311
DNRNHL5	Waikola Low	-25.988	149.151	107	109	0.191	0.325
DNRNHL6	Waikola Low	-25.988	149.151	107	109	0.184	0.350
DNRNH1	Waikola High	-26.030	149.151	108	110	0.206	0.271
DNRNH2	Waikola High	-26.030	149.151	109	111	0.209	0.261
DNRNH3	Waikola High	-26.030	149.151	110	112	0.201	0.289
DNRNH4	Waikola High	-26.030	149.151	111	113	0.193	0.320
DNRNH5	Waikola High	-26.030	149.151	112	114	0.203	0.284
DNRNH6	Waikola High	-26.030	149.151	113	115	0.313	-0.107
CQUYN2	Yuleba North	-26.593	149.466	114	116	0.211	0.255
CQUYN3	Yuleba North	-26.594	149.466	115	117	0.209	0.260
CQUYN1	Yuleba North	-26.594	149.465	116	118	0.205	0.275
CQUYN4	Yuleba North	-26.594	149.466	117	119	0.207	0.268
CQUYN5	Yuleba North	-26.594	149.466	118	120	0.204	0.279
CQUYN6	Yuleba North	-26.594	149.466	119	121	0.196	0.307
CQUYS1	Yuleba South	-26.916	149.753	120	122	0.199	0.296
CQUYS2	Yuleba South	-26.916	149.753	121	123	0.201	0.291
CQUYS3	Yuleba South	-26.916	149.753	122	124	0.208	0.266
CQUYS4	Yuleba South	-26.916	149.753	123	125	0.299	-0.056
CQUYS5	Yuleba South	-26.917	149.753	124	126	0.190	0.328
CQUYS6	Yuleba South	-26.917	149.753	125	127	0.205	0.277
CQUS5	Southwood	-27.811	150.090	126	128	0.152	0.462
CQUS6	Southwood	-27.811	150.089	127	129	0.134	0.527
CQUS3	Southwood	-27.811	150.090	128	130	0.156	0.450
NSW1084051	Newell Hwy x Karingal Rd	-28.864	150.201	129	131	0.201	0.291
NSW1084052	Newell Hwy x Karingal Rd	-28.864	150.201	130	132	0.212	0.251
NSW1084053	Newell Hwy x Karingal Rd	-28.864	150.201	131	133	0.204	0.280
NSW1084054	Newell Hwy x Karingal Rd	-28.864	150.201	132	134	0.194	0.313
NSW1084055	Newell Hwy x Karingal Rd	-28.864	150.201	133	135	0.200	0.292
NSW1084056	Newell Hwy x Karingal Rd	-28.864	150.201	134	136	0.209	0.262
NSW1084188	Careunga NR	-29.022	149.857	135	137	0.210	0.259
NSW1084189	Careunga NR	-29.022	149.857	135	137	0.211	0.255
NSW1084191	Careunga NR	-29.022	149.857	135	137	0.204	0.280
NSW1084192	Careunga NR	-29.022	149.857	135	137	0.204	0.278
NSW1084190	Careunga NR	-29.022	149.857	135	138	0.205	0.276
NSW1084197	Careunga NR	-29.022	149.857	135	139	0.207	0.270
NSW1084095	Warialda Rd	-29.085	150.733	136	140	0.207	0.267
NSW1084098	Warialda Rd	-29.085	150.733	136	140	0.206	0.271
NSW1084099	Warialda Rd	-29.085	150.733	136	140	0.208	0.264
NSW1084100	Warialda Rd	-29.085	150.733	136	140	0.208	0.263
NSW1084096	Warialda Rd	-29.085	150.733	137	141	0.201	0.290
NSW1084097	Warialda Rd	-29.085	150.733	138	142	0.196	0.306
NSW1076115	Croppa Ck	-29.164	150.281	139	143	0.200	0.293
NSW1076114	Croppa Ck	-29.164	150.281	140	144	0.206	0.274
NSW1076113	Croppa Ck	-29.164	150.281	141	145	0.187	0.340
NSW1076137	Croppa Ck	-29.164	150.281	142	146	0.208	0.265
NSW1076138	Croppa Ck	-29.164	150.281	143	147	0.188	0.335
NSW1076116	Croppa Ck	-29.164	150.281	144	148	0.194	0.316
NSW1083848	Ledknapper State Forest	-29.180	146.269	145	149	0.184	0.349
NSW1083851	Ledknapper State Forest	-29.180	146.270	146	150	0.189	0.333
NSW1083852	Ledknapper State Forest	-29.180	146.270	147	151	0.190	0.329
NSW1083850	Ledknapper State Forest	-29.180	146.270	148	152	0.202	0.286
NSW1083847	Ledknapper State Forest	-29.180	146.270	149	153	0.185	0.346
NSW1083849	Ledknapper State Forest	-29.181	146.270	150	154	0.105	0.628
NSW1084182	Gwydir Wetlands SCA	-29.278	149.293	151	155	0.097	0.656
NSW1084186	Gwydir Wetlands SCA	-29.278	149.293	151	155	0.205	0.275
NSW1084183	Gwydir Wetlands SCA	-29.278	149.293	152	156	0.102	0.639
NSW1084184	Gwydir Wetlands SCA	-29.278	149.293	153	157	0.203	0.284

NSW1084185	Gwydir Wetlands SCA	-29.278	149.293	153	157	0.204	0.278
NSW1084187	Gwydir Wetlands SCA	-29.278	149.293	153	158	0.193	0.317
NSW1076470	Biniguy	-29.538	150.214	154	159	0.195	0.312
NSW1076471	Biniguy	-29.538	150.214	154	159	0.199	0.298
NSW1076472	Biniguy	-29.538	150.214	154	159	0.201	0.290
NSW1076473	Biniguy	-29.538	150.214	154	159	0.201	0.288
NSW1076475	Biniguy	-29.538	150.214	154	159	0.203	0.282
NSW1076476	Biniguy	-29.538	150.214	154	159	0.198	0.300
NSW1076945	Gurley	-29.834	149.788	155	160	0.204	0.280
NSW1076949	Gurley	-29.834	149.788	155	160	0.206	0.271
NSW1076953	Gurley	-29.834	149.788	155	160	0.200	0.293
NSW1076969	Gurley	-29.834	149.788	156	161	0.199	0.296
NSW1076972	Gurley	-29.834	149.788	156	162	0.198	0.299
NSW1084028	Gurley	-29.834	149.788	157	163	0.204	0.279
NSW1093768	Brigalow SCA	-30.410	149.594	158	164	0.205	0.274
NSW1093763	Brigalow SCA	-30.411	149.595	159	165	0.193	0.320
NSW1093748	Brigalow SCA	-30.411	149.595	160	166	0.298	-0.053
NSW1093758	Brigalow SCA	-30.411	149.597	161	167	0.165	0.417
NSW1093749	Brigalow SCA	-30.411	149.598	162	168	0.194	0.313
NSW1093734	Brigalow SCA	-30.411	149.598	163	169	0.201	0.291
NSW1083905	Quambone Rd	-30.739	147.801	164	170	0.318	-0.124
NSW1083910	Quambone Rd	-30.740	147.803	164	170	0.318	-0.123
NSW1083909	Quambone Rd	-30.739	147.803	165	171	0.332	-0.172
NSW1083906	Quambone Rd	-30.739	147.801	165	171	0.332	-0.173
NSW1083907	Quambone Rd	-30.740	147.802	165	171	0.320	-0.131
NSW1083911	Quambone Rd	-30.740	147.802	165	171	0.321	-0.133
NSW1172854	Wilgaroon Rd	-31.119	145.847	166	172	0.175	0.382
NSW1172855	Wilgaroon Rd	-31.119	145.847	167	173	0.179	0.369
NSW1172856	Wilgaroon Rd	-31.119	145.847	168	174	0.168	0.406
NSW1172826	Wilgaroon Rd	-31.120	145.848	169	175	0.177	0.376
NSW1172857	Wilgaroon Rd	-31.120	145.847	170	176	0.182	0.358
NSW1172870	Wool Tk	-31.920	145.141	171	177	0.203	0.282
NSW1172874	Wool Tk	-31.920	145.142	171	177	0.198	0.302
NSW1083368	Yathong NR	-32.692	145.624	172	178	0.296	-0.047
NSW1083325	Yathong NR	-32.693	145.624	172	178	0.287	-0.016
NSW1083324	Yathong NR	-32.693	145.623	172	178	0.294	-0.040
NSW1083323	Yathong NR	-32.693	145.623	172	178	0.301	-0.065
NSW1083322	Yathong NR	-32.693	145.623	172	178	0.295	-0.043
NSW1083321	Yathong NR	-32.693	145.623	172	178	0.299	-0.058
NSW1083371	Yathong NR	-32.695	145.626	173	179	0.277	0.021
NSW1083370	Yathong NR	-32.695	145.625	173	179	0.279	0.014
NSW1083372	Yathong NR	-32.695	145.625	174	180	0.292	-0.034
NSW1083326	Yathong NR	-32.695	145.625	174	180	0.292	-0.031
NSW1083327	Yathong NR	-32.695	145.625	174	180	0.284	-0.003
NSW1083369	Yathong NR	-32.695	145.625	174	180	0.295	-0.042

5. References

Arnold, J. B. (2024). *ggthemes: Extra Themes, Scales and Geoms for "ggplot2"* (Version 5.1.0.9000) [Computer software]. https://github.com/jrnold/ggthemes

Gu, Z. (2022). Complex heatmap visualization. *iMeta*, *1*(3), e43. https://doi.org/10.1002/imt2.43

Jaccoud, D., Peng, K., Feinstein, D., & Kilian, A. (2001). Diversity Arrays: A solid state technology for sequence information independent genotyping. *Nucleic Acids Research*, *29*(4), e25. https://doi.org/10.1093/nar/29.4.e25

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

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., Caig, V., Heller-Uszynska, K., Jaccoud, D., Hopper, C., Aschenbrenner-Kilian, M., Evers, M., Peng, K., Cayla, C., Hok, P., & Uszynski, G. (2012). Diversity Arrays Technology: A Generic Genome Profiling Technology on Open Platforms. In F. Pompanon & A. Bonin (Eds.), *Data Production and Analysis in Population Genomics: Methods and Protocols* (pp. 67–89). Humana Press. https://doi.org/10.1007/978-1-61779-870-2_5

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, *26*(22), 2867–2873. https://doi.org/10.1093/bioinformatics/btq559

Rossetto M, Bragg J, Kilian A, McPherson H, van der Merwe M, Wilson PD (2019) Restore and Renew: a genomics-era framework for species provenance delimitation. *Restoration Ecology* 27(3):538–548. DOI: 10.1111/rec.12898

Rossetto M, PD Wilson, J Bragg, J Cohen, M Fahey, JYS Yap, M van der Merwe (2020) Perceptions of Similarity Can Mislead Provenancing Strategies – An Example from Five Co-Distributed *Acacia* Species. *Diversity* 12:306. DOI: 10.3390/d12080306

Weir, B. S., & Hill, W. G. (2002). Estimating F-Statistics. *Annual Review of Genetics*, *36*(1), 721–750. https://doi.org/10.1146/annurev.genet.36.050802.093940

Westgate, M., Stevenson, M., Kellie, D., & Newman, P. (2023). *galah: Atlas of Living Australia (ALA) Data and Resources in R* [Computer software]. https://CRAN.R-project.org/package=galah

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326–3328. https://doi.org/10.1093/bioinformatics/bts606