Final Report

Koala Genetic Analysis:

Yarrabilba Study Site

Prepared by Lyndal Hulse (BAppSc MScAg) and Sean FitzGibbon
for the University of Queensland Koala Ecology Group

1 EXECUTIVE SUMMARY

This report presents the findings of a study into koala population genetics from a research site at Yarrabilba, South East Queensland. Koalas (N = 25) from the site were analysed for genetic diversity, relatedness and population structure, based on 30 microsatellite markers.

The koala population at Yarrabilba showed moderate-to-high genetic diversity, with expected heterozygosity of the population greater than 0.7. When compared with other regional Queensland koala populations using the same genetic markers, the Yarrabilba koala population exhibited the most diverse number of alleles per locus (allelic richness = 7.34).

The results suggest that there is weak to moderate genetic differentiation between the Yarrabilba koala population and regional Queensland koala populations, indicating evidence of gene flow between populations. This limited amount of genetic differentiation between the examined Queensland koala populations may be due to natural movement (immigration/emigration) and possibly human-mediated animal translocation. STRUCTURE analysis identified three genetic clusters of koalas both within the Yarrabilba koala population and regional Queensland populations. Koalas from Yarrabilba showed majority genetic concordance with koalas from the Gold Coast, which is unsurprising given their geographical proximity.

Table of Contents

2 GENETIC ANALYSIS

2.1 Genotypes

Genotypes across 30 microsatellite loci were generated for koalas from the Yarrabilba study site, using genomic DNA extracted from koala tissue (N = 19 individuals) and faecal scats (N = 7 scat samples). Scat samples were collected from the base of trees so the identity of koalas that produced the scats were unknown. Analysis revealed an 85% genotype match between DNA extracted from one tissue sample and one faecal scat sample, suggesting the samples were from the same individual (animal ID 13496 and faecal scat ID 679). Therefore, data generated from the faecal scat were removed from the analysis to avoid duplication of the genotype profile. This resulted in a final genotype dataset for 25 individual koalas.

2.2 Genetic Diversity

Genetic diversity is the variability of genes in a species; high genetic variability is associated with the potential fitness of a population and ultimately its long-term persistence. In population genetics, the concept of heterozygosity is commonly extended to refer to the population as a whole, i.e., the fraction of individuals in a population that are heterozygous for a particular locus. It can also refer to the fraction of loci within an individual that are heterozygous. High heterozygosity (closer to 1.0) indicates high genetic variability, whereas, low heterozygosity (closer to 0.0) means little genetic variability.

Gene diversity is affected by two elements; 1) the number of alleles and 2) the abundance (or evenness) of the alleles. Increases in either of these leads to an increase in the expected heterozygosity. If a population consists of an excess of homozygotes for different alleles this leads to a low observed heterozygosity but does not affect the expected heterozygosity calculated from Hardy-Weinberg Equilibrium.

Analysis of genetic diversity was performed using the software GENALEX version 6.5 (Peakall and Smouse, 2012) to calculate mean number of alleles and observed and expected

heterozygosity. FSTAT (Goudet, 2001) was used to calculate allelic richness, a measure of allelic diversity that takes into account differences in sample sizes by standardising to the smallest number of individuals typed for a locus in a sample, so as to enable comparison among populations. FSTAT was also used to estimate the inbreeding coefficient (F_{IS}) for which a positive value indicates that individuals in a population are more related than would be expected under a model of random mating, and a negative value indicates that individuals in a population are less related than expected.

Genetic diversity values, estimated through expected heterozygosity and allelic richness, were compared between samples obtained from the Yarrabilba koala population and previously examined regional Queensland koala populations, using the same genetic markers (Table 1). These data indicate that the tested Yarrabilba koalas had similar expected heterozygosity to the most diverse of the tested regional koala populations, and was considerably higher than the examined Oakey, St Bees and Hidden Vale populations. In addition, the Yarrabilba population had greater allelic richness (7.34) than all the other seven regional populations. The inbreeding coefficient for Yarrabilba (F_{IS} = 0.168) was slightly higher than the overall mean for the examined populations (Table 1).

Table 1. Comparison of genetic diversity statistics within Queensland koala populations (Allelic richness was estimated for n=13).

Population	N	A_{mean}	A_r	F_{IS}	H_o	H_e
Yarrabilba study site	26	8.92	7.34	0.168	0.630	0.756
Sunshine Coast (SEQ)	171	11.15	7.13	0.146	0.645	0.760
Gold Coast (SEQ)	210	11.77	7.21	0.185	0.627	0.772
Clarke-Connors Range (central Qld)	54	9.65	7.25	0.231	0.582	0.737
Mt Byron (SEQ)	39	8.00	6.37	0.117	0.623	0.702
Oakey (SEQ)	16	6.19	5.88	0.042	0.623	0.662
St Bees Island (central Qld)	40	6.96	5.15	0.140	0.544	0.624
Hidden Vale (SEQ)	26	7.00	5.97	0.098	0.618	0.699

N: Number of individuals sampled **A_{mean}**: Mean number of alleles per locus

A_r: Allelic richness **H_o**: Observed heterozygosity **H_e**: Expected heterozygosity

F_{IS}: Inbreeding coefficient - the proportion of variance in a population that is contained within an individual; F_{IS} >0 indicates high levels of homozygosity and can be indicative of inbreeding.

2.3 Pairwise Genetic Differentiation (F_{ST})

Restrictions to gene flow among populations results in a genetic differentiation or divergence of the populations. F_{ST} is a measure of population genetic differentiation that quantifies the proportion of variance in allele frequencies among populations relative to the total variance. As a measure of genetic differentiation among populations, F_{ST} is calculated to evaluate how genetically different koala populations are to one another. A common reason for populations becoming more genetically different is reduced breeding movements of koalas among populations. The greater the genetic differentiation between populations, the less breeding movements there are between them and the more isolated they are from one another. F_{ST} can range from zero to one, where zero means populations show no genetic separation; a value of 0.25 or greater indicates strong differences among populations.

Assessment of genetic differentiation between koala populations was calculated using FSTAT (Goudet, 2001). Table 2 presents genetic differentiation between the Yarrabilba koala population and pairs of regional Queensland populations. There is moderate to strong differentiation between the Yarrabilba and St Bees koala populations, as expected given the geographical distances between populations and isolation of the island population on St Bees.

Table 2. Pairwise F_{ST} values of Queensland koala populations.

	Sunshine Coast	Gold Coast	Clark Connors	Mt Byron	Oakey	St Bees	Hidden Vale
Yarrabilba	0.061	0.049	0.089	0.088	0.105	0.149	0.065
Sunshine Coast		0.046	0.072	0.056	0.079	0.121	0.057
Gold Coast			0.089	0.085	0.091	0.135	0.065
Clark Connors				0.093	0.100	0.072	0.101
Mt Byron					0.077	0.154	0.099
Oakey						0.169	0.117
St Bees							0.154

<0.05 = **weak** genetic differentiation

0.05-0.15 = **moderate** genetic differentiation

0.15-0.25 = **strong** genetic differentiation

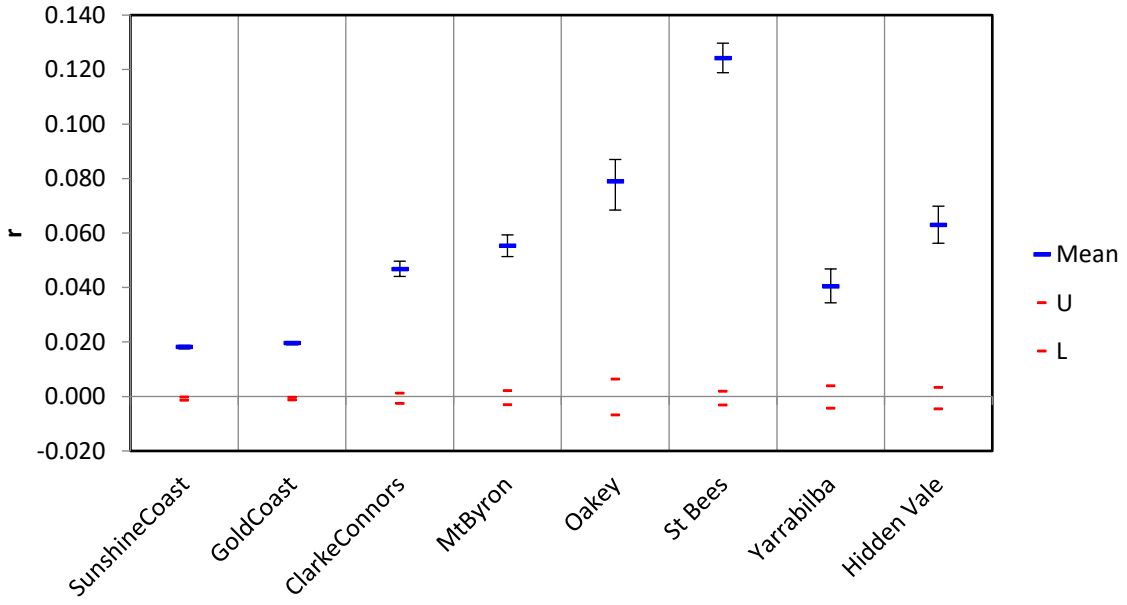
>0.25 = **very strong** genetic differentiation

2.4 Genetic Relatedness

Genetic relatedness was estimated to indicate the proportion of shared ancestry in pairs of individuals. Expected values are 0.5 for parent-offspring or full-sib pairs and 0.25 for half-sib pairs. However, genetic relatedness values will form a distribution around these expected values. Genetic relatedness of within-population individuals was calculated in GENALEX version 6.5 (Peakall and Smouse, 2012) using the Queller and Goodnight estimator of relatedness.

Genetic relatedness was estimated for regional Queensland populations, including Yarrabilba, separately and presented in Figure 1. Noticeably, all koala populations showed a mean relatedness that was higher than the confidence intervals, indicating koalas are significantly more related than expected. In addition, the standard deviation also fell outside the confidence interval for all populations.

Figure 1. Mean genetic relatedness for examined regional koala populations. The red lines indicate the upper (U) and lower (L) 95% confidence interval expected for that population under the null hypothesis of no difference among populations; r = relatedness.



2.5 Population Structure

The clustering of koalas into genetic populations, termed population structuring, was determined using the Bayesian clustering program STRUCTURE version 2.3.4 (Pritchard et al. 2000). STRUCTURE implements a model-based clustering method for inferring population structure using genotype data of unlinked markers. This method demonstrates the presence of population structure, identifies distinct genetic populations and assigns individuals to populations or clusters without any prior information about geographical location. The notion of a genetic cluster is that individuals within the cluster share on average more similar allele frequencies to each other than to those in other clusters.

Analysis of koala population genotype data involved 5 replicates of K = 1 to K = 10 (K = genetic cluster) using 100,000 iterations with 100,000 iterations discarded as burn-in. The number of K clusters was determined using both the maximum likelihood and the deltaK method of Evanno et al. (2005).

STRUCTURE analysis identified three genetic clusters of koalas both within the Yarrabilba koala population and regional Queensland populations. (K = 3, Figures 2 and 3). Koalas from Yarrabilba showed majority concordance with koalas from the Gold Coast, further confirming the least amount of genetic differentiation between the two populations, as presented in Table 2. However, there is some evidence of genetic diversity within all populations, indicating some gene flow between populations.

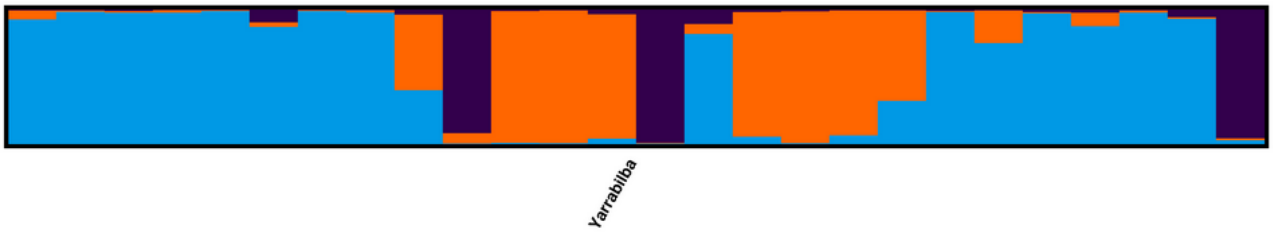


Figure 2. Population substructure of Yarrabilba koala population using STRUCTURE based on 30 loci. K = 3. Each bar represents an individual koala and colours indicate the proportion of the population cluster to which an individual was assigned.

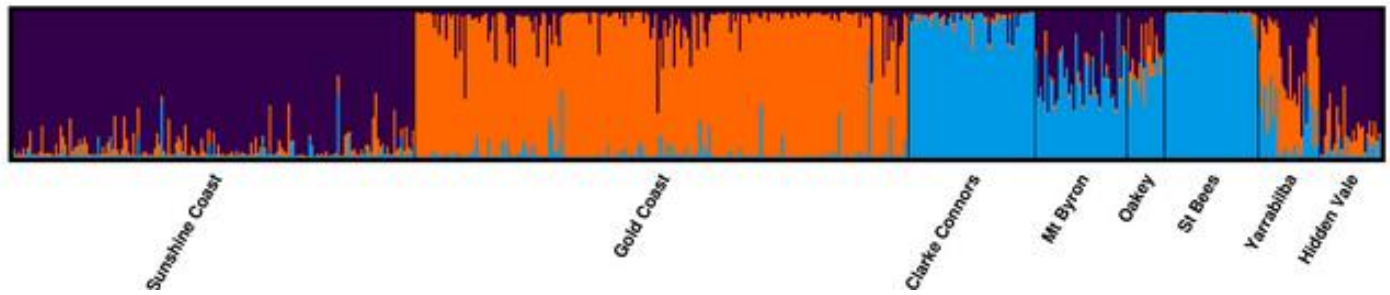


Figure 1. Comparison of population substructure of regional koala populations using STRUCTURE based on 30 loci. K = 3. Each bar represents an individual koala and colours indicate the proportion of the population cluster to which an individual was assigned.

3 REFERENCES

Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 14, 2611-2620.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available at http://www2.unil.ch/popgen/softwares/fstat.htm.

Peakall R and Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28, 2537-2539.