Genetic assessment of koalas from Wang Wauk State Forest for Bunyah Koala Project

Report prepared for Bunyah Restoration Project Incorporated

by Detection Dogs for Conservation, The University of the Sunshine Coast

Dr Ajith Horane Karayalage, Dr Katrin Hohwieler, Dr Alexis Levengood and Dr Romane Cristescu

July 2024

Acknowledgements

We acknowledge the First Nations people of the land where we work, play and live. We acknowledge their spiritual and cultural belonging and recognise their continuing connection to land, waters, and culture. We pay respect to Elders past, present and emerging.

Disclaimer

Document version control

Version	Date	Author	Reviewers
V1	31/05/2024	Ajith Horane Karayalage and Romane Cristescu	Katrin Hohwieler
V2	13/06/2024	Ajith Horane Karayalage	Katrin Hohwieler
V3	14/07/2024	Ajith Horane Karayalage	Katrin Hohwieler and Romane Cristescu
V4_issued	23/07/2024	Ajith Horane Karayalage	Romane Cristescu

Table of Contents

Acronyms

Acronym	Meaning
DArT	Diversity Arrays Technology®, Canberra
DDC	Detection Dogs for Conservation
DNA	Deoxyribonucleic acid
NSW	New South Wales
PCA	Principal component analysis
SE	Standard Error
UniSC	University of the Sunshine Coast

Executive Summary

Bunyah Restoration Project Incorporated contracted the University of the Sunshine Coast's Detection Dogs for Conservation team (UniSC's DDC) to genetically analyse a set of koala scat samples collected for the Bunyah Koala Project – a project primarily aiming to protect the local koala population while demonstrating how farming and conservation can be mutually beneficial. The assessment of koalas in this area will be critical in informing regional koala management plans and guidelines. Samples were collected on lands adjacent to and partially including the Wang Wauk State Forest, an Area of Regional Koala Significance (ARKS) in the Mid North Coast of New South Wales (NSW). Samples had been collected between November 2022 and August 2023 by Bunyah Restoration Project Incorporated and provided to DDC in mid-October 2023. Here, we report analyses and results of koala sex, prevalence of *Chlamydia* and genetic diversity for koalas based on samples provided from the Bunyah Koala Project area. Further samples collected from several other locations in Mid Coast Council in NSW were analysed and are reported separately in Appendix A, due to the fact they were from distant locations from the Bunyah Koala Project area, they were scattered across a large area and they had a small sample size.

A total of 25 samples collected within the Bunyah Koala Project area (hereafter referred as Bunyah) were delivered to the UniSC laboratory, where deoxyribonucleic acid (DNA) was extracted. For 21 samples two or more scats were present in the sample tube, hence DNA was extracted in two replicates using two different scats from the sample tube. The remaining four samples only had one scat in each sample tube allowing for only single DNA extractions. Together, a total of 46 DNA extractions were sent for genotyping to Diversity Arrays Technology® (DArT), in Canberra.

Of the 25 samples, six samples had to be excluded because data quality was insufficient for genetic fingerprinting and further analyses (i.e. DNA extractions failed in genotyping quality control for each extraction from the six samples). The samples likely failed due to not originating from koala. Further, genetic fingerprinting identified three cases where a sample tube contained scats from two different individual koalas, hence replicated DNA extractions of those three samples identified separate individuals for each replicate. Altogether, out of the 25 samples originally presented, 18 unique koalas were identified. Only results from unique kolas

are used in genetic calculation e.g. sex ratio, *Chlamydia* prevalence, heterozygosity and inbreeding. These unique individuals included 12 males and six females with a sex ratio of 1.0:0.5 male to female.

Chlamydia pecorum was detected in 17 of the 18 koalas, resulting in 94% prevalence of infection. While infection does not necessarily develop into disease, the fact that in Bunyah *Chlamydia* is present in all but one of the sampled individuals means this group of koalas could be severely threatened by chlamydial disease.

There were 16 unique koalas with samples of sufficient data quality for genetic diversity estimates. Population structure analyses indicated a panmictic group of koalas, e.g. all koalas are one breeding population. This group of koalas had a comparatively high level of heterozygosity and a low inbreeding coefficient, suggesting a genetically healthy group of koalas.

Limitations

- Quality of DNA derived from non-invasive samples (such as scats) is lower than high quality blood samples, and the quality of non-invasive samples is more variable overall. The DDC has developed techniques that maximise the ability to retain samples with lower quality DNA during molecular research; however, it is common that some samples result in too little genotyped data to be used for subsequent analyses. These samples need to be excluded to ensure the most accurate estimates of genetic measures are obtained.
- Due to the nature of non-invasive sampling, it is common that duplicate samples are collected, i.e. two or more scat samples that originate from the same koala. These need to be identified and removed from analyses so they do not bias the results and produce unreliable findings.
- The prevalence of *Chlamydia* (i.e. percentage of sampled koalas with the pathogen) identified through genetic analyses indicates the presence of infection. However, the presence of chlamydial infection does not necessarily equate to the presence of chlamydial disease and its severity.

- The sample size for Mid Coast Council was small and samples were scattered across a large area distant from/to the Bunyah Koala Project area, restricting a reliable co-analysis. Therefore, samples from Bunyah were analysed separately.
- While the DDC had no influence over the scat collection and initial storage methods, care has been taken to ensure no contamination or damage occurred to the samples once they had been received.

1. Background

Bunyah Restoration Project Incorporated contracted the University of the Sunshine Coast's Detection Dogs for Conservation team (UniSC's DDC) to genetically analyse and report the results of koala sex, prevalence of *Chlamydia* and genetic diversity of koala scat samples collected for the Bunyah Koala Project – a project primarily aiming to protect the local koala population while demonstrating how farming and conservation can be mutually beneficial. The assessment of the Bunyah koala population will be critical in informing regional koala management plans and guidelines. Samples were collected on lands adjacent to and partially including the Wang Wauk State Forest, an Area of Regional Koala Significance (ARKS) in the Mid North Coast of New South Wales (NSW) by Bunyah Restoration Project Incorporated between November 2022 and August 2023 and delivered to UniSC by a commercial courier service in mid-October 2023.

2. Methodology

2.1 Koala scat samples

A total of 35 scat samples were received by DDC from Bunyah Restoration Project Incorporated, including samples collected within the Bunyah Restoration Project area (N = 25) and from several other locations in Mid Coast Council in NSW (N = 10). The samples collected outside of the Bunyah Restoration Project area were analysed and are reported separately in Appendix A, due to the fact they were from distant locations from the Bunyah Koala Project area, they were scattered across a large area and they had a small sample size. The sample locations of the 25 samples collected within the Bunyah Restoration Project area (hereafter referred as Bunyah) between 26/11/2022 and 23/08/2023, are given in Figure 1 and Appendix B, Table B1, which also shows additional sample information. Scat samples were delivered on dry ice in mid-October 2023 and transferred to a -20°C freezer immediately on arrival and stored until processing for deoxyribonucleic acid (DNA) extractions.

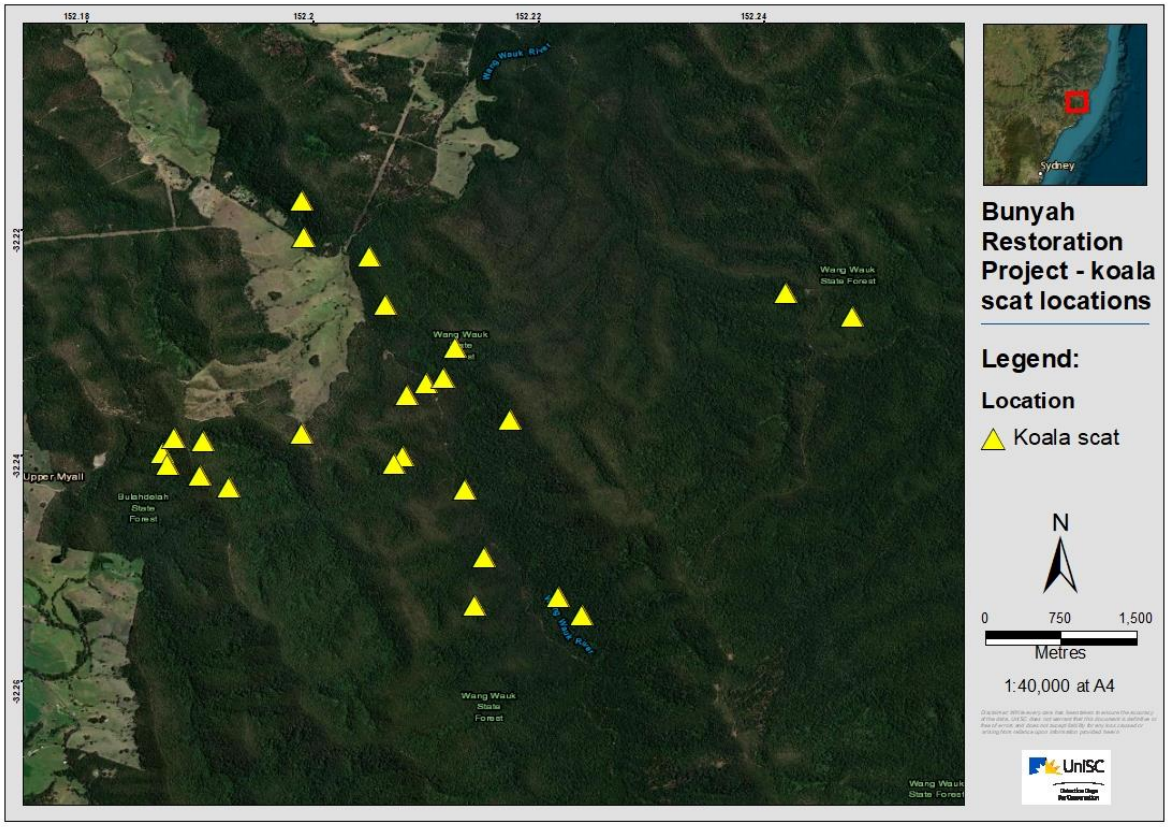


Figure 1: Distribution and location of scats collected from the Bunyah Koala Project area.

DNA quality is generally higher when extracted from fresh koala scats (Schultz, Cristescu et al. 2018). Fresh scats (i.e. when the scat age is estimated to be less than one week old, categories 1 and 2, Table 1) present a shiny mucus layer and a strong smell. However, no records of scat ages were provided for the current set of samples.

Table 1: General guide used to age koala scats in the field.

Scat age categories	Age	Characteristics
1	One day old or less	Very fresh (covered in mucus, wet)
2	Few days old	Fresh (shine and smell)
3	Couple of weeks old	Medium fresh (shine or smelly when broken)
4	Months old	Old (no shine, no smell)
5	More than a few months old	Very old and discoloured

2.2 Extracting DNA

All 25 samples were processed for DNA extraction. Four samples only had one scat in the sampling tube, only allowing for one DNA extraction each. For all others (N = 21) DNA extractions were replicated with a second scat from the sampling tube. In total, 46 DNA extractions were performed (see Appendix C, Table C1 for list of samples and DNA extractions). This was done so that each sample could be genotyped twice in order to maximise availability and quality of genetic data for analyses. We followed the protocol of Schultz, Cristescu et al. (2018) to extract DNA from koala scats. However, instead of scraping the outer layer off the scats, we used a lysis wash to rinse the DNA off the surface of the scats. This faecal sample wash was then processed using the QIAamp PowerFecal Pro DNA Kit (Qiagen), with the following modification to the manufacturer's protocol. After adding the buffer to the faecal sample wash, a one-hour incubation step (65°C) was added, and samples were vortexed for seven minutes at maximum speed using Genie 2 Vortex Mixer (Scientific Industries). Finally, DNA was eluted in 200 µl of elution buffer and concentrated down to a volume of ~30 µl. Extracted DNA was stored at -20°C until it was shipped to Diversity Arrays Technology® (DArT) in Canberra for genotyping.

2.3 Genotyping

DNA aliquots were genotyped using a next-generation sequencing protocol for detecting Single Nucleotide Polymorphisms (SNPs) by DArT (Jaccoud, Peng et al. 2001, Kilian, Wenzl et al. 2012). A targeted approach was chosen (DArTag), where specifically designed molecular probes (i.e. koala-specific capture probes) select small target regions containing sequence variants. A total of 4,393 koala SNPs were genotyped. In addition, sex and *Chlamydia* markers were also genotyped from the same DNA extractions, using the sex- and *Chlamydia pecorum*-specific probes developed in collaboration with DArT and integrated to the DArTag panel.

2.4 Filtering of genetic data

Genetic data were analysed using the R package *dartR* (Gruber, Unmack et al. 2019) in the R environment using R v4.1.0 (R Core Team 2018), unless specified. Genotyped data were filtered to improve the quality of the dataset by removing samples with too little data (i.e. those

with low individual call rate) as well as SNP loci that were not called across most samples (i.e. those with low locus call rate). We applied a stepwise increasing locus call rate threshold, from 0.2 to 0.8 – only retaining those SNPs with at least 80% data. When filtering for individual call rate, different filtering regimes were applied, depending on the analysis. This is because only 200 high-quality loci are needed to identify unique individuals (Schultz, Cristescu et al. 2018); however, many high-quality loci are required to measure genetic diversity. Therefore, to identify unique individuals, where the focus was on maximising the number of individuals that could be used while retaining sufficient high-quality SNPs, samples were filtered for an individual call rate threshold of 0.2. On the other hand, for genetic diversity analyses, where the focus was on maximising the number of high-quality loci while maintaining as many individuals as possible, samples were filtered using a stepwise approach, increasing individual call rate threshold from 0.2 to 0.5 – resulting in only retaining samples with at least 50% data.

Other constant thresholds were applied to remove potentially erroneous loci. This included filtering for allele read depth (minimum threshold of five), minor allele frequency (MAF, minimum threshold of 0.01) and loci appearing on the same contig as another (secondary loci). Because filtering can result in previously polymorphic loci becoming monomorphic, a filter to remove all monomorphic loci was applied at the end of the filtering protocol.

2.5 Genetic fingerprinting

Genetic fingerprinting allows for the allocation of scat samples to individual koalas. The total number of unique individuals identified with this technique was then used for population estimates of sex, *Chlamydia* prevalence, inbreeding and genetic diversity. Genetic fingerprinting enabled the identification and elimination of multiple samples originating from the same individual koala, which would have otherwise biased those estimates.

SNPs filtered for an individual call rate threshold of 0.2 and a locus call rate threshold of 0.8 were used for genetic fingerprinting. Any sample that indicated a genetic relatedness value > 0.75 using the ‘dyadml’ method (Milligan 2003) from the *related* R package (Pew, Muir et al. 2015) was considered a duplicate sample and eliminated from further analyses.

2.6 Sex and sex ratio

Sex of individual koalas was determined through sex-linked genetic markers integrated into the DArTag panel. We then calculated the sex ratio, which is the relationship between the number of males to the number of females. A typical sex ratio in natural, healthy populations is expected to be close to 1:1. However, a good representation of the population, i.e. large sample size and good geographic spread of samples, is required to get a reliable value.

2.7 *Chlamydia* detection

Chlamydia detection in scats was based on the same DNA extraction described above. We then used *Chlamydia pecorum*-specific probes, developed and integrated into the DArTag panel, to determine the presence or absence of chlamydial DNA.

2.8 Population genetic structure and genetic diversity

Data filtered for a locus call rate threshold of 0.8 and an individual call rate threshold of 0.5 were used to measure the population genetic structure and genetic diversity indices. To identify the presence of population structure within the data set, principal component analysis (PCA) and genetic structure analysis were conducted using *dartR* package and STRUCTURE v2.3.4 software (Pritchard, Stephens et al. 2000), respectively. For the latter, the number of genetic clusters (K) was set to vary between 1 to 5 with 10 iterations and the most likely number of clusters was determined based on the best K from a parsimony index (Wang 2019) for the STRUCTURE outputs using KFinder (Wang 2022).

Genetic diversity was calculated using GenAlEx v6.5 (Peakall and Smouse 2012). We calculated three values: observed heterozygosity H_O, which is the level of heterozygosity from the allele frequencies of the population under study; expected heterozygosity H_E (adjusted for small sample size), which is the level of heterozygosity that could be expected based on observed allele frequencies if the population was at the Hardy-Weinberg equilibrium (panmictic population with constant genetic variation across generations); and lastly F_{IS}, also called inbreeding coefficient, which is the proportion of the variance in the subpopulation contained in an individual and which can range from -1 to 1 (the closer to 1, the higher the degree of inbreeding). Note that inbreeding can not only result from non-random mating but

also from small, isolated populations, where all individuals are more closely related than in large populations. Given the increasingly fragmented landscape koalas have to navigate, this second cause of inbreeding is becoming more common and important to investigate.

2.9 Limitations

Genotyping was conducted non-invasively from genetic material contained on the surface of koala scats. This allows for large-scale, relatively cheap, unbiased sampling of DNA compared to other available methods (e.g., catching koalas, anaesthetising them and collecting high-quality samples such as blood or biopsies, or relying on wildlife hospital samples). However, compared to high-quality blood/biopsy samples, DNA present in scat is of lower quantity and quality, which yields lower numbers of high-quality SNPs. DDC was able to optimise scat genotyping for koalas by developing a specific-probe approach, i.e. the DArTag method, which increased genotyping success, and the quality of data. However, data quality of non-invasive samples can only be improved to a certain degree, with some samples still containing insufficient data to be included in further analyses. To maximise data derived from the non-invasive samples, all samples were extracted twice, in all instances where a minimum of two scats per sample (tube) was available.

Presence of duplicate samples (i.e. two or more samples originating from the same individual) can falsely inflate data, and collection of duplicate samples is common in non-invasive sampling methods. These samples need to be identified and removed to avoid producing unreliable findings. For example, if a koala with *Chlamydia* infection is sampled multiple times, it would artificially inflate *Chlamydia* prevalence, or if duplicate samples were kept, as they are genetically identical, they would falsely inflate measures of inbreeding in the population. Here, care has been taken to remove duplicate samples identified through genetic fingerprinting, retaining only the best quality sample from each cluster of duplicate samples for further analyses.

The prevalence of *Chlamydia* (i.e. the percentage of unique koalas with the pathogen) is an important population characteristic for informing conservation management. However, the presence and severity of chlamydial disease varies greatly between individual koalas, as well as between populations (Ellis, Girjes et al. 1993, Waugh, Hanger et al. 2016). Notably,

individual koalas can shed large numbers of *Chlamydia* organisms without clinical signs of disease (Wan, Loader et al. 2011), and populations can have high *Chlamydia* prevalence with minimal detectable health impacts. For instance, in the Mt Lofty ranges, 90% of koalas were *Chlamydia* positive but there was a low prevalence of clinical (symptomatic) disease (Polkinghorne, Hanger et al. 2013); see also Weigler, Girjes et al. (1988). Therefore, quantifying *Chlamydia* pathogen prevalence is only the first step in understanding the threat that this pathogen presents to an individual and a population.

The sample size available for Mid Coast Council was small and samples were scattered across a large geographical area. Further, samples were from locations distant to the Bunyah Koala Project area, preventing a combined analysis with Bunyah samples. They are therefore presented separately (Appendix A).

Sample collection was conducted by Bunyah Restoration Project Incorporated and Mid Coast Council, therefore DDC had no influence over the scat collection and initial storage methods. However, care has been taken to ensure no contamination or damage occurred to the samples once they had been received.

3. Results

3.1 Extraction, quality control and unique individuals

All samples were genotyped using DArTag. However, DNA quality varied, which is common when using non-invasive samples, and samples below analysis-specific quality thresholds were excluded from the analyses. Six samples were excluded from the analyses due to insufficient data, i.e. DNA extractions failed in genotyping for each of the six samples including their replicated extractions (Appendix C, Table C1 and see Table C2 for scat images from failed samples). Based on the visible characteristic of the scats, such as shape of the scats and size of visible fibres, it is likely that they failed due to being from another species.

Data filtration for identifying unique individuals (genetic fingerprinting) retained a total of 1,158 SNPs with an average of 4.6% missing data. Our extraction system (i.e. extracting in duplicates whenever at least two scats are available in one sample tube) and genetic analysis revealed that, for three tubes, two distinct individuals were present. This means that scats from

at least two individuals were collected into same sampling tube. Specifically, for samples K26, K31 and K36, two individuals per tube were identified. Four samples were found to be duplicates (i.e. scats collected from the same, already identified individuals) and were subsequently removed from further analyses, retaining only the best sample from each uniquely identified koala. Overall, out of the 25 samples originally presented, 18 unique koalas were identified (Table 2).

Table 2. List of unique and duplicate samples, as determined by genetic fingerprinting. Sample names with "_A" extension indicate those replicated DNA extraction that resulted in identification of two separate individuals.

Sample name	Duplicate sample name
K13	
K15	
K17	
K20	
K21	K10
K22	
K24	
K26_A	K36
K27	
K30	
K31	K32
K31_A	K26
K33	
K34	
K36_A	
K37	
K38	
K40	

3.2 Sex of unique individuals and sex ratio

Based on the sex-linked markers, of the 18 unique individuals, 12 (67%) were males and six (33%) were females (Table 3 and see Figure 2 for locations of each individual), translating to a sex ratio of 1.0:0.5 male to female, which denotes a male biased sample set.

Table 3. Sex of unique koalas. Sample names with "_A" extension represents unique individual koala from its original sample tube and not identified as a duplicate of its original sample.

Sample name	Sex
K13	M
K15	F
K17	F
K20	M
K21	M
K22	M
K24	F
K26_A	M
K27	F
K30	F
K31	M
K31_A	M
K33	M
K34	M
K36_A	M
K37	F
K38	M
K40	M

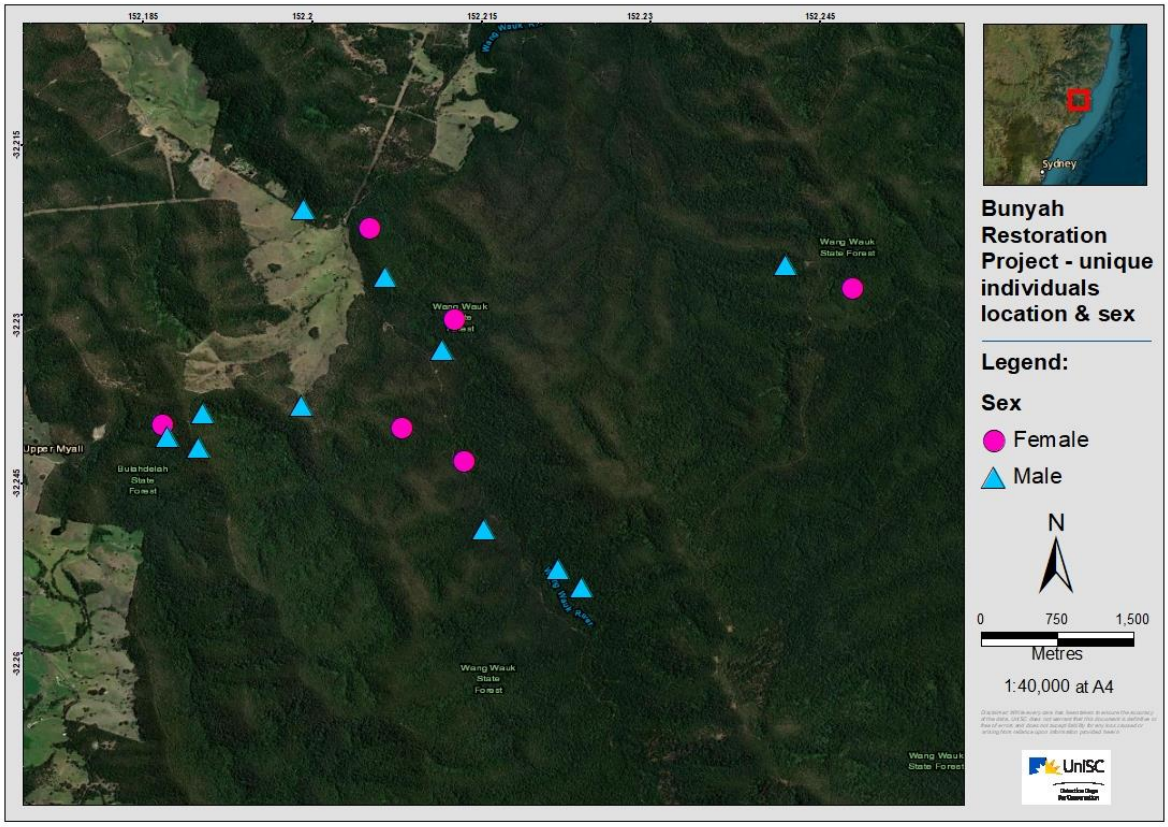


Figure 2: Distribution of unique individuals by sex. Locations for two koalas, namely K31 (male) and K31_A (male) were overlapped in the map.

3.3 *Chlamydia* prevalence

Of the 18 unique individuals, 17 (94%) were positive for *Chlamydia* (Table 4 and see Figure 3 for sample locations), indicating a large proportion of koalas carrying the infection. However, it should be noted that the presence of the *Chlamydia* pathogen does not necessarily equate to clinical signs of disease.

Table 4. *Chlamydia* status of unique koalas by sex. Sample names with "_A" extension represents a unique individual koala not identified as a duplicate of its original sample, but gives the name from its original sample tube (e.g. K26).

Sample name	Sex	*Chlamydia* status
K24	F	Negative
K15	F	Positive
K17	F	Positive
K27	F	Positive
K30	F	Positive
K37	F	Positive

K13	M	Positive
K20	M	Positive
K21	M	Positive
K22	M	Positive
K26_A	M	Positive
K31	M	Positive
K31_A	M	Positive
K33	M	Positive
K34	M	Positive
K36_A	M	Positive
K38	M	Positive
K40	M	Positive

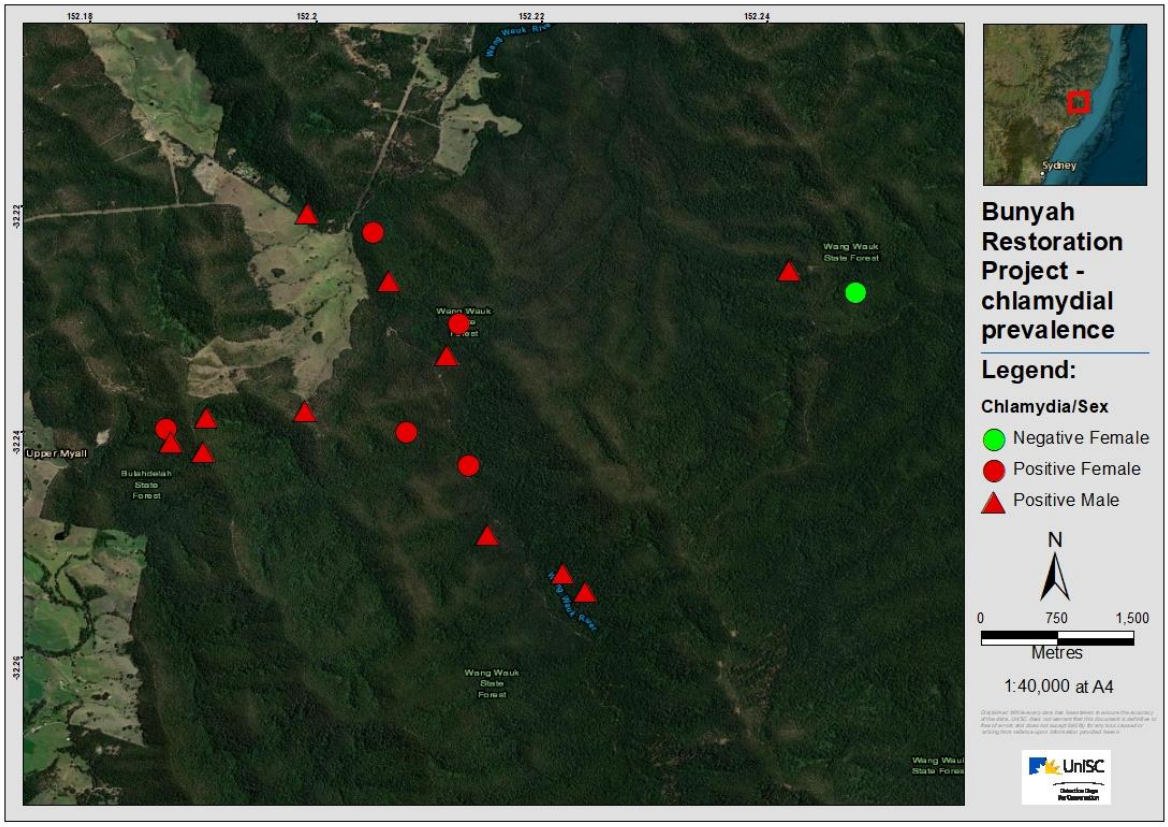


Figure 3: Distribution and *Chlamydia* status of unique individuals.

3.4 Population genetic structure and genetic diversity

After the application of more stringent filtering for the individual call rate, a further two samples were removed from the 18 unique individuals due to insufficient data, retaining 16 samples for analyses. A total of 2,691 loci were retained with 1.5% missing data.

Principal component analysis indicated the presence of at least two (K = 2) genetic clusters (Figure 4), separating two samples (K13 and K15, top left corner of the Figure 4) from rest of the samples. The first two principal components (two PCA axes in Figure 4) explained 17.6% of the total variance among samples. However, based on the best K from parsimony index for STRUCTURE results indicated only a single population (K = 1). Further, no significant difference was found between the two clusters observed in PCA (i.e. samples K13 and K15 vs all other samples) using pairwise F_{ST} (F_{ST} = 0.08, $P \geq 0.05$). Hence this set of samples were considered as one panmictic population for genetic diversity estimates.

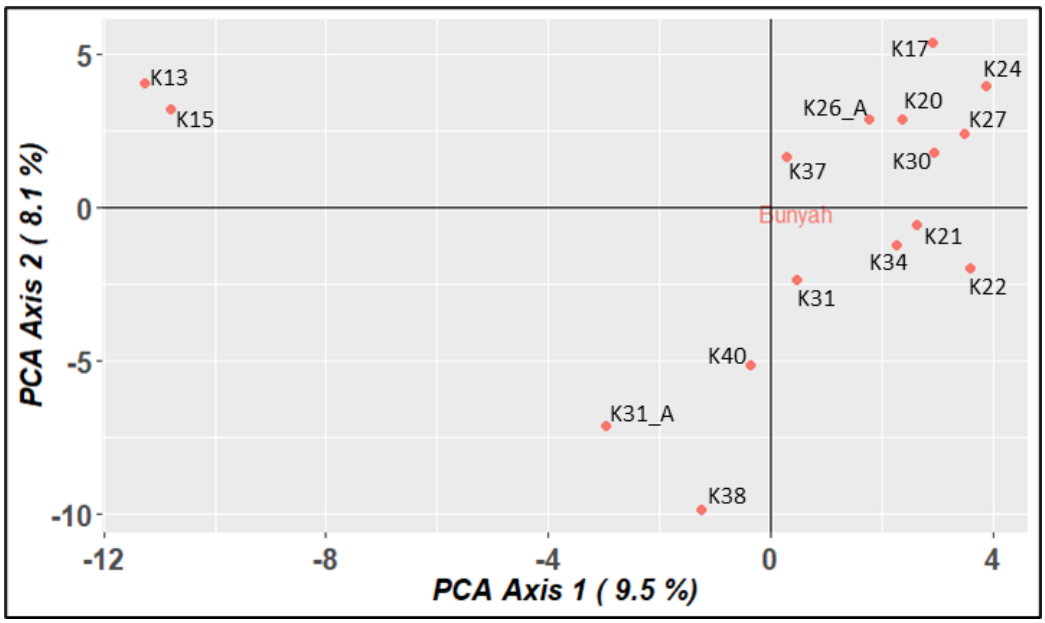


Figure 4: Results of the principal component analysis (PCA) for the 16 unique koalas, separating K13 and K15 samples (placed on the top left corner of the plot) from rest of the samples.

Three genetic diversity indices including observed heterozygosity (H_O), expected heterozygosity (H_E) and inbreeding coefficient (F_{IS}) were calculated and indicated a high level of heterozygosity and very low level of inbreeding (Table 5). These values were compared to those of other koala populations in the discussion (i.e. section 4).

Table 5: Genetic diversity indices for the 16 unique koalas: observed heterozygosity (H_O), expected heterozygosity (H_E) and inbreeding coefficient (F_{IS}). SE: standard error.

Parameter	Mean	SE
H_O	0.307	0.003
H_E	0.329	0.003
F_{IS}	0.066	0.005

4. Discussion

From the provided 25 samples, six samples were failed in genotyping quality control. Sample failures can be due to multiple reasons including highly degraded DNA in the samples, or failure to detect target host DNA, e.g. if a scat sample stems from a different species such as possum. Here based on the observed characteristics of the samples, it is likely some failed samples were from another species.

Three sample tubes contained scats from two different individuals. This is one of the difficulties associated with non-invasive sampling, especially when two or more individuals share the same tree. Generally, physical appearance of scats including scat size (particularly the diameter) and scat age should be considered to separate them at collection, which can minimise, but not completely eliminate, this issue.

4.1 Sex of unique individuals and sex ratio

Overall, a larger proportion (67%) of the individuals in the current sample set were males. Generally, while the sex ratio of a natural, healthy population is expected to be 1:1, a small bias toward females may be advantageous for conservation purposes, as larger female cohorts are associated with larger number of offspring, and therefore a larger population in the next generation. Here, the samples representing this group of koalas indicate a strong male bias, with a male to female ratio of 1.0:0.5. It is important to monitor the sex ratio of koalas from Bunyah to understand the potential causes for the low number of females, and a larger sample size would help to confirm or dispute this observation. A male biased sex ratio can have detrimental consequences for the conservation and management of this population, as females drive population growth, which can place small and isolated populations at risk of extinction (Lopez-Sepulcre, Norris et al. 2009, Grayson, Mitchell et al. 2014).

4.2 *Chlamydia* prevalence

We observed 94% prevalence of *Chlamydia* for koalas in this study, with 17 of the 18 koalas positive for *C. pecorum*. Overall, prevalence was higher than what has been found in some other populations for *C. pecorum* urogenital infections, including Mutdapilly (52%), Coombabah (10%) and Moreton Bay (27%) in Queensland, and Mount Lofty Ranges (47%) in

South Australia (Jackson, White et al. 1999, Nyari, Waugh et al. 2017, Fabijan, Caraguel et al. 2019). Further, the prevalence of *Chlamydia* infection found in this study is higher than that observed for other sites surveyed by DDC, such as Ngunya Jargoon Indigenous Protected Area in NSW (58%) in 2022, and Redland City Council (mainland) in Southeast Queensland (38%) in 2020-21. It is important to note that although the pathogen was detected in 94% of the sampled koalas, this does not necessarily reflect their chlamydial disease status. For instance, some koalas could have recovered from disease but were still carrying *Chlamydia* in their gastrointestinal tracts and others could be carrying the pathogen without any clinical signs (Robbins, Hanger et al. 2019). However, when *Chlamydia* infection does progresses into disease it can cause infertility and overall increased morbidity and mortality (Hulse, Beagley et al. 2021, Pagliarani, Johnston et al. 2022). This could have a large negative impact on the study population.

While we can report on chlamydial infection, veterinary examinations are required to detect and confirm chlamydial disease. Given the very high infection prevalence, an investigation into disease prevalence is strongly encouraged to assess the specific risk that this pathogen poses to this population.

4.3 Population genetic structure and genetic diversity

Although PCA analysis separated two individuals, namely K13 and K15, from rest of the koalas, further analyses did not find significant genetic difference between these two and the remaining koalas. Further, these two koalas were located very closer to the other koalas, at least within less than 1km distance. Hence, while it appears that this group of koalas share genetic materials from more than one ancestral population, we considered them to be one panmictic cluster.

Theoretically, heterozygosity values (i.e. observed and expected heterozygosity) range from zero to one. High heterozygosity means high genetic variability and diversity, and is therefore assumed to indicate higher resilience (e.g. higher chances of adapting to current and future challenges, including disease, which could be problematic in this population based on the high prevalence of *Chlamydia* infection) and evolutionary potential, characterising a genetically healthy population (Orsted, Hoffmann et al. 2019). Another sign of a healthy population is low

inbreeding values (Moss, Arce et al. 2007). In general, if the observed heterozygosity is lower than the expected heterozygosity, the discrepancy is attributed to inbreeding. The koalas studied in this project showed signs of a genetically healthy population, with high levels of genetic diversity (H_O = 0.307 and H_E = 0.329) and very low levels of inbreeding (F_{IS} = 0.066).

The genetic diversity values are best interpreted by comparing them to other populations where diversity measures were calculated using similar methods. In a previous study in the Northern Tablelands, NSW, in 2019–2020, DDC estimated diversity measures using DArTcap, a next generation sequencing method closely related to the current DArTag, for the Armidale/Uralla region and for the Inverell/Delungra region, and for koalas in Redland City Council (mainland), Southeast Queensland in 2020-21 (Table 6). Using the same methods as in this project (DArTag), we estimated diversity measures for koalas in the Ngunya Jargoon Indigenous Protected Area, NSW in 2022 (Table 6). Further comparisons can be made by consulting Table 7, which was taken from Kjeldsen, Zenger et al. (2016). This table shows genetic diversity measures from other wild koala populations across Queensland, New South Wales and Victoria, using a different set of SNPs obtained through double digest restriction-associated genotyping (DArTseq), compared to our koala-specific DArTag panel (a panel we developed to increase non-invasive sample data recovery and consistency).

Table 6: Genetic diversity measured through SNP sequencing in wild koala populations in NSW and QLD by DDC. N = sample size, H_O = observed heterozygosity, H_E = expected heterozygosity, F_{IS} = inbreeding coefficient.

Population	N	H_O	H_E	F_{IS}
Armidale/Uralla, NSW	36	0.23	0.28	0.20
Inverell/Delungra, NSW	40	0.23	0.28	0.18
Redland City Council (mainland), QLD	227	0.24	0.32	0.26
Ngunya Jargoon Indigenous Protected Area, NSW	20	0.30	0.33	0.08

Table 7: Genetic diversity established through double digest restriction-associated SNP sequencing in wild koala populations across QLD, NSW and Victoria. N = sample size, H_O = observed heterozygosity, H_E = expected heterozygosity, F_{IS} = inbreeding coefficient. Table taken from Kjeldsen et al. (2016).

State	Location	N	H_O	H_E	F_{IS} (P <0.01)
QLD	St Bees Island	19	0.29	0.35	0.23
QLD	St Lawrence	19	0.26	0.30	0.20

QLD	Koala Coast	24	0.22	0.30	0.32
QLD	Ipswich	23	0.27	0.31	0.19
NSW	Port Macquarie	45	0.23	0.28	0.21
NSW	Campbelltown	09	0.27	0.33	0.27
VIC	South Gippsland	19	0.24	0.30	0.27
VIC	Cape Otway	13	0.24	0.25	0.11

The observed heterozygosity values in the current study were higher than many populations listed in Table 6 and Table 7 but were comparable with, for instance, koalas in Ngunya Jargoon Indigenous Protected Area, NSW, and St Bees Island (Kjeldsen, Zenger et al. 2016). Most koala populations compared in Table 6 and Table 7 indicate higher F_{IS} than what we found in this study. The F_{IS} values found in this study are very low, a positive characteristic. Overall, such high levels of genetic heterozygosity and low inbreeding are positive indicators for this population.

It should be noted, however, that measures of genetic diversity and inbreeding come with an associated time-lag (Landguth, Cushman et al. 2010) and often, signs of decline in these measures only occur after the population has already experienced a major impact. Therefore, genetic diversity measures might not reflect current issues in a population.

Overall, measures of genetic diversity suggest a genetically healthy population. The study group of koalas had a higher level of heterozygosity than most populations we can compare it to, and a low inbreeding coefficient. However, while these are positive findings, we also found a very high prevalence of *Chlamydia* infections, which can put the population at risk. Therefore, we recommend the following:

1. A study that includes catching koalas for veterinary examinations to further understand chlamydial disease prevalence and its potential impact on koalas in this area (e.g. animal welfare, fertility and reproductive output). This should include treatment of diseased koalas.
2. Subsequent monitoring of koalas in this area for disease, to further understand short- and long-term effects of high *Chlamydia* infection rates and viability of the population.

5. References

Ellis, W. A. H., A. A. Girjes, F. N. Carrick and A. Melzer (1993). "Chlamydial infection in koalas under relatively little alienation pressure." Australian Veterinary Journal **70**(11): 427-428.

Fabijan, J., C. Caraguel, M. Jelocnik, A. Polkinghorne, W. S. J. Boardman, E. Nishimoto, G. Johnsson, R. Molsher, L. Woolford, P. Timms, G. Simmons, F. Hemmatzadeh, D. J. Trott and N. Speight (2019). "*Chlamydia pecorum* prevalence in South Australian koala (*Phascolarctos cinereus*) populations: Identification and modelling of a population free from infection." Scientific Reports **9**.

Grayson, K. L., N. J. Mitchell, J. M. Monks, S. N. Keall, J. N. Wilson and N. J. Nelson (2014). "Sex ratio bias and extinction risk in an isolated population of Tuatara (Sphenodon punctatus)." PLoS One **9**(4): e94214.

Gruber, B., P. Unmack, O. Berry and A. Georges (2019). "Introduction to dartR." User Manual **51**: 1-41.

Hulse, L., K. Beagley, R. Larkin, V. Nicolson, J. Gosálvez and S. Johnston (2021). "The effect of *Chlamydia* infection on koala (*Phascolarctos cinereus*) semen quality." Theriogenology **167**: 99-110.

Jaccoud, D., K. Peng, D. Feinstein and A. Kilian (2001). "Diversity arrays: a solid state technology for sequence information independent genotyping." Nucleic acids research **29**(4): e25-e25.

Jackson, M., N. White, P. Giffard and P. Timms (1999). "Epizootiology of Chlamydia infections in two free-range koala populations." Veterinary Microbiology **65**(4): 255-264.

Kilian, A., P. Wenzl, E. Huttner, J. Carling, L. Xia, H. Blois, V. Caig, K. Heller-Uszynska, D. Jaccoud and C. Hopper (2012). "Diversity arrays technology: a generic genome profiling technology on open platforms." Data production and analysis in population genomics: Methods and protocols: 67-89.

Kjeldsen, S. R., K. R. Zenger, K. Leigh, W. Ellis, J. Tobey, D. Phalen, A. Melzer, S. FitzGibbon and H. W. Raadsma (2016). "Genome-wide SNP loci reveal novel insights into koala (Phascolarctos cinereus) population variability across its range." Conservation Genetics **17**: 337-353.

Landguth, E. L., S. A. Cushman, M. K. Schwartz, K. S. McKelvey, M. Murphy and G. Luikart (2010). "Quantifying the lag time to detect barriers in landscape genetics." Molecular Ecology **19**(19): 4179-4191.

Lopez-Sepulcre, A., K. Norris and H. Kokko (2009). "Reproductive conflict delays the recovery of an endangered social species." J Anim Ecol **78**(1): 219-225.

Milligan, B. G. (2003). "Maximum-likelihood estimation of relatedness." Genetics **163**(3): 1153-1167.

Moss, D. R., S. M. Arce, C. A. Otoshi, R. W. Doyle and S. M. Moss (2007). "Effects of inbreeding on survival and growth of Pacific white shrimp *Penaeus* (*Litopenaeus*) *vannamei*." Aquaculture **272**: S30-S37.

Nyari, S., C. A. Waugh, J. Dong, B. L. Quigley, J. Hanger, J. Loader, A. Polkinghorne and P. Timms (2017). "Epidemiology of chlamydial infection and disease in a free-ranging koala (*Phascolarctos cinereus*) population." PLOS ONE **12**(12): e0190114.

Orsted, M., A. A. Hoffmann, E. Sverrisdóttir, K. L. Nielsen and T. N. Kristensen (2019). "Genomic variation predicts adaptive evolutionary responses better than population bottleneck history." Plos Genetics **15**(6).

Pagliarani, S., S. D. Johnston, K. W. Beagley, L. Hulse and C. Palmieri (2022). "Chlamydiosis and cystic dilatation of the ovarian bursa in the female koala (*Phascolarctos cinereus*): Novel insights into the pathogenesis and mechanisms of formation." Theriogenology **189**: 280-289.

Peakall, R. and P. E. Smouse (2012). "GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update." Bioinformatics **28**(19): 2537-2539.

Pew, J., P. H. Muir, J. Wang and T. R. Frasier (2015). "related: an R package for analysing pairwise relatedness from codominant molecular markers." Molecular Ecology Resources **15**: 557-561.

Polkinghorne, A., J. Hanger and P. Timms (2013). "Recent advances in understanding the biology, epidemiology and control of chlamydial infections in koalas." Veterinary microbiology **165**(3-4): 214-223.

Pritchard, J. K., M. Stephens and P. Donnelly (2000). "Inference of Population Structure Using Multilocus Genotype Data." Genetics **155**(2): 945.

R Core Team (2018). R: a language and environment for statistical computing. Vienna, R Foundation for Statistical Computing.

Robbins, A., J. Hanger, M. Jelocnik, B. L. Quigley and P. Timms (2019). "Longitudinal study of wild koalas (*Phascolarctos cinereus*) reveals chlamydial disease progression in two thirds of infected animals." Scientific Reports **9**(1): 13194.

Schultz, A. J., R. H. Cristescu, B. L. Littleford-Colquhoun, D. Jaccoud and C. H. Frère (2018). "Fresh is best: Accurate SNP genotyping from koala scats." Ecology and evolution **8**(6): 3139-3151.

Wan, C., J. Loader, J. Hanger, K. W. Beagley, P. Timms and A. Polkinghorne (2011). "Using quantitative polymerase chain reaction to correlate Chlamydia pecorum infectious load with ocular, urinary and reproductive tract disease in the koala (Phascolarctos cinereus)." Australian veterinary journal **89**(10): 409-412.

Wang, J. L. (2019). "A parsimony estimator of the number of populations from a STRUCTURE-like analysis." Molecular Ecology Resources **19**(4): 970-981.

Wang, J. L. (2022). "Fast and accurate population admixture inference from genotype data from a few microsatellites to millions of SNPs." Heredity **129**(2): 79-92.

Waugh, C., J. Hanger, P. Timms and A. Polkinghorne (2016). "Koala translocations and Chlamydia: Managing risk in the effort to conserve native species." Biological Conservation **197**: 247-253.

Weigler, B. J., A. A. Girjes, N. A. White, N. D. Kunst, F. N. Carrick and M. F. Lavin (1988). "Aspects of the epidemiology of *Chlamydia psittaci* infection in a population of koalas

(*Phascolarctos cinereus*) in southeastern Queensland, Australia." Journal of Wildlife Diseases **24**(2): 282-291.

6. Appendix A: Genetic assessment of koalas from Mid Coast Council, New South Wales

Background

Koala scat samples collected from several locations (outside of Bunyah) throughout Mid Coast Council, NSW were received by DDC together with the samples from Bunyah, but were analysed and reported separately here due to their geographic distance to the Bunyah samples (Figure A1), the large geographical spread of sample locations and the small sample size. Population genetic indices, such as heterozygosity and inbreeding were not estimated due to the geographical spread and small sample size, meaning that there were likely too fewer samples representing multiple clusters.

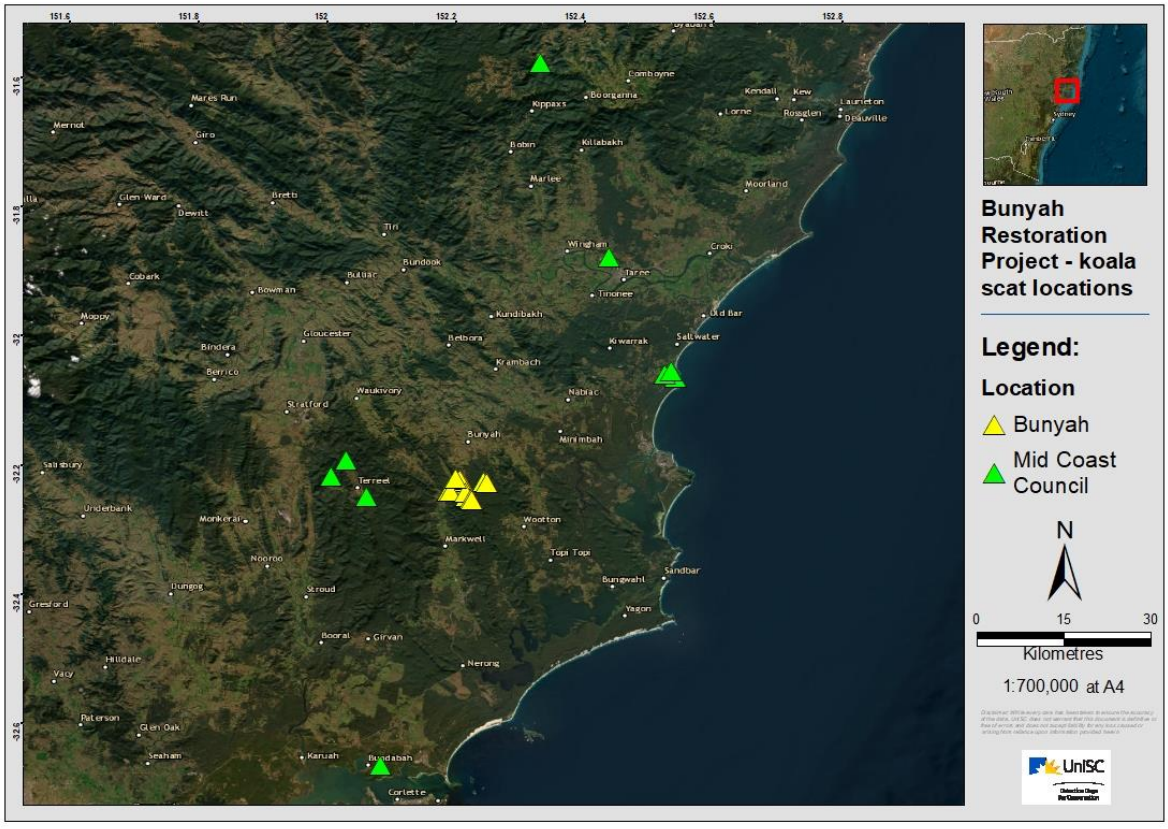


Figure A1. Distribution and location of scats collected from several locations of Mid Coast Council compared to the Bunyah sample locations.

Methods

A total of 10 samples collected between July and September 2023 (Table A1) were delivered to UniSC by a commercial courier service in mid-October 2023. The processing and handling of samples were similar to the methods described in the main report (section 2). All 10 samples were processed for DNA extraction. DNA extractions were replicated with a second scat from the same sampling tube. A total of 20 DNA extractions were sent for genotyping using DArTag as described in the main report section 2. Data were analysed following the same procedure described in the main report section 2.

Table A1. List and overview of koala scat samples collected from Mid Coast Council (N = 10).

Sample name	Date of collection	Method of survey	Location	Latitude	Longitude
KH01	19/07/2023	Detection Dog	Not provided	-31.576093	152.329962
KH02	20/07/2023	Detection Dog	Not provided	-32.032297	152.499580
Kundle	9/08/2023	Detection Dog	Kundle Kundle	-31.878214	152.436915
TG01	30/08/2023	Detection Dog	Tea Gardens	-32.665573	152.082676
WP2	2/08/2023	Detection Dog	Terreel	-32.192161	152.028737
WP21	3/08/2023	Detection Dog	Terreel	-32.248838	152.060861
WP312	1/08/2023	Detection Dog	Terreel	-32.217572	152.005451
HP01	6/09/2023	Detection Dog	Hallidays Point	-32.065386	152.539606
HP02	6/09/2023	Detection Dog	Not provided	-32.058967	152.523737
HP03	3/09/2023	Detection Dog	Hallidays Point	-32.055167	152.532877

Results

One sample, namely KH02 (including its duplicate DNA extraction) failed genotyping (see Table A2 for scat images), and was excluded from further analyses. A total of nine samples were successfully genotyped and passed the quality control for analyses. Data filtration for identifying unique individuals (genetic fingerprinting) retained a total of 1,091 SNPs with an average of 10.5% missing data. No duplicate samples were found within the sample set and all nine were identified as unique koalas.

Table A2. Image of KH02 scat samples that was failed in genotyping.

Sample name	Scat image
KH02	

Based on the sex-linked markers, of the nine unique individuals, seven (78%) were males and two (22%) were females (Table A3 and see Figure A2 for location), however no sex ratio can be calculated given the geographic spread and low sample size. Based on the *Chlamydia*-specific markers, of the nine individuals, five (56%) were *Chlamydia* positive (Table A3 and see Figure A3 for location). A similar level of infection was observed in both male and female koalas (57% vs 50%, respectively).

Table A3. Sex and *Chlamydia* status of unique koalas from Mid Coast Council samples.

Sample name	Sex	*Chlamydia* status
KH01	M	Positive
Kundle	M	Negative
TG01	M	Positive
HP01	M	Positive
HP02	F	Positive
HP03	M	Negative
WP2	F	Negative
WP21	M	Negative
WP312	M	Positive

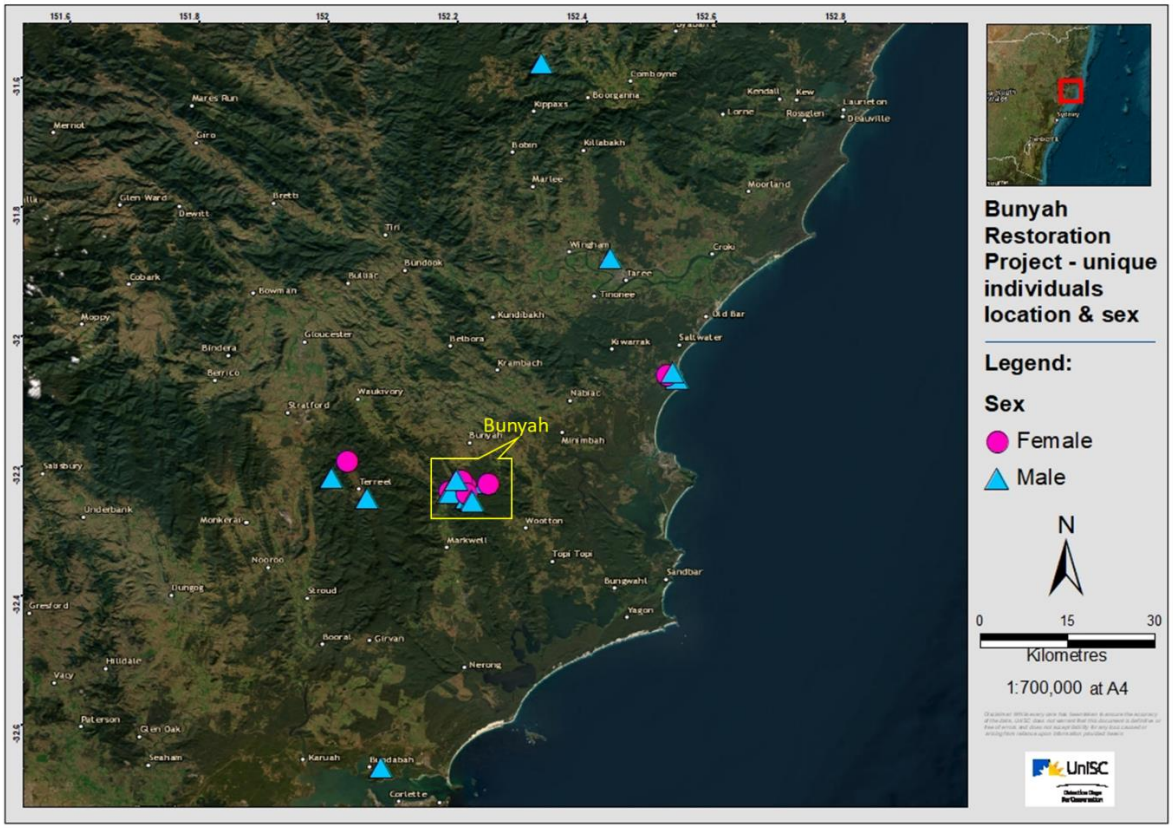


Figure A2. Distribution and sex of unique koalas from Mid Coast Council compared to the Bunyah sample locations.

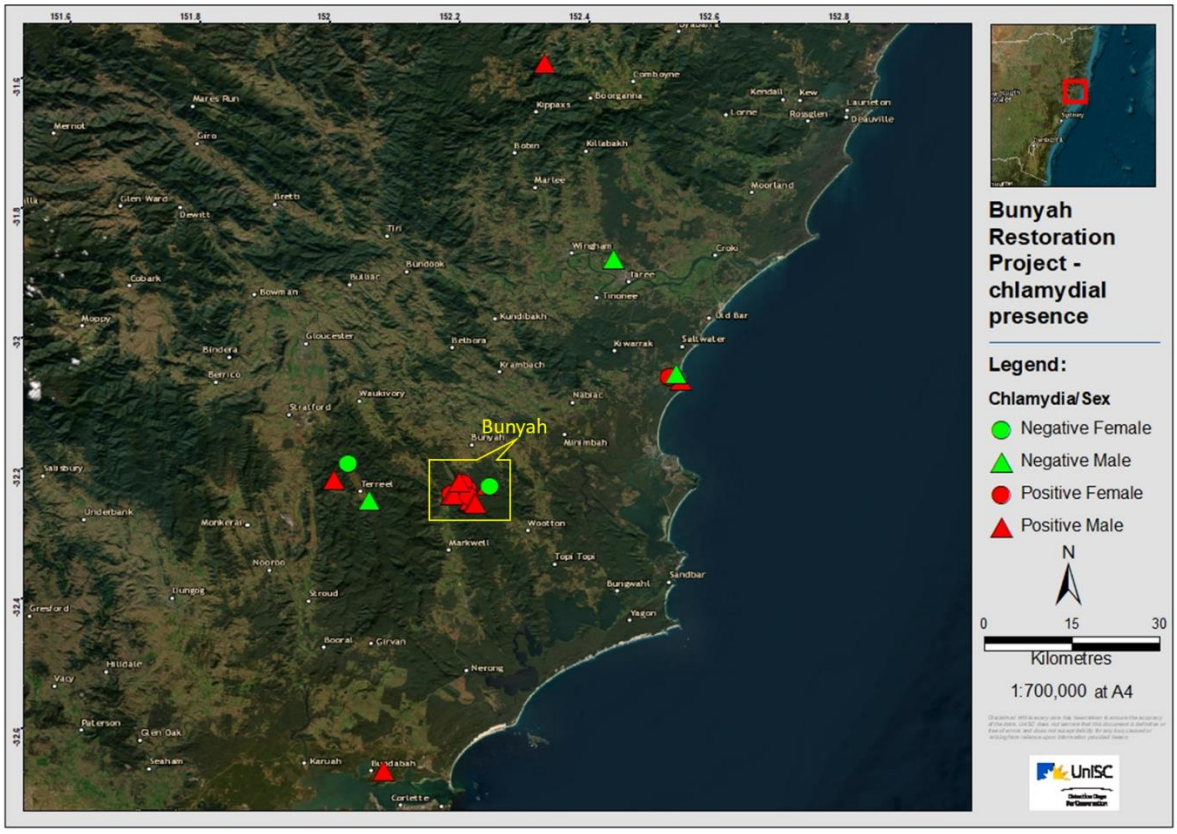


Figure A3. Distribution and *Chlamydia* infection status among unique koalas from Mid Coast Council compared to the Bunyah sample locations.

Discussion

One sample from the provided 10 was failed in genotyping for this set of samples. The sample set from Mid Coast Council also showed a larger proportion of males (seven out of nine), similar to results observed for Bunyah samples. However, considering the geographical spread, these koalas likely do not represent only one population and concluding on a sex ratio would be incorrect; there is also too small a sample size with nine individuals.

The proportion of samples with *Chlamydia* infection (56%) was not as high as in Bunyah samples (94% in Bunyah). To place these koalas into context with the Bunyah koalas, we would recommend additional study aiming for a larger sample size across the geographical spread of the Mid Coast survey. That would enable for interesting measures such as connectivity between the groups to be investigated.

7. Appendix B: Overview of the scat samples and scat collection locations

Table B1. List and overview of koala scat samples collected from Bunyah Koala Project area (N = 25).

Sample name	Date of collection	Method of survey	Location	Latitude	Longitude
K10	22/08/2023	Thermal Drone	Prices Creek	-32.217445	152.198995
K11	23/08/2023	Thermal Drone	Wang Wauk State Forest	-32.233579	152.210115
K12	23/08/2023	Thermal Drone	Wang Wauk State Forest	-32.234598	152.208335
K13	23/08/2023	Thermal Drone	Wang Wauk State Forest	-32.233141	152.211570
K15	23/08/2023	Thermal Drone	Wang Wauk State Forest	-32.230419	152.212625
K16	23/08/2023	Thermal Drone	Wang Wauk State Forest	-32.236798	152.217456
K17	23/08/2023	Thermal Drone	Wang Wauk State Forest	-32.240015	152.207951
K19	23/08/2023	Fresh Scat	Wang Wauk State Forest	-32.240750	152.207240
K20	29/05/2023	Scat	Manning Hill	-32.238080	152.199040
K21	29/05/2023	Scat	Manning Hill	-32.238680	152.190280
K22	30/05/2023	Thermal Drone	Wang Wauk State Forest	-32.225619	152.241972
K24	30/05/2023	Thermal Drone	Wang Wauk State Forest	-32.227679	152.247881
K26	30/05/2023	Thermal Drone	Wang Wauk State Forest	-32.226621	152.206454
K27	30/05/2023	Thermal Drone	Wang Wauk State Forest	-32.222320	152.205073
K30	31/05/2023	Thermal Drone	Wang Wauk State Forest	-32.239747	152.186707
K31	31/05/2023	Thermal Drone	Wang Wauk State Forest	-32.240856	152.187168
K32	31/05/2023	Thermal Drone	Wang Wauk State Forest	-32.238472	152.187754
K33	26/11/2022	Scat	Prices Creek	-32.220620	152.199250
K34	31/05/2023	Thermal Drone	Wang Wauk State Forest	-32.241756	152.189974
K35	31/05/2023	Thermal Drone	Wang Wauk State Forest	-32.242860	152.192564
K36	1/06/2023	Thermal Drone	Wang Wauk State Forest	-32.249008	152.215242
K37	1/06/2023	Thermal Drone	Wang Wauk State Forest	-32.242987	152.213452
K38	1/06/2023	Thermal Drone	Wang Wauk State Forest	-32.252482	152.221822
K39	1/06/2023	Thermal Drone	Wang Wauk State Forest	-32.253278	152.214311
K40	1/06/2023	Thermal Drone	Wang Wauk State Forest	-32.254108	152.223898

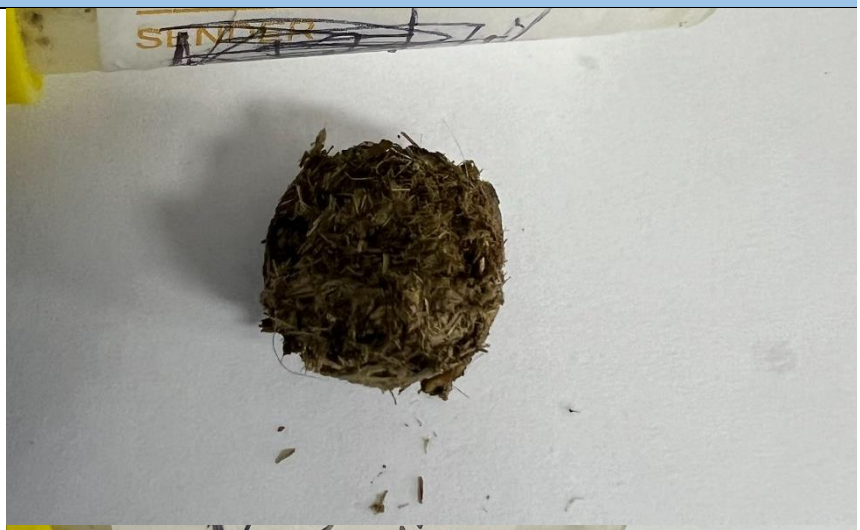
8. Appendix C: Status of samples for genetic analyses

Table C1. List of samples and DNA extractions (N = 46) with its quality control status for genetic analyses. Sample names with "_A" extension indicates the second replicate DNA extraction from a separate scat of its original sample tube and is only available for the samples which had a minimum of two scats.

Sample name	Quality control status for genetic fingerprinting
K10	Passed
K10_A	Passed
K11	Failed
K11_A	Failed
K12	Failed
K13	Passed
K13_A	Passed
K15	Passed
K15_A	Passed
K16	Failed
K17	Passed
K17_A	Passed
K19	Failed
K19_A	Failed
K20	Passed
K20_A	Passed
K21	Passed
K21_A	Passed
K22	Passed
K22_A	Failed
K24	Passed
K24_A	Failed
K26	Passed
K26_A	Passed
K27	Passed
K27_A	Passed
K30	Passed
K30_A	Passed
K31	Passed
K31_A	Passed
K32	Failed
K32_A	Passed
K33	Passed
K34	Passed
K34_A	Passed
K35	Failed

K36	Passed
K36_A	Passed
K37	Passed
K37_A	Passed
K38	Passed
K38_A	Passed
K39	Failed
K39_A	Failed
K40	Passed
K40_A	Passed

Table C2. Images of scat samples that failed genotyping. Note: pictures from K11, K12, K16 and K35 could not be captured as samples were exhausted after DNA extraction.

Sample name	Scat image
K19	
K39	

End of report