Detection Dogs for Conservation

Coffs Harbour Koala Survey – Genotyping from Scats

A report by Detection Dogs for Conservation, University of the Sunshine Coast

for

Jaliigirr Biodiversity Alliance and Canines for Wildlife

Authors:

Dr. Katrin Hohwieler, Dr. Alexis Levengood, and Dr. Romane Cristescu

USC Detection Dogs for Conservation

Disclaimer

This report was prepared in accordance with the scope of work agreed with Canines for Wildlife (the Client) and is subject to the specific time, cost and other constraints as defined by the scope of work.

To prepare this report, USC relied on information supplied by the Client and does not accept responsibility for the accuracy or completeness of this information. USC also relied on information gathered at particular times and under particular conditions and does not accept responsibility for any changes or variances to this information which may have subsequently occurred. Accordingly, the authors of the report provide no guarantee, warrant or representation in respect to the accuracy, adequacy, or completeness of the information, whether generally or for use or reliance in specific circumstances. To the extent permitted by law, the authors exclude liability including any liability for negligence, for any loss, damage, injury, illness howsoever caused, including (with limitation) by the use of, or reliance upon, the information, and whether arising from errors or omissions or otherwise.

Contents

Coffs Harbour Koala Survey – Genotyping from Scats 1
Disclaimer 2
Executive Summary 4
Scope of works 5
Limitations 7
Methods & Results 7
DNA extraction 7
SNP genotyping 7
Filtering of genetic data 10
Number of unique individuals, sex and chlamydia 10
Population structure and genetic measures 14
References 22
Appendix 24

Executive Summary

Jaliigiir Biodiversity Alliance and Canines for Wildlife contracted the University of the Sunshine Coast's Detection Dogs for Conservation team (DDC) to genetically analyse a set of 46 koala scat samples that were collected in the Coffs Harbour region.

The koala scats were processed in the laboratory; DNA was extracted and sent to the genotyping provider Diversity Arrays Technology, Canberra. Out of 46 samples genotyped, 29 samples were of good quality and used to identify unique individuals. A total of six samples were removed due to being a duplicate sample, i.e. when there are two samples stemming from the same koala. Therefore, we were able to confidently identify 23 unique koalas. For these 23 koalas, maps showing sex and distribution of possible chlamydia pathogen detection are provided.

The 20 best samples were used to assess population structure. There is likely sub-structuring as well as cryptic clines in this group of koalas, based on three different methods to estimate population structure. However, with 20 individuals, sample size was very small und thus results need to be interpreted with caution.

Measures of genetic diversity were estimated and compared to values from other populations. Values of heterozygosity were similar to those found in other populations, inbreeding was found to be similar but generally lower than in other populations.

Overall, results presented in this report indicate potentially interesting population patterns that would be worthwhile to further investigate with an increased sample size. Genetic diversity measures were comparable to estimations from other populations.

Scope of works

Jaliigiir Biodiversity Alliance and Canines for Wildlife contracted the University of the Sunshine Coast's Detection Dogs for Conservation team (DDC) to analyse a set of 46 koala scat samples that were collected in the Coffs Harbour region (Fig 1) by Canines for Wildlife Scent Detection Dog team (CFW). These scat samples were sent to the DDC, where DNA was extracted from the scats. The DNA extract was sent for genotyping at Diversity Arrays Technology (DArT) in Canberra. The resulting genotyping data was analysed by the DDC and methods and results are reported in this document.

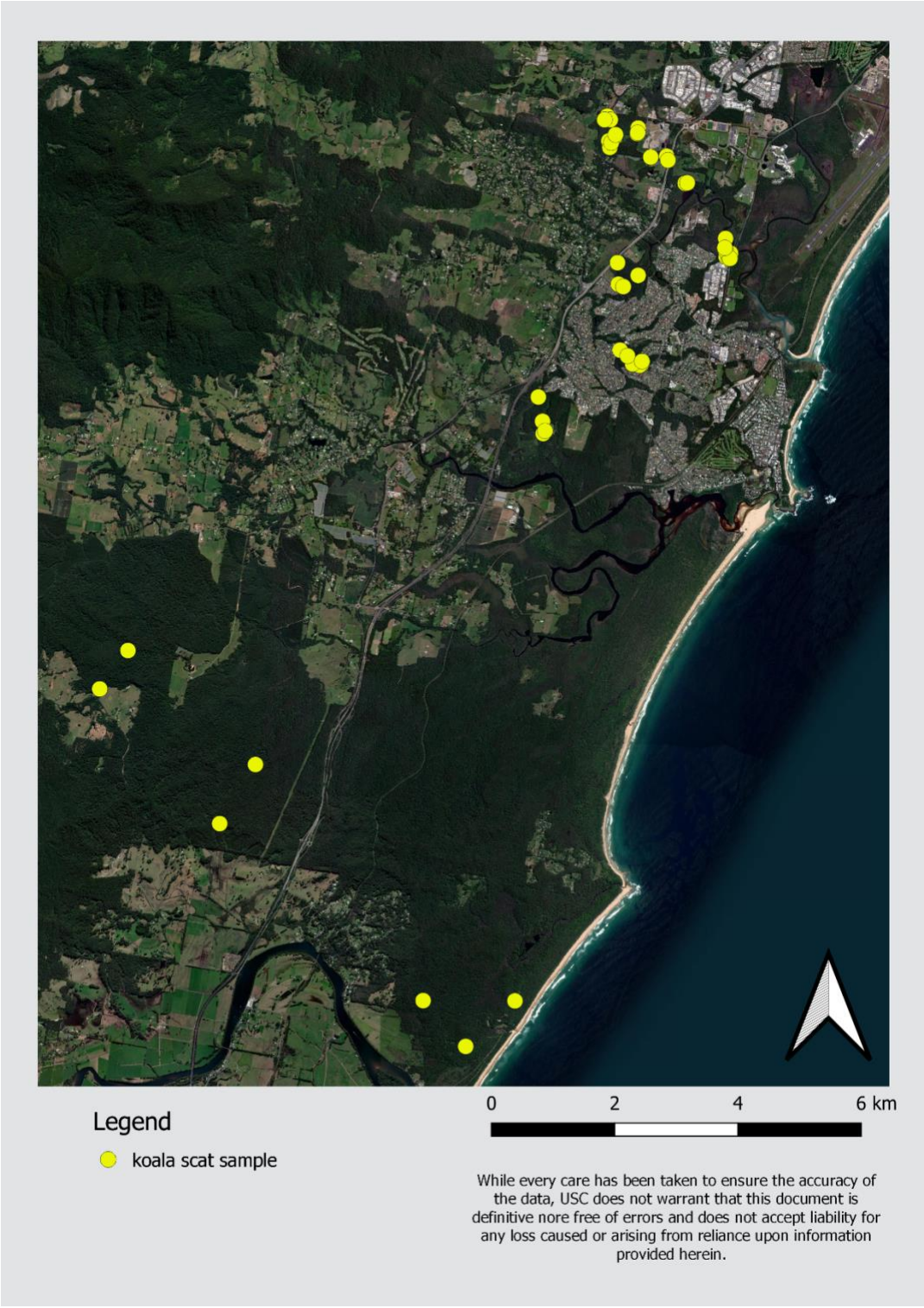

Figure 1: Overview of the locations of 46 koala scat samples collected by Canines for Wildlife and provided to the Detection Dogs for Conservation team for genetic analysis.

Limitations

- Data quality of DNA derived from scats is very variable. The DDC has maximised the ability to retain samples, however, it is common that some samples result in too little genotyped data to be used for subsequent analyses. These need to be excluded to ensure the best possible estimations of genetic measures.
- The DDC had no influence over the scat collection method. The DDC only had influence over storage of samples once received and the DNA extraction process. Best care has been taken to ensure no contamination or damage.
- Due to the nature of non-invasive sampling, it is common that duplicate samples are collected, i.e. two or more samples that originate from the same koala. These need to be identified and removed from analyses to not inflate the results artificially and produce false outcomes. For instance, duplicate samples are genetically the same or very similar, therefore, retaining duplicates can wrongly lead to increased measures of inbreeding.

Methods & Results

DNA extraction

A total of 46 samples were processed for DNA extraction. The outer layer of a scat was scraped off using a sharp razor blade, to capture the epithelial cells. This raw sample containing koala cells was then further processed using the QIAmp PowerFecal Pro DNA Kit (Qiagen), following the manual instructions. An additional one-hour incubation step after adding the buffer to the fecal sample was added. Final DNA isolates were eluted in 100ul of elution buffer and concentrated down to a volume of 50-60ul. The amount of DNA present in extracted samples (both koala and foreign) was extracted using the Thermo Scientific NanoDrop 2000 Spectrometer (Thermo Fisher Scientific, Victoria). Extracted DNA was stored at -80°C until being shipped to DArT.

SNP genotyping

Genetic samples were genotyped using a next-generation sequencing protocol for detecting Single Nucleotide Polymorphisms (SNPs, Kilian et al. (2012)). Specifically, a targeted

approach was chosen (DArTtag), where a molecular probe selects small target regions containing sequence variants. A total of 4393 SNPs were genotyped. All 46 samples were genotyped using DArTtag, however, genotypes differed in quality. Therefore, some samples had to be excluded for any analysis due to insufficient amount of data. Figure 2 shows the location of all samples, but also which samples were used in subsequent analyses (successful genotyping), which samples had to be removed due to being duplicates of the same individual koala (successful genotyping but removed due to duplication) and which samples did not reach a sufficient amount of data due to poor quality (insufficient genotype data). For further details see *Appendix Table 1*.

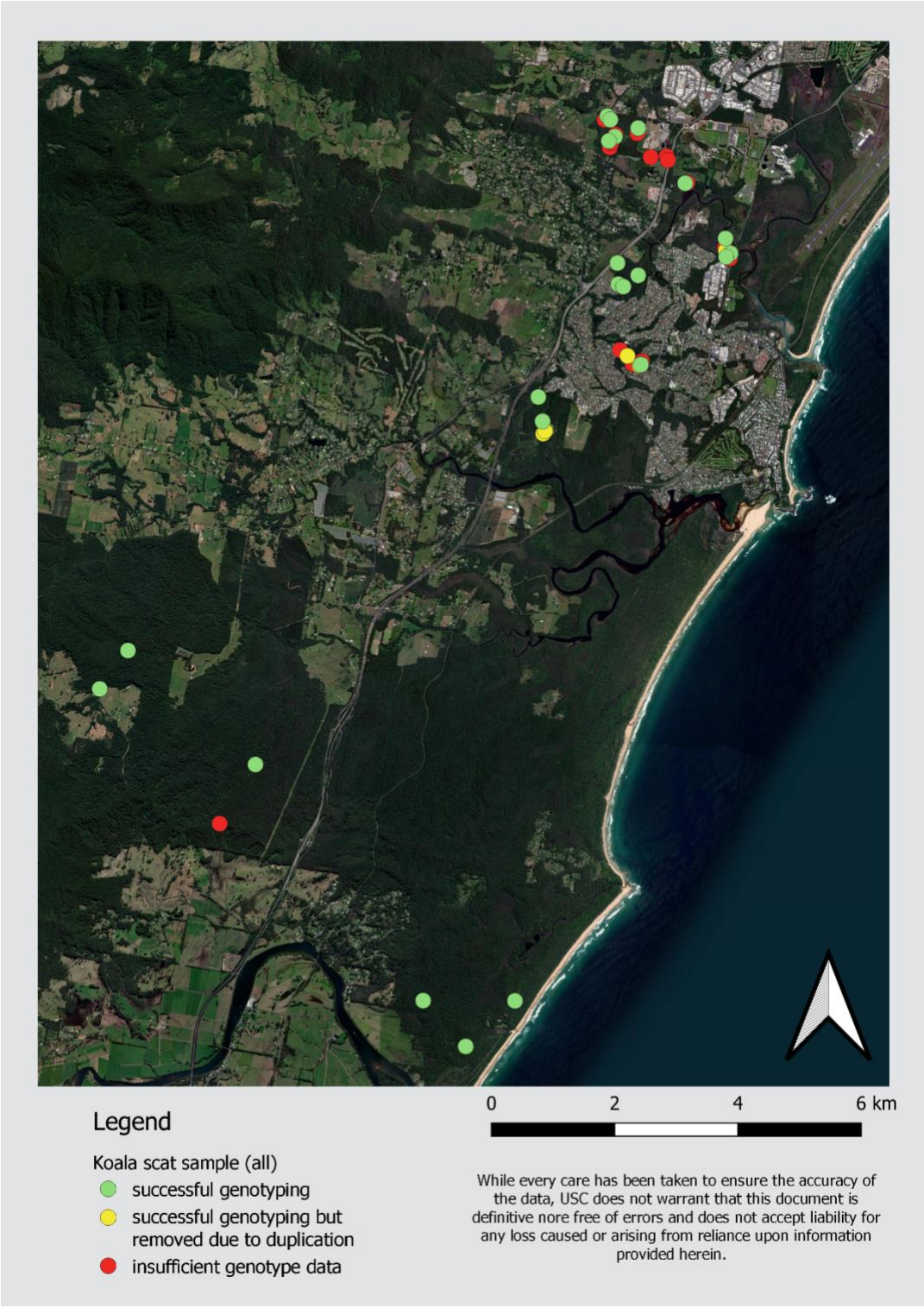

Figure 2: Map showing the location of all 46 koala scat samples and the quality of the genotyping of each sample. Of the 46 samples, 29 were of good quality, of which six were removed due to being a duplicate (shown in green and yellow). Red points show which samples did not present sufficient genotyping data to be considered for subsequent analyses.

Filtering of genetic data

Genotyped data is filtered to improve the quality of the dataset by removing samples with too little data (low *individual call rate*) as well as SNP loci that were not called across most samples (low *locus call rate*). The thresholds for these filters can be adjusted depending on the type of analysis, e.g. only few loci are needed for identifying unique individuals, however, many high-quality loci are needed for measuring genetic diversity. Therefore, different filtering regimes have been applied for different analyses. Other constant thresholds are applied to remove potentially erroneous loci. This included filtering for allele read depth (minimum threshold of 5) and 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 is applied as well. All filtering was done using the R package *dartR* (Gruber et al. 2018).

Number of unique individuals, sex and chlamydia

For identifying unique individuals, only approximately 200 SNPs are required (Schultz et al. 2018). Therefore, the focus was on maximising the number of individuals that can be used while retaining sufficient high-quality SNPs. Additional to the constant filters, we applied a stepwise increasing locus call rate filtering, from 0.2 to 0.9 – resulting in only retaining SNPs with at least 90% data. Samples were also filtered for individual call rate at 0.1. This resulted in the removal of 17 samples due to insufficient amount of data, and 29 genotypes left for this analysis. A total of 331 loci were retained with 5.66% missing data.

Samples stemming from the same individual were identified by calculating relatedness estimates using the related package (Pew et al. 2015). The idea here is that, for instance, parents and their offspring, as well as full siblings (sharing both the same mother and father), share about 50% of the same DNA (*0.5 related*). Half siblings share 25% of the same DNA (*0.25 related*), and comparing DNA of the same individual should result in a very high relatedness value close to 0.9 or higher. Due to generally low data quality, we lowered the threshold and assumed that samples of the same individual would result in a relatedness value of 0.75 or higher (to avoid including duplicate samples whilst risking exclusion of highly related siblings). Multiple different relatedness estimators were investigated to decide on duplicated

samples (Queller and Goodnight 1989, Lynch and Ritland 1999, Wang 2002, Milligan 2003). The following samples are suspected to be duplicates:

Table 1: Pairwise table showing relatedness values between samples that are very likely from the same individual koala. Note that the three samples 7.1, 7.2 and 7.3 all belong to the same individual, whilst the other samples are duplicates. The asterisk indicates which sample was removed based on lower amount of data.

Sample 1	Sample 2	Wang (2002)	Lynch and Ritland (1999)	Queller and Goodnight (1989)	Milligan (2003)
Coffs_1.7	Coffs_1.9*	0.9473	0.9144	0.9462	0.9583
Coffs_2.1	Coffs_2.1.4*	0.9075	0.8907	0.9088	0.9885
Coffs_1.4*	Coffs_1.3	0.9091	0.8823	0.9047	0.9504
Coffs_13.4*	Coffs_13.3	0.8408	0.7925	0.8521	0.9293
Coffs_7.1	Coffs_7.2*	0.9003	0.9141	0.899	1
Coffs_7.3*	Coffs_7.1	0.8946	0.8521	0.8857	0.9469
Coffs_7.3*	Coffs_7.2*	0.7949	0.7735	0.7995	0.8722

Because only one sample for each individual koala can be retained, the sample with the lower amount of data was removed for further analyses (note that the differences were very small). We removed samples Coffs_1.9, Coffs_2.1.4, Coffs_1.4, Coffs_7.2, Coffs_7.3, and Coffs_13.4 (Table 1). For an overview of duplicate locations see Appendix Figure 1.

Therefore, from the 46 samples that were sent from Canines for Wildlife, we were able to confidently identify 23 unique individuals. For these 23 individuals, we are presenting the sex (Fig 3) and whether the chlamydia pathogen was detected (Fig 4). For complete overview see *Appendix Table 1*.

Of the 23 individuals, 12 were male and 11 were female. Therefore, the sex ratio is very close to 1:1, which is considered good. However, the sample size was small.

Out of the 23 koalas, eight were found to possibly be carriers of the chlamydial pathogen, and three were found to very likely be carrying the pathogen. The spread of the chlamydial pathogen was even across the landscape (Fig 4). These numbers have to be interpreted with caution as the presence of the chlamydia pathogen does not equate to clinical signs of disease.

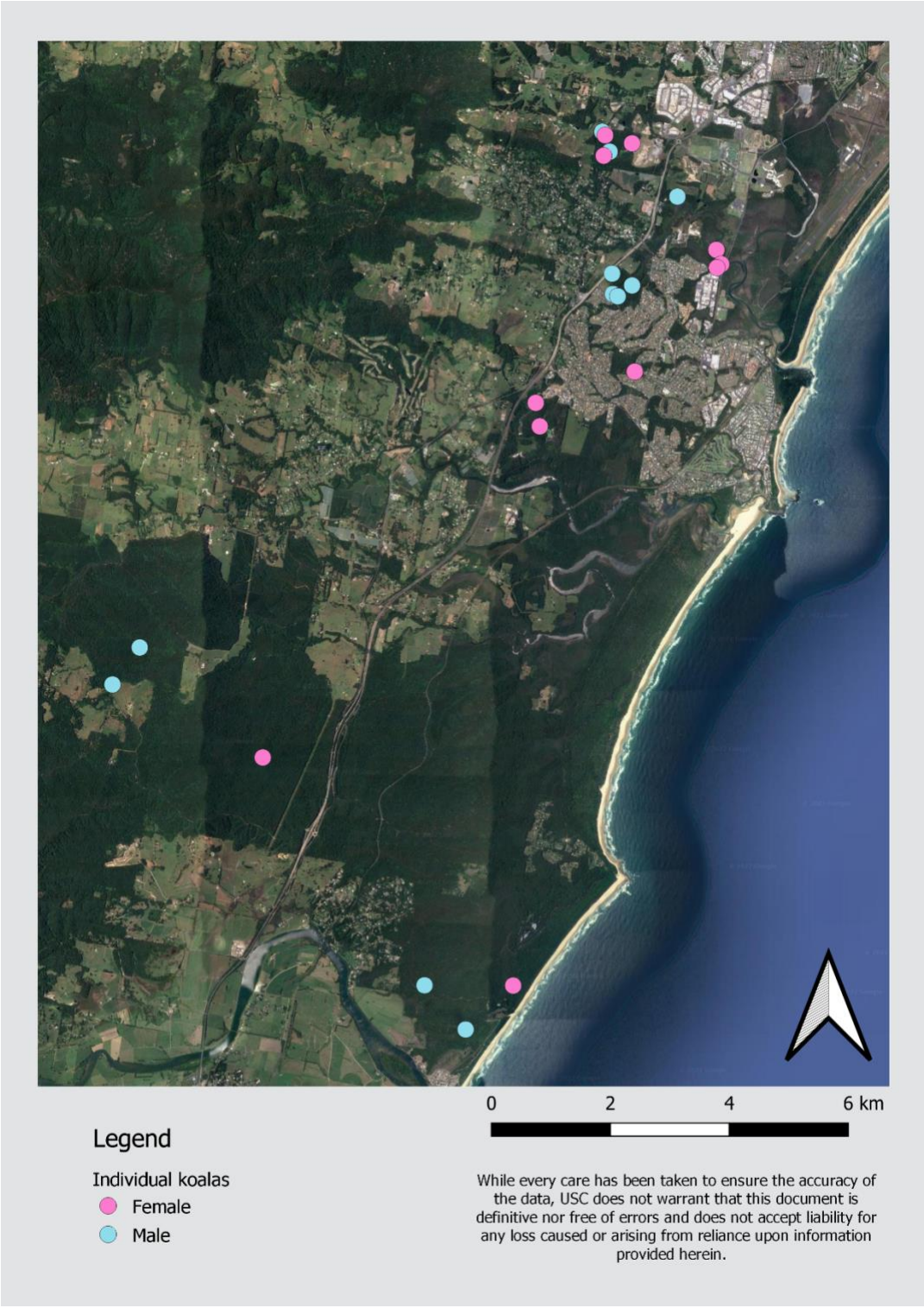

Figure 3: Map is showing 23 koalas identified as unique individuals, presented in pink for female koalas (N = 11) and blue for male koalas (N = 12).

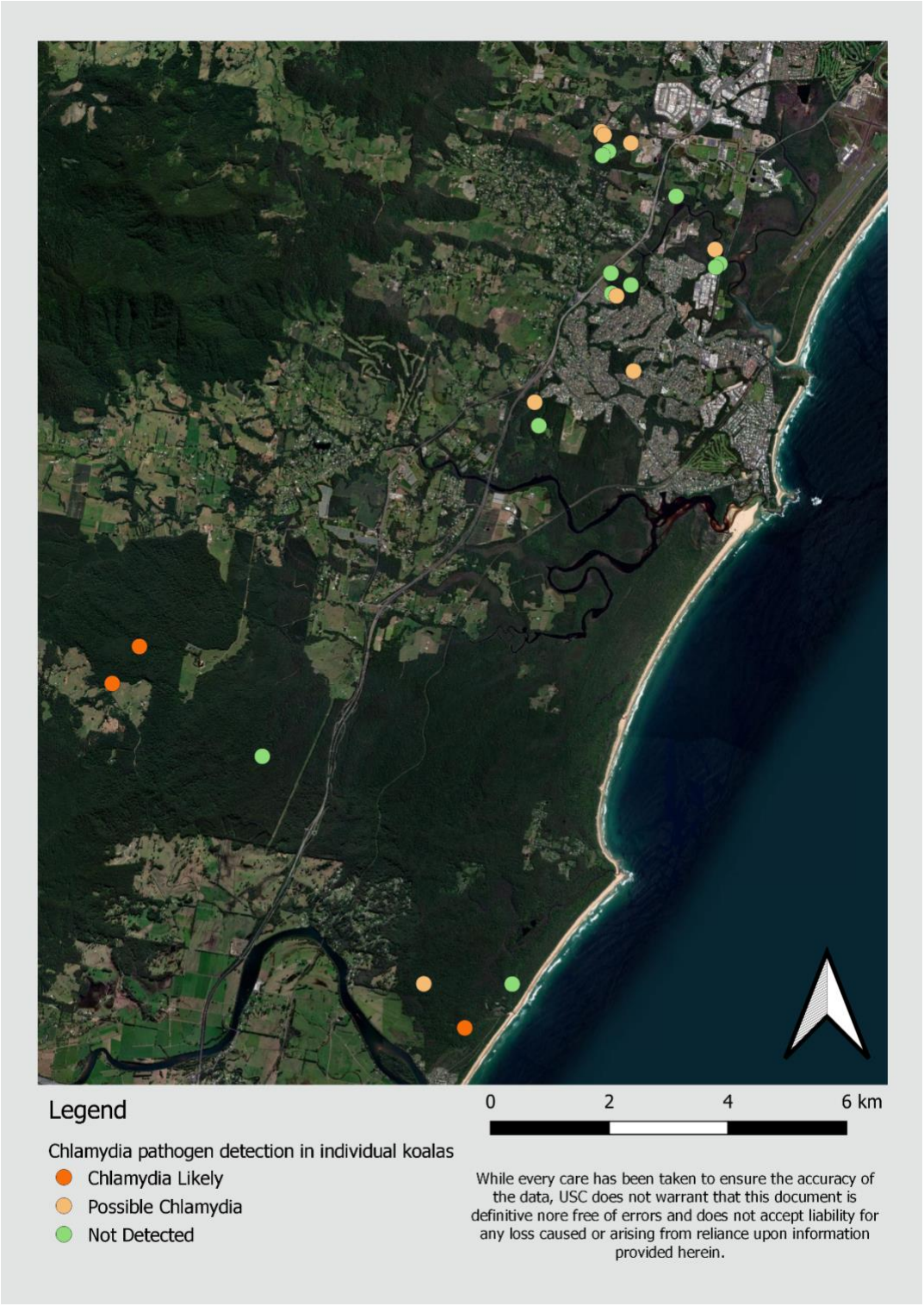

Figure 4: Map of uniquely identified koalas (N=23) and whether any chlamydia pathogen was detected in the genetic sample. Detection of chlamydia pathogen and diagnosis is categorised into three groups: red dots present koalas with high chlamydial presence (N = 3); orange dots present koalas where low thresholds of the pathogen were detected, therefore a chlamydial infection is possible (N = 8); green dots present koalas where a chlamydial infection is unlikely (N = 12).

Population structure and genetic measures

Additional to the constant filters, we applied a stepwise increasing locus as well as individual call rate filtering: from 0.1 to 0.8 for individual call rate (retaining only samples with at least 80% data) and from 0.2 to 0.9 for locus call rate (retaining loci with at least 90% data). This resulted in 26 genotypes with 1117 loci. After removal of the duplicate samples, 20 individual koalas were retained for subsequent analysis with 4.08% missing data. High filtering is required to ensure reliable results for genetic measures. Low quality data can otherwise make a population look less diverse/more inbred than it actually is.

Faststructure

Commonly, a population structure analysis using *STRUCTURE* (Pritchard et al. 2000) or *faststructure* (Raj et al. 2014) is conducted to identify patterns in the population. However, usually a larger number of individuals are required, and results likely only reflect broad scale structure. We used *faststructure* to investigate any structure amongst the 20 individuals but sample size might be insufficient for a reliable result. We investigated a range of possible numbers of ancestral populations, from two to four (K = 2-4) with a simple prior setting. The likely number of ancestral populations found through this analysis was three (Fig 5). Whilst this seems like a high number for a relatively small area, the *faststructure* result reflects similar patterns to the PCA result (Fig 6) in terms of clustering of individuals. Nevertheless, the clusters are close to each other and real differentiation is unlikely. The differences between these groups are likely not strong but could reflect genetic clusters of individuals that are genetically more similar to each other. A larger sample size would be required to do further in depth analyses. It should be noted, that, for instance, inclusion of translocated individuals in a genetic dataset (whether on purpose or unknown) can alter “natural” population genetic patterns and a population might look more differentiated than it would appear without translocated individuals.

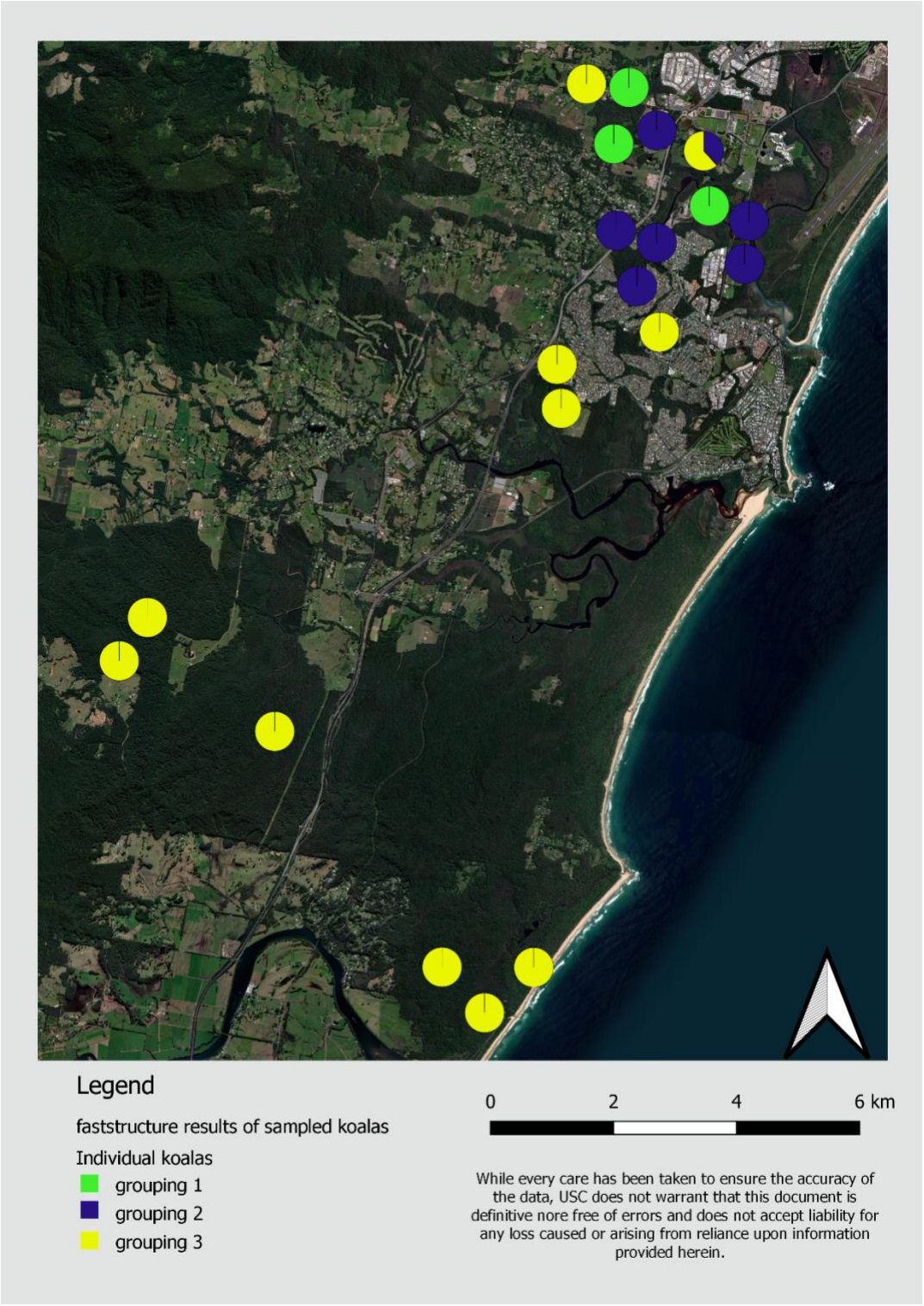

Figure 5: Map presenting results of a faststructure *analysis (Raj et al. 2014) which investigates population structure. The most likely number of ancestral populations (K) was K=3. Likely, these clusters are groupings of individuals that are more similar to each other than to the remaining individuals, however, a sample size of 20 individuals is likely too small to draw meaningful conclusions.*

Principal Component Analysis

A simple Principal Component Analysis (PCA) was also conducted, which is widely used to quantify patterns of population structure (Ma and Amos 2012). A PCA can indicate to what amount genetic variation contributes to any detected population patterns. Here, the PCA shows that there likely is a small amount of structure, as the summarized percentage of the two axes exceeds 10% (Fig 6). However, no distinct structure pattern can be identified from the graph. Nonetheless, the distribution of the individuals across the two axes contains similarities to the *faststructure* results.

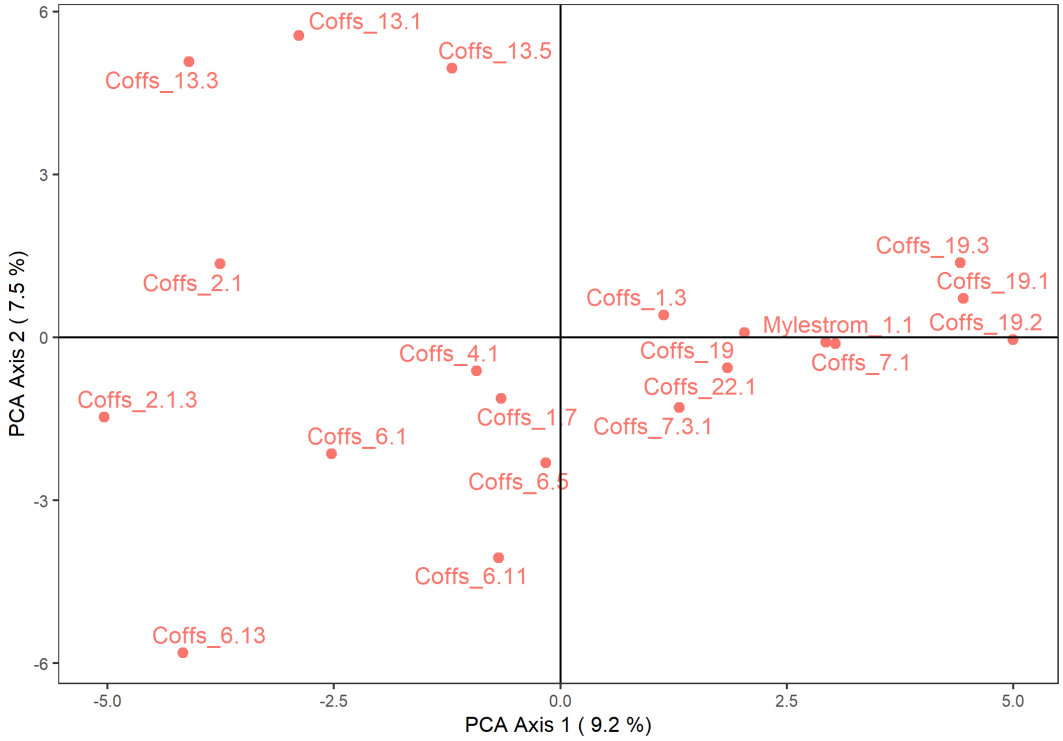

Figure 6: Results of a Principal Component Analysis of 20 individual koalas.

Spatial Principal Component Analysis

We ran a spatial Principal Component Analysis (sPCA, Jombart et al. (2008)) using the R package *adegenet* (Jombart 2008). An sPCA has been shown to detect more fine scale, cryptic patterns of population structure (Jombart et al. 2008, Leys et al. 2016), and found a significant cline between individuals north and south of the Bonville Creek. This result is visualised in Figure 7 where black and white squares represent individual koalas, and the size of the squares correlates with the magnitude of genotype differences. Genotypes in the centre of the

distribution have less extreme values (smaller squares) indicating a cline from north to south rather than strongly differentiated groups (Jombart et al. 2008). For detailed data see *Appendix Table 2*.

This pattern is unlikely to be caused by isolation-by-distance (IBD) as indicated by the results of a Mantel test which was conducted through GenAlEx v. 6.5 (Peakall and Smouse 2012). This test shows that there is no significant genetic differentiation caused by geographic distance ($R^2 = 0.0052$, Fig 8).

Figure 7: Map showing the results of a spatial Principal Component Analysis. A significant global structure was detected between koalas north and south of the Bonville Creek. This likely presents a cline between these individuals and therefore rather cryptic population genetic patterns. Each square represents the score of a genotype and is positioned by its spatial coordinates. The colour together with the size of the square represents the magnitude of the genotype differences.

This is unlikely to be caused by isolation by distance (IBD) as indicated by the mantel test results, which show that there is no significant genetic differentiation caused by geographic distance.

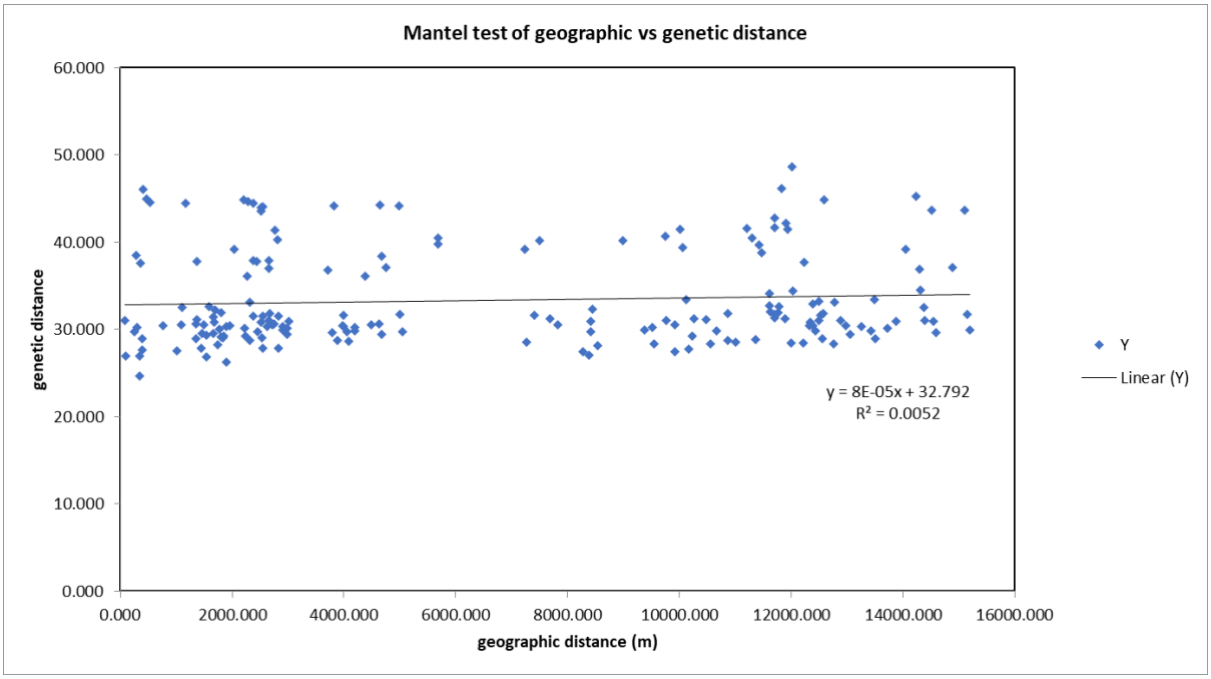

Figure 8: Mantel test conducted on data of 20 individual koalas. Here, geographic distance and genetic distance are compared in a pairwise manner to determine whether genotype differences are cause by geographic distance (isolation-by-distance). In this case, there was no evidence of isolation-by-distance.

Measures of genetic diversity

Genetic diversity was calculated using GenAlEx v. 6.5 (Peakall and Smouse 2012). We calculated three values: observed Heterozygosity H_O, which is the calculated 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); lastly F_{IS}, also called inbreeding coefficient, which is the proportion of the variance in the subpopulation contained in an individual and 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 large populations.

Calculated with 20 individual koalas from the study area, we estimated:

- $H_O = 0.270$,
- $H_E = 0.329$, and
- $F_{IS} = 0.1805$

The numbers are best interpreted by comparing to other populations where diversity measures were calculated using similar methods. For instance, using the same methods, the DDC team estimated diversity measures for Redland City Council (mainland) koalas (N = 227) to be $H_O = 0.237$, $H_E = 0.320$ and $F_{IS} = 0.259$. While H_O and H_E in the 20 Coffs Harbour koalas were similar to the measures estimated for Redlands koalas, the difference between H_O and H_E in Redlands koalas was larger, resulting in a higher inbreeding coefficient. Therefore, inbreeding is estimated to be lower in Coffs Harbour koalas than in koalas in Redland City Council.

Further comparisons can be done by consulting Table 2, which was taken from Kjeldsen et al. (2016). Here, measures are presented from other wild koala populations across Queensland, New South Wales and Victoria.

Table 2: 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, IR = internal relatedness and NeLD = effective population size calculated using linkage equilibrium. Table taken from Kjeldsen et al. (2016).

State	Location	n	Ho	He	Fis ($P < 0.01$)	IR (±SD)	Ne_{LD} (95 %CI)
QLD	St Bees Island	19	0.29	0.35	0.23	0.29 (±0.15)	Infinite (∞)
QLD	St Lawrence	19	0.26	0.30	0.20	0.21 (±0.11)	Infinite (∞)
QLD	Koala Coast	24	0.22	0.30	0.32	0.42 (±0.29)	Infinite (921.20–∞)
QLD	Ipswich	23	0.27	0.31	0.19	0.26 (±0.16)	Infinite (∞)
NSW	Port Macquarie	45	0.23	0.28	0.21	0.25 (±0.15)	116.8 (109.8–124.6)
NSW	Campbelltown	9	0.27	0.33	0.27	0.34 (±0.27)	2.7 (2.4–3.2)
VIC	South Gippsland	19	0.24	0.30	0.27	0.31 (±0.34)	Infinite (∞)
VIC	Cape Otway	13	0.24	0.25	0.11	0.20 (±0.11)	46.7 (40.8–54.4)

References

Gruber, B., P. J. Unmack, O. F. Berry, and A. Georges. 2018. DARTR: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Molecular Ecology Resources **18**:691-699.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics **24**:1403-1405.

Jombart, T., S. Devillard, A.-B. Dufour, and D. Pontier. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity **101**:92-103.

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. Pages 67-89 Data production and analysis in population genomics. Springer.

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.

Leys, M., I. Keller, K. Räsänen, J.-L. Gattolliat, and C. T. Robinson. 2016. Distribution and population genetic variation of cryptic species of the Alpine mayfly Baetis alpinus (Ephemeroptera: Baetidae) in the Central Alps. BMC Evolutionary Biology **16**:77.

Lynch, M., and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. Genetics **152**:1753-1766.

Ma, J., and C. I. Amos. 2012. Principal Components Analysis of Population Admixture. Plos One **7**:e40115.

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

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**:2537-2539.

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

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics **155**:945-959.

Queller, D. C., and K. F. Goodnight. 1989. Estimating relatedness using genetic markers. Evolution **43**:258-275.

Raj, A., M. Stephens, and J. K. Pritchard. 2014. fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. Genetics **197**:573-589.

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

Wang, J. 2002. An estimator for pairwise relatedness using molecular markers. Genetics **160**:1203-1215.

Appendix

Table 1: Table with overview of all samples sent to USC, including sample name, Koala ID given after duplicate identification, whether samples were duplicates, data sufficiency for subsequent analyses, whether data was used in subsequent analyses ("Used"), Sex and Chlamydia information when available, geographical coordinates (provided by Canines for Wildlife)

Sample ID	Koala ID	Duplicate	Data sufficiency	Used	Sex	Chlamydia detection	Lat	Long
Coffs_1.7	1	yes	yes	yes	Female	Possible Chlamydia	-30.3371	153.0927
Coffs_1.9	1	yes	yes - but duplicate	no	Female	Not Detected	-30.3543	153.0761
Coffs_2.1	2	yes	yes	yes	Female	Not Detected	-30.3393	153.0935
Coffs_2.1.4	2	yes	yes - but duplicate	no	Female	Not Detected	-30.3389	153.0927
Coffs_1.3	3	yes	yes	yes	Female	Possible Chlamydia	-30.3555	153.0784
Coffs_1.4	3	yes	yes - but duplicate	no	Female	Not Detected	-30.3556	153.0781
Coffs_13.3	4	yes	yes	yes	Male	Not Detected	-30.3407	153.0744
Coffs_13.4	4	yes	yes - but duplicate	no	Male	Not Detected	-30.3407	153.0744
Coffs_7.1	5	yes	yes	yes	Female	Not Detected	-30.3638	153.0617
Coffs_7.2	5	yes	yes - but duplicate	no	Female	Not Detected	-30.3652	153.0622
Coffs_7.3	5	yes	yes - but duplicate	no	Female	Not Detected	-30.3656	153.0618
Coffs_1.2	6	no	yes	yes	Male	Not Detected	-30.3555	153.0784
Coffs_13.1	7	no	yes	yes	Male	Not Detected	-30.3438	153.0745
Coffs_13.2	8	no	yes	yes	Male	Possible Chlamydia	-30.3441	153.0754
Coffs_13.5	9	no	yes	yes	Male	Not Detected	-30.3425	153.0779
Coffs_19	10	no	yes	yes	Female	Not Detected	-30.4138	153.0132
Coffs_19.1	11	no	yes	yes	Male	Chlamydia Likely	-30.4028	152.9869
Coffs_19.2	12	no	yes	yes	Female	Not Detected	-30.4483	153.0571
Coffs_19.3	13	no	yes	yes	Male	Chlamydia Likely	-30.3972	152.9917
Coffs_2.1.3	14	no	yes	yes	Female	Not Detected	-30.3398	153.0928
Coffs_22.1	15	no	yes	yes	Male	Possible Chlamydia	-30.4482	153.0416
Coffs_4.1	16	no	yes	yes	Male	Not Detected	-30.3291	153.0858
Coffs_6.1	17	no	yes	yes	Female	Possible Chlamydia	-30.321	153.0779
Coffs_6.11	18	no	yes	yes	Male	Possible Chlamydia	-30.3193	153.0727
Coffs_6.13	19	no	yes	yes	Female	Possible Chlamydia	-30.3198	153.0732

Coffs_6.5	20	no	yes	yes	Male	Not Detected	-30.3223	153.074
Coffs_6.6	21	no	yes	yes	Female	Not Detected	-30.3229	153.0729
Coffs_7.3.1	22	no	yes	yes	Female	Possible Chlamydia	-30.3602	153.061
Mylestrom_1.1	23	no	yes	yes	Male	Chlamydia Likely	-30.4549	153.0488
Coffs_1.5	-	no	no	no	-	-	-30.355	153.0785
Coffs_1.6	-	no	no	no	-	-	-30.3554	153.077
Coffs_1.8	-	no	no	no	-	-	-30.3534	153.0748
Coffs_19.4	-	no	no	no	-	-	-30.4224	153.0072
Coffs_2.1.1	-	no	no	no	-	-	-30.34	153.0934
Coffs_2.1.2	-	no	no	no	-	-	-30.3401	153.0932
Coffs_2.1.5	-	no	no	no	-	-	-30.3385	153.0925
Coffs_4.2	-	no	no	no	-	-	-30.329	153.0862
Coffs_6.10	-	no	no	no	-	-	-30.3251	153.0827
Coffs_6.12	-	no	no	no	-	-	-30.3198	153.0722
Coffs_6.14	-	no	no	no	-	-	-30.3253	153.08
Coffs_6.2	-	no	no	no	-	-	-30.3218	153.0777
Coffs_6.3	-	no	no	no	-	-	-30.3217	153.0778
Coffs_6.4	-	no	no	no	-	-	-30.322	153.0741
Coffs_6.7	-	no	no	no	-	-	-30.3232	153.0733
Coffs_6.8	-	no	no	no	-	-	-30.3238	153.0731
Coffs_6.9	-	no	no	no	-	-	-30.3257	153.0829

Table 2: Table with sPCA values for each unique koala used for population genetic analyses

Sample ID	Koala ID	sPCA value
Mylestrom_1.1	23	-5.6556
Coffs_22.1	15	-4.60387
Coffs_19.1	11	-4.44609
Coffs_19.2	12	-4.38355
Coffs_19	10	-3.31423
Coffs_19.3	13	-3.14054
Coffs_7.3.1	22	0.388585
Coffs_7.1	5	0.455042
Coffs_13.1	7	1.455932
Coffs_1.3	3	1.563357
Coffs_2.1.3	14	2.272557
Coffs_6.13	19	2.310041
Coffs_2.1	2	2.315766
Coffs_13.3	4	2.328692
Coffs_13.5	9	2.364646
Coffs_6.1	17	2.364646
Coffs_4.1	16	2.391483
Coffs_6.11	18	2.416676
Coffs_1.7	1	2.447457
Coffs_6.5	20	2.457892

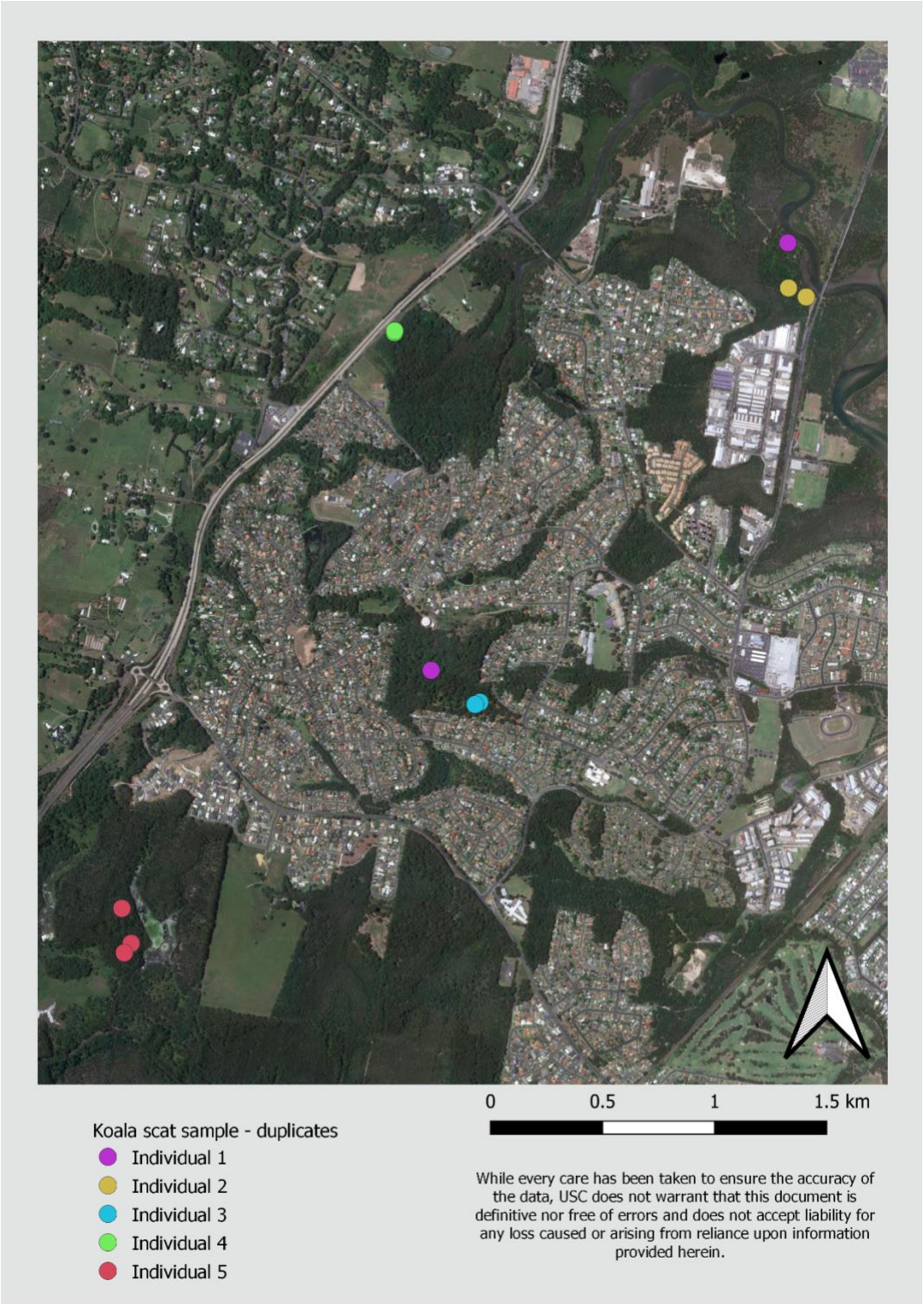

Figure 1: Location of duplicated samples collected from the same individual koala.