

Susceptibility to a sexually-transmitted disease in a wild koala population shows heritable genetic variance but no inbreeding depression

Romane Cristescu¹, Kasha Strickland², Anthony Schultz¹, Loeske Kruuk³, Deirdre de Villiers⁴, and Celine Frere⁵

¹University of the Sunshine Coast

²The University of Edinburgh Institute of Evolutionary Biology

³Australian National University Research School of Biology

⁴Affiliation not available

⁵The University of Queensland

July 25, 2022

Abstract

The koala, one of the most iconic Australian wildlife species, is facing several concomitant threats that are driving population declines. Some threats are well known and have clear methods of prevention (e.g. habitat loss can be reduced with stronger land-clearing control), whereas others are less easily addressed. One of the major current threats to koalas is chlamydial disease, which can have major impacts on individual survival and reproduction rates, and can translate into population declines. Effective management strategies for the disease in the wild are currently lacking, and to date we know little about the determinants of individual susceptibility to disease. Here we used a rare opportunity to investigate the genetic basis of variation in susceptibility to chlamydia using one of the most intensively studied wild koala populations. We combine data from veterinary examinations, chlamydia testing, genetic sampling and movement monitoring. Out of our sample of 342 wild koalas, 60 were found to have chlamydia. Using genotype information on 8649 SNPs to investigate the role of genetic characteristics in determining disease status, we found no evidence of inbreeding depression, but a heritability of 0.14 (95%CI: 0.06 – 0.23) for the probability that koalas had chlamydia. Heritability of susceptibility to chlamydia could be relevant for future disease management in koalas, as it suggests the potential to select for disease resilience through assisted breeding.

Introduction

Disease can pose a serious threat to flora and fauna, recently demonstrated through the global impact of white nose syndrome , sarcoptic mange , and anthrax . Similarly, plant diseases such as chestnut blight and Dutch elm disease have had large-scale impacts on species and ecosystem diversity . It is also likely that threats to wildlife populations from disease will increase in frequency, due to interactions with both climate change and accelerating anthropogenic disturbances . In addition to the threat to biodiversity, wildlife diseases can have substantial economic impacts and affect ecosystem services . Wildlife diseases also pose a growing threat to human health, with an estimated 75% of emerging infectious diseases in humans being of wildlife origin (Merianos 2007).

Managing diseases in wild populations is difficult, in part because effective management of disease requires an understanding of host-pathogen-environment dynamics that may be costly to obtain . Intensive management such as vaccination might only be feasible in certain situations, where vaccines can be distributed in baits that are readily taken (e.g. rabies vaccine for carnivores,) or for small, contained populations . Targeted

disease mitigation strategies in the case of the chytrid fungus in amphibians, for instance, have included preventing pathogen spread to naive populations, establishing ex-situ insurance colonies, and captive breeding of amphibian hosts . To achieve effective management of already-infected populations, however, mitigation strategies need to integrate specific host-pathogen interactions, in particular targeting pathogenicity and host susceptibility . Moreover, as part of the increased appreciation of the value of evolutionary insights for wildlife disease management, there is increasing awareness of the need to understand what factors contribute to susceptibility to disease. This information can be fundamental in predicting whether species and populations will adapt to diseases, and could enable us to assist this process by selecting resilient animals .

Many factors influence the susceptibility of a host to a pathogen. The host's environment will have multiple effects on its susceptibility to disease: for example, the quality and quantity of food resources can influence host condition and hence immunological health , and both the prevalence of the pathogen and characteristics of individuals' social environment can affect infection risk . Host genetics may also play a role in determining individuals' susceptibility to disease . For example, genetic variation may affect susceptibility via additive genetic effects (Hill 2012). Another genetic characteristic related to the ability of a host to cope with disease is inbreeding, the risk of which increases with small populations. Links between inbreeding and lowered disease resilience (i.e., the cumulative impact of resistance and tolerance) may be driven by inbreeding depression, explained through decreased heterozygosity in genes of the immune system (e.g., major histo-compatibility complex), and the loss of alleles linked to resilience . As such, determining the extent of heritability of susceptibility to disease and potential inbreeding depression linked to the disease could be valuable for disease management. In particular, a genetically heritable basis to variation in susceptibility indicates the potential for evolutionary responses, and for selective breeding for resistance, whereas the existence of strong effects of inbreeding would support efforts to improve genetic variance within populations.

In this study, we present quantitative genetic and inbreeding analyses of a disease affecting an Australian mammal, the koala *Phascolarctos cinereus* . Recent declines in population size, including those due to extreme bushfires in 2020, have led to the classification of koalas in Queensland, New South Wales, and the Australian Capital Territory (approximately 60% of their range) as endangered in 2022. Whilst the bushfires decimated large areas of koala habitat, the bacterial pathogen *Chlamydia* has been identified as a major threat to their survival . The bacterium causes chlamydial disease, which primarily affects the ocular and urogenital tracts in koalas and decreases both survival and reproduction rates . The main transmission path is sexual, although some mother-offspring transmission does occur . The severity of chlamydial disease varies greatly both between and within populations , and not all koalas that are infected progress to 'diseased' status (i.e., show clinical signs of disease) . Chlamydial disease is the most common illness causing death of wild koalas and is such a critical threat to koalas that, when adequately addressed, population declines can be reversed . However, management of the disease currently requires costly treatment, including catching koalas and multiple weeks of antibiotic treatment at wildlife hospitals . An understanding of the potential for evolutionary responses could therefore greatly improve management strategies for koala populations.

Using high-resolution data from a wild population of koalas, we investigated the effect of genetic and environmental factors on chlamydia disease. There is a high prevalence of chlamydia disease in the study population (28%,), and chlamydia is a leading cause of mortality (18% of all deaths,). Further, we have previously demonstrated that koalas in this population do not avoid mating with either diseased, or with closely-related individuals . Here we aimed to 1) determine whether there was evidence for inbreeding depression associated with disease susceptibility of koalas, and 2) partition variance in koalas' susceptibility to disease into additive genetic and environmental components. Our analysis provides insight into the extent to which an individual's genes and/or environment affects their susceptibility to disease, and whether a wild population could respond to selection for improved resilience to chlamydial disease.

Methods

Study population

Our study population consisted of 519 koalas that were monitored between 2013 and 2017 as part of a man-

agement program associated with a 13km rail infrastructure project in Moreton Bay Council, Queensland, Australia (area centroid: -27.234° ; 153.036°), and described in detail in . During the five-year program, most of the population ($>95\%$ of individuals in the study area) was intensively monitored. Individuals were first captured between 13/03/2013 and 02/02/2017 using either live traps or flagging pole methods, and were fitted with VHF collars. After their first capture, individuals were then periodically re-caught at least every six months for routine veterinary exams, or earlier if they were observed with visible signs of disease (see below) or injury. Tissue samples were taken and ear tags fitted at first capture, and blood samples were taken during routine veterinary examinations, at first capture and at subsequent check-ups. Blood samples were stored at -20°C , and tissue samples were stored in 70% ethanol. Individuals were given a comprehensive veterinary examination each time they were caught, which included recording of sex, age, and presence of a joey (offspring) for females (either through the female having a back-riding or pouch joey or an elongated teat, or through a pregnancy observed by sonogram). Ageing of koalas relied on examination of their tooth wear . They were then tested for the presence of chlamydia bacterium and inspected for visible signs of chlamydial disease (see below for full details). During the project, individuals were treated for any illness (including chlamydial disease, see below) or injury, and some individuals were also included in a chlamydia vaccination trial, which involved a subset of individuals being vaccinated for chlamydia at their first capture (see). VHF tracking of individuals took place approximately twice a week. By identifying the time of the year in which joeys were born, and accounting for the gestation period (one month), based on the birth of 350 joeys between 2013 and 2016, we determined the breeding season to occur from September to December in this population. The analyses presented here used data from those individuals for which we had complete genetic, spatial and disease status data (described below), a total of $N = 342$. The dataset included some joeys ($N = 4$), because dependent joeys have been found to be infected by chlamydia as young as nine months old (Nyari et al. 2017) and the youngest dependent joey sampled in this survey was 10 months old.

The monitoring program was conducted under animal ethics approvals (Queensland Department of Agriculture and Fisheries CA 2012/03/597, CA 2013/09/719, CA 2014/06/777, CA 2015/03/852, and CA 2016/03/950) and scientific purposes permits (Queensland Department of Environment and Heritage Protection WISP 11525212, WISP 16125415, WISP 13661313, WITK 14173714 and WISP 17273716).

Chlamydial disease

Of the two species of chlamydia that infect koalas (*Chlamydia pecorum* and *C. pneumoniae*), *C. pecorum* is consistently more prevalent (i.e., higher percentage individuals infected) and more pathogenic . *C. pecorum* is therefore the chlamydia species that researchers and veterinarians focus on . At each veterinary examination, koalas were (1) tested for the presence of *C. pecorum* , and (2) assessed for clinical signs of chlamydial disease. Tests for *C. pecorum* were conducted using either quantitative real-time PCR (qPCR) targeting a fragment of the *C. pecorum* 16s rRNA gene , or using the Clearview® *Chlamydia* MF test kit, which has a 60% sensitivity (Inverness Medical, Unipath Ltd). Clinical signs of chlamydial disease are mostly found in the ocular and urogenital regions. The ocular form of the disease varies from inflammation of the mucosal surfaces of the eye, or conjunctival hyperplasia, to the complete opacification and scarring of the cornea, which can cause blindness . In urinary disease, inflammation of the bladder (cystitis) leads to incontinence and staining of the fur around the cloaca, which is also coupled with alopecia and ulceration in severe cases . Finally, infection of the reproductive tract in females causes bursal cysts surrounding the ovary and upper reproductive tract pathology, while in males orchitis and epididymitis can occur . In this study, veterinary diagnostic techniques for assessing the presence of chlamydial disease involved observation of external signs of disease, cystocentesis with observation of the urine sediment, and ultrasound examination of the kidneys, reproductive tract, and bladder (as described in). We initially aimed to independently analyse whether koalas tested positive for the presence of *C. pecorum* , and whether koalas were diagnosed with clinical signs of disease. However, in our data, the majority of individuals that tested positive also displayed clinical signs of disease (87%). We therefore opted to use whether koalas tested positive ('chlamydia status') as our response variable in all analyses.

Although individuals were monitored for multiple years (average = 1.2 years, max = 4 years) and had

repeated measures of disease status gathered at scheduled veterinary exams ($N = 1182$ observations of $N = 342$ koalas, average = 3 observations per koala), we selected to only use chlamydia status at their first capture. This was because during the project individual koalas were treated for any illness or injury they presented at veterinary examinations (including chlamydial disease), and some individuals were also included in a chlamydia vaccination trial. Individual disease status at subsequent veterinary examinations could therefore have been affected by the veterinary treatment, vaccine trial and management of the population.

DNA extraction and genotyping

We conducted DNA extraction and genotyping as previously described for this population. Briefly, we used single nucleotide polymorphism (SNP) genotyping data from blood or tissue samples. DNA was extracted using the DNeasy Blood and Tissue Kit (QIAGEN), following the manufacturer’s protocol, and DNA extracts were stored at -80°C . Genotyping was conducted as per by Diversity Arrays Technology, Canberra, using their proprietary DArTseqTM technology. DArTseqTM utilizes a combination of next-generation sequencing platforms and DArT complexity-reduction methods, and, similar to DArT methods based on array hybridizations, the protocol is optimized for a specific organism and application by selecting the most appropriate complexity reduction method. This is assessed based on minimizing skewed size ranges, non-ideal numbers of fragments, and percentages of repetitive elements. Samples were then processed as per. Genotyping produced a total of 8649 SNPs, which were then filtered to remove SNPs with $\text{MAF} < 1\%$; call rate $< 95\%$; technical replicate scores $< 95\%$; and HWE where $p > .05$. This resulted in a total of 5007 SNPs that were used for subsequent analyses. Observed and expected heterozygosity, as well as the inbreeding coefficient F_{IS} were calculated using the dartR package, following the same procedure as in.

Relatedness assignment

We calculated pairwise relatedness estimates using the R package *Related* using the filtered set of 5007 SNPs, as per. Pairwise relatedness across all pairs of individuals in our genetic dataset was estimated using Queller and Goodnight’s relatedness estimator. Although there are multiple relatedness estimates available, we selected to use Queller and Goodnight’s because it has previously been found to be the most accurate for this dataset in determining true pairwise relationships. Briefly, using simulated genotypes of known relationships, compared the accuracy of different estimates in correctly determining relationships (i.e. parent-offspring or half-sibs), finding that Queller and Goodnight’s estimate was the most accurate. Maternities were identified through genetic parentage analyses, or, rarely, through field observations. Maternity assignment was conducted using a reduced panel of more stringently filtered SNPs ($N = 427$ SNPs, $\text{MAF} > 0.3$, missingness $< 1\%$, consistent marker scores for technical replicate assay pairs of $>99\%$), linkage disequilibrium $r^2 < 0.4$) in *CERVUS* 3.0.7 as per.

Inbreeding Coefficients.

Using the same SNP set ($n = 5007$ SNPs) as for pairwise relatedness, we calculated ‘internal relatedness’ (IR) for each individual. This measure was chosen as it strongly correlates with standard measures of heterozygosity but incorporates further information about the frequency of the alleles in the individual’s genotype. As such, internal relatedness is a common measure of individual heterozygosity used to investigate inbreeding depression. When calculated over multiple loci, individuals’ internal relatedness values in a population are approximately normally distributed with negative values suggesting more ‘outbred’ individuals and positive values being suggestive of inbreeding. We used the *GENHET* v3.1 function in the R statistical environment to estimate IR as $IR = \frac{(2H - \sum f_i)}{(2N - \sum f_i)}$, where H is the number of loci that are homozygous, N is the number of loci and f_i is the frequency of the i th allele contained in the genotype.

Shared environment effects

Phenotypic similarity between individuals in any quantitative trait may be generated by individuals experiencing similar environmental conditions, as well as by their shared genes. This may include, for example, effects of variation in resources (e.g. habitat quality) or of social networks (two individuals with spatial overlap will be more likely to encounter the same conspecifics). We accounted for the possible effects of

shared environments by estimating individuals' overlap in their spatial environments, using data on their home ranges (see model details below). Home ranges were estimated as utilization distributions using VHF radiotracking locations of collared koalas. Tracking data was therefore collected after our measure of infection (i.e. whether koalas tested positive at first capture). However, koala home ranges are thought to be relatively stable through time, even when disturbed by habitat changes. Utilization distributions describe the frequency distribution of individuals' location data, and estimate the probability of an individual occurring across the study site. We included all individuals with at least five sightings over the study period ($N = 342$). The average number of sightings per individual was 186 (interquartile range = 48 – 346). Utilization distributions were estimated using the *adehabitatHR* package in the R statistical environment with a smoothing parameter of 100 meters and grid size of 50 meters. This smoothing parameter was selected to 1) standardize the estimation of home ranges of all individuals, and 2) avoid over- or under-smoothing in utilization distribution estimation for individuals with few sightings. All utilization distributions were visually inspected post estimation in order to ensure accuracy. Home range overlap (abbreviated "HRO" in the results) between all pairs of individuals was then measured using utilization distribution overlap indices, from which a home-range overlap matrix was constructed. This measures the extent to which koalas overlap spatially, although it does not account for the possibility that some pairs of koalas may co-occur spatially but not temporally.

Statistical Analysis.

We first tested whether our genetic dataset contained the power to detect inbreeding (if present) using the R package *inbreedR* to estimate identity disequilibrium (g_2). Specifically, this tested for variance in inbreeding in the population, which is required to detect inbreeding depression. Then, to investigate whether there was evidence for inbreeding depression, additive genetic effects, and/or shared environment effects, we ran a suite of generalized linear mixed effects models using the MCMCglmm package in the R statistical environment. In all models, our response was a binary variable describing whether or not the koala tested positive for *C. pecorum* at first capture (effectively a case/control comparison), and all models were fit with the following fixed effects: age in years at time of capture, season in which the individual was caught (breeding *vs* non-breeding), sex, IR, and the interaction between age and IR. The effect of IR tested for evidence for inbreeding depression, predicting that if there was evidence for inbreeding depression, more inbred individuals would have poorer health outcomes and would therefore be more likely to test positive for *C. pecorum*. We also included an interaction between IR and age to test whether the effect of inbreeding changes with age. Although it is likely that the components of variance also change with age, our data-set structure and size were not suitable for addressing this question: of the 342 individuals in the dataset, only 20 individuals were between 0-1 years old, and 81 between 1 – 2 years old, resulting in very few observations of individuals prior to sexual maturity, making estimating variances at different ages difficult using this dataset. Fixed effects were given flat weakly informative priors, and random effects were given a X_1^2 prior, following the advice of. In this paper, de Villemereuil and colleagues found that when estimating heritability of binary traits, this prior was less sensitive to the inclusion of multiple random effects than alternative priors and performs best with small sample sizes. We ran 1,030,000 iterations per model with a burn in period of 30,000 iterations, sampling at intervals of 1000 iterations; this resulted in low autocorrelation and a sufficient number of iterations for the model to mix and converge. Convergence of models was assessed by examining trace-plots to visualise sampling mixing and by assessing effective sample sizes. Further, we plotted predicted random effect values to visually check for normality. We considered estimates of fixed effects to be different from zero when the 95% credible intervals of the posterior distribution did not overlap with zero. These models were fitted with a threshold distribution and a probit link. Residual variance was fixed at one due to the use of binary data.

The first model we fitted used the full dataset including all individuals for which we had disease, genetic, and spatial data ($N=342$), and decomposed variance not accounted for by the fixed effects into two components: additive genetic effects and shared-environment effects. Additive genetic effects were estimated by fitting the relatedness matrix as a covariance matrix. This was therefore an 'animal model', which extends linear mixed effects models to incorporate relatedness information, and partitions phenotypic variance into additive

genetic effects and other sources of variance . Shared-environment effects were estimated by including the home range overlap matrix as an additional covariance matrix, allowing us to estimate the variance associated with individuals sharing the same environment. Prior to fitting the animal model, we ensured that the relatedness matrix and the home range overlap matrix were not correlated with each other and would not affect variance partitioning. The Pearson’s correlation between the two matrices was 0.3, which we deemed low enough to not affect parameter estimates derived from the model. Nevertheless, we also fitted the model without the spatial overlap matrix and compared the deviance information criterion (DIC) of the two models to further examine evidence for a shared environment effect. Doing so also enabled us to ensure that fitting the spatial overlap matrix did not influence our heritability estimates. Fitting animal models requires positive definite matrices, and as our relatedness matrix was not positive definite, we calculated the nearest positive definite matrix to our observed relatedness matrix using the *corpcor* package in R . We subsequently ensured that the original information contained in the relatedness matrix was unchanged by calculating and visualizing the correlation between the observed and the new matrix ($r^2 = 0.98$). The observed spatial overlap matrix was positive definite and therefore did not require transformation.

The second model we ran was used to investigate the extent to which the probability that a koala tests positive for *C. pecorum* may be caused by vertical transmission from mother to offspring (maternal effects). This is necessary because heritability estimates may be inflated when maternal effects are not accounted for (Kruuk and Hadfield, 2007). To do this, we used a subset of the data used for the first model that included only the individuals for which we knew the mother ($N_{\text{IND}} = 195$ of $N = 106$ mothers). Although analytically possible, we considered this reduced sample size too small and lacking in statistical power to fully partition variance in disease status into additive genetic effects, maternal effects, and shared environment effects. Instead, we aimed to (1) estimate the extent of maternal effect variance, and (2) ensure that our heritability estimates were not inflated by potential mother-offspring transmission of disease. To do this, we again fitted the model with the relatedness matrix (as explained above), but here included maternal ID as an additional random effect (instead of the shared environment effect). The maternal ID term estimates the phenotypic variance that is attributed to individuals sharing the same mother, over and above that due to shared additive genetic effects .

Heritability estimation

We estimated narrow-sense heritability (h^2) as the amount of additive genetic variance divided by the total phenotypic variance (i.e., the sum of the different variance components). Because the model was fitted with binary data, all variance estimates of the model were calculated on a latent scale. It is possible to measure heritability on either the latent trait scale or the observed data-scale, the selection of which depends on inferences being made . Heritability on the latent scale was estimated as $h^2 = \frac{V_A}{V_A + V_{SE} + V_e}$, where V_A is the additive genetic variance, V_{SE} is the variance associated with shared environment, and V_e is the residual variance, which was fixed at 1. In this case, we also estimated heritability on the observed data-scale, as this provides parameter estimates which are directly interpretable in relation to the ecology of the species whilst incorporating other factors (i.e., fixed effects and residual variance) which may influence the probability of contracting chlamydia. To convert parameter estimates, we used the *QGglmm* package in R . This package uses estimates of additive genetic variance, phenotypic variance (sum of all random effects variance + residual variance) and the intercept and converts them to the data-scale, thereby allowing for the calculation of heritability on this scale. We repeated this process for all random effect variances, thus measuring the proportion of phenotypic variance attributed to either shared environment effects or maternal effects (hereafter referred to as intra-class coefficient (ICC)).

Results

A total of 342 individuals for which we had genetic, spatial and disease status data were included in our final dataset. This consisted of 155 males (of which one was a joey <1 year old caught with its mother) and 187 females (of which three were joeys <1 year old caught with their mothers). We were able to detect maternities for 195 individuals, who were the offspring of 106 unique mothers.

Out of the total 342 individuals, there were 60 that tested positive at first capture. Observed heterozygosity was estimated as 0.223, within population genetic diversity (also known as expected heterozygosity) was estimated as 0.278, and the population average inbreeding coefficient, F_{IS} , was estimated as 0.196. The identity disequilibrium proxy (g_2) in our 5007 SNP dataset differed significantly from zero ($g_2 = 0.013 \pm 0.002$ bootstrap confidence interval = 0.0101 – 0.0167, $P(g_2 = 0) = 0.001$), where permutations = 1000 and bootstraps = 1000). This indicated that our dataset met the requirements of variance between individuals in levels of inbreeding, which is required for detecting if there is inbreeding depression in a population. Accordingly, IR varied substantially between individuals, ranging from a minimum estimate of -0.39 (expected when the individuals’ parents are outbred) to 0.38 (expected when the individuals’ parents are related to one another) (mean IR = 0.05, standard deviation = 0.09, Fig 1a). The variance in pairwise relatedness values was 0.004, with approximately 232 pairs of first-degree relatives (i.e., parent-offspring pairs or full siblings) and 680 pairs of second-degree relatives (see Figure 1b).

There was no evidence for sex differences in the probability that a koala tested positive for *C. pecorum* (posterior mode $\beta_{MAAE} = 0.037$, Table 1). Koalas had a higher probability of testing positive for *C. pecorum* in the breeding season compared to the non-breeding season (posterior mode $\beta_{BPEE\Delta INT} = 0.623$, Table 1). The probability of testing positive for *C. pecorum* also increased with age (posterior mode $\beta = 0.145$, Table 1, Fig 2). There was no evidence that internal relatedness affected the probability that koalas tested positive for *C. pecorum* (posterior mode $\beta = 1.071$, Table 1, Figure 3). Furthermore, there was no evidence that the effect of internal relatedness changed with the age of the koala (IR*age interaction, posterior mode $\beta = -0.038$, Table 1). Together, the lack of an association between IR and disease suggests that there was no evidence for inbreeding depression in this population.

The probability of testing positive for *C. pecorum* was associated with moderate levels of additive genetic variance (V_A) on both latent and observed data scales, whereby the posterior distribution was clearly different from zero on both scales (Table 1, see supplement for posteriors). More specifically, the V_A for the probability of testing positive for *C. pecorum* was estimated at 1.35 on the latent scale (posterior mode, 95% $_{CI} = 0.23 - 2.93$, Table 1), and 0.008 on the observed scale (posterior mode, 95% $_{CI} = 0.003 - 0.33$, Table 2). The heritability (h^2) for the probability of being diseased was estimated at 0.57 on the latent scale (posterior mode, 95% $_{CI} = 0.33 - 0.74$), and 0.11 on the observed scale (posterior mode, 95% $_{CI} = 0.06 - 0.23$, Table 1).

We found no evidence that the probability of testing positive for *C. pecorum* was associated with variance in shared environment effects (V_S). The posterior distribution for V_S (measured using the home range overlap matrix) was very low and bordered zero on both the latent and observed data scales (latent scale; posterior mode = <0.001 , 95% $_{CI} = <0.001 - 0.071$; observed data scale; posterior mode = <0.001 , 95% $_{CI} = <0.001 - <0.001$, Table 1). This represented $<1\%$ of variance in probability of being diseased on both latent and data scales. Furthermore, the inclusion of V_S in the model did not improve the fit of the model (DIC $V_A + V_S = 278.69$, DIC $V_A = 277.88$). Finally, V_A estimates did not vary qualitatively between models with and without the home range overlap matrix, suggesting that its inclusion did not affect our reported estimates of h^2 (see supplement).

Using a subset of the data that included only individuals for which we knew the identity of their mothers ($N = 195$), we found only a small maternal effect variance (V_{ME}) in the probability of testing positive for *C. pecorum*. More specifically, the posterior mode of V_{ME} was estimated as 0.005 on the latent scale, and <0.001 on the observed data-scale, and the lower tail of the posterior distribution bordered zero on both scales (see Table 2 and supplement for full posteriors). This corresponded to a low proportion of phenotypic variance attributable to maternal effects (ICC_{ME}), with no indication that ICC_{ME} was statistically different from zero (Table 2). Given the low sample size, it may be that we lacked the statistical power to separate V_A from V_{ME} within a single model. However, when running a model without estimating V_A (i.e., without the relatedness matrix), we found that V_{ME} did not differ (qualitatively) to that estimated when fitting both simultaneously (see supplement). Moreover, the DIC was lowest for the model containing just the relatedness matrix, and highest for the model containing just the maternal effects (DIC: $V_A + V_{ME} = 277.66$, $V_{ME} = 314.20$, $V_A = 140.64$). Together, this suggests that maternal effects explained little to none of the phenotypic

variance in the probability of individuals testing positive for *C. pecorum* . Finally, estimates for V_A and for h^2 did not differ qualitatively between the model containing the maternal effect vs without, suggesting that our reported heritability estimates (Table 1) were not inflated by maternal effects that were unaccounted for.

Discussion

Our results showed that aging and breeding season increased the probability of being diseased in a wild koala population, but there was no evidence of effects of either sex or inbreeding on disease. We did, however, find evidence for heritable genetic variation in susceptibility to disease ($h^2 = 0.11$). We discuss these results in more detail below.

Absence of inbreeding depression

Inbreeding depression can be defined as heterozygosity-fitness correlations and, using this approach, we found no evidence of inbreeding depression in disease (a key fitness trait) in this koala population. Inbreeding depression is difficult to investigate in wild populations, not least because obtaining necessary datasets (including pedigrees and/or genome-wide markers and fitness measures) is both costly and labor intensive. However, given the potential consequences of inbreeding depression on population growth and survival , identifying the risk it poses to populations is critical to implementing informed management. Interestingly, empirical evidence of an association between disease and inbreeding in wild populations is mixed, with effects varying both between and within species , and with pathogen type . In koalas, evidence to date for inbreeding depression is also mixed. For instance, inbreeding depression was found when comparing cryptorchidism levels in island to mainland populations , but the relationship was not evident at the individual level , and did not correspond to decreased fitness, as the most inbred populations were also experiencing exponential population growth .

The effects of inbreeding on individual or population fitness can be dependent on environment, age, sex and the genetic constitution of populations . Interestingly, it has been suggested that koalas may have a reduced susceptibility to inbreeding depression, due to low historical effective population sizes and the elimination of recessive deleterious alleles . Although we did not detect an age-dependent effect of inbreeding in our analyses, there may have been other context dependent effects of inbreeding that we did not measure. As such, although the lack of inbreeding depression could reduce concerns for this population, our results provide no guarantee of future resilience to chlamydial disease, especially as the context in which koalas live rapidly changes under continued environmental degradation.

Heritability of susceptibility to disease in the wild

Measures of heritability for disease susceptibility in wild populations are relatively rare: for example, only between 0.03 and 2.5% of studies of heritability estimates of traits in wild populations were related to disease . However, some recent studies have found that there may be considerable additive genetic variation associated with disease susceptibility in the wild. For instance, heritability of infection risk was found to be 12% for *Mycobacterium bovis* in adult European badgers and 55% for *Mycoplasma ovipneumoniae* in bighorn sheep . Interestingly, we found relatively low environmental effect variance for disease in this population, explaining less than 1% of the total variance. The low shared-environment effect variance found here may reflect the small size of the study area that this population inhabits (of only 13 km length, and several hundred meters width), and therefore relative homogeneity in environmental conditions for all koalas. Nevertheless, considerable variation in susceptibility to disease remains unexplained (in the residual variance), suggesting that there are other environmental factors influencing disease in this population that were not captured by the shared environment effect. Moreover, additive genetic effects in disease are likely not homogeneous across ages . We were not able to formally investigate this, owing to low statistical power at different ages, but this would be an interesting and valuable follow-up study.

Narrow sense heritability (h^2), as calculated in this study, can be used to predict the response to selection across generations (Walsh and Lynch 2018), and can therefore be used to predict outcomes for populations

facing disease outbreaks . Heritability is a population-level parameter which depends on population-specific factors (e.g., allele frequencies or the effects of gene variants), traits, and variation due to environmental factors . Together with recent findings suggesting that polymorphisms in immune genes play an important role in koalas' ability to resolve a chlamydial infection , the additive genetic variance for susceptibility to infection suggests there may be some potential for koalas to respond adaptively to the presence of the pathogen. Given the degree of heritability found in this population, the immediate question is why have koalas in this population not yet adapted to be resilient to chlamydia?

There are several possible reasons for the maintenance of genetic variance in disease susceptibility in this population. First, genetic variance may be maintained through antagonistic pleiotropy with traits associated with other fitness components . Second, predation (predominantly by dogs) accounted for about 63% of all mortality in this population and could therefore have imposed a stronger selective pressure than disease. Third, koalas are hypothesized to have both their own chlamydia species and others more recently transmitted from livestock over the last 200 years (Timms 2005), so – whilst there is not yet any empirical support for this hypothesis – a possible explanation for the current virulence of chlamydia is that new strains have recently been introduced, and there has not been sufficient time for adaptation of resistance to these new strains. Finally, pathogens also adapt in response to host resistance and/or tolerance, and theoretical models suggest that maintenance in genetic variation in resistance to pathogens may be explained by a continuous process of host-pathogen coevolution .

Can our findings be used to improve koala population management?

Management of the threat of chlamydia to koalas has been difficult thus far: vaccine development is promising, but still a work in progress , and treatment is effective only in certain koalas or for a limited time . Therefore, although treatment of isolated koala populations can decrease chlamydia prevalence and reverse population decline , in non-isolated populations into which diseased animals may immigrate, maintaining effective treatment will likely require intensive and repeated interventions. Heritability of susceptibility to disease found here might open the door to enhancing adaptation through genetic rescue or artificial selection . Individuals that are the most resilient to disease could be identified and selected for breeding programs and/or translocations into struggling populations . Strategies such as this are not without risk or concerns. In general, these centre around the potential decrease in genetic diversity that can be associated with selection, ethical and moral concerns regarding anthropogenic interference with natural processes, or whether resources might be better allocated to more urgent or productive causes . One also needs to ensure that efficient selection in the face of current challenges does not compromise resilience to uncertain future challenges .

Disease management is one of the greatest challenges of wildlife management. It is hard to gain knowledge of diseases in the wild, although non-invasive methods are promising. For example, surveillance can be enhanced by the use of detection dogs and pathogen detection from scats (for examples with koalas and chlamydia, see). Ultimately, disease management options must be supported by decision makers and the public, and usually need to balance a multitude of factors – including risk and ethical perception, cost, concern and tolerance about (perceived or real) impacts on wildlife, livestock and human health, values and social acceptability of management actions as well as scientifically expected outcomes and uncertainties. Incorporating information about genetic susceptibility for disease when assessing risk to populations and in conservation planning is therefore important in ensuring we implement effective management strategies.

Acknowledgements

The authors would like to acknowledge the Kabi Kabi, Jinibara and Turrbal people, the Traditional Custodians of the land where this project was delivered, and pay the respects to their Elders past, present and emerging. We recognize their continuous and ongoing relationship with the land, waterways and sea. This project was significantly supported by the Queensland Government (Department of Transport and Main Roads), and specifically the Moreton Bay Rail project. We acknowledge, with thanks, the staff from Endeavour Veterinary Ecology and contractors who worked on the Moreton Bay Rail Koala Management

Program. Without their hard work collecting and curating the vast spatial, veterinary, and genetic data set from the Moreton Bay Rail Koala Management Program, this research would not have been possible. LK was supported by an Australian Research Council Laureate Fellowship.

References

Data Accessibility and Benefit-Sharing Section

All data (individuals’ meta-data, pairwise relatedness matrix and pairwise spatial overlap matrix) used in this manuscript will be made available in the dryad repository upon acceptance for publication.

Author Contributions

RHC, KS and CHF conceived and designed the research; DdV collected data; AS and KS analysed SNP data to generate relatedness and internal relatedness information; KS analysed the data, with input from LK and CHF; KS and RHC wrote the manuscript, and all authors contributed to reviewing and editing.

Figure 1. Distribution of (a) internal relatedness of koalas used in analyses, and (b) pairwise relatedness from the matrix used to estimate additive genetic variance. Pairwise relatedness was estimated using Queller and Goodnight estimate (Queller and Goodnight 1989), where, for instance, full-sibs and parent-offspring pairs would have a value of 0.5, and half-sibs would have a value of 0.25.

Hosted file

image1.emf available at <https://authorea.com/users/466420/articles/578384-susceptibility-to-a-sexually-transmitted-disease-in-a-wild-koala-population-shows-heritable-genetic-variance-but-no-inbreeding-depression>

Figure 2. The effect of age on the probability of testing positive for chlamydia. Data points represent the raw data of whether the koala tested positive at first capture. Regression line derived from model estimates represents the predicted relationship between IR and probability of testing positive.

Hosted file

image2.emf available at <https://authorea.com/users/466420/articles/578384-susceptibility-to-a-sexually-transmitted-disease-in-a-wild-koala-population-shows-heritable-genetic-variance-but-no-inbreeding-depression>

Figure 3. The effect of internal relatedness (IR) of probability of testing positive for chlamydia. Data points represent the raw data of whether the koala tested positive at first capture. Regression line derived from model estimates represents the predicted relationship between IR and probability of testing positive.

Hosted file

image3.emf available at <https://authorea.com/users/466420/articles/578384-susceptibility-to-a-sexually-transmitted-disease-in-a-wild-koala-population-shows-heritable-genetic-variance-but-no-inbreeding-depression>

Table 1. Estimates for both fixed and random effects from a model used to investigate the effect of inbreeding, additive genetic variance (V_A) and variance in shared environment effects (V_S) on the probability of koalas testing positive for *Chlamydia pecorum*.

			Posterior distribution	Posterior distribution	Posterior distribution	Posterior distribution
			mode	mean	CI5%	CI95%
Fixed Effects	β	Intercept	-2.332	-2.436	-3.445	-1.381
		Sex _{MALE}	0.037	0.098	-0.402	0.587
		Age	0.145	0.147	0.033	0.248
		IR	1.071	1.496	-3.402	6.122

			Posterior distribution	Posterior distribution	Posterior distribution	Posterior distribution
<i>Random effects</i>	V_A	Season _{BREEDING}	0.623	0.691	0.217	1.122
		Age*IR	-0.038	-0.133	-0.179	0.852
	h^2	latent	1.348	1.575	0.225	2.935
		data-scale	0.008	0.014	0.003	0.326
	V_S	latent	0.573	0.576	0.33	0.737
		data-scale	0.112	0.142	0.059	0.229
	ICC _S	latent	<0.001	0.019	<0.001	0.071
		data-scale	<0.001	<0.001	<0.001	<0.001
		latent	<0.001	0.009	<0.001	0.034
		data-scale	<0.001	0.002	<0.001	0.008

Values in brackets are 95% credible intervals. ‘Sex’ indicates difference in males relative to females; ‘Age’ is the effect in years; ‘IR’ is an individual’s internal relatedness value; ‘Season’ is the effect of breeding season relative to non-breeding; and Age*IR is the interaction. Variance component estimates (for additive genetic variance V_A and shared environment variance V_S), heritability (h^2) and the proportion of variance attributed to a shared environment effect (ICC_S) are all presented on both latent and observed data-scale. Parameters estimates were converted to the observed data-scale using the `QGglmm` package (see Methods for details of models). N=324 individuals.

Table 2. Estimates for both fixed and random effects from a model used to investigate maternal effects on the probability of koalas testing positive for *Chlamydia pecorum* .

			Posterior distribution	Posterior distribution	Posterior distribution	Posterior distribution
<i>Fixed Effects</i>	β	Intercept	mode	mean	CI5%	CI95%
		Sex _{MALE}	-2.106	-2.482	-3.58	-1.4
		Age	-0.044	0.073	-0.438	0.5
		IR	0.103	0.12	0.001	0.2
		Season _{BREEDING}	0.961	1.381	-2.935	6.2
		Age*IR	0.767	0.698	0.198	1.2
			0.05	-0.134	-1.129	0.9
<i>Random effects</i>	V_A	latent	1.496	1.529	0.316	2.8
		data-scale	0.009	0.009	0.003	0.0
	h^2	latent	0.611	0.532	0.296	0.7
		data-scale	0.128	0.111	0.071	0.1
	V_{ME}	latent	0.005	0.204	<0.001	0.6
		data-scale	<0.001	0.001	<0.001	0.0
	ICC _{ME}	latent	0.001	0.073	0.001	0.2
		data-scale	<0.001	0.015	<0.001	0.0

Values in brackets are 95% credible intervals. ‘Sex’ indicates difference in males relative to females; ‘Age’ is the effect in years; ‘IR’ is an individual’s internal relatedness value; ‘Season’ is the effect of breeding season relative to non-breeding; and Age*IR is the interaction. Random effect variance estimates, genetic heritability (h^2) and proportion of variance attributed to maternal effects (ICC_{ME}) are all presented on both latent and observed data-scale. Parameters estimates were converted to the observed data-scale using `QGglmm` package. N=195 individuals, for whom maternity was known, born to 106 mothers.