Mutation rate estimate and population genomic analysis reveals decline of koalas prior to human arrival

Toby G. L. Kovacs[1], Nicole M. Foley[2], Luke W. Silver[1,3], Elspeth A. McLennan[1], William J. Murphy[2], Carolyn J. Hogg[1,3], Simon Y. W. Ho[1]

[1]School of Life and Environmental Sciences, University of Sydney, Sydney, Australia

[2]Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA

[3]Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Sydney, Sydney, NSW, Australia

Corresponding author: Toby G. L. Kovacs, toby.kovacs@sydney.edu.au

Abstract

The koala (*Phascolarctos cinereus*), an iconic Australian marsupial, has experienced substantial historical and contemporary population declines. Identifying the drivers of these declines has been hindered by limited genomic data and by uncertainty regarding the koala mutation rate. Here, we provide the first direct estimate of the koala mutation rate, based on four parent-offspring trios, yielding a mean of 6.12×10^{-8} mutations per base pair per generation (95% confidence interval: $5.03–7.45\times10^{-8}$). Using this estimate of the rate, we reconstructed the demographic history of koalas using 458 whole-genome sequences sampled across their entire range. Our results refine the estimated timing of past changes in population size, suggesting a large decline beginning ~100 kya, before the arrival of modern humans in Australia. The koala population then split into five genetic populations 6–30 kya, which are now spread across >3500 km on the east coast of Australia. We also use the koala mutation rate to infer recombination maps for each population, confirming lower recombination rates in marsupials than in eutherian mammals. These findings provide critical insights into the evolutionary history of koalas, while highlighting the impact of using species-specific evolutionary rates in the inference of demographic histories and recombination landscapes. Our estimates of the genome-wide mutation rate and population-specific recombination maps for koalas provide valuable resources for future evolutionary and conservation analyses of marsupials.

Introduction

The Australian continent contains a rich diversity of endemic flora and fauna, which endured severe environmental shifts throughout the late Cenozoic. The continent was dominated by wet forests during the Paleogene but underwent drastic changes during the Miocene as the Australian tectonic plate drifted northwards (White 1994). Widespread aridification was marked by the retraction of rainforests to the coasts, the opening of forests, and the expansion of grasslands and deserts (Byrne et al. 2011). Intense glacial and interglacial climate cycles during the Plio-Pleistocene then drove further shifts towards arid and fire-prone environments (White 1994). These changes have underpinned the evolution of Australian biota, driving adaptation to arid environments, the isolation of taxa in remnant mesic habitats, and the extinction of many lineages, especially rainforest specialists (Byrne 2008; Byrne et al. 2008; Byrne et al. 2011). More recently, the first arrival of modern humans 65–47 kya was followed by the extinction of Australia's megafaunal species, changes to fire frequency and habitats (O'Connell and Allen 2015; Clarkson et al. 2017).

Marsupials are among the most recognizable components of the Australian fauna. These include the koala (*Phascolarctos cinereus*), which is the sole extant member of the family Phascolarctidae. The koala lineage diverged from its closest extant living relatives, the wombats, ~36 Mya (Duchêne et al. 2018), but fossil evidence suggests that the late Oligocene and early Miocene contained a large diversity of koala species across multiple genera (Black et al. 2014; Black 2016; Crichton et al. 2023). The modern koala species appeared at least 350 kya, while most of the other species went extinct by the end of the Pleistocene (Black et al. 2014). Bioclimate modelling has estimated that koala habitat on the east coast experienced substantial contractions during the Last Glacial Maximum before expanding to its present-day range during the current interglacial period (Adams-Hosking et al. 2011). Contemporary

koala populations inhabit an enormous range (>3500 km) but have experienced recent declines as a result of land clearing, disease, hunting, feral dog attacks, vehicle strikes, and bushfires (Menkhorst 2008; Adams-Hosking et al. 2016; Cristescu et al. 2023), leading to their listing as Endangered in Queensland (QLD), New South Wales (NSW), and the Australian Capital Territory (ACT) in 2022.

Genetic studies have suggested historically small population sizes in koalas, with museum samples from the 1800s containing low mitochondrial diversity (Tsangaras et al. 2012). Studies of mitochondrial DNA have suggested that population sizes were stable over the last 50,000 years, before expanding after the Last Glacial Maximum (Neaves et al. 2016). Although evidence from exons has pointed to pronounced bottlenecks >300 kya (Lott et al. 2022), whole-genome analyses have suggested that koala populations declined precipitously after modern humans arrived on the continent (Johnson et al. 2018; De Cahsan et al. 2025). This aligns with global patterns of anthropogenic impacts on megafauna, including in Australia (Johnson 2016; Van Der Kaars et al. 2017; Adesanya et al. 2023). However, the demographic inferences from koala genomes relied on mutation rate estimates from distantly related eutherian mammals (humans and mice), introducing uncertainty in the timing of population-size changes. Therefore, a comprehensive analysis using a more relevant mutation rate can provide a clearer picture of the patterns and drivers of demographic change in koalas.

Mutation rates can be estimated using three main methods: (i) the phylogenetic method, which infers the mutation rate from an estimate of the substitution rate; (ii) the mutation-accumulation method, which requires the sampling of a population for many generations; and (iii) the parent-offspring trio method (Pfeifer 2020). The phylogenetic method has historically been the most widely used, given the relative ease of estimating the substitution rate using molecular-dating methods with fossil calibrations. However, the substitution rate is expected to underestimate the mutation rate, because a proportion of *de novo* mutations are removed

by purifying selection over time (Ho et al. 2011; Yoder and Tiley 2021). Mutation-accumulation and trio-based methods tend to yield broadly similar estimates of mutation rates (Wang and Obbard 2023), but the former method is only feasible for model organisms with short generations.

There has been rapid growth in mutation rate estimates from parent-offspring trios, particularly for vertebrate species (Bergeron et al. 2023; Wang and Obbard 2023). Although many estimates have been obtained for eutherian mammals, there have only been two estimates for marsupial species, each based on a single parent-offspring trio. The only estimate from an Australian marsupial (Tasmanian devil, *Sarcophilus harrisii*) was based on individuals from a highly inbred captive population and might not reflect the mutation rate in wild populations (Bergeron et al. 2023). Therefore, obtaining further estimates of mutation rates in Australian marsupials, particularly for the speciose order Diprotodontia, has the potential to produce substantial improvements in demographic and evolutionary inferences.

Another benefit of obtaining a species-specific mutation rate is accurately inferring other evolutionary parameters, such as recombination maps. Until recently, producing recombination maps was a laborious process that involved genotyping individuals of known pedigree through generations to allow the identification of crossover events between chromosomes (Burbrink et al. 2025). This barrier has now been overcome with the availability of reference-quality long-read assemblies for species of conservation concern (R. N. Johnson et al. 2018) and machine-learning approaches that can infer recombination maps from population genomic data (Adrion et al. 2020). These methods have been used to estimate the recombination landscape for a number of non-model species (Bredemeyer et al. 2023; Foley et al. 2024), but their accuracy has been hindered by uncertainty in mutation rates (Adrion et al. 2020).

Recombination rates in marsupials are expected to be lower than those of eutherian mammals, based on their linkage map length (Dumont and Payseur 2008). However, an accurate genomic estimation of the recombination landscape using a relevant mutation rate is required to confirm this. Species-specific recombination maps can also be used to assess whether genes of conservation concern are in regions of low or high recombination. This determines the impact of linked selection (Nordborg et al. 1996) and the likelihood of successfully incorporating diversity through translocations, helping conservation managers better tailor management strategies to a population's needs.

Here we estimate the genomic mutation rate of koalas from four parent-offspring trios and use it to infer demographic history from a large set of whole-genome sequences. We further leverage this mutation rate estimate to generate population-specific recombination maps for koalas and demonstrate their use by querying the recombination context of immune genes. Our study shows that the koala population began to decline prior to the arrival of modern humans on the Australian continent.

Results

Estimation of the koala mutation rate

We sequenced the genomes of 12 koalas from three families, comprising seven parents and five offspring, with an average coverage of 32× across the samples. The inferred relationships in one trio differed from that recorded in the studbook and so this trio could not be used to estimate mutation rates. We identified 100 unique *de novo* mutations in the remaining four offspring after filtering for false discovery following the methods used by Bergeron et al. (2023). We estimated a mean mutation rate of 6.12×10^{-9} mutations/bp/generation (95% confidence interval: $5.03–7.45\times10^{-9}$) across the four trios (Fig. 1*a*), using the false-negative

rate (0.068) and the size of the callable genome (average 68.6% total genome size) (supplementary table S1). The mean estimate can also be expressed as 8.74×10^{-10} mutations/bp/year, assuming a generation length of 7 years (Phillips 2000). The detected mutations had a transition to transversion ratio of 2.23 (Fig. 1*b*).

Estimation of demographic history

We performed population demographic analyses using 458 genomes of wild koalas, including all 413 samples from the Koala Genome Survey (Hogg et al. 2023; McLennan et al. 2024) and 45 additional samples from Kangaroo Island and the Mount Lofty Ranges (South Australia) to fill geographical gaps. Our phylogenetic network of wild and captive koalas showed a weakly structured population, with samples broadly grouping into the five populations that were recognized in a previous study (McLennan et al. 2024): North Queensland (N QLD), South-East Queensland/North New South Wales (SE QLD/ N NSW), Middle New South Wales (M NSW), Southern New South Wales (S NSW), and Victoria (VIC) (Fig. 2*b*). We also show that most of the additional South Australian koalas form a nested clade within the Victorian population. However, a few koalas from the Mount Lofty Ranges were spread throughout the Victorian clade, suggesting that multiple lineages could be present in South Australia.

We estimated the historical demography of the koala using the pairwise sequentially Markovian coalescent (PSMC; Li and Durbin 2011). Our results suggest that koala populations expanded 700 kya and remained stable until ~120 kya, when they experienced severe and prolonged reductions to reach their minimum population size ~60 kya (Fig. 2*a*). This pattern showed little variation within and between populations and did not appear to be influenced by the level of sequencing coverage (supplementary fig. S1). Our results also suggest that population sizes remained small following this bottleneck, although PSMC is not

particularly reliable for inferring demographic trends on recent timeframes (Patton et al. 2019).

When we used the default time intervals designed for analysing human genomic data, several of our samples produced PSMC plots with recent spikes in population size (supplementary fig. S2). These spikes are known artefacts in PSMC analyses and previous studies have shown that they can be removed by subdividing recent time intervals (Hilgers et al. 2025). We also found that, in cases where similar artefacts were produced when grouping the first four intervals (denoted as “4+…”), subdividing the first time interval into two or four intervals (as “2+2+…” and “1+1+1+1+…”) sometimes removed these peaks (supplementary fig. S2). In some cases, however, when grouping the first four intervals did not introduce an artefact, subdividing the first interval introduced an artificial peak in population size, suggesting that this issue should be assessed on a case-by-case basis.

Demographic inference using the sequentially Markovian coalescent can be improved by the addition of multiple samples per population, with SMC++ considered the state-of-the-art method for analysing samples of more than eight genomes (Terhorst et al. 2017). Our estimates of historical population sizes using SMC++ showed increased resolution, with a larger number of shorter time intervals, and with greater similarity in demographic plots between populations when increasing the number of knots (supplementary fig. S3). Our plots using 50 knots showed a similar pattern to that observed in our PSMC results, with koala populations increasing more gradually from 700 to 200 kya before declining between 110 and 40 kya to ~10% of their maximum population size (Fig. 3*a*). Following the bottleneck, the three northern populations (N QLD, SE QLD/N NSW, and M NSW) expanded at 15 kya before stabilizing. The South New South Wales population recovered slightly, whereas the Victorian population does not show any signs of recovery.

We also used SMC++ to estimate the split times between pairs of koala populations. Divergences among the populations were estimated to have occurred in the last 30 ky, following a general isolation-by-distance pattern (Fig. 3*b;* supplementary fig. S4). This suggests that the severe bottleneck occurred in the ancestral population of all extant koalas, with modern lineages subsequently diverging as the population sizes recovered. Based on the pairwise estimates between geographically adjacent populations, we found that the ancestral koala population split at 15–16.5 kya to form the three ancestral populations of i) N QLD and SE QLD / N NSW, ii) M NSW, and iii) S NSW and VIC. The two northernmost and two southernmost populations were the last to split, diverging from the remaining populations at 6.5 and 7.6 kya, respectively.

Estimation of the population specific koala recombination maps

We estimated differing recombination rates between populations using ReLERNN (Adrion et al. 2020) (Fig. 4), with the VIC population having the highest mean recombination rate of 1.40×10^{-9} crossovers/bp/generation (Q1–Q3: $0.83–1.84\times10^{-9}$ crossovers/bp/generation) and the SE QLD/N NSW population with the lowest recombination rate 0.74×10^{-9} crossovers/bp/generation (Q1–Q3: $0.51–0.91\times10^{-9}$ crossovers/bp/generation) (Fig. 4; supplementary table S3). The VIC population had a significant difference in recombination rate compared to the four other populations (Mann-Whitney test, all comparisons $p < 2.2\times10^{-16}$) (Fig. 4). The recombination landscape was highly correlated when using either the pan-mammal or koala-specific mutation rate (M NSW rho=0.74, $p < 2.2\times10^{-16}$; N QLD rho=0.65, $p < 2.2\times10^{-16}$; VIC rho=0.66, $p < 2.2\times10^{-16}$). However, the mean recombination rate was significantly higher when using the pan-mammal compared with the koala-specific mutation rate for the three populations analysed (Mann-Whitney test, all comparisons $p < 2.2\times10^{-16}$).

The components of the koala immune repertoire varied considerably in terms of their recombination context (supplementary table S4). The tissue-resident memory T cells were consistently found in high-recombining genomic regions across all populations. In contrast, major histocompatibility complex (MHC) class III genes were generally found in lower-recombining regions across all populations (supplementary table S4). However, the recombination context of some components differed between populations. In the VIC population, the MHC extended class II components occur in low-recombining genomic regions (1st decile), while the same components are found in predominantly high-recombining regions (9th decile) in the population from SE QLD/N NSW.

DISCUSSION

Germline mutation rate in the koala

Our sequencing and analysis of koala genomes from parent-offspring trios has yielded a mutation rate estimate of 6.12×10^{-9} mutations/bp/generation (95% confidence interval: 5.03–7.45×10^{-9}), similar to those reported for the Tasmanian devil (5.95×10^{-9} mutations/bp/generation) and Virginia opossum (*Didelphis virginiana*; 4.60×10^{-9} mutations/bp/generation) but considerably lower than that inferred for modern humans (11.6×10^{-9} mutations/bp/generation) (Bergeron et al. 2023). The koala yearly mutation rate (0.87×10^{-9} mutations/bp/year) is lower than those of the Tasmanian devil (1.27×10^{-9} mutations/bp/year) and Virginia opossum (2.35×10^{-9} mutations/bp/year), due to the shorter generations of these two other marsupials (Lewin and Eyre-Walker 2025). The yearly koala rate is comparable to those of eutherian mammals with similar generation times but higher than that of humans (0.43×10^{-9} mutations/bp/year; Lewin and Eyre-Walker 2025). Our estimate of the koala mutation rate is the first for Diprotodontia, the largest extant marsupial

order, providing a valuable resource for evolutionary and demographic analyses of this group.

We obtained similar estimates of mutation rates from genomes sequenced from museum skins and from fresh blood. This result demonstrates that older tissue samples can provide a useful genomic resource with strict site filtering, despite theoretical concerns (Bergeron et al. 2022). Although we sequenced to a lower depth than Bergeron et al. (2023), we found similar percentages of the genome as being callable after the same strict filtering. However, our study has also revealed potential risks in relying on studbook records for parental identifications. We also detected high relatedness between the parents in one trio, suggesting that their founding individuals were likely to be related. This led to a small number of sites for which the parents were homozygous for different alleles, meaning that this trio contributed minimally to the calculation of the false-negative rate.

Compared with previous estimates of mutation rates in marsupials, which only used a single trio each (Bergeron et al. 2023), our analysis of four replicate trios has enabled us to more accurately estimate the distribution of mutations across classes. The distribution across mutation classes for koalas was broadly similar to that observed in primates (Fig. 1*b*; Versoza et al. 2025), despite >120 My since the divergence between marsupial and eutherian mammals (Álvarez-Carretero et al. 2022) and mutation spectra diverging between some eutherian mammals (Beichman et al. 2023).

Our study provides an estimate of a key evolutionary parameter that can enhance genomic analyses of koalas and other marsupials. Further estimates of mutation rates from a broader range of marsupial species will allow us to test whether the drivers of mutation rates are consistent with those seen in eutherian mammals. Previous studies have reported relationships between the number of *de novo* mutations and both maternal and paternal age at conception (Jónsson et al. 2017), as well as a paternal bias in mutation contributions (Manuel

et al. 2022). Marsupials could differ in the rate at which mutations accumulate because they give birth after a short gestation period, with most development occurring in the pouch rather than *in utero* (Smith & Keyte 2020). Teasing apart how *de novo* mutations arise during this process could help to reveal the broader determinants of mutation rates. Among species, mutation rate variation is correlated with generation time (Bergeron et al. 2023; Zhu et al. 2025), effective population size (Lynch et al. 2016; Zhu et al. 2025), and genome size (Lynch et al. 2016). It is unclear whether these patterns hold in marsupials, but their contrasting developmental strategies make them an interesting test case for evolutionary theories of mutation rate variation in mammals.

Population declines coincide with glacial environmental shifts

Our results provide a compelling illustration of the impact of the mutation rate used to scale the timing of demographic inferences. Notably, our application of the koala mutation rate led to demographic plots that were drastically different from those scaled using the human mutation rate. Although our PSMC curves have the same shape as those previously inferred for koalas, we found that population declines occurred prior to the arrival of modern humans ~65–47 kya (O'Connell and Allen 2015; Clarkson et al. 2017), whereas previous studies suggested that these declines followed human arrival (Johnson et al. 2018; De Cahsan et al. 2025). There has been much debate over when humans first reached the continent (O'Connell et al. 2018; Bradshaw et al. 2023), but the arrival of humans into southern Australia, and the majority of the koala's range, most likely occurred towards the more recent end of this scale. While both our PSMC and SMC++ results show that the koala population started to decline prior to human arrival, our PSMC estimates suggest sharp bottlenecks while our SMC++ estimates suggest a more gradual decline that partially overlaps with the presence of humans. In addition, compared with previous studies (Johnson et al. 2018; De Cahsan et al. 2025), our

PSMC and SMC++ analyses both infer an earlier expansion of koala populations 700–200 kya, aligning with the oldest known fossil of the present-day koala 350 kya (Black et al. 2014; Fig. 2*a*, Fig. 3*a*).

We estimated that a bottleneck occurred in the ancestral koala population before it split into the present-day populations in the last 30 ky. These results are substantially different from previous reports of the populations diverging 250–200 kya (De Cahsan et al. 2025) or the ancestor diverging into three populations >300 kya (Lott et al. 2022). After the populations began to diverge, we estimated slight recoveries in the three northernmost populations but not in the two southern populations. This pattern is consistent with the present-day diversity in each of the populations, pointing to the influence of pre-European processes. However, it is well documented that southern populations experienced more severe recent bottlenecks due to a prolific fur trade around the turn of the 20th century (Phillips 1990; Menkhorst 2008), leading to decreased diversity and increased inbreeding (McLennan et al. 2024). The impact of these recent processes on SMC++ analyses is not well understood, and they could potentially be driving the observed patterns in historical sizes of the southern populations. This bias could also explain the reduced resolution (number of time intervals) in the analysis of the Victorian population, for which a slightly later bottleneck was inferred than for the other populations. Given that the Victorian population was not isolated until after the bottleneck, the inferred later bottleneck is unlikely to be biologically meaningful.

Population declines during the last 110 ky coincide with the onset of the most recent glacial period. During this period, koala ranges became increasingly restricted to habitat refugia along the eastern and southwest coasts (Adams-Hosking et al. 2011). This is likely to have seen the final separation of the eastern koala populations from the now-extinct western population through the emergence of the Nullarbor (meaning ‘no tree’) Plain, the treeless southern central region of Australia. During the mid-Pleistocene, the Nullarbor Plain

harboured a diverse collection of herbivores that included arboreal species such as koalas and tree kangaroos, suggesting a more diverse flora than at present (Prideaux et al. 2007). This would have provided a connection between eastern and western eucalypt forests, before the landscape transitioned to semi-arid shrubland by ~70 kya to form a vast biogeographical barrier between the western and eastern koala populations (Van Der Kaars et al. 2017; De Deckker et al. 2021; Adesanya et al. 2023). The appearance of this barrier would have led to a substantial reduction in the effective population size of koalas, matching the declines observed in our demographic models.

Koalas would have become increasingly restricted to the east and west coast as they approached the Last Glacial Maximum (Adams-Hosking et al. 2011). This is consistent with the declines inferred from SMC++, which suggest a continued decline until ~20 kya. Modern humans arrived on the Australian continent after the koala population had already started to decline, but humans have been credited with the extinction of the last remaining Western Australian population (Balme et al. 1978; Prideaux et al. 2010). Population expansions and divergences after the Last Glacial Maximum also coincide with the expansion of habitat across the east coast of Australia to the koala's present-day range of more than 3500 km (Adams-Hosking et al. 2011).

Historical population sizes estimated using the sequentially Markovian coalescent (e.g., PSMC and SMC++) rely on the assumption that a population's inverse instantaneous coalescent rate is directly correlated with its effective population size. However, a reduction in the inverse instantaneous coalescent rate can be caused not only by a reduction in population size, but also by increasing structure with reduced migration and gene flow (Mazet et al. 2016; Chikhi et al. 2018; Mather et al. 2020). Therefore, it is possible that the inferences from these methods indicate a decline in connectivity between koala populations during this period instead of reductions in size. However, decreased population size and

increased population structure would both support the notion that koala habitat became more restricted during the last glacial period. It is likely that these two processes occurred in tandem.

Population-specific koala recombination landscape

Our inferred recombination maps contained mean recombination rates ranging from 0.74×10^{-9} to 1.40×10^{-9} crossovers/bp/generation, an order of magnitude lower than the average rate in animals of 25.2×10^{-9} crossovers/bp/generation (Stapley et al. 2017). However, the estimated recombination rates in koalas are more consistent with the low rates inferred for other marsupials from linkage maps, such as 2.3×10^{-9} crossovers/bp/generation in the gray short-tailed opossum (*Monodelphis domestica*) (Samollow et al. 2007, Samollow et al. 2010) and 5.1×10^{-9} crossovers/bp/generation in the tammar wallaby (*Notamacropus eugenii*) (given a genome length of 3.3 Gb; Wang et al. 2011). Reduced recombination rates in marsupials have been explained by fewer double-stranded breaks during meiosis (Marín-Gual et al. 2022), but the evolutionary drivers remain unknown.

When recombination is reduced, the Hill-Robertson effect reduces the effectiveness of selection and the effective population size at linked sites, increasing genetic drift and lowering local levels of diversity (Charlesworth et al., 1993; Nordborg et al., 1996). Therefore, lower recombination rates are expected to further threaten genetically depleted species. We inferred differing recombination rates across the five koala populations, with the highest rate in the VIC population. This result suggests that although the VIC population has the lowest level of genetic diversity due to a recent bottleneck (McLennan et al. 2024), its high rate of genome-wide recombination means that it is likely to experience lower levels of linked selection. This could potentially explain how this population has successfully

increased in population size following its near extinction in the early 20th century (Whisson and Ashman 2020).

Population-specific variation in recombination rates has previously been found in a number of species. In humans and other animals, recombination has been found to be driven by variation in the *PRDM9* alleles which determine the genomic locations of recombination hotspots (Paigen and Petkov 2018). Other causes of population-specific recombination rates include variation in inversions, chromatin structure, and epigenetics (Stapley et al. 2017). Our recombination maps provide a useful resource for determining whether the same drivers of recombination rates occur in marsupials.

Producing a recombination map for a heavily managed threatened species could help to predict the success of mixing genetic diversity through translocation programs, one of the primary objectives of genomics-based conservation management. In sexually reproducing populations, meiotic recombination (along with the random assortment of chromosomes) creates new combinations of alleles in the offspring as homologous chromosomes cross over and exchange genetic material (Burbrink et al. 2025). Traditional conservation management strategies like translocations are most likely to create new combinations of genetic variants in high-recombining parts of the genome, where exchange occurs most frequently. However, if the goal of a translocation is to reduce or eliminate deleterious variation in a gene that resides in a low-recombining genomic region, it will be harder to remove these variants without also losing diversity at linked sites. Our recombination maps have highlighted how much the recombination context can vary between genes and how this pattern can vary across different populations. Our results suggest that translocations of koalas between populations are more likely to generate novel combinations of alleles in tissue-resident memory T (TRM) cell genes compared with MHC class III genes. With greater knowledge about the specific genes or variants conferring resistance and susceptibility to *Chlamydia* infection in koalas, these

resolved recombination maps will help conservation managers to predict whether translocation experiments are likely to affect genomic regions of interest.

Previous estimates of recombination maps using ReLERNN in mammal species for which estimates of mutation rates are unavailable (Bredemeyer et al. 2023; Foley et al. 2024) have used a previously inferred ‘pan-mammal’ rate (Kumar and Subramanian 2002). Although this approach enables robust comparison of the recombination landscape within a species, it does not allow direct comparison with other species. Our results confirm the original assessment that ReLERNN produces correlated recombination landscapes for differing mutation rates, but applying an overestimate of the mutation rate, as is the case when applying the ‘pan-mammal’ rate to koalas, results in overestimation of the recombination rates (Adrion et al. 2020). These outcomes highlight another benefit of producing species-specific mutation rates for a diverse range of organisms.

Although our analysis was based on a high-quality, long-read-based genome assembly, it also highlighted some insufficiencies in this genome assembly strategy for the study of the immune loci. Immune genes are innately difficult to assemble because they are highly heterozygous, often evolve through gene duplication, and contain complex structural rearrangements (Logsdon et al. 2025). Although long-read sequencing technologies have evolved to account for repetitive genomic regions, they still struggle to assemble highly heterozygous parts of diploid genome assemblies (Rice et al. 2020). This was reflected in the fact that some immune loci were assigned to scaffolds that were too small for their rate of recombination to be inferred with confidence. This was an issue for genes within the immunoglobulin kappa locus and genes located within the immunoglobulin heavy-chain locus (supplementary table S4). Our results highlight the need for complete telomere-to-telomere haploid genomic sequences, not only for humans, apes, and agricultural and

companion animals (Nurk et al. 2022; Bredemeyer et al. 2023; Rice et al. 2020), but also for species of conservation concern.

Conclusions

Genome-scale analyses of demographic and evolutionary history are allowing increasingly fine-grained inferences of complex historical processes. We have shown, however, that our ability to do so in non-model organisms, which often contain the most fascinating histories, is hindered by gaps in our knowledge of fundamental evolutionary parameters. With a growing catalogue of mutation rate estimates for organisms across the Tree of Life, our demographic modelling will become increasingly accurate and will allow us to resolve the drivers of historical population-size dynamics.

In the koala, our use of a direct estimate of the mutation rate led to a shift in the estimated timing of population-size changes, with important consequences for identifying the drivers of population declines. In particular, we found that late Pleistocene population contractions occurred prior to the arrival of humans on the continent. This shifts the narrative away from anthropogenic declines of koalas and highlights the substantial ecosystem changes that have occurred during the history of the continent. Broader analyses of the Australian fauna are required to determine whether this pattern is unique to koalas, or if ecosystem changes have driven similar declines in other species.

We have also demonstrated that using species-specific mutation rates allows more accurate estimation of recombination maps. Using these maps, we characterized the recombination context of koala immune genes, identifying genes that are more likely to be affected by linked selection and more likely to benefit from translocations. These results provide a useful context to help conservation managers tailor management strategies to a population’s needs.

Materials and Methods

Whole-genome resequencing of parent-offspring trios

Our data set comprised 12 resequenced genomes from five male and seven female koalas, from five putative parent-offspring trios from three families. Eight individuals forming three trios were sourced from skins at the Australian Museum and from a captive-bred colony at Featherdale Sydney Wildlife Park. Four live individuals forming two trios (two full siblings) were selected from the Taronga Conservation Society Australia's breeding program, with DNA extracted from blood. All genome sequences were produced in the first stage of the Koala Genome Survey (Hogg et al. 2023; McLennan et al. 2024). We downloaded bam files from the Amazon Web Services Open Data Sponsorship program (Australasian Genomes; https://koalagenomes.s3.ap-southeast-2.amazonaws.com/index.html). Details of DNA extraction, sequencing, and read mapping were reported previously (Hogg et al. 2023; McLennan et al. 2024). The average genome coverage across koala samples forming trios was 31×. *De novo* mutations, false-positive rates, and false-negative rates were calculated using the methods described by Bergeron et al. (2022), with the additional false-positive steps introduced by Bergeron et al. (2023).

The Koala Genome Survey variant call format (VCF) file was updated to include all available koala genomes, including captive-bred individuals, as well as wild representatives from South Australian populations on Kangaroo Island (Gates et al. 2025) and the Mount Lofty Ranges. We used individual genomic variant call format (GVCF) files for each sample as input to Illumina's DRAGEN Germline Pipeline (v4.3.6; Illumina), using the koala reference genome (GCA_002099425.1_phaCin_unsw_v4.1; Johnson et al. 2018) to perform joint genotyping across all samples. This uses the DRAGEN gVCF PopGen Genotyper (4.3.6;

Illumina) to produce a single, multi-sample variant calling file (VCF) used for all downstream analyses. Depth (DP) was calculated for all sites in all samples using bcftools (v1.17), by summing the localized allelic depths (LAD). Our unfiltered VCF contained 59,694,393 SNPs. Variants were filtered using bcftools to remove sites that did not pass Illumina's QUAL filter and had read depth <15 or >60, leaving 47,655,085 SNPs. Finally, we filtered out variants that were missing in more than one individual, leaving 17,573,439 SNPs. Custom scripts based on those developed in Bergeron et al. (2022) and Bergeron et al. (2023) for the estimation of mutation rates are available at https://github.com/tobykovacs796/koalamutrate.

Phylogenetic network

The filtered VCF was read into a genlight object in R using *read.PLINK*. Euclidean distances were calculated using *bitwise.dist* in the package poppr, with the "scale_missing" setting. Using these genetic distances, we inferred a phylogenetic network using SplitsTree 6 (Huson and Bryant 2024). Individuals were then assigned to populations based on their maximum ancestry in previous *fast*STRUCTURE analyses (McLennan et al. 2024). Custom scripts for demographic analyses are available at https://github.com/tobykovacs796/koalamutrate.

Demographic inference

Historical population sizes were estimated using a suite of methods that use the sequentially Markovian coalescent (SMC) as well as the site frequency spectrum. SMC methods are generally most accurate between 1000 and 100,000 generations ago (Patton et al. 2019), with PSMC being more accurate for deeper timeframes in this range, and newer methods, such as SMC++, being more accurate for more recent timeframes due to the consideration of the site

frequency spectrum (Terhorst et al. 2017). In order to demonstrate the impact of the mutation rate on demographic inference, we ran analyses using PSMC and SMC++.

We analysed all 458 genomes from wild koalas using PSMC (Li and Durbin 2011). We trimmed the koala reference genome (GCA_002099425.1_phaCin_unsw_v4.1; Johnson et al. 2018) using vcftools (Danecek et al. 2011) to remove all contigs <50 kb and those putatively corresponding to the X chromosome (Hogg et al. 2024). We generated the input file for PSMC (Li and Durbin 2011) by calling variants for each individual using the bam file and trimmed reference genome using samtools mpileup (v1.6) and bcftools (v1.3.1), following the instructions given for PSMC (Li and Durbin 2011). PSMC was run using the time intervals “4+25×2+4+6”, assuming a generation length of 7 years (Phillips 2000), and with our estimate of the per-generation mutation rate. Individual demographic plots were grouped according to the previously assigned genetic clusters of koalas (McLennan et al. 2024). Because some of our initial PSMC plots displayed population spikes in single recent time intervals, we tried splitting the earlier time intervals as done by Hilgers et al. (2025) (2+2+25×2+4+6 and 1+1+1+1+25×2+4+6). This removed the spikes in some plots but added them to other plots where they did not originally occur. Hence, the final plot was selected to minimize these extreme, single-interval peaks, which are a known artefact of these methods.

Historical population sizes were also estimated for each of the five genetic populations using SMC++. Koalas were assigned to a population based on their highest ancestry as identified previously (McLennan et al. 2024). Koalas from Narrandera were not included in any population for SMC++ analyses because they are the result of recent translocations from SE QLD / N NSW and VIC populations (McLennan et al. 2024). Captively bred koalas from Dubbo, which were not included in previous analyses of population structure (McLennan et al. 2024), were placed across M NSW and S NSW based on our phylogenetic network. Because these koalas are likely to contain recently mixed ancestry, they were excluded from

SMC++ analyses. Koalas from South Australia, which were nested within the Victorian population in our phylogenetic network, originated from serial bottlenecking and inbreeding and so were not included in the analysis of Victorian koalas using SMC++.

For our demographic analyses using SMC++ (Terhorst et al. 2017), we used vcftools (Danecek et al. 2011) to filter the VCF to keep only biallelic sites and remove all contigs <50 kb and those that have been putatively identified as forming the X chromosome (Hogg et al. 2024). Historical population sizes were inferred in SMC++ using *estimate*, calculating the composite likelihood across several distinguished lineages. Distinguished lineages were selected based on the largest group of unrelated individuals that could be found in each population, with up to 10 distinguished lineages. We ran SMC++ with default settings with our estimated mutation rate, across the time interval 10^3–10^6 years. Initial analyses using the default number of knots (8) showed limited resolution over key timeframes of interest, so analyses were rerun with the number of knots set to 16, 32, and 50. These results showed increasing resolution with the number of knots (supplementary fig. S3), by introducing more inflection points and inferring different effective population sizes for a larger number of shorter time intervals. Several runs with 32 or 50 knots failed shortly after starting, with the error: *erroneous average coalescent time*. These runs were restarted until they completed successfully. The timing of population separation was inferred in SMC++ (Terhorst et al. 2017) using the function *split* and the marginal estimates for each population when using 50 knots. Custom scripts for demographic analyses are available at https://github.com/tobykovacs796/koalamutrate.

Recombination maps

A subset of unrelated individuals from each of the five populations were selected to infer population-specific recombination landscapes. Relatedness (kinship) between individuals was estimated within each of the five populations identified in McLennan et al. (2024) using vcftools *--relatedness2*, which implements the KING method (Manichaikul et al. 2010). We wrote a script to randomly select individuals and iteratively build a group of unrelated individuals containing five females and five males. The algorithm began by randomly choosing one individual, then repeatedly searching the population for an unrelated individual, adding each as it was found, until either a group of ten was assembled or no additional unrelated individuals could be added. When multiple valid sets of ten were possible, we prioritized individuals to maximize genome coverage and ancestry representation. Variants for these individuals were subsetted from the unfiltered multi-sample VCF. ReLERNN requires that SNPs residing in transposable elements be masked before the analysis. As such, we identified the genome-wide transposable element content of the reference genome using RepeatMasker (version Open 3.2.6 A.F.A. Smit, R. Hubley & P. Green RepeatMasker at http://repeatmasker.org). This program was run using the -qq parameter and using the Zoonomia repeat library (Osmanski et al. 2023). To prepare variants for generating a recombination map, we removed variants overlapping a transposable element annotation in the reference genome. We further filtered variants, removing variants within 5 bp of an indel and those that did not meet the following quality criteria:e'%QUAL<30 | INFO/DP<16 | INFO/DP>62 | QD<2 | FS>60 | SOR>10 | ReadPosRankSum <-8 | MQRankSum <-12.5 | MQ<40' in bcftools (https://github.com/samtools/bcftools). Using VCFtools (https://vcftools.github.io/man_latest.html), we removed indels and subset the data to include only biallelic SNPs for each clade for further analysis. Custom scripts used to contextualize and average the raw ReLERNN output are available at https://github.com/eutherialab/Foley_XLRD.

To model the genome-wide recombination rate for each koala population, we used ReLERNN (https://github.com/kr-colab/ReLERNN), a deep learning approach that uses recurrent neural networks (Adrion, Galloway, and Kern 2020). Our estimate of the mutation rate was incorporated into the analysis. ReLERNN was run using the simulate, train, predict, and bscorrect modules with default settings. Given that we were interested in broadly evaluating the differences in recombination rates among populations, we averaged rates in 2 Mb blocks with a 50 kb step. A non-parametric Mann-Whitney test was used to determine if there was a significant difference between the populations with the highest and lowest inferred recombination rates.

We then investigated whether there was a difference between recombination rates inferred using an empirically estimated mutation rate for koalas compared with a 'pan-mammal' yearly mutation rate of 2.2×10^{9}, estimated in a previous analysis of branch lengths (Kumar and Subramanian 2002). This was multiplied by 7 to produce a per-generation mutation rate for koalas to allow direct comparison with our empirically estimated mutation rate. We reran ReLERNN for a subset of populations using the pan-mammal mutation rate. Tests for normality were conducted using qq-norm in R. Spearman's rank correlation test was used to evaluate differences in the inferred recombination landscapes. A Mann-Whitney test was used to determine if there were differences in the mean of the rates between recombination maps.

Recombination context of koala immune genes

Per-population, raw recombination rates were ordered from lowest to highest and assigned to deciles, where the 1st and 10th deciles represent the lowest and highest recombination rates, respectively. Previous analyses have identified and annotated immune gene clusters in the

koala genome (R. N. Johnson et al. 2018). These included comprehensive annotation of the major histocompatibility complex, T-cell receptors, immunoglobulins, cytokines (R. N. Johnson et al. 2018), additional MHC genes (Silver et al. 2022; Silver et al. 2024), toll-like receptors (Cian et al. 2024), and natural killer cells (Morris et al. 2015). For each gene, we list the recombination decile that corresponds to those genomic coordinates for each of the five koala populations.

Data availability

Raw whole genome re-sequences are available in The National Center for Biotechnology Information (NCBI) under BioProject PRJNA940526 at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA940526. All outputs including the position of *de novo* mutations and relatedness values are available at https://github.com/tobykovacs796/koalamutrate and the supplementary material.

Acknowledgements

We thank Taronga Zoo for taking blood samples from their captive koalas (Animal Ethics Committee approval 4a/06/22) and the Australian Museum and Featherdale Sydney Wildlife Park for providing skins, and all institutions for access to the studbook. We thank Eilish McMaster for help with data visualization, and Zhiliang Chen and the Illumina team for assistance with DRAGEN.

Funding

Genome sequencing for the genome survey koalas and captive koalas was supported by the New South Wales Government and the Australian Government's Bushfire Recovery for

Wildlife and their Habitats program (GA-2000526); further support was provided by The University of Sydney, Amazon Web Services Open Data Sets, Ramaciotti Centre for Genomics, and Illumina. Sequencing for the South Australian koalas was supported by the Australian Research Council (LP210100450). L.W.S. is supported by a fellowship from the Australian Research Council (CE200100012). T.G.L.K. acknowledges the support provided by the Australian Government's Research Training Program award. N.M.F and W.J.M are supported by a National Science Foundation Grant (DEB-2150664).

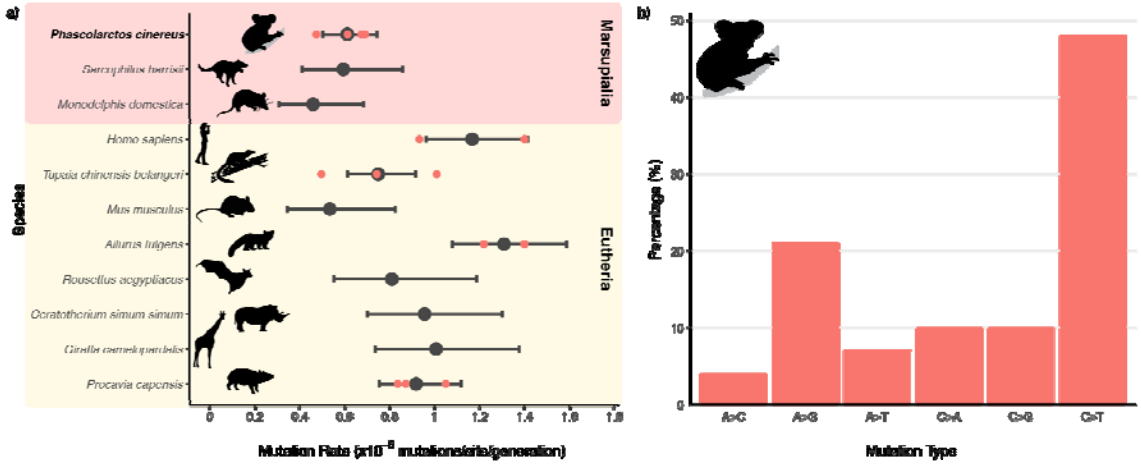

FIG. 1. Germline mutation rate of the koala and other mammals. (*a*) Germline mutation rates inferred in the koala (*Phascolarctos cinereus*) based on four parent-offspring trios. Individual estimates for each trio are shown in orange. The average mutation rate is shown in black and 95% confidence intervals are indicated with error bars. Other mutation rate estimates are from Bergeron et al. (2023) and can be found in supplementary table S2. Silhouettes were sourced from Phylopic (http://phylopic.org). (*b*) Mutation spectrum for koala genomes. The percentage of each mutation type is shown after collapsing reverse complements.

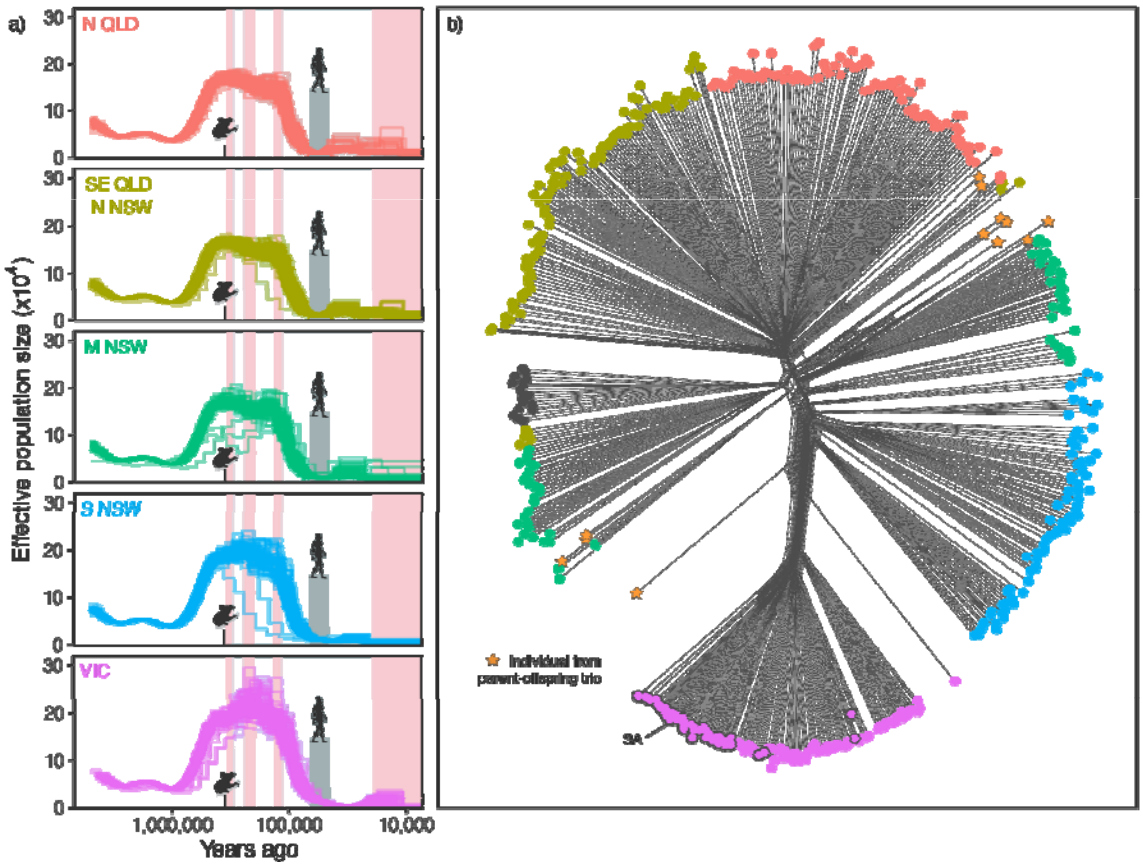

FIG. 2. (*a*) Historical effective population sizes estimated using PSMC for 458 individuals across the five genetic populations. Backgrounds indicate the interglacial (warmer and wetter; pink) and glacial (cooler and drier; blue) periods of the last three glacial cycles. The dark grey column and human silhouette indicate the period in which humans are believed to have arrived and spread across Australia (65–47 kya; O'Connell & Allen 2015; Clarkson et al. 2017). The koala silhouette indicates the earliest known fossil of the modern koala species (Black et al. 2014). (*b*) Phylogenetic network of all 474 koala samples based on Euclidean genetic distances. Colours correspond to the five genetic clusters shown in (*a*). Individuals that form parent-offspring trios used for mutation-rate estimation are indicated by yellow stars. Koalas from Narrandera, which represent a mix of two populations, are shown in dark grey. Grey outlines highlight koalas from South Australia (SA), which are nested within the Victorian population.

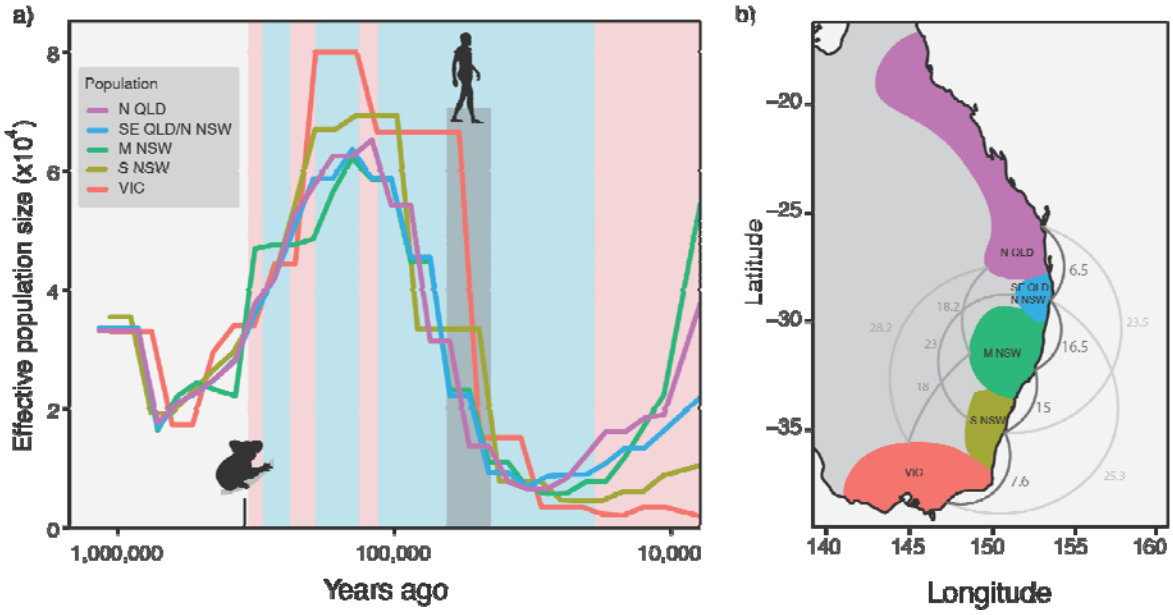

FIG. 3. (*a*) Historical population sizes estimated using SMC++ for each of the five genetic groups of koalas. Pink and blue backgrounds indicate the interglacial (warmer and wetter) and glacial (cooler and drier) periods of the last three glacial cycles. The dark grey column and human silhouette indicate the period in which humans are believed to have arrived and spread across Australia. The koala silhouette indicates the earliest known fossil of the modern koala species (Black et al. 2014). (*b*) Population split times (kya) estimated using SMC++ for all pairs of populations. The split times are darker and larger between more closely located populations.

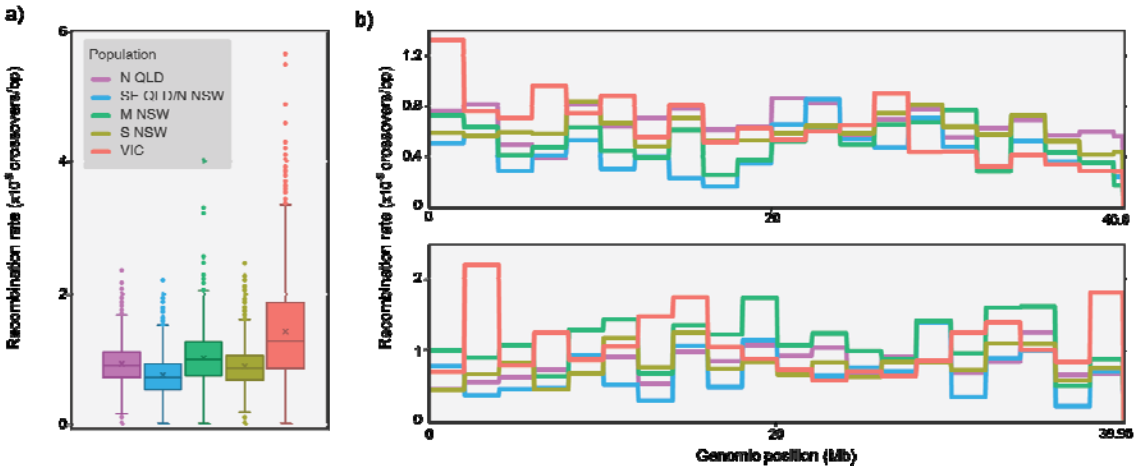

FIG. 4. Variation in recombination rates across koala populations. (*a*) Boxplots showing genome-wide recombination rates. In each box, the × and the horizontal line represent the mean and the median recombination rate, respectively. Whiskers extend to the most extreme data points within 1.5 times the interquartile range from the first and third quartiles. Points beyond this are plotted. (*b*) Variation in recombination rates among populations across the two largest scaffolds.

References

Adams-Hosking C, McBride MF, Baxter G, Burgman M, de Villiers D, Kavanagh R, Lawler I, Lunney D, Melzer A, Menkhorst P, et al. 2016. Use of expert knowledge to elicit population trends for the koala (*Phascolarctos cinereus*). *Divers Distrib* 22:249–262.

Adams-Hosking C, Moss PT, Rhodes JR, Grantham HS, McAlpine C. 2011. Modelling the potential range of the koala at the Last Glacial Maximum: Future conservation implications. *Aust Zool* 35:983–990.

Adesanya M, Charles S, Gallagher R, Kaars S Van Der, Deckker P De, Hua Q, Graeme S. 2023. On the timing of megafaunal extinction and associated floristic consequences in Australia through the lens of functional palaeoecology. *Quat Sci Rev* 316:108263.

Adrion JR, Galloway JG, Kern AD. 2020. Predicting the landscape of recombination using deep learning. *Mol Biol Evol* 37:1790–1808.

Álvarez-Carretero S, Tamuri AU, Battini M, Nascimento FF, Carlisle E, Asher RJ, Yang Z, Donoghue PCJ, dos Reis M. 2022. A species-level timeline of mammal evolution integrating phylogenomic data. *Nature* 602:263–267.

Balme J, Merrilees D, Porter JK. 1978. Late Quaternary mammal remains, spanning about 30,000 years, from excavations in Devil's Lair, Western Australia. *J R Soc West Aust* 61:33–65.

Beichman AC, Robinson J, Lin M, Moreno-Estrada A, Nigenda-Morales S, Harris K. 2023. Evolution of the mutation spectrum across a mammalian phylogeny. *Mol Biol Evol* 40:msad213.

Bergeron LA, Besenbacher S, Turner TN, Versoza CJ, Wang RJ, Price AL, Armstrong E, Riera M, Carlson J, Chen HY, et al. 2022. The Mutationathon highlights the importance of reaching standardization in estimates of pedigree-based germline mutation rates. *elife* 11:e73577.

Bergeron LA, Besenbacher S, Zhang J, Li P, Zheng J, Li P, Bertelsen MF, Quintard B, Hoffman JI, Li Z, et al. 2023. Evolution of the germline mutation rate across vertebrates. *Nature* 615:285–291.

Black KH. 2016. Middle Miocene origins for tough-browse dietary specialisations in the koala (Marsupialia, Phascolarctidae) evolutionary tree: Description of a new genus and species from the Riversleigh World Heritage area. *Mem Mus Vic* 74:255–262.

Black KH, Price GJ, Archer M, Hand SJ. 2014. Bearing up well? Understanding the past, present and future of Australia's koalas. *Gondwana Res* 25:1186–1201.

Bradshaw CJA, Crabtree SA, White DA, Ulm S, Bird MI, Williams AN, Saltré F. 2023. Directionally supervised cellular automaton for the initial peopling of Sahul. *Quat Sci Rev* 303.

Bredemeyer KR, Hillier LD, Harris AJ, Hughes GM, Foley NM, Lawless C, Carroll RA, Storer JM, Batzer MA, Rice ES, et al. 2023. Single-haplotype comparative genomics provides insights into lineage-specific structural variation during cat evolution. *Nat Genet* 55:1953–1963.

Burbrink FT, DeBaun D, Foley NM, Murphy WJ. 2025. Recombination-aware phylogenomics. *Trends Ecol Evol* In press.

Byrne M. 2008. Evidence for multiple refugia at different time scales during Pleistocene climatic oscillations in southern Australia inferred from phylogeography. *Quat Sci Rev* 27:2576–2585.

Byrne M, Steane DA, Joseph L, Yeates DK, Jordan GJ, Crayn D, Aplin K, Cantrill DJ, Cook LG, Crisp MD, et al. 2011. Decline of a biome: Evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. *J Biogeogr* 38:1635–1656.

Byrne M, Yeates DK, Joseph L, Kearney M, Bowler J, Williams MAJ, Cooper S, Donnellan SC, Keogh JS, Leys R, et al. 2008. Birth of a biome: Insights into the assembly and maintenance of the Australian arid zone biota. *Mol Ecol* 17:4398–4417.

Charlesworth B. 2009. Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* 10:195–205.

De Cahsan B, Velasco MS, Westbury M V., Duchêne DA, Sinding MHS, Morales HE, Kalthoff DC, Barnes I, Brace S, Miguez RP, et al. 2025. Road to extinction? Past and present population structure and genomic diversity in the koala. *Mol Biol Evol* 42:1–19.

Chikhi L, Rodríguez W, Grusea S, Santos P, Boitard S, Mazet O. 2018. The IICR (inverse instantaneous coalescence rate) as a summary of genomic diversity: Insights into demographic inference and model choice. *Heredity* 120:13–24.

Clarkson C, Jacobs Z, Marwick B, Fullagar R, Wallis L, Smith M, Roberts RG, Hayes E, Lowe K, Carah X, et al. 2017. Human occupation of northern Australia by 65,000 years ago. *Nature* 547:306–310.

Crichton AI, Beck RMD, Couzens AMC, Worthy TH, Camens AB, Prideaux GJ. 2023. A probable koala from the Oligocene of central Australia provides insights into early diprotodontian evolution. *Sci Rep* 13:14521.

Cristescu RH, Gardiner R, Terraube J, McDonald K, Powell D, Levengood AL, Frère CH. 2023. Difficulties of assessing the impacts of the 2019–2020 bushfires on koalas. *Austral Ecol* 48:12–18.

Cui J, Batley KC, Silver LW, McLennan EA, Hogg CJ, Belov K. 2025. Spatial variation in toll-like receptor diversity in koala populations across their geographic distribution. *Immunogenetics* 77.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27:2156–2158.

De Deckker P, van der Kaars S, Haberle S, Hua Q, Stuut JBW. 2021. The pollen record from marine core MD03-2607 from offshore Kangaroo Island spanning the last 125 ka; implications for vegetation changes across the Murray-Darling Basin. *Aust J Earth Sci* 68:928–951.

Duchêne DA, Bragg JG, Duchêne S, Neaves LE, Potter S, Moritz C, Johnson RN, Ho SYW, Eldridge MDB. 2018. Analysis of phylogenomic tree space resolves relationships among marsupial families. *Syst Biol* 67:400–412.

Dumont BL, Payseur BA. 2008. Evolution of the genomic rate of recombination in mammals. *Evolution* 62:276–294.

Foley NM, Harris AJ, Bredemeyer KR, Ruedi M, Puechmaille SJ, Teeling EC, Criscitiello MF, Murphy WJ. 2024. Karyotypic stasis and swarming influenced the evolution of viral tolerance in a species-rich bat radiation. *Cell Genomics* 4:100482.

Gates K, Sandoval-castillo J, Beaman J, Burke K, Silva D, Saltre F, Belov K, Hogg C, Bradshaw C. 2025. Genomic erosion and inbreeding in an abundant island population of koalas. *Preprint*:www.authorea.com/doi/full/10.22541/au.174119408.80.

Hilgers L, Liu S, Jensen A, Brown T, Cousins T, Schweiger R, Guschanski K, Hiller M. 2025. Avoidable false PSMC population size peaks occur across numerous studies. *Curr Biol* 35:927–930.

Ho SYW, Lanfear R, Bromham L, Phillips MJ, Soubrier J, Rodrigo AG, Cooper A. 2011. Time-dependent rates of molecular evolution. *Mol Ecol* 20:3087–3101.

Hogg CJ, Edwards RJ, Farquharson KA, Silver LW, Brandies P, Peel E, Escalona M, Jaya FR, Thavornkanlapachai R, Batley K, et al. 2024. Extant and extinct bilby genomes

combined with Indigenous knowledge improve conservation of a unique Australian marsupial. *Nat Ecol Evol* 8:1311–1326.

Hogg CJ, Silver L, McLennan EA, Belov K. 2023. Koala Genome Survey: an open data resource to improve conservation planning. *Genes* 14:546.

Huson DH, Bryant D. 2024. The SplitsTree App: interactive analysis and visualization using phylogenetic trees and networks. *Nat Methods* 21:1773–1774.

Johnson CN. 2016. Fire, people and ecosystem change in Pleistocene Australia. *Aust J Bot* 64:643–651.

Johnson RN, O'Meally D, Chen Z, Etherington GJ, Ho SYW, Nash WJ, Grueber CE, Cheng Y, Whittington CM, Dennison S, et al. 2018. Adaptation and conservation insights from the koala genome. *Nat Genet* 50:1102–1111.

Jónsson H, Sulem P, Kehr B, Kristmundsdottir S, Zink F, Hjartarson E, Hardarson MT, Hjorleifsson KE, Eggertsson HP, Gudjonsson SA, et al. 2017. Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature* 549:519–522.

Van Der Kaars S, Miller GH, Turney CSM, Cook EJ, Nürnberg D, Schönfeld J, Kershaw AP, Lehman SJ. 2017. Humans rather than climate the primary cause of Pleistocene megafaunal extinction in Australia. *Nat Commun* 8:1–7.

Kumar S, Subramanian S. 2002. Mutation rates in mammalian genomes. *Proc Natl Acad Sci U S A* 99:803–808.

Lewin L, Eyre-Walker A. 2025. Estimates of the mutation rate per year can explain why the molecular clock depends on generation time. *Mol Biol Evol* 42:msaf069.

Li H, Durbin R. 2011. Inference of human population history from individual whole-genome sequences. *Nature* 475:493–496.

Logsdon GA, Ebert P, Audano PA, Loftus M, Porubsky D, Ebler J, Yilmaz F, Hallast P, Prodanov T, Yoo DA, et al. 2025. Complex genetic variation in nearly complete human genomes. *Nature* 644:430–441.

Lott MJ, Wright BR, Neaves LE, Frankham GJ, Dennison S, Eldridge MDB, Potter S, Alquezar-Planas DE, Hogg CJ, Belov K, et al. 2022. Future proofing the koala: synergising genomic and environmental data for effective species management. *Mol Ecol* 31:3035–3055.

Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, Foster PL. 2016. Genetic drift, selection and the evolution of the mutation rate. *Nat Rev Genet* 17:704–714.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics* 26:2867–2873.

De Manuel M, Wu FL, Przeworski M. 2022. A paternal bias in germline mutation is widespread in amniotes and can arise independently of cell division numbers. *Elife* 11:1–26.

Marín-Gual L, González-Rodelas L, Pujol G, Vara C, Martín-Ruiz M, Berríos S, Fernández-Donoso R, Pask A, Renfree MB, Page J, et al. 2022. Strategies for meiotic sex chromosome dynamics and telomeric elongation in Marsupials. *PLOS Genet* 18:1–26.

Mather N, Traves SM, Ho SYW. 2020. A practical introduction to sequentially Markovian coalescent methods for estimating demographic history from genomic data. *Ecol Evol* 10:579–589.

Mazet O, Rodríguez W, Grusea S, Boitard S, Chikhi L. 2016. On the importance of being structured: Instantaneous coalescence rates and human evolution-lessons for ancestral population size inference? *Heredity* 116:362–371.

McLennan EA, Kovacs TGL, Silver LW, Chen Z, Jaya FR, Ho SYW, Hogg CJ. 2024. Genomics identifies koala populations at risk across eastern Australia. *Ecol Appl* 35:e3062.

Menkhorst P. 2008. Hunted, marooned, re-introduced, contracepted: a history of Koala management in Victoria. In: Lunney D, Munn A, Meikle W, editors. Too close for comfort: contentious issues in human-wildlife encounters. Mosman: Royal Zoological Society of New South Wales. p. 73–92.

Morris KM, Mathew M, Waugh C, Ujvari B, Timms P, Polkinghorne A, Belov K. 2015. Identification, characterisation and expression analysis of natural killer receptor genes in *Chlamydia pecorum* infected koalas (*Phascolarctos cinereus*). *BMC Genomics* 16:1–11.

Neaves LE, Frankham GJ, Dennison S, FitzGibbon S, Flannagan C, Gillett A, Hynes E, Handasyde K, Helgen KM, Tsangaras K, et al. 2016. Phylogeography of the Koala, (*Phascolarctos cinereus*), and harmonising data to inform conservation. *PLOS One* 11:e0162207.

Nordborg M, Charlesworth B, Charlesworth D. 1996. The effect of recombination on background selection. *Genet Res* 67:159–174.

Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze A V, Mikheenko A, Vollger MR, Altemose N, Uralsky L, Gershman A, et al. 2022. The complete sequence of a human genome. *Science* 376:44–53.

O'Connell JF, Allen J. 2015. The process, biotic impact, and global implications of the human colonization of Sahul about 47,000 years ago. *J Archaeol Sci* 56:73–84.

O'Connell JF, Allen J, Williams MAJ, Williams AN, Turney CSM, Spooner NA, Kamminga J, Brown G, Cooper A. 2018. When did homo sapiens first reach Southeast Asia and Sahul? *Proc Natl Acad Sci U S A* 115:8482–8490.

Osmanski AB, Korstian J, Grimshaw JR, Halsey M, Moreno-Santillán DD, Crookshanks C, Roberts J, Garcia C, Johnson MG, Densmore LD, et al. 2023. Insights into mammalian TE diversity through the curation of 248 genome assemblies. *Science* 380:eabn1430.

Paigen K, Petkov PM. 2018. PRDM9 and Its Role in Genetic Recombination. *Trends in Genetics* 34:291–300.

Patton AH, Margres MJ, Stahlke AR, Hendricks S, Lewallen K, Hamede RK, Ruiz-Aravena M, Ryder O, McCallum HI, Jones ME, et al. 2019. Contemporary demographic reconstruction methods are robust to genome assembly quality: A case study in Tasmanian devils. *Mol Biol Evol* 36:2906–2921.

Pfeifer SP. 2020. Spontaneous Mutation Rates. In: Ho SYW, editor. The Molecular Evolutionary Clock. Cham, Switzerland: Springer Nature. p. 35–44.

Phillips B. 1990. Koalas. The little Australians we'd all hate to lose. Canberra: Australian Government Publishing Service

Prideaux GJ, Gully GA, Couzens AMC, Ayliffe LK, Jankowski NR, Jacobs Z, Roberts RG, Hellstrom JC, Gagan MK, Hatcher LM. 2010. Timing and dynamics of Late Pleistocene mammal extinctions in southwestern Australia. *Proc Natl Acad Sci U S A* 107:22157–22162.

Prideaux GJ, Long JA, Ayliffe LK, Hellstrom JC, Pillans B, Boles WE, Hutchinson MN, Roberts RG, Cupper ML, Arnold LJ, et al. 2007. An arid-adapted middle Pleistocene vertebrate fauna from south-central Australia. *Nature* 445:422–425.

Rice ES, Koren S, Rhie A, Heaton MP, Kalbfleisch TS, Hardy T, Hackett PH, Bickhart DM, Rosen BD, Ley B Vander, et al. 2020. Continuous chromosome-scale haplotypes assembled from a single interspecies F1 hybrid of yak and cattle. *Gigascience* 9:1–9.

Terhorst J, Kamm JA, Song YS. 2017. Robust and scalable inference of population history from hundreds of unphased whole-genomes. *Nat Genet* 176:139–148.

Tsangaras K, Ávila-Arcos MC, Ishida Y, Helgen KM, Roca AL, Greenwood AD. 2012. Historically low mitochondrial DNA diversity in koalas (*Phascolarctos cinereus*). *BMC Genet* 13:1–11.

Samollow PB. 2010. Marsupial Linkage Maps. In: Deakin JE, Waters PD, Graves JAM, editors. Marsupial Genetics and Genomics. Dordrecht, The Netherlands; Heidelberg, Germany; London, UK; New York, NY, USA: Springer. p. 75–100.

Samollow PB, Gouin N, Miethke P, Mahaney SM, Kenney M, VandeBerg JL, Graves JAM, Kammerer CM. 2007. A microsatellite-based, physically anchored linkage map for the gray, short-tailed Opossum (*Monodelphis domestica*). *Chromosome Research* 15:269–282.

Silver LW, Cheng Y, Quigley BL, Robbins A, Timms P, Hogg CJ, Belov K. 2022. A targeted approach to investigating immune genes of an iconic Australian marsupial. *Mol Ecol* 31:3286–3303.

Silver LW, McLennan EA, Beaman J, da Silva KB, Timms P, Hogg CJ, Belov K. 2024. Using bioinformatics to investigate functional diversity: a case study of MHC diversity in koalas. *Immunogenetics* 76:381–395.

Smith KK, Keyte AL. 2020. Adaptations of the Marsupial Newborn: Birth as an Extreme Environment. Anatomical Record 303:235–249.

Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017. Variation in recombination frequency and distribution across eukaryotes: Patterns and processes. *Phil. Trans. R. Soc. B* 372:20160455.

Versoza CJ, Ehmke EE, Jensen JD, Pfeifer SP. 2025. Characterizing the rates and patterns of de novo germline mutations in the aye-aye (*Daubentonia madagascariensis*). *Mol Biol Evol* 42: msaf034.

Wang C, Webley L, Wei KJ, Wakefield MJ, Patel HR, Deakin JE, Alsop A, Marshall Graves JA, Cooper DW, Nicholas FW, et al. 2011. A second-generation anchored genetic linkage map of the tammar wallaby (*Macropus eugenii*). *BMC Genet* 12.

Wang Y, Obbard DJ. 2023. Experimental estimates of germline mutation rate in eukaryotes: a phylogenetic meta-analysis. *Evol Lett* 7:216–226.

Whisson DA, Ashman KR. 2020. When an iconic native animal is overabundant: The koala in southern Australia. *Conserv Sci Pract* 2:e188.

White ME. 1994. After the Greening: The Browning of Australia. Kenthurst, NSW: Kangaroo Press

Yoder AD, Tiley GP. 2021. The challenge and promise of estimating the de novo mutation rate from whole-genome comparisons among closely related individuals. *Mol Ecol* 30:6087–6100.

Zhu L, Beichman A, Harris K. 2025. Population size interacts with reproductive longevity to shape the germline mutation rate. *Proc Natl Acad Sci U S A* 122:e2423311122.