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], 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

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 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^{-9} mutations per base pair per generation (95% confidence interval: $5.03–7.45\times10^{-9}$). 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. This decline occurred in the ancestral koala population, which subsequently split into multiple populations 6–30 kya. These findings provide critical insights into the evolutionary history of koalas, while highlighting the impact of using species-specific evolutionary rates in demographic inference in non-model species. Our estimates of the genome-wide mutation rate 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). More recently, the Plio-Pleistocene was dominated by intense glacial and interglacial climate cycles, which drove further shifts towards arid and fire-prone environments (White 1994). These changes have underpinned the evolution of Australian fauna, with adaptation to arid environments, the restriction of some taxa to remnant mesic habitats, and extinctions of various species. More recently, the arrival of modern humans led to widespread habitat modification, predation, and introduction of non-native animals and plants.

Marsupials are among the most recognisable 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 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). Most of these were extinct by the end of the Pleistocene, likely due to the opening of forests and the expansion of grasslands across most of Australia (Black et al. 2014). The modern koala species appeared at least 350 kya and is the only species to persist into the Holocene (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 extant range during the current interglacial period (Adams-Hosking et al. 2011). Contemporary koala populations inhabit an enormous range (>3500 km) but have experienced recent population 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 low 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 mostly remained 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 65–47 kya (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 has the potential to provide a clearer picture of the patterns and drivers of historical population sizes.

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 true 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 in Copenhagen Zoo 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.

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. 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), using the false-negative rate (0.068) and the size of the callable genome (average 68.6% total genome size). 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).

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 recognised in a previous study: 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) (McLennan et al. 2024). 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). 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). 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. S3). 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. S4). 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 stabilising. 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. All of the populations were estimated to have diverged in the last 30 ky, following a general isolation-by-distance pattern (fig. 3*B*). 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 each other at 6.5 and 7.6 kya, respectively.

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} muts/bp/gen) and opossum (*Didelphis virginiana*; 4.60×10^{-9} muts/bp/gen) but considerably lower than that inferred for modern humans (11.6×10^{-9} muts/bp/gen) (Bergeron et al. 2023). The koala yearly mutation rate (0.87×10^{-9} mutations/bp/year) is lower than that of the Tasmanian devil (1.27×10^{-9} muts/bp/y) and opossum (2.35×10^{-9} muts/bp/y) due to their shorter generation times (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} muts/bp/y) (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, confirming 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 had 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; Versoza et al. 2025), despite >120 My since the divergence of 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 has the potential to enhance genomic analyses of koalas and potentially other marsupials. Further estimates of mutation rates across marsupials, including in unsampled families, will allow us to test whether the drivers of mutation rates within species (e.g., paternal and maternal age) and among species (e.g., generation time and genome size) are the same in marsupials as they are in eutherian 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 arrived on 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 support the commencement of declines in the koala population prior to human arrival, our PSMC estimates suggest sharp bottlenecks while our SMC++ estimates suggest a more gradual decline that partially overlaps with the arrival 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, Fig. 3).

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 population 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, suggesting that pre-European processes could have contributed to this contemporary pattern. However, it is well documented that southern populations experienced more severe recent population 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 reduced diversity and recent inbreeding on SMC++ analyses is not well understood, and these contemporary processes could potentially be driving these observed patterns in historical population sizes for 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, it is unlikely that the inferred later bottleneck is 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 isolated 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 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 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, aligning with the declines observed in our demographic models. After this time, 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 3,500 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 an increasingly structured population 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 reduction in connectivity between koala populations during this period instead of population declines. However, decreased population size and increased population structure would both support the notion that koala habitat was increasingly restricted during the last glacial period. It is likely that these two processes occurred in tandem.

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. As mutation rate estimates continue to proliferate 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 Australian fauna 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.

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-sampling 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 localised 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.

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).

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 (SFS). 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 SFS (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 minimise 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 their likely 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, 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.

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 visualisation, 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). LWS is supported by a fellowship from the Australian Research Council (CE200100012).

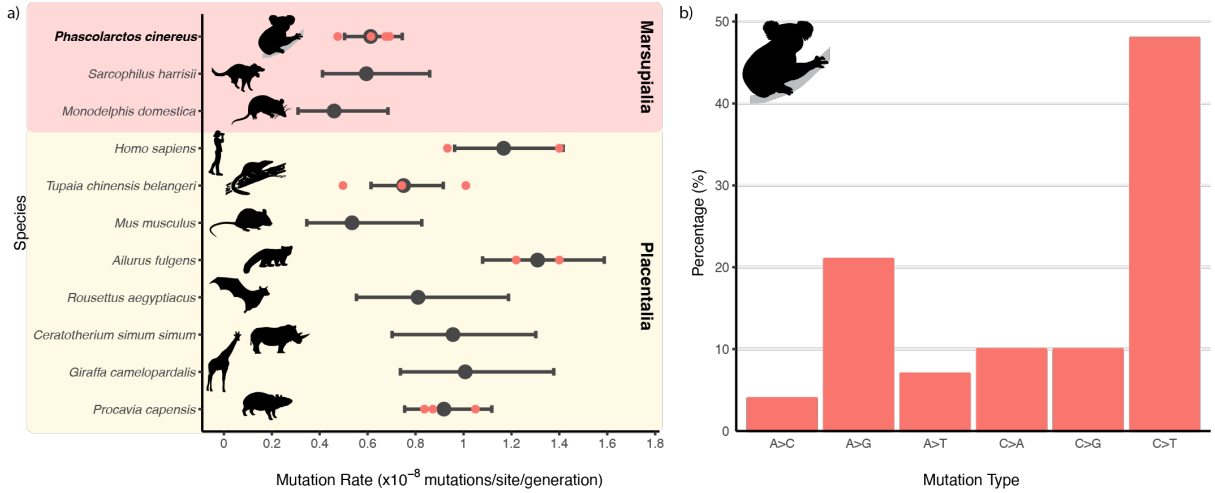

FIG. 1. Germline mutation rate of the koala and other mammals. (*a*) Germline mutation rates inferred in the koala (*Phascolarctos cinereus*) based on four parent-offspring trios. Individual estimates for each trio are shown in orange. The average mutation rate is shown in black and 95% confidence intervals are indicated with error bars. Other mutation rate estimates are from Bergeron et al. (2023). 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.

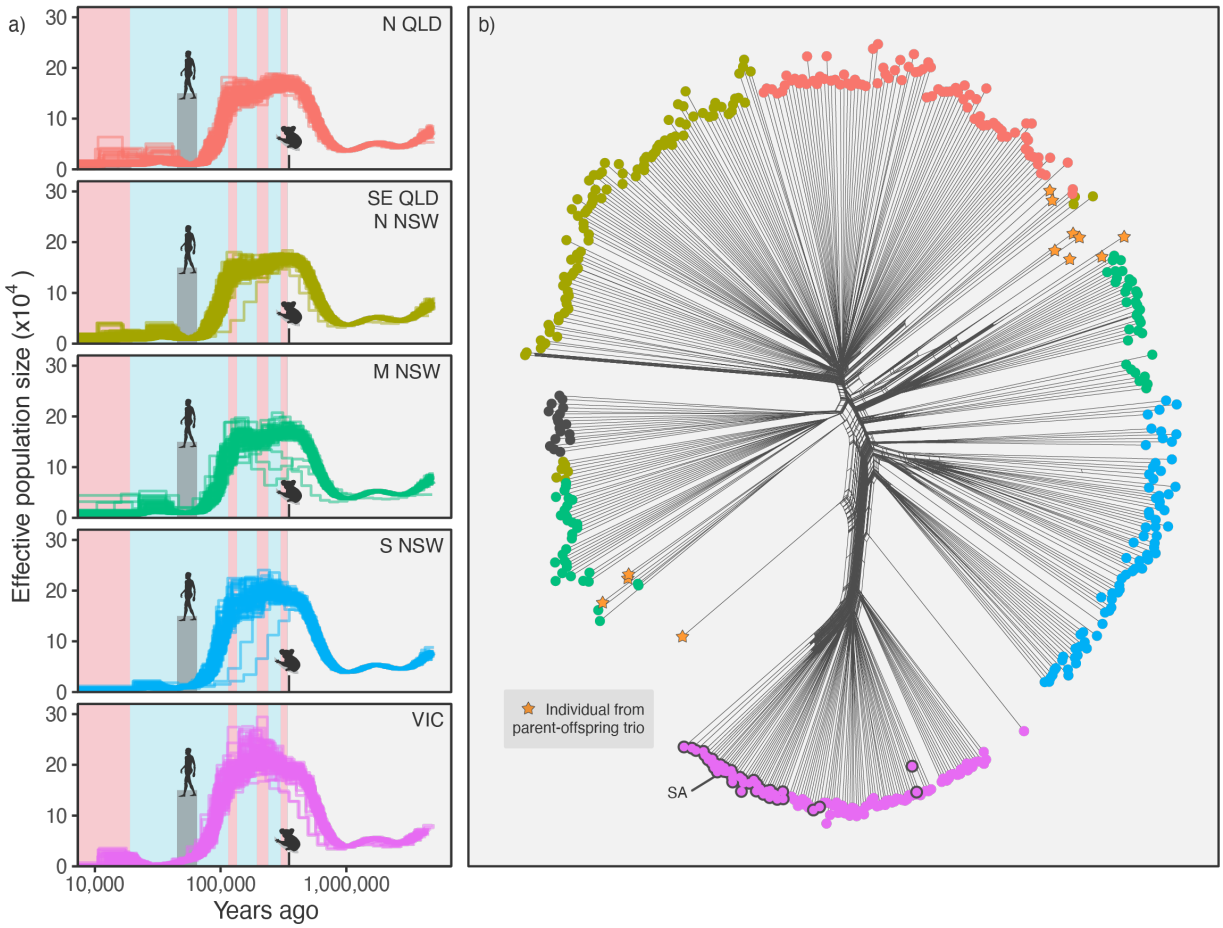

FIG. 2. (*a*) Historical effective population sizes estimated using PSMC for 458 individuals across the five genetic populations. Backgrounds indicate the interglacial (warmer and wetter; pink) and glacial (cooler and drier; blue) periods of the last three glacial cycles. The dark grey column and human 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.

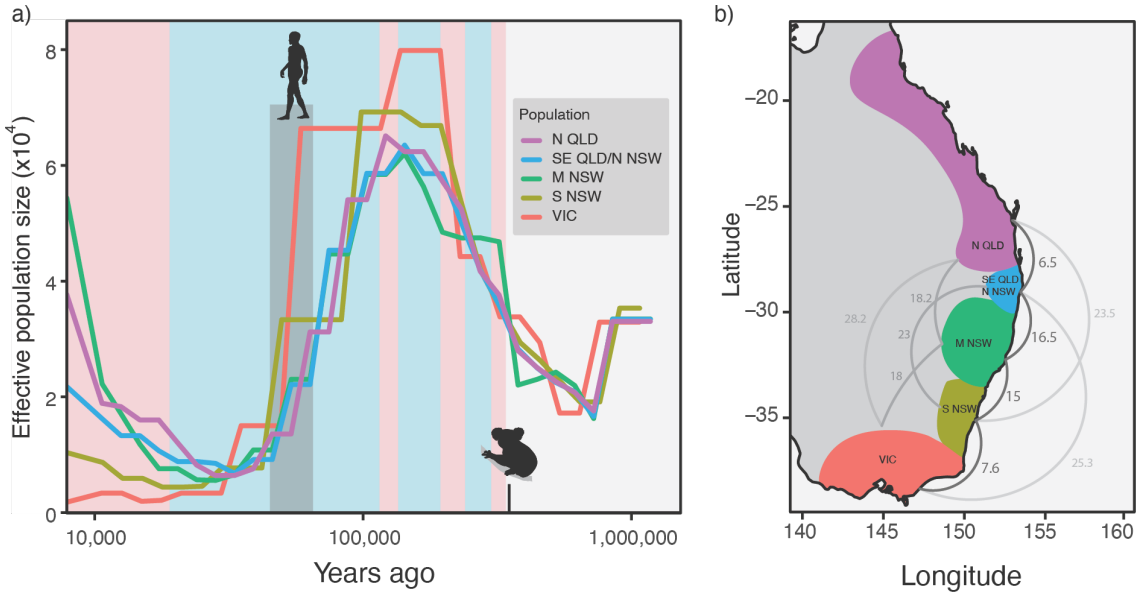

FIG. 3. (*a*) Historical population sizes estimated using SMC++ for each of the five genetic groups of koalas. Pink and blue backgrounds indicate the interglacial (warmer and wetter) and glacial (cooler and drier) periods of the last three glacial cycles. The dark grey column and human silhouette indicate 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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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

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.