Title: Genome of the endangered eastern quoll (*Dasyurus viverrinus*) reveals signatures of historical decline and pelage color evolution.

Gabrielle A. Hartley[1], Stephen R. Frankenberg[2], Natasha M. Robinson[3], Anna J. MacDonald[4], Rodrigo K. Hamede[5], Christopher P. Burridge[5], Menna E. Jones[5], Tim Faulkner[6], Hayley Shute[6], Karrie Rose[7], Robert Brewster[8], Rachel O'Neill[1,9], Marilyn B. Renfree[2], Andrew J. Pask[2,10], Charles Y. Feigin[2*].

1. Institute for Systems Genomics, University of Connecticut, Storrs, CT 06269, USA.
2. School of BioSciences, The University of Melbourne, VIC 3010, Australia.
3. Fenner School of Environment & Society, Australian National University, Canberra, ACT 2601, Australia.
4. Research School of Biology, Australian National University, Canberra, ACT 2601, Australia; Current address: Australian Antarctic Division, Department of Climate Change, Energy, the Environment and Water, Kingston, TAS 7050, Australia.
5. School of Natural Sciences, University of Tasmania, Hobart, TAS 7005, Australia.
6. Australian Reptile Park & Aussie Ark, NSW 2250, Australia.
7. Australian Registry of Wildlife Health, Taronga Conservation Society Australia, Mosman, NSW 2088, Australia.
8. WWF-Australia, PO Box 528, Sydney, NSW 2001, Australia.
9. Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269, USA.
10. Department of Sciences, Museums Victoria, Carlton, VIC 3053, Australia.

*Correspondence: charles.feigin@unimelb.edu.au.

Abstract

The eastern quoll (*Dasyurus viverrinus*) is an endangered marsupial mesopredator native to Australia. Since the extirpation of the last mainland Australian populations in the late 20th century, wild populations of this species have been restricted to two islands at the far southern end of its historical range. Eastern quolls are the subject of captive breeding programs and attempts have recently been made to re-establish a population in mainland Australia through translocations. However, few resources currently exist to guide the genetic management of this species. Here, we generated a chromosome-scale reference genome for the eastern quoll, along with gene annotations supported by multi-tissue transcriptomes. Through comparisons

with related species, we find that our reference genome is among the most complete marsupial assemblies currently available. Using this assembly, we infer the species' demographic history and identify potential evidence of a long-term decline beginning in the late Pleistocene. Finally, we identify a deletion at the *ASIP* locus that likely drives differences in pelage color between the eastern quoll and the closely related Tasmanian devil (*Sarcophilus harrisii*). The genomic resources we present are valuable new tools for evolutionary and conservation genomic studies.

1. Introduction

The eastern quoll (*Dasyurus viverrinus*; Fig. 1a) is a marsupial carnivore native to Australia[2]. A nocturnal mesopredator, the eastern quoll's diet primarily consists of small mammals, reptile and invertebrate prey, as well as opportunistic scavenging on the remains of larger species[4]. The eastern quoll is morphologically, behaviorally, and ecologically distinct from the larger, parapatric spotted-tailed quoll (*Dasyurus maculatus*)[5]. Eastern quolls give birth to exceptionally altricial young that are comparable to mid-gestation fetuses of eutherian mammals[6,7]. Indeed, neonates have structurally immature lungs and perform as much as 95% of their gas exchange through their skin[8-10]. Like other marsupials, their young travel from the birth canal into their mother's pouch on the day of birth and complete much of their development *ex utero*[2,11].

Despite having once ranged across much of southeastern Australia, eastern quolls were likely extirpated from the mainland by the late 20th century[12]. Today, natural populations of this species are restricted to the state of Tasmania on two islands (the Tasmanian main island and Bruny Island; Fig. 1b), which lie at the southern end of their historical range. Moreover, several Tasmanian eastern quoll populations have also undergone significant and ongoing population declines in recent decades and the species was declared 'Endangered' by IUCN in 2016[13-15]. Suggested causes of decline include invasive predators and climatic fluctuation[16,17].

Fig. 1.

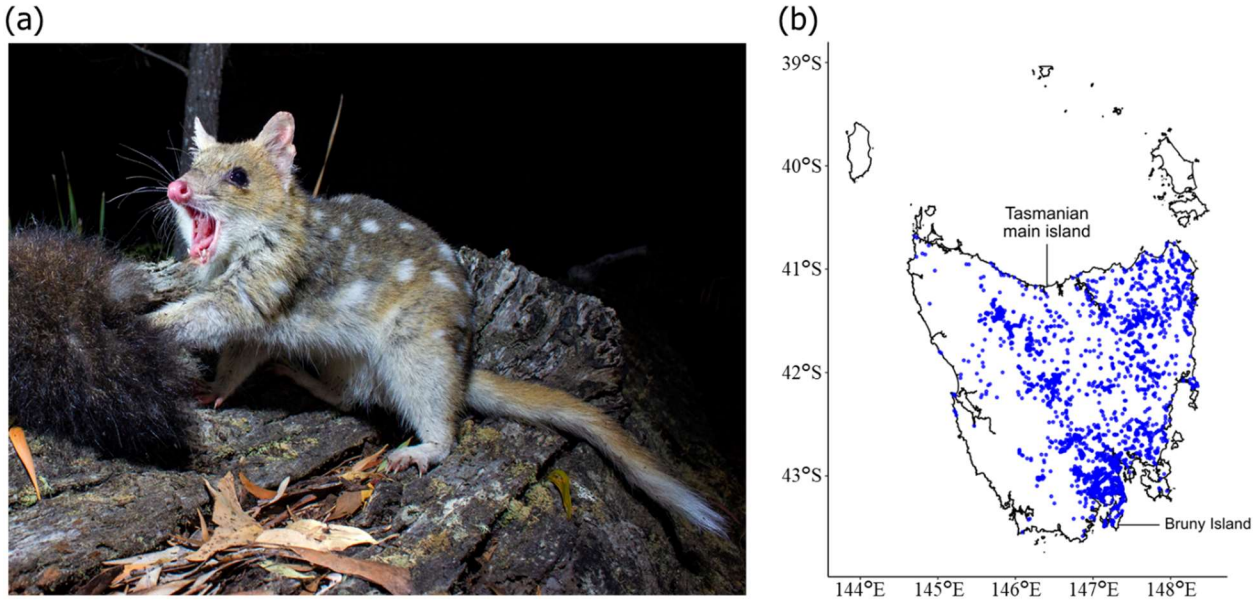

Figure 1. Appearance and distribution of the eastern quoll. a) Photograph of an adult eastern quoll (photo credit Brett Vercoe). b) Map of the state of Tasmania, showing 50 years of eastern quoll sightings across the Tasmanian main island and Bruny Island recorded in the Tasmanian Natural Values Atlas. Individual sightings are shown as blue dots.

The eastern quoll has been flagged as one of 20 priority mammal species in the Australian Government's Threatened Species Strategy, with conservation efforts including population supplementation studies, an extensive captive breeding program spanning multiple sanctuaries, and fenced safe havens free of cats and foxes[18,19]. The species has also recently been the subject of an ambitious project aimed at re-establishing wild populations in parts of its former mainland range, with two pilot releases of captive-bred animals in Booderee National Park on the south coast of New South Wales[20,21]. However, potential challenges to the long-term fitness and adaptive viability of eastern quolls in both Tasmania and mainland Australia remain largely unexplored. The paucity of genomic resources also presents a barrier to conservation genetic management of the species[22]. To address this limitation, we present a high-quality eastern quoll reference genome with transcriptome-based gene annotations.

2. Materials and Methods

2.1 Eastern quoll sightings map

An approximation of the eastern quoll's range within Tasmania was produced by accessing curated sighting data from studies recorded in the Tasmanian Natural Values Atlas (https://www.naturalvaluesatlas.tas.gov.au/), filtering to those within the 50-year period from 1/1/1973 to 1/1/2023. Locations were visualized on a map of Tasmania using a custom R script.

2.1 Tissue sampling

Samples of kidney, heart, liver, and spleen tissue were opportunistically collected as secondary use during the necropsy of an adult, brown morph female quoll named Manda, originating from the Aussie Ark captive breeding sanctuary. Samples were collected under the auspices of the Taronga Conservation Society Australia's Animal Ethics Committee (approval no. 3b1218), pursuant to NSW Office of Environment and Heritage-issued 543 scientific license no. SL100104.. Excess tissue samples from this individual were deposited in the Australian Museum along with the pelt and skeleton under accession number M.52159. Subsampled tissues were snap frozen and were stored at -80°C until used.

Additionally, slices of toe pad were taken and used to isolate primary dermal fibroblasts in culture. Briefly, small pieces of footpad tissue were scored into the base of a 30-mm plastic culture dish. 0.5 mL of DMEM (Gibco 10569044) containing 10% fetal bovine serum and antibiotic/antimycotic) was carefully added to the dish without disrupting the tissue pieces and incubated at 33˚C in 5% CO2. The next day, medium was carefully added up to a total of 2 mL, then replaced every two days. Once the fibroblast outgrowths were 90% confluent, cells were expanded by trypsinization to a T25 flask, then five aliquots cryopreserved in fibroblast medium containing 10% DMSO. Cells were expanded in culture for eight passages before pellets were collected and frozen at -80°C for RNA-seq.

2.2 Genome sequencing and assembly

89 Samples of liver tissue were used by Dovetail Genomics (Scotts Valley, CA USA) for
90 genome sequencing and assembly. High molecular weight DNA was extracted using the Qiagen
91 Blood & Cell Culture DNA miniprep kit (catalogue no: 13323) and quantified using a Qubit 2.0
92 Fluorometer (Life Technologies, Carlsbad, CA, USA). PacBio (Menlo Park, CA, USA) SMRTbell
93 libraries were constructed using the SMRTbell Express Template Prep Kit 2.0. Libraries were
94 bound to polymerase using the Sequel II Binding Kit 2.0 and were sequenced on Sequel II 8M
95 SMRT cells. This yielded approximately 97.68 gigabases of sequence comprised of 5,959,830
96 HiFi reads with an average length of ~16,390nt. HiFi reads were then assembled into contigs
97 using Hifiasm v0.15.4-r343 with default parameters[23]. Unpurged duplicates were then removed
98 using purge_dups (v1.2.5) using automated thresholds[24].
99 A 3D chromatin contact library was produced from additional liver subsamples using
100 Dovetail's Omni-C kit and was sequenced on an Illumina HiSeqX, generating approximately 126
101 gigabases of sequence (~420 million read pairs in 2x150bp format). Chromosome-scale
102 scaffolds were then produced by first aligning Omni-C libraries to the *de novo* contigs with BWA
103 mem and the HiRise Pipeline was used to make contig joins and break putative misjoins[25,26].
104 Contact data were visualized by processing alignments with Pairtools (v1.0.2) and creating a
105 .hic file with juicer_tools pre (v1.22.01) which was loaded into Juicebox (v1.11) and exported to
106 generate a contact map[27-29].

107 **2.3 Genome assessment**

108 Genome assembly metrics related to contiguity, base composition and contig/scaffold
109 lengths were generated using the stats.sh script contained in the bbmap v39.01package[30].
110 Assembly completeness was further assessed via the recovery of benchmarking orthologs
111 using BUSCO v5.4.6 in genome mode with the mammalia_odb10 database[31]. BUSCO gene
112 recovery in the eastern quoll assembly was compared to that of all other marsupial whole-

113 genome assemblies hosted on the NCBI Genomes database and labelled as the representative
114 genome for their species as of April 18th 2023[32-42].

115 **2.4 Inference of chromosome homology**

116 Homology between eastern quoll chromosomes and those of the related Tasmanian devil
117 and yellow-footed antechinus were inferred using gene annotation overlap. Briefly, homologs of
118 Tasmanian devil genes and antechinus genes respectively were identified via liftover to the
119 eastern quoll genome using Liftoff v1.6.3 (parameters -d 4 -a 0.9 -s 0.9)[43]. Annotations were
120 compared between chromosome-scale scaffolds in the eastern quoll and each reference
121 dasyurid's genomes. Chromosomes sharing >= 95% of their gene content were deemed to be
122 homologous. Annotation-based homology assignment was supported by the nearly identical
123 relative sizes of presumptive homologous chromosomes between species, consistent with the
124 exceptionally conserved karyotype previously reported among all examined dasyurids[44,45].

125 **2.5 RNA-sequencing and gene annotation**

126 Frozen samples of eastern quoll heart, kidney, liver, spleen and a pellet of cultured dermal
127 fibroblasts were provided to Psomagen Inc (Rockville, MD USA). RNA-seq libraries were
128 prepared using the TruSeq Stranded mRNA LT sample prep kit (15031047 Rev. E) and
129 sequenced on an Illumina NovaSeq 6000 in 2x150bp format. Residual adapters were removed,
130 and reads were trimmed and filtered for quality using Trimmomatic v0.39 (parameters:
131 SLIDINGWINDOW:5:15, MINLEN:50, AVGQUAL:20, ILLUMINACLIP:2:30:10)[46]. After
132 processing, libraries ranged from approximately 29 to 40 million retained read pairs.

133 To annotate genes in the eastern quoll genome, filtered RNA-seq reads from all five tissues
134 together with RefSeq homologous proteins from seven other marsupial species were provided
135 to the funannotate v1.8.14 pipeline which integrated these with augustus *ab initio* predictions to
136 infer gene models[47]. Within the funannotate pipeline, the following modules were used: train
137 (parameters: --no_trimmomatic, --stranded RF), predict (--augustus_species human, --

138 busco_seed_species human, --optimize_augustus, --busco_db mammalia, --ploidy 1, --
139 organism other, --min_intronlen 10, --max_intronlen 100000, --repeats2evm), and update
140 (default parameters). This approach produced models for 29,622 genes, including 31,319
141 protein-coding transcripts. Of these, we were able to assign gene symbols to 14,293, based on
142 high-confidence 1-to-1 orthology inferred using eggNOG-mapper's Mammalia database, a figure
143 comparable to that of the Tasmanian devil and yellow-footed antechinus annotations produced
144 by RefSeq (15,613 and 15,573 respectively)[48]. The density of annotated genes was visualized
145 using the RIdeogram, providing a histogram of gene counts across 10 megabase (mb) windows
146 on each chromosome[49].

147 **2.6 Repeat masking and annotation**

148 Repeats in the eastern quoll assembly were annotated with RepeatMasker (v4.1.3) using
149 the NCBI BLAST derived search engine rmblast, sensitive settings (-s), and a combined Dfam
150 (v3.6) and Repbase (v20181026) repeat database for marsupials (-species metatheria)[50,51]. The
151 repeat annotations produced were used to hardmask the assembly using Bedtools (v2.29.0)[52].
152 Subsequently, RepeatMasker was performed on the hardmasked assembly using a custom
153 repeat library to identify KERV long terminal repeat elements (LTRs) and other marsupial-
154 derived satellites absent in the above repeat databases. The resulting repeat annotations were
155 combined and summarized using the RepeatMasker utility script buildSummary.pl.

156 **2.7 Historical demography**

157 Haplotype-phased variants for the reference eastern quoll were identified by parsing
158 aligned Omni-C reads with samtools and pairtools and passing them to the Google Deepvariant
159 pipeline which was run with default settings to call and filter variants[27,53,54]. Variant phasing was
160 performed using HapCUT2 (v1.2)[55]. The three largest chromosomes in the eastern quoll
161 genome exceed the maximum scaffold size limitations in the .bai alignment index required for
162 several steps in data processing. Therefore, a second copy of the assembly was made, in which

each of these scaffolds was split into two equal halves using samtools faidx. HiFi reads were then mapped against this copy using minimap2 v 2.24-r1122 (parameters -a -x map-hifi) and filtered with samtools view (parameters -h -q 20 -F 2304)[54,56]. The alignment file was then reduced to only chromosome-scale scaffolds corresponding to the six autosomes and the average mapping depth was calculated with samtools depth. A mapping coverage mask file was then generated using the bamcaller.py included in the MSMC2 package, providing the average coverage depth calculated above. Additionally, the RepeatMasker bed file was used to identify construct a mappability mask which excluded repetitive regions. These masks were provided together with heterozygous SNPs with QUAL >= 10 to the msmc-tools script generate_multihetsep.py (https://github.com/stschiff/msmc-tools). Finally, MSMC2 v2.1.4 was used to infer the eastern quoll's demographic history with 50 EM-iterations (-i 50) and time pattern (-p) 1*2+20*1+1*5, reducing the number of free parameters relative to default settings and grouping the last five time segments to reduce overfitting at extremely ancient time periods. Data were then plotted using the previously-reported generation time of two years and a per-site, per-generation mutation rate of 5.95e-9[57,58].

2.8 Comparative genomics

The *ASIP* sequence was annotated in all available dasyuromorph reference genomes on NCBI via lift-over from the annotation of the yellow-footed antechinus (see: 'Inference of chromosome homology'). Sequences were then extracted with gffread, translated into the conserved reading frame using MACSE v2, and then aligned using the MAFFT web server using default parameters[59-61].

To generate coverage histograms, Tasmanian devil long reads used to generate mSarHar1.11 were accessed from NCBI SRA (Accession: ERR3930603) and aligned against the eastern quoll genome with minimap2 v2.24 (parameter: -x map-ont). Pileups of reads mapped to the *ASIP* locus were then visualized using IGV[62].

188 To perform alignments of genomic sequence surrounding the *ASIP* exon 1 region, *ASIP*
189 exon 2 from the Tasmanian devil was aligned against the eastern quoll (DasViv_v1.0), yellow-
190 footed antechinus (AdamAnt_v2), and Tasmanian devil (mSarHar1.11 and
191 SarHar_Dovetail_2.0) assemblies using blastn with default settings[63]. Samtools faidx was then
192 used to extract the region containing the putative deletion region (i.e., between conserved flank
193 sequences found in all species). These regions were then and were then aligned using the
194 MAFFT web browser with default parameters[54,61].

195 **3. Results**

Fig. 2.

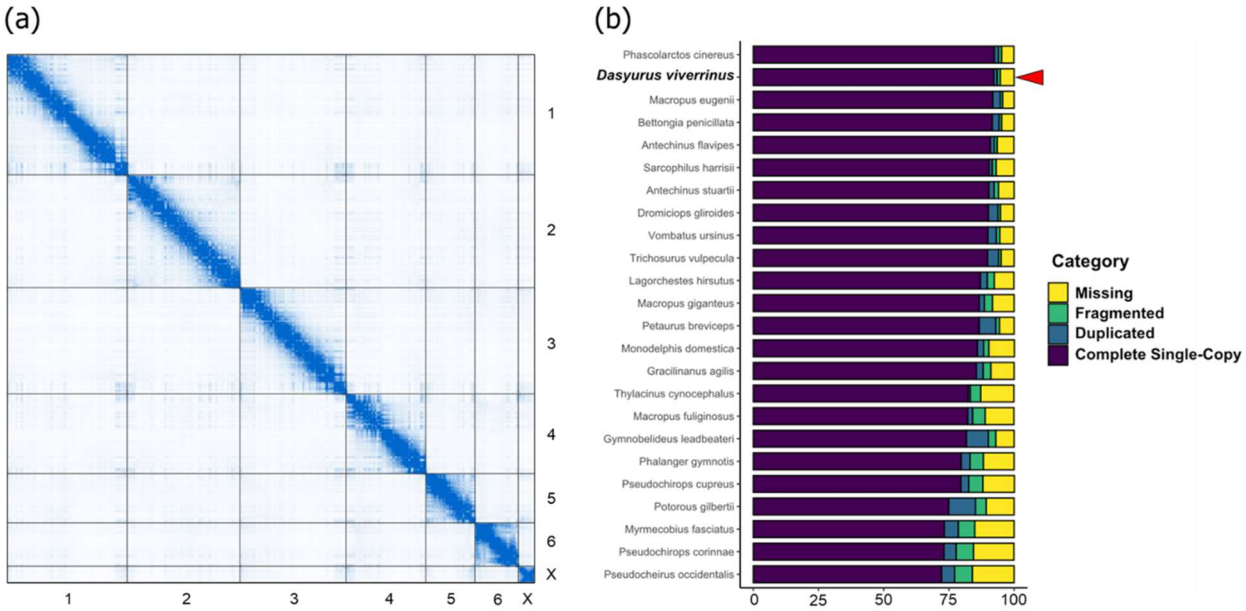

Figure 2. Assessment of genome integrity and completeness. a) Contact map of Omni-C data against chromosome-scale scaffolds in the eastern quoll assembly. The enrichment level of chromosomal contacts is shown as blue pixel intensity b) Stacked bar plot comparing the recovery of mammalian BUSCO genes. Recovery in the eastern quoll is among the highest for sequenced marsupials, with low rates of duplicated or fragmented orthologs.

196 **3.1 The eastern quoll reference genome**

197 Eastern quoll samples were collected opportunistically during the post-mortem examination
198 of an adult female, including several tissues which were frozen and live toe pad which was used
199 to isolate primary fibroblasts in culture (Supplementary Table S1). Using these, we generated a

200 chromosome-scale, de novo reference genome by assembling ~97.68 gigabases (Gb) of Pacific
201 Biosciences high-fidelity (HiFi) long-reads and ~126 Gb of Omni-C long-range chromatin
202 contact data. We called the resulting assembly DasViv_v1.0. The assembly size was ~3.14 Gb,
203 comparable to that of related marsupials, and scaffolds were nearly free of internal gaps
204 (Supplementary Table S2)[36,38]. The assembly was composed of only 76 scaffolds, of which the
205 seven largest corresponded to the conserved karyotype found across all known dasyurids: six
206 autosomes plus the X chromosome (Fig. 2a)[44,45]. Together, the seven chromosome-scale
207 scaffolds accounted for 99.34% of the total assembly size. Homology between eastern quoll
208 chromosome-scale scaffolds and those of the related Tasmanian devil (*Sarcophilus harrisii*) and
209 yellow-footed antechinus (*Antechinus flavipes*) was confirmed by the high overlap of
210 orthologous gene annotations (>95%) and their similar total lengths (Supplementary Table
211 S3)[36,38]. Recovery of complete single-copy mammalian BUSCOs (Benchmarking Universal
212 Single-Copy Orthologs) was 92.2%, second only to the koala (*Phascolarctos cinereus*) among
213 marsupial reference genomes currently available on NCBI (Fig. 2b). Moreover, rates of
214 duplicated and fragmented BUSCO genes were low (1.3% and 1.2%, respectively), reinforcing
215 the completeness and integrity of our assembly.

216 **3.2 Genome annotation**

217 To accompany our assembly, we generated gene annotations by combining evidence from
218 transcriptome data generated from five tissues (Supplementary Table S1), homologous proteins
219 from related marsupials (Supplementary Table S4), as well as *ab initio* predictions. In total, we
220 generated 29,622 gene models (Fig. 3a)[32,36,38,41,64].

221 We also annotated and characterized the repeat content of the eastern quoll genome using
222 RepeatMasker[50]. In total, ~1.476 gigabases were masked as repetitive (47.2% of the assembly),
223 comprising mainly LINEs (30.61%) and SINEs (8.93%; Fig. 3b). L1 repeats constituted the most
224 abundant LINEs in the assembly (20.15%) and MIRs constituted the most abundant SINEs

225 (7.67%). Notably, we identified a small fraction of bases (7,906 in total) corresponding to
226 putative KERV elements, an endogenous retrovirus that has undergone radical expansion in the
227 kangaroo genus *Macropus* (Supplementary Table S5)[65].

Fig. 3.

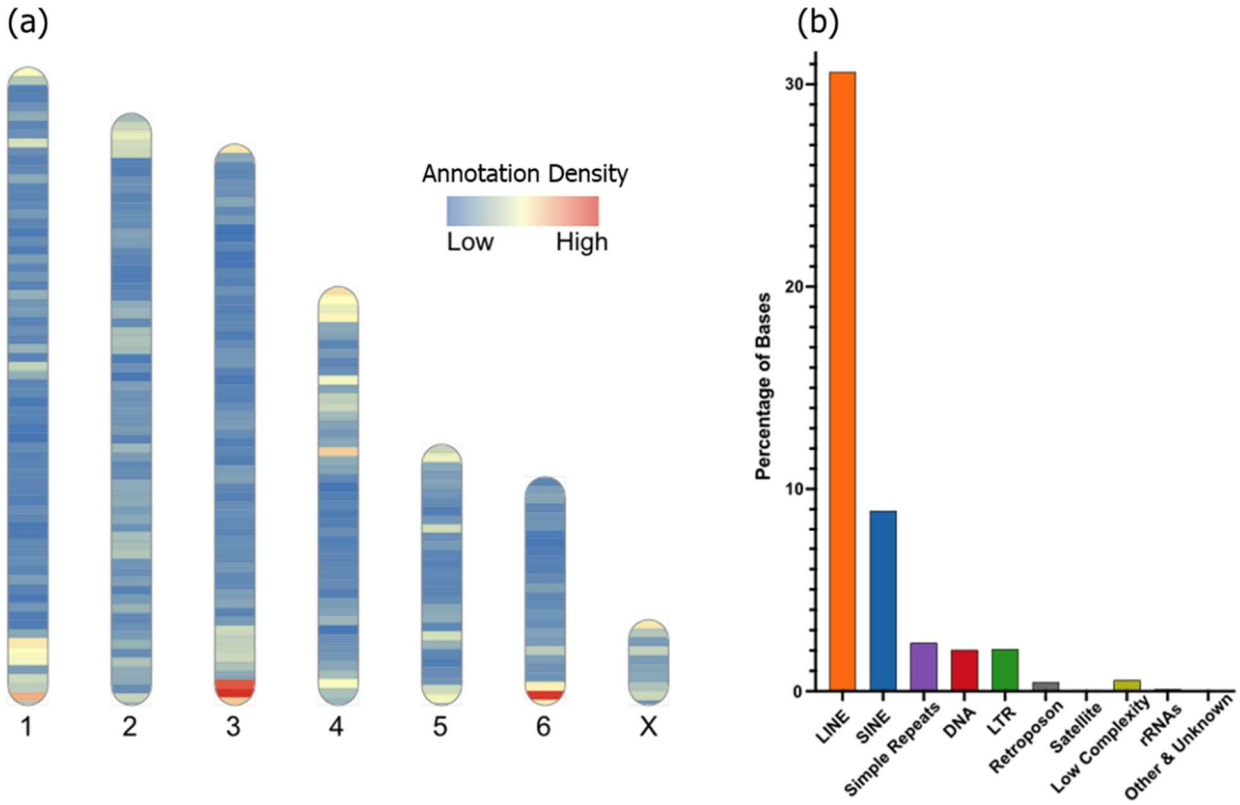

Figure 3. Gene and repeat annotation. a) Heat map illustrating the density of annotated genes across eastern quoll chromosome-scale scaffolds. b) Bar plot showing the distribution of annotated repeats by class.

228 **3.3 Historical demography**

229 Australia has experienced substantial ecological changes over recent geological epochs.
230 For instance, the late Pleistocene saw extensive climatic changes, the periodic isolation of
231 Tasmania from the mainland, and the arrival and dispersal of humans into Australia[1,66,67]. This
232 period also marked the extinction of the remaining megafauna (terrestrial vertebrates >40
233 kilograms)[68]. However, the effects that these potential historical stressors may have had on
234 eastern quoll populations are poorly explored. Therefore, we next sought to examine historical

235 trends in eastern quoll effective population size (N_e) with multiple sequentially Markovian
236 coalescent (MSMC) analysis, using a mutation rate for the closely-related Tasmanian devil
237 measured from parent-offspring trios and a generation time of 2 years[57,58].
238 Our analysis indicated long-term decline in N_e, comprised of two phases (Fig. 4a). The first
239 of the inferred decline phases is indicated to have begun ~300 thousand years ago (kya) and
240 continued until approximately 70 kya. However, recent work on the properties of MSMC analysis
241 in the related Tasmanian devil suggest that accuracy can decline markedly during very recent
242 and very ancient periods (less than 1,000 and greater than 100,000 generations ago,
243 respectively)[3]. Thus, we are cautious in over-interpreting the precise start of this decline phase,
244 which falls outside of this range in our analysis (~2-200 kya). The more recent inferred decline
245 began approximately 28 kya and extended into the Holocene (Fig. 4a). The arrival of humans in
246 Tasmania ~40 kya seems an unlikely candidate to have driven this decline, as they co-existed
247 with eastern quolls for some 12 kyr with apparently little impact on N_e[67]. Curiously, our analyses
248 indicate that both declines may have initiated during periods in which Tasmania was connected
249 to mainland Australia due to lowered sea levels[69], despite the potential for increased gene flow
250 between Tasmanian and mainland populations. It is possible that unfavorable climates during
251 these colder periods created suboptimal environmental conditions in Tasmania and it is unclear
252 whether the Bassian Plain (the land bridge connecting Tasmania and mainland Australia)
253 environments were conducive to gene flow. Indeed, bare-nosed wombat (*Vombatus ursinus*)
254 populations with past territorial connections via the Bassian (i.e. on Tasmania, Flinders Island,
255 and mainland Australia) appear to have maintained genetic isolation[70]. Future resequencing
256 studies of eastern quolls from Tasmanian, Bruny Island, and preserved historical mainland
257 specimens may provide further insights into historical gene flow and population structure, which
258 can confound MSMC and related analyses[71].

Fig. 4.

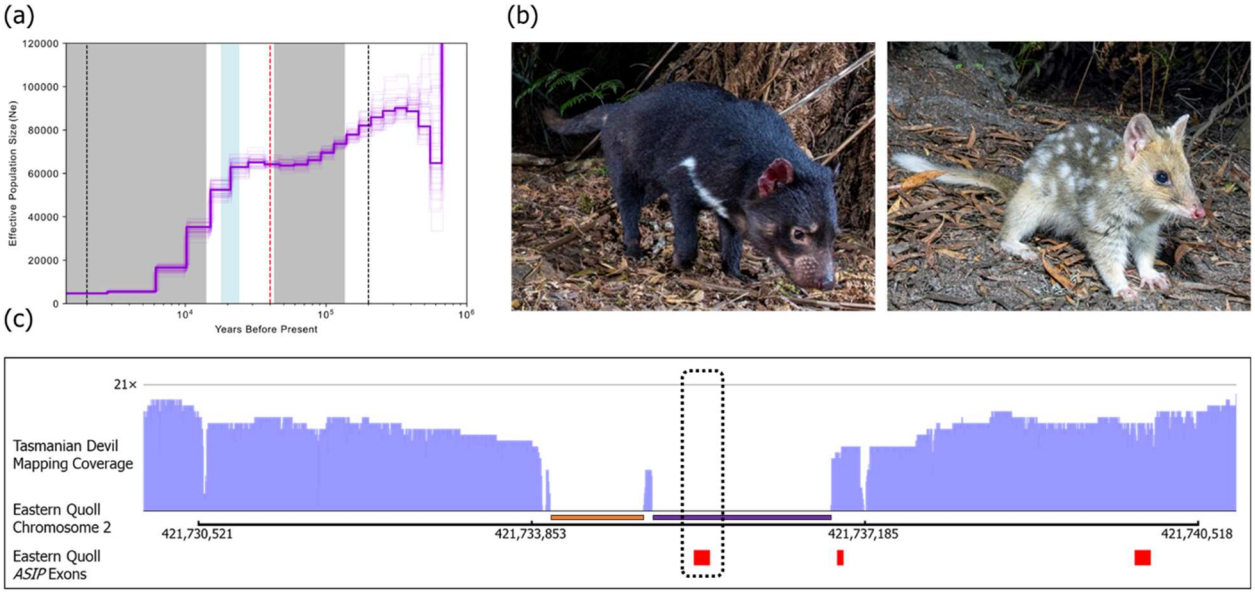

Figure 4. Demography and comparative genomics of pigmentation. a) Step plot illustrating inferred changes in eastern quoll effective population size (Ne) over time. Grey regions indicate periods in which Tasmania was separated from mainland Australia (~135-43 kya and ~14 kya to present). The blue region indicates the coldest period of the Last Glacial Maximum (~24-18 kya)[1]. The dashed red line represents the arrival of humans in Tasmania (~40 kya). Area between the black dashed lines at 2,000 and 200,000 years represent the window within which MSMC analysis is expected to be accurate based on previous studies (1-100 thousand generations before present)[3]. b) Photographs comparing pelage color of the Tasmanian devil and eastern quoll (photo credits Brett Vercoe). c) Mapping coverage of Tasmanian devil long reads (in light purple) across the *ASIP* locus in the eastern quoll genome. Exons of the eastern quoll *ASIP* gene are shown as red blocks. The deletion region is indicated by the absence of mapped Tasmanian devil reads underlined by a purple bar and encompassing exon 1 (dashed box) of *ASIP*. The putative eastern quoll insertion is underlined by an orange bar, upstream of the Tasmanian devil deletion.

3.4 Comparative genomics of pigmentation

The closest living relative to the quolls (genus *Dasyurus*) is the Tasmanian devil. Notably, quolls and Tasmanian devils differ markedly in their coat color and patterning. The quoll background coat color consists of brown dorsal fur, punctuated by spots of white fur and white-to-yellow ventral fur. In contrast, Tasmanian devils have nearly uniform brown/black fur on both dorsum and ventrum, with many individuals bearing unpigmented white patches on the chest, shoulders and/or base of the tail (Fig. 4b). The color of the background dorsal fur of eastern quolls is produced by alternating bands of yellow pheomelanin and dark brown eumelanin in

267 individual hair shafts, a pattern called “agouti”[72]. The agouti pattern, which is common among
268 diverse mammals and likely ancestral among dasyurid marsupials, is known to be regulated in
269 part by the interactions of two key proteins, the melanocortin 1 receptor (MC1R) which promotes
270 eumelanin production and agouti signaling peptide (ASIP), which antagonizes MC1R leading to
271 pheomelanin synthesis[73]. Melanistic morphs of many animals have been shown to be caused by
272 loss-of-function (LOF) mutations in the *ASIP* gene, including multiple independent
273 polymorphisms among *Peromyscus* mice and in wild cats[74,75]. Therefore, we hypothesized that
274 the dark, eumelanin-bearing hair in Tasmanian devils might therefore have evolved through a
275 comparable mechanism.

276 We first sought to compare ASIP orthologs from our eastern quoll genome (DasViv_v1.0)
277 and the reference genomes of the Tasmanian devil (mSarHar1.11) with those from several
278 other dasyuromorph species by aligning their coding sequences (Supplementary Data S1 and
279 Supplementary Table S6)[34,36-38,42]. Notably, while extracting orthologs from each reference
280 genome, we failed to identify the first exon of ASIP from mSarHar1.11, despite the assembly
281 being of high-quality and containing few gaps[38]. To rule out the possibility that this region is
282 present in the Tasmanian devil genome, but not reliably assembled during the construction of
283 mSarHar.1.11, we aligned the nanopore long reads used to produce this genome against our
284 eastern quoll genome. A histogram of read coverage showed two regions in which no
285 Tasmanian devil reads mapped (Fig. 4c). Among these, the second region (~1.8kb long)
286 encompassed the entirety of *ASIP* exon 1. Together, these observations indicated the presence
287 of a potential deletion of the first exon of this gene, including the start codon, supporting the
288 notion of an *ASIP* LOF underlying the Tasmanian devil’s melanistic coat color.

289 To confirm that the putative *ASIP* exon 1 deletion was not unique to the individual animal
290 used to produce the mSarHar1.11 assembly, we next extracted the orthologous genomic region
291 from another Tasmanian devil genome assembly available on NCBI, (SarHar_Dovetail_2.0),
292 which was generated from a different animal and used different sequencing and assembly

293 approaches. Additionally, we extracted this region from a third dasyurid species, the yellow-
294 footed antechinus (AdamAnt_v2) for comparison. Interestingly, alignment of these regions
295 revealed that the first, upstream region which had shown zero Tasmanian devil read mapping
296 coverage likely represents an eastern quoll-specific insertion, as this sequence was not found in
297 either Tasmanian devil or yellow-footed antechinus (Supplementary Data S2). However,
298 consistent with our previous observations, our alignments also revealed identical deletion
299 breakpoints for the putative deletion region encompassing *ASIP* exon 1 in both Tasmanian devil
300 individuals (Supplementary Data S2).

301 Taken together, these results provide strong evidence for the fixation of a loss-of-function
302 mutation in the Tasmanian devil ortholog of *ASIP*. Given the known effects of convergent *ASIP*
303 LOF in other mammals, such a mutation is expected to cause a melanistic (non-agouti) pigment
304 phenotype, thus underpinning the stark difference in background coat color between the
305 Tasmanian devil and quolls.

306 **4. Discussion**

307 Here, we present new, high-quality genomic resources for the endangered eastern quoll.
308 Our chromosome-scale genome assembly closely matches the known chromosome
309 complement of the species which is conserved among dasyurid marsupials. The completeness
310 and contiguity of the assembly exceeds that of most existing marsupial reference genomes as
311 indicated by high BUSCO recovery and low gap percentage. Our comparative analysis of core
312 pigmentation loci reveals the probable basis of pelage variation between the eastern quoll and
313 its close relative the Tasmanian devil, and genome annotations provided here represent a tool
314 for future comparative studies. Moreover, we identify preliminary evidence of historical
315 demographic declines, reinforcing the value of future population genomic studies aimed at
316 defining diversity, population structure, and genetic load in this species.

Data availability

The eastern quoll reference genome and all sequence data used in its generation are available on NCBI under BioProject PRJNA758704. Transcriptome data used in gene annotations are available under BioProject PRJNA963007. Gene annotation GFF files and all original code used in this study can be accessed in a permanent FigShare repository: https://doi.org/10.26188/23501301.v1.

Acknowledgements

We thank Revive & Restore for their funding and organizational support for this project with special thanks to Bridget Baumgartner, Cantata Bio LLC for computational support and sequencing services with special thanks to Mark Daly and Jordan Zhang, Karrie Rose for collecting tissues, Sandy Ingleby and Emma Peel for depositing samples at Australian Museum, the Computational Biology Core in the Institute for Systems Genomics at the University of Connecticut for the use of the Xanadu HPC, Brett Vercoe for the use of his eastern quoll and Tasmanian devil photographs and Elise Ireland for proofreading the manuscript.

Funding

This project was supported by a Wild Genomes grant to C.Y.F., A.J.M., N.M.R., R.K.H., C.P.B. and M.E.J. from Revive & Restore (Contract no. 2020-017). C.Y.F. was supported by NIH NRSA fellowship F32GM139240-01. G.A.H. and R.J.O. are supported by a grant from the US National Institutes of Health R01GM123312-02.

Author contributions

C.Y.F., A.J.M., N.M.R., R.K.H., C.P.B. and M.E.J. conceived the study and acquired funding. C.Y.F. performed assembly quality, gene annotation, historical demographic, and comparative genomic analyses. G.A.H. performed genomic repeat characterization. S.R.F. isolated dermal fibroblasts. R.B., T.F., H.S. K.R. and N.M.R. M.B.R, R.O. and A.J.P. acquired

samples and undertook sample logistics. C.Y.F., G.A.H. and S.R.F. wrote the manuscript with input from all authors.

Conflict of interest

The authors declare no conflict of interest.

References

1. De Deckker, P., Moros, M., Perner, K., et al. 2020, Climatic evolution in the Australian region over the last 94 ka - spanning human occupancy -, and unveiling the Last Glacial Maximum. *Quaternary Science Reviews*, **249**, 106593.
2. Jones, M. E. and Rose, R. K. 2001, Dasyurus viverrinus. *Mammalian Species*, 1-9.
3. Patton, A. H., Margres, M. J., Stahlke, A. R., et al. 2019, Contemporary Demographic Reconstruction Methods Are Robust to Genome Assembly Quality: A Case Study in Tasmanian Devils. *Molecular Biology and Evolution*, **36**, 2906-2921.
4. Jones, M. E. and Barmuta, L. A. 1998, Diet overlap and relative abundance of sympatric dasyurid carnivores: a hypothesis of competition. *Journal of Animal Ecology*, **67**, 410-421.
5. Jones, M. E. and Barmuta, L. A. 2000, Niche Differentiation Among Sympatric Australian Dasyurid Carnivores. *Journal of Mammalogy*, **81**, 434-447.
6. Cook, L. E., Newton, A. H., Hipsley, C. A. and Pask, A. J. 2021, Postnatal development in a marsupial model, the fat-tailed dunnart (Sminthopsis crassicaudata; Dasyuromorphia: Dasyuridae). *Communications Biology*, **4**, 1028.
7. Ferner, K., Schultz, J. A. and Zeller, U. 2017, Comparative anatomy of neonates of the three major mammalian groups (monotremes, marsupials, placentals) and implications for the ancestral mammalian neonate morphotype. *J Anat*, **231**, 798-822.
8. Ferner, K. 2021, Early postnatal lung development in the eastern quoll (Dasyurus viverrinus). *The Anatomical Record*, **304**, 2823-2840.
9. Ferner, K. 2021, Development of the skin in the eastern quoll (Dasyurus viverrinus) with focus on cutaneous gas exchange in the early postnatal period. *J Anat*, **238**, 426-445.
10. Ferner, K., Zeller, U. and Renfree, M. B. 2009, Lung development of monotremes: evidence for the mammalian morphotype. *Anatomical record (Hoboken, N.J. : 2007)*, **292**, 190-201.
11. Tyndale-Biscoe, H. and Renfree, M. 1987, *Reproductive Physiology of Marsupials*. Cambridge University Press: Cambridge.
12. Frankham, G. J., Thompson, S., Ingleby, S., Soderquist, T. and Eldridge, M. D. B. 2017, Does the 'extinct' eastern quoll *Dasyurus viverrinus* persist in Barrington Tops, New South Wales? *Australian Mammalogy*, **39**, 243-247.
13. Burbidge, A. A. and Woinarski, J. 2016, Dasyurus viverrinus. *The IUCN Red List of Threatened Species 2016*, e.T6296A21947190.
14. Fancourt, B. A., Hawkins, C. E. and Nicol, S. C. 2013, Evidence of rapid population decline of the eastern quoll (*Dasyurus viverrinus*) in Tasmania. *Australian Mammalogy*, **35**, 195-205.
15. Cunningham, C. X., Aandahl, Z., Jones, M. E., Hamer, R. and Johnson, C. N. 2023, Regional patterns of continuing decline of the eastern quoll. *Australian Mammalogy*, **45**, 151-159.
16. Fancourt, B. A., Bateman, B. L., VanDerWal, J., et al. 2015, Testing the Role of Climate Change in Species Decline: Is the Eastern Quoll a Victim of a Change in the Weather? *PLoS One*, **10**, e0129420.

17. Fancourt, B. A. 2016, Diagnosing species decline: a contextual review of threats,causes and future directions for management and conservation of the eastern quoll. *Wildlife Research*, **43**, 197-211.
18. Hamer, R. P., Robinson, N., Brewster, R., et al. 2023, Not waiting for the death knell: a pilot study to examine supplementation and survivorship in a declining population of Tasmanian eastern quoll (*Dasyurus viverrinus*). *Australian Mammalogy*, **45**, 171-180.
19. Fraser, H., Legge, S. M., Garnett, S. T., et al. 2022, Application of expert elicitation to estimate population trajectories for species prioritized in Australia's first threatened species strategy. *Biological Conservation*, **274**, 109731.
20. Robinson, N. M., Blanchard, W., MacGregor, C., Brewster, R., Dexter, N. and Lindenmayer, D. B. 2020, Finding food in a novel environment: The diet of a reintroduced endangered meso-predator to mainland Australia, with notes on foraging behaviour. *PLOS ONE*, **15**, e0243937.
21. Robinson, N. M., Blanchard, W., Macgregor, C., Brewster, R., Dexter, N. and Lindenmayer, D. B. 2021, Can evolutionary theories of dispersal and senescence predict postrelease survival, dispersal, and body condition of a reintroduced threatened mammal? *Ecology and Evolution*, **11**, 1002 - 1012.
22. Brandies, P., Peel, E., Hogg, C. J. and Belov, K. 2019, The Value of Reference Genomes in the Conservation of Threatened Species. *Genes (Basel)*, **10**, 846.
23. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. and Li, H. 2021, Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods*, **18**, 170-175.
24. Guan, D., McCarthy, S. A., Wood, J., Howe, K., Wang, Y. and Durbin, R. 2020, Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*, **36**, 2896-2898.
25. Li, H. and Durbin, R. 2009, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**, 1754-1760.
26. Putnam, N. H., O'Connell, B. L., Stites, J. C., et al. 2016, Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Res*, **26**, 342-350.
27. Open2C, Abdennur, N., Fudenberg, G., et al. 2023, Pairtools: from sequencing data to chromosome contacts. *bioRxiv*, 2023.2002.2013.528389.
28. Durand, N. C., Shamim, M. S., Machol, I., et al. 2016, Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst*, **3**, 95-98.
29. Durand, N. C., Robinson, J. T., Shamim, M. S., et al. 2016, Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst*, **3**, 99-101.
30. Bushnell, B. 2014, BBMap: a fast, accurate, splice-aware aligner. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).
31. Seppey, M., Manni, M. and Zdobnov, E. M. 2019, BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods Mol Biol*, **1962**, 227-245.
32. Johnson, R. N., O'Meally, D., Chen, Z., et al. 2018, Adaptation and conservation insights from the koala genome. *Nature genetics*, **50**, 1102-1111.
33. Dudchenko, O., Batra, S. S., Omer, A. D., et al. 2017, De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science*, **356**, 92-95.
34. Feigin, C., Frankenberg, S. and Pask, A. 2022, A Chromosome-Scale Hybrid Genome Assembly of the Extinct Tasmanian Tiger (Thylacinus cynocephalus). *Genome Biology and Evolution*, **14**.
35. Peel, E., Silver, L., Brandies, P., Hogg, C. J. and Belov, K. 2021, A reference genome for the critically endangered woylie, Bettongia penicillata ogilbyi. *GigaByte*, **2021**, gigabyte35.
36. Tian, R., Han, K., Geng, Y., et al. 2022, A chromosome-level genome of Antechinus flavipes provides a reference for an Australian marsupial genus with male death after mating. *Molecular ecology resources*, **22**, 740-754.

430 37. Brandies, P. A., Tang, S., Johnson, R. S. P., Hogg, C. J. and Belov, K. 2020, The first Antechinus
431 reference genome provides a resource for investigating the genetic basis of semelparity and
432 age-related neuropathologies. *Gigabyte*, **2020**, 0.
433 38. Stammnitz, M. R., Gori, K., Kwon, Y. M., et al. 2023, The evolution of two transmissible cancers
434 in Tasmanian devils. *Science*, **380**, 283-293.
435 39. Rhie, A., McCarthy, S. A., Fedrigo, O., et al. 2021, Towards complete and error-free genome
436 assemblies of all vertebrate species. *Nature*, **592**, 737-746.
437 40. Feigin, C. Y., Moreno, J. A., Ramos, R., et al. 2023, Convergent deployment of ancestral functions
438 during the evolution of mammalian flight membranes. *Science Advances*, **9**, eade7511.
439 41. Tian, R., Han, K., Geng, Y., et al. 2021, A Chromosome-Level Genome of the Agile Gracile Mouse
440 Opossum (Gracilinanus agilis). *Genome Biol Evol*, **13**.
441 42. Peel, E., Silver, L., Brandies, P., Hayakawa, T., Belov, K. and Hogg, C. J. 2022, Genome assembly of
442 the numbat (Myrmecobius fasciatus), the only termitivorous marsupial. *bioRxiv*,
443 2022.2002.2013.480287.
444 43. Shumate, A. and Salzberg, S. L. 2021, Liftoff: accurate mapping of gene annotations.
445 *Bioinformatics*, **37**, 1639-1643.
446 44. Rofe, R. and Hayman, D. 1985, G-banding evidence for a conserved complement in the
447 Marsupialia. *Cytogenetics and cell genetics*, **39**, 40-50.
448 45. Deakin, J. E. 2018, Chromosome Evolution in Marsupials. *Genes (Basel)*, **9**, 72.
449 46. Bolger, A. M., Lohse, M. and Usadel, B. 2014, Trimmomatic: a flexible trimmer for Illumina
450 sequence data. *Bioinformatics*, **30**, 2114-2120.
451 47. Palmer, J. M. and Stajich, J. 2020, Funannotate v1.8.1: Eukaryotic genome annotation (v1.8.1).
452 *Zenodo*.
453 48. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. and Huerta-Cepas, J. 2021,
454 eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at
455 the Metagenomic Scale. *Molecular Biology and Evolution*, **38**, 5825-5829.
456 49. Hao, Z., Lv, D., Ge, Y., et al. 2020, RIdeogram: drawing SVG graphics to visualize and map
457 genome-wide data on the idiograms. *PeerJ. Computer science*, **6**, e251.
458 50. Tarailo-Graovac, M. and Chen, N. 2009, Using RepeatMasker to identify repetitive elements in
459 genomic sequences. *Current protocols in bioinformatics*, **Chapter 4**, Unit 4.10.
460 51. Hubley, R., Finn, R. D., Clements, J., et al. 2016, The Dfam database of repetitive DNA families.
461 *Nucleic acids research*, **44**, D81-D89.
462 52. Quinlan, A. R. and Hall, I. M. 2010, BEDTools: a flexible suite of utilities for comparing genomic
463 features. *Bioinformatics*, **26**, 841-842.
464 53. Poplin, R., Chang, P.-C., Alexander, D., et al. 2018, A universal SNP and small-indel variant caller
465 using deep neural networks. *Nature biotechnology*, **36**, 983-987.
466 54. Li, H., Handsaker, B., Wysoker, A., et al. 2009, The Sequence Alignment/Map format and
467 SAMtools. *Bioinformatics*, **25**, 2078-2079.
468 55. Edge, P., Bafna, V. and Bansal, V. 2017, HapCUT2: robust and accurate haplotype assembly for
469 diverse sequencing technologies. *Genome Res*, **27**, 801-812.
470 56. Li, H. 2018, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, **34**, 3094-
471 3100.
472 57. Burbidge, A., Harrison, P. and Woinarski, J. 2014, *The Action Plan for Australian Mammals 2012*.
473 58. Bergeron, L. A., Besenbacher, S., Zheng, J., et al. 2023, Evolution of the germline mutation rate
474 across vertebrates. *Nature*, **615**, 285-291.
475 59. Pertea, G. and Pertea, M. 2020, GFF Utilities: GffRead and GffCompare [version 1; peer review: 3
476 approved]. *F1000Research*, **9**.

60. Ranwez, V., Douzery, E. J. P., Cambon, C., Chantret, N. and Delsuc, F. 2018, MACSE v2: Toolkit for the Alignment of Coding Sequences Accounting for Frameshifts and Stop Codons. *Molecular Biology and Evolution*, **35**, 2582-2584.
61. Katoh, K., Misawa, K., Kuma, K.-i. and Miyata, T. 2002, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**, 3059-3066.
62. Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., et al. 2011, Integrative genomics viewer. *Nature biotechnology*, **29**, 24-26.
63. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. 1990, Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403-410.
64. Mikkelsen, T. S., Wakefield, M. J., Aken, B., et al. 2007, Genome of the marsupial Monodelphis domestica reveals innovation in non-coding sequences. *Nature*, **447**, 167-177.
65. Ferreri, G. C., Brown, J. D., Obergfell, C., et al. 2011, Recent amplification of the kangaroo endogenous retrovirus, KERV, limited to the centromere. *Journal of virology*, **85**, 4761-4771.
66. Clarkson, C., Jacobs, Z., Marwick, B., et al. 2017, Human occupation of northern Australia by 65,000 years ago. *Nature*, **547**, 306-310.
67. Tobler, R., Rohrlach, A., Soubrier, J., et al. 2017, Aboriginal mitogenomes reveal 50,000 years of regionalism in Australia. *Nature*, **544**, 180-184.
68. Hocknull, S. A., Lewis, R., Arnold, L. J., et al. 2020, Extinction of eastern Sahul megafauna coincides with sustained environmental deterioration. *Nature Communications*, **11**, 2250.
69. Lambeck, K. and Chappell, J. 2001, Sea Level Change Through the Last Glacial Cycle. *Science*, **292**, 679-686.
70. Martin, A., Carver, S., Proft, K., et al. 2019, Isolation, marine transgression and translocation of the bare-nosed wombat (Vombatus ursinus). *Evolutionary applications*, **12**, 1114-1123.
71. Mazet, O., Rodríguez, W., Grusea, S., Boitard, S. and 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.
72. Dry, F. W. 1926, The coat of the mouse (Mus musculus). *Journal of Genetics*, **16**, 287-340.
73. McRobie, H. R., King, L. M., Fanutti, C., Symmons, M. F. and Coussons, P. J. 2014, Agouti signalling protein is an inverse agonist to the wildtype and agonist to the melanic variant of the melanocortin-1 receptor in the grey squirrel (Sciurus carolinensis). *FEBS Letters*, **588**, 2335-2343.
74. Kingsley, E. P., Manceau, M., Wiley, C. D. and Hoekstra, H. E. 2009, Melanism in peromyscus is caused by independent mutations in agouti. *PLoS One*, **4**, e6435.
75. Schneider, A., David, V. A., Johnson, W. E., et al. 2012, How the leopard hides its spots: ASIP mutations and melanism in wild cats. *PLoS One*, **7**, e50386.