Title: A chromosome-scale hybrid genome assembly of the extinct Tasmanian tiger (*Thylacinus cynocephalus*)

Authors and Affiliations: Charles Feigin[1,2*], Stephen Frankenberg[1], Andrew Pask[1,3*]

1. School of BioSciences, The University of Melbourne, Parkville, VIC, Australia

2. Department of Molecular Biology, Princeton University, Princeton, NJ, USA

3. Department of Sciences, Museums Victoria, Carlton, VIC, Australia

*Authors for Correspondence:

Charles Feigin, School of BioSciences, The University of Melbourne, Parkville, Victoria, Australia, +1 (203) 297-3130, charles.feigin@unimelb.edu.au

Andrew Pask, School of BioSciences, The University of Melbourne, Parkville, Victoria, Australia, + 61 3 9035 4310, ajpask@unimelb.edu.au

Abstract

The extinct Tasmanian tiger or thylacine (*Thylacinus cynocephalus*) was a large marsupial carnivore native to Australia. Once ranging across parts of the mainland, the species remained only on the island of Tasmania by the time of European colonization. It was driven to extinction in the early 20th century and is an emblem of native species loss in Australia. The thylacine was a striking example of convergent evolution with placental canids, with which it shared a similar skull morphology. Consequently, it has been the subject of extensive study. While the original thylacine assemblies published in 2018 enabled the first exploration of the species' genome biology, further progress is hindered by the lack of high-quality genomic resources. Here, we present a new chromosome-scale hybrid genome assembly for the thylacine, which compares favorably with many recent *de novo* marsupial genomes. Additionally, we provide homology-based gene annotations, characterize the repeat content of the thylacine genome and show that,

consistent with demographic decline, the species possessed a low rate of heterozygosity even compared to extant, threatened marsupials.

Keywords: Thylacine; Tasmanian tiger; *Thylacinus cynocephalus*; genome; Dasyuromorphia

Significance

The lack of high-quality genomes for extinct species inhibits research into their biology. Moreover, marsupials are underrepresented among sequenced genomes. Here, we present a new, chromosome-scale thylacine genome. This high-quality assembly is a valuable new resource for studies on marsupial carnivores.

Introduction

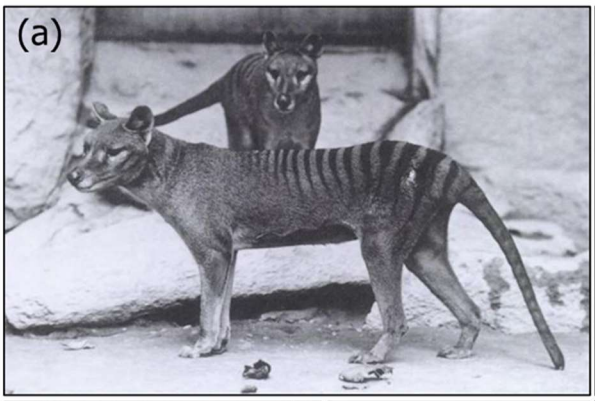
The Tasmanian tiger or thylacine (*Thylacinus cynocephalus*; Fig. 1a) was the largest marsupial predator of the Holocene (Mitchell, et al. 2014; Prowse, et al. 2014). While it once inhabited mainland Australia, by the arrival of European colonists it was restricted to the island of Tasmania (Lambeck and Chappell 2001; Paddle 2000). The thylacine was considered an agricultural pest and targeted by an extermination campaign, incentivized by a £1 bounty (Fig. 1b). The last known individual died in 1936 and the species was declared extinct in 1986 (Paddle 2000). The thylacine was captured in multiple photographs and short films, contributing to its status as an emblem of Australia's high extinction rate among native species (Sleightholme and Campbell 2018; Woinarski, et al. 2015).

The relative abundance of thylacine specimens in museums has facilitated extensive study of its morphology, ecology and evolution (Newton, et al. 2018; Rovinsky, et al. 2021; White, et al. 2018; Wroe, et al. 2007). Recently, it has also become a focal species for genomic research, with the first genome assemblies being published in 2018, using DNA from a >100-year-old ethanol-

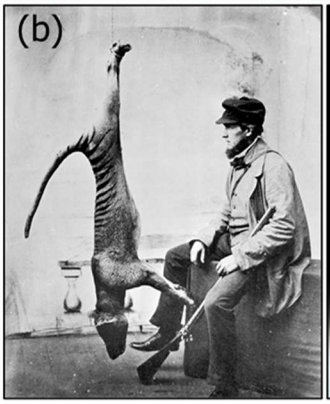
46 preserved pouch young specimen (Fig. 1c) (Feigin, et al. 2018). These assemblies were used to
47 explore the molecular basis of thylacine-canid craniofacial convergence, confirm its
48 phylogenetic relationships, and infer its demographic history (Feigin, et al. 2018). Subsequent
49 studies examined enhancer evolution and characterized the thylacine's immune gene
50 complement (Feigin, et al. 2019; Peel, et al. 2021). However, contiguity of the original
51 assemblies was limited by the fragmentary nature of historical DNA and the absence of high-
52 quality assemblies from related species suitable for reference-guided scaffolding (Feigin, et al.
53 2018). This presents a substantial challenge for continued research into the thylacine's genome
54 biology (Garrett Vieira, et al. 2020; Peel, et al. 2021).

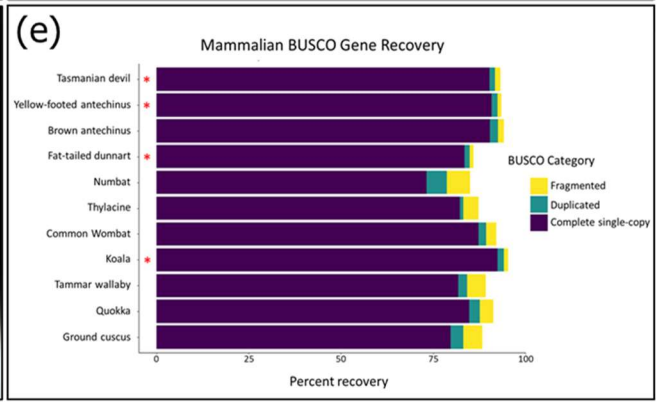
55 The thylacine (family Thylacinidae) represents the closest sister lineage to the families
56 Dasyuridae and Myrmecobiidae (Feigin, et al. 2018; Miller, et al. 2009; Mitchell, et al. 2014).
57 These groups contain numerous species of significant interest to evolutionary, developmental
58 and conservation biology, such as the Tasmanian devil, quolls, dunnarts and the numbat (Cook,
59 et al. 2021; Fancourt 2016; Spencer, et al. 2020; Stahlke, et al. 2021; Wright, et al. 2020).
60 Moreover, the thylacine's exceptional craniofacial similarities with canids, despite their ~160
61 million year divergence, make the species an excellent model system to study the genomic basis
62 of morphological evolution (Bininda-Emonds, et al. 2007; Feigin, et al. 2018; Newton, et al.
63 2021; Rovinsky, et al. 2021). Improved genomic resources for this species are thus of
64 considerable value to the broader genomics community. Here, we leveraged improvements in
65 short read assembly tools and newly-available marsupial reference genomes to produce a
66 chromosome-scale hybrid genome assembly for the thylacine.

Fig. 1



(d)	Contigs	Scaffolds
Total Length	3041 Gb	3.379 Gb
Max Length	165.7 Kb	737.6 Mb
Number	441,756	151,245
N50	15 Kb	629 Mb
N90	9 Kb	479 Mb
G+C %	36.26 %	-
Gap %	-	10%





(a) Adult thylacines in captivity. The thylacine was noted for its canid-like morphology. (b) A wild thylacine killed by a hunter. A bounty on thylacines contributed to their extinction. (c) Thylacine pouch young specimen C5757 (Melbourne Museum; Victoria, Australia) provided DNA used for genome sequencing. (d) Assembly metrics for the improved thylacine genome. (e) Comparison of BUSCO gene recovery from the thylacine genome and several recently-released marsupial assemblies. Asterisk indicates assemblies incorporating long reads.

Results and Discussion

Genome Assembly and Assessment

The new thylacine assembly is composed of 7 large scaffolds, corresponding to each of the 6 dasyuromorph autosomes and the X chromosome (Supplementary Table 1), together comprising ~93.25% of the sequence content (Deakin 2018). The gap-free assembly size is ~3.04Gbp and G+C content is 36.26%, comparable to that of the Tasmanian devil (Fig. 1d, Supplementary Table 2). Scaffold N50 and N90 are high (629Mbp and 479Mbp respectively), reflecting the large size of dasyuromorph autosomes (Deakin 2018). Contig N50 was 5-fold higher than that of

the original *de novo* draft assembly, and similar to that of several other recent marsupial assemblies (Supplementary Table 2). A tail of small scaffolds comprising approximately 205Mbp remained unplaced, contributing to a relatively high gap percentage (~10%; Fig. 1d). Nonetheless, the new assembly represents a dramatic improvement in contiguity.

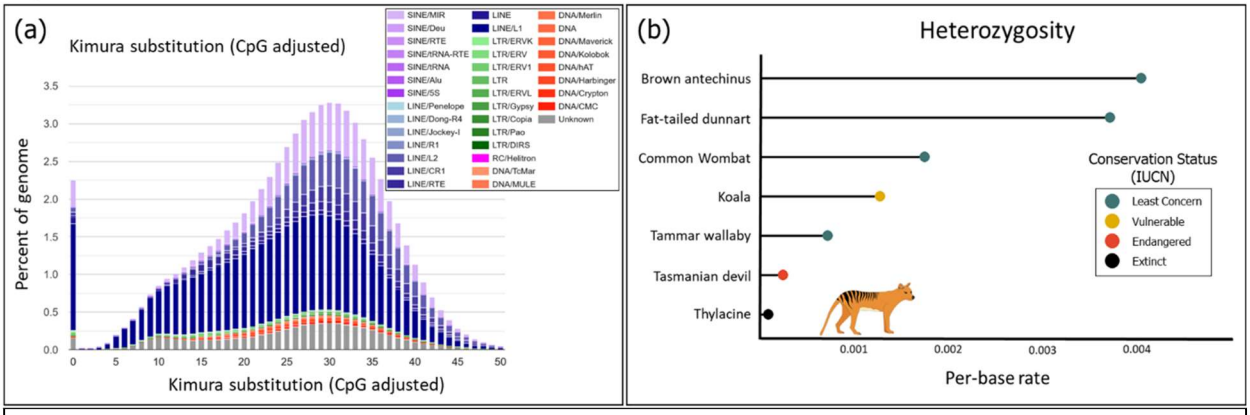
To evaluate the completeness and integrity of the assembly, BUSCO was used to annotate benchmarking mammalian orthologs. This identified 82.3% of BUSCO genes as complete and single-copy, with little duplication (0.9%). Another 4.1% were found as partial copies (Fig. 1e). This is a drastic increase over the original thylacine *de novo* assembly, from which BUSCO recovery was negligible (<10%), owing to low contiguity (Supplementary Table 3). While BUSCO gene recovery compares well with several other recently released marsupial assemblies, particularly those built from short read-based contigs scaffolded with Hi-C, it lags somewhat behind a small number of assemblies built using long reads and Hi-C (Fig. 1e, Supplementary Table 4). Unfortunately, the century-long room-temperature preservation of all existing thylacine tissue samples, and corresponding DNA fragmentation, limits the potential for long read sequencing to be applied productively in this species.

Repeat Classification and Genome Annotation

Repetitive regions in the thylacine genome were annotated with RepeatMasker, using a custom database of species-specific and curated marsupial repeats (Fig. 2a) (Ellinghaus, et al. 2008; Flynn, et al. 2020; Hubley, et al. 2016; Tarailo-Graovac and Chen 2009). Interspersed repeats constituted ~56% of the assembly (Supplementary Table 5). Consistent with the highly conserved genome organization of dasyuromorphs, the thylacine had similar overall repeat composition to its living relatives (Tian, et al. 2022). The dominant repeat class was LINE elements (~36.5%), occurring at a frequency comparable to that of the Tasmanian devil (~39%),

98 though somewhat lower than that of the brown antechinus (~45%) (Tian, et al. 2022).
99 Interestingly, we observed that LTRs were sparse in the thylacine genome (~1.51%) compared to
100 previously studied marsupial species (which ranged from 6.53%-18.89%; Supplementary Table
101 5) (Tian, et al. 2022).

Fig. 2



(a) Interspersed repeat landscape of thylacine genome. The percentage of total genome size and sequence divergence (based on CpG-adjusted Kimura substitution level) are shown for each repeat subclass. (b) Comparison of the per-base rate of heterozygosity in the thylacine and several extant marsupials. The thylacine showed the lowest heterozygosity of examined marsupial species.

102 To provide gene annotations for the new thylacine assembly, we identified orthologs to
103 Tasmanian devil genes using a homology-based annotation liftover procedure (see Methods and
104 Methods). Ortholog recovery was high, with ~96% of gene models being successfully transferred
105 to the thylacine genome, comparable to or exceeding that of other dasyuromorphs
106 (Supplementary Table 6). Interestingly, we observed disparities in the detection of different short
107 RNA classes. In particular, micro-RNAs (miRNAs) showed nearly complete recovery from the
108 thylacine genome (~98%), compared with ~71% of small nucleolar RNAs and just ~37% of
109 small nuclear RNAs (snoRNAs and snRNAs respectively; Supplementary Table 6). A similar
110 pattern was observed among other dasyuromorphs, which showed lower snoRNA and snRNA
111 recovery (particularly in species more distantly-related to the Tasmanian devil), while generally
112 retaining high miRNA recovery (Supplementary Table 6). Taken together, this suggests that

while many miRNAs are ancestral to Dasyuromorphia (hence having orthologs across species) and have remained conserved over time, the evolution of snRNAs and snoRNAs in this lineage has potentially been more dynamic, with accelerated sequence divergence and/or more rapid turnover of individual elements among species.

Genetic Diversity

We next sought to gain insights into the thylacine's genetic diversity prior to its extinction. Previously, multiple sequentially Markovian coalescent (MSMC) analysis was used to infer the demographic history of the thylacine. This uncovered evidence of an extended period of genetic decline predating the arrival of humans in Australia and the thylacine's isolation on Tasmania (Feigin, et al. 2018; Schiffels and Durbin 2014). A decrease in genetic diversity concomitant with such demographic decline may have left the thylacine vulnerable to inbreeding depression, reducing its fitness on the backdrop of pressures imposed by humans. To further explore this possibility, heterozygosity was calculated in non-repetitive regions of the thylacine genome and compared to that of extant marsupials with varying conservation statuses. Consistent with reduced genetic diversity preceding its extinction, the thylacine had the lowest rate of heterozygosity among the marsupials examined, including vulnerable or endangered species (Fig. 2b, Supplementary Table 7).

Conclusions

The quality of the first draft thylacine assemblies limited their utility in genomic research. Gene recovery was severely impaired by low contiguity, and repetitive regions were not adequately represented (Feigin, et al. 2018). By contrast, our new thylacine genome has a ~5-fold larger contig N50, comparable to that of many recent marsupial assemblies. Moreover, we have produced chromosome-scale scaffolds that enable the recovery of numerous genetic elements

with orthologs in related species. This assembly has also permitted the first examination of the repeat composition and heterozygosity of the thylacine genome. Future whole-genome resequencing studies, empowered by this assembly, have the potential to provide population-level insights into the thylacine's demography and level of genetic load prior to its extinction.

Materials and Methods

Genome Assembly

Thylacine reads were accessed from NCBI Sequence Read Archive (SRA; Supplementary Table 8). These data originated from individual C5757, which we previously used to produce the original contig-level *de novo* assembly and a read-mapping-based, reference-guided assembly of non-repetitive regions (Feigin, et al. 2018).

De novo contigs were assembled using MEGAHIT v1.2.9 (Li, et al. 2015) with multiple k-mer lengths (kmers = 21, 29, 39, 59, 79, 99, 119, 141). Purging of redundant haplotypes and short read scaffolding were performed using Redundans v0.14a (parameters: identity = 0.8, overlap = 0.8, minLength = 200bp, joins = 5, limit = 1.0, iterations = 2) (Pryszcz and Gabaldón 2016). Purging removed ~178.5Mbp of sequence.

Dasyuromorphs possess an exceptionally-conserved karyotype (2n = 14), with nearly identical chromosome sizes and g-banding patterns (Deakin 2018; Rofe and Hayman 1985). Moreover, sequence mapability between thylacine and Tasmanian devil is high (Feigin, et al. 2018). Therefore, chromosome-scale thylacine scaffolds were produced by ordering thylacine *de novo* scaffolds and inferring gap sizes through alignment against the recently-available Tasmanian devil reference genome (GCF_902635505.1/mSarHar1.11; (O'Leary, et al. 2016)) using RagTag v2.1.0 (RagTag parameters: scaffold, -f 200, -r, -g 100 -m 10000000; minimap2 v2.22-r1101 parameters: -x asm 10) (Alonge, et al. 2021; Alonge, et al. 2019; Li 2018).

Genome Annotation

Repeat elements were annotated using RepeatMasker v4.1.2 (Flynn, et al. 2020; Tarailo-Graovac and Chen 2009). Custom thylacine repeat libraries were produced with RepeatModeler v2.0.2a and LTRharvest v1.6.2, and were combined with marsupial repeats contained with the Dfam3.2 database (Ellinghaus, et al. 2008; Flynn, et al. 2020; Hubley, et al. 2016). RepeatMasker was then run on each chromosome using this library (Supplementary Table 5). The repeat landscape of the thylacine genome was visualized using the calcDivergenceFromAlign.pl and createRepeatLandscape.pl scripts provided with RepeatMasker. This displays the genome percentage of each repeat subclass, organized by CpG-adjusted kimura substitution level (a distance-based proxy for repeat copy age) (Flynn, et al. 2020; Kimura 1980).

Given the thylacine's extinction, RNA cannot be recovered. However, annotations are essential for many genomic analyses. We therefore employed a homology-based approach implemented in the program liftoff v1.6.1 to predict thylacine orthologs of Tasmanian devil genes (Shumate and Salzberg 2021). Exons from the Tasmanian devil RefSeq annotation were mapped to the thylacine genome assembly with minimap2 (Li 2018; O'Leary, et al. 2016). Thylacine gene models were then produced by linking mapped exons of a common parent feature, retaining only those which preserved the structure of their corresponding Tasmanian devil reference annotation (allowing a distance factor of 4X; parameter -d 4, Supplementary Table 6).

Assembly Evaluation and Comparisons

Assembly completeness and integrity were assessed using Benchmarking Universal Single-Copy Orthologs annotated by BUSCO (v5.2.2) with the mammalian_odb10 ortholog database. These results were compared with several recent *de novo* marsupial genome assemblies (Fig. 1e, Supplementary Tables 3 and 4**)** (Brandies, et al. 2020; Dudchenko, et al. 2017; Johnson, et al.

2018; Peel, et al. 2022; Seppey, et al. 2019; Tian, et al. 2022). Comparison genomes were chosen to represent a variety of marsupial lineages and assembly approaches released within the past 4 years. Genome assembly metrics (Fig. 1d, Supplementary Table 2) were calculated using the stats.sh script in the BBmap package (v37.93**)** (Bushnell 2014).

Heterozygosity

To calculate heterozygosity across species, short reads were aligned to each genome assembly with bwa-mem2 (-M flag; Supplementary Table 4) (Vasimuddin, et al. 2019). Samtools v1.11 was used to filter alignments (view -F 3340 -f 3) and remove duplicates (fixmate -m, markdup -r -S) (Li, et al. 2009). Pileups and variant filtering were performed using bcftools v1.11 mpileup (-q 20 -Q 20 -C 50) call (-m) and view (QUAL > 20, && DP>N && DP<M, where N and M represented 0.5x and 2x the average alignment coverage post-filtering) (Danecek, et al. 2021). Variants within repeats were identified with Red v2.0 and excluded using bedtools v2.27.1, due to low accuracy of read mapping within such regions (Girgis 2015; Quinlan and Hall 2010). This approach was applied to all genomes for this analysis rather than RepeatMasker alone, as Red has similar masking sensitivity to RepeatMasker with orders-of-magnitude lower computational overhead (Girgis 2015). Per-base heterozygosity was taken as the quotient of heterozygous positions and total callable genomic positions (Fig. 2b).

Acknowledgements

We thank the Tasmanian Museum and Art Gallery and Museums Victoria for the use of the images in Fig. 1b & c respectively. This work is supported by Discovery Projects DP210102645 and DP210100505 from the Australian Research Council to AJP and SRF. CYF is supported on Ruth L. Kirschstein National Research Service Award 1F32GM139240-01 by the National Institute of General Medicinal Sciences of the National Institutes of Health. CYF performed all

genomic analyses. AJP and SRF collected and sequenced fat-tailed dunnart samples used in heterozygosity analyses. CYF wrote the manuscript with input from all authors. We thank Elise Ireland for proofreading.

Data Availability

Thylacine assembly, reads and inferred transcripts are submitted under NCBI BioProject PRJNA354646. Dunnart reads have been submitted to NCBI under SUB11101552.

Literature Cited

Alonge M, et al. 2021. Automated assembly scaffolding elevates a new tomato system for high-throughput genome editing. bioRxiv: 2021.2011.2018.469135. doi: 10.1101/2021.11.18.469135

Alonge M, et al. 2019. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. Genome Biol 20: 224. doi: 10.1186/s13059-019-1829-6

Bininda-Emonds OR, et al. 2007. The delayed rise of present-day mammals. Nature 446: 507-512. doi: 10.1038/nature05634

Brandies PA, Tang S, Johnson RSP, Hogg CJ, Belov K 2020. The first Antechinus reference genome provides a resource for investigating the genetic basis of semelparity and age-related neuropathologies. Gigabyte 2020: 0. doi: 10.46471/gigabyte.7

Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner. In: Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).

Cook LE, Newton AH, Hipsley CA, Pask AJ 2021. Postnatal development in a marsupial model, the fat-tailed dunnart (Sminthopsis crassicaudata; Dasyuromorphia: Dasyuridae). Communications Biology 4: 1028. doi: 10.1038/s42003-021-02506-2

Danecek P, et al. 2021. Twelve years of SAMtools and BCFtools. GigaScience 10. doi: 10.1093/gigascience/giab008

228 Deakin JE 2018. Chromosome Evolution in Marsupials. Genes 9: 72. doi:
229 10.3390/genes9020072
230 Dudchenko O, et al. 2017. De novo assembly of the Aedes aegypti genome using Hi-C yields
231 chromosome-length scaffolds. Science. doi: 10.1126/science.aal3327
232 Ellinghaus D, Kurtz S, Willhoeft U 2008. LTRharvest, an efficient and flexible software for de
233 novo detection of LTR retrotransposons. BMC Bioinformatics 9: 18. doi: 10.1186/1471-2105-9-
234 18
235 Fancourt BA 2016. Diagnosing species decline: a contextual review of threats,causes and future
236 directions for management and conservation of the eastern quoll. Wildlife Research 43: 197-211.
237 doi: https://doi.org/10.1071/WR15188
238 Feigin CY, et al. 2018. Genome of the Tasmanian tiger provides insights into the evolution and
239 demography of an extinct marsupial carnivore. Nat Ecol Evol 2: 182-192. doi: 10.1038/s41559-
240 017-0417-y
241 Feigin CY, Newton AH, Pask AJ 2019. Widespread cis-regulatory convergence between the
242 extinct Tasmanian tiger and gray wolf. Genome Research 29: 1648-1658.
243 Flynn JM, et al. 2020. RepeatModeler2 for automated genomic discovery of transposable
244 element families. Proceedings of the National Academy of Sciences 117: 9451-9457. doi:
245 10.1073/pnas.1921046117
246 Garrett Vieira F, Samaniego Castruita JA, Gilbert MTP 2020. Using in silico predicted ancestral
247 genomes to improve the efficiency of paleogenome reconstruction. Ecology and Evolution 10:
248 12700-12709. doi: 10.1002/ece3.6925
249 Girgis HZ 2015. Red: an intelligent, rapid, accurate tool for detecting repeats de-novo on the
250 genomic scale. BMC Bioinformatics 16: 227. doi: 10.1186/s12859-015-0654-5

Hubley R, et al. 2016. The Dfam database of repetitive DNA families. Nucleic Acids Research 44: D81-D89. doi: 10.1093/nar/gkv1272

Johnson RN, et al. 2018. Adaptation and conservation insights from the koala genome. Nature Genetics 50: 1102-1111. doi: 10.1038/s41588-018-0153-5

Kimura M 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120. doi: 10.1007/bf01731581

Lambeck K, Chappell J 2001. Sea Level Change Through the Last Glacial Cycle. Science 292: 679-686. doi: doi:10.1126/science.1059549

Li D, Liu C-M, Luo R, Sadakane K, Lam T-W 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31: 1674-1676. doi: 10.1093/bioinformatics/btv033

Li H 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34: 3094-3100. doi: 10.1093/bioinformatics/bty191

Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079. doi: 10.1093/bioinformatics/btp352

Miller W, et al. 2009. The mitochondrial genome sequence of the Tasmanian tiger (Thylacinus cynocephalus). Genome Research 19: 213-220. doi: 10.1101/gr.082628.108

Mitchell KJ, et al. 2014. Molecular Phylogeny, Biogeography, and Habitat Preference Evolution of Marsupials. Molecular Biology And Evolution 31: 2322-2330. doi: 10.1093/molbev/msu176

Newton AH, et al. 2018. Letting the ‘cat’ out of the bag: pouch young development of the extinct Tasmanian tiger revealed by X-ray computed tomography. Royal Society Open Science 5. doi: 10.1098/rsos.171914

273 Newton AH, Weisbecker V, Pask AJ, Hipsley CA 2021. Ontogenetic origins of cranial
274 convergence between the extinct marsupial thylacine and placental gray wolf. Communications
275 Biology 4: 51. doi: 10.1038/s42003-020-01569-x
276 O'Leary NA, et al. 2016. Reference sequence (RefSeq) database at NCBI: current status,
277 taxonomic expansion, and functional annotation. Nucleic Acids Res 44: D733-745. doi:
278 10.1093/nar/gkv1189
279 Paddle RN. 2000. The last Tasmanian tiger : the history, extinction and myth of the thylacine:
280 Oakleigh, Vic. : Cambridge University Press, 2000.
281 Peel E, Frankenberg S, Hogg CJ, Pask A, Belov K 2021. Annotation of immune genes in the
282 extinct thylacine (Thylacinus cynocephalus). Immunogenetics 73: 263-275. doi:
283 10.1007/s00251-020-01197-z
284 Peel E, et al. 2022. Genome assembly of the numbat (Myrmecobius fasciatus), the only
285 termitivorous marsupial. bioRxiv: 2022.2002.2013.480287. doi: 10.1101/2022.02.13.480287
286 Prowse TA, Johnson CN, Bradshaw CJ, Brook BW 2014. An ecological regime shift resulting
287 from disrupted predator-prey interactions in Holocene Australia. Ecology 95: 693-702. doi:
288 10.1890/13-0746.1
289 Pryszcz LP, Gabaldón T 2016. Redundans: an assembly pipeline for highly heterozygous
290 genomes. Nucleic Acids Research 44: e113-e113. doi: 10.1093/nar/gkw294
291 Quinlan AR, Hall IM 2010. BEDTools: a flexible suite of utilities for comparing genomic
292 features. Bioinformatics 26: 841-842. doi: 10.1093/bioinformatics/btq033
293 Rofe R, Hayman D 1985. G-banding evidence for a conserved complement in the Marsupialia.
294 Cytogenet Cell Genet 39: 40-50. doi: 10.1159/000132101

295 Rovinsky DS, Evans AR, Adams JW 2021. Functional ecological convergence between the
296 thylacine and small prey-focused canids. BMC Ecology and Evolution 21: 58. doi:
297 10.1186/s12862-021-01788-8
298 Schiffels S, Durbin R 2014. Inferring human population size and separation history from
299 multiple genome sequences. Nature Genetics 46: 919-925. doi: 10.1038/ng.3015
300 Seppey M, Manni M, Zdobnov EM 2019. BUSCO: Assessing Genome Assembly and
301 Annotation Completeness. Methods Mol Biol 1962: 227-245. doi: 10.1007/978-1-4939-9173-
302 0_14
303 Shumate A, Salzberg SL 2021. Liftoff: accurate mapping of gene annotations. Bioinformatics.
304 doi: 10.1093/bioinformatics/btaa1016
305 Sleightholme SR, Campbell CR 2018. The International Thylacine Specimen Database (6th
306 Revision-Project Summary & Final Report). Australian Zoologist 39: 480-512.
307 Spencer PBS, McConnell S, Prada D, Friend JA 2020. Parentage assignment using microsatellite
308 DNA typing for the endangered numbat (Myrmecobius fasciatus). Australian Mammalogy 42:
309 240-243. doi: https://doi.org/10.1071/AM19046
310 Stahlke AR, et al. 2021. Contemporary and historical selection in Tasmanian devils (Sarcophilus
311 harrisii) support novel, polygenic response to transmissible cancer. Proc Biol Sci 288: 20210577.
312 doi: 10.1098/rspb.2021.0577
313 Tarailo-Graovac M, Chen N 2009. Using RepeatMasker to identify repetitive elements in
314 genomic sequences. Curr Protoc Bioinformatics Chapter 4: Unit 4.10. doi:
315 10.1002/0471250953.bi0410s25

316 Tian R, et al. 2022. A chromosome-level genome of Antechinus flavipes provides a reference for
317 an Australian marsupial genus with male death after mating. Mol Ecol Resour 22: 740-754. doi:
318 10.1111/1755-0998.13501
319 Vasimuddin M, Misra S, Li H, Aluru S editors. 2019 IEEE International Parallel and Distributed
320 Processing Symposium (IPDPS). 2019 20-24 May 2019.
321 White LC, Saltré F, Bradshaw CJA, Austin JJ 2018. High-quality fossil dates support a
322 synchronous, Late Holocene extinction of devils and thylacines in mainland Australia. Biology
323 Letters 14: 20170642. doi: doi:10.1098/rsbl.2017.0642
324 Woinarski JCZ, Burbidge AA, Harrison PL 2015. Ongoing unraveling of a continental fauna:
325 Decline and extinction of Australian mammals since European settlement. Proceedings of the
326 National Academy of Sciences 112: 4531-4540. doi: 10.1073/pnas.1417301112
327 Wright BR, et al. 2020. A demonstration of conservation genomics for threatened species
328 management. Mol Ecol Resour 20: 1526-1541. doi: 10.1111/1755-0998.13211
329 Wroe S, Clausen P, McHenry C, Moreno K, Cunningham E 2007. Computer simulation of
330 feeding behaviour in the thylacine and dingo as a novel test for convergence and niche overlap.
331 Proceedings. Biological Sciences 274: 2819-2828.