Received: 22 November 2022 | Revised: 16 February 2023 | Accepted: 2 March 2023

DOI: 10.1111/1755-0998.13780

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Genomes of two Extinct-in-the-Wild reptiles from Christmas Island reveal distinct evolutionary histories and conservation insights

Tristram O. Dodge[1,2,3,4] | Katherine A. Farquharson[1,2] | Claire Ford[5] | Lisa Cavanagh[5] | Kristen Schubert[6] | Molly Schumer[3] | Katherine Belov[1,2] | Carolyn J. Hogg[1,2]

[1]School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales, Australia

[2]Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Sydney, Sydney, New South Wales, Australia

[3]Department of Biology, Stanford University, Stanford, California, USA

[4]Australian-American Fulbright Commission, Deakin, Australian Capital Territory, Australia

[5]Taronga Conservation Society Australia, Mosman, New South Wales, Australia

[6]Parks Australia, Christmas Island, Indian Ocean, Australia

Correspondence
Tristram O. Dodge, Department of Biology, Stanford University, Stanford, CA 94305, USA.
Email: todotdge@stanford.edu

Carolyn J. Hogg, School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia.
Email: carolyn.hogg@sydney.edu.au

Funding information
Australian Research Council Centre of Excellence for Innovations in Peptide & Protein Science, Grant/Award Number: CE200100012; Threatened Species Initiative

Handling Editor: Alison Nazareno

Abstract

Genomics can play important roles in biodiversity conservation, especially for Extinct-in-the-Wild species where genetic factors greatly influence risk of total extinction and probability of successful reintroductions. The Christmas Island blue-tailed skink (*Cryptoblepharus egeriae*) and Lister's gecko (*Lepidodactylus listeri*) are two endemic reptile species that went extinct in the wild shortly after the introduction of a predatory snake. After a decade of management, captive populations have expanded from 66 skinks and 43 geckos to several thousand individuals; however, little is known about patterns of genetic variation in these species. Here, we use PacBio HiFi long-read and Hi-C sequencing to generate highly contiguous reference genomes for both reptiles, including the XY chromosome pair in the skink. We then analyse patterns of genetic diversity to infer ancient demography and more recent histories of inbreeding. We observe high genome-wide heterozygosity in the skink (0.007 heterozygous sites per base-pair) and gecko (0.005), consistent with large historical population sizes. However, nearly 10% of the blue-tailed skink reference genome falls within long (>1 Mb) runs of homozygosity (ROH), resulting in homozygosity at all major histocompatibility complex (MHC) loci. In contrast, we detect a single ROH in Lister's gecko. We infer from the ROH lengths that related skinks may have established the captive populations. Despite a shared recent extinction in the wild, our results suggest important differences in these species' histories and implications for management. We show how reference genomes can contribute evolutionary and conservation insights, and we provide resources for future population-level and comparative genomic studies in reptiles.

KEYWORDS
Cryptoblepharus egeriae, *Lepidodactylus listeri*, major histocompatibility complex, runs of homozygosity, sex chromosome

1 | INTRODUCTION

Many species are threatened by human-driven extinction, yet reptiles have historically received less conservation attention as compared to other vertebrate groups (Gibbons et al., 2000). Recently, the first global assessment of reptile extinction risk found 21% of species were broadly considered threatened (Cox et al., 2022). "Extinct-in-the-Wild" (EW), defined as species "known only to survive in cultivation, in captivity or as a naturalized population (or populations) well outside the past range", is the final International Union for the Conservation of Nature (IUCN) category before total extinction (IUCN, 2023). A major conservation goal for EW species is reintroduction to the wild once extrinsic threats to their survival have abated (Smith et al., 2023). However, while in captivity, EW species often face additional genetic threats including loss of diversity (Russello & Jensen, 2021), increased frequency of deleterious variants (Ralls et al., 2000), and inbreeding (Robinson et al., 2021), which can reduce probability of successful reintroduction to the wild and increase the risk of total extinction.

Genetic management of threatened species is key to preserving fitness and future adaptive potential. Pedigree-based approaches have been employed for decades to avoid inbreeding and maximize retention of genetic diversity, but these approaches are imprecise due to linkage and make assumptions regarding equal relatedness of founders that are often violated in practice (Hogg et al., 2019; Knief et al., 2017). Genomic data can have multiple benefits in threatened species management, including the ability to ascertain relatedness of unpedigreed individuals and to directly assess genetic diversity over time (Russello & Jensen, 2021). In addition, it is possible to infer other information relevant to conservation from genomic data including historical demography (Humble et al., 2020), inbreeding levels (Robinson et al., 2021), genetic load (Dussex et al., 2021), disease susceptibility (Lok et al., 2022), and population recovery potential (Robinson et al., 2022).

Such genomic investigations are aided by high-quality reference genomes, which allow for a more comprehensive measurement of genetic diversity and the opportunity to link genotype to phenotype, including in complex regions of the genome (Brandies et al., 2019; Formenti et al., 2022; Paez et al., 2022). While reference genomes from related species can be used for variant discovery and for some population genetic analyses (Galla et al., 2019), this approach may miss species-specific information, especially in parts of the genome where important evolutionary transitions occur, such as rapid sex chromosome turnover that can occur in reptiles, fish, and amphibians (Bachtrog et al., 2014).

Past work generating reference genomes for threatened species has typically relied on additional resequencing of multiple individuals in order to perform analyses relevant to conservation (e.g., Dussex et al., 2021; Humble et al., 2020; Robinson et al., 2021; Wilder et al., 2022; but see Morin et al., 2021; Sutton et al., 2018). However, for EW species, where traditional population-level studies may be difficult due to challenges in acquiring samples, it is valuable to investigate how much information can be obtained from a single, high-quality reference genome. Additionally, some commonly used genomic data types in conservation, such as reduced representation sequencing (e.g., ddRAD, GBS, DArT) and SNP arrays, are limited in their ability to detect informative local patterns of genomic diversity including runs of homozygosity (ROH; Ceballos et al., 2018; Lavanchy & Goudet, 2023). By analysing the same highly accurate long-read sequence data used to assemble the reference genome, it is possible to detect such patterns and thus better design future studies at the population-level. Such studies on maintaining genetic diversity are especially important in EW species, given there is no opportunity to introduce genetic variation from wild populations (i.e., genetic rescue, Frankham, 2015).

Only two reptile species are currently recognized as EW by the IUCN (Smith et al., 2023), both of which are endemic to Christmas Island, Australia (Figure 1a). The Christmas Island blue-tailed skink (*Cryptoblepharus egeriae*; Boulenger, 1888; Figure 1b) was once abundant and occupied diverse habitats across the island (Cogger et al., 1983; Gibson-Hill, 1947). In contrast, Lister's gecko (*Lepidodactylus listeri*; Boulenger, 1888; Figure 1c) was an infrequently observed arboreal specialist restricted to mature rainforest (Cogger et al., 1983; Gibson-Hill, 1947). Both species persisted across Christmas Island for a century after permanent human settlement in the late 1880s, with the skink becoming "hyperabundant" in gardens and phosphate mines (Cogger et al., 1983; Emery, Mitchell, et al., 2021). However, shortly after the introduction of the common wolf snake (*Lycodon capucinus*) in the mid-1980s (Smith, 1988), both species began to decline dramatically (Cogger & Sadlier, 1999). Predation by the wolf snake is considered the most likely driver of native reptile declines on Christmas Island—notably, the last blue-tailed skinks recorded in the wild were found in snake stomachs (Smith et al., 2012)—although predation by introduced giant centipedes (*Scolopendra subspinipes*) and habitat modification by humans and invasive yellow crazy ants (*Anoplolepis gracilipes*) may have contributed (Emery, Mitchell, et al., 2021). Twenty years after declines were first reported, both species were limited to one location on Christmas Island (Figure 1a; Smith et al., 2012). From 2009 to 2010, 66 blue-tailed skinks and 43 Lister's geckos were collected to establish two independently managed captive populations, one on Christmas Island and the other at Taronga Zoo in Sydney, Australia (Smith et al., 2012). Both species were listed as EW in 2014 (IUCN, 2023).

A priority is protecting both species from extinction, with the aim of one day reintroducing self-sustaining wild populations to Christmas Island. Over the past decade, populations of the blue-tailed skink and Lister's gecko have increased under captive management (Andrew et al., 2018). However, immediate threats to their survival in captivity remain. In 2014, an outbreak of the bacterium *Enterococcus lacertideformus* at the Christmas Island breeding facility led to mortality in both species and initiated ongoing biosecurity protocols (Agius et al., 2021; Rose et al., 2017). Furthermore, the inability to manage invasive predators makes a wild release back on Christmas Island not yet feasible (Emery, Valentine, et al., 2021). However, between 2019 and 2021, the blue-tailed skink was

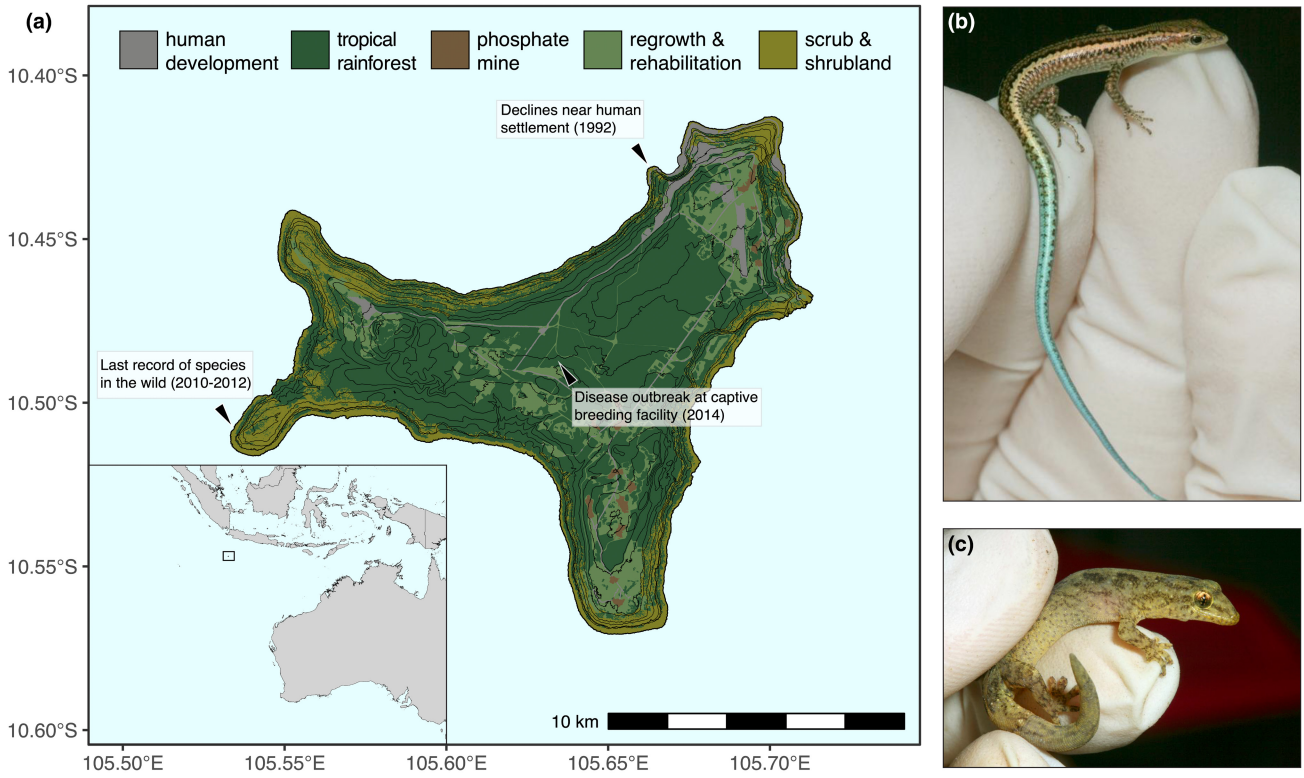

FIGURE 1 Christmas Island and its EW reptiles. (a) Vegetation cover and human-modification of Christmas Island (data from catalogue.data.wa.gov.au) and timeline of key events. Location of Christmas Island in map inset. (b) Photograph of Christmas Island blue-tailed skink (*Cryptoblepharus egeriae*) at Taronga Zoo (provided by Lisa Cavanagh). (c) Photograph of Lister's gecko (*Lepidodactylus listeri*) at Taronga Zoo (provided by Lisa Cavanagh).

introduced to two small islets in the Cocos Archipelago, 1000 km southwest of Christmas Island in an assisted colonization trial. Introduction of Lister's gecko to the Cocos continues to be evaluated for suitability due to the unresolved status of the resident gecko *Lepidodactylus lugubris*.

Although the conservation breeding program has used pedigrees to genetically manage blue-tailed skinks and Lister's geckos for over a decade, no studies of genome-wide variation have been conducted. More generally, development of genomic resources for reptiles has been sparse and uneven (Geneva et al., 2022): to date, only 1% of reptile species have published reference genomes available on NCBI and many taxonomic groups lack a sequenced representative. For instance, skinks (Family: Scincidae) represent 15% of all extant reptiles (Uetz et al., 2022) but are 196 million years diverged from the closest published reference genomes (Zheng & Wiens, 2016). Geckos (Infraorder: Gekkota) represent another 18% of all reptile species (Uetz et al., 2022), but only a single published chromosome-level genome currently exists for the clade (Pinto et al., 2022). Thus, development of high-quality reference genomes for both species would be valuable for both conservation and evolutionary studies.

Here, we assemble and analyse high-quality reference genomes for the Christmas Island blue-tailed skink and the Lister's gecko. Despite both being endemic reptiles from Christmas Island that share a recent history of extinction in the wild and a decade of management in captivity, patterns of diversity across their genomes reveal important differences in their histories and may suggest alternative conservation approaches. We provide a demonstration of the value of a single high-quality genomic resource to investigate multiple conservation and evolutionary questions.

2 | MATERIALS AND METHODS

2.1 | Study system

Christmas Island is a 135 km^2 seamount in the Indian Ocean, located 360 km south of Java (Figure 1a). The island is a drowned coral atoll that reemerged an estimated 5.03 (4.49–5.66) million years before present (Ali & Aitchison, 2020). At the time of first permanent human settlement in the late 19th century, the island was primarily vegetated by tropical rainforest (Figure 1a; Andrews, 1899). However, 25% of the island has since been cleared for phosphate mining operations (Smith et al., 2012). All five of Christmas Island's endemic reptile species are currently either Extinct, Extinct-in-the-Wild, or Endangered (James et al., 2019). The blue-tailed skink is estimated to be 7 million years diverged from its closest known relatives in the mainland Australian *Cryptoblepharus metallicus* group (Oliver et al., 2018). Lister's gecko is estimated to be 23–26 million years diverged from its closest known relatives in the *Lepidodactylus*

lugubris group (Oliver et al., 2018), but this radiation has not been fully molecularly characterized (McDonald et al., 2022).

In 2011, a year after the Christmas Island conservation population was established, 83 blue-tailed skinks (19 wild-born founders and 64 first-generation captive-born descendants) and 52 Lister's geckos (22 wild-born founders and 30 first-generation captive-born descendants) were transferred from the Christmas Island facility to Taronga Zoo in Sydney, Australia to start a second, independently-managed population (Andrew et al., 2018). At Taronga Zoo, blue-tailed skinks have been managed under a Maximum Avoidance of Inbreeding (MAI) scheme (Princée, 1995), while a pedigree-based minimization of kinship approach has been used for Lister's geckos (Andrew et al., 2018). The Taronga populations have expanded over the last decade to a maximum of 425 skinks in October 2018 and 190 geckos in April 2021. The populations on Christmas Island are both managed under MAI and peaked at more than 1417 skinks in August 2017 and 1430 geckos in December 2018. The mean generation time in captivity for both species in the Taronga population is 3–4 years and is expected to be similar for the Christmas Island population (C. Ford, personal communication).

2.2 | Sampling, extraction, and sequencing

Due to their rarity, all individuals were sampled opportunistically after euthanasia for medical reasons. A full list of individuals sequenced is available in Table S1. All sampling and extractions were undertaken at the Taronga Zoo facilities in accordance with Approved Arrangement 7.9 for Zoo Animal Requirements, Department of Agriculture, Water, and the Environment (DAWE) biosecurity protocols. Samples were released from quarantine with permission from DAWE.

We conducted high molecular weight (HMW) DNA extractions from heart, kidney, and muscle tissue for a single male from each species using the Circulomics Nanobind Tissue Big DNA Kit (NB-900-701-001). We submitted pooled HMW DNA to the Australian Genome Research Facility (Brisbane, Australia) for PacBio HiFi library preparation and sequencing. HMW DNA was sheared to 15–20 kb fragments using a Megaruptor2 (Diagenode), and fragments longer than 15 kb were selected with a BluePippin (Sage Science). A 0.5× ampurePB bead clean-up followed by a DNA repair protocol with NEBNext FFPE DNA Repair Mix treatment (NEB #M6630S/L) and a second bead clean-up were then undertaken. These fragments were used as input to the SMRTbell Express Template Prep Kit 2.0 to prepare PacBio HiFi libraries. The resulting blue-tailed skink and Lister's gecko libraries were sequenced on a PacBio Sequel II across two and three single-molecule real-time (SMRT) cells, respectively. Sequencing yielded 58.7 Gb of HiFi data for the skink and 73.6 Gb for the gecko.

To increase assembly contiguity, we subsequently generated Hi-C data for Lister's gecko. Hi-C sequencing guidelines recommend flash-frozen tissue; however, due to biosecurity requirements, we preserved an entire male Lister's gecko in 100% ethanol, which allowed it to be released from quarantine. Heart, liver, and tail muscle tissue were subsequently dissected from this individual and stored in 100% ethanol for shipment to the ACRF Biomolecular Resource Facility (The John Curtin School of Medical Research, Australian National University, Canberra, Australia) for Hi-C library preparation using the Arima Hi-C kit, according to the manufacturer's protocols. Prior to library preparation, the tissue samples were washed twice for 5 min with 1×PBS using a rotator wheel at room temperature. The resulting libraries were sequenced on an Illumina NovaSeq 6000 to 150-base pairs (bp) paired-end (PE) length.

To increase genome annotation quality, we generated transcriptomes for both species (Table S2). Immediately following euthanasia, we preserved three tissues (brain, liver, and gonads) in RNAlater and extracted total RNA using the Qiagen RNeasy Plus Mini Kit. Extracted RNA was sequenced at the Ramaciotti Centre for Genomics (University of New South Wales, Sydney, Australia) with TruSeq mRNA library preparation. All tissue libraries were sequenced using an Illumina NovaSeq 6000 to 100-bp PE length.

2.3 | Read quality assessment and filtering

We assessed quality of the long-read sequencing using NanoPlot (De Coster et al., 2018) (Table S3). The blue-tailed skink HiFi data had a mean read length of 12,556.8 bases and quality score of 35.7 (read N50 of 12,885.0), and the Lister's gecko data had a mean read length of 14,441.2 bases and quality score of 32.8 (read N50 of 14,478.0). We used HiFiAdapterFilt (Sim et al., 2022) with default parameters (44 bp and 97% match) to remove residual adapter contamination present in our PacBio HiFi reads. This removed 0.134% of HiFi reads for the skink and 0.141% for the gecko.

We evaluated transcriptome read quality with FastQC version 0.11.8 (Andrews, 2010; https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and multiQC version 1.13 (Ewels et al., 2016; https://github.com/ewels/MultiQC). We used Trimmomatic version 0.39 (Bolger et al., 2014) to remove lower quality reads with the following flags: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 SLIDINGWINDOW:4:5 LEADING:5 TRAILING:5 MINLEN:25. Over 99.9% of reads for both species passed quality filtering.

2.4 | Genome assembly

We assembled the genome with Hifiasm version 0.16.0 (Cheng et al., 2021; https://github.com/chhylp123/hifiasm) using a 256 GB RAM, 64 vCPU, 3TB storage cloud-based machine with the Pawsey Supercomputing Centre, Perth, Australia. We tested several Hifiasm parameter combinations that are known to increase contiguity at the cost of runtime (Cheng et al., 2021; shown in Table S4). To scaffold the Lister's gecko genome, we followed the Arima genomics pipeline for Hi-C read cleaning and alignment (A160156 v02; https://github.com/ArimaGenomics/mapping_pipeline). Briefly, we mapped our paired-end data as single end and used the Arima script to find

successfully mapped pairs. We used the alignments, sorted by name, as input for genome scaffolding with YaHS (Zhou et al., 2023; https://github.com/c-zhou/yahs). We visualized the Hi-C contact matrix in Juicer (Durand et al., 2016; https://github.com/aidenlab/juicer) and made manual edits to correct two misassemblies identified on scaffolds 3 and 21. We assessed assembly contiguity using QUAST (Gurevich et al., 2013; https://quast.sourceforge.net/quast) and assembly completeness with BUSCO version 5.4.3 (Simão et al., 2015; https://busco.ezlab.org/) using the Vertebrata single copy orthologue database (version 10).

We used MitoHiFi version 2.2 (Uliano-Silva et al., 2021; https://github.com/marcelauliano/MitoHiFi) to assemble a mitochondrial genome for each species from the raw HiFi reads. For the skink and gecko, we downloaded a complete mitochondrial genome from the closest available relative (*Eutropis multifasciata* and *L. lugubris*, respectively) as input to MitoHiFi. We annotated and visualized the mitogenomes with MitoZ (Meng et al., 2019; https://github.com/linzhi2013/MitoZ). We then used BLASTn to search for high-identity mitochondrial contaminants in each genome assembly, all of which were present on short contigs (<30 kb), and we subsequently removed these sequences.

2.5 | Synteny

To assess the extent of chromosome evolution, we generated whole-genome alignments between the blue-tailed skink and two chromosome-level squamate reptile assemblies: the green anole (*Anolis carolinensis*; RefSeq: GCF_000090745.1) and the common wall lizard (*Podarcis muralis*; RefSeq: GCF_004329235.1) using the approximate mapping feature of minimap2 version 2.24-r1122 (Li, 2018; https://github.com/lh3/minimap2). We also generated an alignment between Lister's gecko and the chromosome-level gecko assembly for *Sphaerodactylus townsendi* (Pinto et al., 2022).

2.6 | Annotation

Due to their divergence from other reptile genomes, we built a custom repeat library for each species. We used RepeatModeler version 2.0.1 (Flynn et al., 2020) with the NCBI engine to model the repeats and RepeatMasker version 4.0.9 (Smit & Green, 2013–2015) to identify and mask repetitive regions in each genome. We hard-masked the genomes, excluding simple repeats (--nolow), and soft-masked the genomes (--xsmall) for subsequent downstream analyses. We identified telomeric repeats using RepeatMasker to discover regions matching canonical telomere hexamer sequences (e.g., (TTAGGG)n) and restricted our search to repeats greater than 1000 bp.

To assist with gene annotation, we generated a global transcriptome for each species using three tissues: brain, gonads, and liver on the University of Sydney's Artemis High Performance Computer. RNA reads were mapped to the hard-masked genomes with HISAT2 version 2.1.0 (Kim et al., 2019), using the paired-end read options (specifying -1, -2, and --rna-strandness RF) and the -dta parameter for downstream transcriptome assembly. Aligned reads were merged into tissue-specific transcriptomes with StringTie version 2.1.6 (Shumate et al., 2022), combined into a global transcriptome with TAMA (Kuo et al., 2020) merge, and coding regions were predicted with TransDecoder version 2.0.1 (Haas et al., 2013) following the approach of Peel et al. (2022). We evaluated the quality of these global transcriptomes with BUSCO run in transcriptome mode using the Vertebrata (v10) database. We performed gene annotation with Fgenesh++ version 7.2.2 (Softberry; Solovyev et al., 2006) on the Pawsey Supercomputing Centre's cloud machine using the longest open reading frame for each predicted gene in the global transcriptome, nonmammalian settings, and optimized parameters with the supplied *A. carolinensis* gene-finding matrix. Because automated annotation pipelines may not characterize immune genes well (Peel et al., 2022), we performed manual annotation of the Toll-like receptors (TLR) and major histocompatibility complex (MHC) families using BLASTn searches, using 82 complete reptile TLR transcripts and 34 MHC transcripts available on NCBI (Table S5).

2.7 | Identifying the sex chromosomes

Sex determination systems are highly variable in reptiles, and while both Christmas Island species are thought to have genetic sex determination (L. Cavanagh, personal communication), their sex chromosomes are unknown. We took several complementary approaches to identify sex chromosomes. First, we assessed read-depth (coverage) across the genome to identify hemizygous regions, expected to be present in the heterogametic sex of individuals with somewhat divergent sex chromosomes. We mapped HiFi reads back to each genome with minimap2 --map-hifi and removed secondary alignments with the samtools(Danecek et al., 2021) --x100 flag. We calculated coverage in nonoverlapping 1 Mb windows using BEDTools version 2.29.2 (Quinlan & Hall, 2010) genomecov. However, because alterations in coverage can represent structural variants unlinked to sex, we also performed BLASTn searches using a list of previously identified X-linked genes shared across Scincidae (Kostmann et al., 2021). We also performed BLASTn searches with Y-linked markers from the skinks *Bassiana duperreyi* (Dissanayake et al., 2020) and *Eulamprus heatwolei* (Cornejo-Páramo et al., 2020). A full list of query sequences for BLASTn searches are available in Table S6. Finally, we used minimap2 to create a whole-genome alignment of the alternate haplotigs generated with hifiasm.

2.8 | Variant calling and filtration

From the PacBio HiFi reads mapped back to the genome assemblies, we called variants using GATK4 version 4.2.0.0 (McKenna et al., 2010) HaplotypeCaller and GenotypeGVCFs with the --all-sites parameter to include invariant sites. We filtered variant and invariant classes separately to remove low confidence genotype calls, following

GATK best practices with two modifications: depth filters below 0.66× average coverage and above 1.5× average coverage, as well as QD (QualByDepth) >10. We chose more stringent depth filters than short-read best practices because coverage across the genome was relatively even, and these filters would exclude most hemizygous (0.5× coverage) or duplicated (2× coverage) regions. Additionally, the distribution of QD values was high and the QD >10 filter removed 2% of total variants but removed most variants with skewed allele frequencies. From the resulting variant calls, we calculated genome-wide heterozygosity for all called sites ((SNP + INDEL)/(SNP + INDEL + INVARIANT)). For local analyses of genetic diversity, we used BEDTools to count the number of variants across the genome in nonoverlapping windows containing 1 million and 10,000 called sites.

2.9 | Inferring historical population sizes

We used the pairwise sequentially Markovian coalescent (PSMC; Li & Durbin, 2011) method to infer historic effective population sizes from variants called with GATK4. We tested several modifications of the time-segment parameter -p, but focus the results of PSMC with parameters -N 25 -t 12 -r 5 -p "4 + 30*2 + 4 + 6 + 10" (Schield et al., 2022) in the main text. We made two main modifications to use PSMC with GATK variants called from HiFi reads. First, because the majority of HiFi reads mapped uniquely, we did not make a mappability mask. We did, however, mask homozygous regions greater than 1 Mb (see Section 3). Second, we kept stringent variant filters based on depth and GQ score, but we kept all reference genotype calls because GATK does not calculate the required FQ field. Generation time (θ) and mutation rate (ρ) are required to scale PSMC results. We used a generation time of 3.5 years for both species (Director of National Parks, 2012). Because mutation rate is not well understood in both skinks and geckos, we used 6.125×10^{-9} per site per generation for both species, calculated from the phylogenetically inferred median squamate mutation rate of 1.75×10^{-9} per site per year (Gemmell et al., 2020). We performed 100 bootstrap replicates to assess variability in our N_e estimates using PSMC.

2.10 | Identifying runs of homozygosity (ROH)

Various approaches exist to detect long ROH (>1 Mb), a genomic signature of recent inbreeding; however, it was unclear how such approaches would perform on PacBio HiFi reads mapped back to a reference genome from the same individual. We developed an observational approach to detect ROH in both genome assemblies from the same set of variants called with GATK4. We first excluded INDELs because sequencing errors of this type are more likely to occur in PacBio HiFi reads (Nurk et al., 2020). We also masked the hemizygous regions, which can lead to high ROH calls for reasons other than inbreeding. We then counted SNPs in nonoverlapping windows containing 10,000 called sites (SNP + INVARIANT). The mean 10,000-site window had 61 SNPs for blue-tailed skink and 43 SNPs for Lister's gecko. We called ROH by identifying stretches of ≥100 consecutive windows with fewer than two SNPs per window. To avoid spuriously breaking up long ROH due to erroneous heterozygous calls caused by sequencing or mapping errors in repetitive regions, we implemented a smoothing step whereby adjacent ROH were joined if they were separated by a single window, even if this window had >2 SNPs. While the distribution of ROH lengths is continuous, we focus on ROH >1 Mb, because 1 Mb is a common threshold for analyses of recent inbreeding and makes our reported values more easily comparable with other recent conservation studies (Iannucci et al., 2021; Robinson et al., 2021; Wilder et al., 2022). To assess the robustness of our developed method, we also called ROH with the homozyg function in PLINK version 1.90b6.5 (Purcell et al., 2007; www.cog-genomics.org/plink/1.9/).

To understand if the identity by descent (IBD) blocks detected as ROH originated (i.e., coalesced) before or after the captive breeding program began, we used mean ROH length to estimate the time since these tracts arose (following the approach of Robinson et al., 2021). To infer number of generations since IBD, we used the equation $l = 100 / 2g$, where l is observed ROH length (in cM) and g is the number of generations since coalescence (Thompson, 2013). This model makes the simplifying assumptions that ROH lengths decay exponentially over time and that recombination rate is constant across the genome. The rate of decay over physical distance depends on genetic map length, which is not well characterized in skinks or geckos. To capture this uncertainty, we used two extremes of reported squamate recombination rates. Representing a lower extreme, *A. carolinensis* has on average 23.3 crossovers per meiosis (estimated with foci of MLH1, a DNA repair protein that localizes to mature recombination nodules; Lisachov et al., 2017) and a genome size of 2.15 Gb (Peterson et al., 1994). On the higher reported bound, the Steppe Agama (*Trapelus sanguinolentus*) has on average 38.4 crossovers per meiosis (Lisachov et al., 2019) and an estimated genome size of 1.68 Gb (Vinogradov, 1998). Genetic map lengths in these reptiles were estimated based the equation $m = C / G \times 50$ cM, where m is map distance (in cM/Mb), C is crossovers (MLH1 foci) and G is haploid genome size in Mb (Calderón & Pigozzi, 2006).

3 | RESULTS

3.1 | High quality de novo genome assemblies

We generated contiguous and complete de novo genome assemblies for both the blue-tailed skink and Lister's gecko (Table 1). Using PacBio HiFi long-reads sequenced at ~40× mean genome coverage, the final skink assembly was 1.40 Gb, near the 1.47 Gb haploid size for relative *Cryptoblepharus plagiocephalus* (MacCulloch et al., 1996). The final assembly was 72 contigs with a contig N50 of 109.1 Mb (Figure 2a). 99.0% of the genome was contained in 18 contigs (all >10 Mb), approaching the chromosome number of the blue-tailed

TABLE 1 Genome assembly quality statistics.

	Blue-tailed skink	Lister's gecko
Sequencing technologies	PacBio HiFi	PacBio HiFi + Hi-C
Hifiasm parameters	–k 63 –r 5	–D 20 –k 63 –r 9 --max-kocc 10000 –a 6 –s 0.65 –N 300 --hom-cov 30
Assembly size	1.40 Gb	2.35 Gb
Number of contigs	72	381
Number of contigs >10 Mb	18	48
Contig N50	109.1 Mb	72.3 Mb
Genome in contigs >10 Mb (%)	99.0%	91.7%
Scaffold number	-	401
Number of scaffolds >10 Mb	-	22
Scaffold N50	-	119.9 Mb
Genome in scaffolds >10 Mb (%)	-	96.2%
Complete vertebrate BUSCOs (%)	97.8%	97.6%

Note: Both Christmas Island reptile genome assemblies were highly contiguous and complete, representing a near chromosome-level assembly for the blue-tailed skink and a chromosome-level assembly for the Lister's gecko.

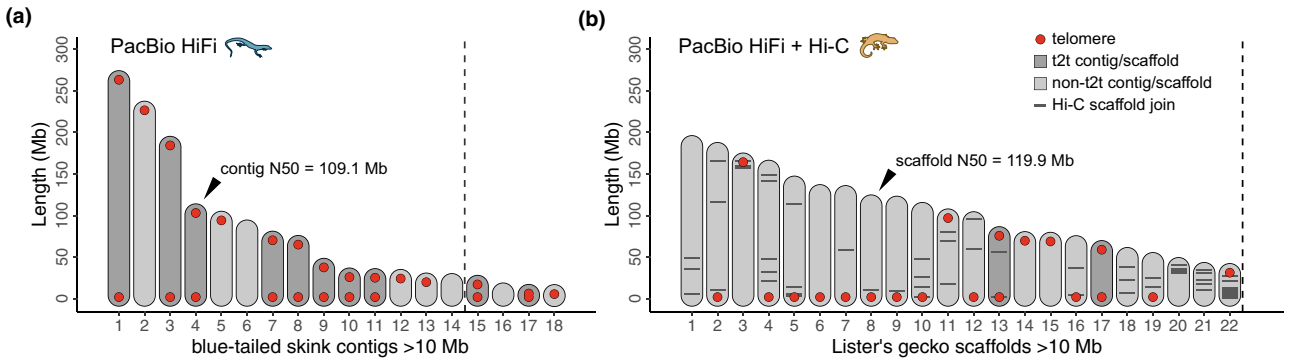

FIGURE 2 Genome assemblies of Christmas Island reptiles. (a) 18 longest contigs of blue-tailed skink genome assembly, representing >99% of assembled sequence, ordered by length. Contig 4 is the shortest contig for which longer and equal length contigs cover ≥50% of the assembly (contig N50). Red dots denote approximate telomere locations. Ten dark grey contigs represent telomere to telomere contigs (i.e., potential chromosomes). Vertical dashed line indicates expected karyotype ($n = 14$) based on closest karyotyped relative *Cryptoblepharus boutonii* (2n=28). (b) Longest 22 scaffolds of Lister's gecko genome assembly, representing >96% of assembled sequence, ordered by length. Scaffold 8 is the shortest scaffold for which longer and equal length scaffolds cover ≥50% of the assembly (scaffold N50). Dark grey scaffolds represent telomere to telomere scaffolds. Horizontal bars denote contig joins based on Hi-C contacts. Vertical dashed line indicates expected karyotype ($n = 22$) based on closest karyotyped relative *Lepidodactylus lugubris* (2n=44 and 3n=66).

skink's closest karyotyped relative *Cryptoblepharus boutonii* (2n=28; Makino & Momma, 1949). Of these 18 contigs, 10 were flanked on both ends by long (>6500 bp) telomere-like repeats, suggesting that these contigs represent full chromosomes (Figure 2a). The blue-tailed skink genome was also highly complete, with 97.8% of vertebrate BUSCOs present and complete (C: 97.8% [S: 97.1%, D: 0.7%], F: 0.7%, M: 1.5%, n: 3354).

For the larger and more complex Lister's gecko genome, and we used PacBio HiFi long reads and Hi-C data to achieve a chromosome-level assembly. The final genome was 2.35 Gb, near average for Gekkonidae (Gregory, 2023). With PacBio HiFi reads, sequenced at ~30× average read-depth, we recovered 381 contigs with a contig N50 of 72.2 Mb. Following Hi-C scaffolding, 96.2% of the Lister's gecko genome was contained in 22 scaffolds (all >30 Mb) with a scaffold N50 of 119.9 Mb (Figure 2b), matching the karyotype of its closest relative, *L. lugubris* (2n=44/3n=66; Volobouev & Pasteur, 1988). Our Hi-C contact map indicated physical proximity within assembled scaffolds, but not between scaffolds (Figure S1). The 22 largest scaffolds had 97.6% of vertebrate BUSCOs present and complete (C: 97.6% [S: 96.3%, D: 1.3%], F: 0.9%, M: 1.5%, n: 3354). The remaining 379 scaffolds contained 14 duplicated BUSCOs and two single-copy BUSCOs not found in the chromosome-level scaffolds. Two Lister's gecko scaffolds were telomere-to-telomere (Figure 2b).

We also assembled mitochondrial genomes for both species. The blue-tailed skink and Lister's gecko mitochondrial genomes

were 17,216 and 16,967 bp long, respectively. Both mitochondrial genomes had the expected 37 genes, including 13 protein-coding genes, two RNAs, and 22 tRNAs (Figure S2).

3.2 | Synteny

Despite nearly 200 million years of divergence, the blue-tailed skink genome displayed strong overall identity with the largest chromosomes of *A. carolinensis*, which are thought to represent an ancestral-like iguanid karyotype (Deakin et al., 2016). The longest three contigs of blue-tailed skink showed identity with three of the four longest *A. carolinensis* chromosomes (Figure S3A). Additionally, contigs 5, 6, and 18 displayed identity with *A. carolinensis* chromosome 2 in an orientation that would create a telomere-to-telomere scaffold in blue-tailed skink. However, contigs shorter than the eighth longest contig in the skink displayed little homology to the *A. carolinensis* genome. There was identity between these smaller skink chromosomes when compared to the genome of the wall lizard *P. muralis* (Figure S3B).

We aligned the Lister's gecko genome to the published chromosome-level genome of the gecko *S. townsendi* (Pinto et al., 2022). Despite having diverged more recently as compared to the blue-tailed skink alignments (119 vs. 196 million years; Zheng & Wiens, 2016), the gecko alignments showed numerous rearrangements, particularly at the putative chromosome ends (Figure S3C), possibly reflecting mapping issues in repetitive, acrocentric regions.

3.3 | Genome annotation

We annotated repetitive regions in both genomes with custom RepeatModeler libraries. The blue-tailed skink assembly was less repetitive than the Lister's gecko assembly, at 44% compared to 55%. Most transposable element (TE) families showed a greater proliferation in the gecko compared to the skink, except for DNA elements (Table S7). A large percentage of repeats in both species were unclassified. In the blue-tailed skink, most telomere-like hexamers were under five repeats long, but 29 of these sequences were greater than 1000 bp, all of which started within 20 bp of a contig end. These 29 putative telomeres had a mean length of 10,066 bp (range: 2153–17,157 bp). In the Lister's gecko, we identified 43 such sequences greater than 1000 bp, all of which started within 135 bp of the end of a scaffold. These 43 putative telomeres in Lister's gecko had a mean length of 10,000 bp (range: 1341–17,822 bp), but we found no evidence that Lister's gecko telomere length was different than that of the blue-tailed skink (Welch's two sample *t*-test, $p = .94$; Figure S4).

Automated gene annotation with FGENESH++, using a combination of transcriptome and protein homology, identified 29,055 genes in the blue-tailed skink and 42,053 genes in Lister's gecko (Table S7). The mean number of exons was 7.84 and 6.82 in the blue-tailed skink and Lister's gecko, respectively. The number of annotated genes is substantially higher than reptile genomes annotated by the NCBI pipeline, a pattern that has previously been described with FGENESH++ software and may result from use of incomplete transcriptomes as evidence (Peel et al., 2021). However, BUSCO analysis of these transcriptomes indicated RNA sequencing captures most core vertebrate genes, with 94.1% of expected genes present and complete for the skink and 92.4% for the gecko. Manual immune gene annotation yielded the expected squamate TLR genes in both species: *TLRs 2-7* and *TLR13*. Using the same approach to annotate the MHC class I and II families, we identified a 3.0 Mb region on contig 2 and a 0.5 Mb region on contig 25 in the blue-tailed skink as well as a nearly 4 Mb region on scaffold 7 containing all MHC class I and II genes in the Lister's gecko.

3.4 | Identification of sex chromosomes

Our read-depth analysis identified a 15 Mb region of half-fold coverage on the eighth longest contig of the blue-tailed skink genome (Figure 3a,b). Such a pattern is consistent with a XY sex determination system but could also be explained by a non-sex-linked structural variant or regions that are difficult to assemble, thus requiring additional evidence. Of 10 X-linked genes shared across multiple skink families (Kostmann et al., 2021), we recovered alignments to six of these transcripts in the blue-tailed skink, all of which fell in this region of half coverage (Figure 3b), identifying this contig as the putative X chromosome. From the blue-tailed skink genome-assembly graph, we recovered a nearly chromosome-length alternate haplotype representing the putative Y chromosome. This region was largely syntenic to the X chromosome but contained a 10 Mb region that was highly divergent from the 15 Mb half-coverage region on the X (Figure 3c). The center of this region showed very little homology between the X and Y, but there was greater sequence identity on both sides despite the presence of a chromosomal inversion (Figure 3c). We found that few Y-linked markers identified in other skinks BLASTed uniquely to the Y chromosome in the blue-tailed skink. Only one of seven Y-linked PCR markers identified in the skink *Bassiana duperreyi* (Dissanayake et al., 2020), and four of 382 Y-linked transcripts in the skink *Eulamprus heatwolei* (Cornejo-Páramo et al., 2020) blasted uniquely to the blue-tailed skink Y chromosome. In contrast, we found multiple small regions of lower coverage in Lister's gecko, particularly on the ends of putative chromosomes (Figure 3d). Given that there were no clear candidate regions based on coverage, we did not pursue additional analyses in the gecko.

3.5 | High heterozygosity and large historical population sizes

After calling and filtering variants with GATK4, we discovered 9.28 million variants for the blue-tailed skink and 9.41 million variants for Lister's gecko. We estimated autosomal heterozygosity to be 0.007 heterozygous sites per base pair in the blue-tailed skink and 0.005

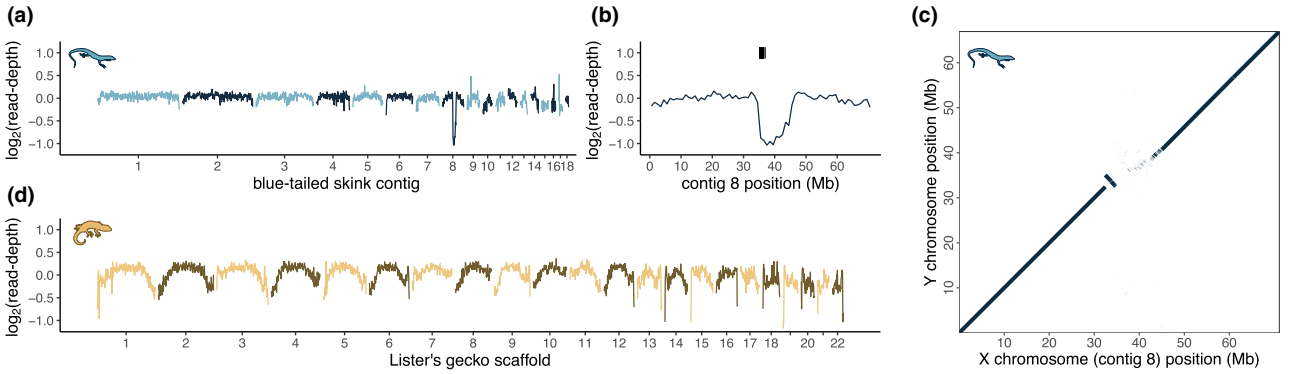

FIGURE 3 Identification of a XY chromosome pair in the blue-tailed skink. (a) Relative read-depth plotted across genome for the blue-tailed skink shows one region of half coverage. Contigs are coloured in an alternating pattern to aid visualization. (b) Relative read-depth across contig 8, the presumptive X chromosome. Black bars above denote BLAST hits for previously identified X-linked genes conserved in across all major skink lineages (Kostmann et al., 2021), which all fall within the region of half read-depth. (c) Alignment between blue-tailed skink X (contig 8) and putative Y chromosome, showing low identity in the centre as well as a chromosomal inversion. (d) Relative read-depth across Lister's gecko genome shows no clear signal of a sex chromosome. Scaffolds are coloured in an alternating pattern to aid visualization. (a and d) are scaled to their respective genome sizes (a: 1.40 Gb, d: 2.35 Gb).

heterozygous sites per base pair in Lister's gecko. In the skink, heterozygosity was relatively uniform across the chromosomes and slightly elevated at the tips except for long stretches of very low heterozygosity (Figure 4a,b, Figure S5). The gecko showed larger variation in heterozygosity across chromosomes with more elevated heterozygosity towards chromosome ends (Figure 4c,d, Figure S6).

From these variant calls, we used PSMC to infer historical effective population sizes (Figure 4e). We inferred that the skink had a larger historical N_e than the gecko, but both were high until relatively recently, peaking at ~965,000 individuals in the skink and ~265,000 in the gecko. However, due to uncertainty in mutation rates and generation times, precise estimates may not be accurate. Additionally, PSMC is known to lose resolution at more recent timescales, e.g. ~20,000 years in humans (Li & Durbin, 2011; Nadachowska-Brzyska et al., 2022), so more recent population size inferences should be interpreted with caution. We also note the overlap in population sizes 5 million years before present does not reflect coalescence between species, as these species diverged some 200 million years ago (Zheng & Wiens, 2016). Our PSMC results were largely robust to different time segment choices (Figure S7).

3.6 | Runs of homozygosity

Using our custom ROH discovery approach, we identified 33 ROH longer than 1 Mb in the skink, covering 135.8 Mb of the genome (F_{ROH} = 0.0963; Figure 4a,f). The mean ROH size in the skink was 4.1 Mb and the longest was 15.1 Mb. In contrast, we detected a single ROH in the gecko (1.7 Mb; F_{ROH} = 0.0007; Figure 4b,f). We uncovered broadly similar patterns using PLINK (Figures S5 and S6); however, many of these ROH appeared spuriously broken or truncated, so we focus here on results from our custom approach. Using the highest and lowest recombination rates known in reptiles and making the simplifying assumption that recombination rate is constant across the genome, we estimate that IBD tracts detected as ROH originated between 11 and 24 generations ago (Figure 4g). A single 6 Mb ROH on contig 2 in the blue-tailed skink genome overlapped with all the MHC class I and II genes on contig 2 and a 7 Mb ROH overlapped with *TLR4* on contig 11 (Figure S8).

4 | DISCUSSION

Here, we generate genomic resources for the Christmas Island blue-tailed skink and Lister's gecko, the only two reptiles currently classified as Extinct-in-the-Wild. These represent some of the highest quality reptile assemblies to date in terms of both contiguity and completeness (Geneva et al., 2022). Our extensive analyses show that long-reads can hold a wealth of information relevant to conservation even when population-level data are not available. Our study marks an important step towards understanding the genomic makeup of two understudied reptile clades, Scincidae and Gekkota, provides insights into the histories of two species of high conservation value, and serves as a springboard for future studies by evolutionary and conservation biologists.

4.1 | Assembly and genome organization

For the smaller and less repetitive blue-tailed skink genome, PacBio HiFi sequencing alone resulted in a nearly chromosome-level assembly, with 10 telomere-to-telomere contigs of the 14 expected chromosome pairs in *Cryptoblepharus* (Makino & Momma, 1949). However, the larger and more repetitive Lister's gecko genome required Hi-C sequencing in addition to PacBio HiFi to achieve chromosome-level scaffolds matching the expected 22 acrocentric chromosome pairs in *Lepidodactylus* (Volobouev & Pasteur, 1988). The Vertebrate Genomes Project aims to generate gapless, telomere-to-telomere

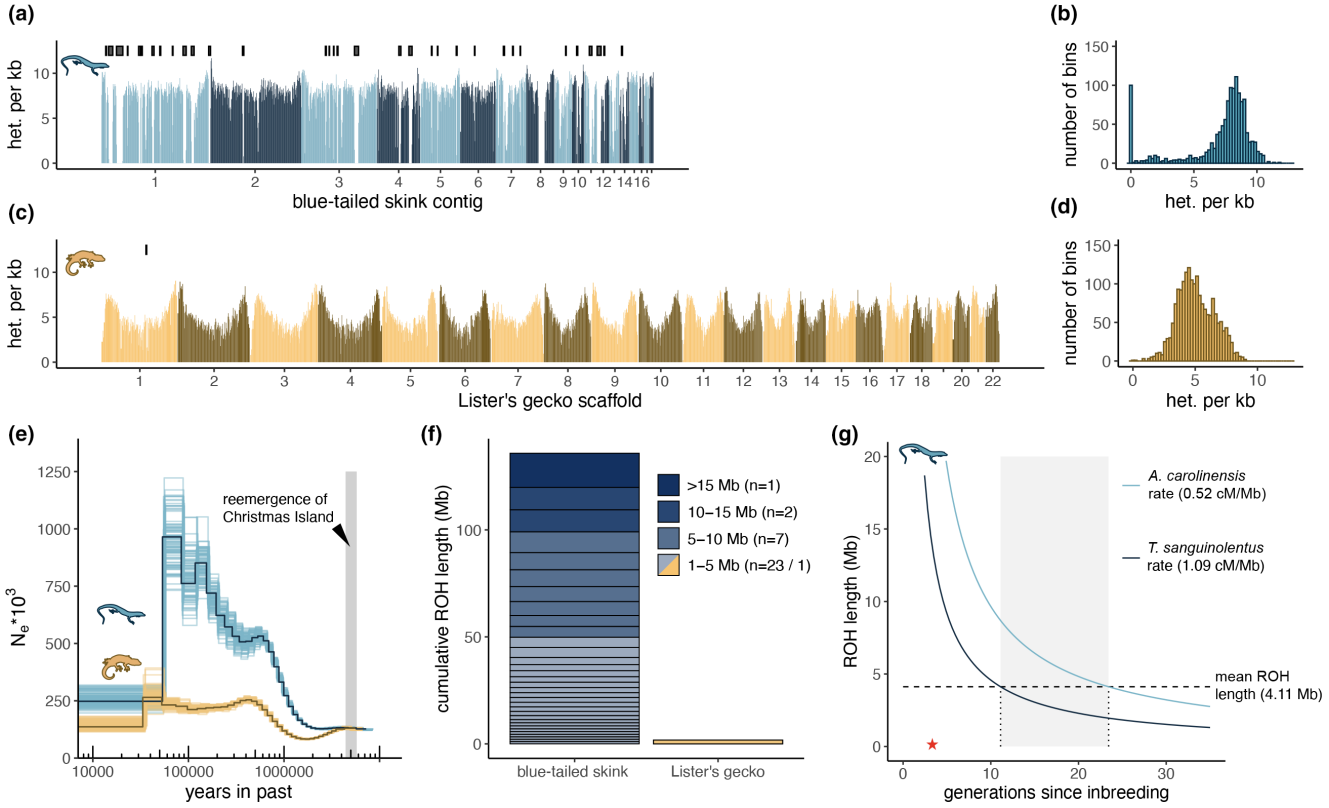

FIGURE 4 Patterns of heterozygosity reveal ancient demography and recent history of inbreeding. (a) Heterozygous sites per kilobase (kb) in bins of 1 million called sites across blue-tailed skink genome. Contigs are coloured in an alternating pattern to aid visualization. Locations of 33 called runs of homozygosity (ROH) displayed in grey blocks above contigs. The 15 Mb X-linked structural variant on contig 8 is masked. (b) Histogram of heterozygous sites per kb in bins of 1 million called sites across skink genome. (c) Heterozygous sites per kb in bins of 1 million called sites across Lister's gecko genome. Scaffolds are coloured in an alternating pattern to aid visualization. Locations of 1 ROH displayed in grey block above scaffold 1. (d) Histogram of heterozygous sites per kb in bins of 1 million called sites across gecko genome. (e) PSMC plot, blue represents skink, yellow represents gecko, with lighter colours denoting 100 bootstraps. PSMC plot is scaled with generation time $\theta = 3.5$ and median squamate mutation rate of $\rho = 6.125 \times 10^{-9}$. Grey bar is estimated reemergence of Christmas Island from Indian Ocean 5 million years before present (Ali & Aitchison, 2020). Overlap in population size ~5 million years ago does not reflect true coalescence, as species diverged >200 million years before present. (f) Cumulative lengths of skink and gecko ROH, coloured by 4 length categories: 1–5 Mb, 5–10 Mb, 10–15 Mb, >15 Mb. (g) Models used to estimate ROH length decay per generation with recombination rates from *Anolis carolinensis* and *Trapelus sanguinolentus*. Grey box shows estimated number of generations ago IBD segments detected as ROH originated. Red star shows average number of generations in captivity for blue-tailed skinks (1.5–3.5 generations based on pedigree). (a and c) are scaled to their respective genome sizes (a: 1.40 Gb, c: 2.35 Gb) to permit ROH comparison.

assemblies of all vertebrate species (Rhie et al., 2021); however, the first human genome meeting these criteria was sequenced only recently from a haploid cell line (Nurk et al., 2022), showcasing the ambition of these goals. Our nearly chromosome-level skink genome, generated with only long-read sequencing, suggests some clades may be less challenging to complete than others.

Long-read assemblies often better resolve repetitive regions of the genome, including transposable elements and telomeres (Nurk et al., 2022; Rhie et al., 2021). Compared to other short-read gecko assemblies of similar or larger size (Liu et al., 2015; Xiong et al., 2016), Lister's gecko had higher transposable element content which could reflect lineage-specific differences in repeat content or be caused by technical factors (e.g., fewer collapses of high-identity repeats during assembly). Telomeres represent another repetitive region improved by long-read assembly. Presence of interstitial telomere-like repeats, which arise due to karyotypic rearrangements, is variable among reptiles (Rovatsos et al., 2015). For instance, fewer than half of all skinks and geckos studied to date have these repeats (Rovatsos et al., 2015). We found no evidence of such interstitial sequences in blue-tailed skink or Lister's gecko, suggesting some karyotypic stability over recent evolutionary timescales.

4.2 | Sex chromosomes

Sex determination systems are variable across reptile clades, and the study of sex chromosome evolution in skinks and geckos has been hampered by lack of high-quality reference genomes. In the blue-tailed skink, we generated phased X and Y haplotypes (Figure 3c). A recent study has shown the skink XY chromosome pair is probably shared across all major skink lineages and originated 85–150 million years ago, despite being cytologically indistinguishable in most

species (Kostmann et al., 2021). While prior genomic studies of skink sex chromosome evolution have used the reference genomes of the lizards *A. carolinensis* (Cornejo-Páramo et al., 2020) and *P. muralis* (Kostmann et al., 2021), these species are known to have distinct sex chromosomes from skinks and are phylogenetically distant. In the alignment of the blue-tailed skink XY chromosome pair, we observe increasingly less sequence identity towards the centre of the sex chromosomes (Figure 3c), a signal of progressive degeneration.

We were not able to confidently identify a sex-chromosome pair in Lister's gecko. We opportunistically sequenced a male Lister's gecko and found no clear signal of hemizygosity (Figure 3d). We note this gecko's closest relative, *L. lugubris*, possesses a ZW system with chromosome 1 being the most likely sex chromosome based on R-banding pattern (Volobouev & Pasteur, 1988), although geckos have undergone over 25 transitions between temperature-dependent sex determination (TSD), male heterogamety (XY), and female heterogamety (ZW) in their evolution (Gamble, 2010). As additional skink and gecko genomes are sequenced, comparisons between species will help characterize sex chromosome evolution in these clades.

Molecular sexing is a priority in reptile conservation management, given that sexes are often indistinguishable as juveniles. In the blue-tailed skink, an early male sex-bias at Taronga Zoo later was determined to have resulted from male aggression, which has since been resolved by adjusting sex ratios in holding tanks (L. Cavanagh, personal communication). Molecular sexing in reptiles has been difficult due to variable sex determining mechanisms, frequency of homomorphic sex chromosomes, and rapid evolution of new sex determination systems. In the blue-tailed skink, we found homology of several previously identified X-linked genes shared across multiple Scincidae groups (Kostmann et al., 2021), indicating broad utility of these markers for qPCR-based molecular sexing. However, previously identified Y-linked markers in other skink species (Cornejo-Páramo et al., 2020; Dissanayake et al., 2020) showed much less identity. These phased X and Y haplotypes can be used to develop Y-linked PCR markers for the blue-tailed skink, which would allow for molecular sexing at lower cost and in easier compliance with quarantine procedures.

4.3 | Patterns of heterozygosity reveal distinct ancient and recent histories

Both species studied here were extremely heterozygous at a genome-wide level compared to other vertebrates, particularly species threatened by extinction (Green et al., 2014; Iannucci et al., 2021; Morin et al., 2021). Heterozygosity was elevated on the ends of chromosomes for both species (Figure 4a,c), mirroring variation in GC content (Figure S9). These patterns are probably driven by differences in recombination rates and background selection across the genome, which may also differ by species (Robinson et al., 2021; Schield et al., 2020). High heterozygosity in these reptiles is reflective of large historical N_e, which we inferred to approach 1 million in the skink and 250,000 in the gecko. An important limitation, however, is that mutation rates underpinning N_e estimates are not well known for skinks or geckos, so precise estimates of population size remain uncertain. In early reptile surveys, the blue-tailed skink was noted as "fairly common" across a range of habitats including coastal thickets and cliffs and was somewhat less common in primary rainforest (Cogger et al., 1983; Gibson-Hill, 1947), and it later became "hyper-abundant" in human modified areas including the main settlement in the northeast (Figure 1a; Cogger et al., 1983; Emery, Mitchell, et al., 2021; Emery, Valentine, et al., 2021). In contrast, Lister's gecko was probably somewhat abundant in primary rainforest (Cogger et al., 1983), although its nocturnal and arboreal life-history probably contributed to it being one of the least commonly observed reptiles on Christmas Island (Gibson-Hill, 1947). While habitat change over time may have mediated changes in historical N_e, patterns of past vegetation change on Christmas Island are not well understood, which limits current insights into mechanisms driving such fluctuations. It is clear, however, that in contrast to many endangered species, these Christmas Island reptiles appeared to have large, genetically diverse populations until a recent collapse, which showcases that even genetically diverse species can be vulnerable to rapid declines brought about by novel invasive predators or disease.

In both cases, high-quality genome assemblies allow for quantification of local patterns of heterozygosity, which illuminate patterns that contrast with the genome-wide heterozygosity estimates and provide a fuller picture of genetic diversity in both species. Despite being more heterozygous genome-wide, 10% of the blue-tailed skink reference genome fell in long runs of homozygosity (ROH), a genomic signature of recent inbreeding. A similar pattern of long ROH in an otherwise highly heterozygous genome has previously been reported in the California condor (*Gymnogyps californianus*), another formerly EW species (Robinson et al., 2021).

To discern if ROH were caused by IBD tracts that coalesced after the skinks were brought into captivity or if they were older, perhaps originating in the wild during population declines, we took advantage of ROH lengths to estimate the number of generations since IBD blocks arose. Recombination rates and genetic map lengths have not been well characterized in skinks, and we account for this uncertainty by using the lowest and highest reported squamate recombination rates (Lisachov et al., 2017, 2019) which led to estimates of identity by descent between 11 and 24 generations ago. Despite uncertainty in global recombination rates, and with the caveat that recombination events are unlikely to be uniformly distributed across the genome (Lisachov et al., 2017, 2019; Schield et al., 2020), these estimates substantially pre-date the number of generations the blue-tailed skink reference individual has been bred in captivity (1.5–3.5 generations based on pedigree). Thus, the IBD segments detected as ROH in this individual probably did not originate in captivity and instead originated before the skink went extinct in the wild. Future population-level resequencing of the founders, which is currently planned, could help discern if the founding individuals were inbred themselves, or merely related to

each other when brought into captivity. Additionally, while we did not detect high ROH abundance in the reference Lister's gecko, we cannot rule out the possibility that other geckos in the population are inbred. Indeed, it is quite possible that the extent of inbreeding in both species may be variable due to stochasticity imposed by small population size.

By sequencing one individual of each species from the Taronga Zoo population, we were unable to capture potential variation in inbreeding at the population level. However, by analysing ROH in the reference genomes, we can better target questions for future resequencing studies. For example, resequencing additional individuals from the Taronga Zoo and Christmas Island populations would help clarify variation in ROH abundance and sharing. Assessing trends in inbreeding within and between the captive populations could inform whether separate management should be continued or if gene flow between captive populations should be instituted. While inbreeding investigations would be partially retrospective in the case of Christmas Island reptiles, knowledge of founder relatedness would be helpful in reducing future inbreeding and modelling admixture scenarios for the assisted colonization of the Cocos Islands (Emery, Mitchell, et al., 2021).

4.4 | Immune gene diversity

Homozygosity may lead to inbreeding depression if it exposes recessive deleterious alleles or overlaps loci where there is a heterozygote advantage. For instance, homozygosity at immune genes reduces pathogen recognition ability and immune response under the heterozygote advantage hypothesis (Sommer, 2005). MHC and TLR genes contribute to reptiles' strong innate immune systems (Rios & Zimmerman, 2015), and confer resistance to bacterial infections, including some *Enterococcus* (Baik et al., 2008; Xiong et al., 2022). In the blue-tailed skink, we found ROH that overlapped all MHC class I and II genes, as well as 1 of 7 TLR genes (*TLR4*; Figure S8). An outbreak of *E. lacertideformus* in 2014 at the Christmas Island facility caused the deaths of 40 blue-tailed skinks and 40 Lister's geckos (Agius et al., 2021; Rose et al., 2017). This deadly pathogen continues to circulate in feral reptiles across the island and poses a threat to both species. The Taronga Zoo population has been under quarantine since 2011 and has not been exposed to *E. lacertideformus*, which may have relaxed selection on immune genes.

Homozygosity at MHC may also reduce mating success in captivity if these species have MHC-dependent mate choice. Prior studies of MHC-dependent mate choice in reptiles have generally found support for assortative mating based on MHC diversity or dissimilarity (Han et al., 2019; Miller et al., 2009; Olsson et al., 2003; but see Hacking et al., 2018), including in some skinks (Pearson et al., 2017). Importantly, mate choice preferences for increased diversity and dissimilarity become more pronounced as more MHC markers are considered (Kamiya et al., 2014). Thus, long ROH causing homozygosity at all MHC class I & II genes could increase magnitude of mate choice effects. In captive populations where mating opportunities are chosen based on overall pedigree relatedness, sequencing to determine MHC genotype could be undertaken to maintain immune gene diversity for both pathogen recognition and possible mate choice. Further investigation of population-level immune gene diversity beyond this individual is a priority.

5 | CONCLUSIONS

We generated high-quality genomic resources for the only two reptile species currently classified as EW by IUCN: the Christmas Island blue-tailed skink and the Lister's gecko. These species join a short list of other EW and formerly EW vertebrates with reference genomes including the ʻAlalā (*Corvus hawaiiensis*; Sutton et al., 2018), California condor (*Gymnogyps californianus*; Robinson et al., 2021), milu (*Elaphurus davidianus*; Zhang et al., 2018), Monterrey platyfish (*Xiphophorus couchianus*; Shen et al., 2016), and scimitar-horned oryx (*Oryx dammah*; Humble et al., 2020). We show here that analysing a single high quality reference genome per species can provide a wealth of information relevant to conservation when population-level data are not yet available. Such analyses of genomic diversity can be an important first step, serving to both complement and guide future resequencing studies, which can ascertain parameters at the population-level. Here, the comparison of the two genomes was valuable in showing that despite a shared recent history of captive management, these species may differ particularly in reference to levels of recent inbreeding. Our results suggest a need to sequence blue-tailed skinks before making breeding decisions to maintain genetic diversity, especially at functional loci such as MHC, which may play a role current and future resilience to disease. These genomes provide a resource for future population-level studies to assess the extent of inbreeding and immune gene diversity in an extinct-in-the-wild program and set the stage to include skinks and geckos in comparative genomic studies of reptiles.

AUTHOR CONTRIBUTIONS

Tristram O. Dodge planned the analysis, performed formal analysis, drafted the manuscript, and designed the figures. Katherine A. Farquharson contributed to the planning of the analysis, performed formal analysis, facilitated access to computational resources, and provided supervision. Claire Ford obtained funding, provided information on the captive breeding programs and species biology, and facilitated access to samples. Lisa Cavanagh provided information on the captive breeding programs and species biology and facilitated access to samples. Kristen Schubert provided information on the assisted colonization and species biology. Molly Schumer contributed to the interpretation of results and provided supervision. Katherine Belov obtained funding and provided supervision and resources. Carolyn J. Hogg conceived the study, obtained funding, and provided supervision and resources. All authors critically revised the manuscript.

ACKNOWLEDGEMENTS

This research was made possible through funding from the Threatened Species Initiative (TSI), NCRIS funded Bioplatforms Australia, and the Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science (CE200100012). This study was supported by resources provided by the Pawsey Supercomputing Research Centre with funding from the Australian Government and the Government of Western Australia, as well as the Sydney Informatics Hub through use of the University of Sydney's high performance computing cluster, Artemis. TOD completed this work while undertaking a Fulbright Future Scholarship funded by the Kinghorn Foundation. We thank Elspeth McLennan and Emma Peel for assistance with DNA and RNA extractions, Merly Escalona for advice on Hi-C scaffolding, and Heng Li for advice on variant calling. We thank Parks Australia for their support and ongoing efforts on the conservation of these two species. We thank members of the Schumer Laboratory for comments on an earlier draft.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts to declare.

DATA AVAILABILITY STATEMENT

Raw sequence genomic and transcriptomic reads and their associated metadata have been made available through the Bioplatforms Australia Data Portal (https://data.bioplatforms.com/) under the Threatened Species Initiative. The final genome assemblies, annotations and global transcriptomes are publicly available on Amazon Web Services Australasian Genomes Open Data Store (https://awgg-lab.github.io/australasiangenomes/). The genome assemblies and raw transcriptome files are also available on NCBI and the SRA, under BioProject PRJNA924831 (blue-tailed skink) and PRJNA926684 (Lister's gecko). Code underlying analyses is available at https://github.com/tododge/christmas_island_reptiles.

BENEFIT-SHARING STATEMENT

A research collaboration was developed with threatened species managers from Taronga Conservation Society Australia and Parks Australia; all collaborators are included as coauthors. The results of the research were shared prior to publication with the conservation and government organizations involved in managing the species. This research addresses a priority concern, being the lack of genetic resources for two Extinct-in-the-Wild reptiles. The sharing of our data and results on public databases as described above will accrue benefits for the species studied as well as the study of reptiles more broadly.

ORCID

Tristram O. Dodge https://orcid.org/0000-0001-8382-431X
Katherine A. Farquharson https://orcid.org/0000-0002-9009-7453
Molly Schumer https://orcid.org/0000-0002-2075-5668
Carolyn J. Hogg https://orcid.org/0000-0002-6328-398X

REFERENCES

Agius, J. E., Phalen, D. N., Rose, K., & Eden, J.-S. (2021). Genomic insights into the pathogenicity of a novel biofilm-forming *Enterococcus* sp. Bacteria (*Enterococcus lacertideformus*) identified in reptiles. *Frontiers in Microbiology*, *12*, 635208. https://doi.org/10.3389/fmicb.2021.635208

Ali, J. R., & Aitchison, J. C. (2020). Time of re-emergence of Christmas Island and its biogeographical significance. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *537*, 109396. https://doi.org/10.1016/j.palaeo.2019.109396

Andrew, P., Cogger, H., Driscoll, D., Flakus, S., Harlow, P., Maple, D., Misso, M., Pink, C., Retallick, K., Rose, K., Tiernan, B., West, J., & Woinarski, J. C. Z. (2018). Somewhat saved: A captive breeding programme for two endemic Christmas Island lizard species, now extinct in the wild. *Oryx*, *52*(1), 171–174. https://doi.org/10.1017/S0030605316001071

Andrews, C. W. (1899). A description of Christmas Island (Indian Ocean). *The Geographical Journal*, *13*(1), 17–35. https://doi.org/10.2307/1774789

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data* (0.11.8). https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bachtrog, D., Mank, J. E., Peichel, C. L., Kirkpatrick, M., Otto, S. P., Ashman, T.-L., Hahn, M. W., Kitano, J., Mayrose, I., Ming, R., Perrin, N., Ross, L., Valenzuela, N., Vamosi, J. C., & The Tree of Sex Consortium. (2014). Sex determination: Why so many ways of doing it? *PLoS Biology*, *12*(7), e1001899. https://doi.org/10.1371/journal.pbio.1001899

Baik, J. E., Ryu, Y. H., Han, J. Y., Im, J., Kum, K.-Y., Yun, C.-H., Lee, K., & Han, S. H. (2008). Lipoteichoic acid partially contributes to the inflammatory responses to *Enterococcus faecalis*. *Journal of Endodontics*, *34*(8), 975–982. https://doi.org/10.1016/j.joen.2008.05.005

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boulenger, G. A. (1888). On the reptiles of Christmas Island. *Proceedings of the Zoological Society of London*, *56*(1), 534–536.

Brandies, P., Peel, E., Hogg, C. J., & Belov, K. (2019). The value of reference genomes in the conservation of threatened species. *Genes*, *10*(11), 846. https://doi.org/10.3390/genes10110846

Cavanagh, L. Taronga Conservation Society Australia, personal communication.

Calderón, P. L., & Pigozzi, M. I. (2006). MLH1-focus mapping in birds shows equal recombination between sexes and diversity of crossover patterns. *Chromosome Research*, *14*(6), 605–612. https://doi.org/10.1007/s10577-006-1059-0

Ceballos, F. C., Hazelhurst, S., & Ramsay, M. (2018). Assessing runs of Homozygosity: A comparison of SNP Array and whole genome sequence low coverage data. *BMC Genomics*, *19*(1), 106. https://doi.org/10.1186/s12864-018-4489-0

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., & Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods*, *18*(2), 170–175. https://doi.org/10.1038/s41592-020-01056-5

Cogger, H. G., & Sadlier, R. (1999). *The terrestrial reptiles of Christmas Island: A reappraisal of their status*. Unpublished report to Parks Australia, Canberra (Australian Museum, Sydney).

Cogger, H. G., Sadlier, R., & Cameron, E. E. (1983). *The terrestrial reptiles of Australia's Island territories*. Australian National Parks and Wildlife Service.

Cornejo-Páramo, P., Dissanayake, D. S. B., Lira-Noriega, A., Martínez-Pacheco, M. L., Acosta, A., Ramírez-Suástegui, C., Méndez-de-la-Cruz, F. R., Székely, T., Urrutia, A. O., Georges, A., & Cortez, D. (2020). Viviparous reptile regarded to have temperature-dependent sex determination has old XY chromosomes. *Genome Biology and Evolution*, *12*(6), 924–930. https://doi.org/10.1093/gbe/evaa104

Cox, N., Young, B. E., Bowles, P., Fernandez, M., Marin, J., Rapacciuolo, G., Böhm, M., Brooks, T. M., Hedges, S. B., Hilton-Taylor, C., Hoffmann, M., Jenkins, R. K. B., Tognelli, M. F., Alexander, G. J., Allison, A., Ananjeva, N. B., Auliya, M., Avila, L. J., Chapple, D. G., … Xie, Y. (2022). A global reptile assessment highlights shared conservation needs of tetrapods. *Nature*, *605*(7909), 285–290. https://doi.org/10.1038/s41586-022-04664-7

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, *34*(15), 2666–2669. https://doi.org/10.1093/bioinformatics/bty149

Deakin, J. E., Edwards, M. J., Patel, H., O'Meally, D., Lian, J., Stenhouse, R., Ryan, S., Livernois, A. M., Azad, B., Holleley, C. E., Li, Q., & Georges, A. (2016). Anchoring genome sequence to chromosomes of the central bearded dragon (*Pogona vitticeps*) enables reconstruction of ancestral squamate macrochromosomes and identifies sequence content of the Z chromosome. *BMC Genomics*, *17*(1), 447. https://doi.org/10.1186/s12864-016-2774-3

Director of National Parks. (2012). *Threatened species nomination submission for bluetailed skink Cryptoblepharus egeriae*. http://www.environment.gov.au/biodiversity/threatened/species/pubs/cryptoblepharus-egeriae-nomination.pdf

Dissanayake, D. S. B., Holleley, C. E., Hill, L. K., O'Meally, D., Deakin, J. E., & Georges, A. (2020). Identification of Y chromosome markers in the eastern three-lined skink (*Bassiana duperreyi*) using in silico whole genome subtraction. *BMC Genomics*, *21*(1), 667. https://doi.org/10.1186/s12864-020-07071-2

Durand, N. C., Shamim, M. S., Machol, I., Rao, S. S. P., Huntley, M. H., Lander, E. S., & Aiden, E. L. (2016). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Systems*, *3*(1), 95–98. https://doi.org/10.1016/j.cels.2016.07.002

Dussex, N., van der Valk, T., Morales, H. E., Wheat, C. W., Díez-del-Molino, D., von Seth, J., Foster, Y., Kutschera, V. E., Guschanski, K., Rhie, A., Phillippy, A. M., Korlach, J., Howe, K., Chow, W., Pelan, S., Mendes Damas, J. D., Lewin, H. A., Hastie, A. R., Formenti, G., … Dalén, L. (2021). Population genomics of the critically endangered kākāpō. *Cell Genomics*, *1*(1), 100002. https://doi.org/10.1016/j.xgen.2021.100002

Emery, J.-P., Mitchell, N. J., Cogger, H., Agius, J., Andrew, P., Arnall, S., Detto, T., Driscoll, D. A., Flakus, S., Green, P., Harlow, P., McFadden, M., Pink, C., Retallick, K., Rose, K., Sleeth, M., Tiernan, B., Valentine, L. E., & Woinarski, J. Z. (2021). The lost lizards of Christmas Island: A retrospective assessment of factors driving the collapse of a native reptile community. *Conservation Science and Practice*, *3*(2), e358. https://doi.org/10.1111/csp2.358

Emery, J.-P., Valentine, L. E., Hitchen, Y., & Mitchell, N. (2021). Survival of an extinct in the wild skink from Christmas Island is reduced by an invasive centipede: Implications for future reintroductions. *Biological Invasions*, *23*(2), 581–592. https://doi.org/10.1007/s10530-020-02386-3

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit, A. F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(17), 9451–9457. https://doi.org/10.1073/pnas.1921046117

Ford, C. Taronga Conservation Society Australia, personal communication.

Formenti, G., Theissinger, K., Fernandes, C., Bista, I., Bombarely, A., Bleidorn, C., Ciofi, C., Crottini, A., Godoy, J. A., Höglund, J., Malukiewicz, J., Mouton, A., Oomen, R. A., Paez, S., Palsbøll, P. J., Pampoulie, C., Ruiz-López, M. J., Svardal, H., Theofanopoulou, C., … Zammit, G. (2022). The era of reference genomes in conservation genomics. *Trends in Ecology & Evolution*, *37*(3), 197–202. https://doi.org/10.1016/j.tree.2021.11.008

Frankham, R. (2015). Genetic rescue of small inbred populations: Meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*, *24*(11), 2610–2618. https://doi.org/10.1111/mec.13139

Galla, S. J., Forsdick, N. J., Brown, L., Hoeppner, M. P., Knapp, M., Maloney, R. F., Moraga, R., Santure, A. W., & Steeves, T. E. (2019). Reference genomes from distantly related species can be used for discovery of single nucleotide polymorphisms to inform conservation management. *Genes*, *10*(1), 9. https://doi.org/10.3390/genes10010009

Gamble, T. (2010). A review of sex determining mechanisms in Geckos (Gekkota: Squamata). *Sexual Development*, *4*(1–2), 88–103. https://doi.org/10.1159/000289578

Gemmell, N. J., Rutherford, K., Prost, S., Tollis, M., Winter, D., Macey, J. R., Adelson, D. L., Suh, A., Bertozzi, T., Grau, J. H., Organ, C., Gardner, P. P., Muffato, M., Patricio, M., Billis, K., Martin, F. J., Flicek, P., Petersen, B., Kang, L., … Stone, C. (2020). The tuatara genome reveals ancient features of amniote evolution. *Nature*, *584*(7821), 403–409. https://doi.org/10.1038/s41586-020-2561-9

Geneva, A. J., Park, S., Bock, D. G., de Mello, P. L. H., Sarigol, F., Tollis, M., Donihue, C. M., Reynolds, R. G., Feiner, N., Rasys, A. M., Lauderdale, J. D., Minchey, S. G., Alcala, A. J., Infante, C. R., Kolbe, J. J., Schluter, D., Menke, D. B., & Losos, J. B. (2022). Chromosome-scale genome assembly of the brown anole (*Anolis sagrei*), an emerging model species. *Communications Biology*, *5*(1), 1126. https://doi.org/10.1038/s42003-022-04074-5

Gibbons, J. W., Scott, D. E., Ryan, T. J., Buhlmann, K. A., Tuberville, T. D., Metts, B. S., Greene, J. L., Mills, T., Leiden, Y., Poppy, S., & Winne, C. T. (2000). The Global Decline of Reptiles, Déjà Vu Amphibians: Reptile species are declining on a global scale. Six significant threats to reptile populations are habitat loss and degradation, introduced invasive species, environmental pollution, disease, unsustainable use, and global climate change. *BioScience*, *50*(8), 653–666. https://doi.org/10.1641/0006-3568(2000)050[0653:TGDORD]2.0.CO;2

Gibson-Hill, C. A. (1947). Contributions to the natural history of Christmas Island, in the Indian Ocean. *Bulletin of the Raffles Museum*, *18*, 6–7.

Green, R. E., Braun, E. L., Armstrong, J., Earl, D., Nguyen, N., Hickey, G., Vandewege, M. W., St. John, J. A., Capella-Gutiérrez, S., Castoe, T. A., Kern, C., Fujita, M. K., Opazo, J. C., Jurka, J., Kojima, K. K., Caballero, J., Hubley, R. M., Smit, A. F., Platt, R. N., … Ray, D. A. (2014). Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. *Science*, *346*(6215), 1254449. https://doi.org/10.1126/science.1254449

Gregory, T. R. (2023, February 5). *Animal genome size database*. https://www.genomesize.com/

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hacking, J. D., Stuart-Fox, D., Godfrey, S. S., & Gardner, M. G. (2018). Specific MHC class I supertype associated with parasite infection and color morph in a wild lizard population. *Ecology and Evolution*, *8*(19), 9920–9933. https://doi.org/10.1002/ece3.4479

Han, Q.-H., Sun, R.-N., Yang, H.-Q., Wang, Z.-W., Wan, Q.-H., & Fang, S.-G. (2019). MHC class I diversity predicts non-random mating

in Chinese alligators (*Alligator sinensis*). *Heredity*, *122*(6), 809–818. https://doi.org/10.1038/s41437-018-0177-8

Hogg, C. J., Wright, B., Morris, K. M., Lee, A. V., Ivy, J. A., Grueber, C. E., & Belov, K. (2019). Founder relationships and conservation management: Empirical kinships reveal the effect on breeding programmes when founders are assumed to be unrelated. *Animal Conservation*, *22*(4), 348–361. https://doi.org/10.1111/acv.12463

Humble, E., Dobrynin, P., Senn, H., Chuven, J., Scott, A. F., Mohr, D. W., Dudchenko, O., Omer, A. D., Colaric, Z., Lieberman Aiden, E., Al Dhaheri, S. S., Wildt, D., Oliaji, S., Tamazian, G., Pukazhenthi, B., Ogden, R., & Koepfli, K.-P. (2020). Chromosomal-level genome assembly of the scimitar-horned oryx: Insights into diversity and demography of a species extinct in the wild. *Molecular Ecology Resources*, *20*(6), 1668–1681. https://doi.org/10.1111/1755-0998.13181

Iannucci, A., Benazzo, A., Natali, C., Arida, E. A., Zein, M. S. A., Jessop, T. S., Bertorelle, G., & Ciofi, C. (2021). Population structure, genomic diversity and demographic history of Komodo dragons inferred from whole-genome sequencing. *Molecular Ecology*, *30*(23), 6309–6324. https://doi.org/10.1111/mec.16121

IUCN. (2023, February 11). *The IUCN Red List of Threatened Species*. Version 2022-2. https://www.iucnredlist.org/en

James, D., Green, P. T., Humphreys, W. F., & Woinarski, J. C. Z. (2019). Endemic species of Christmas Island, Indian Ocean. *Records of the Western Australian Museum*, *34*(2), 55–114. https://doi.org/10.18195/issn.0312-3162.34(2).2019.055-114

Kamiya, T., O'Dwyer, K., Westerdahl, H., Senior, A., & Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: The role of diversity and dissimilarity. *Molecular Ecology*, *23*(21), 5151–5163. https://doi.org/10.1111/mec.12934

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907–915. https://doi.org/10.1038/s41587-019-0201-4

Knief, U., Kempenaers, B., & Forstmeier, W. (2017). Meiotic recombination shapes precision of pedigree- and marker-based estimates of inbreeding. *Heredity*, *118*(3), 239–248. https://doi.org/10.1038/hdy.2016.95

Kostmann, A., Kratochvíl, L., & Rovatsos, M. (2021). Poorly differentiated XX/XY sex chromosomes are widely shared across skink radiation. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1943), 20202139. https://doi.org/10.1098/rspb.2020.2139

Kuo, R. I., Cheng, Y., Zhang, R., Brown, J. W. S., Smith, J., Archibald, A. L., & Burt, D. W. (2020). Illuminating the dark side of the human transcriptome with long read transcript sequencing. *BMC Genomics*, *21*(1), 751. https://doi.org/10.1186/s12864-020-07123-7

Lavanchy, E., & Goudet, J. (2023). Effect of reduced genomic representation on using runs of homozygosity for inbreeding characterization. *Molecular Ecology Resources*, 1–16. https://doi.org/10.1111/1755-0998.13755

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, *475*(7357), 493–496. https://doi.org/10.1038/nature10231

Lisachov, A. P., Tishakova, K. V., Tsepilov, Y. A., & Borodin, P. M. (2019). Male meiotic recombination in the steppe agama, *Trapelus sanguinolentus* (Agamidae, Iguania, Reptilia). *Cytogenetic and Genome Research*, *157*(1–2), 107–114. https://doi.org/10.1159/000496078

Lisachov, A. P., Trifonov, V. A., Giovannotti, M., Ferguson-Smith, M. A., & Borodin, P. M. (2017). Immunocytological analysis of meiotic recombination in two anole lizards (Squamata, Dactyloidae). *Comparative Cytogenetics*, *11*(1), 129–141. https://doi.org/10.3897/CompCytogen.v11i1.10916

Liu, Y., Zhou, Q., Wang, Y., Luo, L., Yang, J., Yang, L., Liu, M., Li, Y., Qian, T., Zheng, Y., Li, M., Li, J., Gu, Y., Han, Z., Xu, M., Wang, Y., Zhu, C., Yu, B., Yang, Y., … Gu, X. (2015). Gekko japonicus genome reveals evolution of adhesive toe pads and tail regeneration. *Nature Communications*, *6*(1), 10033. https://doi.org/10.1038/ncomms10033

Lok, S., Lau, T. N. H., Trost, B., Tong, A. H. Y., Wintle, R. F., Engstrom, M. D., Stacy, E., Waits, L. P., Scrafford, M., & Scherer, S. W. (2022). Chromosomal-level reference genome assembly of the North American wolverine (*Gulo gulo luscus*): A resource for conservation genomics. *G3: Genes, Genomes, Genetics*, *12*(8), jkac138. https://doi.org/10.1093/g3journal/jkac138

MacCulloch, R. D., Upton, D. E., & Murphy, R. W. (1996). Trends in nuclear DNA content among amphibians and reptiles. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *113*(3), 601–605. https://doi.org/10.1016/0305-0491(95)02033-0

Makino, S., & Momma, E. (1949). An idiogram study of the chromosomes in some species of reptiles. *Cytologia*, *15*(1–2), 96–108. https://doi.org/10.1508/cytologia.15.96

McDonald, P. J., Brown, R. M., Kraus, F., Bowles, P., Arifin, U., Eliades, S. J., Fisher, R. N., Gaulke, M., Grismer, L. L., Ineich, I., Karin, B. R., Meneses, C. G., Richards, S. J., Sanguila, M. B., Siler, C. D., & Oliver, P. M. (2022). Cryptic extinction risk in a western Pacific lizard radiation. *Biodiversity and Conservation*, *31*(8), 2045–2062. https://doi.org/10.1007/s10531-022-02412-x

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Meng, G., Li, Y., Yang, C., & Liu, S. (2019). MitoZ: A toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Research*, *47*(11), e63. https://doi.org/10.1093/nar/gkz173

Miller, H. C., Moore, J. A., Nelson, N. J., & Daugherty, C. H. (2009). Influence of major histocompatibility complex genotype on mating success in a free-ranging reptile population. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1662), 1695–1704. https://doi.org/10.1098/rspb.2008.1840

Morin, P. A., Archer, F. I., Avila, C. D., Balacco, J. R., Bukhman, Y. V., Chow, W., Fedrigo, O., Formenti, G., Fronczek, J. A., Fungtammasan, A., Gulland, F. M. D., Haase, B., Peter Heide-Jorgensen, M., Houck, M. L., Howe, K., Misuraca, A. C., Mountcastle, J., Musser, W., Paez, S., … Jarvis, E. D. (2021). Reference genome and demographic history of the most endangered marine mammal, the vaquita. *Molecular Ecology Resources*, *21*(4), 1008–1020. https://doi.org/10.1111/1755-0998.13284

Nadachowska-Brzyska, K., Konczal, M., & Babik, W. (2022). Navigating the temporal continuum of effective population size. *Methods in Ecology and Evolution*, *13*(1), 22–41. https://doi.org/10.1111/2041-210X.13740

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S. J., Diekhans, M., Logsdon, G. A., Alonge, M., Antonarakis, S. E., Borchers, M., Bouffard, G. G., Brooks, S. Y., … Phillippy, A. M. (2022). The complete sequence of a human genome. *Science*, *376*(6588), 44–53. https://doi.org/10.1126/science.abj6987

Nurk, S., Walenz, B. P., Rhie, A., Vollger, M. R., Logsdon, G. A., Grothe, R., Miga, K. H., Eichler, E. E., Phillippy, A. M., & Koren, S. (2020). HiCanu: Accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. *Genome Research*, *30*(9), 1291–1305. https://doi.org/10.1101/gr.263566.120

Oliver, P. M., Blom, M. P. K., Cogger, H. G., Fisher, R. N., Richmond, J. Q., & Woinarski, J. C. Z. (2018). Insular biogeographic origins and high phylogenetic distinctiveness for a recently depleted lizard fauna from Christmas Island, Australia. *Biology Letters*, *14*(6), 20170696. https://doi.org/10.1098/rsbl.2017.0696

Olsson, M., Madsen, T., Nordby, J., Wapstra, E., Ujvari, B., & Wittsell, H. (2003). Major histocompatibility complex and mate choice in sand

lizards. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *270*(suppl_2), S254–S256. https://doi.org/10.1098/rsbl.2003.0079

Paez, S., Kraus, R. H. S., Shapiro, B., Gilbert, M. T. P., Jarvis, E. D., & Vertebrate Genomes Project Conservation Group. (2022). Reference genomes for conservation. *Science*, *377*(6604), 364–366. https://doi.org/10.1126/science.abm8127

Pearson, S. K., Godfrey, S. S., Schwensow, N., Bull, C. M., & Gardner, M. G. (2017). Genes and group membership predict Gidgee skink (*Egernia stokesii*) reproductive pairs. *Journal of Heredity*, *108*(4), 369–378. https://doi.org/10.1093/jhered/esx026

Peel, E., Silver, L., Brandies, P., Hogg, C. J., Belov, K., Silver, L., Brandies, P., Hogg, C. J., & Belov, K. (2021). A reference genome for the critically endangered woylie, *Bettongia penicillata ogilbyi*. *GigaByte*, *2021*, 1–15. https://doi.org/10.46471/gigabyte.35

Peel, E., Silver, L., Brandies, P., Zhu, Y., Cheng, Y., Hogg, C. J., & Belov, K. (2022). Best genome sequencing strategies for annotation of complex immune gene families in wildlife. *GigaScience*, *11*, giac100. https://doi.org/10.1093/gigascience/giac100

Peterson, D. G., Stack, S. M., Healy, J. L., Donohoe, B. S., & Anderson, L. K. (1994). The relationship between synaptonemal complex length and genome size in four vertebrate classes (Osteicthyes, Reptilia, Aves, Mammalia). *Chromosome Research*, *2*(2), 153–162. https://doi.org/10.1007/BF01553494

Pinto, B. J., Keating, S. E., Nielsen, S. V., Scantlebury, D. P., Daza, J. D., & Gamble, T. (2022). Chromosome-level genome assembly reveals dynamic sex chromosomes in neotropical leaf-litter geckos (Sphaerodactylidae: Sphaerodactylus). *Journal of Heredity*, *113*(3), 272–287. https://doi.org/10.1093/jhered/esac016

Princée, F. P. G. (1995). Overcoming the constraints of social structure and incomplete pedigree data through low-intensity genetic management. In J. D. Ballou & M. E. Gilpin (Eds.), *Population management for survival and recovery: Analytical methods and strategies in small population conservation* (pp. 124–154). Columbia University Press.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Ralls, K., Ballou, J. D., Rideout, B. A., & Frankham, R. (2000). Genetic management of chondrodystrophy in California condors. *Animal Conservation*, *3*(2), 145–153. https://doi.org/10.1111/j.1469-1795.2000.tb00239.x

Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., Lee, C., Ko, B. J., Chaisson, M., Gedman, G. L., Cantin, L. J., Thibaud-Nissen, F., Haggerty, L., Bista, I., Smith, M., … Jarvis, E. D. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, *592*(7856), 737–746. https://doi.org/10.1038/s41586-021-03451-0

Rios, F. M., & Zimmerman, L. M. (2015). Immunology of reptiles. In *eLS* (pp. 1–7). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0026260

Robinson, J. A., Bowie, R. C. K., Dudchenko, O., Aiden, E. L., Hendrickson, S. L., Steiner, C. C., Ryder, O. A., Mindell, D. P., & Wall, J. D. (2021). Genome-wide diversity in the California condor tracks its prehistoric abundance and decline. *Current Biology*, *31*(13), 2939–2946. e5. https://doi.org/10.1016/j.cub.2021.04.035

Robinson, J. A., Kyriazis, C. C., Nigenda-Morales, S. F., Beichman, A. C., Rojas-Bracho, L., Robertson, K. M., Fontaine, M. C., Wayne, R. K., Lohmueller, K. E., Taylor, B. L., & Morin, P. A. (2022). The critically endangered vaquita is not doomed to extinction by inbreeding depression. *Science*, *376*, 635–639. https://doi.org/10.1126/science.abm1742

Rose, K., Agius, J., Hall, J., Thompson, P., Eden, J.-S., Srivastava, M., Tiernan, B., Jenkins, C., & Phalen, D. (2017). Emergent multisystemic *Enterococcus* infection threatens endangered Christmas Island reptile populations. *PLoS One*, *12*(7), e0181240. https://doi.org/10.1371/journal.pone.0181240

Rovatsos, M., Kratochvíl, L., Altmanová, M., & Pokorná, M. J. (2015). Interstitial telomeric motifs in squamate reptiles: When the exceptions outnumber the rule. *PLoS One*, *10*(8), e0134985. https://doi.org/10.1371/journal.pone.0134985

Russello, M. A., & Jensen, E. L. (2021). Ex situ wildlife conservation in the age of population genomics. In P. A. Hohenlohe & O. P. Rajora (Eds.), *Population genomics: Wildlife* (pp. 473–492). Springer International Publishing. https://doi.org/10.1007/13836_2018_44

Schield, D. R., Pasquesi, G. I. M., Perry, B. W., Adams, R. H., Nikolakis, Z. L., Westfall, A. K., Orton, R. W., Meik, J. M., Mackessy, S. P., & Castoe, T. A. (2020). Snake recombination landscapes are concentrated in functional regions despite PRDM9. *Molecular Biology and Evolution*, *37*(5), 1272–1294. https://doi.org/10.1093/molbev/msaa003

Schield, D. R., Perry, B. W., Adams, R. H., Holding, M. L., Nikolakis, Z. L., Gopalan, S. S., Smith, C. F., Parker, J. M., Meik, J. M., DeGiorgio, M., Mackessy, S. P., & Castoe, T. A. (2022). The roles of balancing selection and recombination in the evolution of rattlesnake venom. *Nature Ecology & Evolution*, *6*(9), 1367–1380. https://doi.org/10.1038/s41559-022-01829-5

Shen, Y., Chalopin, D., Garcia, T., Boswell, M., Boswell, W., Shiryev, S. A., Agarwala, R., Volff, J.-N., Postlethwait, J. H., Schartl, M., Minx, P., Warren, W. C., & Walter, R. B. (2016). *X. couchianus* and *X. hellerii* genome models provide genomic variation insight among *Xiphophorus* species. *BMC Genomics*, *17*(1), 37. https://doi.org/10.1186/s12864-015-2361-z

Shumate, A., Wong, B., Pertea, G., & Pertea, M. (2022). Improved transcriptome assembly using a hybrid of long and short reads with StringTie. *PLoS Computational Biology*, *18*(6), e1009730. https://doi.org/10.1371/journal.pcbi.1009730

Sim, S. B., Corpuz, R. L., Simmonds, T. J., & Geib, S. M. (2022). HiFiAdapterFilt, a memory efficient read processing pipeline, prevents occurrence of adapter sequence in PacBio HiFi reads and their negative impacts on genome assembly. *BMC Genomics*, *23*(1), 157. https://doi.org/10.1186/s12864-022-08375-1

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Smit, A. F. A., & Green, P. T. (2013–2015). *RepeatMasker open-4.0*. http://www.repeatmasker.org/RepeatMasker/

Smith, D., Abeli, T., Bruns, E. B., Dalrymple, S. E., Foster, J., Gilbert, T. C., Hogg, C. J., Lloyd, N. A., Meyer, A., Moehrenschlager, A., Murrell, O., Rodriguez, J. P., Smith, P. P., Terry, A., & Ewen, J. G. (2023). Extinct in the wild: The precarious state of Earth's most threatened group of species. *Science*, *379*, eadd2889.

Smith, L. A. (1988). Lycodon aulicus capucinus a colubrid snake introduced to Christmas Island, Indian Ocean. *Records of the Western Australian Museum*, *14*(2), 251–252.

Smith, M. J., Cogger, H., Tiernan, B., Maple, D., Boland, C., Napier, F., Detto, T., & Smith, P. (2012). An oceanic Island reptile community under threat: The decline of reptiles on Christmas Island, Indian Ocean. *Herpetological Conservation and Biology*, *7*(2), 206–218.

Solovyev, V., Kosarev, P., Seledsov, I., & Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biology*, *7*(1), S10.1–S10.12. https://doi.org/10.1186/gb-2006-7-s1-s10

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, *2*(1), 16. https://doi.org/10.1186/1742-9994-2-16

Sutton, J. T., Helmkampf, M., Steiner, C. C., Bellinger, M. R., Korlach, J., Hall, R., Baybayan, P., Muehling, J., Gu, J., Kingan, S., Masuda, B. M., & Ryder, O. A. (2018). A high-quality, long-read de novo

genome assembly to aid conservation of Hawaii's last remaining crow species. *Genes*, *9*(8), 393. https://doi.org/10.3390/genes9080393

Thompson, E. A. (2013). Identity by descent: Variation in meiosis, across genomes, and in populations. *Genetics*, *194*(2), 301–326. https://doi.org/10.1534/genetics.112.148825

Uetz, P., Freed, P., Aguilar, R., & Hošek, J. (2022). The reptile database. http://www.reptile-database.org/

Uliano-Silva, M., Nunes, J. G. F., Krasheninnikova, K., & McCarthy, S. A. (2021). *marcelauliano/MitoHiFi: Mitohifi_v2.0*. Zenodo. https://doi.org/10.5281/zenodo.5205678

Vinogradov, A. E. (1998). Genome size and GC-percent in vertebrates as determined by flow cytometry: The triangular relationship. *Cytometry*, *31*(2), 100–109. https://doi.org/10.1002/(SICI)1097-0320(19980201)31:2<100::AID-CYTO5>3.0.CO;2-Q

Volobouev, V., & Pasteur, G. (1988). Presumptive sex chromosomes of a unisexual homomorphic species of lizards, *Lepidodactylus lugubris*. *Heredity*, *60*(3), 463–467. https://doi.org/10.1038/hdy.1988.65

Wilder, A. P., Dudchenko, O., Curry, C., Korody, M., Turbek, S. P., Daly, M., Misuraca, A., Wang, G., Khan, R., Weisz, D., Fronczek, J., Aiden, E. L., Houck, M. L., Shier, D. M., Ryder, O. A., & Steiner, C. C. (2022). A chromosome-length reference genome for the endangered pacific pocket mouse reveals recent inbreeding in a historically large population. *Genome Biology and Evolution*, *14*(8), evac122. https://doi.org/10.1093/gbe/evac122

Xiong, X., Tian, S., Yang, P., Lebreton, F., Bao, H., Sheng, K., Yin, L., Chen, P., Zhang, J., Qi, W., Ruan, J., Wu, H., Chen, H., Breault, D. T., Wu, H., Earl, A. M., Gilmore, M. S., Abraham, J., & Dong, M. (2022). Emerging enterococcus pore-forming toxins with MHC/HLA-I as receptors. *Cell*, *185*(7), 1157–1171.e22. https://doi.org/10.1016/j.cell.2022.02.002

Xiong, Z., Li, F., Li, Q., Zhou, L., Gamble, T., Zheng, J., Kui, L., Li, C., Li, S., Yang, H., & Zhang, G. (2016). Draft genome of the leopard gecko, *Eublepharis macularius*. *GigaScience*, *5*(1), 47. https://doi.org/10.1186/s13742-016-0151-4

Zhang, C., Chen, L., Zhou, Y., Wang, K., Chemnick, L. G., Ryder, O. A., Wang, W., Zhang, G., & Qiu, Q. (2018). Draft genome of the milu (*Elaphurus davidianus*). *GigaScience*, *7*(2), gix130. https://doi.org/10.1093/gigascience/gix130

Zheng, Y., & Wiens, J. J. (2016). Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species. *Molecular Phylogenetics and Evolution*, *94*, 537–547. https://doi.org/10.1016/j.ympev.2015.10.009

Zhou, C., McCarthy, S. A., & Durbin, R. (2023). YaHS: Yet another Hi-C scaffolding tool. *Bioinformatics*, *39*(1), btac808. https://doi.org/10.1093/bioinformatics/btac808

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Dodge, T. O., Farquharson, K. A., Ford, C., Cavanagh, L., Schubert, K., Schumer, M., Belov, K., & Hogg, C. J. (2023). Genomes of two Extinct-in-the-Wild reptiles from Christmas Island reveal distinct evolutionary histories and conservation insights. *Molecular Ecology Resources*, *00*, 1–17. https://doi.org/10.1111/1755-0998.13780