Putting genomics to work in threatened species management

Holly Vigo Nelson

Bachelor of Animal and Veterinary Bioscience (Hons I)

School of Life and Environmental Sciences

Faculty of Science

THE UNIVERSITY OF SYDNEY

2024

A thesis submitted to fulfill the requirements for the degree of Doctor of Philosophy

Declaration

This thesis is submitted to the University of Sydney in fulfilment of the requirement for the degree of Doctor of Philosophy. I certify that the content of this thesis is my own work and that all assistance received in preparing this body of work, including published and unpublished works, have been acknowledged throughout. This thesis has not been submitted for any other degree.

Holly Nelson 26/11/2024

Acknowledgements

There is an endless list of people and animals that I would like to thank for their support, guidance, and kindness throughout my PhD. Firstly, I'd like to thank my supervisors, Carolyn Hogg, Kathy Belov, Arthur Georges, and Jane DeGabriel. All of you have been such incredible role models and I have so much respect and admiration for each of you. An additional thanks to Carolyn for holding the fort these past few years, this body of work would not have been possible without your guidance or voice telling me to keep going when I needed it. And to Arthur, your sense of humour was always much appreciated for some comic relief and your quick turnaround with feedback remains unmatched — sending papers to you at 6 pm on a Friday and receiving them back with revisions by 8 pm was truly remarkable.

To past and present members of the Australasian Wildlife Genomics Group, it has been an absolute pleasure working with you. You guys have provided ears and brains when I needed them along with constant reminders that there are people just as weird as I am. To Kate and Elle, thank you for your postdoc wisdom and teachings. This thesis wouldn't have been possible without your patience, advice, and brains. To Andrea—my equally delusional partner-in-crime who started and will finish her PhD alongside me — thank you. There's something comforting in having a fellow PhDer-in-crime who was as equally unhinged by the end. I'd like to say a special thank you to my now very close friends that I've met on my PhD journey – Luke, Kimberley, Lucy, and Toby. To Luke, everyone needs one of you during their PhD. You always made time, always offered an ear, and provided the best company on coffee acquisition walks. Kimberley, thank you for always being the rational angel on my shoulder and always providing impeccable banter. Lucy, thank you for being someone who was always happy to hear me talk about my turtles but most importantly, thank you for sharing your streaming services with me to provide distractions during my downtime. And finally, Toby, for your quality character and very dry sense of humour - you lightened the mood when it was needed.

To Chris, Dan, Shane, Mel, Courtney, and the rest of the NSW DCCEEW team in Coffs Harbour, it has been an absolute pleasure working with you all. I am so grateful for how accommodating

you have been with my research. No matter the time of day, you guys were always there to answer my queries and questions. I couldn't have asked for a better recovery team to work alongside. To Jarrad and the team at Symbio Wildlife Park, I don't think there is a way to thank you enough for your kindness and willingness to collaborate and assist me when I needed. You and the turts made so much of this work possible.

To my family. First, to my amazing parents who have always given me so much freedom and support to pursue my passions. You have always been invested in my success even if you're still unsure about what I have spent the last three and a half years of my life doing other than "something to do with genetics and turtles". Thank you to you and Abuelos for instilling in me a strong sense of discipline and work ethic and providing me with countless opportunities to excel and be the best person I can be. I am forever grateful. To my furry angel, Ruby. Although you couldn't see me through my final year, I will never forget your crazy and happy greetings after a big day or your warm, furry cuddles when you knew I needed them. To Poppy, the last thing I thought I needed in the last 6 months of my PhD was the responsibility of a puppy, but I think it was exactly what I needed. Thank you for making me feel lighter. To my dearest and late Abuela, you told me throughout most of my life to "estudio mucho y hacer bien" (study lots and do well). So, if this body of work doesn't satisfy that request, then I honestly don't know what will, because a Nobel Prize is not on the to-do list – perhaps a postdoc in Germany? In all seriousness, I do hope I have made you proud up there.

Next, I'd like to say a very special thank you to my nearest and dearest amigos outside of academia – Anna, Georgie, Keaara, Kirrily, Alex and many more. I am incredibly lucky to have had each of you. Throughout this journey you have provided love, support, ears, rants and bants, laughs, and kept life fun! You have helped keep me sane by providing a work-life balance. You are all incredible humans and inspirational young women that I regard so highly. I am so grateful to have such high-quality people in my life, and I'm excited to see all the amazing things you'll achieve in the future.

And last, with a large emphasis on "but certainly not least", to my dance family, MADSOC. Although at times it's felt like I have been a full-time dancer rather than a full-time thesis

procurer, you guys have kept me grounded by constantly providing a life outside this body of work. After making it through an insurmountable number of laughs and silliness, fake tans and false eyelashes, injuries and bruised bodies, fouettés, wedgies, burnout, tears, rehearsals, intra and interstate comps, pre-sunrise call times and midnight award ceremonies, wins, losses, and back-to-back performances – I wouldn't take back a single second of balancing both worlds. You have been my second family. Even though you haven't known what has occupied my time off stage, aside from a "PhD" or being the "turtle girl", you have always given me such a strong sense of belonging, safety, and unwavering support. You are all such incredible young women. My heart is full. GO LIONS!

I would also like to acknowledge the funding and support that made this research possible. Thanks to the NSW DCCEEW Saving our Species program, Taronga Conservation Society, Symbio Wildlife Park, and the NCRIS funded Bioplatforms Australia Threatened Species Initiative. Computing resources were provided by the Pawsey Supercomputing Research Centre; Bioplatforms Australia; Galaxy Australia; and the University of Sydney's High Performance Computing cluster Artemis. Thanks to the Australian National University's Biomolecular Resource Facility and Australasian Wildlife Genomics Group for assistance with extractions and laboratory work. Thanks also to the University of Canberra Wildlife Tissue Collection (UC<Aus>) for samples and data. Funding for this work was provided by the Australian Government's Bushfire Recovery Scheme (ERF-WRR2-020) to the Threatened Species Initiative and HVN scholarship funding was provided by the NSW DCCEEW and the University of Sydney. The research reported in this thesis was supported by the award of a Research Training Program recipient offset to the PhD candidate. Cover photo credit to Alex Pike. Photos throughout credit to Shane Ruming and Alex Pike.

Table of Contents

Author attribution statements

A note on the style and layout of this thesis

This thesis consists of a general introduction (Chapter 1), three research chapters (Chapters 2–4), and a general summary of the research undertaken, including the main conclusion and implications of this thesis (Chapter 5). Each research chapter in this thesis is presented in the format of a manuscript prepared for academic journals. As a result, there is some repetition of background material across chapters. All manuscripts have been published. The final published versions, included in the appendices, may differ slightly from research chapter manuscripts due to amendments made during the peer review process. Authorship attribution statements are provided below. The inclusion of published manuscripts follows the University of Sydney's Thesis and Examination of Higher Degrees by Research Policy 2015 guidelines for a thesis with publications. Table and figure numbers in each manuscript have been updated to reflect the corresponding chapter number. Supplementary material for each chapter are provided in the appendices with table and figure numbers updated to reflect the corresponding appendix. PDF versions of published chapters and additional publications and reports that were contributed to during my candidature are presented in the appendices.

Published manuscripts included in this thesis

Chapter 2 *of this this thesis includes the publication:* **Nelson, H. V.**, Farquharson, K. A., Georges, A., Mclennan, E. A., DeGabriel, J. L., Giese, M., Ormond, C., Mcfadden, M., Skidmore, A., Prangell, J., Belov, K., & Hogg, C. J. 2024. A genomic framework to assist conservation breeding and translocation success: A case study of a critically endangered turtle', *Conservation Science and Practice*, p.e13204.

Carolyn Hogg and Holly Nelson conceived and designed the project. Arthur Georges, Mel Giese, Christopher Ormond, Michael McFadden, Adam Skidmore and Jarrad Prangell provided samples and information on *M. georgesi*. Carolyn Hogg obtained funding. Carolyn Hogg and Elspeth McLennan undertook laboratory work to prepare genome samples. Holly Nelson created the reference genome and performed downstream analyses with the assistance of

Katherine Farquharson and Elspeth McLennan. Holly Nelson drafted the manuscript. All authors revised the manuscript.

Chapter 3 *of this this thesis includes the publication:* **Nelson, H.V**., Georges, A., Farquharson, K.A., McLennan, E.A., DeGabriel, J.L., Belov, K. and Hogg, C.J., 2025. A genomic-based workflow for eDNA assay development for a critically endangered turtle, *Myuchelys georgesi*, *Ecology and Evolution*, *15*(1), p.e70798.

Holly Nelson and Carolyn Hogg conceptualised this study. Holly Nelson conducted all field and laboratory work, conducted all analyses and wrote the first draft. Carolyn Hogg obtained funding. Arthur Georges provided information on *M. georgesi* and mitochondrial data, and Katherine Farquharson assisted with initial bioinformatic analyses. Elspeth McLennan and Jane DeGabriel helped revise the manuscript with all authors.

Chapter 4 *of this this thesis includes the publication*: **Nelson, H. V.**, Silver, S., Kovacs, T. G. L., Mclennan, E. A., Georges, A., DeGabriel, J. L., Hogg, C. J., Belov, K., 2025. Genome-wide diversity and MHC characterisation in a critically endangered freshwater turtle susceptible to disease, *Immunogenetics, 77*(1), pp.1-17.

Holly Nelson, Carolyn Hogg and Katherine Belov conceptualised this study. Holly Nelson conducted laboratory work, conducted all analyses and wrote the first draft. Carolyn Hogg obtained funding and Arthur Georges provided samples. Luke Silver and Toby Kovacs assisted with bioinformatic analyses. Elspeth McLennan and Jane DeGabriel helped revise the manuscript with all authors.

Other publications contributed to during candidature

Bower, D.S., Mcknight, D.T., Sullivan, K., Macdonald, S.L., Georges, A., Clulow, S., Mathwin, R., Connell, M.J., **Nelson, H.V.**, Santoro, A. And Nordstrom, B., 2023. Opportunities for research and conservation of freshwater turtles in Australia. *Austral ecology*, *48*(8), pp.1483-1491.

Scientific research reports contributed to during candidature

McLennan, E., **Nelson, H V.**, Hogg, C, 2022. Bellinger River turtle Genetic Management Preliminary Report for Symbio Wildlife Park and Taronga Zoo

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Holly Nelson 26/11/2024

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Professor Kathy Belov 26/11/2024

Presentations

2022

Faculty of Science, School of Life and Environmental Sciences HDR Showcase, Sydney, 31st May 2021; seminar presentation.

55th Annual General Meeting of the Australian Society of Herpetologists, Adelaide, 11th– 14th July 2022; seminar presentation

Bellinger River turtle Recovery Team workshop, Coffs Harbour, 30th July – 1st June 2022; seminar presentation

Pacific Biosciences HiFi Sequencing Experience Tour, Sydney, 24th November 2022; seminar presentation

2023

Faculty of Science, School of Life and Environmental Sciences HDR Showcase, Sydney, 31st May 2023; poster presentation

Pacific Biosciences HiFi Sequencing Experience Tour, Sydney, 28th July 2023; seminar presentation

2024

Rainforest Connections Conference 2024, Ballina, 4th– 7th June; seminar presentation

Genetics Society of Australasia Annual Conference, Sydney, 1st– 3rd July; seminar presentation

Faculty of Science, School of Life and Environmental Sciences HDR Showcase, Sydney, November 2024; seminar presentation

Awards

The University of Sydney, School of Life and Environmental Sciences HDR Showcase, prize for best oral presentation, 2021

The University of Sydney, Faculty of Science Postgraduate Research Support Scheme, 2022

Rainforest Connections Conference, prize for best student presentation, 2024

The University of Sydney, School of Life and Environmental Sciences HDR Showcase, prize for best oral presentation, 2024

Abstract

Earth's sixth mass extinction event threatens over one million species. Small populations are most at risk of extinction due to genomic erosion and decreased adaptive potential. Conservation genetic management of threatened populations is pivotal to species recovery, yet implementation of genetic data into recovery programs lags behind ecological and biological disciplines, making genetic initiatives more crucial than ever. Technological advancements and reductions in sequencing costs have increased our ability to genetically monitor and manage wildlife, resulting in what has been termed the "genomics era" of conservation genetics. Consortia such as the Earth BioGenome Project and the Genome 10K project are generating reference genomes for a diverse array of non-model organisms however, there remains an implementation gap between the generation of genomic data and resources, and their downstream applications into conservation management applications. Much of this space is the result of an inability to transform and translate genetic and genomic data into management information and tools that can be applied by recovery programs.

Here, I generate and utilise a high-quality reference genome. In combination with a variety of bioinformatic and sequencing approaches, I use the genome to produce a suite of downstream genomic resources that can be applied by threatened species recovery programs. I used the critically endangered and disease susceptible Bellinger River turtle (*Myuchelys georgesi*) as my case study species. My thesis aims to (i) generate a high-quality reference genome for a critically endangered species, (ii) use the reference genome in combination with existing genetic datasets to answer important conservation questions relating to genetic diversity and relatedness in captive and wild populations, (iii) develop an eDNA assay for species detection using a bioinformatic approach, and (iv) undertake high-resolution immune gene investigations to better understand adaptive potential in a disease susceptible species.

In Chapter 2, I generated the first chromosome-level reference genome for a Gondwanan turtle species. I aligned previously generated reduced-representation data to the developed reference genome to examine temporal changes in genetic diversity. As genetic diversity is

essential for species survival, I examined both historical and current genetic diversity and assessed relatedness among captive founders using four temporal and spatial groups*:* pre-disease outbreak, post-disease outbreak and two captive populations. I found that genetic diversity has decreased in the wild since disease outbreak. I also detected significantly higher levels of relatedness within compared to between captive populations, providing opportunities to mate with unrelated individuals, mitigating future inbreeding. I developed a framework, modelled on the six stages proposed by Frankham and colleagues, to help managers incorporate genetic data into actionable conservation strategies for any species. This framework, using *M. georgesi* as a case study, aimed to assist managers in effectively integrating genetic approaches into conservation breeding programs.

In Chapter 3, I designed the first eDNA markers for the non-invasive detection of *M. georgesi* to assist managers with detection of the evasive species in areas with restricted accessibility or in areas beyond the species current know distribution. I present a practical workflow for the development of eDNA markers using a novel, reference genome-derived approach involving bioinformatic assembly of a complete mitochondrial genome (mitogenome). I validated the *in silico* approach by comparing a bioinformatically assembled mitochondrial genome (mitogenome) with a Sanger sequenced mitogenome and found zero base pair mismatches. Using the bioinformatically extracted mitogenome, I designed two primers targeting a fragment of the cytochrome oxidase 1 (CO1) gene and cytochrome B (CytB) gene, which were successfully validated *in silico*, *in vitro*, and *in situ*. This work employed a bioinformatic approach for eDNA marker design for incorporation into *M. georgesi* management, offering detailed methodologies and a visual workflow applicable to other species with available genomic data.

In Chapter 4, I extended beyond putatively neutral genetic diversity investigations undertaken in Chapter 2 and employ a targeted gene approach. Given the high susceptibility of *M. georgesi* to disease, I investigated functional gene diversity, specifically, immune genes in the major histocompatibility complex gene family. Using the reference genome, I undertook the first manual immune gene annotation in a testudines species and identified five MHC class I and ten MHC class II. I characterised genome-wide and MHC diversity from pre and post-

disease using 37 re-sequenced genomes and found low levels of genome-wide diversity and no significant shift in MHC diversity over time. As *M. georgesi* and *Emydura Macquarii* hybrids are known to be less susceptible to virus outbreaks, I also undertook a comparative analysis of MHC diversity between pure and backcross animals and found significantly higher variant numbers within MHC genes of backcross animals relative to pure animals. These results highlight the importance of maintaining diversity at MHC genes and the potential role of hybridisation in increasing adaptive potential in disease susceptible species.

Bridging the gap between geneticists and conservation action is crucial for maximising the uptake of genomic resources for the management of threatened species. My thesis showcases the implementation of reference genomes and other genome-wide data into conservation recovery programs. By leveraging cutting-edge genomic technologies alongside existing genomic resources, I demonstrate the power of large datasets in identifying molecular information that can be transformed into practical conservation recommendations and tools.

Abbreviations

A_D	allelic diversity
A_R	allelic richness
AusARG	Australian Amphibian and Reptile Genomics
AWS	Amazon web services
BAM	binary alignment map
BCF	binary counterpart format
BED	browser extensible data
Bp	basepair
BUSCO	benchmarking universal single-copy orthologs
BWA	Burrows-Wheeler aligner
CI	confidence interval
CPU	central processing unit
DAPC	discriminant analysis of principal component
DArTseq	Diversity arrays technologies sequencing
DCCEEW	Department of Climate Change, Energy, the Environment and Water
DNA	deoxyribonucleic acid
eDNA	environmental deoxyribonucleic acid
F1	first generation hybrid
F2	second generation hybrid
F_{IS}	inbreeding coefficient
F_{ROH}	genomic inbreeding coefficient
F_{ST}	pairwise fixation index=
G1	first generation
GATK	genomic analysis toolkit
Gb	gigabase pairs
GO	gene ontology
H_E	expected heterozygosity
HiFi	high fidelity
HMW	high molecular weight
H_O	expected homozygosity
HPC	high performance computer
H_S	standardised heterozygosity
IBD	isolation by distance
IGV	integrative genomics viewer
IQR	interquartile range
IUCN	International Union for Conservation of Nature
K	genetic clusters
LD	linkage disequilibrium
MAF	minor allele frequency

Mb	megabase pairs
MEGA	molecular evolutionary genetic analysis
MK	mean kinship
MHC	major histocompatibility complex
mtDNA	mitochondrial deoxyribonucleic acid
NCBI	National Center for Biotechnology Information
N	sample or population size
N_e	effective population size
ns	non-synonymous
NSW	New South Wales
ORF	open reading frame
P_A	private alleles
PacBio	Pacific Biosciences
PCA	principal components analyses
PCoI	principal coordinates axis 1
PCoII	principal coordinates axis 2
PCoA	principal coordinates analyses
PCR	polymerase chain reaction
PSMC	pairwise sequentially Markovian coalescent
QC	quality control
RAM	random access memory
RNA	ribonucleic acid
ROH	runs of homozygosity
RRS	reduced representation sequencing
SAM	sequence alignment map
SD	standard deviation
SE	standard error
SMRT	single-molecule real-time
SNP	single nucleotide polymorphism
SSC	Species Survival Commission
TSI	Threatened Species Initiative
UC	University of Canberra
VCF	variant call format
WGR	whole genome re-sequencing

Chapter 1 - Thesis Introduction

1.1 The Anthropocene and small populations

We are in the middle of a global biodiversity crisis. Anthropogenetic impacts are key drivers behind species' declines (Jaureguiberry et al., 2022). These declines pose significant threats to the functioning of a diverse range of ecosystems that support life, including humans (Díaz et al., 2019). Today, extinctions are occurring over shorter timeframes (Novacek and Cleland, 2001) with extinction rates between 100 to 1,000 times higher than baseline rates due to habitat loss and fragmentation, climate change, invasive species, pollution, over-exploitation, and emerging infectious diseases (Frankham, 2010; Baronsky, 2011; Wintle, 2019). The loss of local populations results in an overall decline in biodiversity due to the loss of unique genetic or phenotypic traits, even if the species and population recovers or persists elsewhere (Brito and Fernandez, 2000). This means we have likely lost more global biodiversity than is currently recognised (Exposito-Alonso et al., 2022). In the last 200 years, Australia's biodiversity has undergone one of the worst declines of any country (Woinarski et al., 2015). Since European colonisation, over 2,000 taxa have been listed as extinct or threatened with extinction (Environment Protection and Biodiversity Act, 1999). As it is currently unknown how many species exist in Australia, this is likely to be an underestimate. Australia is recognised as a 'megadiverse' country, meaning it hosts a significant proportion of Earth's species and has high levels of endemism (Common and Norton, 1992). The island continent has been geographically isolated for over 50 million years, which has led to the evolution of numerous unique species found nowhere else, with 87% of mammals, 93% of reptiles, and 94% of frogs in Australia considered endemic (Chapman, 2009).

Extinction risk is elevated in small, isolated populations owing to a myriad of processes (Klütsch and Laikre, 2021). Gilpin and Soule (1986) proposed four different stochastic processes, or "vortices", that threaten small populations including demographic stochasticity; environmental stochasticity; genetic stochasticity; and loss of adaptive potential. Demographic stochasticity is the difference in life-history variables such as sex ratio and age structure where variation influences birth, mortality, immigration, and emigration rates. Environmental stochasticity refers to environmental factors such as droughts, flooding, bushfires, and disease that can cause large-scale population declines in short timeframes. Genetic stochasticity, or genetic drift, refers to random fluctuations in allele frequencies owing

to chance events that have more pronounced effects in small populations. Loss of adaptive potential, often included under genetic stochasticity, refers to the loss of a populations or species' ability to respond adaptively to future environmental changes due to population-wide loss of genetic variation. The decline to extinction of small populations from these four processes and the interactions between them is known as an "extinction vortex" (Figure 1.1). Therefore, characterising and understanding the stochastic pressures impacting small, threatened populations enables researchers and conservation managers to develop strategies to help manage them.

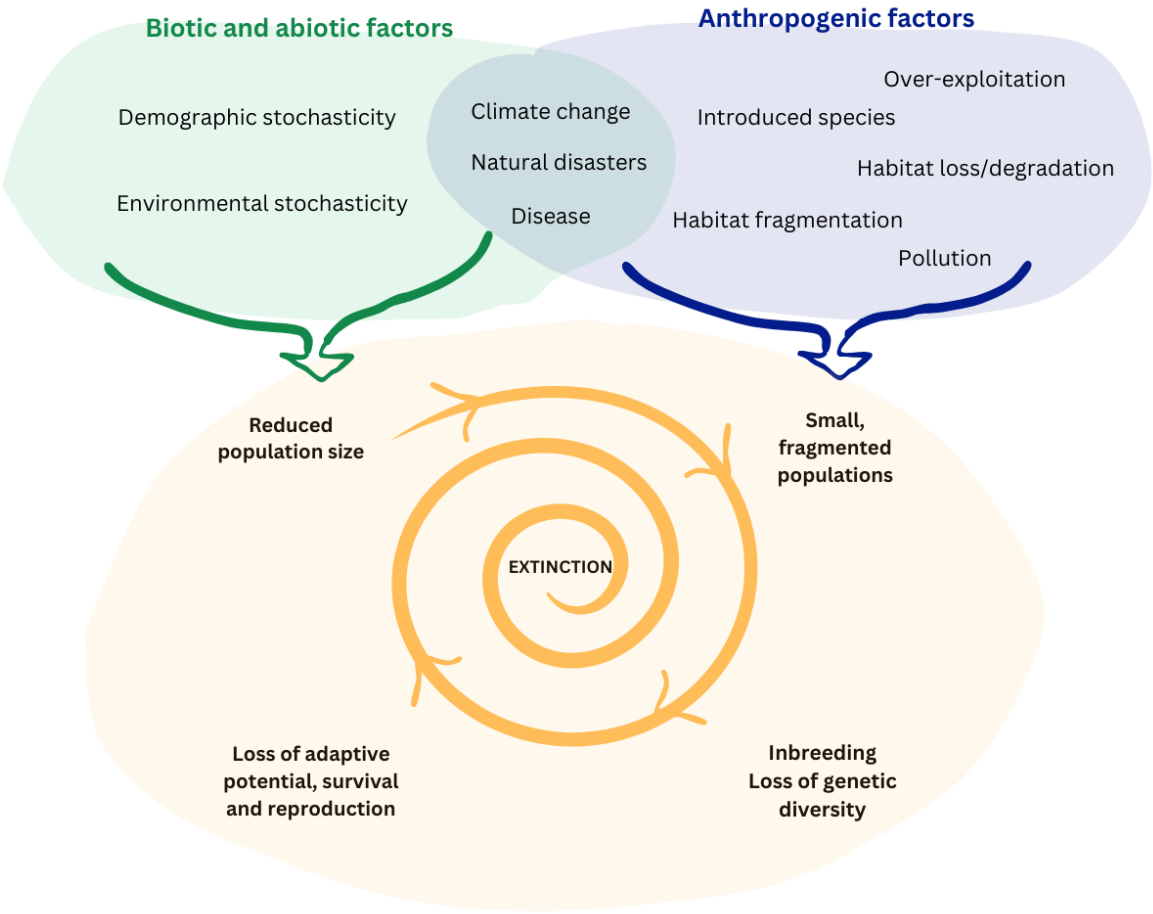


Figure 1.1. A model of the extinction vortex. The vortex shows natural (biotic and abiotic) and anthropogenic factors that work simultaneously to threaten small populations. Natural factors that are exacerbated by human activities are shown between factors. Once a population enters the vortex, the cascade of events resulting in genetic drift and bottlenecks can prevent recovery, resulting in extinction. Adapted from (Gilpin, 1986). Generated using canva.com.

Although a range of factors contribute to the extinction vortex of small populations, one factor becoming more prevalent – as habitats are reduced and species that have not coevolved come into contact – is infectious diseases (Myers et al., 2013, Cohen et al., 2020, Guégan et al., 2020,

Baker et al., 2022, Menajovsky et al., 2023). Both native and invasive species can act as reservoirs/hosts of disease through natural or human-mediated dispersal into habitats outside their historical range (Crowl et al., 2008, Gaertner et al., 2017, Wilkinson et al., 2018). Although an escalating threat to biodiversity globally, most studies focus on infectious diseases in human and domestic animals, with fewer studies undertaken on understanding and characterising diseases in wildlife and the adaptive potential of endangered species to infectious diseases. Assessing the impacts of emerging infectious diseases in wildlife populations can be problematic as there is often little data on the biology, ecology, and genetics of the host species (Robinson et al., 2010).

The emergence of novel infectious diseases is becoming one of the most significant stochastic threats faced by wildlife and small populations. Notable wildlife diseases include chytridiomycosis in amphibians (Weldon et al., 2004), white-nose syndrome in bats (Blehert et al., 2009), and avian influenza (Roberton et al., 2006). However, many emerging infectious diseases remain unidentified, or poorly understood, due to limited data on the underlying pathogens or their hosts. The risks posed by emerging infectious diseases are amplified by the shrinking gap between human and natural environments, thereby increasing the urgency to better understand wildlife diseases. Understanding the interactions between wildlife as both hosts and vectors of disease is essential for One Health, as highlighted by the COVID-19 pandemic (Hudson et al., 2002, Hu et al., 2013, Brearley et al., 2013, Baker et al., 2022). Although infectious diseases are a naturally occurring process in ecosystems, anthropogenic changes to landscapes in combination with changing climate conditions have emerged as significant factors driving the emergence of new infectious diseases and spillover events across species (Baker et al., 2022). Managing declining populations in the presence of infectious diseases, in addition to all the other drivers of the extinction vortex, is still a burgeoning field and a topic of this thesis.

1.2 Conserving dwindling populations

When anthropogenic factors, such as disease outbreaks, drive small and fragmented populations into an irreversible extinction vortex, intensive management strategies become necessary to prevent extinctions (Figure 1.1). Intensive management can involve

translocations, breeding in fenced areas of wild habitat, supplementary feeding, captive and hand rearing of young, and captive/conservation breeding (Leus, 2011). In Australia, these actions are often managed under a conservation action/recovery plan that provide structured approaches for the conservation and recovery of species at risk of extinction. Conservation breeding and release programs (CBRP) have become a valuable tool for managing small and fragmented wild populations (Conway, 2011, Grueber et al., 2019). Programs can assist conservation management efforts via scientific research, public education, and as genetic reservoirs for reinforcing dwindling wild populations (Pritchard et al., 2012, Ochoa et al., 2016). As such, programs require a multi-disciplinary approach across science, interactive management, politics, environmental education, habitat preservation and habitat restoration (Mallinson, 1995). Programs typically aim to: (i) establish and grow a captive insurance population, (ii) supplement an existing population, and/or (iii) re-establish a population that is extinct in the wild.

The development of CBRPs is challenging and has received criticism due to high costs, the diversion of resources away from habitat protection, and the potential of species adaptation to captivity (Balmford et al., 1995, Snyder et al., 1996, Harley et al., 2018). Because of this, programs have usually been considered as a last resort and are commonly implemented when populations are nearing extinction. Still, CBRPs have played a major role in species that would otherwise likely have gone extinct, as well as species that have shown a status improvement in IUCN Red List reassessments (Hoffmann et al., 2010, Conde et al., 2011, Conde et al., 2013). The success of conservation breeding programs to stave off extinction can be attested to the number of animals bred for release and the level of genetic diversity retained by the program (Hogg, 2013), with successful cases including the Californian condor (*Gymnogyps californianus*) (Ralls and Ballou, 2004), and the Arabian oryx (*Oryx leucoryx*) (Price, 1989), and more recent cases including the Vietnam pheasant (*Lophura edwardsi*) (Collar, 2020), and the orange-bellied parrot (*Neophema chrysogaster*) (Pritchard et al., 2022). Given current extinction trends, the number of species requiring conservation breeding programs is likely to increase over coming decades (Conde et al., 2013, Steventon et al., 2024). Several key factors have been shown to be critical to the success of a conservation breeding program including ongoing investment into infrastructure and commitment to the program (Harley et al. 2018), use of molecular genetics to understand relationships of founding individuals (Hogg et al.,

2017a), and a commitment to maintaining the species within captivity until the conservation breeding and release program is no longer needed (Short, 2009).

The value of genetic information for conservation actions is globally recognised through the new Global Biodiversity Framework that has the maintenance of genetic diversity, of both domesticated and wild species, as a key component of its 2050 goals (Convention on Biological Diversity, 2020). Conservation genetics focuses on understanding the genetic diversity, population structure, and evolutionary processes of small populations or species (Allendorf et al., 2012) and can involve: resolving taxonomic uncertainties; defining evolutionary diverged units; forensic analyses; management of invasive species; molecularly obtaining species information (sex, population size, demographic history, mating system, population structure, gene flow, parentage, diet, disease); and genetic management to minimise inbreeding and loss of genetic diversity in wild and captive populations (Frankham et al., 2010). The main goal of conservation action plans is to maintain the long-term viability of a species. Maintaining genetic diversity is an important component of population viability as it helps to mitigate negative effects associated with inbreeding and provides populations with the potential to adapt to changing environments (Lacy, 1987, Ballou et al., 2010, Frankham et al., 2010). Understanding a species' inherent genetic diversity, in addition to their historical diversity and future potential, is of utmost importance in species conservation. When genetic variation is present, populations can adapt to environmental stressors and changes, but without it, populations are more susceptible to extinction risk through loss of genetic diversity, inbreeding depression, and loss of evolutionary potential (Frankham, 2010).

There are two principal types of genetic diversity: neutral and adaptive genetic diversity. Neutral diversity is often considered the variation in regions between the functional genes, and adaptive diversity is the variation within a functional gene. Together, coding and non-coding variations allow species to adapt to both immediate environmental challenges (through functional protein changes) and longer-term shifts in gene expression or regulatory networks (Kashi and King, 2006, Racimo and Schraiber, 2014). Variation detected outside non-coding regions of the genome (neutral diversity) has formed the foundation of conservation genetic research (Holderegger et al., 2006). Neutral genetic variation is often used to provide

estimates on basic features of wildlife populations, including inbreeding and genetic diversity, demographic history, taxonomy, evolutionary significant units, and population structure (Frankham, 1996, Reed and Frankham, 2001, Charlesworth et al., 2017, Yıldırım et al., 2018). It also provides opportunities to detect fine-scale patterns of variation on an individual level, enabling more informed conservation decisions of intensely managed species through kinship and pedigree analyses (Ballou, 1995, Lacy et al., 1995, Ivy et al., 2009). Adaptive or functional genetic diversity refers to variants within a gene (or quantitative trait) that affect an organism's adaptive potential. Variation within functional regions is essential for populations to respond to natural and anthropogenic factors (Figure 1.1), and influences populations' ability to persist despite environmental shifts or changes in selective pressures (Holderegger et al., 2006, Hoffmann et al., 2017). Neutral genetic variation is often used as a proxy for adaptive potential due to the strong correlations that have been shown between whole genome diversity and neutral diversity (McLennan et al., 2019, Wright et al., 2020). However, there has been discussion that levels of neutral diversity may not be representative of adaptive potential and that genetic diversity at specific loci may be a more important measure of adaptive potential (Teixeira & Huber, 2021). For example, Manlik et al. (2019) examined functional genes in the major histocompatibility complex (MHC) immune gene family and neutral microsatellite diversity in two bottlenose dolphin (*Tursiops aduncus*) populations; one with high reproductive success and fitness, Shark Bay; and the other with lower reproductive success and fitness, Bunbury. Their findings showed that the Shark Bay population had higher levels of MHC diversity compared to Bunbury, whereas neutral microsatellite diversity was similar between the two populations, noting that this study only used 23 microsatellites and so was likely to be under-powered. Understanding adaptive diversity provides insight on how populations might adapt to novel environments and how prone they may be to stochastic events. As threats, like climate change and disease outbreaks, are difficult to mitigate, conservation strategies that promote adaptation and resilience may be more effective and self-sustaining in the long-term (Kosch et al., 2022). Therefore, it is important to identify, characterise and conserve functional diversity to preserve the potential for future adaptation (Hoelzel et al., 2019). It is also important to ensure conservation decisions are based on genomic data and consider functional genetic variation, as well as assessing population level diversity metrics such as genome-wide heterozygosity, inbreeding and population differentiation.

1.3 The conservation genomics era

Since the publication of the human genome in 2003 (Consortium, 2004), investment and innovations in genomic technologies and computational methods have revolutionised genetic research across the tree of life (Hu et al., 2021). This has sparked what has been termed the "genomics era", enabling genetic research to expand beyond first-generation approaches such as microsatellite marker analyses (Supple and Shapiro, 2018). Initial high throughput sequencing (HTS) used short-reads to develop genome-wide markers, through reduced representation sequencing (RRS), that captured single-nucleotide polymorphisms (SNPs) across putatively neutral regions of the genome (Davey et al., 2011). This data has been useful for investigating population genetic metrics such as heterozygosity, inbreeding, and relatedness, and is a cost-effective approach for informing conservation actions (Kraus et al., 2015, Kleinman-Ruiz et al., 2017, Delord et al., 2018). More recently, the advent of long-read sequencing, such as PacBio (Pacific Biosciences) high fidelity (HiFi) (California, USA), whole-genome high-throughput chromosome conformation capture (Hi-C), and Oxford Nanopore sequencing (Oxford, UK), has facilitated more comprehensive and contiguous reference genome assemblies. Combining these with short-read transcriptomics and re-sequenced genomes has revolutionised how we can investigate species of conservation concern at both their neutral and functional diversity (Larson et al., 2019, Lewis et al., 2022).

A reference genome is a set of nucleotide sequences, typically from a single individual, that serves as a genetic representation of the structure and organisation of the species' genome (Formenti et al., 2022). Although reference genomes have no direct impact on conservation outcomes, highly contiguous, accurate, and annotated chromosome-level assemblies can greatly enhance genomic studies by providing a foundational blueprint for a suite of downstream applications (Hogg, 2024). Reference genomes provide data for (i) population analyses, (ii) aligning and calling reduced representation sequencing (RRS) data within the same or closely related species to improve the reliability of variant calls and downstream inferences (Torkamaneh et al., 2016, Shafer et al., 2017, Galla et al., 2018, Wright et al., 2019), (iii) developing targeted single nucleotide polymorphism (SNP) panels (Kraus et al., 2015, Wright et al., 2015) and (iv) complete assembly of functional gene families such as MHC

(Horton et al., 2004, Jain et al., 2018, Peel et al., 2022). RRS data is commonly used globally for the assessment and management of a range of wildlife species, such as the little owl (*Athene noctua*) (Pellegrino et al., 2016), white-tailed deer (*Odocoileus virginianus*) (Chafin et al., 2021), and yellow-footed rock-wallabies (*Petrogale xanthopus*) (Smith et al., 2023). More recently, the combination of high-quality reference genomes and re-sequenced genomes data has allowed the identification of functional gene variation that may influence phenotypic or disease traits, fitness, and adaptive potential of a species in species (Onley et al., 2021, Theissinger et al., 2023), such as the black flying fox (*Pteropus alecto*) (He et al., 2023a), white eared-pheasant (*Crossoptilon crossoptilon*) (Wu et al., 2024), and fishing cat (*Prionailurus viverrinus*) (Carroll et al., 2024).

Genetic advancements have also facilitated widespread adoption of molecular tools, including metabarcoding where DNA barcodes are used to identify putative species within a mixed sample, which has become increasingly common in assessing community diversity and abundance across various spatial scales in the form of dietary analyses (Díaz-Abad et al., 2022, Menning et al., 2023), genetic diversity of populations (Adams et al., 2023, Andres et al., 2023), microbiome studies (Cabodevilla et al., 2023, Meng et al., 2024), and for detection of rare and invasive species through environmental DNA (eDNA) (Rees et al., 2014, Ardura et al., 2015, Ruppert et al., 2019, Beng and Corlett, 2020, Lam et al., 2022). eDNA refers to DNA that is deposited from an organism into its environment (water, soil, air) that can then be detected using molecular approaches such as metabarcoding (Barnes and Turner, 2016). Initially developed to study microbial diversity in the environment, the technique was first proposed in the 1980's as a way of obtaining microorganisms' DNA from environmental samples (Olsen et al., 1986). With the emergence of high-throughput sequencing, relatively cheap costs, and the non-invasive nature, eDNA investigations for threatened species management has grown exponentially (Sahu et al., 2023, Takahashi et al., 2023).

The use of eDNA analyses for resolving questions around threatened species distribution and site occupancy can significantly enhance the precision of conservation assessments and strategies for species recovery. This is achieved through detection of individual species using primers or assays designed specifically for that species or DNA metabarcoding to detect entire

communities within a taxonomic group, such as vertebrates, insects, or eukaryotes (Mauvisseau et al., 2019, Lopes et al., 2021, Valdivia-Carrillo et al., 2021). To develop taxonomically universal or specific markers, a representation of the existing nucleotide diversity must be considered to ensure (i) variation within markers to distinguish among species/taxa and (ii) the development of primers that work for the group or species targeted while avoiding the amplification of unwanted taxa. Species delimitation using mitochondrial DNA (mtDNA) is a widely used and effective method for identifying and defining species. However, relying on a single locus, such as DNA barcoding, can also lead to a bias towards overestimating species diversity (Hickerson et al., 2006, Davis et al., 2024). There is often not enough representation of the DNA regions in public databases, which adds to the need for experiments to increase the representation of species. Even when markers are already developed, species-specific reference sequences might be missing from databases and need to be produced before the methodology can be used for species detection.

1.4 Implementation space

Although the usefulness of having genetic and genomic data for threatened species management is widely acknowledged (Hoban et al., 2022), challenges remain in bridging the gap between current academic/research knowledge and the practical application of genetics in management practices and actions. The data output by genetic and genomic sequencing and use of bioinformatic software often requires interpretation from experts in the field of genomics, resulting in a disconnect between research investigation and applied management (Hogg, 2024). These discrepancies can stem from multiple sources including the relatively young age of conservation genetics and genomics compared to disciplines like ecology and biology, resulting in limited genetic literacy of tertiary trained managers, and insufficient translation of findings to policymakers for incorporation into guidelines and frameworks (Britt et al., 2018; Habel et al., 2013; Hoban et al., 2013; Shafer et al., 2015; Klütsch and Laikre 2021). As a result, there can be discrepancies between the conservation priorities of managers and the research focus of conservation geneticists. With the advent of NGS and transition from genetic to genomic studies, there is a growing impetus for scientists to use new technologies that may not have immediate practical applications. There is also an expectation for researchers to publish scientific articles involving novel genetic and genomic

methodologies and these results are not necessarily aimed at answering conservation questions (Taylor et al., 2017). Practitioners often lack access to scientific articles, making it difficult for species managers to remain up to date on conservation genetic research (Hogg et al., 2017b). The transition from genetic to genomic studies and the growing reliance on bioinformatic analyses of large datasets puts pressure on researchers to continually upskill (Schweizer et al., 2021). Additionally, this approach requires significant compute power to execute, that are often inaccessible to managers (Fuentes-Pardo and Ruzzante, 2017).

Several studies have suggested the need for collaborative meeting points where genetic knowledge can be translated and practical issues addressed (Cook et al. 2013; Hoban et al. 2013a, b; Sandström et al. 2016, 2019; Shafer et al. 2015; Holderegger et al. 2019; Hogg, 2024). Translational research is an interdisciplinary approach to conservation that seeks to bridge the gap between scientific knowledge and practical applications (Enquist et al., 2017). To maximise interdisciplinary contributions to threatened species conservation, it is desirable for researchers and managers to engage directly to implement research findings into management (Taylor et al., 2017). Engagement between researchers and practitioners starting from study design and then running onto on-ground implementation is an effective approach to close the knowledge-implementation gap in the future (Ainsworth et al., 2020, Jarvis et al., 2020).

Over the last decade, several international consortia under the Earth BioGenome Project (PRJNA533106) have been established to create high-quality reference genomes and genomic data for species across the tree of life (Lewin et al., 2018). In Australia, several national initiatives (Threatened Species Initiative (PRJNA1075750) (Hogg et al., 2022), Oz Mammals Genomics (PRJNA1075707) (Eldridge et al., 2020), and Amphibian and Reptile Genome Initiative (PRJNA1075730) (https://doi.org/10.25953/xer9-2e12)) have also been established under Bioplatforms Australia to generate genomic data for Australian wildlife. For example, the Threatened Species Initiative aims to connect managers with researchers, including bioinformaticians and geneticists, who are trained in complex areas where they can interpret and translate research data outputs. Additionally, they are putting tools in the hands of conservation managers by developing user-friendly workflows, pipelines, training modules

and workshops to improve the integration of genetics into management actions (Hogg et al., *2024*).

1.5 Reptiles and the plight of freshwater testudines

Modern reptiles are a diverse class of organisms encompassing Squamata (lizards and snakes), Testudines (turtles and tortoises), Crocodilia (crocodiles, alligators, caimans, and gharials), and Rhynchocephalia (tuatara) (Sues, 2019). Several anthropogenetic factors are resulting in declines of reptile populations worldwide including habitat loss, unsustainable removal for food, traditional medicines, fashion, exotic pet trade, environmental contamination, climate change, invasive species, and disease (Gibbons et al., 2000, Todd et al., 2010). Long generation times, limited dispersal capabilities and temperature dependant physiological functions make them ill-adapted and highly susceptible to environmental change (Li et al., 2024). Relative to other groups such as birds, mammals, and economically important fish, reptiles often garner fewer conservation resources (Olson and Pilliod, 2022). Even when budgets are diverted to non-model organisms, species in this class are generally overlooked due to greater interest in more 'iconic' fauna like mammal and avian species (Small, 2012, Olson and Pilliod, 2022). Although interest in the preservation of biodiversity, and consequently interest in reptile conservation, is growing, the idea of reptiles as less important, simplistic, expendable proto-animals remains deep-rooted in society (de Miranda, 2017, Todd et al., 2010). Consequently, many reptile species lack basic information on ecology, status and abundance, and geographic distribution, impeding conservation actions (Guedes et al., 2023).

Freshwater reptile species are particularly at risk of extinction. Freshwater environments cover 0.8% of the Earth's surface yet are home to 6% of all species, making them biodiversity hotspots (Hawkesworth and Kalin-Arroyo 1995; Dudgeon et al. 2006). One in three freshwater species globally are now threatened with extinction (Collen et al., 2014). The high level of connectivity of freshwater ecosystems means that fragmentation can have profound effects (Revenga et al., 2005) and threats such as pollution, invasive species, and disease can be easily transported across waterbodies (Dudgeon et al., 2006, Darwall et al., 2009). This lends urgency to the study of diversity and of the relative risk of extinction of species in freshwater ecosystems. Most freshwater species are endemic to a single river, lake or basin, making them

highly specialised to their environment which reduces their ability to escape environmental change and increases their sensitivity to threatened processes (Reid et al., 2019).

Freshwater turtles (Order Testudines) are long-lived species with species divergence times dating back 220-million years (Joyce et al., 2013, Shaffer et al., 2017, Lyson and Bever, 2020, Thomson et al., 2021). Turtles have high annual fecundity and high mortality rates in early years and late onset of reproductive maturity (Iverson, 1991). Approximately 60% of all turtle species are listed as being endangered, threatened, or vulnerable, making them one of the most threatened taxa globally (Figure 1.2) (Van Dyke et al., 2018).

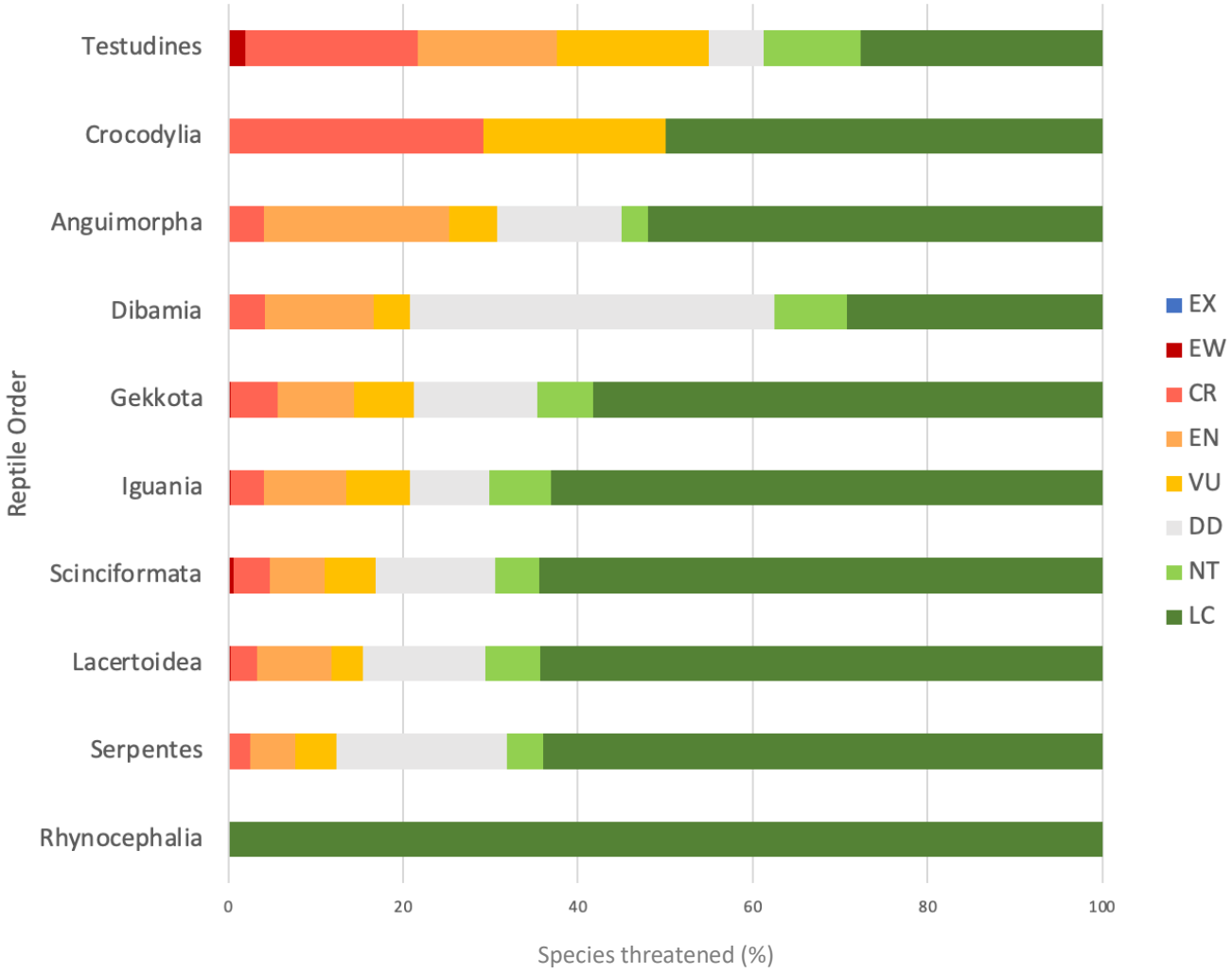


Figure 1.2. Percentage of species threatened by reptile order. EX, extinct; EW, extinct in the wild; CR, critically endangered; EN, endangered; VU, vulnerable; DD, data deficient; NT, near threatened; LC, least concern. Data adapted from Cox et al. (2022).

The testudines suborder Pleurodira consists of side-necked turtles that occupy freshwater habitats in the southern hemisphere including Australia, New Guinea, Indonesia, Africa, and South America (Bourque, 2016). Many species in this suborder remain underrepresented in literature compared to their northern hemisphere counterparts (Suborder Cryptodira) (Card et al., 2023). Approximately 45% of Australian freshwater turtle species are listed as threatened by the IUCN (Van Dyke et al., 2018), or by the Australian Federal and/or State Governments. Major threats include nest predation by native and introduced predators, habitat modification, water quality, road mortality, invasive species, hybridisation, and disease outbreaks (Van Dyke et al., 2018). A number of diseases that affect turtle survival have been documented throughout the literature including *Fusarium sp.* that cause significant declines in turtle hatchling success (Smyth et al., 2020), mycotic shell disease (Nardoni and Mancianti, 2023), ranavirus (Johnson et al., 2008), herpesvirus (Alfaro-Núñez et al., 2014, Ossiboff et al., 2015, Sim et al., 2015, Mullin et al., 2020), iridovirus (Chen et al., 1999, Marschang et al., 1999) and nidovirus (Zhang et al., 2018).

1.6 The Bellinger River turtle

The Bellinger River turtle (*Myuchelys georgesi)* is a species of short-necked (Family Chelidae) turtle and is one of two turtle species that is listed as critically endangered in Australia. The species was discovered by Cann (1997) in the Bellinger River and a short section of its main tributary, the Kalang River in New South Wales, Australia (Figure 1.3), although the species has not been observed in the Kalang since 2007 (Georges et al., 2011). The species is a medium-sized omnivorous turtle with a current known distribution that is restricted to a 60 km range of the Bellinger catchment in north-eastern New South Wales (NSW), Australia (Figure 1.3) (Cann et al., 2015, Zhang et al., 2018). The Bellinger River runs perpendicular down the Great Dividing Range, flowing unimpeded to the east for 110 km. The river is predominantly freshwater aside from 20 km of estuary at the mouth of the river. The Bellinger catchment and several small freshwater catchments in NSW remain isolated and relatively untouched which has facilitated unique habitat specialisation and catchment-specific speciation (Spencer et al., 2014). *M. georgesi* has adapted to upstream regions of the Bellinger River, preferring deeper waterholes surrounded by bedrock (Spencer et al., 2014). Males reach reproductive maturity at 5 years while females reach reproductive maturity at roughly

7 years of age (Blamires et al., 2005), and the turtles are estimated to live beyond 30 years. Females lay one clutch of 10-15 eggs per year on land between October and early November.

In 2005, the population was estimated to be *ca* 4500 ± 1400 individuals (Blamires et al., 2005). Although a relatively untouched catchment, the only known population of *M. georgesi* is threatened by riparian habitat loss, predation by introduced and native predators, reduced water quality, hybridisation due to human mediated dispersal of the Murray River turtle (*Emydura macquarii*) and a novel disease outbreak (Spencer et al., 2014, Georges et al., 2018, Zhang et al., 2018, Chessman et al., 2020).

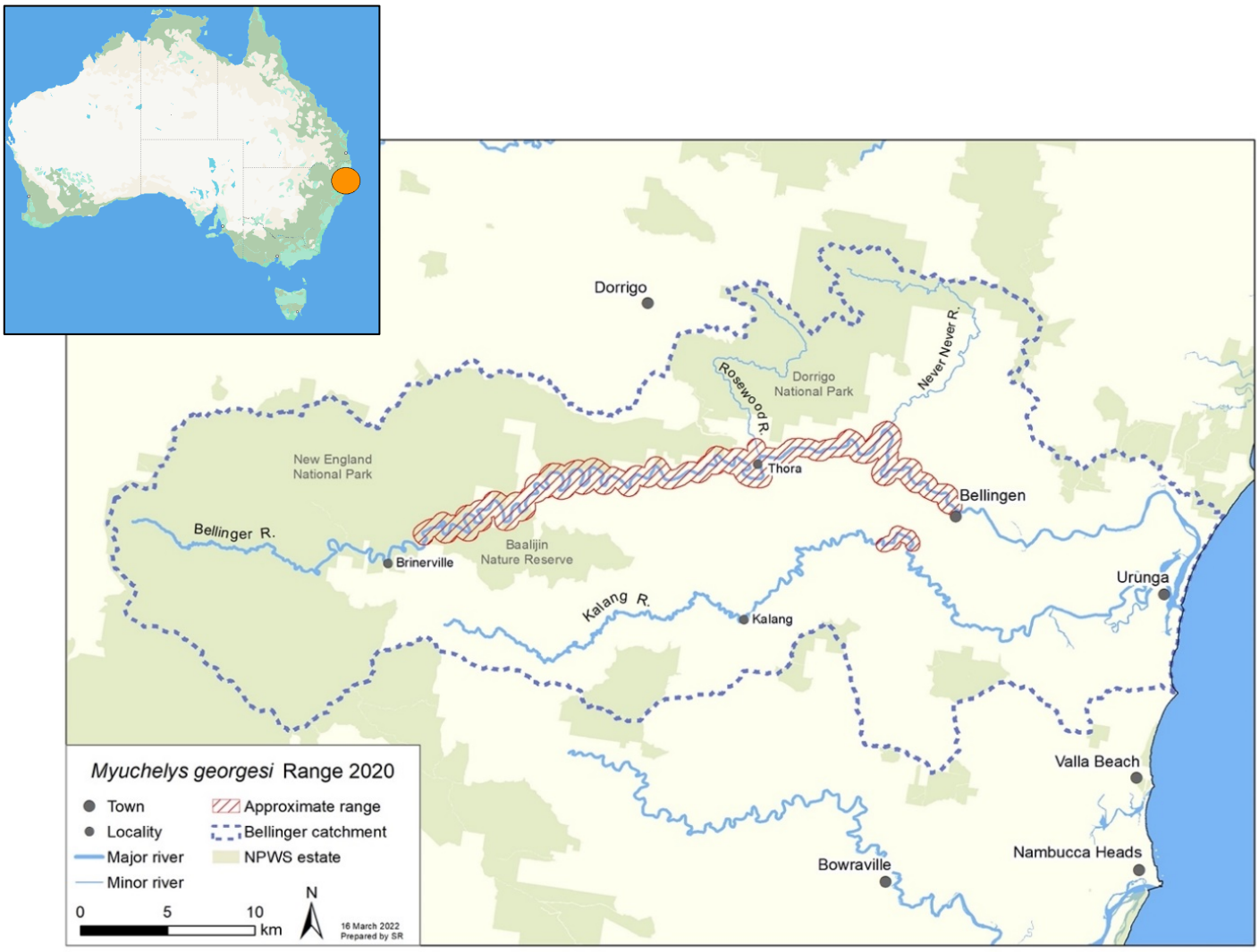


Figure 1.3. An Inset of Australia with a map of the Bellinger catchment showing the current known distribution of the Bellinger River turtle within the Bellinger and Kalang Rivers. Source: NSW Department of Climate Change, Energy, the Environment, and Water.

In 2015, a novel nidovirus, the Bellinger River virus (BRV), resulted in the death of more than 90% of individuals, with mortalities occurring mostly among adults (Zhang et al., 2018, Chessman et al., 2020, Parrish et al., 2024). During this time, it is estimated that the population declined to less than 150 individuals (Spencer et al., 2018, Chessman et al., 2020). The BRV was initially detected in the lower reaches of the river before spreading upstream over 2 months. Disease symptoms included swollen eyelids, sinuses and oral cavities, conjunctivitis, and tan foci on the skin of the ventral thighs (Figure 1.4) (Zhang et al., 2018). Histopathology results of necropsied animals showed severe inflammation of the spleen, kidneys, and liver (Zhang et al., 2018). The other two locally occurring species (*E. Macquarii* and *Chelonia longicollis*) did not display symptoms of the disease. Since the 2015 outbreak, two lesser outbreaks have occurred in 2022 and 2024 during which a handful of fatalities were documented. It is suspected that these outbreaks are from novel nidovirus strains. Although there has been evidence of individuals recovering from latter outbreaks, there is little data on subsequent outbreaks. Additionally, knowledge on how the outbreaks have affected the species' distribution throughout the river remains limited due to accessibility constraints in densely vegetated areas and areas of the river that are adjacent to private land.

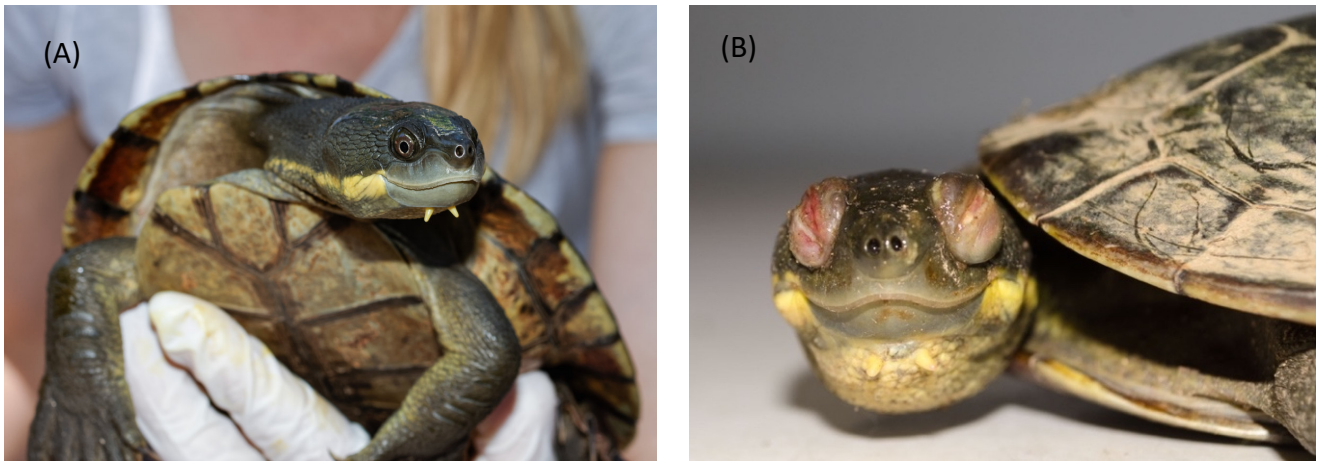


Figure 1.4. *Myuchelys georgesi* showing **(A)** A healthy adult being processed during annual surveys. **(B)** An adult infected with Bellinger River virus showing severe ocular symptoms. Images: Shane Ruming.

The high susceptibility of *M. georgesi* to BRV and subsequent virus outbreaks puts the species at high risk of extinction. As a result of the decline, *M. georgesi* was declared 'Critically Endangered' in both NSW under the Threatened Species Conservation Act 1995 (NSW

Scientific Committee 2016); and nationally under the Environmental Protection and Biodiversity Conservation Act 1999 (Commonwealth of Australia 1999); in addition to the IUCN/SSC Tortoise and Freshwater Turtle Specialist Group Draft Red List (Rhodin et al. 2017).

The species is currently managed by NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW) and includes a conservation breeding and release program that commenced in 2015. The first captive colony was founded at Western Sydney University in 2015, and was later relocated to Taronga Conservation Society Australia, Mosman, Australia in 2015 (*N* = 16 founding individuals). The Taronga population was founded from emergency intakes, with seven females and nine males collected from two sites in upper reaches of the Bellinger River where the virus had not yet reached. This was followed by a second intake in 2017 to Symbio Wildlife Park, Helensburgh, Australia (*N* = 19) (Figure 1.2A). The Symbio population was founded post-virus from six females and thirteen males collected from four sites in the lower reaches of the river (Figure 1.1A). To date, the program has conducted four rounds of reinforcements by releasing 174 G1 juveniles into the Bellinger River, significantly boosting wild population numbers to an estimated 300 wild individuals.

Genetic sampling of the species dates back to 1986 (Georges and Adams, 1992) with subsequent sampling in 2007 for phylogenetic and hybridisation research (Georges et al., 2011, Cann et al., 2015, Georges et al., 2018). When established in 2015, the contemporary wild population and conservation breeding program had limited population genetic data. This lack of data for genetic management of the species could potentially have long-term implications on the retention of genetic diversity and population viability for the species. Since the outbreak, RRS data has been generated for the contemporary wild population followed by a call for development of a high-quality reference genome for the species by the recovery team in 2020, to undertake and output genetic and genomic data for the species.

In September of 2022, a review of the Bellinger River turtle conservation action plan brought together managers, researchers, members of the community and other stakeholders for a two-day workshop. This collaborative meeting point facilitated a “management driven”

research approach through the generation of genetic and genomic management questions that form the aims and objectives of this thesis.

1.7 Research Aims:

The objective of this thesis was to establish a research program in collaboration with the conservation managers to ensure that the outcomes were focussed on developing tools for future management actions and providing critical research data to inform the current and ongoing management of the conservation breeding and release program, following the principles outlined in Hogg et al. (2017). This was achieved by establishing a collaboration between the University of Sydney, the University of Canberra and the NSW DCCEEW Saving our Species program. This collaboration ensured that all research questions were informed by management questions, management funding was able to leverage research funding, and research outcomes were communicated directly to the NSW DCCEEW team and the conservation breeding centres to maximise benefit of the results to the breeding program.

Here, I provide worked examples of how a reference genome, next generation sequencing data, and bioinformatic-based tools can be directly applied to conservation decisions of threatened species. Using the critically endangered Bellinger River turtle (*M. georgesi*) as case study, specifically this thesis has the following aims:

1) To generate a reference genome for a critically endangered species with limited genomic data to provide a genomic resource for genetic and genomic analyses and a valuable resource for ongoing molecular management of the species.
2) To use the reference genome and putatively neutral, genome-wide markers to:
    a. Generate baseline population genomic data for monitoring of the wild and captive populations
    b. Develop a framework that can be applied across conservation breeding programs to provide researchers and captive managers with information on how and when to implement genetic and genomic data into the breeding program.
3) To utilise the reference genome to undertake a novel approach to eDNA assay development, that is translatable across taxa, through bioinformatic assembly of a complete mitochondrial genome; and

4) To employ the reference genome to undertake high-resolution, functional gene analyses to annotate and characterise immune gene diversity in a disease susceptible species.

This thesis addresses the above objective and aims as follows:

- In **Chapter 2**, I generate a chromosome-level reference genome for *M. georgesi* and utilise the reference genome together with additional genomic data, reduced representation sequencing (RRS), to investigate genetic diversity across time and space in captive and wild individuals (*N* = 166). Using these analyses and the *M. georgesi* conservation breeding program as a case study, I provide a generalised framework for researchers and conservation practitioners wanting to integrate genetic and genomic data into conservation breeding programs.
- In **Chapter 3**, I use the reference genome generated in Chapter 2 to develop an eDNA assay by bioinformatically extracting a complete mitochondrial genome. I design two sets of PCR primers for Cytochrome Oxidase 1 and Cytochrome B genes that I successfully validate across three levels; *in silico; in vitro*; and *in situ*.
- In **Chapter 4**, I extend beyond neutral diversity investigations undertaken in Chapter 2 and use whole genome re-sequencing data to explore functional gene diversity, specifically the major histocompatibility complex (MHC) immune gene family, in *M. georgesi.* Again, using the reference genome and a targeted gene approach, I undertake the first manual MHC gene annotation in a freshwater turtle species. I use 35 re-sequenced genomes to investigate genome-wide and MHC variation across temporal and spatial scales in both pure *M. georgesi* and second-generation hybrids.
- In **Chapter 5**, I provide a synopsis of my work where I draw together the outcomes of the individual chapters, highlight the implications of this thesis, address key gaps, and provide suggestions for future directions.

Together, this thesis showcases how a reference genome and genomic data can be transformed by researchers into applied outcomes for threatened species management and help close the space between research and implementation.

Chapter 2 - A genomic framework to assist conservation breeding and translocation success: A case study of a critically endangered turtle

2.1 Manuscript

Published in *Conservation Science and Practice*

A genomic framework to assist conservation breeding and translocation success: A case study of a critically endangered turtle

Holly V. Nelson[1], Katherine A. Farquharson[1, 2], Arthur Georges[3], Elspeth A. McLennan[1], Jane L. DeGabriel[4], Melissa Giese[5], Christopher Ormond[5], Michael McFadden[6], Adam Skidmore[6], Jarrad Prangell[7], Katherine Belov[1,2], and Carolyn J. Hogg[1,2 *]

[1] School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia.

[2] Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Sydney, Sydney, NSW 2006, Australia

[3] Institute for Applied Ecology, University of Canberra, Bruce, ACT 2617, Australia

[4] NSW Department of Climate Change, Energy, the Environment and Water, Parramatta, NSW 2150, Australia

[5] NSW Department of Climate Change, Energy, the Environment and Water, Coffs Harbour, NSW 2450, Australia

[6] Taronga Conservation Society Australia, Mosman, NSW 2088, Australia

[7] Symbio Wildlife Park, Helensburgh, NSW 2508, Australia

2.2 Abstract

Conservation breeding programs are an effective approach to addressing biodiversity loss. Captive populations are managed to maintain genetic diversity, yet there remains an "implementation gap" in effectively translating molecular genetic data into management. Technological advancements are facilitating rapid generation of genetic data, increasing accessibility for breeding programs. In 2010, Frankham and colleagues proposed a six-stage process for establishing successful conservation breeding and release programs. Here we describe the conservation breeding program for the critically endangered Bellinger River turtle (*Myuchelys georgesi*) and characterise the value of genetic sampling for informing management actions. By generating a chromosome-level genome and population genetic data, we investigated past and present diversity and assessed relatedness among captive

founders. We present a framework modelled on Frankham and colleagues six stages to assist managers in implementing genetic data into actionable conservation strategies. This framework, and worked case study, for managers aims to better guide implementation of genetic approaches into conservation breeding programs.

2.3 Introduction

The six stages of conservation breeding

Conservation breeding programs are a valuable method for managing threatened species (Conway, 2011, Grueber et al., 2019). Such programs can assist biodiversity conservation via scientific research, public education, and as genetic reservoirs for reinforcing dwindling wild populations (Pritchard et al., 2012, Ochoa et al., 2016). High-profile cases of successful conservation breeding and release programs include the black-footed ferret (*Mustela nigripes*) (Wisely et al., 2003), Californian condor (*Gymnogyps californianus*) (Ralls and Ballou, 2004), and the Arabian oryx (*Oryx leucoryx*) (Price, 1989). Despite these examples and other notable successes, including several in Australia (Andrew et al., 2018, Scheele et al., 2021, Heinsohn et al., 2022), conservation breeding programs remain an intensive and expensive management approach (Conde et al., 2011).

Frankham et al. (2010) described a six-stage process of establishing successful conservation breeding and release programs; (1) recognising decline of the wild population and its genetic consequences; (2) founding one or more captive populations; (3) expanding captive populations to a secure size; (4) managing the captive population over generations; (5) choosing individuals for reintroduction; and (6) managing the translocated population in the wild. Priorities throughout this process include developing husbandry techniques, rapid reproduction, disease mitigation, and genetic management (Frankham et al., 2010). Traditionally, genetic management of captive populations has been based on pedigrees from studbook records and the underlying assumption that founders are neither related nor inbred, which is often not the case (Lacy, 1987, Hogg et al., 2019). This assumption means initial breeding events may result in inadvertent inbreeding and diversity loss for populations already experiencing limited genetic variation (Frankham et al., 2017, Barrett et al., 2022).

Advancements in genetic sequencing technologies and bioinformatic tools are making it feasible to integrate molecular data into conservation breeding programs to determine levels of relatedness and genetic diversity metrics.

High-throughput sequencing has resulted in the generation of large amounts of data and the emergence of reference genomes for conservation management. Reference genomes provide data for a range of investigations including designing species-specific microsatellite markers for population analyses, developing targeted single nucleotide polymorphism (SNP) panels, aligning and calling reduced representation sequencing (RRS) data within the same or closely related species, exploratory genome-wide analyses and high-resolution functional gene investigations such as complex immune gene families (Galla et al., 2018, Brandies et al., 2019, Peel et al., 2022). The data output by high-throughput sequencing often requires interpretation from experts in the field of genomics resulting in a "research implementation gap" (Taylor et al., 2017). Translational research is an interdisciplinary approach to conservation that seeks to bridge the gap between scientific knowledge and practical applications (Enquist et al., 2017). To maximise interdisciplinary contributions to conservation breeding programs, there is growing responsibility for scientists to engage with managers directly to implement research findings into management. Here, we present a case study that epitomises Frankham and colleagues six-stage process of establishing a conservation breeding and reintroduction program and show how the integration of a multidisciplinary approach has benefited a critically endangered turtle species.

Our case study species

Turtles are among the most threatened vertebrate taxa globally (Van Dyke et al., 2018). Over 20% of turtle species are listed as Critically Endangered by the International Union for the Conservation of Nature (IUCN; (McCallum, 2021), with declines greatly reducing turtles' contributions to ecological processes and food webs (Chessman et al., 2020). Pleurodira, a 200-million-year-old suborder of freshwater turtles found only in Australia, New Guinea, Africa, and South America remains highly under-represented in the peer-reviewed literature. Within this suborder, the Bellinger River turtle (*Myuchelys georgesi)* is a Critically Endangered

species in the family Chelidae. The species is a medium-sized omnivorous turtle with a current known distribution that is restricted to a 60 km range of the Bellinger catchment in north-eastern New South Wales (NSW), Australia (Figure 2.1A) (Cann et al., 2015, Zhang et al., 2018). The Bellinger catchment and several small freshwater catchments in NSW remain isolated and relatively untouched which has facilitated unique habitat specialisation and catchment-specific speciation (Spencer et al., 2014). *M. georgesi* has adapted to up-stream regions of the Bellinger River, preferring deeper waterholes surrounded by bedrock (Spencer et al., 2014). Although a relatively untouched catchment, the only known population is threatened by riparian habitat loss as a result of private landownership, predation by introduced and native predators, reduced water quality, hybridisation due to human mediated dispersal of the Murray River turtle (*Emydura macquarii*) and a novel disease outbreak (Spencer et al., 2014, Georges et al., 2018, Zhang et al., 2018, Chessman et al., 2020).

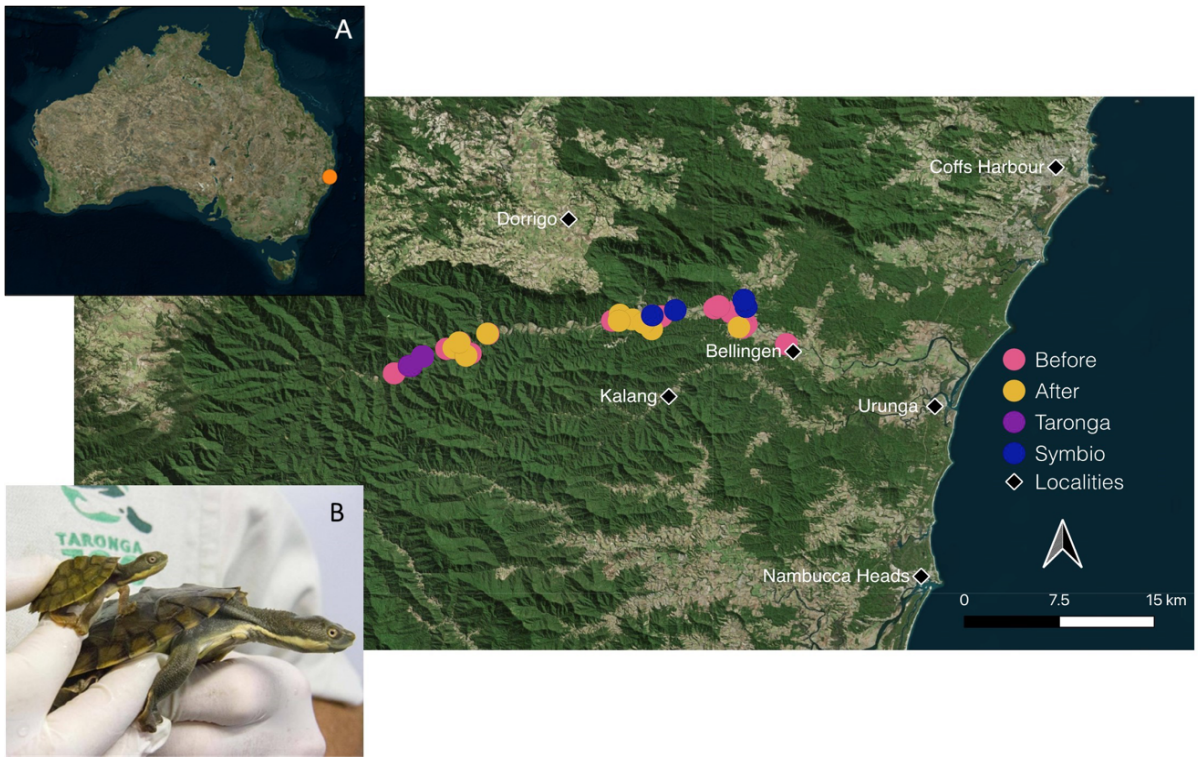


Figure 2.1. (A) An inset of Australia with a map of the Bellinger River basin showing the locations of historic (Before: 2007, pink) and contemporary (After: 2015-2020, yellow) samples, including the founder collection locations for the two conservation breeding programs (Taronga: 2015, purple; and Symbio: 2017, blue) (NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW) (unpublished data) sample points. Note: Sample locations have been obscured as *Myuchelys georgesi* is listed as a Category 2 species in the DCCEEW sensitive species data policy. **(B)** A captive *M. georgesi* hatchling and adult. Photo: Amy Russell.

In 2015, a species-specific nidovirus resulted in the death of more than 90% of individuals, with mortalities occurring mostly among adults (Figure 2.1B) (Zhang et al., 2018, Chessman et al., 2020). During this time, it is estimated that the population declined from approximately 3000 to less than 150 individuals (Spencer et al., 2018, Chessman et al., 2020). To date, there has been limited evidence of individuals recovering from the disease and no records of breeding in the river since the outbreak. Additionally, knowledge on how the outbreak affected the species' distribution throughout the river remains limited due to accessibility constraints.

Species recovery is currently managed by NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW) and includes a conservation breeding program that commenced in 2015, with a captive colony founded at Western Sydney University that was later relocated to Taronga Conservation Society Australia, Mosman, Australia (*N* = 16 individuals). The Taronga population was founded from emergency intakes, with seven females and nine males collected from two sites in the upper reaches of the Bellinger River where the virus had not yet reached (Figure 2.1A). This was followed by a second intake in 2017 to Symbio Wildlife Park, Helensburgh, Australia (*N* = 19) (Figure 2.2A). The Symbio population was founded post-virus from six females and thirteen males collected from four sites in the lower reaches of the river (Figure 2.1A).

The species recovery program mirrors the principles of Frankham *et al.*'s six-stage process (Figure 2.2A) and has developed successful husbandry, breeding, and disease mitigation protocols (Taronga Conservation Society Australia, 2023). While Frankham's stages were written in 2010 for reintroductions, the underlying principles can be applied to various conservation relocation/translocation types, including conservation introductions and reinforcements (IUCN/SSC, 2013). Genetic sampling of the species dates back to 1986 (Figure 2.2A) (Georges and Adams, 1992) but the conservation breeding and release program initially had limited genetic data. This lack of data for genetic management of the species could potentially have long-term implications on the retention of genetic diversity and population viability for the species.

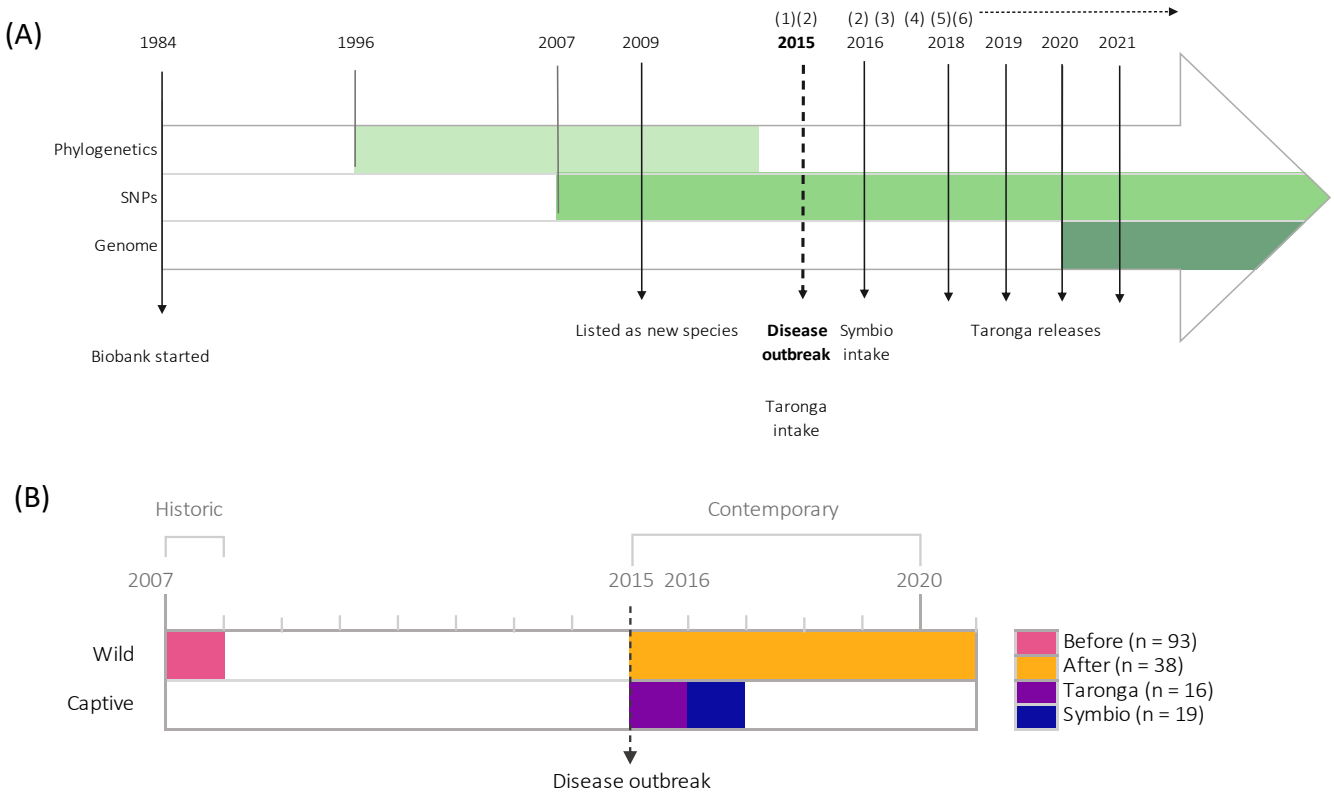


Figure 2.2. (A) Timeline of the collection and implementation of genetic data into *Myuchelys georgesi* management including disease outbreaks, founder intakes and captive release events. Numbers in brackets above the timeline indicate completion of Frankham et al. (2010) six-stage process for conservation breeding and reintroduction programs; (1) Recognising decline of the wild population and its genetic consequences; (2) Founding one or more captive populations; (3) Expanding captive populations to a secure size; (4) Managing the captive population over generations; (5) Choosing individuals for reintroduction; and (6) Managing the reintroduced population in the wild. **(B)** Historic and contemporary samples used for comparative analyses of wild and captive individuals.

Here, we generate a comprehensive genetic toolkit for the Bellinger River turtle, translating our genetic findings into management recommendations for the conservation breeding and release program (Table 2.1). To achieve this, we assembled the first chromosome-level genome for the suborder Pleurodira and aligned population genetic data to (1) investigate historic and contemporary diversity and differentiation, (2) identify levels of founder relatedness within and between the Taronga and Symbio populations, and (3) develop an easy-to-follow framework for managers to translate research into management actions.

2.4 Methods

Reference genome

Comprehensive details of genome assembly and annotation are provided in the Supplementary Material. In summary, we conducted high molecular weight DNA extractions from the heart tissue of a male *M. georgesi* using the Nanobind Tissue Big DNA kit following the manufacturer's protocol (Circulomics, Pacific Biosciences, California, United States of America). PacBio HiFi Single-Molecule Real-Time (SMRT) bell libraries were sequenced across two SMRT cells on the PacBio Sequel II at the Australian Genome Research Facility (Brisbane, Australia). We assembled the HiFi genome using Hifiasm v.0.16.0 (Cheng et al., 2021), and scaffolded it using Hi-C data produced on a Illumina Novaseq 6000 and the YaHS v.1.1 scaffolding pipeline (Zhou et al., 2022). We extracted RNA from brain, liver, and spleen tissue of a female *M. georgesi* using the Qiagen Rneasy Mini Kit (Qiagen, Hidlen, Germany) following the manufacturer's protocol. Transcriptomes were sequenced at the Ramaciotti Centre for Genomics (The University of New South Wales, Sydney, Australia) on an Illumina NovaSeq 6000 S1 flow cell as 100 bp paired-end reads. We annotated the genome with FGENESH++ using a global transcriptome assembly that was generated from the brain, liver, and spleen transcriptomes (Table A1.2.1).

Table 2.1. A framework for integrating genetic data into conservation breeding management, mirroring Frankham et al. (2010) six-stage process of establishing successful conservation breeding and reintroduction (or other conservation translocation) program. Bold text highlights the genetic contributions of our study. Note: The steps in the *Myuchelys georgesi* case study were not undertaken chronologically at each stage, as we have retrospectively integrated genetic inputs into the program. We suggest that other conservation breeding programs make efforts to follow the order we recommend.

Stage	Genetic input	Management outcome	*M. georgesi* program
1. Recognising decline of the wild population and its genetic consequences	• Biobanking • Reference genome generation • DNA sampling of declining population (blood/tissue)	• Setting genetic foundations and acquiring samples for subsequent steps • Baseline analyses investigating wild population genetic diversity, differentiation and temporal changes wrought by the decline	• Tissue samples were collected in 1986 (A. Georges; Figure 2.2B) and biobanked at The University of Canberra • **A chromosome-level reference genome has been assembled and annotated for downstream analyses** • Commencement of annual surveying and sampling of wild population by NSW DCCEEW • **Baseline analyses of genetic metrics (H_S, H_E, H_O, F_{IS}, A_R, P_A,N_e, F_{ST})[1]**
2. Founding one or more captive populations	• 20-30 contributing founders • DNA sampling of all founders • DNA sampling of contemporary wild individuals	• **Identifying founder relationships for baseline studbook data** • **Ensuring no hybrids or backcross individuals in captivity** • **Ensuring captive populations are representative of wild diversity**	• Two captive populations established from 35 individuals (Taronga Conservation Society and Symbio Wildlife Park) • Tissues collected from wild and founding individuals (NSW DCCEEW; Figure 2.2B) • The founders gathered by NSW DCCEEW and various institutions were sourced from opposite ends of the species distribution and are housed separately • Genetic analyses identified the presence of hybrids which were then removed
3. Expanding captive populations to a secure size	• Establishing a studbook with known founder relatedness • DNA sampling of each generation • DNA sampling of new founder intakes	• Provides data for stage 4	• Tissue collected from G1 for parentage analyses (Georges, 2020) • Development of studbook • Breeding program successfully increased numbers from 35 to *ca.* 299 individuals (including releases)

4. Managing the captive population over generations	• Maintaining a studbook with known founder relatedness • Breeding between captive populations • Introducing genetic material via new founders	• **Maintaining/increasing genetic diversity** • **Minimising MK**[1] • Monitor for genetic drift • Increasing genetic diversity with new individuals • Ensuring individuals are representative of wild diversity	• Continuation of progeny sampling • **Analyses to monitor genetic metrics (H_S, H_E, H_O, F_{IS}, A_R, P_A, N_e, F_{ST})**[1] • **Implementation of founder relatedness results by breeding individuals with low MK** and integrating genetic data into studbook
5. Choosing individuals for reintroduction/release	• Results from analyses in steps 2-4 i.e. reintroduction/translocation cohorts with wide ranging diversity	• Reintroduce/translocate genetically diverse individuals • Insight on where to reintroduce/translocate based on wild diversity and population structure	• Retrospective integration of genetic diversity results from this study to inform reintroduction/translocation decisions
6. Managing the reintroduced (or translocated) population in the wild	• Routine DNA sampling of wild population • Knowledge of captive and wild genetics from earlier stages	• Maintain wild diversity post-release • Ensure wild is not 'swamped' by different genotypes from a single captive population • Ensure captive-released animals are breeding with wild animals and contributing to the next generation	• Annual sampling of wild population by NSW DCCEEW • **Genetic diversity of wild and captive populations undertaken in this study**

[1] NSW DCCEEW New South Wales Department of Climate Change, Energy, the Environment and Water; MK Mean kinship; H_S Standardised heterozygosity; H_E Expected heterozygosity; H_O Observed heterozygosity; F_{IS} Inbreeding coefficient; A_R Allelic richness; P_A Private alleles; N_e Effective population size; F_{ST} differentiation.

Genetic analyses for management

Two-week river-wide, randomised, stratified surveys were conducted in April of 2007 and October and November of 2015, 2016, 2019, and 2020 by the NSW DCCEEW. October and November coincided with the beginning of the breeding season, at which time gravidity of adult females could be confirmed. Turtles were identified using scute notching, and their carapace width and length were measured. They were also weighed, bled or had skin biopsied, body condition checked, and swabbed for viral screening (Chessman et al., 2020). DNA samples were collected by extracting blood from the jugular vein or by removing part of the trailing webbing of the clawless toe on the hindfoot (Georges et al., 2018).

Blood and skin biopsies from 166 individuals were collected across 33 sites during surveys (2007, 2015-2020) for the purposes of population genetic analyses (Figure 2.1A, Figure 2.2A, Table A1.2.6). The population genetics samples were stored in 75% ethanol at −20°C in the University of Canberra Wildlife Tissue Collection (GenBank UC<Aus>). As described in Georges et al. (2018), samples were sequenced over multiple runs using high coverage DArTseq™ (Diversity Arrays Technology PL, Canberra, Australia), a form of RRS. We aligned raw DArT sequences to the repeat masked genome generated in this study (Figure A1.3.1, Table A1.2.2) and called SNPs using Stacks v2.61 (Catchen et al., 2013, Rochette et al., 2019). The 'populations' module was then run with the following parameters: minimum samples per population 30% (−r 0.3); minimum minor allele frequency (MAF) of 0.01 (--min_maf 0.01); and --write_random_snp.

To partition our data, we investigated genetic clusters using a variational Bayesian framework in fast*STRUCTURE* v1.0 (Raj et al., 2014) and visualised the results using DISTRUCT v1.1 (Rosenberg, 2004). K=1-4 clusters were tested based on the demographic characteristics of the species including population size and range (Figure A1.3.2). We used the fast*STRUCTURE* "chooseK.py" script to decide the optimal K. As no genetic clusters were detected (optimal value of K was one), we partitioned the SNP dataset output by the 'populations' module according to collection time relative to the disease outbreak and current location (Figure 2.2B). These predetermined groups were: (1) wild individuals sampled before the disease

outbreak in 2007, *N* = 92 (hereafter “Before”); (2) wild individuals sampled after the disease outbreak between 2015-2020, *N* = 38 (hereafter “After”); (3) Taronga founders sampled in 2015, *N* = 16; and (4) Symbio founders sampled in 2017, *N* = 19. Variant filtering was carried out on three datasets: (i) all groups consisting of the four predetermined groups (*N* = 166), (ii) wild groups only (Before and After; *N* = 131), and (iii) captive populations only (Taronga and Symbio; *N* = 35). The SNP datasets were filtered on minimum average read depth (>2.5x), coverage difference between the reference and alternate alleles (<80%), call rate (>75%), retention of loci with heterozygosity <80%, and reproducibility (>90%) calculated using 51 technical replicates performed by DArT. Where samples were grouped across different sequencing plates, we found no evidence to suggest that batch effects influenced our results (Figure A1.3.3).

To investigate variation in the ‘wild’ and ‘all groups’ datasets, we undertook exploratory principal coordinates analyses (PCoA) using Euclidean distance via the R v4.3.0 package dartR v1.9 (Gruber et al., 2018, R Core Team, 2023) and visually inferred putative genetic differentiation between groups using eigenvalues in adegenet v2.1.3 (Jombart, 2008). We also applied a PCA to the dataset using the glPca() function in adegenet. We calculated pairwise fixation indices (F_{ST}) between all groups using hierfstat v0.5-11 (Goudet 2005). To identify relationships between geographic distance and genetic distance we used the Before dataset to perform an isolation-by-distance (IBD) mantel test in dartR using the gl.ibd() function with 999 permutations.

We calculated standardised heterozygosity (H_s) using the genhet function in R (Coulon, 2010) for all groups (where 1 is the average and so a value greater than 1 is more diverse than average); observed (H_o) and expected heterozygosity (H_E) using GenAlEx v6.5 (Peakall and Smouse, 2006) and visualised individual H_o distributions by group using the boxplot() functions in R. We calculated autosomal H_o and H_E (Schmidt et al., 2021) by re-running the Stacks ‘populations’ module with the parameter: minimum samples per population 75% (−r 0.75), and without the parameters: minimum minor allele frequency (MAF) of 0.01 (--min_maf 0.01); and --write_random_snp to retain both variant and invariant loci (Schmidt et

al., 2021). We filtered the Stacks output on minimum average read depth (>2.5x), coverage difference between the reference and alternate alleles (<80%), call rate (>75%), and reproducibility (>90%) calculated using 51 technical replicates performed by DArT. Loci on sex chromosomes (scaffold 4) were removed from the dataset (Martinez et al., 2008). The resulting 11,208 loci (at both variant and invariant sites) were used to calculate autosomal H_O and H_E using GenAlEx v6.5 (Peakall and Smouse, 2006).

We performed a t-test to test whether H_S was significantly different between wild groups before and after the disease and used the Bartlett's test of homogeneity of variances base function to test whether individual observed heterozygosity was significantly different from expected heterozygosity in R. We calculated inbreeding coefficients (F_{IS}) and the associated 95% CI using the diveRsity v1.9.9 package (Keenan et al., 2013) and PopGenReport v3.0.7 (Adamack and Gruber, 2014) to calculate allelic richness (A_R) in R. We calculated the number of private alleles in each group compared to all other groups and pairwise private alleles between groups using the gl.report.pa() function in the dartR package. To calculate molecular relatedness, we ran simulations in COANCESTRY v1.0.1.10 (Wang, 2011) to determine the most appropriate moments estimator as per Hogg et al. (2019). We selected TrioML for final analyses. We set COANCESTRY parameters to account for inbreeding, with the number of reference individuals and bootstrapping samples set to 100 for all groups and between captive groups. We calculated mean kinship (MK) by dividing the TrioML value by two, representing the average relationship of each animal to all others within the sample set. We estimated MK for individuals within each group (MK_{WITHIN}) and between captive groups ($MK_{BETWEEN}$) using the captive dataset. We calculated effective population size (N_e) for wild groups using NeEstimator v2.1 (Do et al., 2014) with values reported for the no singleton alleles analysis and the associated jack-knifed 95% confidence intervals (Jones et al., 2016). We excluded captive groups from N_e analyses due to small sample sizes resulting in infinite confidence intervals.

2.5 Results

Reference genome

The final genome assembly was 2.0 Gb in size, consisted of 129 scaffolds, had a contig N50 of 56.12 Mb, and scaffold N50 of 123.4 Mb (Figure A1.3.1A, Table A1.2.2). BUSCOv5.2.2 identified 95.4% complete vertebrata genes, 94.9% of which were single copy and 0.5% were duplicated, 2.1% were fragmented, and 2.5% were missing (Figure A1.3.1A, Table A1.2.2). The genome statistics and distinct chromosome-length scaffolds (Figure A1.3.1B) confirm chromosome level-completeness. We used this genome to align and call variants using the population genetics data. Comprehensive details of genome assembly and annotation are provided in the Supplementary Material.

Genetic analyses for management

We aligned the high density DArTseq data to the reference genome. Our initial analyses showed evidence of an *E. macquarii* (*N* = 1), *M. georgesi* and *E. macquarii* hybrid (*N* = 1), and backcross (*N* = 2) individuals in the wild groups (Figure A1.3.4) (Georges et al., 2018). In accordance with earlier findings, we did not detect hybrids in either captive population having been identified and removed when the populations were founded (Georges et al., 2018). The *E. macquarii*, *M. georgesi* and *E. macquarii* hybrid, and backcross individuals were removed from the dataset and SNPs were re-called. The re-called dataset yielded a Stacks output of 2,172 SNPs. Refiltering on SNP calls for each group in R resulted in reduced representation datasets of 460 genome-wide SNPs in the 'all groups' dataset, 473 genome-wide SNPs in the 'wild' dataset, and 227 genome-wide SNPs in the 'captive' dataset. Our exploratory PCoA using the wild only dataset revealed minor levels of genetic structuring between Before and After individuals, with up to 3.5% of variance explained by PCo I and 2.5% by PCo II (Figure 2.3A). The addition of captive individuals in the second PCoA, revealed clustering of captive groups with Before individuals (Figure 2.3B). Up to 5.3% of the variation was explained by PCoI, and PCo II (3.7% variation) with both axes primarily separating out After individuals (Figure 2.3B). An almost identical clustering result was seen when a PCA was applied to the dataset (data not shown).

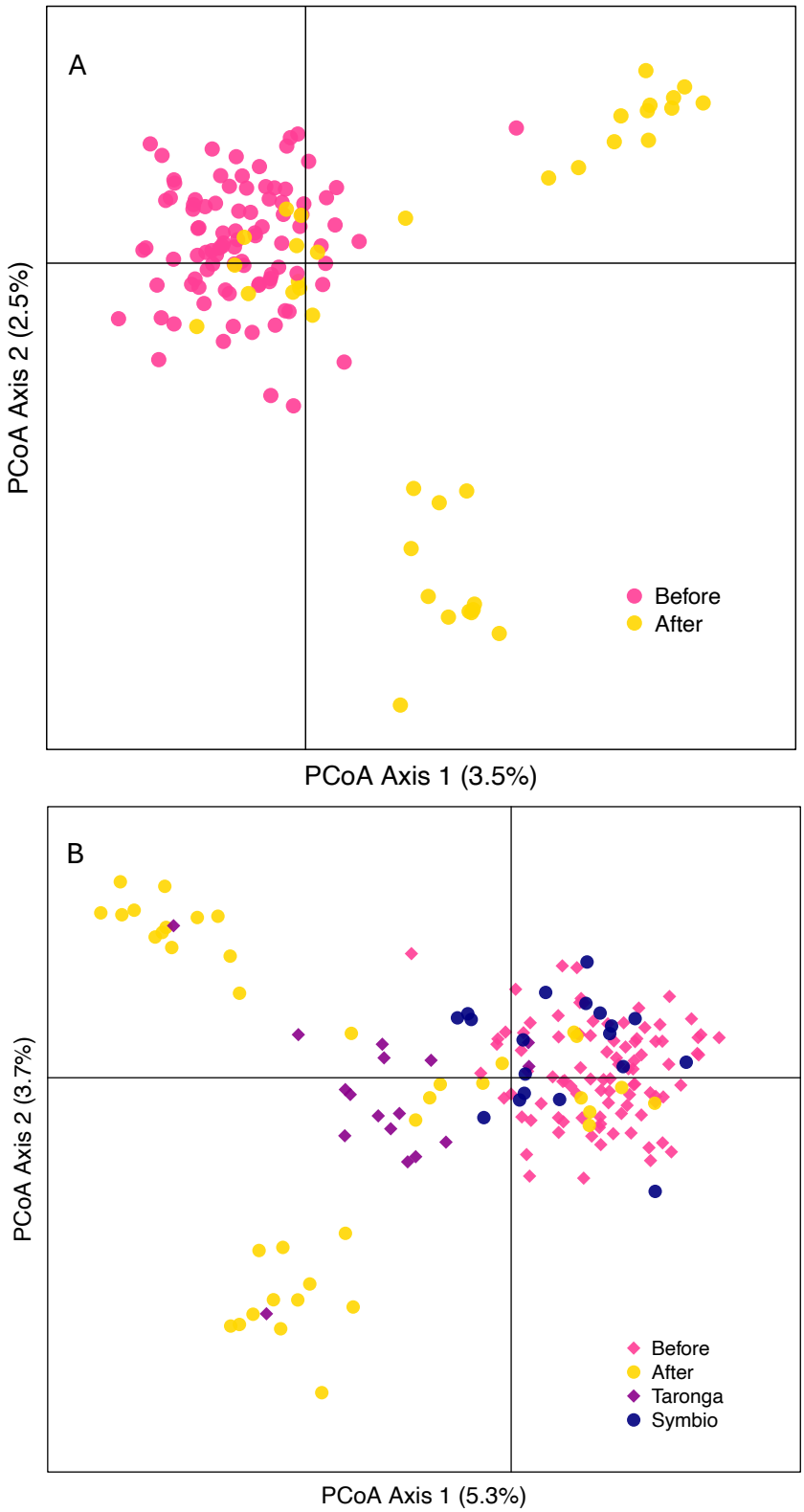


Figure 2.3. (A) Principal coordinates analysis (PCoA) of wild individuals before and after the disease outbreak (*N* = 131) using 474 genome-wide SNP markers. **(B)** PCoA of genome-wide diversity of all individuals (*N* = 166) using 460 SNP markers.

We found no notable differences in H_s between captive and wild groups (Table 2.2A) and no significant differences in the wild across time (t=-0.902, df=128, P=0.369). SNP and autosomal H_O and H_E was similar across all groups (Table 2.2A, Figure A1.3.5) with significantly higher observed than expected heterozygosity across all groups, indicating an excess of heterozygotes. There were no statistically significant levels of inbreeding (F_{IS}) observed as confidence intervals encompassed zero for all groups (Table 2.2A). A_R was also similar with no distinct differences across groups (Table 2.2A). P_A were only present in the Before group when comparing across all groups (Table 2.2A). Pairwise P_A numbers were consistently higher for Before and lowest for Taronga and Symbio, respectively (Table A1.2.3). MK_{WITHIN} ranged from 0.019±0.048 to 0.048±0.150, with the highest value observed in the After population (Table 2.2A). $MK_{BETWEEN}$ for Taronga and Symbio using the captive dataset was 0.012±0.040, where 0.1250 is the equivalent of a half-sibling relationship. N_e estimates for Before and After were 148.9 (120.4-191.8) and 11.9 (8.2-17.1) respectively (Table 2.2A). F_{ST} values between all groups ranged from 0.005 (0.001-0.009) to 0.027 (0.020-0.034), with statistically significant F_{ST} observed between the Before and After groups (Table 2.2B). Our IBD mantel test found no correlation between geographic distance and genetic distance in the species (r = 0.364, P = 0.258).

Table 2.2. (A) Population genetic indices of our sample groups, including standardised (H_S), observed (H_O), and expected (H_E) SNP and autosomal heterozygosity, inbreeding coefficient (F_{IS}), allelic richness (A_R), private alleles (P_A), mean kinship within populations (MK_{WITHIN}), and effective population size (N_e). SD = standard deviation, CI = 95% lower and upper confidence intervals. Significantly higher H_O to H_E denoted by *. **(B)** Population differentiation (F_{ST}) including 95% CI between all groups.

(A)

Group	n	H_S (± SD) SNP	H_O (± SE) SNP	H_E (± SE) SNP	H_O (± SE) Autosomal	H_E (± SE) Autosomal	F_{IS} (95% CI)	A_R	P_A	MK_{WITHIN} (± SD)	N_E (95% CI)
Before	92	1.007 (0.097)	0.290 (0.009)*	0.280 (0.008)	0.0180 (0.0007)*	0.0148 (0.0005)	-0.035 (-0.051--0.020)	1.727	11	0.019 (0.047)	148.9 (120.4-191.8)
After	38	0.990 (0.097)	0.286 (0.010)*	0.262 (0.008)	0.0170 (0.0006)*	0.0138 (0.0005)	-0.094 (-0.125--0.064)	1.688	0	0.048 (0.150)	11.9 (8.2-17.1)
Symbio	19	0.969 (0.102)	0.279 (0.011)*	0.252 (0.009)	0.0174 (0.0007)*	0.0136 (0.0005)	-0.181 (-0.149--0.075)	1.669	0	0.023 (0.072)	-
Taronga	17	1.009 (0.098)	0.293 (0.012)*	0.248 (0.009)	0.0165 (0.0007)*	0.0130 (0.0005)	-0.108 (-0.253--0.121)	1.658	0	0.027 (0.076)	-

(B)

	Before	After	Symbio
After	0.026 (0.020-0.031)		
Symbio	0.005 (0.001-0.009)	0.027 (0.020-0.034)	
Taronga	0.010 (0.004-0.016)	0.002 (-0.002-0.006)	0.019 (0.011-0.027)

2.6 Discussion

A worked example of how conservation managers can apply the framework

Please refer to Table 2.1 for a summary of the framework for integrating genetic data into conservation breeding management, particularly in relation to the genetic input, management outcomes and how this applies to our case study.

Stage 1–Recognising decline of the wild population and its genetic consequences

"When recognising declines in wild populations, the collection and preservation of DNA samples, such as blood and tissue, in a biobank can provide essential genetic data for future research" (Frankham et al., 2010). Through collection and preservation of DNA, we can also identify genetic consequences resulting from threatening processes through analysis of temporal data, as well as sample provision for development of a reference genome for neutral and adaptive genomic investigations (Sunde et al., 2022).

The reference genome for *M. georgesi* was created using tissue collected and stored in a biobank from 1986. The sampling by managers in 2007 and 2015-2022 (Figure 2.2B) has provided population genetic data for researchers to investigate the genetic consequences of the nidovirus outbreak and other threatening processes, including evidence of shifts in wild genetic diversity since the disease event (Table 2.2A, Figure 2.3A). We developed and utilised our reference genome during establishment stages 2 and 6 for Symbio and Taronga, respectively. If financial resources and collaborative opportunities are available, we recommend that a reference genome be developed during stage 1 for conservation breeding programs so genetic output can be utilised as early as possible and to improve the reliability of RRS data variant calls to improve downstream inferences (Torkamaneh et al., 2016, Shafer et al., 2017, Wright et al., 2019). We also recommend samples from the declining wild population be collected to provide baseline genetic data and to readily assess genetic consequences of population declines. This may encompass decreases in diversity, the risk of bottlenecks, the potential for inbreeding depression, and the occurrence of genetic drift.

Addressing these issues early is crucial to prevent the need for more extensive interventions due to a potential delayed response in observable genetic changes.

We examined population genetic data from samples that were previously sequenced for earlier analyses (Georges et al., 2018) (Figure 2.2B). Using our reference genome and set of reliable genome-wide SNP markers, we found no significant differences in H_s between all groups (P=0.369) indicating consistent levels of heterozygosity in the population since the nidovirus outbreak. SNP and autosomal observed heterozygosity were significantly higher than expected heterozygosity across all groups (Table 2.2A). While bottlenecks typically reduce both H_o and H_E without necessarily causing a deviation between them, this pattern may reflect a loss of rare alleles that disproportionately lowers H_E, or a transient excess of heterozygotes due to the population not yet reaching Hardy-Weinberg equilibrium following the mortality event. However, the large drop in N_e suggests a bottleneck is likely. The relatively low genome-wide diversity compared to other species also suggests historical bottlenecks have already occurred in this species (Georges, 2020). It has also been suggested that associative overdominance (neutral loci becoming effectively over-dominant as a result of disequilibrium with a locus under selection) may conserve genetic variation in small populations compared to expectations from neutral theory (Frydenberg, 1963, Rumball et al., 1994, Gilligan et al., 2005, Schou et al., 2017), potentially contributing to the significantly higher levels of observed to expected heterozygosity in all groups however, long-term monitoring is needed to confirm this. Notably higher SNP heterozygosity compared to autosomal heterozygosity estimates are due to the inclusion of invariant sites and the species small population size (Schmidt et al., 2021), with autosomal heterozygosity estimates likely providing a more accurate representation of low genome-wide diversity in the species (Schmidt et al., 2021). A_R was consistent across groups, with slightly lower values observed for the After, Symbio, and Taronga groups compared to Before (Table 2.2A). The lower number of unique P_A in the After group compared to Before highlights potential losses in alleles overtime, although discrepancies in sample size may not have captured the range of alleles currently present in the wild (Table A1.2.3). Although alleles present in the After group do not appear to be captured in captivity, each captive population contains multiple alleles not observed in the current wild population (Table A1.2.3). As there has been minimal

evidence of wild clutches hatching since the outbreak, the After group may still reflect the diversity of the larger pre-disease population suggesting that genetic consequences of the outbreak may not be evident until post-disease G1 individuals can be analysed. We have advised species managers that continual monitoring once wild individuals start reproducing will be crucial in identifying long-term trends.

In this case study, we demonstrate the value of biobanking during the early stages of a wild population decline, and we advocate for continual monitoring in threatened species including DNA samples collection and biobanking to capture temporal trends in genetic diversity.

Stage 2–Founding one or more captive populations

"When founding one or more captive populations, a fully representative sample that encompasses wild diversity is needed to maximise captive population viability" (Frankham et al., 2010). Frankham and colleagues proposed that 20-30 contributing founders are sufficient to create a genetically diverse population that is representative of wild diversity, although molecular analyses are needed to confirm this hypothesis. DNA samples from, (1) all founders are needed and, if feasible, (2) contemporary samples collected across the wild population should be sequenced. The data generated from these samples can be used in stages 3 and 4 for, (1) diversity and founder kinship analyses that can inform management decisions to maximise diversity in captivity, prevent inbreeding and to detect hybrids, or backcross, individuals among founders, and (2) contemporary wild samples that provide information to ensure that the genetic diversity of the wild population is reflected in captive individuals.

In line with stage 2 management outcomes, we observed minor but non-significant shifts in genetic structure since the outbreak, where Taronga and Symbio are shown to be most representative of historical diversity (Before) (Figure 2.3B). The clustering of captive groups with Before in the PCoA suggests that Taronga and Symbio represent the genetic profile of the once larger population which may be useful in reinforcing the current wild population. The small number of Taronga individuals that clustered with After suggests greater representation

of current wild variation within Taronga that is consistent with the low F_{ST} values between the two groups (Table 2.2B). However, as only a small amount of variation is explained by the PCoA axes (<4%), long-term analyses of the offspring of the outbreak survivors will be useful in confirming these trends. MK_{WITHIN} in the wild is higher after the disease outbreak with Taronga and Symbio falling between historical and contemporary levels. Populations that sustain high MK_{WITHIN} over generations are expected to experience more rapid changes in allele frequencies and lower adaptation potential in future (Frankham, 1996). In this instance, the conservation breeding program will play a crucial role in strengthening the wild population post-bottleneck by providing opportunities to mate with unrelated individuals, mitigating future inbreeding.

In our case study, molecular analyses to assess genetic representation of the captive populations was conducted post-founding due to a lack of genetic resources at the time. Ideally, stage 2 should be implemented at founding to quickly address genetic concerns prior to breeding. In situations where genetic analysis cannot occur at the time of founding, we recommend collection of DNA from all founders for future analysis.

Stage 3 –Expanding captive populations to a secure size

"During the expansion phase of conservation breeding, priority is on rapid population growth rather than intense genetic management" (Frankham et al., 2010). During stage 3, maintaining an accurate studbook is crucial but may not be effective enough to mitigate long-term founder effects and prevent inbreeding depression when assuming founders are unrelated (Ivy and Lacy, 2010, Ivy et al., 2009). During this stage, DNA samples from each generation and new intakes are essential for molecular identification of relatives to prevent long-term founder effects and inbreeding in subsequent stages (Hogg et al., 2019). This is important in cases where paternity determination can be difficult, including group-housed enclosures, in species where females retain sperm (Gist and Jones, 1987, Sever and Hamlett, 2002), and when mixed parental clutches are present. This information allows for more informed genetic selection of mates and alleviates the impact of founder effects in subsequent generations. Additionally, this information allows managers to monitor novel genetics

introduced into the population through new intakes and provides data for long-term monitoring of genetic drift.

Through expert collaboration and development of effective breeding protocols (Taronga Conservation Society Australia, 2023), the number of *M. georgesi* individuals in captivity has rapidly grown from 35 to *ca.* 299 individuals in eight years, increasing the global population size from <150 to approximately 450 individuals. Throughout stage 3, a studbook has been maintained to inform breeding, established before the availability of genetic data. Before our study, the DNA sequences of Taronga's founders and offspring were used for internal parentage analysis to inform the studbook (Georges, 2020). As the turtles are typically housed in groups where females have been observed to retain sperm, determining paternity through traditional means may sometimes be challenging with the absence of molecular information. Due to low numbers in the wild, there have been no new founder intakes to the captive program. In accordance with the species' conservation action plan (Jakob-Hoff et al., 2017), we anticipate DNA from new wild individuals adopted into the breeding program will be analysed and integrated into the studbook (Table A1.2.4, Table A1.2.5).

Our case study and previous work demonstrates that rapid population growth can go hand in hand with the integration of traditional pedigree and molecular genetic data.

Stage 4 – Managing captive populations over generations

"Loss of genetic diversity and inbreeding is exacerbated over generations in small captive populations" (Frankham, 1995). When managing captive populations over generations, focus shifts from rapid reproduction in stage 3 to mitigating genetic issues in stage 4" (Frankham et al., 2010). Samples collected in stage 3 can be used to explore a standardised set of genetically measured Essential Biodiversity Variables in stage 4 (H_S, H_E, H_O, F_{IS}, A_R, P_A, N_e and F_{ST}) to provide comparable data for monitoring and management (Hoban et al., 2022). It may be preferable to undertake these analyses during stage 3 to identify genetic changes early on and ascertain if genetic interventions are required, such as the introduction of new individuals to

increase genetic diversity, or breeding between populations to reduce inbreeding (Hoffmann et al., 2021, Kinghorn and Kinghorn, 2021). In cases where founder relationships are unclear during initial breeding, it is essential to retrospectively incorporate these data to help inform future breeding decisions (Hogg et al., 2019), and to apply genetic principles such as not breeding individuals captured close together, until genetic data can be incorporated.

After population expansion and establishment of a studbook in stage 3, *M. georgesi* managers sought genetic expertise for continued genetic monitoring of captive individuals. This was an opportune time as individuals only had one generation of captive ancestry, allowing our data to be proactively implemented to mediate future genetic issues. Given that the lifespan of *M. georgesi* likely surpasses 30 years and individuals typically do not reach sexual maturity until approximately 6-12 years of age, this may prove more challenging in species with shorter lifespans and short generation times. With access to genetic data, our analyses revealed consistent genetic diversity metrics in both Taronga and Symbio, with Taronga showing slightly higher diversity metrics throughout all analyses (Table 2.2A, Table 2.2B).

MK values are useful for determining how related individuals are within, and between, populations and for indicating which animals should be considered for breeding/translocation to minimise relatedness and maximise genetic diversity (Frankham et al., 2017b). When choosing breeding pairs, individuals with high MK estimates should not be paired, and breeding of their progeny should be carefully considered. To assist the stage 4 outcome of managing captive populations over generations using our MK analyses (Table A1.2.4, Table A1.2.5), we suggest breeding individuals that are on average less related to each other within Taronga and Symbio to prevent founder effects and inbreeding depression (Lacy, 1987, Frankham, 2008, Frankham et al., 2010). As we observed lower levels of MK between ($MK_{BETWEEN}$) Taronga and Symbio compared to within (MK_{WITHIN}) (Table 2.2A, Table A1.2.4, Table A1.2.5), we have also suggested to the recovery team and captive managers that breeding between Taronga and Symbio (when sexually mature) could result in a decrease in average kinship, a reduction in inbred individuals, and an increase in diversity (Hoffmann et al., 2021, Kinghorn and Kinghorn, 2021), as observed in other species (del Mar Ortega-

Villaizan et al., 2011, Thavornkanlapachai et al., 2021). Given that only Taronga individuals have reached reproductive maturity and animals born in captivity (G1) have yet to breed, genetic monitoring of each generation is necessary to detect any genetic drift in captivity (Gilligan and Frankham, 2003) in addition to diversity and MK monitoring. This is being implemented through genetic sampling of captive born individuals.

We demonstrate in our case study how the use of genome-wide SNP data can provide fine-scale insights to support mitigation of genetic issues in captive populations. The generation of a reference genome in this study will also provide support for future investigations into adaptive potential including immune gene diversity.

Stage 5 – Choosing individuals for reintroduction/release

To choose individuals most suitable for release, genetic information from stages 2-4 can be used to select individuals with wide-ranging levels of differentiation and/or low levels of relatedness. Before translocating captive individuals, it is essential to understand the genetic metrics of both the wild and captive population to prevent the introduction of identical or closely related genotypes from captivity into the wild, particularly when there may be more diverse individuals available in captivity (Hogg et al., 2020). Analyses of wild populations in stages 1-2 can be used to inform release site selection by releasing individuals at sites where genetic differentiation within a species or population is high and inbreeding or outbreeding depression is unlikely (Grueber et al., 2018, Nistelberger et al., 2023). We recommend performing this genetic admixture to maximise diversity in the wild but using populations that have exchanged genes within the last 500 years and where there is little evidence of local adaptation to minimise the risk of outbreeding depression (Frankham et al., 2011).

For *M. georgesi*, the availability of contemporary wild samples meant that we could identify that the combined captive populations are representative of historical wild diversity before the outbreak (Figure 2.3B) (stage 2). As such, releases from both Taronga and Symbio will be essential in supplementing pre-disease diversity into the contemporary wild population.

Additionally, our MK estimates provide data for breeding decisions in the captive breeding program and minimising the release of inbred individuals into the wild (Table A1.2.4, Table A1.2.5). To date, site selection for release has not been genetically informed due to insufficient data and has instead been selected based on ease of access and ability to obtain landholder approvals.

Our case study supports the use of genetically informed translocations by providing insight on the genetics at different sites throughout the river (Figure 2.1, Figure 2.3B). For example, the release of Taronga and eventually Symbio individuals at the opposing end of the river from their collection site (Figure 2.1A) could facilitate admixture with genotypes in lower and upper regions, respectively.

Stage 6 – Managing the reintroduced population in the wild

Due to persistent threats in their natural environment, reintroduction of certain species or reinforcement of wild populations might not be possible for an extended period. However, over time, captive populations may undergo genetic changes that make them better suited to their captive conditions but less adapted to the challenges of the wild (Frankham, 2008, Christie et al., 2012). This adaptation to captivity can result in reduced fitness in the wild due to a loss of genetic diversity, the accumulation of deleterious mutations, or the fixation of alleles that are advantageous in captivity but not in the wild (Lacy, 1987, Frankham et al., 2017). To minimise genetic adaptation to captivity and ensure long-term viability of the remaining wild population (Frankham, 2008, Crates et al., 2023), it is commonly suggested that releases be carried out within a few generations (Williams and Hoffman, 2009). To evaluate the effects of conservation translocations on genetic diversity, it is crucial to have ongoing monitoring of the reintroduced or reinforced population, at least until the population is self-sustaining with an improved conservation status (IUCN/SSC, 2013). Apart from techniques like radiotracking, this requires the collection of DNA samples from the wild population over numerous generations and analysing them for any changes. For example, this can be achieved by mirroring the temporal comparative analyses in this study (2.2A, Table 2.2B).

It is recommended that 1-10 individuals every 2-3 generations is sufficient in increasing genetic diversity (Allendorf, 1983, Lacy, 1987, Lande and Barrowdough, 1987, Backus et al., 1995). Prior to reinforcing the wild population, the *M. georgesi* recovery program undertook measures including predator management, habitat restoration, water quality assessments, and community engagement to minimise environmental stressors (Jakob-Hoff et al., 2017). The program has conducted four rounds of reinforcements by releasing 82 F1 juveniles of unknown sex into the Bellinger River (Figure 2.2A), significantly boosting wild population numbers.

To evaluate the short-term survival and movement of the released individuals, a select number have been radio tracked. Additionally, annual surveys are conducted to sample animals, with DNA biobanked for future sequencing. Given that genetic changes may not be immediately apparent in a long-lived species, long-term monitoring with a focus on sampling new and juvenile individuals, is critical for detecting alterations in the species. Long-term genetic monitoring will also play a critical role monitoring levels of hybridisation and introgression which simultaneously threatens the persistence and locally adapted genetic identity of the species (Georges et al., 2018).

Integration of genetic data into management actions

In combination with husbandry, breeding, and disease mitigation, incorporating genetic data throughout conservation breeding establishment and management is valuable for maximising long-term viability in captivity and to guide conservation translocations to the wild. Here we have provided a worked example of geneticists working with an established breeding program to answer genetic questions posed by managers using DNA samples they provided. We communicated our findings to managers during scheduled meetings and in-person workshops where our recommendations have been integrated into the species Conservation Action Plan (Jakob-Hoff, *unpublished*).

In summary, for the Bellinger River turtle we recommend that genetic diversity be maintained, or potentially increased, by:

(i) Breeding individuals that are on average less related to each other within Taronga and Symbio to prevent founder effects and inbreeding depression.

(ii) Breeding between Taronga and Symbio to minimise founder effects and inbreeding depression.

(iii) Encouraging translocations from both Taronga and Symbio to ensure representation of genotypes from both captive populations in the wild.

(iv) Alternating or varying release locations during each release round to facilitate genetic admixture between captive bred and wild individuals as there is some genetic differentiation between the wild and captive populations (Figure 2.3).

Here we aimed to provide a generalised checklist based on our own example of genetic data integration, that can be adapted for other conservation breeding programs. Integration of genetics into management activities may need to occur retrospectively depending on the establishment stage of the conservation breeding program when the genetic data is generated (Table 2.1). Instigating discussions between scientists and managers can facilitate productive dialogue, allowing for a better comprehension and adaptation of each other's work and tools (Hogg et al., 2017b). This cyclical approach leads to ongoing improvement and enhancement of conservation strategies.

Future directions

As emerging infectious diseases are causing rises in extinction risk (Piotrowski et al., 2004, Blehert et al., 2009), our capacity to understand genetics and genomics is also increasing. The DNA of disease-susceptible species provides valuable insight on species resilience, with genomics giving us the tools to unlock the answers. For *M. georgesi,* as translocation candidates are juveniles, there remains an unknown risk that once individuals mature, they will succumb to the virus in the wild, compromising the translocation program. Another unknown threat is whether the few adult survivors of the virus are genetically predisposed to resist the virus, or they merely avoided exposure (Zhang et al., 2018). To answer these

questions, our high-throughput sequencing efforts and chromosome-level genome provides a valuable genomic tool for future functional gene research on *M. georgesi* and other Australian turtle species. Beyond neutral diversity investigations, high-throughput and genome-wide sequencing provides high-resolution data for immune-gene investigations, that rely on high-quality assemblies and genome-wide data (Peel et al., 2022). Each year the cost of sequencing and bioinformatic analyses becomes cheaper and more streamlined permitting studies like ours (Wright et al. 2019). Once streamlined sampling and basic genetic management become conventional practice in breeding programs, genetic research can progress towards higher resolution conservation efforts aimed at preserving the adaptive potential and functional diversity within captive programs.

Chapter 3 - A genomic-based workflow for eDNA assay development for a critically endangered turtle, *Myuchelys georgesi*

3.1 Manuscript

Published in *Ecology and Evolution*

A genomic-based workflow for eDNA assay development for a critically endangered turtle, *Myuchelys georgesi*

Holly V. Nelson[1], Jarrad Prangell[2], Arthur Georges[3], Elspeth A. McLennan[1], Jane L. DeGabriel[4], Katherine Belov[1,5], and Carolyn J. Hogg[1,5 *]

[1] School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia.

[2] Symbio Wildlife Park, Helensburgh, NSW 2508, Australia

[3] Institute for Applied Ecology, University of Canberra, Bruce, ACT 2617, Australia

[4] NSW Department of Climate Change, the Environment, Energy and Water, Parramatta, NSW 2150, Australia

[5] Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Sydney, Sydney, NSW 2006, Australia

3.2 Abstract

Environmental DNA (eDNA) analysis has become a popular conservation tool for detecting rare and elusive species. eDNA assays typically target mitochondrial DNA (mtDNA) due to its high copy number per cell and its ability to persist in the environment longer than nuclear DNA. Consequently, the development of eDNA assays has relied on mitochondrial reference sequences available in online databases, or in cases where such data are unavailable, de novo DNA extraction and sequencing of mtDNA. In this study, we designed eDNA primers for the critically endangered Bellinger River turtle (*Myuchelys georgesi*) using a bioinformatically assembled mitochondrial genome (mitogenome) derived from a reference genome. We confirmed the accuracy of this assembled mitogenome by comparing it to a Sanger-sequenced mitogenome of the same species, and no base pair mismatches were detected. Using the bioinformatically extracted mitogenome, we designed two 20 bp primers that target a 152-base-pair-long fragment of the cytochrome oxidase 1 (CO1) gene and a 186-base-pair-long fragment of the cytochrome B (CytB) gene. Both primers were successfully validated *in silico*, *in vitro*, and *in situ*.

3.3 Introduction

In recent years conservation geneticists have made substantial progress in understanding how to apply genetic data to conservation actions for threatened species (Hohenlohe et al., 2021). The prevalence of cost-effective, non-invasive molecular tools like environmental DNA (eDNA) assays have become increasingly common in detecting invasive species, assessing community diversity across various spatial scales, and monitoring rare or cryptic species (Rees et al., 2014, Ardura et al., 2015, Ruppert et al., 2019, Lam et al., 2022). eDNA refers to extra organismal genetic material the comprises of molecules that have been shed into the environment by decaying bodies, leaves, blood, pollen, seeds, urine, faces, skin, hairs, and other types of organismal material (Freeland, 2017), that can be extracted from environmental samples such as soil, water and air (Barnes et al., 2014, Rees et al., 2014). The presence of eDNA can be detected using DNA metabarcoding for detection of entire communities or species-specific primers or assays to detect a target species (Mauvisseau et al., 2019, Lopes et al., 2021, Valdivia-Carrillo et al., 2021). eDNA assays commonly target and amplify a short fragment of mitochondrial DNA (mtDNA) through polymerase chain reaction (PCR). mtDNA is commonly targeted as it is highly abundant in cells and can persist in environments longer than nuclear DNA (nuDNA) (Wilcox et al., 2016, Bylemans et al., 2018).

The development of species-specific eDNA markers relies on the species' mtDNA being sequenced. Conventional methods for acquiring mtDNA sequence data have involved tissue acquisition, DNA extraction, designing universal primers or primers of a closely related species, long-range polymerase chain reactions (PCRs), shotgun sequencing, followed by bioinformatic assembly (Kundu et al., 2020, Chen et al., 2021, Tessler et al., 2023). The advent of high-throughput parallel sequencing (HTS), reductions in sequencing costs, and lower input DNA requirements, as well as improved bioinformatic pipelines, have given rise to the genomics era where traditional genetic approaches are being replaced by whole-genome approaches to conservation genetic research (Satam et al., 2023). While genomic data alone have no direct impact on conservation outcomes, they provide a foundational blueprint that that can be harnessed by geneticists and conservationists for a range of downstream applications (Hogg et al., 2022). These can include; aiding in the identification of genetic

variants for population genetic analyses (Brandies et al., 2019); investigations into functionally important genetic variation such as immune genes (Peel et al., 2022), development of PCR primers and recently *in silico* extraction of complete mitochondrial genomes (hereafter "mitogenomes") (Meng et al., 2019, Uliano-Silva et al., 2021).

The Bellinger River turtle (*Myuchelys georgesi*) is a species of short-necked turtle (Family Chelidae) and is one of two turtle species that is listed as Critically Endangered in Australia under the Environmental Protection and Biodiversity Conservation Act 1999 (Commonwealth of Australia 1999). The species has a current known distribution that is restricted to 60 km of the Bellinger River and a short section of its main tributary, the Kalang River, in north-eastern New South Wales (NSW), Australia (Cann et al., 2015). However, the species has not been recorded in the Kalang since 2007 (Georges et al., 2011). *M. georgesi* is a rare and cryptic species that has adapted to up-stream regions of the Bellinger River, preferring deeper waterholes surrounded by bedrock making them difficult to survey using conventional diving and trapping methods (Spencer et al., 2014). In 2015, a novel nidovirus outbreak resulted in the estimated death of more than 90% of individuals, further contributing to the species' rarity (Zhang et al., 2018, Chessman et al., 2020). The species also faces threats from competition with another locally occurring species, the Murray River turtle (*Emydura macquarii*). Implementation of eDNA analyses in both known and data deficient areas of the catchment (e.g. Kalang River) is currently listed in the species Conservation Action Plan to inform survey site selection (Jakob-Hoff R et al., unpublished), yet no such tool currently exists.

Given the growing application of both eDNA and genomic data in conservation management, we used a PacBio HiFi reference genome to develop species-specific eDNA markers for *M. georgesi.* We also provide comprehensive methodologies and a visual workflow for other threatened species, with reference genomes or genomic data, that would benefit from an eDNA assay using this approach.

3.4 Methods

Mitogenome Assembly

We previously assembled a chromosome-level reference genome for *M. georgesi* using PacBio High Fidelity (HiFi) (CA, United States) sequencing (Nelson et al., 2024). HiFi sequencing is a type of long-read data that is generated by circular consensus sequencing (CCS). Raw CCS reads can be as long as 15,000 to 20,000 base pairs, allowing full-length mitogenome sequences to be captured within a single read. To generate HiFi sequence data, high molecular weight DNA was extracted from the heart tissue of a male *M. georgesi* using the Nanobind Tissue Big DNA kit following the manufacturer's protocol (Circulomics, Pacific Biosciences, California, United States of America). PacBio HiFi Single-Molecule Real-Time (SMRT) bell libraries were sequenced at the Australian Genome Research Facility (Brisbane, Australia). The HiFi genome was assembled using Hifiasm v.0.16.0 (Cheng et al., 2021). To obtain a complete mitogenome (i.e., the entire mitochondrial DNA), we bioinformatically extracted the mtDNA sequence from the HiFi genome fasta file (a text-based file format containing nucleotide sequences) using MitoHifi v2 (Uliano-Silva et al., 2021). The -c flag was used to identify and annotate the mitogenome from genome scaffolds, rather than assembling it from raw reads with the -r flag. MitoHiFi also requires a mitochondrial reference sequence as input in either fasta or GenBank format (e.g. https://www.ncbi.nlm.nih.gov/genbank/samplerecord/). MitoHiFi provides an internal script (findMitoReference.py) that can be used to find and download the most closely related mitogenome for the species of interest. For this study, we manually obtained reference sequences from the NCBI for the Green Sea Turtle (*Chelonia mydas*) (NC_000886) (Kumazawa and Nishida, 1999), Murray River Turtle (*Emydura macquarii*) (NC_041302.1) (unpublished), and a previously Sanger sequenced *M. georgesi* mitogenome (NC_042474.1) (unpublished). These sequences were used to evaluate whether levels of divergence between reference and target species affected assembly quality. The mitogenome was visualised using Proksee (Grant et al., 2023) (Figure 3.1). To confirm efficacy of the bioinformatically extracted mitogenome, we used MEGA11 (Tamura et al., 2021) to align the assembly to a Sanger sequenced *M. georgesi* mitogenome for structural comparison and mismatches between sequences (NC_042474.1).

Species-specific primer development and validation

Using the annotated fasta file output by MitoHiFi, we located genetic sequences labelled "CO1" and "CytB" and used the complete sequence (Figure A2.3.1, Figure A2.3.2) as input into Primer3Plus v3.3.0 (Untergasser et al., 2012) to design forward and reverse primer sequences with 0 base pair mismatches with the CO1 and CytB gene sequences. These genes were used as they are known to be highly variable among closely related species providing greater specificity for species-specific eDNA assays compared to mitochondrial genes with lower inter-specific variation (Moritz et al., 1987, Meyer, 1994, Johns and Avise, 1998, Hebert et al., 2003). Forward and reverse primers output by Primer3Plus were individually input into OligoAnalyzer (https://sg.idtdna.com/calc/analyzer) for quality checks using the hairpin and homodimer options to ensure efficiency and sensitivity of primer binding. To ensure primer stability and minimise the likelihood of hairpin structure formation (when complementary base-pair sequences create a loop), we used a ΔG (Gibbs free energy change) threshold of -4.5 kcal/mol. For homodimers (annealing of identical primer sequences) we ensured primers had no more than 3 complementary bases. The melting temperature (T_m) for all primer sequences fell between 59.6˚C – 60.1˚C . Final primer pairs are provided in Table 3.1 and Table A2.2.1. The specificity and sensitivity of primer sets were evaluated at three stages: *in silico*, *in vitro*, and *in situ*.

In silico validation

To confirm specificity *in silico*, the alignment search tool Basic Local Alignment Search Tool (BLAST) was used to confirm percent of sequence similarity with other species (https://blast.ncbi.nlm.nih.gov/). To visually confirm specificity and optimal primer design against another locally occurring species (*E. macquarii*) we used MEGA v11 to align both our assembled and *E. macquarii* mitogenomes (NC_041302.1) and ensured 2-3 mismatches between primer design sequences (De Brauwer et al., 2022b). Custom primer sets were ordered using ThermoFisher Scientific (MA, United States) custom DNA oligos synthesis service.

In vitro validation

To evaluate specificity and amplification efficiency *in vitro*, we carried out tests using tissue-derived DNA from *M. georgesi* and *E. macquarii* (Table A2.2.2) using conventional PCR. Heart tissue belonging to a female *M. georgesi* that required medical euthanasia in 2021 (C10031) was flash frozen at -80°C at Taronga Zoo and stored at -80°C at the University of Sydney. *E. macquarii* skin tissue was acquired from the trailing webbing of the hindfoot of a wild individual in 2015 (UC<Aus>AA063724) and stored at -80˚C in the University of Canberra Wildlife Tissue Collection (GenBank UC<Aus>). To prevent contamination during lab procedures, equipment was decontaminated in an autoclave and benchtops cleaned with 80% ethanol. DNA (A2.2.3) was extracted using the Qiagen DNeasy Mini Kit (Qiagen, Germany) following the manufacturer's protocol, except for a final elution in 100 µL buffer AE (Qiagen). Quality (fragmentation) and concentration of DNA were assessed using a combination of a Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific) and 1.5% agarose/TBE gel electrophoresis stained with SYBR safe (Life Technologies), alongside a 1 kb size standard (Bioline) and run for 55 min at 100 V. Samples yielding high concentrations of DNA were used for subsequent PCR amplification assays.

For PCR set-up, 0.25 µM of CytB and CO1 forward and reverse primers were used. 0.25 µM of 12Sv5F/12Sv5R universal vertebrate primers were used as a positive control by amplifying a ~100 bp fragment of the V5 loop of the 12S mitochondrial gene (Riaz et al., 2011). The final PCR reaction consisted of 3 µL of *M. georgesi* DNA template or negative extraction control (*E. macquarii* DNA template, ddH_2O), 25 µL of Bioline MyTaq Mix (Bioline, UK), 2.5 µM of forward and reverse primers (either 12Sv5, CytB or CO1), and 17 µL of nuclease free water to make a total volume of 50 µL.

Real-time PCR cycling was carried out on a T100 Thermal Cycler (BioRad). Cycling conditions were 10 min for enzyme activation at 95˚C , 35 cycles of denaturation at 95˚C for 30 sec, annealing at 50˚C for 30 sec, extension at 72˚C for 30 sec, and a final extension at 72˚C for 10 min. Amplification was confirmed using 1.5% agarose/TBE gel electrophoresis stained with SYBR safe (Life Technologies), alongside a 1 kb size standard (Bioline) and run for 55 min at

100 V. Bands were visualised under ultraviolet light using a ChemiDoc XRS+system (BioRad) and images were analysed with ImageLab (BioRad).

In situ validation

M. georgesi eDNA water samples (positive controls) were obtained from three 4000 L, closed-system tanks at Symbio Wildlife Park in Helensburgh, Australia, each housing four or five animals. From each tank, two 500 mL water samples were collected. For negative controls, we collected two 500 mL water samples from a 2000 L pond containing four *E. macquarii* and two Eastern long-necked turtles (*Chelondina longicollis*). We transported the water samples on ice and stored them briefly at -2 ˚C before filtering within 1-2 hours of collection. Negative control samples were handled and stored separately to prevent contamination.

A 47 mm Whatman membrane filter paper with a pore size of 0.45 µm was dampened with deionised water before the 500 mL water samples were filtered through. The filtration system included a 50 mm Büchner funnel, adaptor, 500 mL Büchner flask, rubber tubing and a diaphragm pump (KNF, California, USA). The filter papers were then placed in individual resealable bags and frozen at -80˚C prior to DNA extraction the following day.

eDNA extractions were conducted on samples (Table A2.2.2) using Qiagen's DNeasy Blood & Tissue Kit (Qiagen, Germany). DNA extraction followed the protocol of Renshaw et al. (2015) with minor adjustments. Briefly, each filter paper was halved and finely cut before being placed in separate 2 mL screw-cap tubes. 540 µL buffer ATL and 60 µL (rather than the recommended 180 and 20 µL, respectively) of Proteinase K were added to submerse each half filter and incubated at 65˚C for one hour. Following lysis, the paper was tightly pressed to the bottom of the tube, and supernatant transferred to a new 2 mL screw-cap tube. 630 µL Buffer AL and 630 µL of ethanol were added and mixed thoroughly with a vortex. The lysates from each half were then combined by passing the mixtures through the same DNeasy Mini spin column, resulting in six rounds of centrifugation and discarded flow-through. Total eDNA was rinsed with 500 µL of AW1 and AW2 solutions respectively and eluted in 100 µL buffer AE

(Qiagen). eDNA concentration was quantified using a Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific). All eDNA extractions were placed in a freezer (-20˚C) for 12 hr until PCR analysis.

Following our *in vitro* validation protocol, 0.25 µM of CytB and CO1 forward and reverse primers were used for PCR set-up. 0.25 µM of 12Sv5F/12Sv5R universal vertebrate primers were used as a positive control. PCR mixes consisted of 3 µL of *M. georgesi* eDNA template or negative extraction control (*E. macquarii* and *C. longicollis* eDNA template or ddH_2O), 25 µL of Bioline MyTaq Mix (Bioline, UK), 2.5 µM of forward and reverse primers, and 17 µL of nuclease free water to make a total volume of 50 µL. Real-time PCR cycling, agarose gel electrophoresis, and image analysis was conducted using the same methods described in *in vitro* validation above.

3.5 Results

Myuchelys georgesi mitogenome

The complete mitochondrial sequence was extracted from scaffold 9 of our reference assembly and yielded a complete length of 16,490bp (Figure 3.1). The same mitogenome was assembled when the Green Sea Turtle, Murray River Turtle and Bellinger River Turtle mitogenomes were used as reference sequences, confirming that varying levels of divergence between reference input and target species does not affect final assembly quality. The size and structure of the mitochondrial genome is comparable to other chelid turtles (Fielder et al., 2012, Zhang et al., 2017), which includes 37 genes consisting of 22 transfer RNA (tRNA) genes, 13 protein coding genes, 2 ribosomal RNA (rRNA) genes, plus a non-coding region (CR) (Table A2.2.4). Additional details can be found in the Appendix 2. Visual alignment of the bioinformatically assembled mitogenome to the Sanger sequenced *M. georgesi* mitogenome (NC_042474.1) using MEGA v11 showed a 100% sequence identity match, confirming efficacy of the *in silico*-based mitogenome.

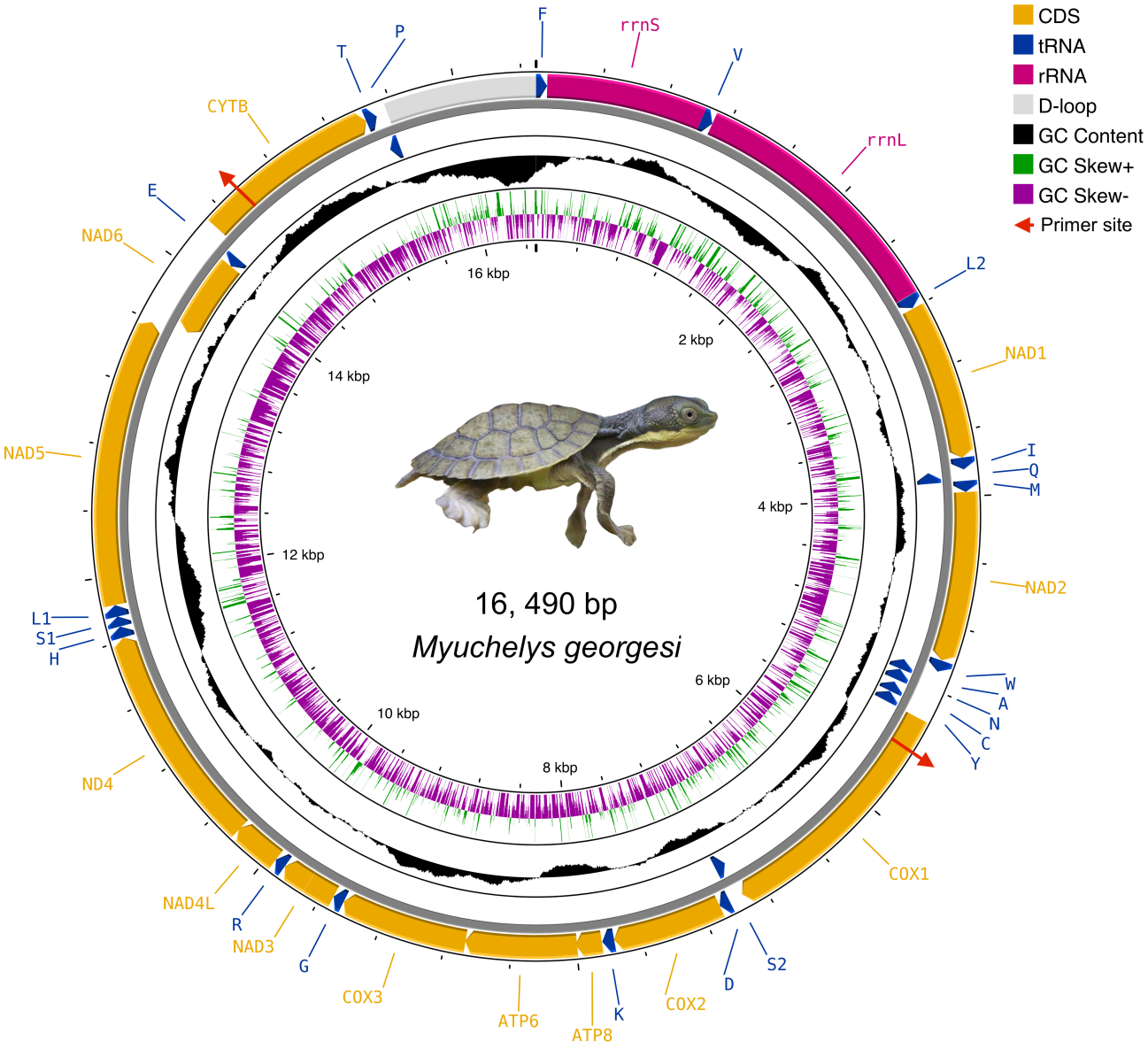


Figure 3.1. The mitochondrial genome of *Myuchelys georgesi* extracted using MitoHiFi (Uliano-Silva et al., 2021). tRNAs are labelled according to their single-letter abbreviation. Arrows indicate direction of gene transcription. Protein coding genes are shown in yellow, rRNA genes in pink, tRNA genes in blue, and the 920bp non-coding region between P and F in white. The GC-skew depicting the deviation from the average in the complete mitogenome is depicted in green (positive) and maroon (negative), and the GC content is depicted in black. The figure was generated using the CGview server (http://stothard.afns.ualberta.ca/cgview_server/). *M. georgesi* photo credit to Paul Fahy.

Primer design and validation

In silico primer assessment found greater species-specificity of the CO1 primers compared to CytB as BLAST results returned lower percentage identity with other species. Both CO1 and

CytB primers successfully amplified *M. georgesi* tissue samples (Figure 2A; lanes 1–2, 4–5, 7–8). Both sets of primers showed no amplification on *E. macquarii* tissue (Figure 2A; lanes 10–11 and 13–14), confirming the species-specificity of the primers against the other established, locally occurring species. The 12Sv5 positive control amplified across both species, indicating the presence of mitochondrial DNA in the tissues (Figure 2A; lanes 3, 6, 9, 12 and 15), while no amplification was observed for the ddH2O negative controls (Figure 3.2A; lanes 16–17). *In situ* evaluation showed PCR products for both primers successfully amplified *M. georgesi* eDNA collected on cellulose ester filters from tank water (Figure 3.2B; lanes 1–2, 4–5, 7–8). Primer sets did not amplify eDNA from tank water containing *E. macquarii* or *C. longicollis* (Figure 2B; lanes 10–11), confirming species-specificity of primers against other locally occurring species. Amplification of the positive 12Sv5 control across tank water confirmed the presence of mtDNA in all samples (Figure 2B; lanes 3, 6, 9 and 12) while no amplification was observed for the ddH2O negative control (Figure 2B; lanes 13–14).

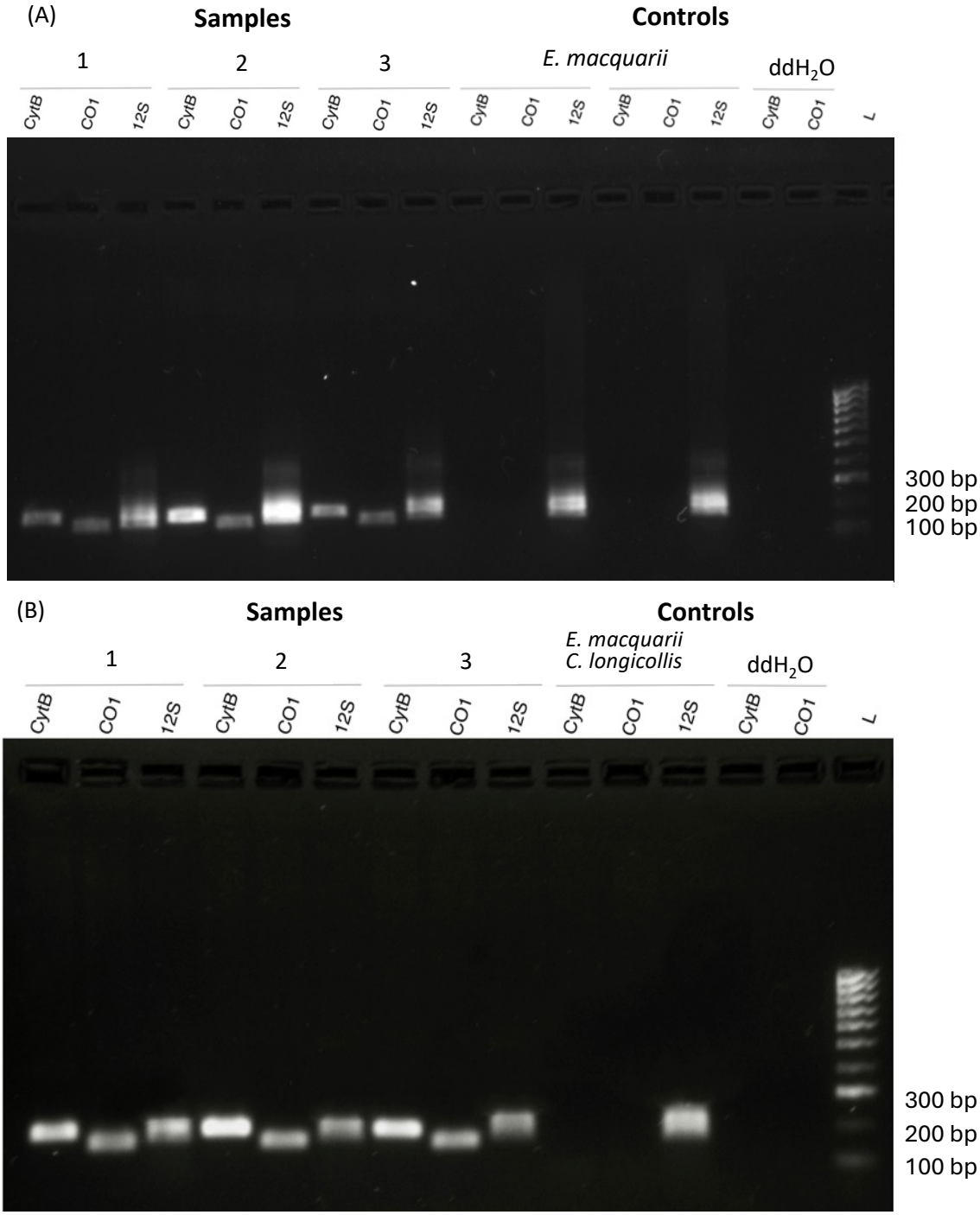


Figure 3.2. 1.5% agarose gel and TBE stained with SYBR safe, showing **(A)** *in vitro* amplicon products of tissue derived DNA for *Myuchelys georgesi* with CytB, CO1 and 12Sv5 control (lanes 1–9), *Emydura macquarii* with CytB, CO1 and 12Sv5 control (lanes 10–15), and ddH_2O with CytB and CO1 (16–17). **(B)** *In situ* amplicon products of tank water derived eDNA for *M. georgesi* with CytB, CO1 and 12Sv5 control (lanes 1–9), *E. macquarii* and *Chelondina longicollis* with CytB, CO1 and 12Sv5 control (lanes 10–12), and ddH_2O with CytB and CO1 (lanes 13–14).

Table 3.1. Primers designed (CytB and CO1) and used (12S) in this study for amplification of *Myuchelys georgesi* mitochondrial eDNA. T_m Melting temperature

Gene	Name	Forward/ Reverse	Nucleotide sequence	Primer length (bp)	Amplicon size (bp)	T_m (˚C)
CytB	MG_CB	Forward	AATCTCCCACATCCAACGAG	20	186	59.9
		Reverse	ATGCGGTGGCTATGACTAGG			60.1
CO1	MG_C1	Forward	ACATTGGCACCCTCTACCTG	20	152	60
		Reverse	AATTAAGGCGTGGGCTGTAA			59.6
12S	12Sv5	Forward	TAGAACAGGCTCCTCTAG	18	~100	(Riaz et al., 2011)
		Reverse	TTAGATACCCCACTATGC			

3.6 Discussion

We developed the first eDNA markers for detection of *M. georgesi* using an existing long-read PacBio HiFi reference genome. The 100% sequence identity match between the Sanger sequenced and bioinformatically assembled mitogenome (NC_042474.1) and successful amplification of mtDNA across *in silico*, *in vitro*, and *in situ* validations highlights the efficacy of genomic data-derived mitogenome assemblies, without the need for targeted mitochondrial DNA tissue extraction and sequencing.

We provide comprehensive methodologies for our workflow for other taxa that may benefit from this approach (Figure 3.3). Conventional approaches rely on the availability of mitochondrial sequence data from online databases or de novo extraction, sequencing and assembly when sequence data is not available (Schmidt et al., 2016, Zhang et al., 2017, Kundu et al., 2019, Frandsen et al., 2020, Kundu et al., 2020, Chen et al., 2021). For conservation programs with genomic resources but lacking mitochondrial or nuclear sequence data, bioinformatics offers an avenue for developing a widely used conservation genetic tool.

When developing a species-specific eDNA assay, it is essential to have DNA sequence information unique to your target organism. The most efficient approach is identifying if relevant sequence data is available in online repositories (Figure 3.3; step 1) such as the National Centre for Biotechnology Information (NCBI), the Barcode of Life Data System (BOLD), and the European Molecular Biology Laboratory online repositories. The Sanger sequenced *M. georgesi* mitogenome used as our positive control for the bioinformatic mitogenome extraction was obtained from the NCBI database using '*Myuchelys georgesi* mitochondrion' as search terms. When relevant mitochondrial sequence data are not available in online repositories or if gene regions are missing for species or taxa (Freeland, 2017, Nordstrom et al., 2022), genomic data (reference genome or raw HTS) provides an *in silico* alternative (Figure 3.3; step 2). For example, the availability of a reference genome has allowed for bioinformatic extraction of the mitogenomes for several cryptic and threatened species lacking mitochondrial sequence data including the Kroombit tinker frog (*Taudactylus Pleione*) (Farquharson et al., 2023), blue-tailed skink (*Cryptoblepharus egeriae*) (Dodge et al., 2023), Lister's gecko (*Lepidodactylus listeri*) (Dodge et al., 2023), and southern stuttering barred frog (*Mixophyes balbus*) (Tang et al., 2024), providing capacity for the development of species-specific eDNA assays in the future. Although our approach leverages PacBio HiFi sequencing data and MitoHiFi mitochondrial genome assembly program for bioinformatic extraction (Figure 3.3; step 3a), a suite of bioinformatic tools are available for extraction and assembly of mitogenomes from a range of HTS data types (Table 3.2). Additionally, some of these tools can take raw HTS sequencing data as input and do not require a reference genome.

Table 3.2. Bioinformatic tools for complete mitochondrial genome assembly using next generation sequencing (NGS) data (including whole genomes) for downstream eDNA assay development. (WGS) Whole genome sequencing. CCS Close consensus sequencing, CLR continuous long reads.

Name	Data input described	Reference
MitoHiFi	PacBio high fidelity (HiFi)(CCS) WGS data	(Uliano-Silva et al., 2021)
PMAT	PacBio high fidelity (HiFi)(CCS) WGS data	(Bi et al., 2024)
SMART	Low-coverage WGS	(Alqahtani and Măndoiu, 2020)
MitoZ	Short WGS raw reads	(Meng et al., 2019)
Norgal	Short WGS raw reads	(Al-Nakeeb et al., 2017)
MITObim	Short WGS raw reads	(Hahn et al., 2013)

If mitochondrial or genomic sequence data does not exist (Figure 3.3; step 1 and 2), conventional approaches involving acquisition of genetic material; DNA extraction; sequencing; and assembly are needed to undertake species' assay design (Figure 3.3; step 3b). Although targeted mitochondrial sequencing may be effective when programs have limited funds available (Schmidt et al., 2024), these approaches often require substantial time and resources to undertake so likely cost the same as whole genome sequencing when labour costs are accounted for. For example, completion of the existing *M. georgesi* mitogenome following the methods of Zhang et al. (2017) used Sanger sequencing and long-range PCR, took 12 weeks to complete, costing $15,000 AUD in labour and $1500 in lab consumables (Arthur Georges pers comm., 2024). By-passing these steps, when genomic data is available, can save conservation programs time and money that can be invested elsewhere. For example, bioinformatic extraction of the mitogenome from a 1.9GB genome required 30 minutes, 1 CPU, and 5.3GB of memory, offering a high cost-effectiveness in terms of labour, data acquisition and analysis. As the costs associated with genome assembly decrease, a 3GB long-read genome can cost ~$5,005 in sequencing, ~$600 in labour and ~$200 in consumables (Elspeth McLennan pers comm., 2024). Additionally, completion of a reference genome can only require 2 days of laboratory work, 6 weeks of sequencing, and 2 days for bioinformatic

assembly. Although costs are not directly comparable, investment in genomic data provides a resource for a plethora of downstream applications beyond mitochondrial and eDNA (Formenti et al., 2022, De León et al., 2023, Schneider, 2023, Brandies et al., 2019).

The key aspect of an eDNA assay is primer design (Figure 3.3; step 4). As mentioned in step 1, primers are often developed using available reference sequences in online databases however, regions may be missing for species or taxa (Freeland, 2017, Nordstrom et al., 2022). For example, 12S, 16S, 18S sequence data is less often available compared to COI and CytB sequence information (Lacoursière-Roussel et al., 2016). An advantage of a reference genome-derived approach is that it provides researchers and managers with a complete or close to complete mitochondrial sequence. This enables the design of molecular markers for any gene in the mitogenome and provides the option to expand into nuclear marker design (McCauley et al., 2024). *In silico*, *in vitro*, and *in situ* validation methodologies (Figure 3.3; step 4A-C) should follow a standardised approach (Nordstrom et al., 2022). Since the rapid uptake of eDNA analysis, comprehensive eDNA guidelines for assay development and validation have been developed to assist researchers and managers in developing eDNA across a range of taxa and ecosystems that can be adapted to the habitat and biology of the target species (Goldberg and Strickler, 2017, De Brauwer et al., 2022a, De Brauwer et al., 2022b).

In summary our methodologies and workflow for *M. georgesi* consist of four stages; (i) identifying availability of a mitogenome (Figure 3.3; step 1); (ii) identifying availability of a reference genome or genomic data when mitochondrial sequence does not exist in online repositories (Figure 3.3; step 2); (iii) bioinformatic assembly of a mitogenome from a reference genome (Figure 3.3; step 3a); and (iv) primer design and *in silico*, *in vitro*, and *in situ* validation (Figure 3.3; step 4).

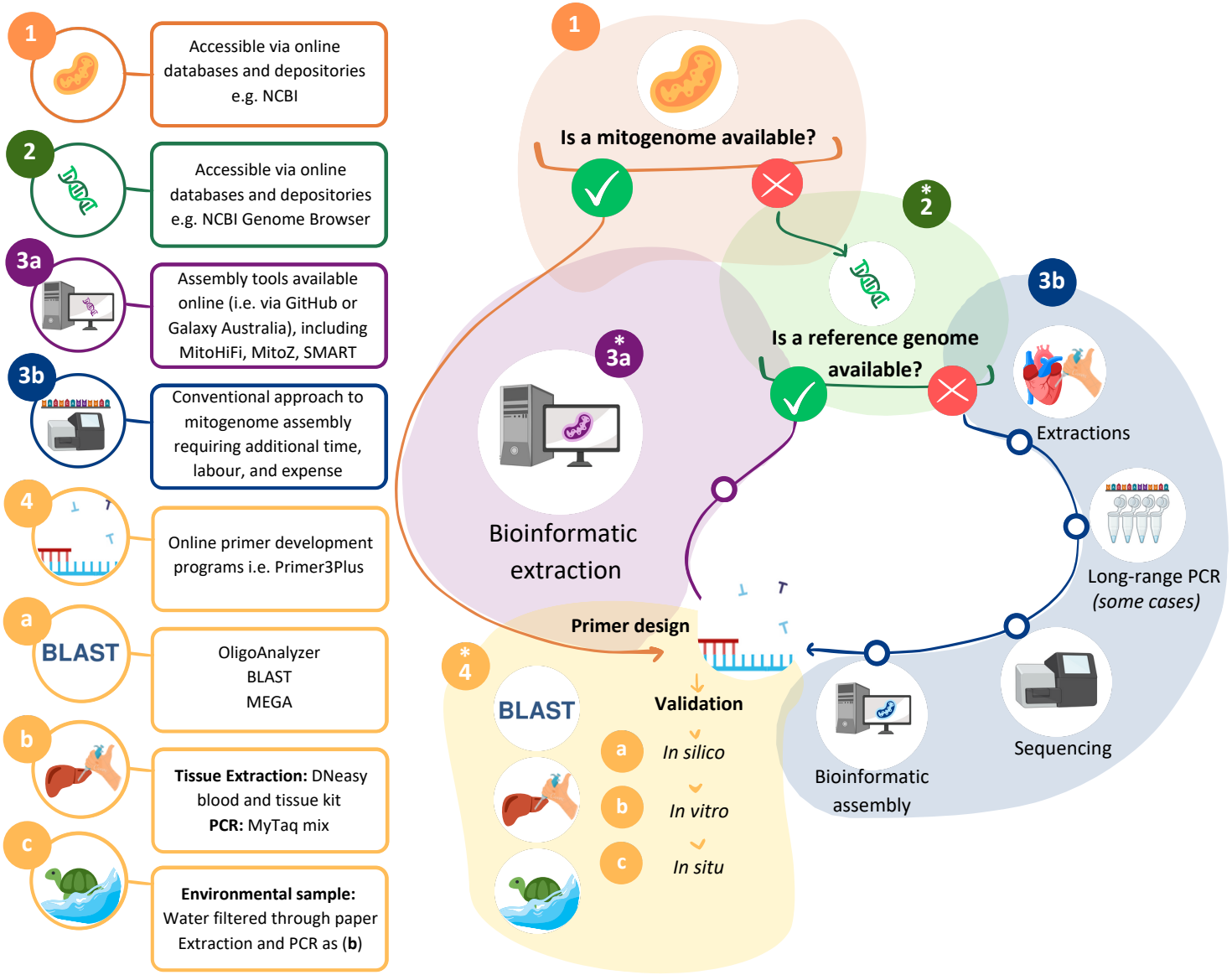


Figure 3.3. Reference genome derived eDNA assay workflow used for species-specific primer development in *Myuchelys georgesi*. NCBI National Center for Biotechnology Information, BLAST Basic Local Alignment Search Tool, MEGA Molecular Evolutionary Genetic Analysis, PCR Polymerase Chain Reaction. Steps undertaken in this study are denoted by *. Image created using Biorender.com and Canva.com.

Our results provide *M. georgesi* managers with an eDNA assay that can be implemented into species monitoring. The assay can assist managers in resolving questions around distribution within the Bellinger River catchment, including reaches in the upper catchment and the Kalang River, and inform survey site selection through identification of occupancy hotspots. Future work is needed to evaluate efficacy of primers on Bellinger River water samples as environmental barriers such as water flow, sediment composition, and microbial and enzyme activity (Barnes et al., 2014, Stoeckle et al., 2017, Stewart, 2019) may influence detection. The technique will be useful for initially be used to identify areas to perform more intensive diving and trapping surveys, providing the species with a multifaceted detection and survey approach (Villacorta-Rath et al., 2022, Lam et al., 2022, Nordstrom et al., 2022, Carvalho et al., 2022).

As conservation genetics moves into the genomics-era, genomic data is becoming increasingly available for non-model organisms, making it important to leverage and apply the information genomic resources provide. We use a reference genome-based approach to develop an eDNA assay for *M. georgesi*. The development of species-specific eDNA primers provides a valuable tool for managers in assessing population dynamics of this rare species, supporting informed management decisions and guiding future conservation efforts.

Chapter 4 - Insights into whole-genome and MHC diversity of a critically endangered turtle

4.1 Manuscript

Published in *Immunogenetics*

Genome-wide diversity and MHC characterisation in a critically endangered freshwater turtle susceptible to disease

Holly V. Nelson[1], Luke Silver[1,2], Toby G. L. Kovacs, [1] Elspeth A. McLennan[1], Arthur Georges[3], Jane L. DeGabriel[4], Carolyn J. Hogg[1,2]* and Katherine Belov[1,2]

[1] School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia.

[2] Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Sydney, Sydney, NSW 2006, Australia

[3] Institute for Applied Ecology, University of Canberra, Bruce, ACT 2617, Australia

Science, The University of Sydney, Sydney, NSW 2006, Australia

[4] NSW Department of Climate Change, the Environment, Energy and Water, Parramatta, NSW 2150, Australia

4.2 Abstract

Small, isolated populations are often vulnerable to increased inbreeding and genetic drift, both of which elevate the risk of extinction. The critically endangered Bellinger River turtle (*Myuchelys georgesi*) is restricted to a single river catchment in New South Wales, Australia. The only extant wild population and the captive breeding program face significant threats from viral outbreaks, most notably a nidovirus outbreak in 2015 that caused a 90% population decline. To better understand the factors contributing to this species' disease susceptibility, we analysed 35 re-sequenced genomes to assess genome-wide diversity in pure *M. georgesi* (*N* = 31) and disease-resilient offspring of F1 hybrids (*M. georgesi* × *Emydura macquarii*) backcrossed to pure *M. georgesi* (*N* = 4). We manually annotated the major histocompatibility complex (MHC), identifying five MHC class I and ten MHC class II genes and investigated genetic diversity across both classes. Our results show that both genome-wide diversity and immunogenetic diversity at MHC loci are lower in pure *M. georgesi* compared to the disease-resilient hybrids. However, the variation observed within the core MHC region, extending

across chromosome 10, exceeds that of all other macrochromosomes. Additionally, no significant short-term changes in either genome-wide or immunogenetic diversity were detected following the 2015 nidovirus outbreak. Demographic history reconstructions indicate a sustained, long-term decline in effective population size since the last interglacial period, accompanied by more recent steep declines. These patterns suggest that prolonged isolation and reduced population size have significantly influenced the dynamics of genome-wide diversity. It is likely that contemporary stressors, including the recent nidovirus outbreak, are acting on an already genetically depleted population. This study provides a more thorough understanding of both genome-wide and immune gene diversity to inform future conservation management actions for the species.

4.3 Introduction

The emergence of novel infectious diseases is a major threat faced by species of conservation concern (Smith et al., 2009, Daszak et al., 2000). Warming climates, pollution, and introduced species that harbour invasive pathogens are facilitating the spread of disease across wildlife populations (Anderson et al., 2004). Managing declining populations in the presence of infectious disease, in addition to other anthropogenic threats, is of growing conservation concern. Understanding the mechanisms of resilience and resistance, and the degree to which populations can adapt to change is a key step towards preventing extinctions (Auteri and Knowles, 2020). Another important aspect of conservation management is maintaining the adaptive potential of a species. Adaptive potential is the ability of a population or species to adapt to changing environmental conditions, such as disease, habitat modifications, and climate change (Holderegger et al., 2006, Hoffmann et al., 2017). Species recovery programs are increasingly using genomic data to understand the adaptive potential of populations through functional gene analyses (McLennan et al., 2024), with growing emphasis on how managers can maintain or potentially increase adaptive potential in wild populations and conservation breeding programs (Farquharson et al., 2022). Advances in sequencing technologies and bioinformatic tools, and reductions in sequencing costs, have made whole-genome data more accessible for threatened species recovery programs. The combination of high-quality reference genomes with whole-genome sequencing data facilitates high-

resolution analyses of neutral genetic variation, while also enabling the characterization of functional gene families and the reconstruction of historical demographic trends (Theissinger et al., 2023). For example, Magid et al. (2022) used 66 re-sequenced genomes to investigate TLR immune gene diversity in shore plovers (*Thinornis novaeseelandiae*) and found low levels of diversity.

The major histocompatibility complex (MHC) is a crucial gene family within the adaptive immune system, primarily responsible for antigen presentation (Piertney and Oliver, 2006). MHC molecules bind and present pathogen-derived peptides to T cells, facilitating the recognition and initiation of immune responses to pathogens and ensuring the specificity of adaptive immunity (Neefjes et al., 2011). Variation within MHC genes has frequently been linked with species' susceptibility to disease. Higher levels of heterozygosity at MHC loci is thought to enhance a hosts ability to respond to a broader range of pathogen-derived antigens (Hughes and Yeager, 1998, Penn et al., 2002). For example, Schmid et al. (2023) found that allelic diversity correlated with coronavirus susceptibility in *Hipposideros* bat species. The association between MHC genes and disease resilience has been observed across multiple wildlife species including the northern leopard frog (*Rana pipiens*) (Trujillo et al., 2021), guignas (*Leopardus guigna*) (Napolitano et al., 2023) and desert bighorn sheep (*Ovis canadensis nelsoni*) (Dugovich et al., 2023), with disease-resilient populations exhibiting greater heterozygosity at MHC loci. The importance of the MHC in the immune cascade makes them excellent candidates to start to understand a populations' adaptive genetic diversity. For many threatened species, the lack of a high-quality reference genome has limited our ability to investigate the repertoire and variation present within the MHC region. Studies on the MHC region in non-avian reptiles have been carried out in the komodo dragon (*Varanus komodoensis*) (Reed and Settlage, 2021) and the tuatara (*Sphenodon punctatus*) (Miller et al., 2015). Recent genomic analyses in two anolis lizard species (*Anolis carolinensis* and *Anolis sagrei*) (Card et al., 2022) and the Chinese alligator (*Alligator sinensis*) (He et al., 2022b), have provided a more comprehensive characterisation of reptile MHC.

The critically endangered Bellinger River turtle (*Myuchelys georgesi*) is restricted to a 60 km stretch of Bellinger River, on the mid-north-coast of NSW, Australia and has undergone declines of over 90% because of disease outbreaks in recent years (Zhang et al., 2018, Chessman et al., 2020). The species consists of a single wild population making it highly susceptible to stochastic events. Survival of the species may depend upon its adaptive potential and genetically informed management actions. Our previous work has shown the species exhibits low genome-wide variation, as measured by reduced representation sequencing data (Nelson et al., 2024). This reduced variation at 460 neutral markers is likely a consequence of the species restricted geographic range, limited dispersal capacity, and long generation time (Spencer et al., 2014, Spencer et al., 2018, Chessman et al., 2020, Nelson et al., 2024). Additional analyses hypothesize that the species has undergone a number of historical bottlenecks, which have likely contributed to its low genetic diversity and ongoing disease susceptibility (Georges, 2020).

In 2015, a species-specific nidovirus outbreak (the Bellinger River virus; BRV) caused the population to crash from *ca.* 4000 to *ca.* 200 individuals and resulted in a significant decline in genetic diversity, as measured by reduced representation sequence data (Zhang et al., 2018, Chessman et al., 2020, Nelson et al., 2024). The near extinction of the species led to urgent conservation efforts including the development of a *M. georgesi* conservation action plan (Jakob-Hoff et al., 2017), a range of community engagement initiatives, habitat restoration projects, and the establishment of a conservation breeding program comprised of two populations founded from 16 and 19 wild individuals, respectively. Since 2015, two lesser-documented viral outbreaks have occurred in the river, one in January 2022 (Parrish et al., 2024) and a second in May 2024 (NSW DCCEEW, *pers. comm*., 2024). The ongoing disease outbreaks put the species at high risk of extinction and require investigation into potential genetic mechanisms that have led to this susceptibility.

In contrast, the Murray River turtle (*Emydura macquarii*) appears to be resilient to the viruses that infected *M. georgesi*. *E. macquarii* is widespread along the east coast of Australia and it is hypothesized that it was introduced to the Bellinger River by humans over the past two

decades (Spencer et al., 2018). *E. macquarii* is known to hybridise with *M. georgesi* (Georges et al., 2018). Pure *E. macquarii*, F1 (pure *M. georgesi* x pure *E. macquarii*) and backcross F2 hybrids (pure *M. georgesi* x F1 hybrid) have tested positive to the recent viruses, but do not exhibit symptoms or succumb to the disease (Zhang et al., 2018). Given the high susceptibility of *M. georgesi* to nidovirus infections, conservation managers are looking for long-term viable options without the need for continual intensive management. A common strategy to enhance genetic diversity and potentially mitigate disease susceptibility is the introduction of individuals from genetically diverse or distinct populations (Frankham, 2015, Frankham et al., 2017). Since this is not feasible for *M. georgesi*, alternative strategies may be required to enhance genetic diversity for long-term population viability such as interspecific hybridisation (Baack and Rieseberg, 2007), or the reintroduction of historical genetic variation lost over time through methods such as back-breeding, cloning, or genome editing (Shapiro, 2017).

In this study, we undertook the first MHC gene annotation in a freshwater turtle species by characterising the genetic features of MHCI and MHCII genes in *M. georgesi*. Using whole genome re-sequencing of 12 contemporary turtles sampled after the disease outbreak, 19 historic turtles sampled before the disease, and four disease-resilient second-generation hybrids we investigated patterns of genome-wide and immunogenetic diversity between (i) pre and post disease animals and (ii) pure *M. georgesi* and F2 hybrids.

4.4 Methods

Immune gene annotation

We used a homology-based approach via BLAST v2.3.30 (Camacho et al., 2009) to manually characterise all MHC class I and II genes in the *M. georgesi* genome. To annotate class I genes we acquired query sequences from the National Centre for Biotechnology Information (NCBI) including the tawny dragon (*Ctenophorus decresii*) (KY905241.1), caiman (*Caiman crocodilus*) (KF769542.1), marine iguana (*Amblyrhynchus cristatus*) (EU839663.1), galapagos land iguana (*Conolophus subcristatus*) (EU604313.1), tuatara (*Sphenodon punctatus*) (DQ145788.1), and green sea turtle (*Chelonia mydas*) (OK135213.1) (Tabls S1). To annotate class II genes we

acquired query sequences from NCBI including the Chinese softshell turtle (*Pelodiscus sinensis*) (MT834970.1), marine iguana (*Amblyrhynchus cristatus*) (FJ623752.1), and green-rumped parrotlet (*Forpus passerines*) (EF710746.1) (Table A3.1.1). Query sequences were used to search the *M. georgesi* genome (rMyuGeo1.pri.20230808) and a global transcriptome (Nelson et al., 2024) using BLASTn and tBLASTn, respectively with an e-value threshold of 1e-10). Exon splicing sites were manually checked by visualising against the reference genome, global transcriptome and automated annotation in IGV v.2.16.0 (Robinson et al., 2011). Nucleotide sequences for each gene were then extracted from the reference genome using bedtools v2.29.2 (Quinlan and Hall, 2010) and input into MEGA v11 (Tamura et al., 2021). Nucleotide sequences were converted to protein sequences to ensure no stop codons were present within the coding sequence. Nucleotide similarity between genes was calculated using EMBL-EBI Clustal Omega (Madeira et al., 2019). To identify both classical and non-classical MHC genes, we analysed gene expression by aligning transcripts from brain, liver, and spleen tissues to MHC genes using IGV v.2.16.0. Phylogenetic relationships were estimated for each class using sequences acquired from a maximum likelihood (ML) analysis performed in IQ- TREE2 (Minh et al., 2020). When possible, complete coding sequences for reptiles and amphibians were acquired from NCBI (Table A3.1.2, Table A3.1.3). The ModelFinder option (-m MFP) within IQTREE2 was used to select the best-fitting substitution model according to the Bayesian information criterion (Kalyaanamoorthy et al., 2017). Node support was assessed using the ultrafast bootstrap (-bb 1000) approximation and the like approximate likelihood-ratio test (-alrt 1000) (Guindon et al., 2010, Hoang et al., 2018).

Re-sequenced genome sampling, extraction, and sequencing

Re-sequenced genomes were generated using samples collected by the NSW Department of Climate Change, Energy, the Environment, and Water (DCCEEW) during surveys before the disease outbreak in April 2007, after the outbreak in November 2019, and opportunistically from second-generation hybrids in November 2019 (Table A3.1.4). DNA samples were collected by removing part of the trailing webbing of the clawless toe on the hindfoot or by extracting blood from the jugular vein (Georges et al., 2018) (Table A3.1.4). Samples were stored in 75% ethanol at −20°C in the University of Canberra Wildlife Tissue Collection

(GenBank UC<Aus>). For DNA extractions we performed a high salt method on 19 skin tissue biopsies and 16 dried bloods on Whatman card samples following a modified protocol from Aljanabi and Martinez (1997). DNA concentration and quality were assessed using a Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific) and 0.95% agarose gel electrophoresis for 30 minutes at 90V. To maximise DNA quality for WGR, we undertook an additional DNA repair step using a FFPE DNA repair protocol (New England Biosciences). Repaired DNA concentration and quality were assessed using a Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific) and 0.95% agarose gel electrophoresis for 30 minutes at 90V. Samples were sent to Ramaciotti Centre for Genomics (University of New South Wales, Australia) for WGS on an Illumina NovaSeq 6000, using a TruSeq DNA PCR free library prep kit across six lanes.

Re-sequenced genome alignment and variant calling

Raw fastq reads for 35 individuals were quality checked and trimmed using fastqc v0.11.8 (Andrews, 2010) and trimmomatic v0.39 (Bolger and Giorgi, 2014) with the parameters ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 SLIDINGWINDOW:4:5 LEADING:5 TRAILING:5 MINLEN:25. Reads were aligned to the reference genome (Nelson et al., 2024; rMyuGeo1.pri.20230808) using Burrows-Wheeler aligner (BWA) v0.7.17 (Li and Durbin, 2009) 'mem' function with default parameters. The resulting alignment files were sorted into bam format using samtools sort v1.6 and alignment rates calculated using samtools flagstat v1.6 (Li et al., 2009). As individuals were sequenced across multiple lanes, bam files pertaining to a single individual were merged using samtools merge and duplicates marked and removed using picard v2.21.9 MarkDuplicates (http://broadinstitute.github.io/picard/). For downstream analyses, we partitioned our data into three putative groups (two temporal, reflective of groups investigated by Nelson et al, 2024); wild individuals sampled before the disease outbreak in 2007, *N* = 19 (hereafter "Before"); wild individuals sampled after the disease outbreak between 2015-2020, *N* = 12 (hereafter "After"); and offspring of F1 hybrids (*M. georgesi* × *Emydura macquarii*) backcrossed to pure *M. georgesi N* = 4 (hereafter "Backcross").

Genome-wide diversity

Coverage, genome-wide diversity, individual-level observed and expected autosomal heterozygosity and runs of homozygosity (ROH) were calculated across each genome using ROHan (Renaud et al., 2019). ROHan combines a local Bayesian model and hidden Markov model (HMM) to identify autosomal heterozygosity and ROHs from bam files of individually mapped genomes (Renaud et al., 2019). Analyses were run on the nine macrochromosomes, excluding the sex chromosome (scaffold 4), with the parameters --rohmu 5e-5, --TsTv 1.965, -t 4 and –size 100,000 (100kb windows) on an Amazon Web Services ubuntu 20.04 LTS cloud machine (r5.8x large, 32 vCPU, 256 Gb RAM, 1TB attached storage). Inbreeding coefficient based on ROH (F_{ROH}) (McQuillan et al., 2008) and average individual-level heterozygosity was calculated for each group using the hmmrohl and summary texts output by ROHan, respectively. Results were visualised in R v4.3.0 using dplyr() v1.1.2. and ggplot2() v3.5.0. We used a Bartlett's test of homogeneity of variances base function to test whether observed heterozygosity was significantly different from expected heterozygosity with a significance threshold of α=0.05.

Demographic reconstructions

Changes in ancient effective population size (N_e) over time were assessed with the pairwise sequentially Markovian coalescent (PSMC) model (Li and Durbin, 2011) using bam files for 31 pure *M. georgesi*. Files were filtered on sites with a coverage below one third or two times the sample's average coverage were removed. For N_e estimates we used scaffolds > 50kb and excluded the sex chromosome. Consensus genome sequences in fastq format were generated using the 'mpileup' command in BCFtools 1.3.1 and the included script 'vcfutils.pl'. PSMC v0.6.5-r67 was run with the following default parameters, -N25 -t15 -r5 -p "4+25*2+4+6". Results were scaled by a mutation rate of 4.61 $x10^{-9}$ substitutions per site per generation based on estimates for *Chrysemys picta* (Bergeron et al., 2023). As precise details of the life history of *M. georgesi* are not available, we used a generation time of 14 years based on manager recommendations. To investigate recent historical patterns of demography we calculated N_e using GoNe (Santiago et al., 2020). We used `bcftools merge` to combine all single-sample gVCFs into a population-level gVCF, followed by `bcftools call` to retain only

variable sites. For stringent filtering, we kept variants with a minor allele frequency > 0.04, an average depth between 5 and 200, a genotype quality of at least 10, a minimum quality score of 10, and no more than 10% missing data. We conducted multiple iterations of GoNe to determine the optimal values for *hc* (0.01, 0.05, 0.1) and various recombination rate estimates (1, 2, 3). The optimal values were identified as a recombination rate of 3 cM/Mb and an *hc* of 0.01. N_e outputs were visualised using ggplot2() v3.5.0 in R v4.3.0.

MHC gene analyses

To investigate temporal MHC diversity, we generated a joint-genotyped file using gatk GenomicsDBImport to generate a sample database and gatk GenotypeGVCF to generate a whole genome joint genotyped multi sample vcf file. We filtered the joint genotyped multi sample vcf to include only biallelic SNPs found within MHC exons (see results) using vcftools v0.1.14 (Danecek et al., 2011) and gatk VariantFiltration and SelectVariants v4.2.0.0 (McKenna et al., 2010). We then used gatk VariantFiltration (McKenna et al., 2010) to remove variants with QUAL < 80, MQ <40, -12.5 < MQRankSum <12.5 and -8 < ReadPosRankSum < 8 and bcftools v1.3.1 (Danecek et al., 2021, Li, 2011) to remove sites with an average depth across all samples of < 10 and sites with an allelic balance > 0.9. Next, we used vcftools to retain SNPs present at a MAF <0.01. Finally, we We then calculated the number of SNPs in each gene and determined if they resulted in synonymous or non-synonymous amino acid substitutions using Geneious Prime 2020 (https://www.geneious.com/) To identify full length allele sequences for each MHC gene in each individual, we undertook phasing. First, we used gatk FastaAlternateReferenceMaker (v4.2.0.0) (McKenna et al., 2010) to generate single sample consensus fasta sequences for each gene. SeqPHASE (Flot, 2010) was used to convert the fasta sequences into phase format then PHASE (v.2.1.1) (Stephens et al., 2001, Stephens and Scheet, 2005) was run to generate alleles and SeqPHASE (Flot, 2010) used to convert phase output into phases fasta sequences to give 70 sequences for each gene (two per individual). To assess interspecific genetic differentiation within the 15 annotated genes, we used variants across the complete MHC I and MHC II exon sequences in temporal (Before and After) and all groups (Before, After, and Backcross). We calculated observed (H_O) and expected heterozygosity (H_E) for each gene across all groups using GenAlEx v6.5 (Peakall and Smouse,

2006) and visualised average individual H_o by group using the boxplot() functions in R. We used the Bartlett's test of homogeneity of variances base function to test whether individual observed heterozygosity was significantly different from expected heterozygosity in R. To assess genetic differentiation among individuals, we generated principal component analysis (PCoA) plots using our genome-wide (pure and backcross) and MHC (pure only) joint genotyped VCFs, using the adegenet package (Jombart, 2008) in R (v4.1.1) (R Core Team, 2023).

4.5 Results

MHC annotation

Through manual annotation of immune-related genes, we identified five MHC class I loci (*Myge-UA*, *Myge-UB*, *Myge-UC*, *Myge-UD*, *Myge-UE;* labelled from most to least number of detected variants), and ten MHC class II loci (*Myge-DAA1*, *Myge-DAB1*, *Myge-DAA2*, *Myge-DAB2*, *Myge-DAA3*, *Myge-DAB3*, *Myge-DAA4*, *Myge-DAB4*, *Myge-DAA5*, *Myge-DAB5*), all located within a 272,213 bp region on chromosome 10. The gene nomenclature adheres to the conventions established by Miller et al. (2005) and Miller et al. (2015) for a reptilian species, however, limited annotations and published sequences in other reptile species made comparative interpretation challenging. The *M. georgesi* MHC class II genes are densely clustered, consisting of alpha (here labelled *DAA*) and beta (here labelled *DAB*) chains encoded on the 3' to 5' and 5' to 3' strands, respectively (Table 4.1). Structurally, the class II genes exhibit high conservation, except for *Myge-DAB5* which shows a putative loss of exon 5 with no transcriptional evidence across three tissue types (Figure 4.1). Similarly, MHC class I genes display a largely conserved architecture, except for *Myge-UA*, which has notably larger introns 1 and 2 compared to other MHC class I genes and contains an additional exon (Figure 4.1, Table A3.1.5). Differentiation analyses revealed high sequence similarity between *Myge-UB* and *Myge-UE*, whereas *Myge-UD*, *Myge-UA*, and *Myge-UC* exhibited closer sequence similarity (Table Table A3.1.6A). As expected, lower levels of differentiation were observed within the MHC II alpha (*DAA*) and beta (*DAB*) gene groups compared to between these gene groups (Table A3.1.6B). Transcriptome data did not provide a clear distinction between classical and non-classical Class I genes, as transcript levels across the three tissue types

(brain, liver, and spleen) were consistent for all genes, suggesting that additional tissue types may be needed to identify if any of the identified Class I genes may have non classical functions.

Table 4.1. Manually annotated MHC I and MHC II genes on chromosome 10 used for downstream analyses. Start and end coordinates refer to whole genome coordinates. Coordinates for each exon are presented in table A3.1.5.

Class	Gene	Chromosome	Exons	Start	End	Strand
MHC I	*Myge_UB*	10	8	30471574	30493484	-
	Myge_UE	10	7	30500310	30508525	-
	Myge_UC	10	7	30533405	30538940	-
	Myge_UD	10	7	30567953	30577665	-
	Myge_UA	10	7	30583497	30591368	-
MHC II	*Myge_DAA1*	10	4	30660982	30663487	-
	Myge_DAB1	10	6	30665109	30669452	+
	Myge_DAA2	10	4	30677698	30680223	-
	Myge_DAB2	10	6	30681654	30686193	+
	Myge_DAA3	10	4	30698489	30700988	-
	Myge_DAB3	10	6	30702397	30706348	+
	Myge_DAA4	10	4	30717695	30720191	-
	Myge_DAB4	10	6	30721446	30725542	+
	Myge_DAA5	10	4	30735202	30737782	-
	Myge_DAB5	10	5	30739701	30743969	+

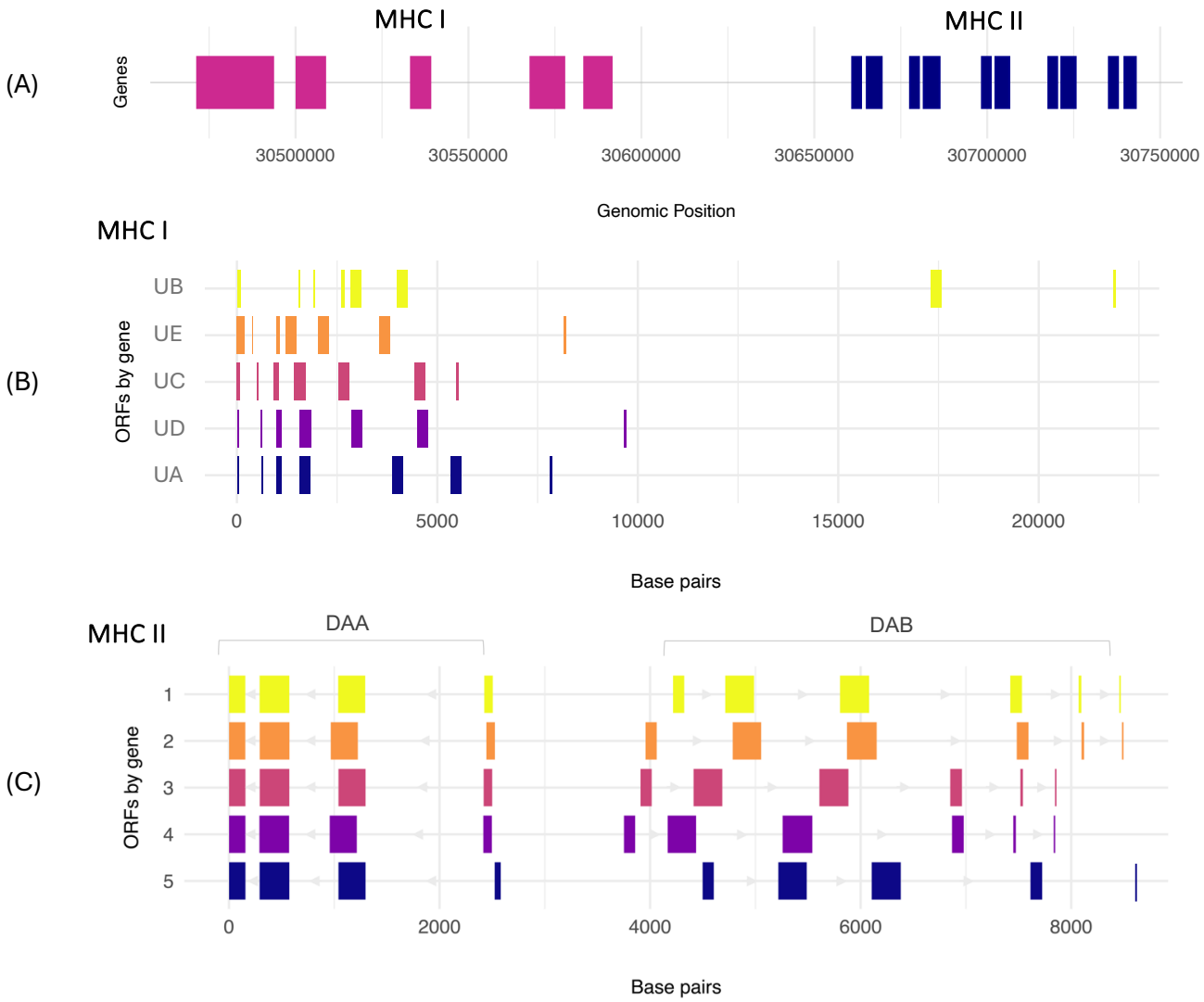


Figure 4.1. Genomic architecture of Major Histocompatibility Complex (MHC) genes. **(A)** Genomic architecture of the core Major Histocompatibility Complex (MHC) region, including class I and class II genes, on chromosome 10. **(B)** and **(C)** Genomic coordinates and structural organisation of manually annotated *Myuchelys georgesi* (*Myge*) MHC genes presented as open reading frames (ORFs), in order of genomic position along chromosome 10.

To contextualise the genetic relationships among MHC homologs in *M. georgesi* and extend these comparisons across non-avian and avian reptiles, we constructed phylogenetic trees using complete coding sequences of MHC I and MHCIIB sequences, following the approach of He et al. (2022b). The resulting phylogenies revealed that *M. georgesi* and members of the Crocodilia form a monophyletic group, distinct from other reptilian orders, including

Squamata, Amphibia, Aves, and Rhynchocephalia in the MHC class I tree, and from Squamata, Aves and Rhynchocephalia in the MHC class II tree (Figure A3.2.2, Figure A3.2.3).

Genome-wide diversity and demographic history

All whole-genome samples had an average alignment rate of 99.6% and 99.5% for pure and backcross animals, respectively, across ten macrochromosomes. ROH were distributed across all macrochromosomes (Figure 4.2A). The longest ROH was identified on chromosome 5 in purebred individuals and covered 36% of the chromosome's total length (Figure 4.2A). In contrast, chromosome 10 exhibited the lowest density of ROH relative to its size (Figure 4.2A), with the MHC region located outside of ROH for almost all individuals (Figure 4.2B). The pure samples had 90% of their genome in ROH, characterised by short (<2 Mb) and long (>2 Mb) ROH as defined by Ceballos et al. (2018), whereas only an extremely small number of short ROH were detected in backcross individuals (Figure 4.2A, Table A3.1.7, Figure A3.2.4, Figure A3.2.5). Small ROH measuring 0-500 kb, 500-100 b, and 1000-2000 kb had similar frequencies across groups (Figure A3.2.5A). Although the counts of small and long ROH were fairly similar, the proportion of the genome in long runs was considerably higher than in small runs. The inbreeding coefficient based on ROH (FROH) was significantly greater in ROH >2000kb (78%) for both temporal groups, where FROH is expressed as the percentage of the total length of the 9 macrochromosomes (Figure A3.2.5B). This is consistent with the expectation that larger ROH occupy a greater proportion of the genome. The Before and After samples had an identical proportion of the genome in ROH and comparable ROH lengths. The average number of ROH per individual was also consistent between the two temporal groups. No significant differences in nucleotide diversity across the genome were observed between the Before and After individuals, regardless of whether ROH were included or excluded. Both temporal groups displayed significantly lower nucleotide diversity compared to the backcross individuals (Table A3.1.7).

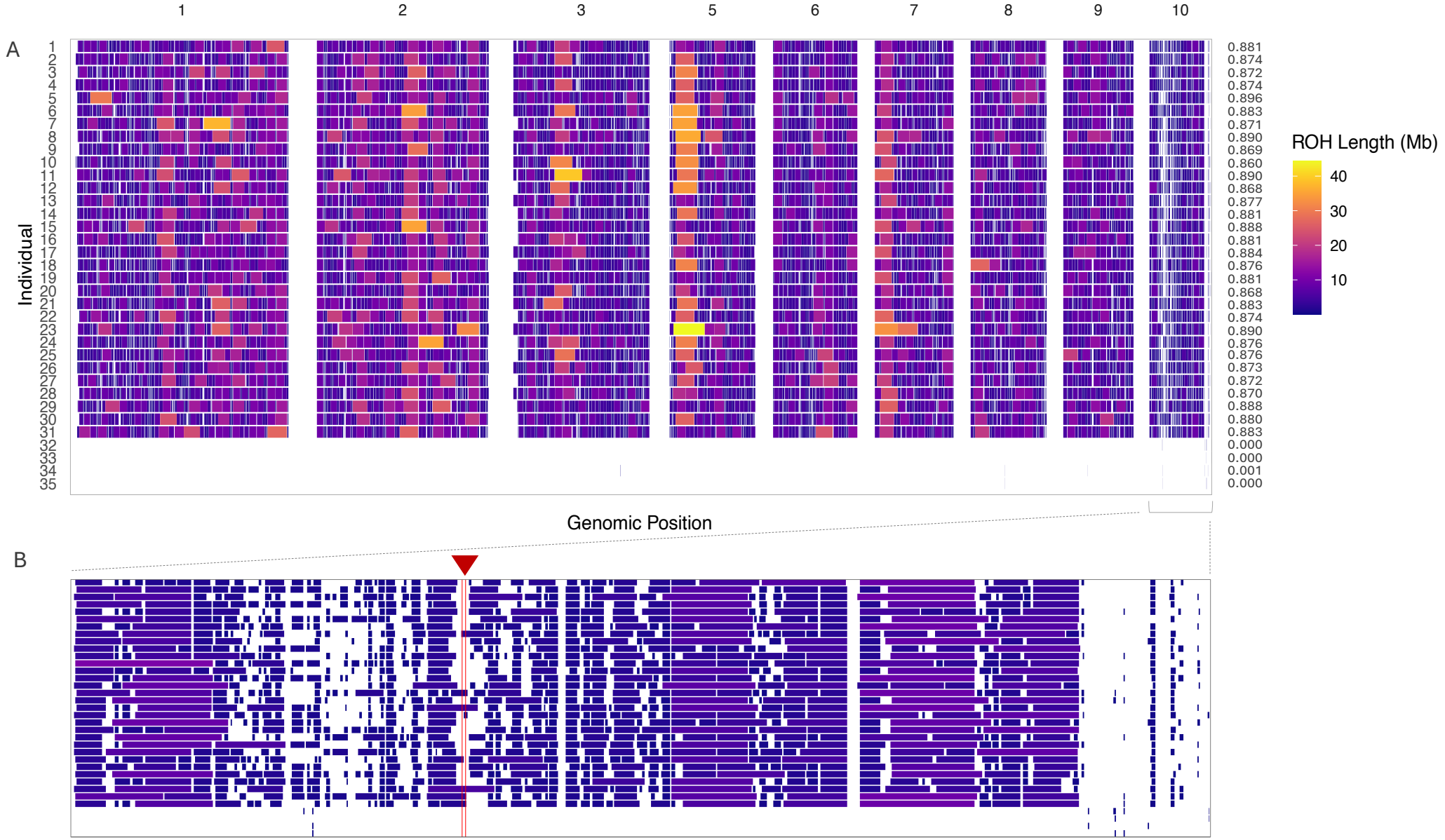


Figure 4.2. Runs of homozygosity (ROH) heatmap illustrating **(A)** the chromosomal distribution, length, and the proportion of the genome in ROH across nine *Myuchelys georgesi* macrochromosomes in Before (individuals 1–19), After (individuals 20–31), and backcross hybrids (individuals 32–35). **(B)** Enlarged view of chromosome 10. Major histocompatibility complex (MHC) core region marked by a red arrow and parallel lines.

We identified significantly lower levels of genome-wide heterozygosity in pure versus backcross animals (Pure: $1.18x10^{-4}$; Backcross: $6.8x10^{-3}$, t=-16.01, P=0.001) (Figure 4.3A). No significant difference was observed between temporal groups (Before: $1.17x10^{-4}$; After: $1.21x10^{-4}$, t=-1.40, P=0.1745) (Figure 4.3B).

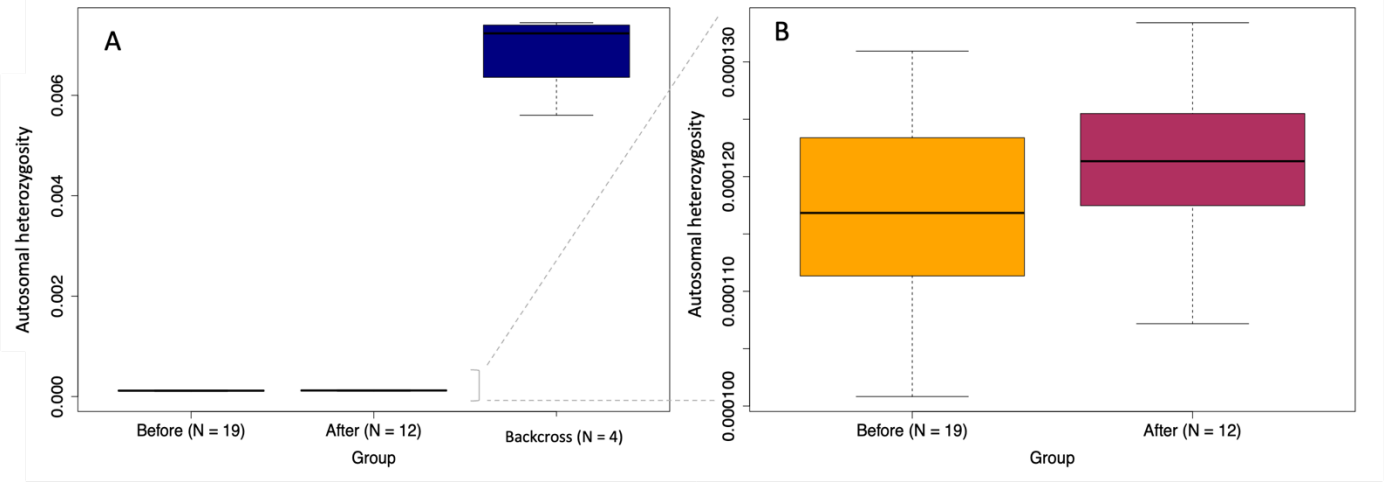


Figure 4.3. Genome-wide heterozygosity estimates for *Myuchelys georgesi* calculated using Rohan. Boxes represent the interquartile range (IQR) with whiskers extending to the upper and lower ranges and the bold line representing the group mean. **(A)** Autosomal heterozygosity estimates for all groups including Before, After, and Backcross. **(B)** Rescaled autosomal heterozygosity estimates of pure *M. georgesi* Before and After groups plotted in A.

Our demographic history reconstructions using PSMC analysis indicate a gradual, long-term decline in effective population size beginning approximately 110,000 years ago, coinciding with the last glacial period (Figure 4.4A). The largest effective population size estimates correspond to a period before the last interglacial, around 110,000 years ago. These estimates were consistent across all 31 individuals analysed. Estimates predating this timeframe are likely less reliable due to including the accumulation of errors in coalescent events in deep time and limited resolution in ancient periods; therefore, they should be approached with caution. GoNe analyses showed there has been a decline in effective population size from an estimate N_e of 700, 50 generations ago to 100 in the current generation (Figure 4.4B).

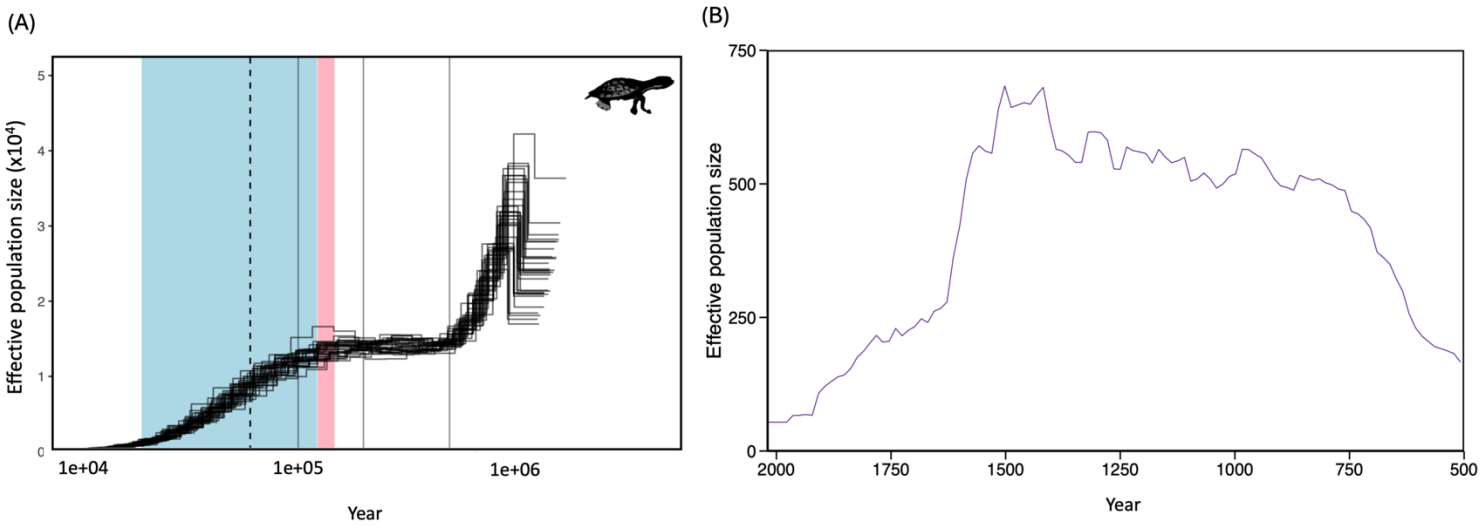


Figure 4.4. Changes in effective population size (N_e) over time estimated using **(A)** PSMC and **(B)** GoNe for *Myuchelys georgesi*. The x axes indicate time before present in years, and the y axis indicates the effective population size. A generation time of 14 years was used for both analyses. For PSMC analyses, axes were scaled by mutation rate of 4.61 x10^{-9} substitutions per site per generation. PSMC Rectangles correspond to the last inter-glacial (red, warm) and glacial period (blue, cold). Vertical dotted line represents arrival of humans to Australia.

MHC diversity

Following the identification and filtering of MHC variants, we identified 257 biallelic SNPs across all MHC exons in the Before group, including 121 SNPs in class I and 136 SNPs in class II genes (Table 4.2). In the After group, 246 biallelic SNPs were identified across all MHC exons, comprising 120 SNPs in class I and 124 SNPs in class II genes (Table 4.2). The SNPs were unevenly distributed among MHC genes, with the highest number of variants contained in *Myge-UA* (*N* = 66) and the lowest number of variants contained in *Myge-DAB3* which was monomorphic (Table 4.2). Overall, the 14 genes with variants had an average of 18 exonic SNPs (range: 3-66), most of which were in exons 2, 3 and 4 across both classes. Among the 257 SNPs identified in the 'Before' animals, 171 (67%) were predicted to result in non-synonymous substitutions. In the 'After' animals, 154 of the 246 (63%) detected variants were also predicted to be non-synonymous. Non-synonymous variation was observed across 14 genes in the 'Before' group and 13 genes in the 'After' group, respectively. A total of 15 and

14 alleles were observed in MHC I and 35 and 24 alleles were observed in MHC II in Before and After groups, respectively (Table 4.2). The Before population had one private allele in *Myge-UE* and 11 private alleles across *Myge-DAA1*, *Myge-DAA2, Myge-DAB4, Myge-DAA5, Myge-DAB5* that were unique to that population. We found no significant differences in allelic diversity between temporal groups (MHC I: t = 1.8711, df = 8, p-value = 0.09823; MHC II: t = 1.688, df = 16, p-value = 0.1108) (Table 4.2). On average, class I genes had higher numbers of SNPs, non-synonymous SNPs, and fewer alleles compared to class II. SNP H_O and H_E was similar across temporal groups (Figure 4.5, Table A3.1.8) with slightly lower observed than expected heterozygosity at MHC loci. Higher number of variants corresponded with higher levels of heterozygosity across both classes in pure *M. georgesi*. Significantly higher observed than expected heterozygosity was observed across Backcross MHC loci, indicating an excess of heterozygotes expected with early generation hybrids (Feuerstein et al., 2024). Overall heterozygosity across all MHC genes were significantly higher for backcross animals compared to pure *M. georgesi* (Figure 4.5, Table A3.1.8). PCoA analyses revealed no clear temporal differentiation in MHC genes, although clustering of individuals was observed, potentially reflecting the presence of shared alleles (Figure A3.2.6). Slightly greater differentiation was observed in the "Before" group for MHC I genes (Figure A3.2.6A), indicating a minor degree of temporal structuring. In contrast, genome-wide analyses showed no discernible clustering patterns. The inclusion of backcross individuals resulted in tighter clustering of purebred animals, reflecting the nearly identical genome-wide diversity within the species and the absence of distinct genetic structuring (Figure A3.2.7A). This observation was further validated by the exclusion of backcross individuals, which revealed no significant clustering or differentiation among pure groups, suggesting minimal genetic divergence over time (Figure A3.2.7B).

Table 4.2. Temporal diversity statistics for each MHC gene investigated including number of non-synonymous SNPs (ns); ratio of synonymous to non-synonymous SNPs (dN/dS); number of alleles; and allelic diversity. Before *N* = 19, After *N* = 12.

Class	Gene	ORF length (bp)	Before SNPs	After SNPs	Before SNPs (ns)	After SNPs (ns)	Before dN/dS	After dN/dS	Before no. alleles	After no. alleles	Before allelic diversity	After allelic diversity
MHC I	*Myge-UA*	1063	66	66	40	40	0.606	0.606	3	3	0.610	0.513
	Myge-UB	1124	23	23	16	16	0.696	0.696	2	2	0.492	0.444
	Myge-UC	1096	18	18	15	15	0.833	0.833	4	4	0.495	0.513
	Myge-UD	1096	8	8	6	6	0.750	0.750	2	2	0.499	0.444
	Myge-UE	1196	6	5	1	0	0.167	0.000	4	3	0.512	0.405
MHC II	*Myge-DAA1*	774	16	13	13	9	0.813	0.563	7	5	0.697	0.506
	Myge-DAB1	805	12	12	1	1	0.083	0.083	2	2	0.493	0.444
	Myge-DAA2	774	3	3	3	3	1.000	1.000	3	2	0.370	0.180
	Myge-DAB2	807	16	11	12	9	0.750	0.563	3	3	0.553	0.537
	Myge-DAA3	774	4	4	4	4	1.000	1.000	2	2	0.347	0.297
	Myge-DAB3	804	0	0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Myge-DAA4	775	29	27	22	21	0.759	0.724	3	3	0.590	0.509
	Myge-DAB4	807	17	15	13	11	0.765	0.647	5	4	0.674	0.675
	Myge-DAA5	751	27	27	17	17	0.630	0.630	7	4	0.763	0.555
	Myge-DAB5	777	12	12	8	8	0.667	0.667	3	2	0.562	0.486

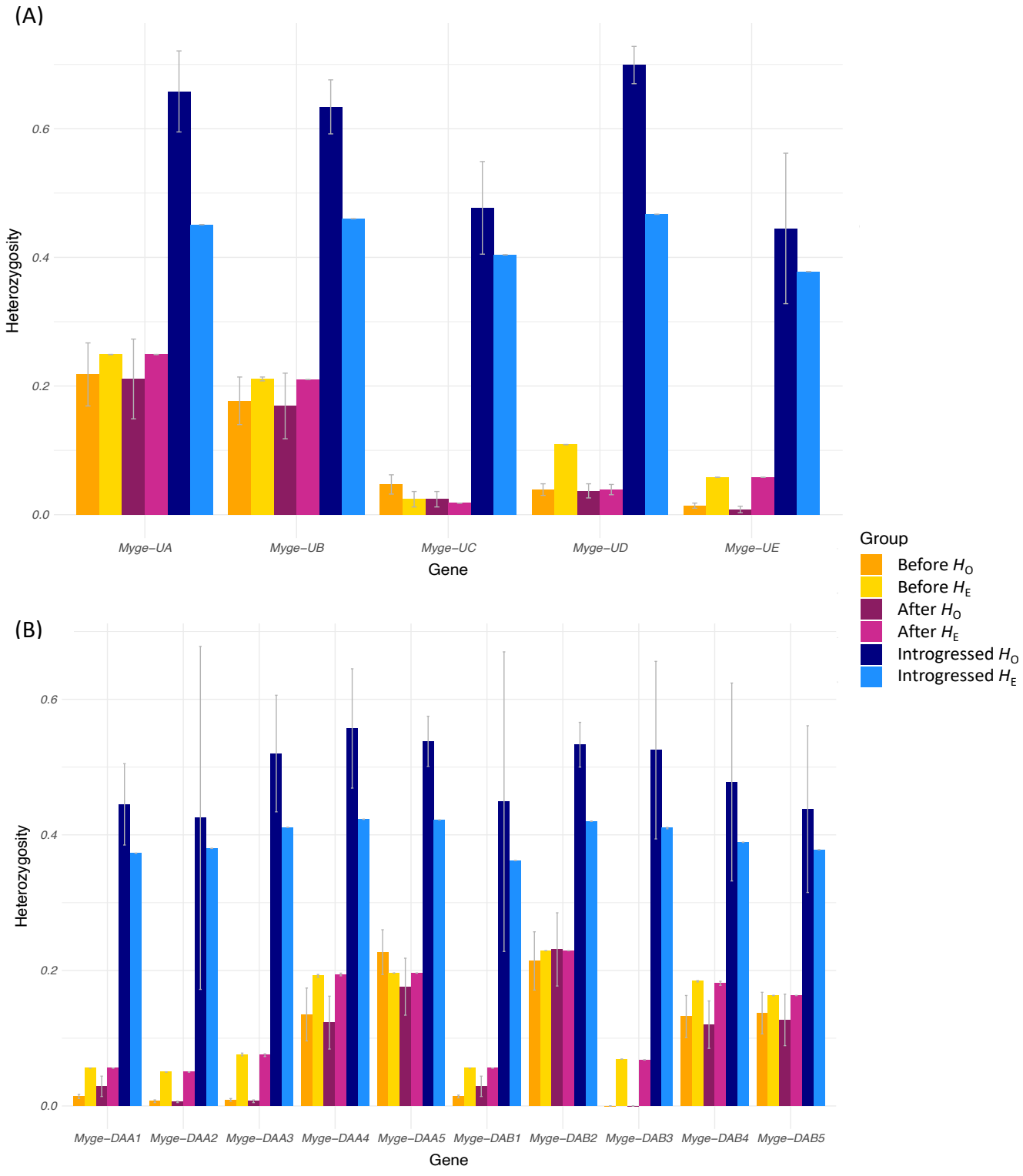


Figure 4.5. Observed (H_O) and expected (H_E) heterozygosity statistics of MHC I **(A)** and MHC II **(B)** in pure *Myuchelys georgesi*; Before (N = 19) and After (N = 12); and Backcross (Introgressed) (N = 4). Standard error bars are shown in light grey. Breakdown of gene metrics presented in Table A3.1.8.

4.6 Discussion

Here we provide a critical first step towards understanding the genome-wide and immune gene diversity in the Bellinger River turtle. We found low levels of genome-wide diversity in pure *M. georgesi* compared to backcross animals and evidence of long-term, continual declines in effective population size. We present the first comprehensive overview of the location and structure of the MHC region in a freshwater turtle species. Based on our manual annotations of the MHC genes, we observed relatively low levels of genetic variation within this gene region in pure *M. georgesi* relative to disease-resilient hybrids. However, chromosome 10, which encompasses the MHC region, exhibited higher levels of variation compared to other macrochromosomes. Consistent across both genome-wide and immune gene analyses, no significant changes in genetic diversity were detected before and after the disease outbreak.

Leveraging a chromosome-level genome assembly and RNAseq data, we identified MHC genes localised within a single core region, consistent with observations in many amniote species (Figure 4.1A). Like *M. georgesi,* these studies revealed a single-core MHC region with linked class I and class II subregions. While the relocation of MHC genes to other genomic regions has been observed in reptiles—for instance, the displacement of class I and class II genes from the core MHC region in tuatara (*Sphenodon punctatus*) and anole (*Anolis carolinensis*) (Miller et al., 2015, Gemmell et al., 2020, Card et al., 2022) —we did not find evidence for extensive duplication of functional MHC genes outside of the core region, aside from incomplete pseudogenes lacking major exons. This suggests a highly conserved MHC region with potential constraints on duplication to other regions of the genome in *M. georgesi*. The genomic organisation of each class forms two distinct clusters observed within the single-core MHC region as observed in the Chinese alligator (He et al., 2022b) and komodo dragon (Reed and Settlage, 2021), with no interspersal of classes as observed in the tuatara (Gemmell et al., 2020), and some amphibian and mammalian species (He et al., 2023b, Peel et al., 2019). These topologies align with the known evolutionary relationships within Reptilia (Zardoya and Meyer, 2001), supporting both the conservation of MHC loci and their divergence along lineage-specific trajectories. Phylogenetic analyses of immune genes similarly reflect these established relationships, with both MHC classes showing the closest divergence from

Crocodilia (Figure A3.2.2, Figure A3.2.2). This raises the question of whether these clades represent classical and non-classical genes or suggest specialised roles within the *M. georgesi* immune system. In humans, for instance, classical HLA (human leukocyte antigen) class Ia genes, such as HLA-A, are involved in antigen presentation to immune cells, are highly polymorphic, and play a critical role in adaptive immunity (Le Bouteiller and Lenfant, 1996). In contrast, non-classical HLA class Ib genes such as HLA-G and HLA-E exhibit specialised immunomodulatory functions, are highly conserved and characterised by limited polymorphism, and are predominantly expressed in immunologically significant tissues (LeMaoult et al., 2003). Consistent with the genomic architecture observed in amphibians, MHC class I introns in *M. georgesi* are notably longer than those of MHC class II, where MHC II DAA and MHC II DAB are tightly linked (He et al., 2023b). These key features suggest that the MHC region in *M. georgesi* is relatively simple as the core MHC I and MHC II regions tightly clustered and adjacent in the genome. This simplified genomic arrangement appears to have been largely conserved throughout the evolutionary history of tetrapods, including avian-reptiles (He et al., 2022a), non-avian reptiles (Miller et al., 2015, Card et al., 2022, He et al., 2022b), and mammals (Silver et al., 2024), particularly in MHC II genes.

Our demographic reconstructions indicate that the effective population size of *M. georgesi* has been declining since the last interglacial period (Figure 4.4), potentially due to limited gene flow and accumulation of ROH. Based on mitochondrial data, *M. georgesi* is estimated to have diverged from its closest relative, *E. macquarii*, approximately 6.1 million years ago during a period of aridification (Le et al., 2013). Geographic isolation and habitat fragmentation likely contributed to the demographic trends identified in our PSMC analyses. The low N_e estimates (<500) from approximately 10,000 years ago (Figure 4.4A) suggest that climatic changes during the last glacial period may have influenced the genetic structure of *M. georgesi,* similar to patterns observed in other freshwater turtle species. However, unlike these common turtle species, which eventually reach an equilibrium in N_e (Hilgers et al., 2024), *M. georgesi* has continued to experience an ongoing decline. The glacial period likely reduced habitat availability and fragmented freshwater populations, leading to isolation, speciation, and long-term declines in effective population size. The species isolation, combined with life-history traits, have likely intensified the impact of accumulated inbreeding,

resulting in reduced fitness and an elevated risk of extinction over time (Frankham et al., 2017). The long-term isolation and gradual decline in effective population size represent key findings for the species, likely exacerbating the critically low levels of contemporary genome-wide diversity observed in heterozygosity and ROH analyses. Consequently, this accumulation of inbreeding effects, compounded by geographic barriers to dispersal, may have driven the species into an extinction vortex well before the onset of the Anthropocene. This is supported by the large proportion of the genome found in ROH (Figure 4.2A), as long-term inbreeding and gene flow can directly influence ROH abundance (Ceballos et al., 2018, Foote et al., 2021, Mooney et al., 2021, Hewett et al., 2023). Short ROH are indicative of background relatedness or inbreeding resulting from distant common ancestry, while long ROH reflect recent parental relatedness or occur in genomic regions with low recombination rates (McQuillan et al., 2008, Pemberton et al., 2012). In *M. georgesi*, the ROH distribution includes both short (<2 Mb) and long (>2 Mb) segments (Figure 4.2A, Figure A3.2.4, Figure A3.2.5). The elevated levels of homozygosity observed in the contemporary population are likely the consequence of both historical inbreeding among distant ancestors and recent background relatedness. This is further supported by the recent declines identified through GoNe analyses, which have likely contributed to the significant proportion of the genome comprising of long ROH (Figure 4.4B, Table A3.1.7, Figure A3.2.5B) (Pemberton et al., 2012, Kardos and Shafer, 2018). ROH, which are typically regions identical by descent (Ceballos et al., 2018), are likely absent in backcross animals due to increased genetic diversity and reduced inbreeding in parental *E. macquarii*. Additionally, the genetic divergence (7.8%) between *M. georgesi* and *E. macquarii* likely contributes to the prevalence of heterozygous offspring, as sequence differentiation between the species prevents the formation of homozygous regions (Fielder et al., 2012). The absence of significant differences in both the number and length of ROH and in autosomal heterozygosity across temporal groups suggests that genomic signatures associated with the recent disease outbreak are not yet detectable at the genome-wide level in a species with long generation times. However, these signatures may emerge in subsequent generations, potentially increasing the species' trajectory towards extinction, emphasising the importance of ongoing genetic monitoring.

Here we found significantly higher levels of expected (H_E) than observed (H_O) at MHC loci. We previously found higher H_O than H_E heterozygosity in the species using 460 putatively neutral SNP markers (Nelson et al., 2024). Although neutral genetic variation is frequently employed as a proxy for adaptive potential, our findings support the argument that neutral diversity may not fully capture adaptive potential, and that genetic diversity at specific functional loci are a more informative indicator of a population's adaptive potential (Teixeira and Huber, 2021, Holderegger et al., 2006).

MHC variants play a critical role in a wide range of biological traits. For instance (i) immune recognition where a single amino acid change in the antigen-binding region of the DRB*1302 allele in humans abolishes malaria recognition (Frank, 2002, Summers et al., 2003); (ii) susceptibility to infectious and autoimmune diseases, as observed in MHC IIB heterozygotes, which exhibit lower Ranavirus infection intensity compared to homozygotes in larval wood frogs (*Rana sylvatica*) (Savage et al., 2019); (iii) individual odours and mating preferences seen in song sparrows (*Melospiza melodia*) (Grieves et al., 2019) and other aves (Leclaire et al., 2014), where preen oil odour is used to discern MHC similarity and diversity of potential kin or mates and; (iv) pregnancy outcomes observed in giant pandas (*Ailuropoda melanoleuca*), where mating pairs with MHC dissimilarity exhibit higher reproductive success (Sommer, 2005, Zhu et al., 2019). High levels of inbreeding often lead to an overall decrease in MHC variants, loss of rare and potentially advantageous alleles, and a decreased ability to adapt to novel or rapidly evolving pathogens (Altizer et al., 2003, Spielman et al., 2004). Despite the low genome-wide diversity observed in *M. georgesi*, MHC core regions remains relatively conserved, displaying a uniform architecture and notable number of polymorphisms when compared to broader taxa (Silver et al., 2024). The lack of ROH seen in the core MHC region, and chromosome 10 in general, relative to other macrochromosomes suggests that maintenance of the MHC core region has disrupted ROH. This may indicate balancing selection maintaining MHC diversity, as genetic diversity across the genome was lost (Oliver and Piertney, 2012, Buzan et al., 2022). Consequently, low levels of variation in MHC genes may not be the primary driver contributing to the species' disease susceptibility as initially hypothesised. Instead, the interspecific diversity introduced by *E. macquarii* across the genome more broadly, may be a significant contributor to disease resilience as reflected in

the higher survival rates observed during surveys of backcross individuals compared to their pure *M. georgesi* counterparts (Parrish et al., 2024, Yang et al., 2024).

Individuals from the After disease group exhibit a reduced number of SNPs, non-synonymous mutations, and alleles at six out of the 15 MHC genes when compared to the Before group. These observations may reflect a possible sampling artifact or could represent early evidence of allele loss in the wild, with some of the remaining alleles potentially conferring resistance to recent nidovirus infections, as seen in amphibians with chytridiomycosis (Fu et al., 2023). However, longer-term monitoring of subsequent generations will be necessary to validate these findings.

Several limitations must be considered when interpreting these results. The first limitation is the scarcity of complete coding sequences for testudines and other reptilian species, which hinders the ability to draw high-resolution conclusions about the evolutionary histories and classification of MHC homologs in *M. georgesi* and testudines more broadly. Compared to the extensive genomic and immunological resources available for other tetrapod clades, such as marsupials (Belov et al., 2013), reptiles, particularly testudines, still lack comprehensive immune gene resources. This deficiency in data restricts the ability to resolve evolutionary relationships of ancient lineages effectively. As next-generation sequencing (NGS) becomes more common for non-model organisms, the increasing availability of genomic data and resources will help address the current deficiency in MHC nomenclature and classification.

Second is the accuracy of mapping of backcrossed *E. macquarii* to the *M. georgesi* genome. A significant limitation is the lack of annotated MHC genes for *E. macquarii*, hindering the accurate comparison of genetic diversity between the two species. Despite achieving high mapping rates for both pure and backcross animals, the lack of MHC locus annotations in *E. macquarii* makes it difficult to accurately assess locus homology, which is critical for understanding the genetic dynamics of hybridisation. Our assumption that both species share a similar number of loci may not hold, as differences in genome structure, such as duplications or gene family expansions, regions with complex architecture, and differences in repeat

content between species may lead to underrepresentation or misidentification of loci in one or both species. Furthermore, it is challenging to determine whether discrepancies in results arise from misalignments caused by structural variations or paralogous sequences that have interfered with correct locus identification. The presence of uncharacterized *E. macquarii* MHC loci could have influenced the comparative patterns of MHC diversity observed in this study. Without further investigation, it is challenging to determine whether the observed variation within annotated genes is contributing to disease resilience or if it is primarily due to differences in MHC copy number in *E. macquarii*.

To address these limitations, future research should focus on annotating MHC genes in *E. macquarii* using long read sequencing to allow for a more accurate comparison of the genomic effects of hybridisation, both across the genome and specifically within MHC loci. Additionally, sampling first-generation hybrids and additional backcross individuals will offer a more comprehensive view of genomic dynamics, facilitating a deeper understanding of the extent of introgression and its potential implications for genome-wide and immune function, and overall fitness.

4.7 Conclusion

Our study reveals higher levels of diversity at MHC genes compared to neutral SNPs across the genome of this critically endangered turtle. Using re-sequenced genomes, we have been able to reconstruct the population's demographic history showing a marked ongoing, gradual decline in effective population size, indicating a long-standing trajectory towards genetic depletion. Genome-wide diversity appears to have diminished over time and will possibly be exacerbated by the significant population crash caused by the 2015 nidovirus outbreak although the species' long generation time is likely masking any further declines at this time. Here we have presented a foundational comprehensive overview the MHC region in a freshwater turtle species providing baseline information for future studies in testudines. Our findings emphasise the urgent need for conservation strategies that address the long-term impacts of genomic erosion and prioritise the preservation of the remaining MHC gene diversity.

Chapter 5 - General discussion, future directions and conclusions

In this thesis, I demonstrate the application of genomic technologies, including the use of a reference genome and next-generation sequencing (NGS), to address specific conservation needs and management-driven questions for the conservation breeding and recovery program of the Bellinger River turtle *M. georgesi*. Although this body of work focuses on a single species, the approaches and methodologies outlined (Figure 5.1) have broad applicability across diverse taxa, offering valuable insights for conservation planning for species facing similar challenges.

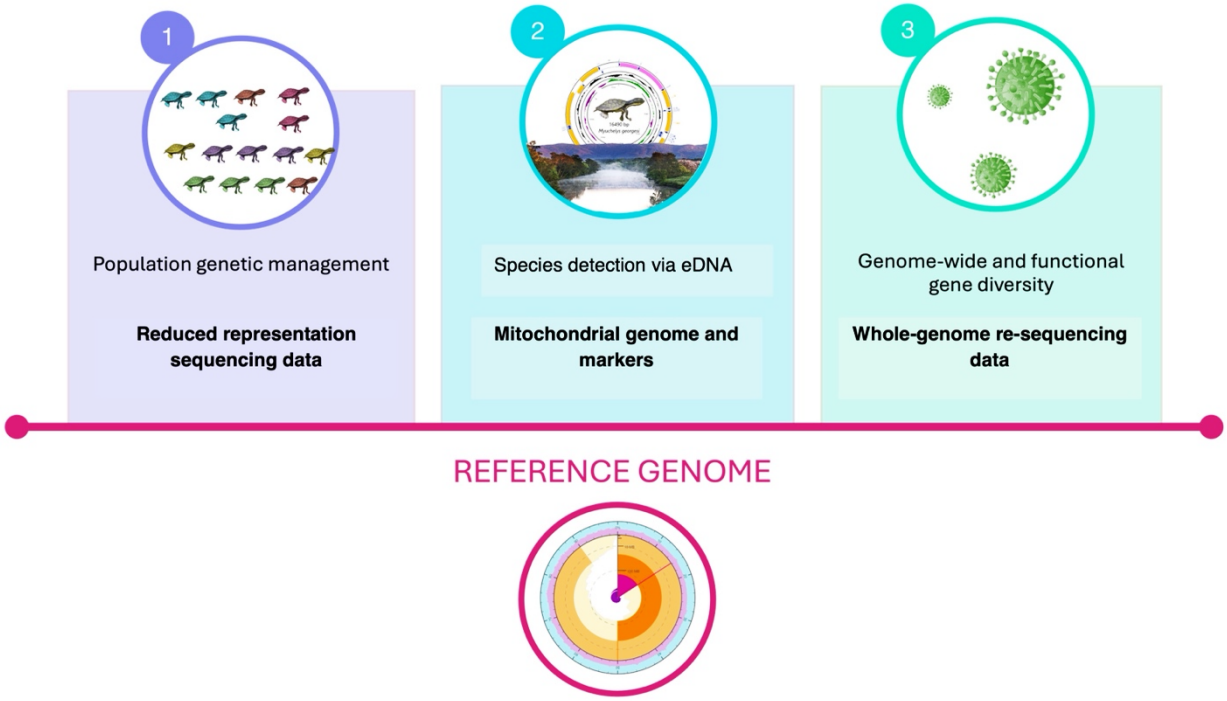


Figure 5.1. Overview of the key management focus areas and associated data types utilised throughout this thesis. All data chapters leveraged the reference genome assembled in Chapter 2 as a fundamental resource. The methodologies presented in this thesis are designed to be broadly applicable across diverse taxa, with relevance extending beyond *Myuchelys georgesi*.

My main findings were:

(a) The wild population of *M. georgesi* shows low but stable genetic diversity, with no detectable bottleneck from the 2015 Bellinger River virus outbreak, and the combined captive populations effectively represent historical diversity, with lower mean kinship observed between populations than within each.

(b) I developed a framework extending on the six-stage model by Frankham and colleagues, providing managers with steps for integrating genetic data into conservation breeding programs. This was published in *Conservation Science and Practice* in 2024 (Chapter 2).

(c) I constructed a mitogenome for eDNA primer development, improving molecular monitoring of *M. georgesi* in the Bellinger River. This work has been accepted in *Ecology and Evolution* (Chapter 3).

(d) I annotated the Major Histocompatibility Complex (MHC), identifying five MHC class I and ten MHC class II genes on chromosome 10 and show that *M. georgesi* has a simple, conserved MHC core region. This work is under review in *Immunogenetics* (Chapter 4).

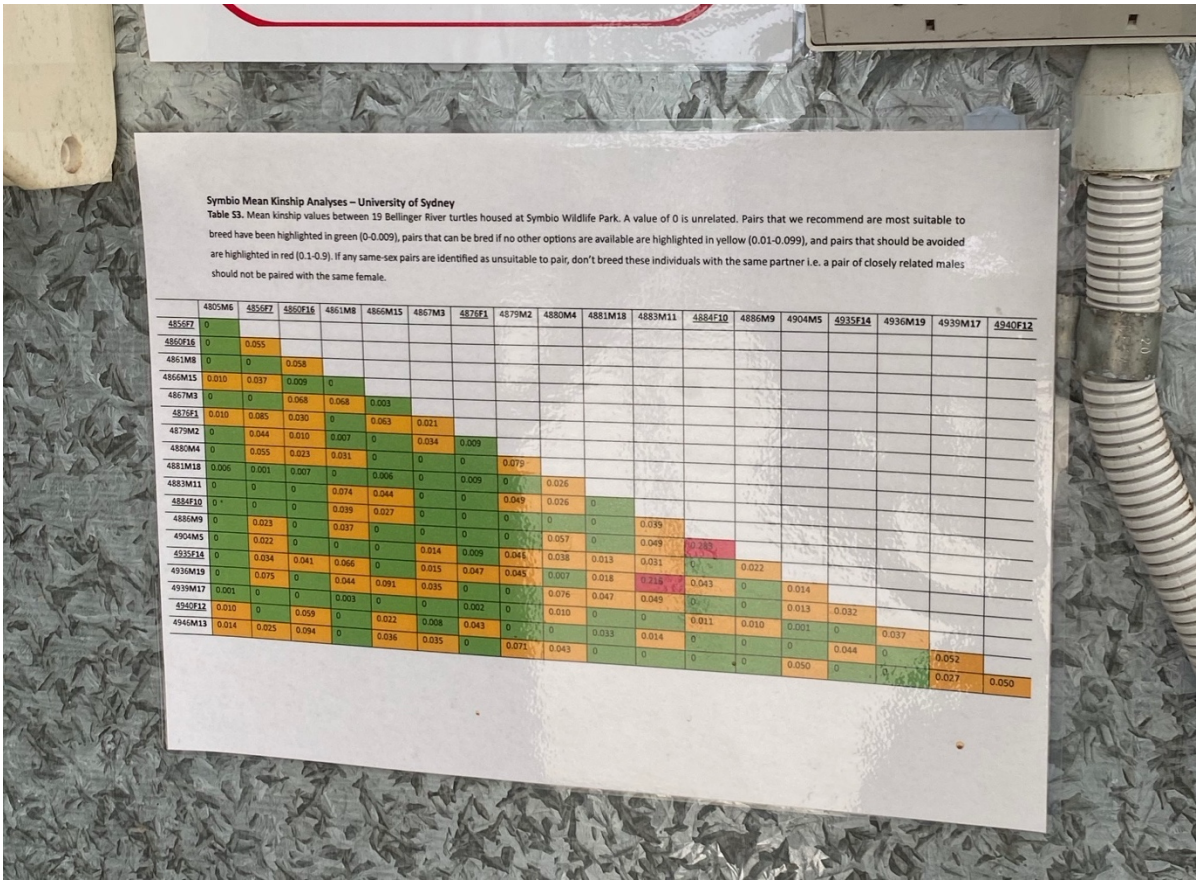
(e) Comparisons with disease-resilient hybrids show lower genome-wide and MHC diversity in pure *M. georgesi*, with no significant short-term genetic changes post-2015 virus outbreak (Chapter 4).

(f) Demographic history reconstructions indicate a gradual decline in effective population size since the last interglacial period, shaped by historical and environmental pressures (Chapter 4).

These findings provide valuable genetic and genomic insights for *M. georgesi* and represent a significant addition to the molecular data available for freshwater turtles (testudines). The methods employed to achieve these findings demonstrate how genomics can be effectively leveraged to answer conservation questions for populations in decline. In this chapter, I generalise the findings for *M. georgesi* where possible to establish relevance to conservation efforts more broadly, explore limitations and avenues for future research, and provide concluding remarks.

5.1 Management recommendations and their implementation for *Myuchelys georgesi*

A fulfilling aspect of this thesis has been the development of collaborative relationships with the recovery team and managers, including captive managers at Symbio Wildlife Park and Taronga Conservation Society Australia, and the NSW DCCEEW field teams, who have supported this research and facilitated its integration into practical management strategies (Figure 5.2) by recognising the critical role of genetic data in conservation planning and integrating genomic objectives into the Bellinger River Turtle conservation action plan in 2022 (Jakob-Hoff et al., unpublished). The multiple complementary perspectives brought to the table by researchers, policy makers and on ground management teams was an essential element in the success of this project.



Symbio Mean Kinship Analyses – University of Sydney

Table S3. Mean kinship values between 19 Bellinger River turtles housed at Symbio Wildlife Park. A value of 0 is unrelated. Pairs that we recommend are most suitable to breed have been highlighted in green (0-0.009), pairs that can be bred if no other options are available are highlighted in yellow (0.01-0.099), and pairs that should be avoided are highlighted in red (0.1-0.9). If any same-sex pairs are identified as unsuitable to pair, don't breed these individuals with the same partner i.e. a pair of closely related males should not be paired with the same female.

	4805M6	4856F7	4860F16	4861M8	4866M15	4867M3	4876F1	4879M2	4880M4	4881M18	4883M11	4884F10	4886M9	4904M5	4935F14	4936M19	4939M17	4940F12
4856F7	0																	
4860F16	0	0.055																
4861M8	0	0	0.058															
4866M15	0.010	0.037	0.009	0														
4867M3	0	0	0.068	0.068	0.003													
4876F1	0.010	0.085	0.030	0	0.063	0.021												
4879M2	0	0.044	0.010	0.007	0	0.034	0.009											
4880M4	0	0.055	0.023	0.031	0	0	0	0.079										
4881M18	0.006	0.001	0.007	0	0.006	0	0.009	0	0.026									
4883M11	0	0	0	0.074	0.044	0	0	0.049	0.026	0								
4884F10	0	0	0	0.039	0.027	0	0	0	0	0	0.039							
4886M9	0	0.023	0	0.037	0	0	0	0	0.057	0	0.049	0.283						
4904M5	0	0.022	0	0	0	0.014	0.009	0.041	0.038	0.013	0.031	0	0.022					
4935F14	0	0.034	0.041	0.066	0	0.015	0.047	0.045	0.007	0.018	0.215	0.043	0	0.014				
4936M19	0	0.075	0	0.044	0.091	0.035	0	0	0.076	0.047	0.049	0	0	0.013	0.032			
4939M17	0.001	0	0	0.003	0	0	0.002	0	0.010	0	0	0.011	0.010	0.001	0	0.037		
4940F12	0.010	0	0.059	0	0.022	0.008	0.043	0	0	0.033	0.014	0	0	0	0.044	0	0.052	
4946M13	0.014	0.025	0.094	0	0.036	0.035	0	0.071	0.043	0	0	0	0	0.050	0	0	0.027	0.050

Figure 5.2. Mean kinship table for 19 Bellinger River turtles at the captive breeding facility at Symbio Wildlife Park, Helensburgh, New South Wales, Australia

Throughout Chapters 2 to 4, I employed a range of genomic technologies and bioinformatic methodologies to generate genetic data to inform and achieve the objectives established at the workshop which included deciphering captive founder relationships (Chapter 2), developing eDNA primers (Chapter 3), and characterising MHC diversity (Chapter 4) (Jakob-Hoff et al., unpublished). Several of my findings have informed key recommendations aimed at managing the genetic integrity of captive and wild populations.

Recommendation: Maintaining genetic diversity in captivity

(i) Breed individuals within both Taronga and Symbio that have low mean kinship, including across captive populations to minimise founder effects and prevent inbreeding depression.

When appropriately managed to maintain genetic diversity and minimise inbreeding, CBRP can be a valuable for dwindling populations. Although captive breeding removes species from their threatening processes, captive breeding programs often limit population sizes, which—especially if sourced from bottlenecked populations—can compromise the maintenance of genetic diversity if not managed effectively. Furthermore, unequal founder contributions can create a conflict between minimising mean kinship and minimising inbreeding. For example, matings between underrepresented and overrepresented founder lines may reduce inbreeding in the short term, but they constrain future options for expanding the underrepresented lineage without simultaneously increasing the contribution of the overrepresented lineage. My results have shown significantly low levels of genome-wide diversity in *M. georgesi*, likely due to long-term geographic isolation and historical bottlenecks (Chapter 2, Chapter 4), therefore careful genetic management and breeding decisions are needed to prevent further declines in diversity. Methods for demographic and genetic recovery typically involve reinforcement translocations sourcing individuals from different populations (Glassock et al., 2021; Grueber et al., 2017; Heber et al., 2013; McLennan et al., 2020, Magid et al., 2022). As *M. georgesi* has only one extant wild population, animals cannot be sourced from other populations to boost genetic diversity, therefore it is essential to implement a carefully managed breeding program within the single existing population to maintain genetic diversity.

I have shown that there are lower levels of empirical mean kinship across captive populations relative to within each population (Chapter 2). I recommend prioritising breeding pairs with the lowest mean kinship within each captive group, as well as across populations, to maximise genetic diversity in future captive generations. By focusing on low-kinship pairings identified in this study, there is potential to counteract genetic drift, enhance disease resilience, and prevent further diversity loss. Although immediate increases in genetic diversity may be limited due to the long generation times of *M. georgesi*, admixture over successive generations could strengthen the population's genetic resilience.

Since the commencement of this work, the wild population of *M. georgesi* has nearly doubled, thanks to the successful release of captive-bred individuals from Taronga highlighting the importance of this research as a genetic reference for the expanding population. Consequently, I recommend ongoing, long-term genetic monitoring to ensure that no genetic drift occurs in the wild due to overrepresentation of Taronga-bred individuals. This will provide the opportunity to study the impact of captive breeding and release on genetic diversity, particularly whether these efforts have altered, enhanced or stabilised genetic diversity as the wild population grows.

Recommendations: Genetically informed releases

(ii) Translocate from both Taronga and Symbio to ensure representation of genotypes from both captive populations in the wild (Chapter 2).

(iii) Vary release locations in each round to promote genetic admixture between captive-bred and wild individuals, addressing existing genetic differentiation between these populations (Chapter 2).

(iv) Uptake of molecular monitoring of *M. georgesi* to resolve questions around distribution and inform release site selection using eDNA (Chapter 3).

In this study, I demonstrated that the combined genetic diversity of the captive populations effectively represents the historical diversity observed in wild populations prior to the disease

outbreak. To ensure that this diversity is maintained in the wild, I recommend reintroductions of individuals from Symbio, as this would incorporate genotypes from both captive populations into the wild population. Introducing individuals from Taronga Zoo and, subsequently, Symbio at release sites on opposite ends of the river from their original collection points could facilitate natural admixture, promoting genetic diversity across upper and lower river regions. Additionally, I suggest adoption of eDNA detection tools developed in Chapter 3 to optimise both survey and release site selection. These tools can help identify areas where more intensive surveys—such as diving and trapping—are warranted, including locations like the Kalang River, to address questions regarding the species' current distribution. This approach will enable more precise identification of suitable habitats, ensuring that turtles are released in areas that encompass their maximum historical range, thereby maximising genetic admixture across the species' distribution (Chapter 3).

Management beyond neutral diversity

Spencer et al. (2018) proposed that low genetic diversity may contribute to *M. georgesi*'s heightened susceptibility to disease, prompting the incorporation of MHC-related objectives into the species' conservation action plan. Addressing complex questions concerning genetic diversity and functional gene families, and their associations with specific traits, such as disease susceptibility, often requires a high-quality reference genome (Chapter 2) (Peel et al., 2022), whole-genome resequencing data (WGR) (Chapter 4) or targeted sequencing (Fuentes-Pardo and Ruzzante, 2017). Additionally, baseline data on functional genes in the species or closely related species, such as immune gene annotations, are needed before downstream correlations between genetic diversity and susceptibility can be made. This work provides the first detailed understanding of immunogenetic architecture in a freshwater turtle species and offers markers for future monitoring of immune diversity in both wild and captive *M. georgesi* populations (Chapter 4). The WGR I generated identified significant differences in genetic variation between purebred and hybrid individuals.

The observed resilience of disease-resistant hybrids (Parrish et al., 2024) highlights potential benefits of hybrid vigour as a potential future management strategy, through introduction of

additional loci or disease resilient alleles (In Chapter 4). Human-mediated outbreeding - using the species' closest relative (*E. macquarii*) - could serve as a genetic rescue strategy for *M. georgesi*. However, this approach warrants caution, as the benefits of hybrid vigour, including the introduction of additional loci, often diminish in subsequent generations due to the risk of epistatic interaction breakdown and a return to higher homozygosity levels (Frankham et al., 2011, Nichols et al., 2024). For this strategy to sustain elevated genetic diversity over the long term, the consistent introduction of new individuals would be necessary. Additionally, such a program would require careful management and continuous monitoring of developmental and reproductive outcomes to prevent hybrid swamping, which could lead to the eventual loss of the parental *M. georgesi* genome (Tensen and Fischer, 2024).

As genetic and genomic technologies progress, alternative genetic rescue techniques may allow the reintroduction of genetic variation that once existed in the historical populations or from other species using synthetic biology methods (Kosch et al., 2022). Recent advancements in gene-editing technologies, particularly through clustered regularly interspaced short palindromic repeats (CRISPR-Cas9) technology, have shown substantial potential in model organisms (Chojnacka-Puchta and Sawicka, 2020, Zhang et al., 2021) and may soon be adaptable for enhancing disease resistance in non-model species, such as *M. georgesi*, through the introduction of disease resilient alleles from *E. macquarii*. Given the critically low genetic diversity and the high homozygosity (90%) in the species' genome, future investment in both hybridisation and biotechnology approaches is recommended, though both require further research and refinement.

The long-term decline in effective population size identified using WGR data in Chapter 4, has likely led to the low levels of genetic diversity that is unlikely to recover through traditional management techniques alone. Life-history traits—such as low reproductive rates, delayed maturation, and extended generation times—present substantial obstacles to demographic and genetic recovery, increasing the likelihood of ongoing decline in an extinction vortex for the species (Howey and Dinkelacker, 2013). Therefore, approaches incorporating hybridisation or advanced biotechnology may be essential for the long-term viability of *M. georgesi*.

5.2 Genomic resources for species management

A genetic roadmap for dwindling populations

Although the value of genetic and genomic data in managing threatened species is well-recognised, translating this research into actionable conservation strategies remains a significant challenge. A key barrier is that practitioners often lack access to up-to-date, peer-reviewed literature, which hinders their ability to stay informed about recent advancements in the field, and when and how to effectively integrate genetic data into CBRPs. To address this, the framework I developed in Chapter 2 provides guidance for translating genetic research into practical management actions. This framework includes a generalised checklist, informed by my example of integrating genetic data into the *M. georgesi* CBRP, which can be adapted for use in other species. The framework is structured around Frankham's six stages of establishing a captive breeding program, allowing it to accommodate each stage of the program's development. This structure also provides flexibility for the retrospective integration of genetic data where needed, ensuring that genetic insights can be effectively incorporated at various stages in the program. It serves as a practical tool for managers to better understand the genetic considerations involved in breeding programs and to approach researchers with the relevant expertise to undertake necessary analyses. Additionally, the framework offers geneticists a structured set of guidelines on how to align their work with conservation management goals, ensuring that genetic data is effectively utilised to improve species conservation. This roadmap helps bridge the gap between scientific research and on-the-ground conservation efforts, making genetic insights more accessible and applicable in real-world management contexts.

For dwindling populations, I show how a high-quality reference genome for the species of interest is a useful tool for high-resolution genomic investigations and how it can provide a blue-print for undertaking analyses to answer management-driven questions. Despite the importance of such a resource, prior to this study, no chromosome-level reference genome was available for any of the approximately 60 species within the family Chelidae. Using a preserved sample from 1986—demonstrating the efficacy of long-term biobanking—I

successfully generated a chromosome-level genome assembly (Chapter 2). This assembly now represents a pivotal genomic resource for advancing evolutionary, ecological, and conservation research in this lineage of side-necked turtles and for testudines more broadly. During this thesis I built on the reference genome I generated (Fig 5.1), and then leveraged two other next-generation sequencing (NGS) approaches—reduced representation sequencing (RRS) (Chapter 2) and whole-genome resequencing (WGR) (Chapter 4)—as well generating mitochondrial data (Chapter 3) to accompany the reference genome for downstream applications, highlighting the advantages of having a reference genome.

Conservation genomics of freshwater testudines

Reptile populations are rapidly declining due to anthropogenic threats such as climate change, habitat destruction, pollution, and disease. With their long generation times, limited dispersal abilities, and temperature-dependent physiology, reptiles are especially vulnerable to environmental changes (Li et al., 2024). However, reptiles often receive less conservation attention compared to other taxa, resulting in limited data availability in the scientific literature and online databases (Olson and Pilliod, 2022). The importance of genomic approaches is amplified as emerging infectious disease, as rising temperatures and environmental stressors can facilitate the spread of pathogens, while also exacerbating the physiological burden on already vulnerable populations. Genomic tools provide a crucial framework for understanding how host-pathogen interactions evolve under these conditions and for anticipating which species or populations may be most at risk.

The data I generated for *M. georgesi* significantly enhances the genomic data available for testudines. As the most complete genome for this group to date, it provides a critical reference for studying genetic diversity, genomic structure, and phylogenetic relationships among species within this suborder for testudines species more broadly, especially those lacking comprehensive genomic resources. The data can be used to gain a deeper understanding of lineage divergence and evolutionary adaptations and help elucidate how environmental pressures, such as disease, influence genetic diversity and evolutionary mechanisms. For example, identification of conserved genetic elements across side-necked turtles may indicate

functional constraints crucial for adaptive potential, which can be used to inform genetic-based conservation breeding decisions.

The immune gene annotations generated for *M. georgesi* represent the first for a Testudines species, offering novel insights into MHC evolution. This dataset provides a critical baseline for comparative studies, enabling the identification of genetic markers associated with immune responses. This is particularly relevant as the emergence of novel infectious diseases becomes an increasingly significant stochastic threat to wildlife, especially small and vulnerable populations. Furthermore, my results can inform conservation strategies for other threatened reptile species by providing reference points for genetic diversity and inbreeding levels (Chapter 2), as well as identifying critical loci that can be monitored for adaptive traits (Chapter 4).

As conservation strategies increasingly prioritise genetic diversity at adaptive loci to foster disease-resilient populations through carefully managed breeding programs (Chhina et al., 2024), the role of genomics in understanding species resilience, tracking pathogen evolution, and informing species management will become critical in addressing the growing prevalence of disease-driven declines in wildlife populations globally. By integrating the genomic data generated in this study into the expanding collection of publicly accessible genomic resources in online databases, this work contributes valuable data for the conservation of other at-risk populations facing similar disease threats, providing a foundation for further research and management strategies.

5.3 Key limitations

The generation time of Myuchelys georgesi

The relatively long generation time of *M. georgesi* limits the feasibility of short-term temporal analyses, making it challenging to track changes in genetic diversity, population structure, and adaptive potential across generations within a limited timeframe. Longitudinal studies encompassing multiple generations are essential for detecting and interpreting evolutionary

responses to selective pressures, such as the 2015 disease outbreak and subsequent outbreaks, over time. Without such long-term data, conclusions about the potential for adaptive changes in response to conservation actions remain speculative.

eDNA markers in the field

Although eDNA markers were developed as part of this research, they have not yet been tested in the river. Field trials of these markers are crucial to validate their effectiveness for species detection and monitoring of population distribution. The lack of field validation leaves uncertainty regarding the practical utility of these markers under real-world conditions, potentially delaying their application in ongoing conservation monitoring efforts.

Metadata and disease resilience

Our study lacked metadata on individual resilience to disease, meaning we could not correlate alleles with disease resilience. Identifying resilient individuals and correlating their genotypes with levels of resilience would be invaluable for management strategies aimed at decreasing disease susceptibility. The absence of these data limits our understanding of genetic factors that may contribute to disease resilience, particularly in pure populations of *M. georgesi*, thereby constraining adaptive potential breeding decisions in the CBRP.

Hybrid analyses

Determining whether discrepancies between pure and backcross animals arise from misalignments due to structural variations or interference from paralogous sequences affecting accurate sequence alignment across backcross genomes can be difficult with short-read data. Synteny analyses between pure species could offer valuable insights into chromosomal rearrangements. Additionally, this study uses short-read data to focus on biallelic SNPs as proxies for genetic diversity. Exploring structural variants and indels represents another promising approach to gain deeper insights into genomic diversity within and between species. These elements can span multiple alleles, genes, and gene regions simultaneously, potentially exerting a more substantial influence on fitness (Wold et al.,

2021). As we did not conduct large-scale, high-resolution MHC analyses in hybrids—due to the potential for bias from unidentified copy number variants or unannotated loci in pure *E. macquarii*—future functional studies are needed. Specifically, annotating MHC genes in *E. macquarii* and applying targeted amplicon sequencing would enable more detailed and accurate comparisons of the genomic impacts of hybridisation at MHC loci, including gene duplications, copy number variation, and expansions in gene families (Peel et al., 2022). Additionally, greater sequencing effort of first-generation hybrids and additional backcross individuals will offer a comprehensive view of genomic dynamics, facilitating a deeper understanding of the extent of introgression and its potential implications for genome-wide and immune function, and overall fitness across a larger sample set.

5.4 Where to from here?

The emergence of novel infectious diseases is rapidly becoming a major stochastic threat to wildlife, particularly for small and vulnerable populations. Detailed immune gene characterisation across diverse taxa is increasingly important to understand and mitigate these risks – this study provides a foundational resource toward that goal. The characterisation of the core MHC region presented here offers the first complete MHC reference for a testudines species, filling a critical gap in reptilian genomics for future studies. Due to the limited availability and quality of immune gene data in other testudines, this research lays essential groundwork for comparative analyses that were not possible until now. These baseline data are particularly valuable as they enable future studies to identify disease resistance mechanisms and support conservation strategies for turtles and reptiles facing similar disease threats.

As new disease outbreaks continue to emerge in the Bellinger River, continual research on the immune responses of *M. georgesi* will be increasingly critical. Future research should focus on genome-wide and immune-related gene diversity, particularly in pure *E. macquarii*, as undertaken by Bentley et al. (2022) in green sea turtle (*Chelonia mydas*) and leatherback turtle (*Dermochelys coriacea*). Comparative synteny analyses between *E. macquarii* and *M. georgesi*, including assessments of structural rearrangements in hybrid and backcross

individuals, would provide further insight into genetic diversity and the mechanisms underlying *E. macquarii*'s resilience and *M. georgesi*'s susceptibility to disease. Future research should also prioritise the collection of samples and accompanying metadata, from *M. georgesi* that exhibit heightened resilience to recent viral outbreaks and those that succumb to disease. Such a sample set could be used to identify potential alleles or variants associated with pathogen resilience, enabling the development of management strategies that target resilience at a functional level, extending beyond neutral genetic markers alone (Chapter 1).

Lastly, hybridisation between *M. georgesi* and *E. macquarii*, raises significant questions about species boundaries and conservation priorities. Given their ability to hybridise easily, it prompts the question of whether these species should be considered separate, especially if gene flow is frequent and results in viable offspring. As hybrids likely fill similar ecological roles as pure *M. georgesi* individuals, *E. macquarii* offers a potential source of genetic variation, making hybridisation a possible strategy for genetic rescue in critically endangered species. However, the conservation concern lies in the potential loss of the unique genetic and evolutionary lineage of *M. georgesi* and the risk that ongoing hybridisation could undermine local adaptations. These considerations highlight the complexity of managing hybrid populations and emphasise the need for conservation planning that balances genetic, ecological, and evolutionary considerations among managers, scientists, and the local community.

5.5 Conclusion

The collaborative framework established in this thesis has enabled a research program that closely aligns with conservation management objectives, ensuring that the genetic findings directly support actionable strategies for future conservation efforts. This body of work provides essential genomic data to guide the ongoing and future management of the *M. georgesi* CBRP. The collaboration between the University of Sydney, the University of Canberra, and the NSW DCCEEW has ensured that research objectives were aligned with management needs. This work demonstrates how a reference genome and accompanying genomic data can be effectively harnessed and translated into actionable management

outcomes. Beyond generating genomic data, tools, and recommendations for *M. georgesi*, this thesis offers valuable examples, methodologies, a framework, and novel genomic insights to assist researchers and conservation practitioners working on other testudines, reptiles, and threatened species more broadly.

As we continue to expand the global database of genomic resources for both threatened and non-threatened species, it is vital for researchers and conservation managers to work together to utilise this data effectively to achieve conservation outcomes. As the global biodiversity crisis intensifies, the role of genomics within a multidisciplinary conservation framework becomes ever more essential for safeguarding the biodiversity of our planet. Although genomic data alone cannot rescue species such as *M. georgesi*, our conservation strategies are increasingly reliant on empirical data and evidence-based approaches. The more comprehensive our understanding, the more effectively we can advocate for endangered species, gain political support and funding for conservation initiatives, and inform the development of conservation action plans.

Appendix 1: Supplementary Material to Chapter 2

This appendix relates to Chapter 2 – A genomic framework to assist conservation breeding and translocation success: A case study of a critically endangered turtle

A1.1 Supplementary methods

A1.1.1 Genome

Sample collection

Heart tissue from a wild male Bellinger River Turtle (*Myuchelys georgesi*) medically euthanised in 1986 (UC_0152) and stored at -80°C in the University of Canberra Wildlife Tissue Collection was used for Pacific Biosciences (PacBio) HiFi sequencing for the generation of a reference genome. Heart tissue belonging to a female that required medical euthanasia in 2021 (C10031) was flash frozen at -80°C at Taronga Zoo and was used for the Arima Hi-C sequencing to scaffold the reference genome. Brain, liver, and spleen tissues from the same individual (C10031) were also collected and flash frozen at -80°C for RNA sequencing of transcriptomes for annotation of the reference genome.

Genome sequencing, assembly, and annotation

High molecular weight (HMW) DNA was extracted from the male heart tissue (UC_0152) using the Nanobind Tissue Big DNA kit (Circulomics, Pacific Biosciences, California, United States of America) following the manufacturer's protocol and quality assessed using a Qubit 2.0 Fluorometer (Thermo Fisher) and 1% agarose gel electrophoresis (90V for 30 minutes). The HMW DNA was submitted to the Australian Genome Research Facility (AGRF, Brisbane, QLD) for PacBio HiFi sequencing. DNA was sheared using the Megaruptor2 kit (Diagenode) to generate 15-20kb fragments. The BluePippin Single-molecule real-time (SMRTbell) library kit was used to select DNA fragments greater than 15kb for input to the Single-Molecule Real-Time SMRTbell express template prep kit 2.0 (PacBio). The resulting PacBio HiFi SMRTbell libraries were sequenced across two SMRT cells on the PacBio Sequel II. This resulted in 40Gb of raw data. To scaffold the genome, a whole-genome high-throughput chromosome conformation capture (Hi-C) library was prepared using heart tissue (C10031) with the Arima

Hi-C 2.0 kit according to the manufacturers protocol and sequenced as 150 bp paired end reads on an Illumina Novaseq 6000 platform at the Biomolecular Resource Facility (Canberra, Australia), and resulted in approximately 80Gb of raw Hi-C paired-end sequence data. All data is summarised in Table A1.2.1.

Bedtools v2.30.0 (Quinlan and Hall, 2010) and HiFiAdapterFilt v2.0 (Sim et al., 2022) was used to filter for Q20 HiFi reads and remove adapter sequences, respectively. Hifiasm v0.16 (Cheng et al., 2021) was used to perform *de novo* genome assembly using the filtered reads on the Nimbus cloud service provided by Pawsey Supercomputing Centre, Perth, Australia (virtual machine – 64 vCPUs; 256GB RAM; 3TB Storage). Two parameters were altered for the assembly; rounds of assembly graph cleaning (-a) was increased to 6; and the similarity threshold for duplicate haplotigs (-s) was set to 0.65. The remaining parameters were left as default. Raw Hi-C data were mapped to the PacBio HiFi assembly using BWA v0.7.17 following the Arima Hi-C mapping pipeline v02 (https://github.com/ArimaGenomics/mapping_pipeline). To generate the scaffolded genome the resulting BAM file and PacBio HiFi genome FASTA file was input into the YaHS v.1.1 scaffolding pipeline (Zhou et al., 2022) using default parameters. Juicer Tools v1.19.02 (Durand et al., 2016b) was then run to generate genome-wide contact frequency information which was visualised as a contact map using Juicebox v2.20 (Durand et al., 2016a). The contact map was manually curated by adjusting three misassembles resulting in the final 129 scaffold assembly which was output as a FASTA file.

BBmap (Bushnell, 2014) was used to generate genome assembly statistics and BUSCO (Benchmarking Universal Single-Copy Orthologs) v5.2.2 (Simão et al., 2015) to assess functional completeness on Galaxy Australia (usegalaxy.org.au) by searching the genome for complete single-copy gene orthologs against the vertebrata database (odbv10).

For repeat identification, genome annotation, and to improve transcriptome alignment rates, a *de novo* database was generated using RepeatModeler v2.0.1 (Flynn et al., 2020). The database was then used to predict, annotate, and mask repetitive elements in the genome

using RepeatMasker v4.0.9 (Smit et al., 2015). The genome was annotated using FGENESH++ v7.2.2 (Solovyev et al., 2006) on the Pawsey Supercomputing Centre's cloud machine using transcriptome data and optimised gene finding parameters from the green sea turtle (*Chelonia mydas*) supplied matrix (Softberry) using non-mammalian settings.

Transcriptome sequencing, assembly, and annotation

Total RNA (excluding micro-RNA) was extracted from brain, liver, and spleen tissue (individual C10031) using the Qiagen RNeasy Mini Kit before quantifying it using the Agilent Bioanalyzer RNA 6000 Nano kit. TruSeq Stranded mRNA-seq library preparation was undertaken on the three tissue types at the Ramaciotti Centre for Genomics (University of New South Wales, NSW, Australia), and sequenced as 100 bp paired-end reads on a NovaSeq 6000 S1 flow cell. This generated 24.9Gb of raw data (Table S1).

FastQC v0.11.8 (Andrews, 2010) was used to quality check the raw RNA-seq data and Trimmomatic v0.39 (Bolger and Giorgi, 2014) was used to trim the reads using default Trinity parameters. Illumina TruSeq sequencing adapters and reads below 25 bp were then removed from the dataset. Reads were also quality trimmed and removed if the average quality score fell below 5, both within a 4 bp sliding window and at the 5' and 3' ends of the read. 99.99% of reads were retained post-trimming.

To construct the global transcriptome, trimmed reads from the brain, liver, and spleen tissues (C10031) were aligned to the repeat masked genome using hisat2v.2.1.0 with default parameters. The resulting SAM files were sorted using SAMtools v1.9 (Li et al., 2009) to produce BAM files which we then input to StringTie v2.1.6 (Pertea et al., 2015) to assemble transcripts. Aligned reads from the three tissue types were merged into the global transcriptome using tama merge (Kuo et al., 2020) with a 5' threshold of 3 and a 3' threshold of 500. Transcripts found only in a single tissue type or that had low expression (a fragments per kilobase of transcript per million [FPKM] of < 0.1) were removed. CPC2 beta (Kang et al., 2017) was then used to predict if a transcript was a coding RNA. To identify candidate coding regions and open reading frames within the transcripts TransDecoder v2.0.1 (Haas and

Papanicolaou, 2016) was used with default parameters. To assess functional completeness of the candidate peptide sequences we used BUSCOv5.2.2 against the vertebrata database (odbv10) (Manni et al., 2021).

Assembly and annotation statistics

The final genome assembly was 2.0Gb in size, consisted of 129 scaffolds and 186 contigs, had a scaffold N50 of 123.5Mb, and a contig N50 of 56Mb (Figure A1.3.1). The largest scaffold was 311.5Mb and longest contig was 121.5Mb. BUSCOv5.2.2 identified 95.4% complete vertebrata genes, 94.9% of which were single copy and 0.5% were duplicated, 2.1% were fragmented, and 2.5% were missing (Table A1.2.2, Figure A1.3.1A). Repeat analysis identified a total of 745Mb of non-redundant repetitive sequences, accounting for 37.64% of the genome (Table S3). The Hi-C contact map suggests that the genome is chromosome-level based on the length and completeness of our scaffold. Additionally our genome has a chromosome configuration that is consistent with other turtle species (Figure A1.3.1B) (Todd et al., 2022). The global transcriptome had 95.9% complete vertebrata genes. In total, FGENESH++ annotated 24,782 non-redundant protein-coding genes. All statistics are summarised in Table S2.

Reduced representation sequencing and filtering

DNA was extracted using the Qiagen DNeasy Blood & Tissue Kit and sequenced using DArTseq™ (Diversity Arrays Technology Pty Ltd; DArT) (Kilian et al., 2012). DNA fragments were double digested using the restriction enzymes *PstI* and *SphI* sequenced on an Illumina HiSeq 2500, resulting in 77 bp or 83 bp single-end reads. Raw DArTseq reads were used as input into Stacks v.2.53 for variant calling (Catchen et al., 2013, Rochette et al., 2019). Reads were cleaned with *process_radtags* to remove uncalled bases or low-quality scores, barcodes, and truncate reads to 60 bp for consistency and to ensure removal of adaptor contamination. Adapter contamination was removed using Trimmomatic v0.39 (Bolger and Giorgi, 2014) via the *ILLUMINACLIP* parameter. Single-end reads were aligned to the reference genome generated in this study using Burrows-Wheeler aligner (BWA) v.0.7.17 'aln' function (Li and Durbin, 2009). To remove adapters from 4 bp and >4 bp barcode reads, -B 4 and -B 5 parameters of BWA were used, respectively. The BWA 'samse' function was used to produce

alignments in SAM format, before being converted to BAM format and sorted and indexed using SAMtools v.1.7. The ‘gstacks’ module was used to call SNPs across the sorted BAM files using default parameters. The ‘populations’ module was then run with the following parameters: minimum samples per population 30% (−r 0.3); minimum minor allele frequency (MAF) of 0.01 (--min_maf 0.01); and --write_random_snp.

A custom R v4.3.0 script, following Wright et al. (2019), was used to perform further filtering as follows. To extract genotypes and associated locus metadata, the Stacks VCF file was converted using the vcfR package (Knaus and Grünwald, 2017, Jombart, 2008). Filtering was carried out on three datasets: (i) all groups consisting of the four predetermined groups (*N* = 166), (ii) wild groups only (*N* = 131), and (iii) captive populations only (*N* = 35). The SNP datasets were filtered on minimum average read depth (>2.5x), coverage difference between the reference and alternate alleles (<80%), call rate (>75%), heterozygosity (<80%), and reproducibility (>90%) calculated using 51 technical replicates performed by DArT that were merged in R. The gl.report.ld() function was used in dartR to calculate pairwise population based Linkage Disequilibrium across all loci. As only 0.01% of pairwise comparisons fell above the D statistic threshold of 0.2, LD was not considered to influence the results of the analysis. The average depth of SNP alleles was 8.1x. The final reduced representation dataset included 460 genome-wide SNPs in the ‘all groups’ dataset, 473 neutral SNPs in the ‘wild’ dataset, and 227 genome-wide SNPs in the ‘captive’ dataset. The ‘all groups’ dataset used to calculate genetic diversity metrics had a final missing data value of 6.63%.

A1.2 Supplementary tables

Table A1.2.1. Data used for analyses.

Sequencing	Platform	Tissue type	Total raw data (Gb)	Year sampled	n
HiFi long reads	PacBio Sequel II	Heart	37	1986	1
Arima Hi-C scaffolding	Illumina NovaSeq 6000	Heart	80	2022	1
RNA seq short reads	Illumina NovaSeq 6000	Heart, liver, spleen	24.9	2021	3
DArT seq short reads	HiSeq 2500	Blood, skin	9.8	2007, 2015-2020	166

Table A1.2.2. Genome assembly and annotation statistics and repeat composition. (C) Complete, (S) Single-copy, (D) Duplicated, (F) Fragmented, (M) Missing.

Genome assembly quality metrics					
	HiFi Read coverage				25X
	Number of contigs				186
	Contig N50 (Mb)				56.18
	Longest contig (Mb)				121.5
	Number of scaffolds				129
	Scaffold N50 (Mb)				123.5
	Largest scaffold (Mb)				311.5
	Size of final assembly (Gb)				2.0
	GC content (%)				43.21
BUSCO completeness % (vertebrata) *N* = 3354	**C**	**S**	**D**	**F**	**M**
Genome	95.4	94.9	0.5	2.1	2.5
Global transcriptome	95.9	27.8	68.1	1.5	2.6
Annotated genes					24, 782

Table A1.2.3. Pairwise private alleles, based on the 'all groups' SNP dataset. Pop = population, *N* = sample size, fixed = fixed differences, priv1 = private alleles in pop1, priv2 = private alleles in pop2, Chao1 = private alleles in pop1 with correction for sample size, Chao2 = private alleles in pop2 with correction for sample size, Total = total number of private alleles for both populations

pop1	pop2	N1	N2	fixed	priv1	priv2	Chao1	Chao2	Total
After	Before	38	93	0	3	22	0	0	25
After	Symbio	38	19	0	46	8	0	0	54
After	Taronga	38	17	0	51	4	0	0	55
Before	Symbio	93	19	0	60	3	0	0	63
Before	Taronga	93	17	0	68	2	0	0	70
Symbio	Taronga	19	17	0	32	23	0	0	55

Table A1.2.4. Mean kinship values between 19 Bellinger River turtles housed at Symbio Wildlife Park. A value of 0 is unrelated. Pairs that we recommend are most suitable to breed have been highlighted in green (0-0.009), pairs that can be bred if no other options are available are highlighted in yellow (0.01-0.099), and pairs that should be avoided are highlighted in red (0.1-0.9). If any closely related same-sex individuals are identified as unsuitable to pair, don't breed these individuals with the same partner i.e. a pair of closely related males should not be paired with the same female.

	4805	4856	4860	4861	4866	4867	4876	4879	4880	4881	4883	4884	4886	4904	4935	4936	4939	4940
4856	0																	
4860	0	0.055																
4861	0	0	0.058															
4866	0.010	0.037	0.009	0														
4867	0	0	0.068	0.068	0.003													
4876	0.010	0.085	0.030	0	0.063	0.021												
4879	0	0.044	0.010	0.007	0	0.034	0.009											
4880	0	0.055	0.023	0.031	0	0	0	0.079										
4881	0.006	0.001	0.007	0	0.006	0	0.009	0	0.026									
4883	0	0	0	0.074	0.044	0	0	0.049	0.026	0								
4884	0	0	0	0.039	0.027	0	0	0	0	0	0.039							
4886	0	0.023	0	0.037	0	0	0	0	0.057	0	0.049	0.283						
4904	0	0.022	0	0	0	0.014	0.009	0.046	0.038	0.013	0.031	0	0.022					
4935	0	0.034	0.041	0.066	0	0.015	0.047	0.045	0.007	0.018	0.216	0.043	0	0.014				
4936	0	0.075	0	0.044	0.091	0.035	0	0	0.076	0.047	0.049	0	0	0.013	0.032			
4939	0.001	0	0	0.003	0	0	0.002	0	0.010	0	0	0.011	0.010	0.001	0	0.037		
4940	0.010	0	0.059	0	0.022	0.008	0.043	0	0	0.033	0.014	0	0	0	0.044	0	0.052	
4946	0.014	0.025	0.094	0	0.036	0.035	0	0.071	0.043	0	0	0	0	0.050	0	0	0.027	0.050

Table A1.2.5. Mean kinship values between 16 Bellinger River turtles housed at Taronga Conservation Society Australia. A value of 0 is unrelated. Pairs that we recommend are most suitable to breed have been highlighted in green (0-0.009), pairs that can be bred if no other options are available are highlighted in yellow (0.01-0.099), and pairs that should be avoided are highlighted in red (0.1-0.9). If any closely related same-sex individuals are identified as unsuitable to pair, don't breed these individuals with the same partner i.e. a pair of closely related males should not be paired with the same female.

	70200	70201	70202	70203	70204	70205	70206	70207	70208	70209	70210	70211	70212	70214	70216
70201	0.036														
70202	0.050	0.010													
70203	0.094	0.100	0												
70204	0	0	0	0											
70205	0.023	0	0.025	0	0.026										
70206	0.073	0.036	0.255	0	0	0.035									
70207	0.055	0	0.022	0.018	0	0.045	0.001								
70208	0.090	0.040	0.075	0	0.003	0	0.170	0.134							
70209	0.025	0.041	0.051	0	0	0.056	0	0.020	0						
70210	0	0.053	0.030	0.035	0	0	0.019	0.034	0	0					
70211	0.029	0	0.020	0	0.014	0.032	0	0.069	0.051	0.008	0.084				
70212	0.012	0.031	0.041	0.070	0.048	0.044	0.069	0.005	0.056	0.011	0.029	0.001			
70214	0	0.057	0	0.008	0.102	0.063	0.072	0.035	0.051	0	0	0.012	0.016		
70216	0	0	0.033	0	0.009	0	0	0.085	0	0.018	0	0.044	0.024	0	
70217	0	0	0	0.004	0.005	0	0	0.032	0.009	0.026	0	0	0	0	0

Table A1.2.6. Metadata of samples used for analyses.

Specimen Identifier	Year	Sampled by	Data Type	Group
UC_0152 (genbank UC<Aus>)	1986	University of Canberra	Genome	
C10031	2021	University of Sydney	Genome	
C10031	2021	University of Sydney	RNA	
AA036206	2007	NSW DPE	SNP	Before
AA036208	2007	NSW DPE	SNP	Before
AA036210	2007	NSW DPE	SNP	Before
AA036214	2007	NSW DPE	SNP	Before
AA036216	2007	NSW DPE	SNP	Before
AA036218	2007	NSW DPE	SNP	Before
AA036220	2007	NSW DPE	SNP	Before
AA036222	2007	NSW DPE	SNP	Before
AA036226	2007	NSW DPE	SNP	Before
AA036228	2007	NSW DPE	SNP	Before
AA036258	2007	NSW DPE	SNP	Before
AA036270	2007	NSW DPE	SNP	Before
AA036276	2007	NSW DPE	SNP	Before
AA036280	2007	NSW DPE	SNP	Before
AA036286	2007	NSW DPE	SNP	Before
AA036288	2007	NSW DPE	SNP	Before
AA036816	2007	NSW DPE	SNP	Before
AA036818	2007	NSW DPE	SNP	Before
AA036820	2007	NSW DPE	SNP	Before
AA036826	2007	NSW DPE	SNP	Before
AA036828	2007	NSW DPE	SNP	Before
AA036830	2007	NSW DPE	SNP	Before
AA036832	2007	NSW DPE	SNP	Before
AA036834	2007	NSW DPE	SNP	Before
AA036836	2007	NSW DPE	SNP	Before
AA036838	2007	NSW DPE	SNP	Before
AA036840	2007	NSW DPE	SNP	Before

AA036842	2007	NSW DPE	SNP	Before
AA036844	2007	NSW DPE	SNP	Before
AA036846	2007	NSW DPE	SNP	Before
AA036848	2007	NSW DPE	SNP	Before
AA036850	2007	NSW DPE	SNP	Before
AA036886	2007	NSW DPE	SNP	Before
AA036890	2007	NSW DPE	SNP	Before
AA036892	2007	NSW DPE	SNP	Before
AA036896	2007	NSW DPE	SNP	Before
AA036898	2007	NSW DPE	SNP	Before
AA036900	2007	NSW DPE	SNP	Before
AA048002	2007	NSW DPE	SNP	Before
AA048004	2007	NSW DPE	SNP	Before
AA048006	2007	NSW DPE	SNP	Before
AA048008	2007	NSW DPE	SNP	Before
AA048010	2007	NSW DPE	SNP	Before
AA048012	2007	NSW DPE	SNP	Before
AA048016	2007	NSW DPE	SNP	Before
AA048018	2007	NSW DPE	SNP	Before
AA048020	2007	NSW DPE	SNP	Before
AA048022	2007	NSW DPE	SNP	Before
AA048024	2007	NSW DPE	SNP	Before
AA048034	2007	NSW DPE	SNP	Before
AA048036	2007	NSW DPE	SNP	Before
AA048038	2007	NSW DPE	SNP	Before
AA048040	2007	NSW DPE	SNP	Before
AA048044	2007	NSW DPE	SNP	Before
AA048048	2007	NSW DPE	SNP	Before
AA048058	2007	NSW DPE	SNP	Before
AA048060	2007	NSW DPE	SNP	Before
AA048064	2007	NSW DPE	SNP	Before
AA048068	2007	NSW DPE	SNP	Before
AA048078	2007	NSW DPE	SNP	Before

AA048082	2007	NSW DPE	SNP	Before
AA048084	2007	NSW DPE	SNP	Before
AA048094	2007	NSW DPE	SNP	Before
AA048100	2007	NSW DPE	SNP	Before
AA048106	2007	NSW DPE	SNP	Before
AA048108	2007	NSW DPE	SNP	Before
AA048122	2007	NSW DPE	SNP	Before
AA048132	2007	NSW DPE	SNP	Before
AA048134	2007	NSW DPE	SNP	Before
AA048136	2007	NSW DPE	SNP	Before
AA064808	2007	NSW DPE	SNP	Before
AA036140	2007	NSW DPE	SNP	Before
AA036145	2007	NSW DPE	SNP	Before
AA036185	2007	NSW DPE	SNP	Before
AA036191	2007	NSW DPE	SNP	Before
AA036197	2007	NSW DPE	SNP	Before
AA036199	2007	NSW DPE	SNP	Before
AA036202	2007	NSW DPE	SNP	Before
AA036204	2007	NSW DPE	SNP	Before
AA036290	2007	NSW DPE	SNP	Before
AA036292	2007	NSW DPE	SNP	Before
AA036294	2007	NSW DPE	SNP	Before
AA036300	2007	NSW DPE	SNP	Before
AA036666	2007	NSW DPE	SNP	Before
AA036668	2007	NSW DPE	SNP	Before
AA036670	2007	NSW DPE	SNP	Before
AA036682	2007	NSW DPE	SNP	Before
AA036692	2007	NSW DPE	SNP	Before
AA036696	2007	NSW DPE	SNP	Before
AA036814	2007	NSW DPE	SNP	Before
AA064807	-	NSW DPE	SNP	Before
AA064810	-	NSW DPE	SNP	Before
AA064811	-	NSW DPE	SNP	Before

BRST_10757.2	2015	NSW DPE	SNP	After
BRST_4815	2016	NSW DPE	SNP	After
BRST_4817	2016	NSW DPE	SNP	After
BRST_4950	2016	NSW DPE	SNP	After
BRST_10758.12	2016	NSW DPE	SNP	After
BRST_10758.14	2016	NSW DPE	SNP	After
BRST_10758.6	2016	NSW DPE	SNP	After
BRST_10829.6	2016	NSW DPE	SNP	After
BRST_4914	2019	NSW DPE	SNP	After
BRST_5085	2019	NSW DPE	SNP	After
BRST_6095	2019	NSW DPE	SNP	After
BRST_6096	2019	NSW DPE	SNP	After
BRST_6143	2019	NSW DPE	SNP	After
B70146	2020	NSW DPE	SNP	After
B70148	2020	NSW DPE	SNP	After
B80239	2020	NSW DPE	SNP	After
B80240	2020	NSW DPE	SNP	After
B80241	2020	NSW DPE	SNP	After
B80242	2020	NSW DPE	SNP	After
B80243	2020	NSW DPE	SNP	After
B80245	2020	NSW DPE	SNP	After
B80249	2020	NSW DPE	SNP	After
B90118	2020	NSW DPE	SNP	After
B90123	2020	NSW DPE	SNP	After
B90130	2020	NSW DPE	SNP	After
B90131	2020	NSW DPE	SNP	After
B90132	2020	NSW DPE	SNP	After
B90133	2020	NSW DPE	SNP	After
B90134	2020	NSW DPE	SNP	After
B90135	2020	NSW DPE	SNP	After
B90136	2020	NSW DPE	SNP	After
B90138	2020	NSW DPE	SNP	After
B90139	2020	NSW DPE	SNP	After

B90140	2020	NSW DPE	SNP	After
B90141	2020	NSW DPE	SNP	After
B90142	2020	NSW DPE	SNP	After
B90143	2020	NSW DPE	SNP	After
B90127	2020	NSW DPE	SNP	After
AA070200	2015	NSW DPE	SNP	Taronga
AA070201	2015	NSW DPE	SNP	Taronga
AA070202	2015	NSW DPE	SNP	Taronga
AA070203	2015	NSW DPE	SNP	Taronga
AA070204	2015	NSW DPE	SNP	Taronga
AA070205	2015	NSW DPE	SNP	Taronga
AA070206	2015	NSW DPE	SNP	Taronga
AA070207	2015	NSW DPE	SNP	Taronga
AA070208	2015	NSW DPE	SNP	Taronga
AA070209	2015	NSW DPE	SNP	Taronga
AA070210	2015	NSW DPE	SNP	Taronga
AA070211	2015	NSW DPE	SNP	Taronga
AA070212	2015	NSW DPE	SNP	Taronga
AA070214	2015	NSW DPE	SNP	Taronga
AA070216	2015	NSW DPE	SNP	Taronga
AA070217	2015	NSW DPE	SNP	Taronga
BRST_4856_C	2016	NSW DPE	SNP	Symbio
BRST_4860_C	2016	NSW DPE	SNP	Symbio
BRST_4861_C	2016	NSW DPE	SNP	Symbio
BRST_4866_C	2016	NSW DPE	SNP	Symbio
BRST_4867_C	2016	NSW DPE	SNP	Symbio
BRST_4876_C	2016	NSW DPE	SNP	Symbio
BRST_4879_C	2016	NSW DPE	SNP	Symbio
BRST_4805_C	2016	NSW DPE	SNP	Symbio
BRST_4880_C	2016	NSW DPE	SNP	Symbio
BRST_4881_C	2016	NSW DPE	SNP	Symbio
BRST_4883_C	2016	NSW DPE	SNP	Symbio
BRST_4884_C	2016	NSW DPE	SNP	Symbio

BRST_4886_C	2016	NSW DPE	SNP	Symbio
BRST_4904_C	2016	NSW DPE	SNP	Symbio
BRST_4935_C	2016	NSW DPE	SNP	Symbio
BRST_4936_C	2016	NSW DPE	SNP	Symbio
BRST_4939_C	2016	NSW DPE	SNP	Symbio
BRST_4940_C	2016	NSW DPE	SNP	Symbio
BRST_4946_C	2016	NSW DPE	SNP	Symbio

A1.3 Supplementary figures

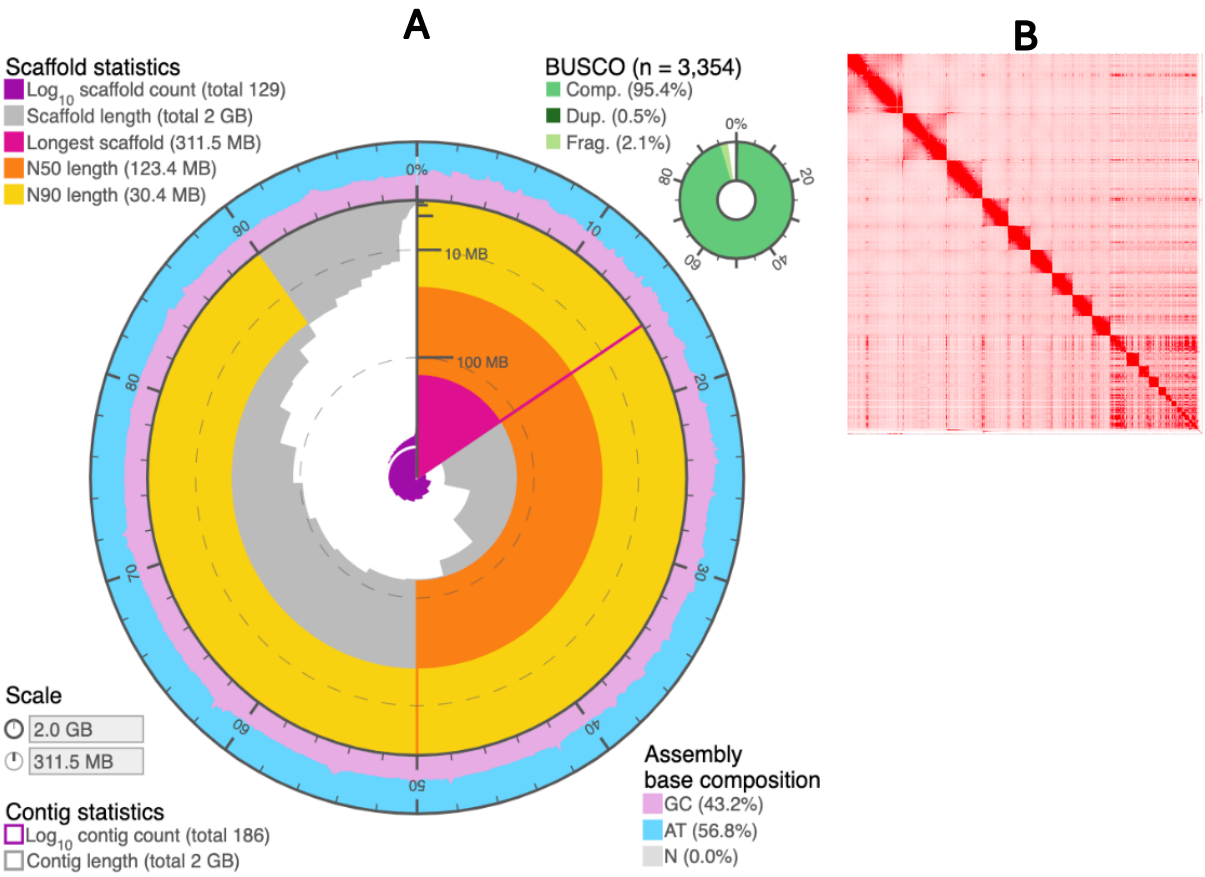


Figure A1.3.1. Visual overview of genome assembly metrics. **(A)** Circular plot generated using assembly-stats (10.5281/zenodo.594927) depicting quality metrics presented in Table A.1.2.2. The circle represents the full length of the assembly (2.0Gb). Length related metrics are represented in the middle; the dark pink line shows the longest scaffold; followed by all other scaffolds in size-order moving clockwise shown in yellow. The purple spiral displays the cumulative scaffold count. The N50 and N90 values are represented in dark and light orange, respectively. The exterior pink and blue circles show mean, maximum, and minimum GC vs. AT levels. BUSCOv5.3.2 (vertebrata_odb10) scores are shown in the top right corner including complete (single and duplicated) (Comp.), duplicated (Dup.), and Fragmented (Frag.). **(B)** Hi-C contact frequency matrices after manual curation with Juicer (Durand et al., 2016b). Scaffolds were sorted according to their length. The map interprets the physical contact of genomic regions in 3D space with each cell corresponding to sequence data that supports the association between two regions.

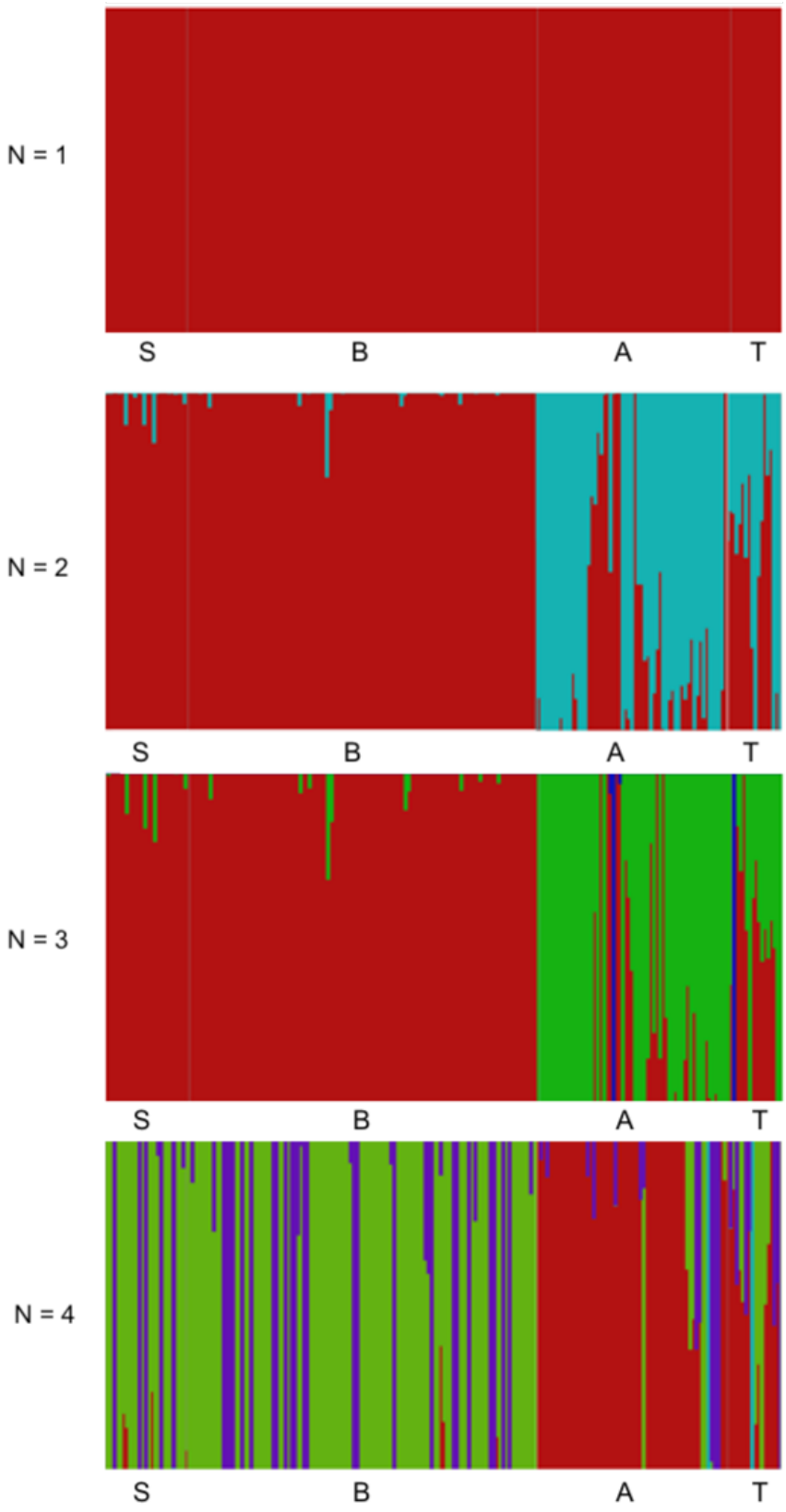


Figure A1.3.2. fast*STRUCTURE* plots investigating K=1-4 genetic clusters (*N* = 166; SNPs=460). (S) Symbio (B) Before (A) After (T) Taronga.

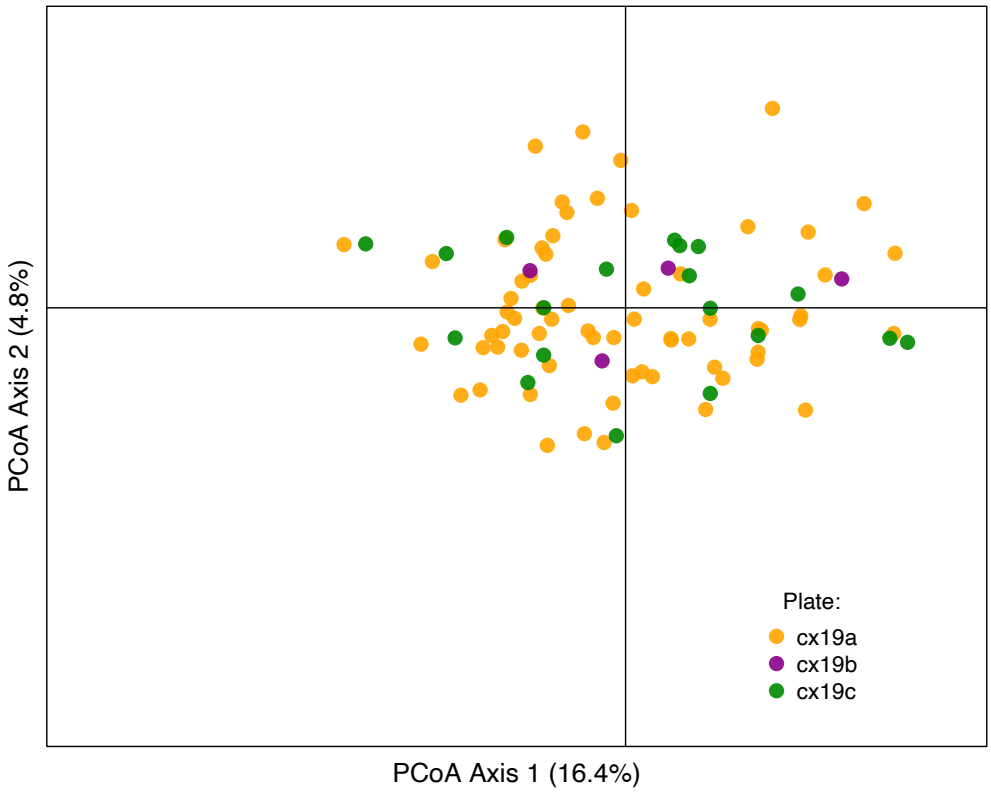


Figure A1.3.3. PCoA analyses investigating batch effects of samples grouped across different sequencing plates.

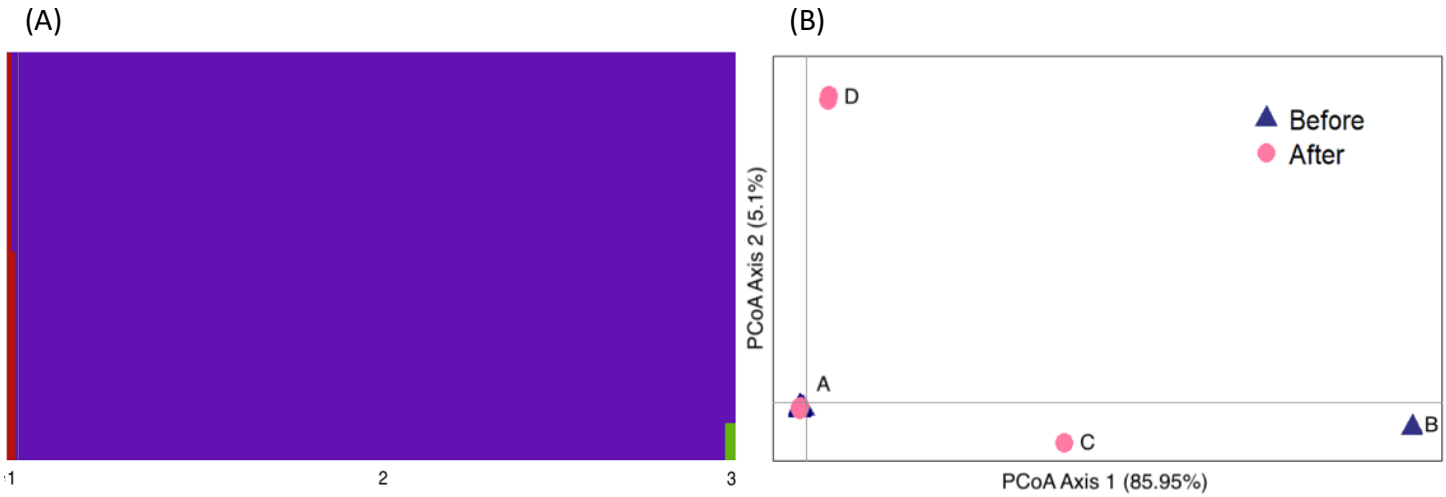


Figure A1.3.4. (A) Preliminary fast*STRUCTURE* plot (K=3) showing evidence of (1) an *Emydura macquarii*, and a *Myuchelys georgesi* and *E. macquarii* hybrid, (2) *M. georgesi*, and (3) backcross animals. **(B)** Preliminary PCoA analyses of Before and After individuals (A; *N* = 130). Outlier samples were identified as *E. macquarii* (B; *N* = 1), *E. macquarii* and *M. georgesi* hybrid (C; *N* = 1), and backcross animals (D; *N* = 2).

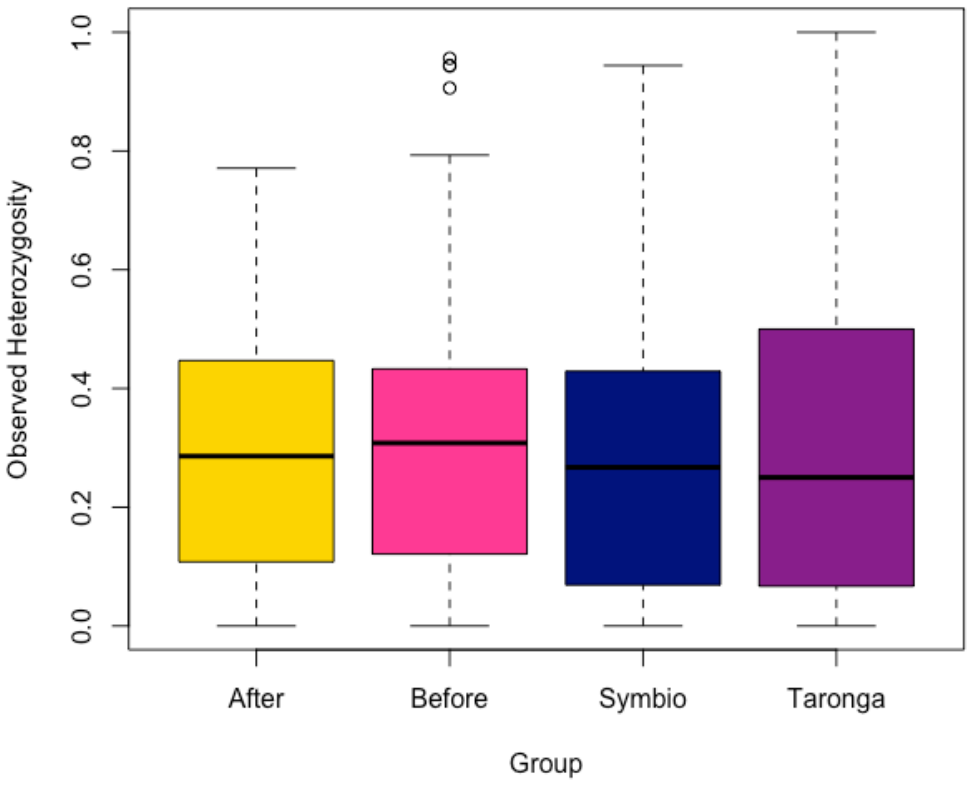


Figure A1.3.5. Distributions of SNP-based observed heterozygosity (H_o) by group.

Appendix 2: Supplementary Material to Chapter 3

This appendix relates to Chapter 3 – A genomic-based workflow for eDNA assay development for threatened species

A2.1 Supplementary methods

A2.1.1 Mitogenome

9 genes are encoded on the light strand including tRNA-Gln, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr, tRNA-Ser, ND6, tRNA-Glu, and tRNA-pro. The remaining 28 genes are encoded on the heavy strand. A total of 7 overlapping regions with a total length of 15 bp were identified in *M. georgesi,* with the longest region of overlap (5 bp) observed between NADH dehydrogenase subunit 4L (nad4l) and NADH dehydrogenase subunit 4 (nad4). 21 intergenic spacer regions with a total length of 1701 bp were observed. The longest region (1241 bp) was observed between NADH dehydrogenase subunit 5 (nad5) and NADH dehydrogenase subunit 6 (nad6) (Table A2.2.4). The total length on PCGs was 12901 bp, representing 78.1% of the complete mitogenome. The ATG start codon initiated the majority of the PCGs, except for ATP synthase F0 subunit 6 (*atp6*), cytochrome c oxidase subunit III (*cox3*), NADH dehydrogenase subunit 4L (*nad4l*), and Cytochrome b (*cytb*) that were initiated by ATA.

A2.2 Supplementary tables

Table A2.2.1. Final CO1 and CytB primer statistics

Gene	Primer	Start	Length	Tm	GC%	Molecular weight	DeltaG
COI	*Forward* acattggcaccctctacctg	41	20 bp	60	55	6012.9 g/mole	-3.14 kcal/mol
	Reverse aattaaggcgtgggctgtaa	192	20 bp	59.6	45	6221.1 g/mole	5.36 kcal/mol
CytB	*Forward* aatctcccacatccaacgag	186	20 bp	59.9	50	6015 g/mole	-3.61 k/mol
	Reverse atgcggtggctatgactagg	373	20 bp	60.1	55	6213.1 g/mole	-4.16 k/mol

Table A2.2.2. *In situ* validation samples

Source	Size	Species	N	Samples
Tank 1	4000L	*Myuchelys georgesi*	5	2 x 500 mL (1L)
Tank 2	4000L	*Myuchelys georgesi*	5	2 x 500 mL (1L)
Tank 3	4000L	*Myuchelys georgesi*	5	2 x 500 mL (1L)
Pond (control)	2000L	*Emydura macquarii*	4	2 x 500 mL (1L)
		Chelodina longicollis	2	
Pond (control)	2000L	*Emydura macquarii*	4	2 x 500 mL (1L)
		Chelodina longicollis	2	

Table A2.2.3. *In vitro* validation tissue specimens

Specimen ID	Species	Sample ID	Date	Type	Sample Location
UC<Aus>AA063726	*Myuchelys georgesi*	AA063727	2015	Heart	Bellinger River
UC<Aus> AA060000	*Emydura macquarii*	AA060000	2015	Skin	Bellinger River

Table A2.2.4. List of annotated mitochondrial genes of *Myuchelys georgesi*

Gene	*Direction*	*Location*	*Size*	*Anti-codon*	*Start codon*	*Stop codon*	*Intergenic nucleotides*
trnF	+	1-67	67	GAA			0
rrnS	+	68-1038	971				0
trnV	+	1039-1108	70	TAC			0
rrnL	+	1109-2720	1612				2
trnL2	+	2723-2797	75	TAA			4
nad1	+	2802-3764	963		ATG	TAG	-1
trnI	+	3764-3836	73	GAT			-1
trnQ	-	3836-3906	71	TTG			-1
trnM	+	3906-3974	69	CAT			0
nad2	+	3975-5018	1044		ATG	TAA	1
trnW	+	5020-5091	72	TCA			2
trnA	-	5094-5162	69	TGC			1
trnN	-	5164-5237	74	GTT			19
trnC	-	5257-5324	68	GCA			1
trnY	-	5326-5395	70	GTA			1
cox1	+	5397-6935	1539		ATG	TAA	2
trnS2	-	6938-7009	72	TGA			12
trnD	+	7023-7089	67	GTC			0
cox2	+	7090-7776	687		ATG	TAG	2
trnK	+	7769-7842	74	TTT			1
atp8	+	7844-7981	138		ATG	AGA	89
atp6	+	8071-8682	612		ATA	TAG	-2
cox3	+	8679-9455	777		ATA	TGA	7
trnG	+	9463-9531	69	TCC			175
nad3	+	9707-9874	168		GTA	TGA	7
trnR	+	9882-9949	68	TCG			0
nad4l	+	9950-10246	297		ATA	TAA	-5
nad4	+	10240-11625	1386		ATG	TAA	-4
trnH	+	11621-11690	70	GTG			0
trnS1	+	11691-11753	63	GCT			-1
trnL1	+	11753-11824	72	TAG			0
nad5	+	11825-12946	1122		ATG	ACC	1241
nad6	-	14188-13658	529		ATG	TTA	1
trnE	-	14190-14258	69	TTC			58
cytB	+	14317-15372	1056		ATA	T(CC)	74
trnT	+	15447-15516	70	TGT			1
trnP	-	15518-15585	68	TGG			

A2.3 Supplementary figures

Figure A2.3.1. Complete nucleotide sequence of Cytochrome Oxidase 1 gene output by MitoHiFi and used as input into Primer3Plus. Forward and reverse primer pairs are highlighted in grey.

>CO1

tattaaaccgctgactattttctactaaccataaagacattggcaccctctacctgatttttggggcctgagcaggaataatcggaaca
gctcttagtctattaattcgaacagaactaagccaaccaggtcccctattaggagatgatcaagtatacaatgtaattgttacagccca
cgccttaattataatctttttcatagttatacctattataattggtggatttggaaactgattagttccaataataattggatcgccagatat
agcatttccacgaatgaataatataagcttttgacttctaccaccatcactactactacttctagcctcctctggtatcgaagccggagct
ggaacagggtgaactgtatatccccctttggctggaaatatagcccacgccggagcttctgtcgacctaactatcttctccctacactta
gccggggcgtcctcaattttaggggccatcaacttcatcaccaccgcaattaatataaaaaccccatcaatatcacaataccaaacac
cacttttcgtatgatcagtacttatcacagctgtattactattactctcccttccagtacttgctgcaggcattactatacttttaacagacc
gaaacctaaatacaaccttctttgatccatctggtggaggagacccaatcctataccaacacctattttgattctttggccaccccgaag
tatacattctcatccttcccggcttcggaataatttctcacgtcgttgcttattatactggtaaaaaagaaccattcgggtacatgggaat
agtctgagcaataatatcaatcggattcttaggattcatcgtttgagctcaccatatattcacagtaggaatagacgtagacacccgag
cctactttacatcagcaacaataatcattgctatccccacaggtgtcaaagtatttagctgactagccaccctccacggaggaataatt
aaatgagatgcccctatactatgagctctgggatttatttttctatttaccattggaggactaacaggtattgtattagctaactcatcact
agatatcgtactacacgatacatactatgtagtagcacacttccattatgtactatcaatgggggccgtatttgccattatagccggattt
acccactgattcccattatttacaggattctcactaaaccaaacatgggcaaaactacaattcgtagtaatattctttggcgtaaacata
acattcttcccacaacacttcctaggtttagctggtataccccgacgatactcagattatccagacgcctacacaatatgaaactccatc
tcatcaattggatcaataatctccatagcagcagtaatcataatactagtcattatctgagaggccttctcatcaaagcgaaaaatagt
actgattgaaccacccctaatcaacgtagaatgactaagcggttgtccgccatccagtcacacctatgaagaatccgcacatatattat
aa

Figure A2.3.2. Complete nucleotide sequence of Cytochrome B gene output by MitoHiFi and used as input into Primer3Plus. Forward and reverse primer pairs are highlighted in grey.

>CytB

ataaaaaacacaaatccactattaaaaattattaacaacaccttcatcgatctccccacccctccaacatctccgccctatgaaactt
cggatcactactaggaatatgcctaattctacaactagccacaggaatcttcctagctatacactactcgcccgatatctccatagcatt
ctcatcaatctcccacatccaacgagacgtccaatatggttgactgattcgaaacatacacgccaacggtgcttcactatttttcatatg
catttacctccatattggacgtggaatctactacggttcctacctctacaataaaacctgaaacactggagtaattctactattcctagtc
atagccaccgcattcgtgggctacgtactaccatgaggccaaatatcattctgaggggctacagtaatcaccaaccttctatcagccat
cccatatgcaggccctacacttgtagaatgaatctgaggaggattctccgtagacaacgccaccctgactcgattcttcacattccactt
tctaatcccattcgctatcctaggaataaccatactacacctcctactactacacgaaacaggatcaaacaacccaacaggattaaac
tcaaactgcgacaagatcccattccacccatatttttcttacaaagacctactaggcctcatcctaataatcatatgcctgctcaccctca
ccctattttacccaaacctactaggagatccagacaacttcacaccagccaacccactaaccaccccaccacacatcaaaccagagt
gatacttcctattcgcttacacaatcctacgttcaatccccaacaaactaggtggcgtcctagccctattcatatcaatcctagtactact
aaccataccaatactccacctatcaaaacaacgaacaaccacattccgaccaatagcacaaatcctcttctgatgcttaaccactgac
ctactaatcctaacatgaatcggaggccaaccagtagaagacccattcatcctcattggacaaatcgcctccctactatacttcaccat
catcttcatcatcacacctataacaagcttaattgaaaacaaaatactaaaccaat

Appendix 3: Supplementary Material to Chapter 4

This appendix relates to Chapter 4 – A genomic-based workflow for eDNA assay development for threatened species

A3.1 Supplementary tables

Table A3.1.1. Reference species and sequences for Major Histocompatibility Complex (MHC) BLAST search in *Myuchelys georgesi*. Complete coding sequences (complete) were used when available. Partial coding sequences (partial) were also included. The corresponding MHC class and shared common ancestor group are also listed.

Class	Reference MHC species	Project	Sequence type	Shared taxonomic group
MHC I	Tawny dragon (*Ctenophorus decresii*)	KY905241.1	Complete	Class Reptilia
MHC I	Caiman (*Caiman crocodilus*)	KF769542.1	Complete	Class Reptilia
MHC I	Marine iguana (*Amblyrhynchus cristatus*)	EU839663.1	Complete	Class Reptilia
MHC I	Galapagos land iguana (*Conolophus subcristatus*)	EU604313.1	Complete	Class Reptilia
MHC I	Tuatara (*Sphenodon punctatus*)	DQ145788.1	Complete	Class Reptilia
MHC I	Green sea turtle (*Chelonia mydas*)	OK135213.1	Partial	Order testudines
MHC II	Chinese softshell turtle (*Pelodiscus sinensis*)	MT834970.1	Complete	Order testudines
MHC II	Marine iguana (*Amblyrhynchus cristatus*)	FJ623752.1	Complete	Order Reptilia
MHC II	Green-rumped parrotlet (*Forpus passerines*)	EF710746.1	Partial	Sauropsida

Table A3.1.2. Genbank Accession numbers of complete MHC I sequences used for phylogenetic comparisons.

Accession	Species	Common name	Classification	Name on tree
L20733.1	*Xenopus laevis*	African clawed frog	Amphibia	*Xela-U*
HQ158307.1	*Crocodylus porosus*	Saltwater crocodile	Crocodilia	*Crpo-U*
KF769542.1	*Caiman crocodilus*	Caiman	Crocodilia	*Cacr-U*
JU175072.1	*Crotalus adamanteus*	Eastern Diamondback rattlesnake	Squamata	*Crad-U*
KM515952.1	*Tiliqua rugosa*	Shingleback	Squamata	*Tiru-UB*
KY905241.1	*Ctenosaura decressii*	Tawny dragon	Squamata	*Ctde-UA*
EU604313.1	*Conolophus subcristatus*	Galapagos iguana	Squamata	*Cosu-UB*
EU604308.1	*Amblyrhynchus cristatus*	Marine iguana	Squamata	*Amcr-UB*
DQ145788.1	*Sphenodon punctatus*	Tuatara	Rhynchocephalia	*Sppu-U*01*
DQ145789.1	*Sphenodon punctatus*	Tuatara	Rhynchocephalia	*Sppu-U*02*
X12780.1	*Gallus gallus*	Chicken	Aves	*Gaga_F10*
MT260395.1	*Anas platyrhynchos*	Mallard	Aves	*Anpl-UAA*
MN339476.1	*Ginglymostoma cirratum*	Nurse shark	Orectolobiformes	*Gici-UDA*

Table A3.1.3. Genbank Accession numbers of complete MHC II beta sequences used for phylogenetic comparisons

Accession	Species	Common name	Classification	Name on tree
KP118841.1	*Crocodylus porosus*	Saltwater crocodile	Crocodilia	*Crpo-DAB1*
KP118846.1	*Crocodylus porosus*	Saltwater crocodile	Crocodilia	*Crpo-DAB2*
AF256652.1	*Caiman crocodilus*	Caiman	Crocodilia	*Cacr-beta*
XM_008121064.1	*Anolis carolinensis*	Green anole	Squamata	*Anca-beta*
FJ623746.1	*Amblyrhynchus cristatus*	Marine iguana	Squamata	*Amcr-DAB1*
FJ623750.1	*Amblyrhynchus cristatus*	Marine iguana	Squamata	*Amcr-DAB3*
DQ124231.1	*Sphenodon punctatus*	Tuatara	Rhynchocephalia	*Sppu-DAB1*
DQ124232.1_	*Sphenodon punctatus*	Tuatara	Rhynchocephalia	*Sppu-DAB2*
DQ124233.1	*Sphenodon punctatus*	Tuatara	Rhynchocephalia	*Sppu-DBB*
NM_001312902.2	*Gallus gallus*	Chicken	Aves	*Gaga-DMB1*
KJ162461.1	*Columba livia*	Rock dove	Aves	*Coli-DAB*
EU442606.2	*Tyto alba*	Barn owl	Aves	*Tyal-DAB1*
MN364682.1	*Alosa sapidissima*	American Shad	Clupeiformes	*Alsa-beta*

Table A3.1.4. Metadata of samples used for analyses. MG *Myuchelys georgesi*, EM *Emydura macquarii*

Specimen ID	Sample ID	Sample type	Species	Year	Group	Sequencing
UC<Aus>AA036038	AA036038	skin	MG	2007	Before	WGR
UC<Aus>AA036098	AA036098	skin	MG	2007	Before	WGR
UC<Aus>AA036219	AA036219	skin	MG	2007	Before	WGR
UC<Aus>AA036245	AA036245	skin	MG	2007	Before	WGR
UC<Aus>AA036250	AA036250	skin	MG	2007	Before	WGR
UC<Aus>AA036253	AA036253	skin	MG	2007	Before	WGR
UC<Aus>AA036275	AA036275	skin	MG	2007	Before	WGR
UC<Aus>AA036814	AA036814	skin	MG	2007	Before	WGR
UC<Aus>AA036815	AA036815	skin	MG	2007	Before	WGR
UC<Aus>AA036831	AA036831	skin	MG	2007	Before	WGR
UC<Aus>AA036869	AA036869	skin	MG	2007	Before	WGR
UC<Aus>AA036889	AA036889	skin	MG	2007	Before	WGR
UC<Aus>AA048006	AA048006	skin	MG	2007	Before	WGR
UC<Aus>AA048016	AA048016	skin	MG	2007	Before	WGR
UC<Aus>AA048033	AA048033	skin	MG	2007	Before	WGR
UC<Aus>AA048041	AA048041	skin	MG	2007	Before	WGR
UC<Aus>AA048057	AA048057	skin	MG	2007	Before	WGR
UC<Aus>AA048084	AA048084	skin	MG	2007	Before	WGR
UC<Aus>AA048177	AA048177	skin	MG	2007	Before	WGR
UC<Aus>BRST_5085	BRST_5085_04	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_5246	BRST_5246_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_5247	BRST_5247_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_5248	BRST_5248_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_5252	BRST_5252_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_6013	BRST_6013_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_6017	BRST_6017_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_6023	BRST_6023_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_6024	BRST_6024_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_6093	BRST_6093_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_6096	BRST_6096_01	Blood stain	MG	2019	After	WGR
UC<Aus>BRST_5251	BRST_5251_02	Blood stain	MG x EM	2019	Backcross	WGR
UC<Aus>BRST_6029	BRST_6029_01	Blood stain	MG x EM	2019	Backcross	WGR
UC<Aus>BRST_6115	BRST_6115_01	Blood stain	MG x EM	2019	Backcross	WGR
UC<Aus>BRST_6099	BRST_6099_01	Blood stain	MG x EM	2019	Backcross	WGR

Table A3.1.5. Complete list of annotated MHC exons in order of genomic position along chromosome 10. ORF Open reading frame

Class	Gene	ORF length (bp)	Exon total	Exon	Start	End	Strand
MHC I	*Myge-UB*	1124	8	1	30471574	30471666	-
				2	30473115	30473143	-
				3	30473497	30473530	-
				4	30474174	30474264	-
				5	30474404	30474680	-
				6	30475552	30475826	-
				7	30488887	30489157	-
				8	30493423	30493484	-
	Myge-UE	1196	7	1	30500310	30500506	-
				2	30500690	30500713	-
				3	30501301	30501391	-
				4	30501531	30501807	-
				5	30502340	30502614	-
				6	30503854	30504130	-
				7	30508464	30508525	-
	Myge-UC	1096	7	1	30533405	30533478	-
				2	30533917	30533949	-
				3	30534318	30534434	-
				4	30534837	30535115	-
				5	30535937	30536212	-
				6	30537832	30538093	-
				7	30538879	30538940	-
	Myge-UD	1063	7	1	30567953	30567993	-
				2	30568557	30568589	-
				3	30568935	30569051	-
				4	30569521	30569799	-
				5	30570806	30571081	-
				6	30572447	30572708	-
				7	30577604	30577665	-
	Myge-UA	1063	7	1	30583497	30583537	-
				2	30584108	30584140	-
				3	30584486	30584602	-
				4	30585060	30585337	-
				5	30587357	30587633	-
				6	30588820	30589081	-
				7	30591307	30591368	-
MHC II	*Myge-DAA1*	774	4	1	30660982	30661136	-
				2	30661272	30661554	-
				3	30662018	30662275	-
				4	30663406	30663487	-
	Myge-DAB1	805	6	1	30665199	30665305	+
				2	30665695	30665968	+
				3	30666785	30667061	+
				4	30668402	30668513	+

			5	30669052	30669077	+
			6	30669438	30669452	+
Myge-DAA2	774	4	1	30677698	30677852	-
			2	30677989	30678271	-
			3	30678664	30678921	-
			4	30680142	30680223	-
Myge-DAB2	807	6	1	30681654	30681760	+
			2	30682482	30682752	+
			3	30683567	30683849	+
			4	30685180	30685291	+
			5	30685795	30685819	+
			6	30686179	30686193	+
Myge-DAA3	774	4	1	30698489	30698644	-
			2	30698780	30699061	-
			3	30699528	30699785	-
			4	30700907	30700988	-
Myge-DAB3	804	6	1	30702397	30702503	+
			2	30702901	30703174	+
			3	30704096	30704372	+
			4	30705339	30705450	+
			5	30706005	30706029	+
			6	30706334	30706348	+
Myge-DAA4	775	4	1	30717695	30717850	-
			2	30717983	30718265	-
			3	30718651	30718908	-
			4	30720110	30720191	-
Myge-DAB4	807	6	1	30721446	30721552	+
			2	30721861	30722131	+
			3	30722953	30723235	+
			4	30724562	30724673	+
			5	30725144	30725168	+
			6	30725528	30725542	+
Myge-DAA5	751	4	1	30735202	30735357	-
			2	30735492	30735774	-
			3	30736241	30736498	-
			4	30737725	30737782	-
Myge-DAB5	777	5	1	30739700	30739806	+
			2	30740419	30740689	+
			3	30741306	30741582	+
			4	30742814	30742925	+
			5	30743771	30743785	+

Table A3.1.6. Differentiation statistics between **(A)** MHC I and **(B)** MHC II *Myge* genes generated using EMBL-EBI Clustal Omega.

(A)

	UA	***UB***	***UC***	***UD***	***UE***
UA	0				
UB	0.523	0			
UC	0.115	0.474	0		
UD	0.134	0.541	0.120	0	
UE	0.506	0.187	0.495	0.541	0

(B)

	DAA1	***DAB1***	***DAA2***	***DAB2***	***DAA3***	***DAB3***	***DAA4***	***DAB4***	***DAA5***	***DAB5***
DAA1	0									
DAB1	0.421	0								
DAA2	0.088	0.448	0							
DAB2	0.503	0.093	0.48	0						
DAA3	0.046	0.481	0.081	0.510	0					
DAB3	0.494	0.083	0.499	0.069	0.493	0				
DAA4	0.090	0.571	0.092	0.505	0.076	0.591	0			
DAB4	0.474	0.118	0.478	0.103	0.497	0.112	0.504	0		
DAA5	0.060	0.450	0.077	0.507	0.064	0.493	0.094	0.502	0	
DAB5	0.467	0.104	0.492	0.080	0.559	0.101	0.521	0.095	0.485	0

Table A3.1.7. Summary of runs of homozygosity (ROH) metrics for individuals in Before and After groups. Each row represents a different individual with details on total ROH length; percentage of genomic segments in ROH; F_{ROH} (proportion of the genome in ROH); Short F_{ROH} (<2mb); and long F_{ROH} (>2mb). π genome-wide nucleotide diversity including ROH.

Individual	ID	Group	Segments in ROH (%)	Total ROH Length	F_{ROH}	F_{ROH} < 2Mb	F_{ROH} > 2Mb	π
1	AA036038	Before	91.44	1245800000	0.881166358	0.112745158	0.768421201	8.00068e-05
2	AA036098	Before	90.54	1235700000	0.874022531	0.129296203	0.744726327	8.00148e-05
3	AA036219	Before	90.39	1232700000	0.871900602	0.100579432	0.77132117	8.00064e-05
4	AA036245	Before	90.56	1235600000	0.8739518	0.129084011	0.744867789	8.00042e-05
5	AA036250	Before	92.43	1266200000	0.895595475	0.077167482	0.818427993	8.00008e-05
6	AA036253	Before	91.58	1247900000	0.882651709	0.096264843	0.786386866	8.00165e-05
7	AA036275	Before	90.16	1231200000	0.870839637	0.104540366	0.766299272	8.00015e-05
8	AA036814	Before	91.89	1257700000	0.889583343	0.10192332	0.787660023	8.00000e-05
9	AA036815	Before	90.66	1228800000	0.869142094	0.120525564	0.74861653	8.00026e-05
10	AA036831	Before	89.76	1216500000	0.860442186	0.12901328	0.731428906	8.00014e-05
11	AA036869	Before	92.13	1258700000	0.890290653	0.099023351	0.791267302	8.0002e-05
12	AA036889	Before	90.13	1227700000	0.868364054	0.112391503	0.755972551	8.00006e-05
13	AA036889	Before	90.98	1240400000	0.877346886	0.125971848	0.751375038	8.00028e-05
14	AA048016	Before	90.87	1245700000	0.881095627	0.115645127	0.7654505	8.00083e-05
15	AA048033	Before	91.58	1255800000	0.888239455	0.094496569	0.793742886	8.00001e-05
16	AA048041	Before	91.13	1245700000	0.881095627	0.105389137	0.77570649	8.00001e-05

17	AA048057	Before	91.61	1249900000	0.884066328	0.105106214	0.778960114	8.00029e-05
18	AA048084	Before	90.89	1238100000	0.875720074	0.10927934	0.766440734	8.00019e-05
19	AA048177	Before	91.11	1245800000	0.881166358	0.104469635	0.776696723	8.00005e-05
20	MG6143	After	90.34	1227600000	0.868293323	0.119676792	0.74861653	8.00076e-05
21	MG5085	After	91.55	1248500000	0.883076094	0.105035483	0.778040612	8.00014e-05
22	MG5246	After	90.42	1235400000	0.873810338	0.107723259	0.766087079	8.00017e-05
23	MG5247	After	92.02	1258900000	0.890432115	0.103479401	0.786952713	8.00063e-05
24	MG5248	After	91.01	1238300000	0.875861536	0.0945673	0.781294236	8.00029e-05
25	MG5252	After	90.58	1239000000	0.876356653	0.099589198	0.776767454	8.00075e-05
26	MG6013	After	90.69	1233600000	0.872537181	0.128023046	0.744514134	8.00034e-05
27	MG6017	After	90.33	1232900000	0.872042064	0.108359838	0.763682226	8.00026e-05
28	MG6023	After	90.39	1230700000	0.870485983	0.125476731	0.745009251	8.00114e-05
29	MG6024	After	91.35	1254900000	0.887602876	0.10121601	0.786386866	8.00009e-05
30	MG6093	After	91.24	1244700000	0.880388318	0.106450102	0.773938216	8.00017e-05
31	MG6095	After	91.21	1247900000	0.882651709	0.10920861	0.773443099	8.00002e-05
32	MG5251	Backcross	0.057	700000	0.000495117	0.000495117	0	0.00484852
33	MG6029	Backcross	0.021	200000	0.000141462	0.000141462	0	0.00485748
34	MG6096	Backcross	0.071	800000	0.000565848	0.000565848	0	0.00442102
35	MG6099	Backcross	0.035	400000	0.000282924	0.000282924	0	0.00483882

4

Table A3.1.8. Observed and expected heterozygosity for MHC Class I and II in the Before, After, and Backcross groups. "All" refers to statistics calculated across all exons for the respective class. SE indicates standard error, and CI refers to the 95% confidence intervals (lower and upper bounds).

Class	MHC Class	Before H_O (± SE)	Before H_E (± SE)	After H_O (± SE)	After H_E (± SE)	Backcross H_O (± SE)	Backcross H_E (± SE)
MHC I	*Myge-UA*	0.177 (0.037)	0.211 (0.003)	0.169 (0.051)	0.210 (0.000)	0.634 (0.042)	0.460 (0.000)
	Myge-UB	0.014 (0.004)	0.058 (0.000)	0.008 (0.005)	0.058 (0.000)	0.445 (0.117)	0.378 (0.000)
	Myge-UC	0.047 (0.015)	0.024 (0.012)	0.024 (0.012)	0.018 (0.000)	0.477 (0.072)	0.404 (0.000)
	Myge-UD	0.039 (0.009)	0.109 (0.000)	0.037 (0.011)	0.039 (0.008)	0.699 (0.029)	0.467 (0.000)
	Myge-UE	0.218 (0.049)	0.249 (0.000)	0.211 (0.062)	0.249 (0.000)	0.658 (0.063)	0.451 (0.000)
	All	0.118 (0.018)	0.148 (0.001)	0.100 (0.022)	0.147 (0.001)	0.471 (0.009)	0.395 (0.001)
MHC II	*Myge-DAA*	0.014 (0.003)	0.056 (0.000)	0.029 (0.015)	0.056 (0.000)	0.445 (0.060)	0.373 (0.000)
	Myge-DAB	0.014 (0.002)	0.056 (0.000)	0.029 (0.015)	0.056 (0.000)	0.449 (0.221)	0.362 (0.00)
	Myge-DAA	0.007 (0.002)	0.050 (0.000)	0.006 (0.001)	0.050 (0.000)	0.425 (0.253)	0.380 (0.000)
	Myge-DAB	0.214 (0.043)	0.229 (0.000)	0.231 (0.054)	0.229 (0.000)	0.533 (0.033)	0.420 (0.000)
	Myge-DAA	0.009 (0.002)	0.076 (0.002)	0.007 (0.002)	0.075 (0.002)	0.520 (0.086)	0.411 (0.000)
	Myge-DAB	0	0.069 (0.000)	0	0.068 (0.000)	0.525 (0.131)	0.410 (0.001)
	Myge-DAA	0.135 (0.039)	0.192 (0.002)	0.123 (0.039)	0.194 (0.002)	0.557 (0.088)	0.423 (0.000)
	Myge-DAB	0.132 (0.031)	0.184 (0.001)	0.120 (0.035)	0.181 (0.003)	0.478 (0.146)	0.389 (0.000)
	Myge-DAA	0.227 (0.033)	0.196 (0.000)	0.176 (0.042)	0.196 (0.000)	0.538 (0.037)	0.422 (0.000)
	Myge-DAB	0.137 (0.0307)	0.163 (0.000)	0.127 (0.038)	0.163 (0.000)	0.438 (0.123)	0.378 (0.000)
	All	0.122 (0.019)	0.148 (0.001)	0.102 (0.023)	0.147 (0.001)	0.474 (0.008)	0.385 (0.001)

A3.2 Supplementary figures

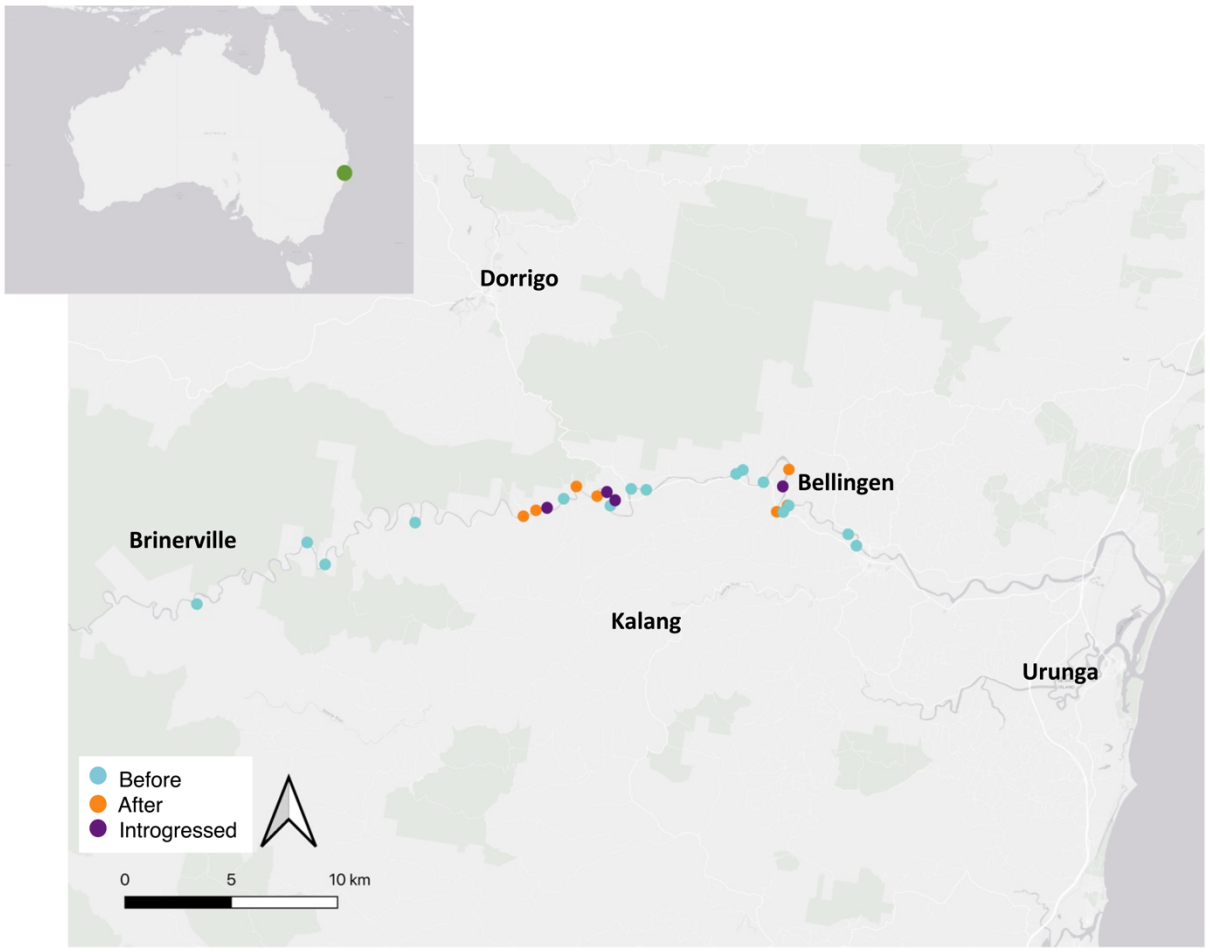


Figure A3.2.1. An inset of Australia with a map of the Bellinger River basin showing the locations of historic (Before: 2007, blue), contemporary (After: 2015-2020, orange) and Backcross (2015-2020, purple) samples collected across the species' range (NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW) (unpublished data).

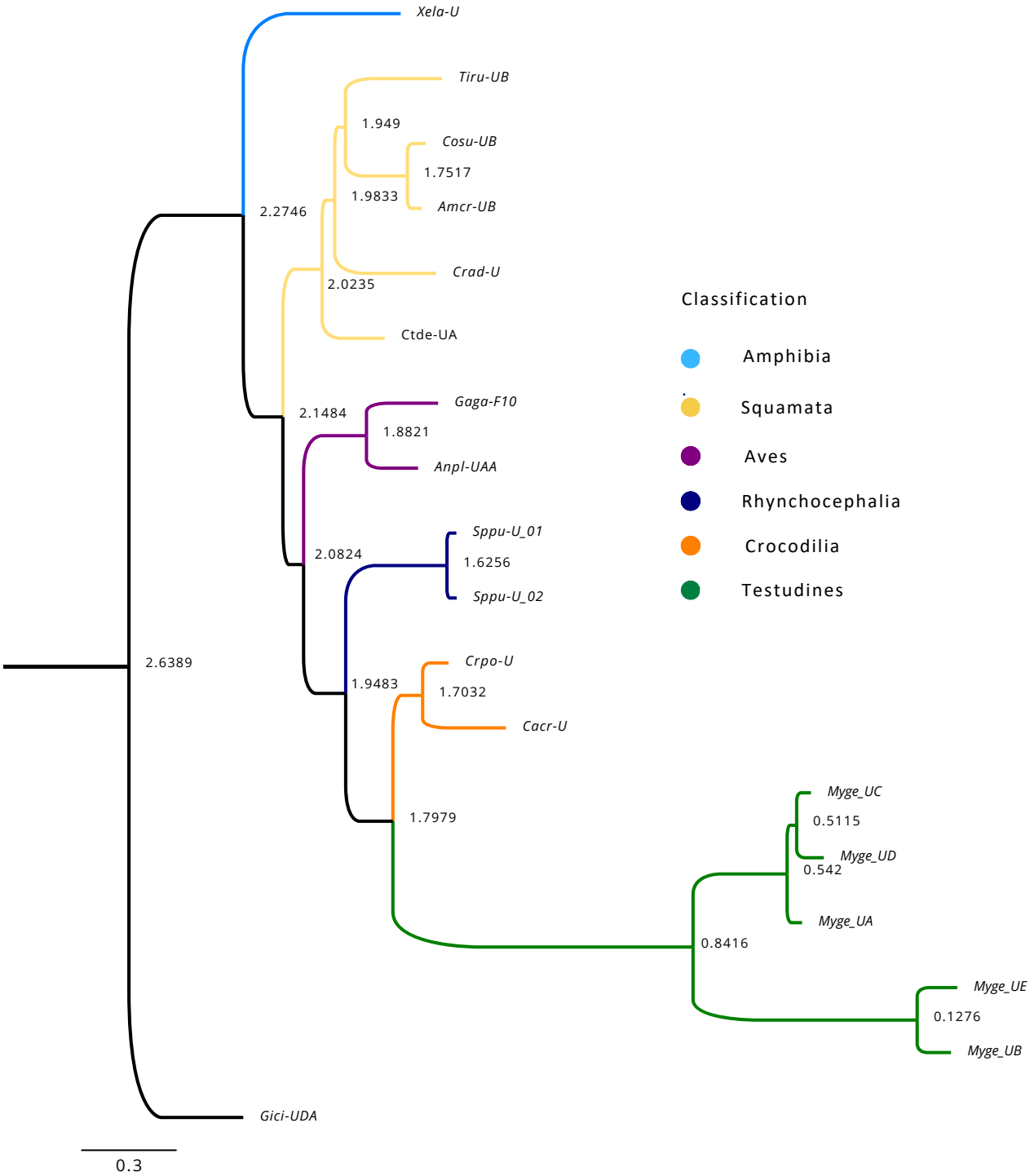


Figure A3.2.2. Maximum likelihood analysis of MHC I sequences for non-avian and avian reptiles. Node ages based on 1000 replicates are shown. Nurse shark (*Ginglymostoma cirratum*) (*Gici-UDA*) was used as an outgroup.

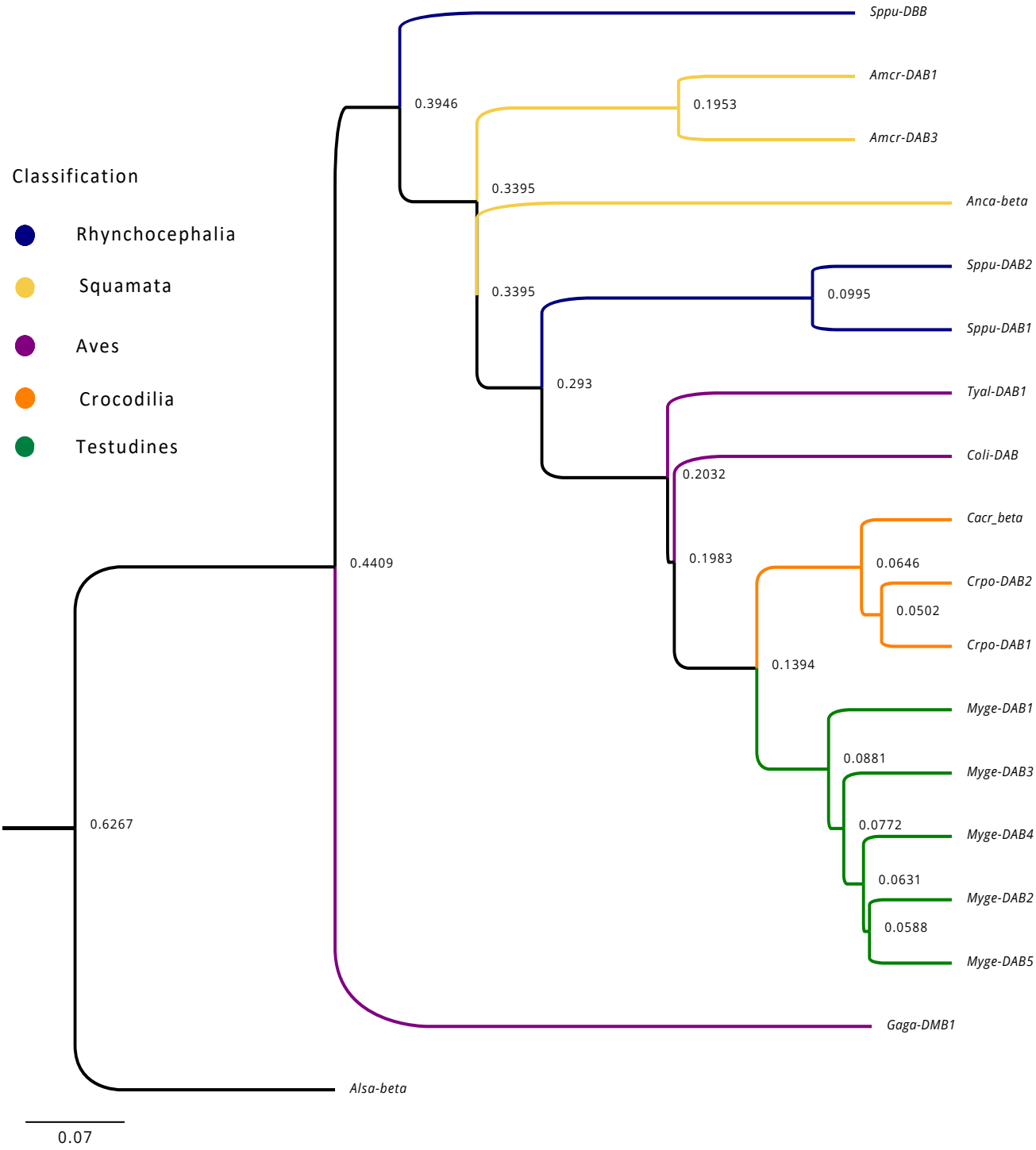


Figure A3.2.3. Maximum likelihood analysis of complete MHC II beta sequences for non-avian and avian reptiles. Node ages based on 1000 replicates are shown. American shad (*Alosa sapidissima*) (*Alsa-beta*) was used as an outgroup. Chicken DMB and Tuatara DBB sequences were included for DMB and DBB-like clades, respectively.

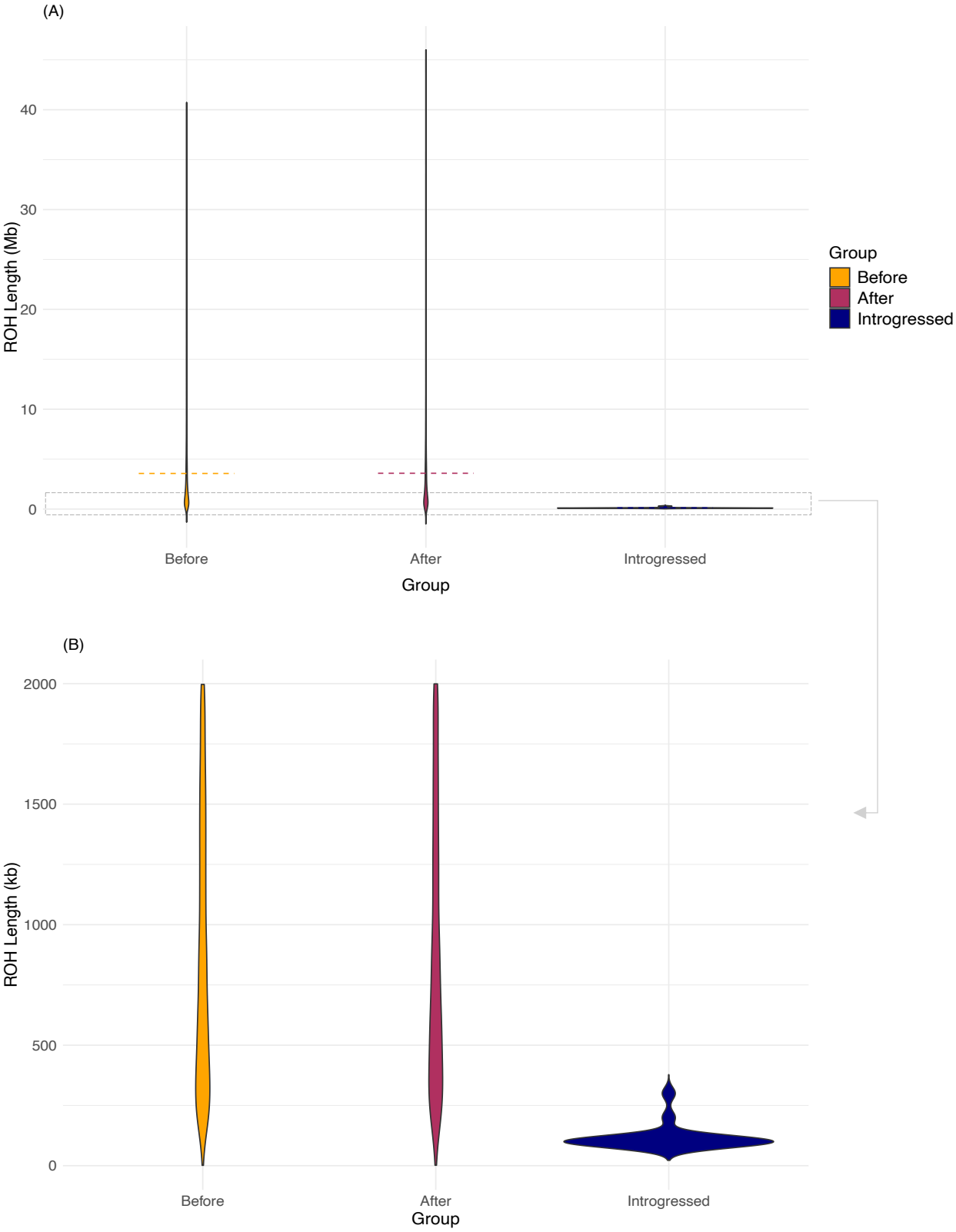


Figure A3.2.4. Violin plots showing the distribution of run lengths of homozygosity (ROH) calculated using ROHan across the Before, After, and Backcross groups. **(A)** Distribution of all ROH lengths, with a dotted line indicating the mean ROH length. **(B)** An enlarged view showing distribution of short ROH with lengths less than 2 Mb.

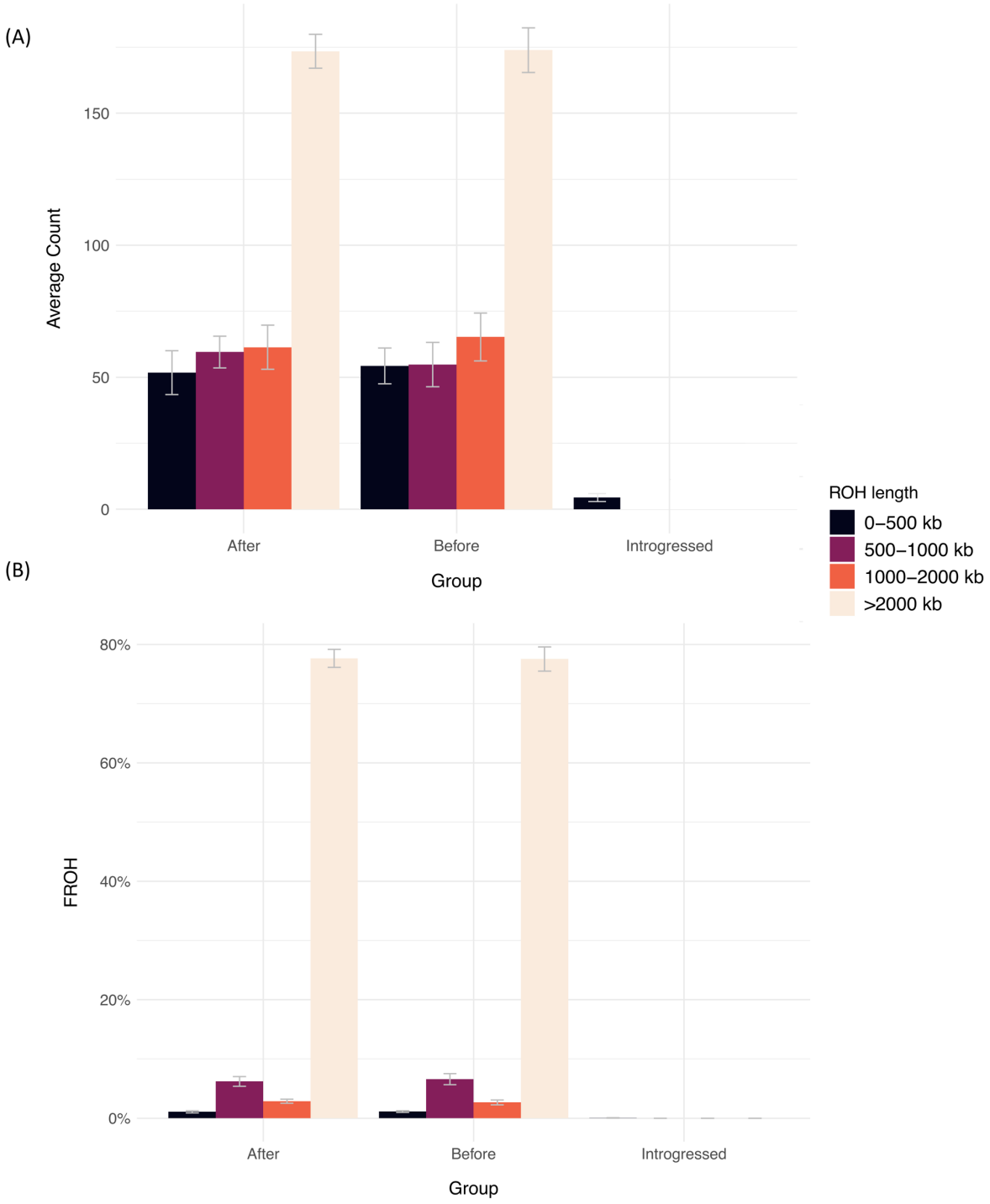


Figure A3.2.5. (A) Average count of runs of homozygosity (ROH) across 0-500 kb, 500-1000 kb, 1000-2000 kb, and >2000 kb length categories for individuals in Before, After, and Backcross groups. **(B)** Inbreeding coefficient based on ROH (F_{ROH}) across 0-500 kb, 500-1000 kb, 1000-2000 kb, and >2000 kb length categories for individuals in Before, After, and Backcross groups. The F_{ROH} is represented as a percentage of the total length of the 9 macrochromosomes.

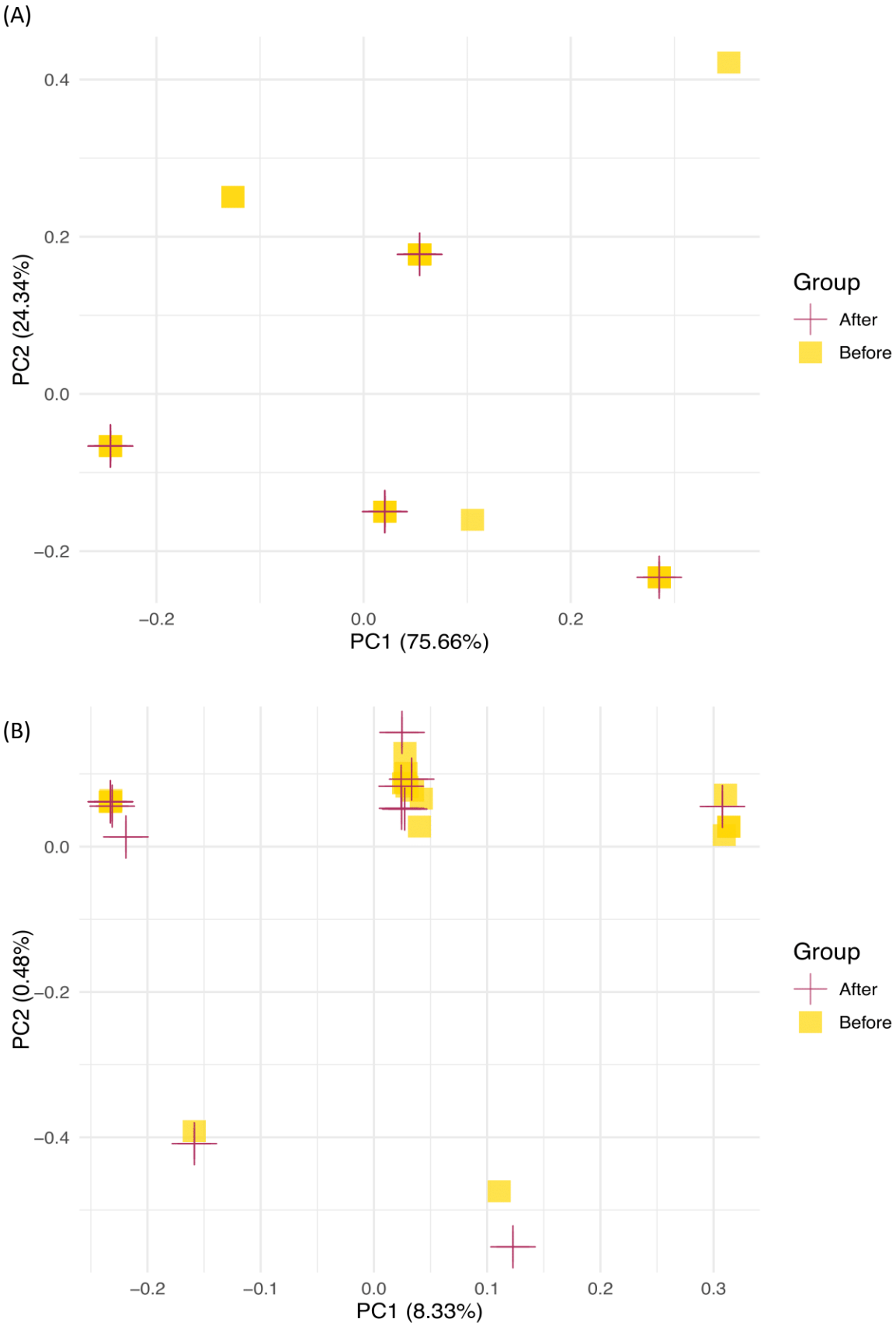


Figure A3.2.6. Principal component analysis plots of SNPs in pure *Myuchelys georgesi* from the Before (N = 19) and After (N = 12) groups in **(A)** MHC I and **(B)** MHC II.

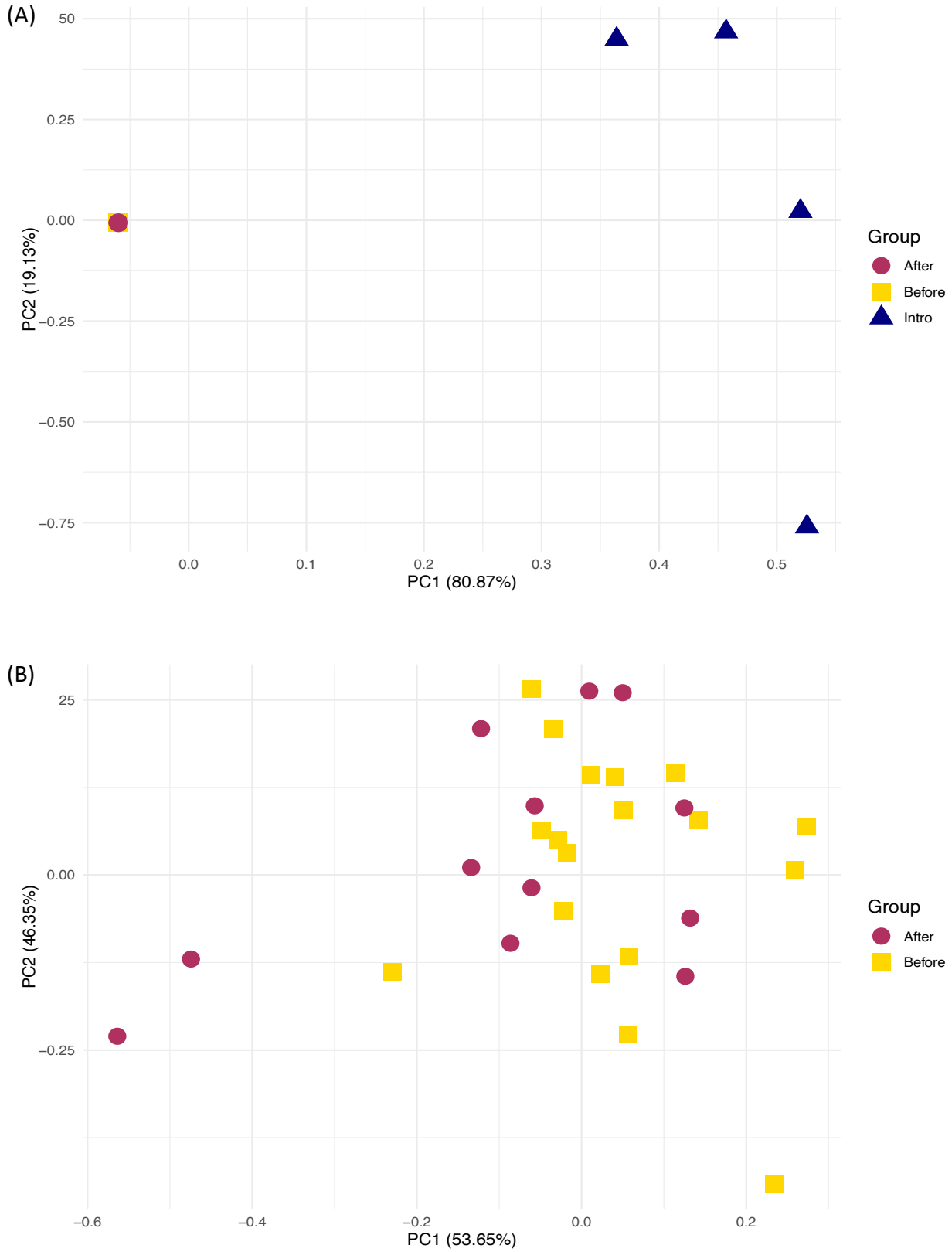


Figure A3.2.7. Principal component analysis plots of genome-wide SNPs in **(A)** pure *Myuchelys georgesi* Before (N = 19), After (N = 12), and F2 backcross animals (N = 4), and **(B)** pure *M. georgesi* Before (N = 19) and After (N = 12).

All exon sequences for MHC class I genes annotated using the reference genome

> Myge-UA

CAGGATGAAGAGCTGGAGGCAGAATCTTCTCCGGTCTTTGGGCTGCCGGAGTGTAGCCGCGTTTGGG
GGTGCTGCTGATTTTCTTCTCCACAGAATGACGCTGGCCCCTACGGCTCCAGCCAGAACCAGCACGGC
CAGGACGATGGCGGCCAGGACCCCAGGGGACAGGGGGCTCTTCTTCCCAGGGGCCAGGCGAGTGT
GGGGTCTGATAGGCTGCTGTGCTCCACCCGGCAGGCATAGCGGTGACTGTCCTCCTGCTGCGGGGAG
ATCTCCAGGGATGACTGCGTGTAGTAGGTGCCATCAGCATTGGGCAGGATCCCACTGGAGTGCTTCTC
CACCAGGATGTCTTCTCCATCCCGCACCCAGGAGACGTGGATGGGACGTGGGTAAAAGCCCCTGGCG
CGGCAGGAGAGGGTGACAGAGCTGTCTGTAGAGTCTGTGTAGGAAACCGAGACCTCGGGGGGTCCT
GCTGCTCCAGCACCGTCCTCCCCTGCTGCACCAGGCTCTGCAGAGTCCCCAGGCACTCATGCTGCAGG
TACTGCCGGACAAACTGAGTCCAGGCCTTGCCCGTCTCCCAGCTCTGCTTCTGGGGGAAGGCCGGCT
GCACGGCCGCGACCCACGTCCCCGTCTGGTTGTCAAAGCTGATGAAGTCCTTCCCGTCGTAGGCATAC
TGGAACTTCGGGTCCACAAGCACCTGGCCGCTCAGGGCACAGCTCACGTGAACCTGCTCGGTGTGAA
TCCCGCCCGTCTGGTTGTGCAGCTGCATCCACCACCTGGTATTGACTTTACAGCCCTCCTCGTATCCCCA
GAACTCCTGGGTCTTTTCCTTGAGATACTCAACACCCACGGCCTGTGTCGCCCAATCCTGGGTAGGTCT
GACCTCTCGCGTGTCGCTGCTGTAGTACGCGATCTTCACGTCATCCAGCTTGGCAATCATGAAGTAGCG
ATTGGTCCCCTCCTCGTCCATGACTGCTGTGACTAGCACGGCCAGGCTGTGGTGCCCGTCAGCCGCGG
CCGGTGCCAGGATTGCCCCGCAGAGCGCCAGGAGCAGCCCCCACGCCAT

>Myge-UB

CATCTGCAGCCCAGGCGCCTGCACTGGGAAGTGGAAGCTGCAGTTGCGGTCTCAGAGCCGGCGTTTC
GCTGGGTACAATCACGGCAAAGGGCGTACCGGAGCTGGCAGAGCCGTCGTGAGCTGAAGCCGCTTT
GTAGCCGCCTTTCTTCTTCCTGAGGCAGAGAACAGCTCCCGCAATGACAGCCACCAGCACGAGAACC
CCGAGAACGATCCCCACGATCAGCGTCACGCTGGACTTGGGCGCCCAGGGCACTCTCAGATCCTCCGT
CAGGCTGGGGTGCTCCACGCAGCAGGTATAATCCGTCTCTTTGCTGGGGTCGATCTCTATGGTCGCCC
GGGTCTGGTAGGTCCCGTCCCCGCTGGGAACCACTCCCAATTGCATCGTCTCCTGGGGCACGGCCACC
CCCTTCTTCAGCCACACGACAGCGACGCTGCGCGGGTAGAAGCCGTGCACCCGGCAGGAGAGGGTG
GTGAGCCCGTCCCCGCTCGGCCGGTCGCTCACCTGCACCCGCGGTCGCTTTCCTCTGCAGAGCGTCCT
TCCCATACTCCAGGTATTTTTTCAGCCAGTCGATACAGATCTCCTCCAGGTAATTCTTTCGCCTTTGAAG
ATGGTCCTTATCAGCCTCCCATTTCCTCTTGGTGACTTGCGCACCAGCATCTGCTGCAGTGTACGTCATG
GTGTCCTTGTCGAAGCTAACGAAGTCTCTCCCGTCATAGGCATACTGGTGAAAAGCCCCAGTGGTGCC
GTCATCCCGGAGATCACAGCCATACATATTCTGCCAGGTGTGAAAGCCCGCGCTCTGGTTGTAGCGCCC
GCGCAGCGTGTTCAGGTCCACGCGGAAGATGGCCTGCCAGCCCTGCAAGTTCTGCGTCTCCCGCTCC
CAGTACTGCGCGTCCTCGATCCGCGCCGTCCAGTCCGCGCGCGGCTCCGCGCTGCCCCGCGCGCTGTC
GTAGTCCATGAAGCGCTGCCCGTCCACGGACCCCACCGTGATGAAGTCGGGCAGCCCCGGGCCGGGC
TCCGACACGCCCGTGTAGAAATAGCGCAGAGAGTGTGTGCCGGGGGCGCAAAGGGCCCCCGGCAGG
GCCACGGCCCCCAGCAGCAGGAGACGCAGGGCCAGCGCCAT

> Myge-UC

TAGCTCCAGTGTCCCAGGCTCCGGGTGTCTGATCCGGATGAAGAGCTGGAGGCAGAATCTTCTCCGAT
CTTTGGGCTGCTGGAGCGTAGATGGGTTTGGAGGGGCTGCTGATTTTCTTCTCCACAGGATGACGCT
GGCCCCTACGGCTCCAGCCAGAACCAGCAAGGCCAGGACGATGGTGGCCAGGACCCCAGGGGACA
GGGGCCCCTTCTTCCCAGGGGCCAGGCGAGCGTGGGCTCTGACAGGCTGCTGTGCTCCACCCGGCA

GGCGTAGCGGTGCCTGTCCTCCTGCTGCGGGGAGATCTCCAGGGACGACTGCGTGTAGTAGGTGCCA
TCGGCGTTGGGCAGGATCCCGCTGGAGTCTGTCTCCACCAGGATGTCTTCTCCATCCCGCACCCAGGA
GACGTGGATGGGACGTGGGTGAAAGCCACTGACACGGCAGGAGAGGGTGACAGAGCCGCTGGGG
GTGTCTCTGCGGGAAACCGAGACCACGGGGGGCACTGCTGCTCCAGCACCGCCCTCCCCTGCTGCAC
CAGGCTCTGCAGAGTCCCCAGGCACTCGTGCTGCAGATACTGCTGGACAAACTTGGTCCAAGTCTTGC
CCGTCTCCCAGCGCTGCTTAGGGGCGAAGGCCAGCTGCACGGCCGCGACCCACGTCCCCGTCTGGTT
GTCAAAGCTGATGAAGTCCCTCCCGTTGTAGGCGAACTGAAACCTCGGGTCCACGGGGGCCTGGCCG
CTCAGTGCACAGCTGACGTGAACCTGCTCGGTGTGAAACCCGCCCGTCTGGTTGTGCAGCTGCATCCA
GCGCCTGGTCTCGACTTTAGAGCCCTCCTCATGCTTCCAGAACTGCTGGGTCTTTTCCTGGAGATACTC
AACACCCACGGCCTGCGCCACCCACTGCTGGGTGGGTCTGACCTCCCGCATCTCGCTGCTGTAGTACG
CGATCTTCACGTCATCCAGCTGGGTGATCATGATGAAGTGGTGGGTCCCCTCCTCGTTGATGACTGCTG
TGACTAGCGAGGCCAGGCTGTGGTGCCTGTCAGCCGCAGCTGCTGCCAGGATTCCCCAGCAGAGCCC
CAGGAGAAGACCCCAGGCCAT

> *Myge-UD*

GATGAAGAGCTGGAGGCAGAATCTTCTCCGGTCTTTGGGCTGCTGGAGCGTAGCCGGGTTTTGGGG
GGCTGCTGATTTCCTTCTCCACAGGATGACGCTGGCCCCTATGGCTCCAGCCAGAACCAGCACGGCCA
GGACGATGGCAGCCAGGACCCCAGGGGACAGGGGCCCCTTCTTCCCAGGGGGCAGGCGAGCGTGG
GCTCTGATAGGCTGCTGTGCTCCACCCGGCAGGCGTAGCGGTGCCTGTCCTCCTGCTGCGGGGAGATC
TCCAGGGACGACTGCGTGTAGTAGGTGCCATCAGCGTTGGGCAGGATCCCGCTGGAGTCTGTCTCCA
CCAGGATGTTTTCTCCGTCCCGCACCCAGGAGACGTGGATGGGACGTGGGTGAAAGCCACTGGCGC
GGCAGGAGAGTGTGACAGAGCCGCTGGGGGTGTCTCTGCGGGAAACTGAGACCACAGGAGGTACT
GCTGCTCCAGCACCGCCTTCCCCTGCTGCATCAGGCTCCGCAGGGTCCAGAGGCACTCATGCTGCAGG
TACTGCTGGACAAACTGAGTCCAGGTCTTGCCCGTCTCCCAGCGCTGCTTCTGAAGGAAGGCCGTCTG
CACGGCCGCGACCCACGTCCCCGTCTGGTTGTCAAAGCTGATGAAGTCTCTCCCATCGTAGGCGTACT
GGTACCTTGCATCCAGGGGGTCCTGGCCACTCAGGGCACAGCCCACGTGAATCTGCGAGCTGTGAAT
CCCACCCGTATGGTTGTGCAGCTGCATCCTCCACCTGGTCCCGCCTTTAAAGCCCTCCTCGTGCCCCCA
GAACTGCTGGGTCTTTTCCTTGATATACTGAGCGCCCACGGCCTGTGCTGCCCACTGCTGGGTGGGTC
TGACCTCTCGCGTGTCGCTGCTGTAGTACGCGATCTTAACATCATCCAGTTCGGCGATCATGATGAAGT
GATAGGTCGCATCTTCGTTGATGATTCCTGTGACTAGCACGGCCAGGCTGTGACGCTCAGCAGCCGCT
GCCGCCCCCAGCAGCGCCCAGCAGAGCCCCAGGGCGACACCCAGGGCCAT

> *Myge-UE*

CACTGCACATCATCTACTGATGCTGGCTGAGCTGTGAGGGGAGAAGGAAGCAGGGTCAGTGTCTCGC
AGAACGCAGGGTCCCCGTCCTCCCCTGCCCCACACTCGTGACACACGGAGACGCCCCAGTAACGGAG
CCGGATGTCGGTGTTTTGCCGTTGTGCCGGGTACCCAGCGATGGGCGACCTATACGGAGCAGCTGCA
GCCGCTTTGTAGCCGCCTTGAGGCAGAGAACAGCTCCCGCGATGACAGCCGCCAGCACGAGAACCCC
GAGAACGATCCCCACGATCAGCGTCACGCTGGACTTGGGCGCCCAGGGCACTCTCAGATCCTCCATCA
GGCTGGGGTGCTCCACGCAGCAGGTATAATCCGTCTCTTTGCTGGGGTCGATCTCTATGGTCGCCCGG
GTCTGATAGGTCCCGTCCCCGCTGGGAAGCACGTCCCACTGCATTGTCTCCTGGGGCACGGCCACCCC
CTTCTTCAGCCACACAACGGCGACGCTGCGCGGGTAGAAGCCGTGCACCCGGCAGGAGAGGGTGGT
GAGCCCGTCCCGGCTTGGCCGGTCGCTCACCTGCACCCGCGGTTGCTCTCCTCTGCAGCGTCTCCTGC
CCATACGCCAGGTACTTCTCCAGCCACTCGATGCAGGTCCCCTCCAGGTAGGCTCTCCACCCCCGAGTG
AAGCTCCTGTCAGCATCCCAGTTCCGTTTGCTGATCTGAGCCGCGTCGTCTGCCGCTACCCAGGTCTCT

CGCTCCTTGTCCAGGCTGAGGAAGTCTCGCCCGTCATACGCCTCCTGGTAAAACCCCCCGATGGAGCC
GTCACCGTGTAGCTCACAGCCGTACATCCTTTGCAGAGTGTGAAATCCCGCGCTGTGGTTGTAGCGCC
CGCGCAGGGTGTTCAGGTTGGCTCGGTACAAGGCCTGCACGCCCTGGGAGACCTGCGTCTCCCCGGC
CCAGTACTGCGCGTCCAGGCTCCGCACCATCCACTCTGCGCGCGGCTCCGCTCTGCGTGTCTCGCTGTC
GTAATGCAGGATGCGCTGCCCGTCCACGGAGCCGACGATGCTGAACTCGGGCAGCCTCGGGCCGGG
CTCCGACACCCCCGTGTAGAAATTGCGCAGCGAGTGCAGGCCGGGCCCGGGGGCGCTACGGGCCCC
CGGCAGGGCCACGGCGCCCAGCAGCAGCAGGCGCAGGGCCAGCGCCAT

All exon sequences for MHC class II genes annotated using the reference genome

> Myge-DAA1

TCACAAGGGGCCCCGCGGGTTGCGGGCAGCGTTCATCTTCATCCCCTTGATGATGAGGATGGTGCCG
GCGATGATGCCGATGAGGCCCACGGCCAGGCCCAGGGCGCACACCAGGGTCTCTGTGGTCTCGGGG
ACGGGGGTGGGCACCTGGGGCTCCCAGTGCTTCAGGAAGGGCTGGGCCAGCCCCGCGTGCTGCACC
CGGCAGTCGTAGAACTCGTCCTGCCGGGGGATGAAGGGCAGGTAGGAGAACTTGCGGAAGGCGTTG
TCCTGGCGGGGGTAGAAGTCGGTCTCGAAGACGCCCCCCGTCACCTCCTGCCCGTTCTTCAGCCACGT
CACACTGAGCGCAGGCGGCGAGAACTTGTCCACGAAGCAGATCAGGACGTTGGGCTCGCCCAGCTC
CACGGGGCCTTTGGTGAACACGGTCACCTCGGGGGGCAGTTCTGGGCCCGCGTGTAGTTGGACATCT
GGATCATGGTCTCCAGGTTTGCTTTGTCCACAGCGATGTTGCCCAGGGCGCCCTGCGCCTCGAAGCTG
GTGAAGGTGCCGAAGTCGGGCAGGCGCCAGACGGTCTCCTTCTTCTCCAGGTCCACGTGGAAGATCT
CATCCTGGTCAAACTCGAACATGAACTCCCCGGACTCCCGCTGGGACTGGGCCGTGCGCTGGTAGAA
CTCCGCCTGGGACAGCAGGTTCTCCGTGTCACTGCCCCAGCGCCCAGCAGCGCCAGCAGGGTGAGCA
GGGCCAGCTGGGCCATGGGGACGCCCCGTCCTGCGCCCAT

> Myge-DAB1

ATGGGGCCGGGTCGGAGCCTGGGGGCCGGGAGCCGCTGGGCTGGGGCTCTGCTCGTGACCCTGAC
GGTGCTGGGAACCCGCCTGGCTCATTGCACGGAGCCCGAAGGGCGGTTCCTGCACCAGGCGAAGCA
CGACTGTTACTTCACCAACGGCTCCGAGCGGGTCCGGTTCGTGGATCGGTACATCTACGACCGGCAGC
AGATCGTGCACTTCGACAGCGACGTGGGGCGGTACGTGGCGGACACGGCGATGGGCCGGCCCACAG
CCGAGTATTGGAACAAGGACCCGGCCGTGCTGGCGGAGAGGCGGGCCAGGGTGGACACGTTCTGCC
GGTACAACTACGGCGTGGCCCAGACGGGGAAGGTGGTCGGCCGCGCAGTTCAGCCTGAGGTGACAG
TTTTCCCCACCAAATCGGGGTCCCAGCCCCACCTGCTGGTTTGCTCCGTGACGGGGTTTTACCCCGGG
GAGATCGAGATCAAGTGGCTGAAGAACGGGCAGGAGCAGACGGCCGGGGTGGTGTCCACGGAGCT
GCTCCAGAACGGAGACTGGACCTTCCAGATCCTGGTGATGCTGGAGACGAGCCCCCGCCGCGGGGA
CGTCTACGCCTGCCAGGTGCAGCACGTCAGCCTGCCCGAGCCCCTCACCGTGCGCTGGGAGGCGCAG
TCTGACACCGCCCGGAGCAAGATGCTGACGGGGGTCGGGGGCTTCGTGCTGGGCCTGATCTTCCTGG
TGCCAGGACTGGCCATCTACCTGAGGAACAAGAAAGGGCGCCCCCTTCCCCAACCTGCAGGGCTCCT
GAGTTAGA

> Myge_DAA2

TCACAAGGGGCCCCGAGGGTTGCGGGCGGCGTTCATCTTCATCCCCTTGATGATGAGGATGGTGCCG
GCGACGATGCCGATGAGGCCCACGGCCAGGCCCAGGGCGCACACCAGGGTCTCTGTGGTCTCGGGG
ACCGGGGTGGGCACCTGGGGCTCCCAGTGCTTCAGGAAGGGCTGGGCCAGCCCCGCGTGCTGCACC

CGGCAGTCGTAGAACTCGTCCTGCCGCGGGATGAAGGGCAGGTAGGAGAACTTGCGGAAGGCGTTG
TCCTGGCGGGGGTAGAAGTCGGTCTCGGAGACGCCCCCCGTCACCTCCTGCCCGTTCTTCAGCCACGT
CACGCTGAGCGCGGGCGGGAAGAACTTGTCCATGAAGCAGATCAGGACGTTGGGCTCGCCCAGCTC
CACGGGGTCTTCGGGGAACACGGTCACCTCGGGGGGCACGTTCTGGGCCCGCGTGTGGTTGGACCT
CTTGATCAGGATCTCCAGGTTGTTCTTGCCCGTAGCCGCGTTGCCCAGAGCGAACTGTGCATCGAAGC
TGGAGAACTCGCCAAATTTGGGCAGGCGCCAGACCGTCTCCTTCTTCTCCAGGTCCACGTAGAACATC
TCATCCTGGTCGAACCCCTGGATGAACTCCCCGGACTCCTGCTGGGACTGGGCCGTGCGCTGGTAGAA
CTCTGCCTGGAAGAGCACGTGCTCCACTTTCACTGCCCCGGTGCCCGGCAGGGCCAGCAGGGCGAG
CAGGGCCAGCTGGGCCATGGGGACGCCCCGTCCTGCGCTCAT

> *Myge_DAB2*

ATGGGGCCGGGTCGGAGCCTGGGGGCCGGGAGCCGCTGGGCTGGGGCTCTGCTCGTGACCCTGAC
GGTGCTGAGAACCCGCCTGGCTCATTGCACGGAGCCCGAAGGGCGGTTCGTGCTCCAGTGGAAGGG
CGACTGTTATTTCACCAACGGCACCGAGCGGGTCCGGCTGCTGGGCCGGTACATCTACAACCAGCAGC
AGATCGCGCACTTCGACAGCGCCGTGGGGCTGTGGGTGCCCGACACGGAGCTGGGCCGGCCCGACG
CCGAGTATTGGAACAAGGACCCGGCCGAGCTGGCGCGCAATCGGGCCGATGTGGACCGGTTCTGCC
GGCACAACTACGAGGTGGACGCGCCTTTCACCGTGGAGCGCAGAGTTCAGCCCAAGGTGAAAGTTT
CCCCCACCAAATCGGGGTCCCAGCCCCACGCCCACCTGCTGGTTTGCTCCGTGACGGGGTTTTACCCC
GGGGAGATCCAGATCACGTGGCTGAAGAACGGGCAGGAGCAGACGGCCGGGGTGGTGTCCACGGA
GCTGCTCCAGAACGGAGACTGGACCTTCCAGATCCTGGTGATGCTGGAGATGAGCCCCCGGCGCGGG
GACGTCTACGCCTGCCAGGTGCAGCACGTCAGCCTGCCCGAGCCCCTCACCGTGCGCTGGGAGGCGC
AGTCTGACGCCGCCCGGAGCAAGATGCTGACGGGGGTCGGGGGCTTCGTGCTGGGCCTGATCTTCCT
GGCGCCGGGACTGGCCATCTACCTGCGGAACAAGAAAGGGCATCCCCTTCCCCAACCTGCAGGGCTC
CTGAGTTAG

> *Myge_DAA3*

TCACAAGGGGCCCCGCGGGTTGCGGGCGGCGTTCATCTTCATCCCCTTGATGATGAGGATGGTGCCG
GCGATGATGCCGATGAGGCCCACGGCCAGGCCCAGGGCGCACACCAGCGTCTCTGTGGTCTCGGGG
ACGGGGGTGGGCACCTGGGGCTCCCAGTGCTTCAGGAAGGGCTGGGCCAGCCCCGCGTGCTGCACC
CGGCAGTCGTAGAACTCGTCCTGCCGCGGGATGAAGGGCAGGTAGGAGAACTTGCGGAAGGCGTTG
TCCTGGCGGGGGTAGAAGTCGGTCTCGAAGACGCCCCCCACCACCTCCTGCCCGTTCTTCAGCCACGT
CATGCTGAGCACGGGCGGGAAGAACTTGTCCACGAAGCAGATCAGGACGTTGGGCTCCCCCAGCTCC
ACGGGGCCTTTGGTGAACACGGTCACCTCGGGGGGCACGTTCTGGGCCCGCGTGTAGTTGGATCTCT
TGATCAGGATCTCCAGGTTTGATTTGTCTGTGGCAATGTTGCCCAGGGCGCCCTGTGCCTCGAAGCTG
GTGAACTTGCCGAAGTCGGGCAGACGCCAGATGGTCTCCTTCTTCTCCAGGTCCACGTGGAAGATCTC
GTCCTGGTCCAACTCGAACATGAACTCCCCGGACTCCTGCTGGGACTGGTCCGTGCGCTGGTAGAACT
CCGCCTGGGACAGCAGGTTCTCCGCTGTCACTGCCCCAGTGCCCGGCAGCGCCAGCAGGGTGAGCA
GGGCCAGCTGGGCCATGGGGACGCCCCGTCCTGCGCCCAT

> *Myge_DAB3*

ATGGGGCCGGGTCGGAGCCTGGGGGCCGGGAGCCGCTGGGCTGGGGCTCTGCTCGTGACCCTGTCG
GTGCTGAGAACGCGCCTGGCTCATTGCACGGAGCCCGAAGGCCGGTTCGTGTATCAGTGGAAGCACG
AGTGTCAGTTCACCAACGGCACCGAGCGGGTCCGGCTCCTGCTGCGGTACATCTACAACCAGCAGCA

GCTCGCGCACTTCGACAGCGACCTGGGGCTGTGGGTGGCGGACACGGAGCTGGGCCGGCCCGACG
CCGAGTTTTGGAACAAGGACCCGGCCGTCCTGGCGCGCAATCGGGCCGATGTGGACCGGTTCTGCCG
GCACAACTATGGGGTGGCCCAGACGGGGAAGATGGTCGGCCGCGCAGTTCAGCCCCAGGTGACAGT
TTTCCCCACCAAGTCGGGGTCCCAGCCCCACCTGCTGGTTTGCTCCGTGACGGGGTTTTACCCCGGGG
GGATCCAGATCAAGTGGCTGAAGAATGGGCAGGAGCAGACGGCCGGAGTGGTGTCCACGGAGCTGC
TCCAGAACGGAGACTGGACCTTCCAGATCCTGGTGATGCTGGAGATGAGCCCCCGGCGCGGGGACAT
CTACGCCTGCCAGGTGCAGCACGTCAGCCTGCCCGAGCCCCTCACCGTGCGCTGGGAGGCGCAGTCT
GACGCCGCCCGGAGCAAGATGCTGACGGGGGTCGGGGGCTTCGTGCTGGGCCTGATCTTCCTGGCG
CCGGGACTGGCCATCTACCTGCGGAACAAGAAAGGGCGCCCCGTTCCCCAACCTGCAGGGCTCCTGC
GTTAG

> Myge_DAA4

TCACAAGGGGCCCCGCGGGTTGCGGGCGGTGTTCATCTTCATCCCCTTGATGATGAGGATGGTGCCG
GCGATGATGCCGATGAGGCCCATGGCCAGGCCCAGGGCGCACACCAGGGTCTCTGTGGTCTCGGGG
ACGGGGGTGGGCACCTGGGGCTCCCAGTGCTTCAGGAAGGGCTGGGCCAGCCCCGCGTGCTGCACC
CGGCAGTCGTAGAACTCGTCCTGCCGGGGGATGAAGGGCAGGTAGGAGAACTTGCGGAAGGCGTTG
TCCTGGCGGGGGTAGAAGTCGGTCTCGGAGATACCCCCCATCACCTCCTGCCCGTTCTTCAGCCACGT
CACGCTGAGCACGGGTGGGAAGAACTTGTCCACGAAGCAGATCAGGACGTTGGGCTCGCCCAGCTC
CACGGGGCCTTTGGGGAACACGGTCACCTCGGGGGGATCGTTCGGGGCCCGCGTGCGGTTCAATCT
TTTGATCATGATCTCCAGGTTGTTCTTGCCCACGGCCATGTTGCTCAGGGCGCCCTGTGCCTCAAAGCT
GGCAAACTTGCCAAAGTAGGGCAGGCGCCAGACAGTCTCCTTCTTCTCCAGGTCCACATGCAAGATCT
CGTCCTGGTCGAACTCGAACATGAACTCCTCGGACTCCTGCTGGGATGGTTCCGTGCGCTGGTAGAAC
TCCACCTGGGCCGACATGTGTTCCACTGTCACTGCCCCGGTGCCCGGCAGCGCCAGCAGGGCGAGCA
GGGCCAGCTGGGCCATGGGGACGCCCCGTCCTGCGCCCAT

> Myge_DAB4

ATGGGGCCGGGTCGGAGCCTGGGGGCCGGGAGCCACTGGGCTGGGGCTCTGCTCGTGACCCTGAC
GGTGCTGAGAACGCGCCTGGCTCATTGCACGGAGCCCCCCGAGCATTTCGTGCACCAGGCGAAGTTC
GAGTGTCACTTCACCAACGGCACCGAGGGGTTCCGGTTCGTGCACCGGCACATCTACAACCGGCAGC
AGTTCGCGCACTTCGACAGCGACCTGGGGCTGTACGTGGCGGACTCGGAGCTGGGCCGAATCGATGT
GGAGACCTGGAACAAGGACCCGGCCGAGCTGGCGTACCGACGGGCCGGGGCGGAACGGTTCTGCC
GGACCAACTACGAGGCGGATGAGCCCTACACCATCGACCGCAGAGATCAGCCCGAGGTGACAGTTTC
CCCCACCAAATCGGGGTCCCAGCCCCACCCCCACCTGCTGGTTTGCTCCGTGATGGGGTTTTACCCCG
GGGGGATCCAGATCAAGTGGCTGAAGAACGGGCAGGAGCAGACGGCCGGGGTGGTGTCCACGGAG
CTGCTCCAGAACGGAGACTGGACCTTCCAGATCCTGGTGATGCTGGAGATGAGCCCCCGGCGCGGGG
ACGTCTACGCCTGCCAGGTGCAGCACGTCAGCCTGCCCGAGCCCCTCACTGTGCACTGGGAGGCGCA
GTCTGACGCCGCCCGGAGCAAGATGCTGACAGGTGTTGGGGGCTTCGTGCTGGGGCTGATCTTCCTG
GCGCCGGGGCTGGCCATCTACCTGCGGAACAAGAAAGGGCGCCCCCTTCCCCAACCTGCAGGGCTCC
TGAGTTAG

> Myge_DAA5

TCACAAGGGGCCGCGCGGGTTGCGGGCAGCGTTCATCTTCATCCCCTTGATGATGAGGATGGTGCCG
GCGACGATGCCGATGAGGCCCACGGCCAGGCCCAGGGCGCACACCAGGGTCTCTGTGGTCTCGGGG

ACGGGGGTGGGCACCTGGGGCTCCCAGTGCTTCAGGAAGGGCTGGGCCAGCCCCGCGTGCTGCACC
CGGCAGTCGTAGAACTCGTCCTGCCGGGGGATGAAGGGCAGGTAGGAGAACTTGCGGAAGGCGTTG
TCGTGGCGGGGATAGAAGTCGGTCTCGAAGACGCCCCCCGTCACCTCCTGCCCGTTCTTCAGCCACGT
CACGCTGAGCGCGGGCGGCGAGAACTTGTCCACGAAGCAGATCAGGACGTTGGGCTCGCCCAGCTC
CACGGGGCCTTTGGTGAACACGGTCACCTCGGGGGGCACATTCTGGGCCCACGTGCGGTTGGATCTC
TTGATCAGGATCTCCAGGTTTGATTTGTCCGTGGCGATGTTGCCCAGGGCGCCCTGCGCCTCAAAGCT
GGTGAAGGTGCCGAAGTCGGGCAGGCGCCAGACGGTCTCCTTCTTCTCCAGGTCCACGTGGAAGGT
CTCGTCCTGGTCGAACTCCTGCATGAACTGCCCGGACTCCTGCTGGGACTGGGCCGTGCGCTGGTAGA
ACTCCGCATGGATGTGCACGTGCTCCACTGTGACTGCCCCGGTGCCTGGCAGGGCCAGCAGGGCGAG
CAGGGCCAGCTGGGCCAT

> Myge_DAB5

ATGGGGCCGGGTCGGAGCCTGGGGGCCGGGAGCCGCTGGGCTGGGGCTCTGCTCGTGACCCTGAC
GGTGCTGAGAACCCGCCTGGCTCATTGCACGGAGCCCCCACGGCGGTTCCTGCACCAGTTCAAGGGC
GAGTGTCAGTACATCAACGGCACGGAGCGGGTCCGGTTCGTGTACCGGCACATCTACGACCGGCAGC
AGATCGCGCACTTCGACAGCGACGTGGGCCGGTTCGTGGCGGACACGGAGCTGGGCCGGCCCGACG
TGGAGTACTGGAACAACGACCCGGCCTTACTGGCGGAGAGGCGGGCCGAGGTGGACCGGTTCTGCC
GGCACAACTACGGGGTGTACGAGCCTTTCACCATCGACCGCAGAGTTGAGCCCGAGGTGACAGTTTC
CCCCACCAAATCGGCGTCCCAGACCCACCTGCTGGTTTGCTCCGTGACGGGGTTTTACCCCGGGGGG
ATCCAGATCAAGTGGCTGAAGAACGGGCAGGAGCAGACGGCCGGGGTGGTGTCCACGGAGCTGCTC
CAGAACGGAGACTGGACCTTCCAGATCCTGGTGATGCTGGAGATGAGCCCCCGGCGCGGGGACGTC
TACGCCTGCCAGGTGCAGCACGTCAGCCTGCCCGAGCCCCTCACCGTGCGCTGGGAGGCGCAGTCTG
ACGCCGCCCGGAGCAAGATGCTGACGGGGGTCGGGGGCTTCGTGCTGGGGCTGAACTTCCTGGCGC
CGGGACTGGCCATCTACCTGCGGAACAAGAAAGGGCTCCTGAGTTAG

Appendix 4: PDF versions of publications and reports

A4.1 A genomic framework to assist conservation breeding and translocation success: A case study of a critically endangered turtle.

The published PDF version of Chapter 2 of this thesis is presented on the following pages.

Received: 3 December 2023 | Revised: 7 July 2024 | Accepted: 16 July 2024

DOI: 10.1111/csp2.13204

CONTRIBUTED PAPER

Conservation Science and Practice
A journal of the Society for Conservation Biology
Open Access
WILEY

A genomic framework to assist conservation breeding and translocation success: A case study of a critically endangered turtle

Holly V. Nelson[1] | **Katherine A. Farquharson**[1,2] | **Arthur Georges**[3] | **Elspeth A. McLennan**[1] | **Jane L. DeGabriel**[4] | **Melissa Giese**[5] | **Christopher Ormond**[5] | **Michael McFadden**[6] | **Adam Skidmore**[6] | **Jarrad Prangell**[7] | **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]Institute for Applied Ecology, University of Canberra, Bruce, Australian Capital Territory, Australia

[4]NSW Department of Climate Change, Energy, the Environment and Water, Parramatta, New South Wales, Australia

[5]NSW Department of Climate Change, Energy, the Environment and Water, Coffs Harbour, New South Wales, Australia

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

[7]Symbio Wildlife Park, Helensburgh, New South Wales, Australia

Correspondence
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
NSW Department of Planning and Environment, SoS PhD Scholarship; Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Grant/Award Number: CE200100012; Australian Government Bushfire Recovery Scheme, Grant/Award Number: ERF-WRR2-020

Abstract

Conservation breeding programs are an effective approach to addressing biodiversity loss. Captive populations are managed to maintain genetic diversity, yet there remains an "implementation gap" in effectively translating molecular genetic data into management. Technological advancements are facilitating rapid generation of genetic data, increasing accessibility for breeding programs. In 2010, Frankham and colleagues proposed a six-stage process for establishing successful conservation breeding and release programs. Here, we describe the conservation breeding program for the critically endangered Bellinger River turtle (*Myuchelys georgesi*) and characterize the value of genetic sampling for informing management actions. By generating a chromosome-level genome and population genetic data, we investigated past and present diversity and assessed relatedness among captive founders. We present a framework modeled on Frankham and colleagues six stages to assist managers in implementing genetic data into actionable conservation strategies. This framework, and worked case study, for managers aims to better guide implementation of genetic approaches into conservation breeding programs.

KEYWORDS
Bellinger River turtle, conservation breeding, *Myuchelys georgesi*, population genetics

1 | INTRODUCTION

1.1 | The six stages of conservation breeding

Conservation breeding programs are a valuable method for managing threatened species (Conway, 2011, Grueber et al., 2019). Such programs can assist biodiversity conservation via scientific research, public education, and as genetic reservoirs for reinforcing dwindling wild populations (Ochoa et al., 2016; Pritchard et al., 2012). High-profile cases of successful conservation breeding and release programs include the black-footed ferret (*Mustela nigripes*) (Wisely et al., 2003), Californian condor (*Gymnogyps californianus*) (Ralls & Ballou, 2004), and the Arabian oryx (*Oryx leucoryx*) (Price, 1989). Despite these examples and other notable successes, including several in Australia (Andrew et al., 2018; Heinsohn et al., 2022; Scheele et al., 2021), conservation breeding programs remain an intensive and expensive management approach (Conde et al., 2011).

Frankham et al. (2010) described a six-stage process of establishing successful conservation breeding and release programs; (1) recognizing decline of the wild population and its genetic consequences; (2) founding one or more captive populations; (3) expanding captive populations to a secure size; (4) managing the captive population over generations; (5) choosing individuals for reintroduction; and (6) managing the translocated population in the wild. Priorities throughout this process include developing husbandry techniques, rapid reproduction, disease mitigation, and genetic management (Frankham et al., 2010). Traditionally, genetic management of captive populations has been based on pedigrees from studbook records and the underlying assumption that founders are neither related nor inbred, which is often not the case (Hogg et al., 2019; Lacy, 1987). This assumption means initial breeding events may result in inadvertent inbreeding and diversity loss for populations already experiencing limited genetic variation (Barrett et al., 2022; Frankham et al., 2017). Advancements in genetic sequencing technologies and bioinformatic tools are making it feasible to integrate molecular data into conservation breeding programs to determine levels of relatedness and genetic diversity metrics.

High-throughput sequencing has resulted in the generation of large amounts of data and the emergence of reference genomes for conservation management. Reference genomes provide data for a range of investigations including designing species-specific microsatellite markers for population analyses, developing targeted single nucleotide polymorphism (SNP) panels, aligning and calling reduced representation sequencing (RRS) data within the same or closely related species, exploratory genome-wide analyses and high-resolution functional gene investigations such as complex immune gene families (Brandies et al., 2019; Galla et al., 2018; Peel et al., 2022). The data output by high-throughput sequencing often requires interpretation from experts in the field of genomics resulting in a "research implementation gap" (Taylor et al., 2017). Translational research is an interdisciplinary approach to conservation that seeks to bridge the gap between scientific knowledge and practical applications (Enquist et al., 2017). To maximize interdisciplinary contributions to conservation breeding programs, there is growing responsibility for scientists to engage with managers directly to implement research findings into management. Here, we present a case study that epitomizes Frankham and colleagues six-stage process of establishing a conservation breeding and reintroduction program and show how the integration of a multidisciplinary approach has benefited a critically endangered turtle species.

1.2 | Our case study species

Turtles are among the most threatened vertebrate taxa globally (Van Dyke et al., 2018). Over 20% of turtle species are listed as Critically Endangered by the International Union for the Conservation of Nature (IUCN; McCallum, 2021), with declines greatly reducing turtles' contributions to ecological processes and food webs (Chessman et al., 2020). Pleurodira, a 200-million-year-old suborder of freshwater turtles found only in Australia, New Guinea, Africa, and South America remains highly under-represented in the peer-reviewed literature. Within this suborder, the Bellinger River turtle (*Myuchelys georgesi*) is a critically endangered species in the family Chelidae. The species is a medium-sized omnivorous turtle with a current known distribution that is restricted to a 60 km range of the Bellinger catchment in north-eastern New South Wales (NSW), Australia (Figure 1A) (Cann et al., 2015; Zhang et al., 2018). The Bellinger catchment and several small freshwater catchments in NSW remain isolated and relatively untouched, which has facilitated unique habitat specialization and

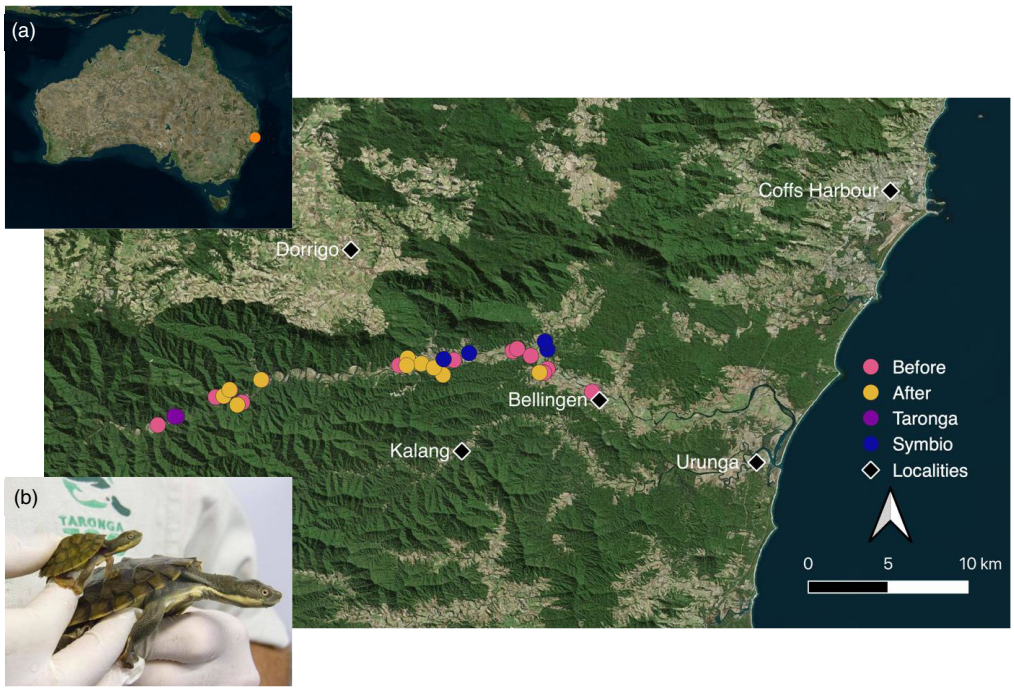


FIGURE 1 (A) An inset of Australia with a map of the Bellinger River basin showing the locations of historic (Before: 2007, pink) and contemporary (After: 2015–2020, yellow) samples, including the founder collection locations for the two conservation breeding programs (Taronga: 2015, purple; and Symbio: 2017, blue) (NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW) (unpublished data) sample points. Note that the sample locations have been obscured as *M. georgesi* is listed as a Category 2 species in the DCCEEW sensitive species data policy. (B) A captive *M. georgesi* hatchling and adult. Photo: Amy Russell.

catchment-specific speciation (Spencer et al., 2014). *M. georgesi* has adapted to up-stream regions of the Bellinger River, preferring deeper waterholes surrounded by bedrock (Spencer et al., 2014). Although a relatively untouched catchment, the only known population is threatened by riparian habitat loss as a result of private landownership, predation by introduced and native predators, reduced water quality, hybridisation due to human mediated dispersal of the Murray River turtle (*Emydura macquarii*) and a novel disease outbreak (Chessman et al., 2020; Georges et al., 2018; Spencer et al., 2014; Zhang et al., 2018).

In 2015, a species-specific nidovirus resulted in the death of more than 90% of individuals, with mortalities occurring mostly among adults (Figure 1B) (Chessman et al., 2020, Zhang et al., 2018). During this time, it is estimated that the population declined from approximately 3000 to less than 150 individuals (Chessman et al., 2020; Spencer et al., 2018). To date, there has been limited evidence of individuals recovering from the disease and no records of breeding in the river since the outbreak. Additionally, knowledge on how the outbreak affected the species' distribution throughout the river remains limited due to accessibility constraints.

Species recovery is currently managed by NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW) and includes a conservation breeding program that commenced in 2015, with a captive colony founded at Western Sydney University that was later relocated to Taronga Conservation Society Australia, Mosman, Australia ($N = 16$ individuals). The Taronga population was founded from emergency intakes, with seven females and nine males collected from two sites in the upper reaches of the Bellinger River where the virus had not yet reached (Figure 1A). This was followed by a second intake in 2017 to Symbio Wildlife Park, Helensburg, Australia ($N = 19$) (Figure 2A). The Symbio population was founded post-virus from six females and 13 males collected from four sites in the lower reaches of the river (Figure 1A).

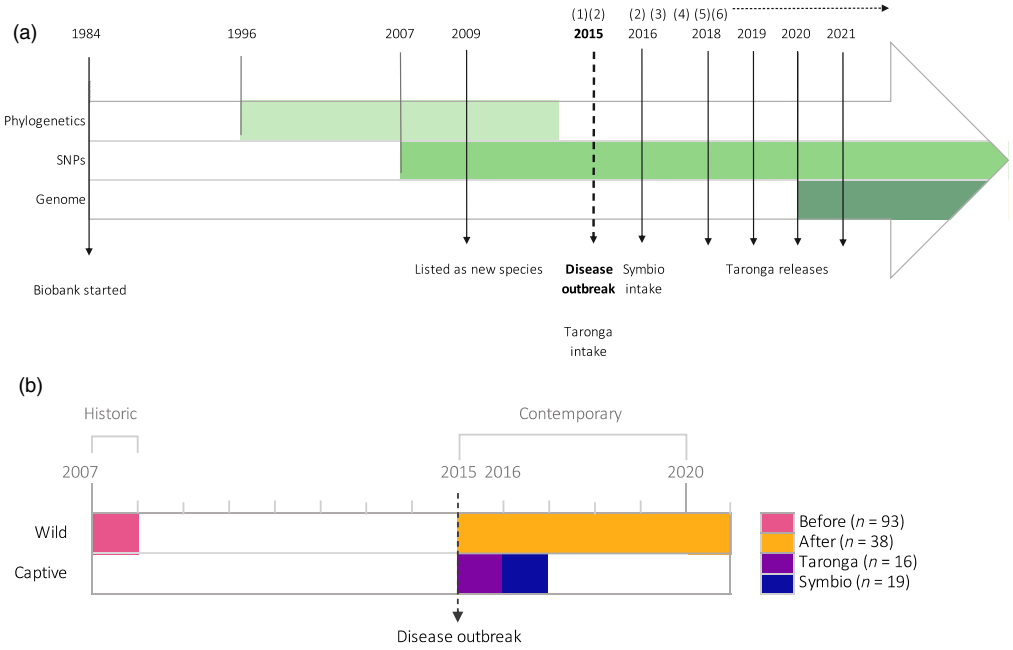


FIGURE 2 (A) Timeline of the collection and implementation of genetic data into *M. georgesi* management including disease outbreaks, founder intakes and captive release events. Numbers in brackets above the timeline indicate completion of Frankham et al. (2010) six-stage process for conservation breeding and reintroduction programs; (1) Recognizing decline of the wild population and its genetic consequences; (2) Founding one or more captive populations; (3) Expanding captive populations to a secure size; (4) Managing the captive population over generations; (5) Choosing individuals for reintroduction; and (6) Managing the reintroduced population in the wild. (B) Historic and contemporary samples used for comparative analyses of wild and captive individuals.

The species recovery program mirrors the principles of Frankham et al.'s six-stage process (Figure 2A) and has developed successful husbandry, breeding, and disease mitigation protocols (Taronga Conservation Society Australia, 2023). While Frankham's stages were written in 2010 for reintroductions, the underlying principles can be applied to various conservation translocation types, including conservation introductions and reinforcements (IUCN/SSC, 2013). Genetic sampling of the species dates back to 1986 (Figure 2A) (Georges & Adams, 1992) but the conservation breeding and release program initially had limited genetic data. This lack of data for genetic management of the species could potentially have long-term implications on the retention of genetic diversity and population viability for the species.

Here, we generate a comprehensive genetic toolkit for the Bellinger River turtle, translating our genetic findings into management recommendations for the conservation breeding and release program (Table 1). To achieve this, we assembled the first chromosome-level genome for the suborder Pleurodira and aligned population genetic data to (1) investigate historic and contemporary diversity and differentiation, (2) identify levels of founder relatedness within and between the Taronga and Symbio populations, and (3) develop an easy-to-follow framework for managers to translate research into management actions.

2 | METHODS

2.1 | Reference genome

Comprehensive details of genome assembly and annotation are provided in the Data S1. In summary, we conducted high molecular weight DNA extractions from the heart tissue of a male *M. georgesi* using the Nanobind Tissue Big DNA kit following the manufacturer's protocol (Circulomics, Pacific Biosciences, California, United States of America). PacBio HiFi Single-Molecule Real-Time (SMRT) bell libraries were sequenced across

TABLE 1 A framework for integrating genetic data into conservation breeding management, mirroring Frankham et al. (2010) six-stage process of establishing successful conservation breeding and reintroduction (or other conservation translocation) program.

Stage	Genetic input	Management outcome	*M. Georgesi* program
(1) Recognizing decline of the wild population and its genetic consequences	• Biobanking • Reference genome generation • DNA sampling of declining population (blood/tissue)	• Setting genetic foundations and acquiring samples for subsequent steps • Baseline analyses investigating wild population genetic diversity, differentiation and temporal changes wrought by the decline	• Tissue samples were collected in 1986 (A. Georges; Figure 2B) and biobanked at The University of Canberra • **A chromosome-level reference genome has been assembled and annotated for downstream analyses** • Commencement of annual surveying and sampling of wild population by NSW DCCEEW • **Baseline analyses of genetic metrics (H_S, H_E, H_O, F_{IS}, A_R, P_A,N_E, F_{ST})**[a]
(2) Founding one or more captive populations	• 20–30 contributing founders • DNA sampling of all founders • DNA sampling of contemporary wild individuals	• **Identifying founder relationships for baseline studbook data** • **Ensuring no hybrids or introgressed individuals in captivity** • **Ensuring captive populations are representative of wild diversity**	• Two captive populations established from 35 individuals (Taronga Conservation Society and Symbio Wildlife Park) • Tissues collected from wild and founding individuals (NSW DCCEEW; Figure 2B) • The founders gathered by NSW DCCEEW and various institutions were sourced from opposite ends of the species distribution and are housed separately • Genetic analyses identified the presence of hybrids which were then removed
(3) Expanding captive populations to a secure size	• Establishing a studbook with known founder relatedness • DNA sampling of each generation • DNA sampling of new founder intakes	• Provides data for stage 4	• Tissue collected from F1 for parentage analyses (Georges, 2020) • Development of studbook • Breeding program successfully increased numbers from 35 to ca. 299 individuals (including releases)
(4) Managing the captive population over generations	• Maintaining a studbook with known founder relatedness • Breeding between captive populations • Introducing genetic material via new founders	• **Maintaining/increasing genetic diversity** • **Minimizing MK**[a] • Monitor for genetic drift • Increasing genetic diversity with new individuals • Ensuring individuals are representative of wild diversity	• Continuation of progeny sampling • **Analyses to monitor genetic metrics (H_S, H_E, H_O, F_{IS}, A_R, P_A, N_E, F_{ST})**[a] • **Implementation of founder relatedness results by breeding individuals with low MK** and integrating genetic data into studbook
(5) Choosing individuals for reintroduction/ release	• Results from analyses in steps 2–4, that is, reintroduction/ translocation cohorts with wide ranging diversity	• Reintroduce/translocate genetically diverse individuals • Insight on where to reintroduce/ translocate based on wild diversity and population structure	• Retrospective integration of genetic diversity results from this study to inform reintroduction/translocation decisions

(Continues)

TABLE 1 (Continued)

Stage	Genetic input	Management outcome	*M. Georgesi* program
(6) Managing the reintroduced (or translocated) population in the wild	• Routine DNA sampling of wild population • Knowledge of captive and wild genetics from earlier stages	• Maintain wild diversity post-release • Ensure wild is not 'swamped' by different genotypes from a single captive population • Ensure captive-released animals are breeding with wild animals and contributing to the next generation	• Annual sampling of wild population by NSW DCCEEW • **Genetic diversity of wild and captive populations undertaken in this study**

Note: Bold text highlights the genetic contributions of our study. The steps in the *Myuchelys georgesi* case study were not undertaken chronologically at each stage, as we have retrospectively integrated genetic inputs into the program. We suggest that other conservation breeding programs make efforts to follow the order we recommend.

[a]NSW DCCEEW New South Wales Department of Climate Change, Energy, the Environment and Water; MK Mean kinship.

Abbreviations: H_S, Standardized heterozygosity; H_E, Expected heterozygosity; H_O, Observed heterozygosity; F_{IS}, Inbreeding coefficient; A_R, Allelic richness; P_A, Private alleles; N_E, Effective population size; F_{ST}, differentiation.

two SMRT cells on the PacBio Sequel II at the Australian Genome Research Facility (Brisbane, Australia). We assembled the HiFi genome using Hifiasm v.0.16.0 (Cheng et al., 2021), and scaffolded it using Hi-C data produced on a Illumina Novaseq 6000 and the YaHS v.1.1 scaffolding pipeline (Zhou et al., 2022). We extracted RNA from brain, liver, and spleen tissue of a female *M. georgesi* using the Qiagen Rneasy Mini Kit (Qiagen, Hidlen, Germany) following the manufacturer's protocol. Transcriptomes were sequenced at the Ramaciotti Centre for Genomics (The University of New South Wales, Sydney, Australia) on an Illumina NovaSeq 6000 S1 flow cell as 100 bp paired-end reads. We annotated the genome with FGENESH++ using a global transcriptome assembly that was generated from the brain, liver, and spleen transcriptomes (Table S1).

2.2 | Genetic analyses for management

Two-week river-wide, randomized, stratified surveys were conducted in April of 2007 and October and November of 2015, 2016, 2019, and 2020 by the NSW DCCEEW. October and November coincided with the beginning of the breeding season, at which time gravidity of adult females could be confirmed. Turtles were identified using scute notching, and their carapace width and length were measured. They were also weighed, bled or had skin biopsied, body condition checked, and swabbed for viral screening (Chessman et al., 2020). DNA samples were collected by extracting blood from the jugular vein or by removing part of the trailing webbing of the clawless toe on the hindfoot (Georges et al., 2018).

Blood and skin biopsies from 166 individuals were collected across 33 sites during surveys (2007, 2015–2020) for the purposes of population genetic analyses (Figures 1A and 2A, Table S5). The population genetics samples were stored in 75% ethanol at −20°C in the University of Canberra Wildlife Tissue Collection (GenBank UC < Aus>). As described in Georges et al. (2018), samples were sequenced over multiple runs using high coverage DArTseq™ (Diversity Arrays Technology PL, Canberra, Australia), a form of RRS. We aligned raw DArT sequences to the repeat masked genome generated in this study (Figure S2, Table S2) and called SNPs using Stacks v2.61 (Catchen et al., 2013; Rochette et al., 2019). The 'populations' module was then run with the following parameters: minimum samples per population 30% (−r 0.3); minimum minor allele frequency (MAF) of 0.01 (--min_maf 0.01); and --write_random_snp.

To partition our data, we investigated genetic clusters using a variational Bayesian framework in fast*STRUCTURE* v1.0 (Raj et al., 2014) and visualized the results using DISTRUCT v1.1 (Rosenberg, 2004). $K = 1$–4 clusters were tested based on the demographic characteristics of the species including population size and range (Figure S3). We used the fast*STRUCTURE* "chooseK.py" script to decide the optimal K. As no genetic clusters were detected (optimal value of K was one), we partitioned the SNP dataset output by the 'populations' module according to collection time relative to the disease outbreak and current location (Figure 2B). These predetermined groups were: (1) wild individuals sampled before the disease outbreak in 2007, $N = 92$ (hereafter "Before"); (2) wild individuals sampled after the disease outbreak between 2015 and 2020, $N = 38$ (hereafter "After"); (3) Taronga founders sampled in 2015, $N = 16$; and (4) Symbio founders sampled in 2017, $N = 19$. Variant filtering was carried out on three datasets: (i) all groups consisting of the four predetermined groups

($N = 166$), (ii) wild groups only (Before and After; $N = 131$), and (iii) captive populations only (Taronga and Symbio; $N = 35$). The SNP datasets were filtered on minimum average read depth (>2.5×), coverage difference between the reference and alternate alleles (<80%), call rate (>75%), retention of loci with heterozygosity <80%, and reproducibility (>90%) calculated using 51 technical replicates performed by DArT. Where samples were grouped across different sequencing plates, we found no evidence to suggest that batch effects influenced our results (Figure S1).

To investigate variation in the 'wild' and 'all groups' datasets, we undertook exploratory principal coordinates analyses (PCoA) using Euclidean distance via the R v4.3.0 package dartR v1.9 (Gruber et al., 2018; R Core Team, 2023) and visually inferred putative genetic differentiation between groups using eigenvalues in adegenet v2.1.3 (Jombart, 2008). We also applied a PCA to the dataset using the glPca() function in adegenet. We calculated pairwise fixation indices (F_{ST}) between all groups using hierfstat v0.5–11 (Goudet 2005). To identify relationships between geographic distance and genetic distance we used the Before dataset to perform an isolation-by-distance (IBD) mantel test in dartR using the gl.ibd() function with 999 permutations.

We calculated standardized heterozygosity (H_s) using the genhet function in R (Coulon, 2010) for all groups (where 1 is the average and so a value greater than 1 is more diverse than average); observed (H_o) and expected heterozygosity (H_E) using GenAlEx v6.5 (Peakall & Smouse, 2006) and visualized individual H_o distributions by group using the boxplot() functions in R. We calculated autosomal H_o and H_E (Schmidt et al., 2021) by re-running the Stacks 'populations' module with the parameter: minimum samples per population 75% (−r 0.75), and without the parameters: minimum minor allele frequency (MAF) of 0.01 (--min_maf 0.01); and --write_random_snp to retain both variant and invariant loci (Schmidt et al., 2021). We filtered the Stacks output on minimum average read depth (>2.5×), coverage difference between the reference and alternate alleles (<80%), call rate (>75%), and reproducibility (>90%) calculated using 51 technical replicates performed by DArT. Loci on sex chromosomes (scaffold 4) were removed from the dataset (Martinez et al., 2008). The resulting 11,208 loci (at both variant and invariant sites) were used to calculate autosomal H_o and H_E using GenAlEx v6.5 (Peakall & Smouse, 2006).

We performed a t-test to test whether H_s was significantly different between wild groups before and after the disease and used the Bartlett's test of homogeneity of variances base function to test whether individual observed heterozygosity was significantly different from expected heterozygosity in R. We calculated inbreeding coefficients (F_{IS}) and the associated 95% CI using the diveRsity v1.9.9 package (Keenan et al., 2013) and PopGenReport v3.0.7 (Adamack & Gruber, 2014) to calculate allelic richness (A_R) in R. We calculated the number of private alleles in each group compared to all other groups and pairwise private alleles between groups using the gl.report.pa() function in the dartR package. To calculate molecular relatedness, we ran simulations in COANCESTRY v1.0.1.10 (Wang, 2011) to determine the most appropriate moments estimator as per Hogg et al. (2019). We selected TrioML for final analyses. We set COANCESTRY parameters to account for inbreeding, with the number of reference individuals and bootstrapping samples set to 100 for all groups and between captive groups. We calculated mean kinship (MK) by dividing the TrioML value by two, representing the average relationship of each animal to all others within the sample set. We estimated MK for individuals within each group (MK_{WITHIN}) and between captive groups ($MK_{BETWEEN}$) using the captive dataset. We calculated effective population size (N_E) for wild groups using NeEstimator v2.1 (Do et al., 2014) with values reported for the no singleton alleles analysis and the associated jack-knifed 95% confidence intervals (Jones et al., 2016). We excluded captive groups from N_E analyses due to small sample sizes resulting in infinite confidence intervals.

3 | RESULTS

3.1 | Reference genome

The final genome assembly was 2.0 Gb in size, consisted of 129 scaffolds, had a contig N50 of 56.12 Mb, and scaffold N50 of 123.4 Mb (Figure S2A, Table S2). BUSCOv5.2.2 identified 95.4% complete vertebrata genes, 94.9% of which were single copy and 0.5% were duplicated, 2.1% were fragmented, and 2.5% were missing (Figure S2A, Table S2). The genome statistics and distinct chromosome-length scaffolds (Figure S2B) confirm chromosome level-completeness. We used this genome to align and call variants using the population genetics data. Comprehensive details of genome assembly and annotation are provided in the Supplementary Material (Figure S2).

3.2 | Genetic analyses for management

We aligned the high density DArTseq data to the reference genome. Our initial analyses showed evidence of an *E. macquarii* ($N = 1$), *M. georgesi* and *E. macquarii* hybrid

($N = 1$), and introgressed ($N = 2$) individuals in the wild groups (Figure S4) (Georges et al., 2018). In accordance with earlier findings, we did not detect hybrids in either captive population having been identified and removed when the populations were founded (Georges et al., 2018). The *E. macquarii*, *M. georgesi* and *E. macquarii* hybrid, and introgressed individuals were removed from the dataset and SNPs were re-called. The re-called dataset yielded a Stacks output of 2172 SNPs. Refiltering on SNP calls for each group in R resulted in reduced representation datasets of 460 genome-wide SNPs in the 'all groups' dataset, 473 genome-wide SNPs in the 'wild' dataset, and 227 genome-wide SNPs in the 'captive' dataset. Our exploratory PCoA using the wild only dataset revealed minor levels of genetic structuring between Before and After individuals, with up to 3.5% of variance explained by PCo I and 2.5% by PCo II (Figure 3A). The addition of captive individuals in the second PCoA, revealed clustering of captive groups with Before individuals (Figure 3B). Up to 5.3% of the variation was explained by PCoI, and PCo II (3.7% variation) with both axes primarily separating out After individuals (Figure 3B). An almost identical clustering result was seen when a PCA was applied to the dataset (data not shown).

We found no notable differences in H_s between captive and wild groups (Table 2A) and no significant differences in the wild across time ($t = -0.902$, $df = 128$, $p = 0.369$). SNP and autosomal H_O and H_E was similar across all groups (Table 2A, Figure S5) with significantly higher observed than expected heterozygosity across all groups, indicating an excess of heterozygotes. There were no statistically significant levels of inbreeding (F_{IS}) observed as confidence intervals encompassed zero for all groups (Table 2A). A_R was also similar with no distinct differences across groups (Table 2A). P_A were only present in the Before group when comparing across all groups (Table 2A). Pairwise P_A numbers were consistently higher for Before and lowest for Taronga and Symbio, respectively (Table S3). MK_{WITHIN} ranged from 0.019 ± 0.048 to 0.048 ± 0.150, with the highest value observed in the After population (Table 2A). $MK_{BETWEEN}$ for Taronga and Symbio using the captive dataset was 0.012 ± 0.040, where 0.1250 is the equivalent of a half-sibling relationship. N_E estimates for Before and After were 148.9 (120.4–191.8) and 11.9 (8.2–17.1) respectively (Table 2A). F_{ST} values between all groups ranged from 0.005 (0.001–0.009) to 0.027 (0.020–0.034), with statistically significant F_{ST} observed between the Before and After groups (Table 2B). Our IBD mantel test found no correlation between geographic distance and genetic distance in the species ($r = 0.364$, $p = 0.258$).

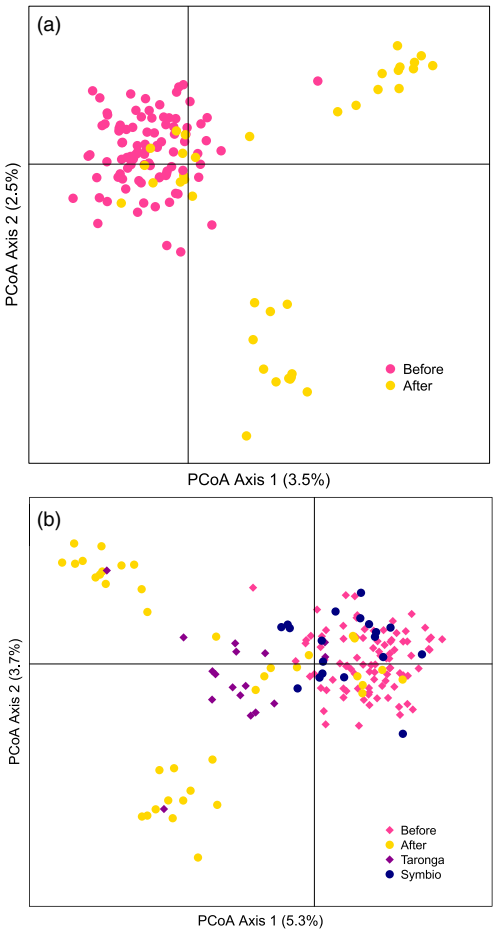


FIGURE 3 (A) Principal coordinates analysis (PCoA) of wild individuals before and after the disease outbreak ($N = 131$) using 474 genome-wide SNP markers. (B) PCoA of genome-wide diversity of all individuals ($N = 166$) using 460 SNP markers.

4 | DISCUSSION

4.1 | A worked example of how conservation managers can apply the framework

Please refer to Table 1 for a summary of the framework for integrating genetic data into conservation breeding management, particularly in relation to the genetic input, management outcomes and how this applies to our case study.

TABLE 2 (A) Population genetic indices of our sample groups, including standardized (H_S), observed (H_O), and expected (H_E) SNP and autosomal heterozygosity, inbreeding coefficient (F_{IS}), allelic richness (A_R), private alleles (P_A), mean kinship within populations (MK_{WITHIN}), and effective population size (N_E).

Group	*n*	H_S (±SD) SNP	H_O (±SE) SNP	H_E (±SE) SNP	H_O (±SE) Autosomal	H_E (±SE) Autosomal	F_{IS} (95% CI)	A_R	P_A	MK_{WITHIN} (±SD)	N_E (95% CI)
Before	92	1.007 (0.097)	0.290 (0.009)*	0.280 (0.008)	0.0180 (0.0007)*	0.0148 (0.0005)	−0.035 (−0.051 to 0.020)	1.727	11	0.019 (0.047)	148.9 (120.4–191.8)
After	38	0.990 (0.097)	0.286 (0.010)*	0.262 (0.008)	0.0170 (0.0006)*	0.0138 (0.0005)	−0.094 (−0.125 to 0.064)	1.688	0	0.048 (0.150)	11.9 (8.2–17.1)
Symbio	19	0.969 (0.102)	0.279 (0.011)*	0.252 (0.009)	0.0174 (0.0007)*	0.0136 (0.0005)	−0.181 (−0.149 to 0.075)	1.669	0	0.023 (0.072)	—
Taronga	17	1.009 (0.098)	0.293 (0.012)*	0.248 (0.009)	0.0165 (0.0007)*	0.0130 (0.0005)	−0.108 (−0.253 to 0.121)	1.658	0	0.027 (0.076)	—

	Before	After	Symbio
After	0.026 (0.020–0.031)		
Symbio	0.005 (0.001–0.009)	0.027 (0.020–0.034)	
Taronga	0.010 (0.004–0.016)	0.002 (−0.002–0.006)	0.019 (0.011–0.027)

Note: Significantly higher HO to HE denoted by "*". (B) Population differentiation (FST) including 95% CI between all groups.
Abbreviations: CI, 95% lower and upper confidence intervals; SD, standard deviation.

4.1.1 | Stage 1—Recognizing decline of the wild population and its genetic consequences

"*When recognising declines in wild populations, the collection and preservation of DNA samples, such as blood and tissue, in a biobank can provide essential genetic data for future research*" (Frankham et al., 2010). Through collection and preservation of DNA, we can also identify genetic consequences resulting from threatening processes through analysis of temporal data, as well as sample provision for development of a reference genome for neutral and adaptive genomic investigations (Sunde et al., 2022).

The reference genome for *M. georgesi* was created using tissue collected and stored in a biobank from 1986. The sampling by managers in 2007 and 2015–2022 (Figure 2B) has provided population genetic data for researchers to investigate the genetic consequences of the nidovirus outbreak and other threatening processes, including evidence of shifts in wild genetic diversity since the disease event (Table 2A, Figure 3A). We developed and utilized our reference genome during establishment stages 2 and 6 for Symbio and Taronga, respectively. If financial resources and collaborative opportunities are available, we recommend that a reference genome be developed during stage 1 for conservation breeding programs so genetic output can be utilized as early as possible and to improve the reliability of RRS data variant calls to improve downstream inferences (Shafer et al., 2017; Torkamaneh et al., 2016; Wright et al., 2019). We also recommend samples from the declining wild population be collected to provide baseline genetic data and to readily assess genetic consequences of population declines. This may encompass decreases in diversity, the risk of bottlenecks, the potential for inbreeding depression, and the occurrence of genetic drift. Addressing these issues early is crucial to prevent the need for more extensive interventions due to a potential delayed response in observable genetic changes.

We examined population genetic data from samples that were previously sequenced for earlier analyses (Georges et al., 2018) (Figure 2B). Using our reference genome and set of reliable genome-wide SNP markers, we found no significant differences in H_s between all groups ($p = 0.369$) indicating consistent levels of heterozygosity in the population since the nidovirus outbreak. SNP and autosomal observed heterozygosity were significantly higher than expected heterozygosity in all groups (Table 2A), suggesting that the population bottleneck caused by the mass mortality event has either not resulted in a corresponding genetic bottleneck or it is a signature of a bottleneck-induced heterozygosity excess, likely to be detectable for a few generations until a new equilibrium between mutation and drift is reached (Cornuet & Luikart, 1996). This is also suggested by the large drop in N_E. The relatively low genome-wide diversity compared to other species also suggests historical bottlenecks have already occurred in this species (Georges, 2020). It has also been suggested that associative overdominance (neutral loci becoming effectively

over-dominant as a result of disequilibrium with a locus under selection) may conserve genetic variation in small populations compared to expectations from neutral theory (Frydenberg, 1963; Gilligan et al., 2005; Rumball et al., 1994; Schou et al., 2017), potentially contributing to the significantly higher levels of observed to expected heterozygosity in all groups however, long-term monitoring is needed to confirm this. Notably higher SNP heterozygosity compared to autosomal heterozygosity estimates are due to the inclusion of invariant sites and the species small population size (Schmidt et al., 2021), with autosomal heterozygosity estimates likely providing a more accurate representation of low genome-wide diversity in the species (Schmidt et al., 2021). A_R was consistent across groups, with slightly lower values observed for the After, Symbio, and Taronga groups compared to Before (Table 2A). The lower number of unique P_A in the After group compared to Before highlights potential losses in alleles overtime, although discrepancies in sample size may not have captured the range of alleles currently present in the wild (Table S3). Although alleles present in the After group do not appear to be captured in captivity, each captive population contains multiple alleles not observed in the current wild population (Table S3). As there has been minimal evidence of wild clutches hatching since the outbreak, the After group may still reflect the diversity of the larger pre-disease population suggesting that genetic consequences of the outbreak may not be evident until post-disease F1 individuals can be analyzed. We have advised species managers that continual monitoring once wild individuals start reproducing will be crucial in identifying long-term trends.

In this case study, we demonstrate the value of biobanking during the early stages of a wild population decline, and we advocate for continual monitoring in threatened species including DNA samples collection and biobanking to capture temporal trends in genetic diversity.

4.1.2 | Stage 2—Founding one or more captive populations

"*When founding one or more captive populations, a fully representative sample that encompasses wild diversity is needed to maximise captive population viability*" (Frankham et al., 2010). Frankham and colleagues proposed that 20–30 contributing founders are sufficient to create a genetically diverse population that is representative of wild diversity, although molecular analyses are needed to confirm this hypothesis. DNA samples from, (1) all founders are needed and, if feasible, (2) contemporary samples collected across the wild population should be sequenced. The data generated from these samples can be used in stages 3 and 4 for, (1) diversity and founder kinship analyses that can inform management decisions to maximize diversity in captivity, prevent inbreeding and to detect hybrids, or introgressed, individuals among founders, and (2) contemporary wild samples that provide information to ensure that the genetic diversity of the wild population is reflected in captive individuals.

In line with stage 2 management outcomes, we observed minor but non-significant shifts in genetic structure since the outbreak, where Taronga and Symbio are shown to be most representative of historical diversity (Before) (Figure 3B). The clustering of captive groups with Before in the PCoA suggests that Taronga and Symbio represent the genetic profile of the once larger population which may be useful in reinforcing the current wild population. The small number of Taronga individuals that clustered with after suggests greater representation of current wild variation within Taronga that is consistent with the low F_{ST} values between the two groups (Table 2B). However, as only a small amount of variation is explained by the PCoA axes (<4%), long-term analyses of the offspring of the outbreak survivors will be useful in confirming these trends. MK_{WITHIN} in the wild is higher after the disease outbreak with Taronga and Symbio falling between historical and contemporary levels. Populations that sustain high MK_{WITHIN} over generations are expected to experience more rapid changes in allele frequencies and lower adaptation potential in future (Frankham, 1996). In this instance, the conservation breeding program will play a crucial role in strengthening the wild population post-bottleneck by providing opportunities to mate with unrelated individuals, mitigating future inbreeding.

In our case study, molecular analyses to assess genetic representation of the captive populations was conducted post-founding due to a lack of genetic resources at the time. Ideally, stage 2 should be implemented at founding to quickly address genetic concerns prior to breeding. In situations where genetic analysis cannot occur at the time of founding, we recommend collection of DNA from all founders for future analysis.

4.1.3 | Stage 3—Expanding captive populations to a secure size

"*During the expansion phase of conservation breeding, priority is on rapid population growth rather than intense genetic management*" (Frankham et al., 2010). During Stage 3, maintaining an accurate studbook is crucial but may not be effective enough to mitigate long-term

founder effects and prevent inbreeding depression when assuming founders are unrelated (Ivy et al., 2009; Ivy & Lacy, 2010). During this stage, DNA samples from each generation and new intakes are essential for molecular identification of relatives to prevent long-term founder effects and inbreeding in subsequent stages (Hogg et al., 2019). This is important in cases where paternity determination can be difficult, including group-housed enclosures, in species where females retain sperm (Gist & Jones, 1987; Sever & Hamlett, 2002), and when mixed parental clutches are present. This information allows for more informed genetic selection of mates and alleviates the impact of founder effects in subsequent generations. Additionally, this information allows managers to monitor novel genetics introduced into the population through new intakes and provides data for long-term monitoring of genetic drift.

Through expert collaboration and development of effective breeding protocols (Taronga Conservation Society Australia, 2023), the number of *M. georgesi* individuals in captivity has rapidly grown from 35 to ca. Two hundred and ninety nine individuals in 80 years, increasing the global population size from <150 to approximately 450 individuals. Throughout Stage 3, a studbook has been maintained to inform breeding, established before the availability of genetic data. Before our study, the DNA sequences of Taronga's founders and offspring were used for internal parentage analysis to inform the studbook (Georges, 2020). As the turtles are typically housed in groups where females have been observed to retain sperm, determining paternity through traditional means may sometimes be challenging with the absence of molecular information. Due to low numbers in the wild, there have been no new founder intakes to the captive program. In accordance with the species' conservation action plan (Jakob-Hoff et al., 2017), we anticipate DNA from new wild individuals adopted into the breeding program will be analyzed and integrated into the studbook (Table S4, Table S5).

Our case study and previous work demonstrates that rapid population growth can go hand in hand with the integration of traditional pedigree and molecular genetic data.

4.1.4 | Stage 4—Managing captive populations over generations

"*Loss of genetic diversity and inbreeding is exacerbated over generations in small captive populations*" (Frankham, 1995). *When managing captive populations over generations, focus shifts from rapid reproduction in Stage 3 to mitigating genetic issues in stage 4*" (Frankham et al., 2010). Samples collected in Stage 3 can be used to explore a standardized set of genetically measured Essential Biodiversity Variables in stage 4 (H_S, H_E, H_O, F_{IS}, A_R, P_A, N_E and F_{ST}) to provide comparable data for monitoring and management (Hoban et al., 2022). It may be preferable to undertake these analyses during Stage 3 to identify genetic changes early on and ascertain if genetic interventions are required, such as the introduction of new individuals to increase genetic diversity, or breeding between populations to reduce inbreeding (Hoffmann et al., 2021, Kinghorn and Kinghorn, 2021). In cases where founder relationships are unclear during initial breeding, it is essential to retrospectively incorporate these data to help inform future breeding decisions (Hogg et al., 2019), and to apply genetic principles such as not breeding individuals captured close together, until genetic data can be incorporated.

After population expansion and establishment of a studbook in Stage 3, *M. georgesi* managers sought genetic expertise for continued genetic monitoring of captive individuals. This was an opportune time as individuals only had one generation of captive ancestry, allowing our data to be proactively implemented to mediate future genetic issues. Given that the lifespan of *M. georgesi* likely surpasses 30 years and individuals typically do not reach sexual maturity until approximately 6–12 years of age, this may prove more challenging in species with shorter lifespans and short generation times. With access to genetic data, our analyses revealed consistent genetic diversity metrics in both Taronga and Symbio, with Taronga showing slightly higher diversity metrics throughout all analyses (Table 2A,B).

MK values are useful for determining how related individuals are within, and between, populations and for indicating, which animals should be considered for breeding/translocation to minimize relatedness and maximize genetic diversity (Frankham et al., 2017). When choosing breeding pairs, individuals with high MK estimates should not be paired, and breeding of their progeny should be carefully considered. To assist the stage 4 outcome of managing captive populations over generations using our MK analyses (Table S4, Table S5), we suggest breeding individuals that are on average less related to each other within Taronga and Symbio to prevent founder effects and inbreeding depression (Frankham, 2008; Frankham et al., 2010; Lacy, 1987). As we observed lower levels of MK between ($MK_{BETWEEN}$) Taronga and Symbio compared to within (MK_{WITHIN}) (Table 2A, Table S4, Table S5), we have also suggested to the recovery team and captive managers that breeding between Taronga and Symbio (when sexually mature) could result in a decrease in average kinship, a reduction in inbred individuals, and an increase in diversity

(Hoffmann et al., 2021, Kinghorn and Kinghorn, 2021), as observed in other species (del Mar Ortega-Villaizan et al., 2011; Thavornkanlapachai et al., 2021). Given that only Taronga individuals have reached reproductive maturity and animals born in captivity (F1) have yet to breed, genetic monitoring of each generation is necessary to detect any genetic drift in captivity (Gilligan & Frankham, 2003) in addition to diversity and MK monitoring. This is being implemented through genetic sampling of captive born individuals.

We demonstrate in our case study how the use of genome-wide SNP data can provide fine-scale insights to support mitigation of genetic issues in captive populations. The generation of a reference genome in this study will also provide support for future investigations into adaptive potential including immune gene diversity.

4.1.5 | Stage 5—Choosing individuals for reintroduction/release

To choose individuals most suitable for release, genetic information from stages 2–4 can be used to select individuals with wide-ranging levels of differentiation and/or low levels of relatedness. Before translocating captive individuals, it is essential to understand the genetic metrics of both the wild and captive population to prevent the introduction of identical or closely related genotypes from captivity into the wild, particularly when there may be more diverse individuals available in captivity (Hogg et al., 2020). Analyses of wild populations in stages 1–2 can be used to inform release site selection by releasing individuals at sites where genetic differentiation within a species or population is high and inbreeding or outbreeding depression is unlikely (Grueber et al., 2018; Nistelberger et al., 2023). We recommend performing this genetic admixture to maximize diversity in the wild but using populations that have exchanged genes within the last 500 years and where there is little evidence of local adaptation to minimize the risk of outbreeding depression (Frankham et al., 2011).

For *M. georgesi*, the availability of contemporary wild samples meant that we could identify that the combined captive populations are representative of historical wild diversity before the outbreak (Figure 3B) (stage 2). As such, releases from both Taronga and Symbio will be essential in supplementing pre-disease diversity into the contemporary wild population. Additionally, our MK estimates provide data for breeding decisions in the captive breeding program and minimizing the release of inbred individuals into the wild (Table S4, Table S5). To date, site selection for release has not been genetically informed due to insufficient data and has instead been selected based on ease of access and ability to obtain landholder approvals.

Our case study supports the use of genetically informed translocations by providing insight on the genetics at different sites throughout the river (Figures 1 and 3B). For example, the release of Taronga and eventually Symbio individuals at the opposing end of the river from their collection site (Figure 1A) could facilitate admixture with genotypes in lower and upper regions, respectively.

4.1.6 | Stage 6— Managing the reintroduced population in the wild

Due to persistent threats in their natural environment, reintroduction of certain species or reinforcement of wild populations might not be possible for an extended period. However, over time, captive populations may undergo genetic changes that make them better suited to their captive conditions but less adapted to the challenges of the wild (Christie et al., 2012; Frankham, 2008). This adaptation to captivity can result in reduced fitness in the wild due to a loss of genetic diversity, the accumulation of deleterious mutations, or the fixation of alleles that are advantageous in captivity but not in the wild (Frankham et al., 2017; Lacy, 1987). To minimize genetic adaptation to captivity and ensure long-term viability of the remaining wild population (Crates et al., 2023; Frankham, 2008), it is commonly suggested that releases be carried out within a few generations (Williams & Hoffman, 2009). To evaluate the effects of conservation translocations on genetic diversity, it is crucial to have ongoing monitoring of the reintroduced or reinforced population, at least until the population is self-sustaining with an improved conservation status (IUCN/SSC, 2013). Apart from techniques like radiotracking, this requires the collection of DNA samples from the wild population over numerous generations and analyzing them for any changes. For example, this can be achieved by mirroring the temporal comparative analyses in this study (Table 2A,B).

It is recommended that 1–10 individuals every 2–3 generations is sufficient in increasing genetic diversity (Allendorf, 1983; Backus et al., 1995; Lacy, 1987; Lande & Barrowdough, 1987). Prior to reinforcing the wild population, the *M. georgesi* recovery program undertook measures including predator management, habitat restoration, water quality assessments, and community engagement to minimize environmental stressors (Jakob-Hoff et al., 2017). The program has conducted four rounds of reinforcements by releasing 82 F1 juveniles of unknown sex into the Bellinger River (Figure 2A), significantly boosting wild population numbers.

To evaluate the short-term survival and movement of the released individuals, a select number have been radio tracked. Additionally, annual surveys are conducted to sample animals, with DNA biobanked for future sequencing. Given that genetic changes may not be immediately apparent in a long-lived species, long-term monitoring with a focus on sampling new and juvenile individuals, is critical for detecting alterations in the species. Long-term genetic monitoring will also play a critical role monitoring levels of hybridisation and introgression which simultaneously threatens the persistence and locally adapted genetic identity of the species (Georges et al., 2018).

4.2 | Integration of genetic data into management actions

In combination with husbandry, breeding, and disease mitigation, incorporating genetic data throughout conservation breeding establishment and management is valuable for maximizing long-term viability in captivity and to guide conservation translocations to the wild. Here, we have provided a worked example of geneticists working with an established breeding program to answer genetic questions posed by managers using DNA samples they provided. We communicated our findings to managers during scheduled meetings and in-person workshops where our recommendations have been integrated into the species Conservation Action Plan (Jakob-Hoff, *unpublished*).

In summary, for the Bellinger River turtle we recommend that genetic diversity be maintained, or potentially increased, by:

i. Breeding individuals that are on average less related to each other within Taronga and Symbio to prevent founder effects and inbreeding depression.
ii. Breeding between Taronga and Symbio to minimize founder effects and inbreeding depression.
iii. Encouraging translocations from both Taronga and Symbio to ensure representation of genotypes from both captive populations in the wild.
iv. Alternating or varying release locations during each release round to facilitate genetic admixture between captive bred and wild individuals as there is some genetic differentiation between the wild and captive populations (Figure 3).

Here we aimed to provide a generalized checklist based on our own example of genetic data integration, which can be adapted for other conservation breeding programs. Integration of genetics into management activities may need to occur retrospectively depending on the establishment stage of the conservation breeding program when the genetic data is generated (Table 1). Instigating discussions between scientists and managers can facilitate productive dialogue, allowing for a better comprehension and adaptation of each other's work and tools (Hogg et al., 2017). This cyclical approach leads to ongoing improvement and enhancement of conservation strategies.

4.3 | Future directions

As emerging infectious diseases are causing rises in extinction risk (Blehert et al., 2009; Piotrowski et al., 2004), our capacity to understand genetics and genomics is also increasing. The DNA of disease-susceptible species provides valuable insight on species resilience, with genomics giving us the tools to unlock the answers. For *M. georgesi*, as translocation candidates are juveniles, there remains an unknown risk that once individuals mature, they will succumb to the virus in the wild, compromising the translocation program. Another unknown threat is whether the few adult survivors of the virus are genetically predisposed to resist the virus, or they merely avoided exposure (Zhang et al., 2018). To answer these questions, our high-throughput sequencing efforts and chromosome-level genome provides a valuable genomic tool for future functional gene research on *M. georgesi* and other Australian turtle species. Beyond neutral diversity investigations, high-throughput and genome-wide sequencing provides high-resolution data for immune-gene investigations, that rely on high-quality assemblies and genome-wide data (Peel et al., 2022). Each year the cost of sequencing and bioinformatic analyses becomes cheaper and more streamlined permitting studies like ours (Wright et al., 2019). Once streamlined sampling and basic genetic management become conventional practise in breeding programs, genetic research can progress towards higher resolution conservation efforts aimed at preserving the adaptive potential and functional diversity within captive programs.

ACKNOWLEDGMENTS

We acknowledge Gumbaynggirr people and their ongoing connection to Country where the turtles live and pay respect to their Elders past and present. This research was made possible by the NSW DCCEEW Saving our Species program, Taronga Conservation Society, Symbio Wildlife Park, and the NCRIS funded Bioplatforms Australia Threatened Species Initiative. Computing resources were provided by the Pawsey Supercomputing Research Centre; Bioplatforms Australia; Galaxy

Australia; and the University of Sydney's High Performance Computing cluster Artemis. Thanks to the Australian National University's Biomolecular Resource Facility and Australasian Wildlife Genomics Group for assistance with extractions and laboratory work. Thanks also to the University of Canberra Wildlife Tissue Collection (UC < Aus>) for samples and data. Funding for this work was provided by the Australian Government's Bushfire Recovery Scheme (ERF-WRR2-020) to the Threatened Species Initiative, the ARC Centre for Excellence for Innovations in Peptide and Protein Science (CE200100012), and HVN scholarship funding was provided by the NSW DCCEEW.

CONFLICT OF INTEREST STATEMENT

This submission is the original work of all authors. All other work is acknowledged in the manuscript. The authors declare they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The male *M. georgesi* reference genome assembly and all raw sequencing reads including the 3-tissue transcriptome RNA-seq reads are available from NCBI under BioProject PRJNA1003540.

ETHICS STATEMENT

Collection of samples was conducted in accordance with the conditions of NSW DCCEEW Animal Ethics Committee (AEC151201-3, AEC160503-01 and AEC180904-5) and Scientific Licenses (MWL00102467, SL101672 and SL10255), and the Taronga Conservation Society Australia Scientific License SL101204.

ORCID

Carolyn J. Hogg https://orcid.org/0000-0002-6328-398X

REFERENCES

Adamack, A. T., & Gruber, B. (2014). PopGenReport: Simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution*, *5*, 384–387.

Allendorf, F. W. (1983). Isolation, gene flow and genetic differentiation among populations. *Genetics and Conservation*, *18*, 51–65.

Andrew, P., Cogger, H., Driscoll, D., Flakus, S., Harlow, P., Maple, D., Misso, M., Pink, C., Retallick, K., & Rose, K. (2018). Somewhat saved: A captive breeding programme for two endemic Christmas Island lizard species, now extinct in the wild. *Oryx*, *52*, 171–174.

Backus, V. L., Bryant, E. H., Hughes, C. R., & Meffert, L. M. (1995). Effect of migration or inbreeding followed by selection on low-founder-number populations: Implications for captive breeding programs. *Conservation Biology*, *9*, 1216–1224.

Barrett, K. G., Amaral, G., Elphinstone, M., Mcadie, M. L., Davis, C. S., Janes, J. K., Carnio, J., Moehrenschlager, A., & Gorrell, J. C. (2022). Genetic management on the brink of extinction: Sequencing microsatellites does not improve estimates of inbreeding in wild and captive Vancouver Island marmots (Marmota vancouverensis). *Conservation Genetics*, *23*, 417–428.

Blehert, D. S., Hicks, A. C., Behr, M., Meteyer, C. U., Berlowski-Zier, B. M., Buckles, E. L., Coleman, J. T., Darling, S. R., Gargas, A., & Niver, R. (2009). Bat white-nose syndrome: An emerging fungal pathogen? *Science*, *323*, 227.

Brandies, P., Peel, E., Hogg, C. J., & Belov, K. (2019). The value of reference genomes in the conservation of threatened species. *Genes*, *10*, 846.

Cann, J., Spencer, R.-J., Welsh, M., & Georges, A. (2015). Myuchelys georgesi (Cann 1997)—Bellinger River Turtle. Conservation Biology of Freshwater Turtles and Tortoises. http://georges.biomatix.org/storage/app/uploads/public/58c/e5f/ce6/58ce5fce69444270331471.pdf

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.

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*, 170–175.

Chessman, B. C., Mcgilvray, G., Ruming, S., Jones, H. A., Petrov, K., Fielder, D. P., Spencer, R. J., & Georges, A. (2020). On a razor's edge: Status and prospects of the critically endangered Bellinger River snapping turtle, on a razor's edge: Status and prospects of the critically endangered Bellinger River snapping turtle,Myuchelys georgesi, Myuchelys georgesi. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *30*, 586–600.

Christie, M. R., Marine, M. L., French, R. A., & Blouin, M. S. (2012). Genetic adaptation to captivity can occur in a single generation. *Proceedings of the National Academy of Sciences*, *109*, 238–242.

Conde, D. A., Flesness, N., Colchero, F., Jones, O. R., & Scheuerlein, A. (2011). An emerging role of zoos to conserve biodiversity. *Science*, *331*, 1390–1391.

Conway, W. G. (2011). *Buying time for wild animals with zoos*. Wiley Online Library, Buying time for wild animals with zoos.

Cornuet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, *144*, 2001–2014.

Coulon, A. (2010). GENHET: An easy-to-use R function to estimate individual heterozygosity. *Molecular Ecology Resources*, *10*, 167–169.

Crates, R., Stojanovic, D., & Heinsohn, R. (2023). The phenotypic costs of captivity. *Biological Reviews*, *98*, 434–449.

Do, C., Waples, R. S., Peel, D., Macbeth, G., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, *14*, 209–214.

Enquist, C. A., Jackson, S. T., Garfin, G. M., Davis, F. W., Gerber, L. R., Littell, J. A., Tank, J. L., Terando, A. J., Wall, T. U., & Halpern, B. (2017). Foundations of translational ecology. *Frontiers in Ecology and the Environment*, *15*, 541–550.

Frankham, R. (1995). Conservation genetics. *Annual Review of Genetics*, *29*, 305–327.

Frankham, R. (1996). Relationship of genetic variation to population size in wildlife. *Conservation Biology*, *10*, 1500–1508.

Frankham, R. (2008). Genetic adaptation to captivity in species conservation programs. *Molecular Ecology*, *17*, 325–333.

Frankham, R., Ballou, J. D., Eldridge, M. D., Lacy, R. C., Ralls, K., Dudash, M. R., & Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology*, *25*, 465–475.

Frankham, R., Ballou, J. D., Ralls, K., Eldridge, M., Dudash, M. R., Fenster, C. B., Lacy, R. C., & Sunnucks, P. (2017). *Genetic management of fragmented animal and plant populations*. Oxford University Press.

Frankham, R., Ballou, S. E. J. D., Briscoe, D. A., & Ballou, J. D. (2010). *Introduction to conservation genetics*. Cambridge university press.

Frydenberg, O. (1963). Population studies of a lethal mutant in Drosophila melanogaster: I. Behaviour in populations with discrete generations. *Hereditas*, *50*, 89–116.

Galla, S. J., Forsdick, N. J., Brown, L., Hoeppner, M. P., Knapp, M., Maloney, R. F., Moraga, R., Santure, A. W., & Steeves, T. E. (2018). Reference genomes from distantly related species can be used for discovery of single nucleotide polymorphisms to inform conservation management. *Genes*, *10*, 9.

Georges, A. (2020). *Assessment of hybridization and introgression in Bellinger River turtles, and assessment of parentage of hatchlings destined for release in the recovery program of Myuchelys georgesi.: Biodiversity and conservation, NSW Department of Planning*. Industry and Environment. http://georges.biomatix.org/storage/app/uploads/public/5f4/f7a/bf4/5f4f7abf48b9f904504354.pdf

Georges, A., & Adams, M. (1992). A phylogeny for Australian chelid turtles based on allozyme electrophoresis. *Australian Journal of Zoology*, *40*, 453–476.

Georges, A., Spencer, R.-J., Kilian, A., Welsh, M., & Zhang, X. (2018). Assault from all sides: Hybridization and introgression threaten the already critically endangered Myuchelys georgesi (Chelonia: Chelidae). *Endangered Species Research*, *37*, 239–247.

Gilligan, D. M., Briscoe, D. A., & Frankham, R. (2005). Comparative losses of quantitative and molecular genetic variation in finite populations of Drosophila melanogaster. *Genetics Research*, *85*, 47–55.

Gilligan, D. M., & Frankham, R. (2003). Dynamics of genetic adaptation to captivity. *Conservation Genetics*, *4*, 189–197.

Gist, D. H., & Jones, J. M. (1987). Storage of sperm in the reptilian oviduct. *Scanning Microscopy*, *1*, 31.

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*, 184–186.

Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). Dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources*, *18*, 691–699.

Grueber, C., Peel, E., Wright, B., Hogg, C., & Belov, K. (2019). A Tasmanian devil breeding program to support wild recovery. *Reproduction, Fertility and Development*, *31*, 1296–1304.

Grueber, C. E., Fox, S., Mclennan, E. A., Gooley, R. M., Pemberton, D., Hogg, C. J., & Belov, K. (2018). Complex problems need detailed solutions: Harnessing multiple data types to inform genetic management in the wild. *Evolutionary Applications*, *12*, 280–291.

Heinsohn, R., Lacy, R., Elphinstone, A., Ingwersen, D., Pitcher, B. J., Roderick, M., Schmelitschek, E., Van Sluys, M., Stojanovic, D., & Tripovich, J. (2022). Population viability in data deficient nomadic species: What it will take to save regent honeyeaters from extinction. *Biological Conservation*, *266*, 109430.

Hoban, S., Archer, F. I., Bertola, L. D., Bragg, J. G., Breed, M. F., Bruford, M. W., Coleman, M. A., Ekblom, R., Funk, W. C., Grueber, C. E., Hand, B. K., Jaffe, R., Jensen, E., Johnson, J. S., Kershaw, F., Liggins, L., Macdonald, A. J., Mergeay, J., Miller, J. M., ... Hunter, M. E. (2022). Global genetic diversity status and trends: Towards a suite of essential biodiversity variables (EBVs) for genetic composition. *Biological Reviews of the Cambridge Philosophical Society*, *97*, 1511–1538.

Hoffmann, A. A., Miller, A. D., & Weeks, A. R. (2021). Genetic mixing for population management: From genetic rescue to provenancing. *Evolutionary Applications*, *14*, 634–652.

Hogg, C., Mclennan, E., Wise, P., Lee, A., Pemberton, D., Fox, S., Belov, K., & Grueber, C. (2020). Preserving the demographic and genetic integrity of a single source population during multiple translocations. *Biological Conservation*, *241*, 108318.

Hogg, C., Wright, B., Morris, K., Lee, A., Ivy, J., Grueber, C., & 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*, 348–361.

Hogg, C. J., Grueber, C. E., Pemberton, D., Fox, S., Lee, A. V., Ivy, J. A., & Belov, K. (2017). "Devil Tools & Tech": A synergy of conservation research and management practice. *Conservation Letters*, *10*, 133–138.

IUCN/SSC. (2013). *Guidelines for reintroductions and other conservation translocations. Version 1.0*. IUCN Species Survival Commission.

Ivy, J. A., & Lacy, R. C. (2010). Using molecular methods to improve the genetic management of captive breeding programs for threatened species. In *Molecular Approaches in Natural Resource Conservation and Management*, 267–295. DeWoody, J. A., Bickham, J. W., Michler, C. H., Nichols, K. M., Rhodes, O. E., & Woeste, K. E. (Eds). New York: Cambridge University Press.

Ivy, J. A., Miller, A., Lacy, R. C., & Dewoody, J. A. (2009). Methods and prospects for using molecular data in captive breeding programs: An empirical example using Parma wallabies (Macropus parma). *Journal of Heredity*, *100*, 441–454.

Jakob-Hoff, R. M., Lees, C. M., & McGilvray, G. (2017). *Status review, disease risk analysis and con- servation action plan for the Bellinger River snapping turtle*. IUCN SSC Conservation Breeding Specialist Group.

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.

Jones, A., Ovenden, J., & Wang, Y. (2016). Improved confidence intervals for the linkage disequilibrium method for estimating effective population size. *Heredity*, *117*, 217–223.

Keenan, K., Mcginnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, *4*, 782–788.

Kinghorn, B. P., & Kinghorn, A. J. (2021). Management of inbreeding and coancestry to target short-term and long-term genetic gains. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics*, *24*, 191–194.

Lacy, R. C. (1987). Loss of genetic diversity from managed populations: Interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology*, *1*, 143–158.

Lande, R., & Barrowdough, G. (1987). Effective population size, genetic variation, and their use in population. *Viable Populations for Conservation*, *87*, 624–635.

Martinez, P. A., Ezaz, T., Valenzuela, N., Georges, A., & Marshall Graves, J. A. (2008). An XX/XY heteromorphic sex chromosome system in the Australian chelid turtle Emydura macquarii: A new piece in the puzzle of sex chromosome evolution in turtles. *Chromosome Research*, *16*, 815–825.

Mccallum, M. L. (2021). Turtle biodiversity losses suggest coming sixth mass extinction. *Biodiversity and Conservation*, *30*, 1257–1275.

Nistelberger, H. M., Roycroft, E., Macdonald, A. J., Mcarthur, S., White, L. C., Grady, P. G., Pierson, J., Sims, C., Cowen, S., & Moseby, K. (2023). Genetic mixing in conservation translocations increases diversity of a keystone threatened species,Bettongia lesueur. In *Genetic mixing in conservation translocations increases diversity of a keystone threatened species*. Bettongia lesueur.

Ochoa, A., Wells, S. A., West, G., Al-Smadi, M. E., Redondo, S. A., Sexton, S. R., & Culver, M. (2016). Can captive populations function as sources of genetic variation for reintroductions into the wild? A case study of the Arabian oryx from the Phoenix zoo and the Shaumari wildlife reserve, Jordan. *Conservation Genetics*, *17*, 1145–1155.

Ortega-Villaizan, M. D. M., Noguchi, D., & Taniguchi, N. (2011). Minimization of genetic diversity loss of endangered fish species captive broodstocks by means of minimal kinship selective crossbreeding. *Aquaculture*, *318*, 239–243.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*, 288–295.

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.

Piotrowski, J. S., Annis, S. L., & Longcore, J. E. (2004). Physiology of Batrachochytrium dendrobatidis, a chytrid pathogen of amphibians. *Mycologia*, *96*, 9–15.

Price, M. R. S. (1989). *Animal reintroductions: The Arabian oryx in Oman*. Cambridge University Press.

Pritchard, D. J., Fa, J. E., Oldfield, S., & Harrop, S. R. (2012). Bring the captive closer to the wild: Redefining the role of ex situ conservation. *Oryx*, *46*, 18–23.

R CORE TEAM. (2023). *R: A language and environment for statistical Computing*_. *2023.06.0 + 421*. R Foundation for Statistical Computing.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, *197*, 573–589.

Ralls, K., & Ballou, J. D. (2004). Genetic status and management of California condors. *The Condor*, *106*, 215–228.

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*, 4737–4754.

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*, 137–138.

Rumball, W., Franklin, I., Frankham, R., & Sheldon, B. (1994). Decline in heterozygosity under full-sib and double first-cousin inbreeding in Drosophila melanogaster. *Genetics*, *136*, 1039–1049.

Scheele, B. C., Hollanders, M., Hoffmann, E. P., Newell, D. A., Lindenmayer, D. B., Mcfadden, M., Gilbert, D. J., & Grogan, L. F. (2021). Conservation translocations for amphibian species threatened by chytrid fungus: A review, conceptual framework, and recommendations. *Conservation Science and Practice*, *3*, e524.

Schmidt, T. L., Jasper, M. E., Weeks, A. R., & Hoffmann, A. A. (2021). Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods in Ecology and Evolution*, *12*, 1888–1898.

Schou, M. F., Loeschcke, V., Bechsgaard, J., Schlötterer, C., & Kristensen, T. N. (2017). Unexpected high genetic diversity in small populations suggests maintenance by associative overdominance. *Molecular Ecology*, *26*, 6510–6523.

Sever, D. M., & Hamlett, W. C. (2002). Female sperm storage in reptiles. *Journal of Experimental Zoology*, *292*, 187–199.

Shafer, A. B., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, *8*, 907–917.

Spencer, R.-J., Georges, A., Lim, D., Welsh, M., & Reid, A. M. (2014). The risk of inter-specific competition in Australian short-necked turtles. *Ecological Research*, *29*, 767–777.

Spencer, R.-J., Van Dyke, J., Petrov, K., Ferronato, B., Mcdougall, F., Austin, M., Keitel, C., & Georges, A. (2018). Profiling a possible rapid extinction event in a long-lived species. *Biological Conservation*, *221*, 190–197.

Sunde, J., Yildirim, Y., Tibblin, P., Bekkevold, D., Skov, C., Nordahl, O., Larsson, P., & Forsman, A. (2022). Drivers of neutral and adaptive differentiation in pike (Esox lucius) populations from contrasting environments. *Molecular Ecology*, *31*, 1093–1110.

Taronga Conservation Society Australia. (2023). *Standard operating procedure—Bellinger River turtle Myuchelys georgesi*. Taronga Zoo.

Taylor, H. R., Dussex, N., & Heezik, V. A. N. (2017). Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Global Ecology and Conservation*, *10*, 231–242.

Thavornkanlapachai, R., Mills, H. R., Ottewell, K., Friend, J. A., & Kennington, W. J. (2021). Temporal variation in the genetic composition of an endangered marsupial reflects reintroduction history. *Diversity*, *13*, 257.

Torkamaneh, D., Laroche, J., & Belzile, F. (2016). Genome-wide SNP calling from genotyping by sequencing (GBS) data: A comparison of seven pipelines and two sequencing technologies. *PLoS One*, *11*, e0161333.

Van Dyke, J. U., Ferronato, B. D. O., & Spencer, R.-J. (2018). Current conservation status of Australian freshwater turtles. *Australian Journal of Zoology*, *66*, 1–3.

Wang, J. (2011). COANCESTRY: A program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, *11*, 141–145.

Williams, S. E., & Hoffman, E. A. (2009). Minimizing genetic adaptation in captive breeding programs: A review. *Biological Conservation*, *142*, 2388–2400.

Wisely, S. M., Mcdonald, D. B., & Buskirk, S. W. (2003). Evaluation of the genetic management of the endangered black-footed ferret (Mustela nigripes). *Zoo Biology: Published in Affiliation with the American Zoo and Aquarium Association*, *22*, 287–298.

Wright, B., Farquharson, K. A., Mclennan, E. A., Belov, K., Hogg, C. J., & Grueber, C. E. (2019). From reference genomes to population genomics: Comparing three reference-aligned reduced-representation sequencing pipelines in two wildlife species. *BMC Genomics*, *20*, 453.

Zhang, J., Finlaison, D. S., Frost, M. J., Gestier, S., Gu, X., Hall, J., Jenkins, C., Parrish, K., Read, A. J., Srivastava, M., Rose, K., & Kirkland, P. D. (2018). Identification of a novel nidovirus as a potential cause of large scale mortalities in the endangered Bellinger River snapping turtle (Myuchelys georgesi). *PLoS One*, *13*, e0205209.

Zhou, C., Mccarthy, S. A. & Durbin, R. (2022). YaHS: Yet another hi-C scaffolding tool. bioRxiv.

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: Nelson, H. V., Farquharson, K. A., Georges, A., McLennan, E. A., DeGabriel, J. L., Giese, M., Ormond, C., McFadden, M., Skidmore, A., Prangell, J., Belov, K., & Hogg, C. J. (2024). A genomic framework to assist conservation breeding and translocation success: A case study of a critically endangered turtle. *Conservation Science and Practice*, e13204. https://doi.org/10.1111/csp2.13204

A4.2 A genomic-based workflow for eDNA assay development for a critically endangered turtle, *Myuchelys georgesi*.

The published PDF version of Chapter 3 of this thesis is presented on the following pages.

Ecology and Evolution

WILEY

Ecology and Evolution Open Access

RESEARCH ARTICLE OPEN ACCESS

A Genomic-Based Workflow for eDNA Assay Development for a Critically Endangered Turtle, *Myuchelys georgesi*

Holly V. Nelson[1] | Arthur Georges[2] | Katherine A. Farquharson[1,3,4] | Elspeth A. McLennan[1] | Jane L. DeGabriel[4] | Katherine Belov[1,3] | Carolyn J. Hogg[1,3]

[1]School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales, Australia | [2]Institute for Applied Ecology, University of Canberra, Bruce, Australian Capital Territory, Australia | [3]Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Sydney, Sydney, New South Wales, Australia | [4]NSW Department of Climate Change, The Environment, Energy and Water, Parramatta, New South Wales, Australia

Correspondence: Carolyn J. Hogg (carolyn.hogg@sydney.edu.au)

Received: 24 November 2024 | **Revised:** 5 December 2024 | **Accepted:** 10 December 2024

Funding: This work was supported by NSW Department of Climate Change, the Environment, Energy and Water; Australian Government's Bushfire Recovery Scheme (ERF-WRR2-020) and Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science (CE200100012).

Keywords: Bellinger River turtle | eDNA | genomic data | *Myuchelys georgesi* | reference genome

ABSTRACT

Environmental DNA (eDNA) analysis has become a popular conservation tool for detecting rare and elusive species. eDNA assays typically target mitochondrial DNA (mtDNA) due to its high copy number per cell and its ability to persist in the environment longer than nuclear DNA. Consequently, the development of eDNA assays has relied on mitochondrial reference sequences available in online databases, or in cases where such data are unavailable, de novo DNA extraction and sequencing of mtDNA. In this study, we designed eDNA primers for the critically endangered Bellinger River turtle (*Myuchelys georgesi*) using a bioinformatically assembled mitochondrial genome (mitogenome) derived from a reference genome. We confirmed the accuracy of this assembled mitogenome by comparing it to a Sanger-sequenced mitogenome of the same species, and no base pair mismatches were detected. Using the bioinformatically extracted mitogenome, we designed two 20 bp primers that target a 152-base-pair-long fragment of the cytochrome oxidase 1 (CO1) gene and a 186-base-pair-long fragment of the cytochrome B (CytB) gene. Both primers were successfully validated *in silico*, *in vitro*, and *in situ*.

1 | Introduction

In recent years, conservation geneticists have made substantial progress in understanding how to apply genetic data to conservation actions for threatened species (Hohenlohe, Funk, and Rajora 2021). The prevalence of cost-effective, non-invasive molecular tools like environmental DNA (eDNA) assays have become increasingly common in detecting invasive species, assessing community diversity across various spatial scales, and monitoring rare or cryptic species (Rees et al. 2014; Ardura et al. 2015; Ruppert, Kline, and Rahman 2019; Lam, Sung, and Fong 2022). eDNA refers to extra organismal genetic material the comprises of molecules that have been shed into the environment by decaying bodies, leaves, blood, pollen, seeds, urine, faeces, skin, hairs and other types of organismal material (Freeland 2017), that can be extracted from environmental samples such as soil, water and air (Barnes et al. 2014; Rees et al. 2014). The presence of eDNA can be detected using DNA metabarcoding for detection of entire communities or species-specific primers or assays to detect a target species (Mauvisseau et al. 2019; Lopes et al. 2021; Valdivia-Carrillo et al. 2021). eDNA assays commonly target and amplify a short fragment of mitochondrial DNA (mtDNA) through polymerase chain reaction (PCR). mtDNA is commonly targeted as it is highly abundant in

cells and can persist in environments longer than nuclear DNA (nuDNA) (Wilcox et al. 2016; Bylemans et al. 2018).

Species-specific eDNA marker development relies on the availability of mtDNA sequences in online databases, as demonstrated in recent studies on diamondback terrapins (*Malaclemys terrapin*) and red eared slider turtles (*Trachemys script elegans*) (Fields et al. 2024); reef sharks (*Carcharhinus amblyrhynchos*) (Dunn et al. 2023); and Atlantic wolf-fish (*Anarhichas lupus*) (Chevrinais and Parent 2023). For species without publicly available mtDNA sequences, sequencing is required to facilitate marker development. Conventional methods for generating mtDNA sequence data have involved tissue acquisition, DNA extraction, designing universal primers, or primers of a closely related species, long-range polymerase chain reactions (PCRs), shotgun sequencing, followed by bioinformatic assembly (Kundu et al. 2020; Chen et al. 2021; Tessler et al. 2023). The advent of high-throughput parallel sequencing (HTS), reductions in sequencing costs and lower input DNA requirements, as well as improved bioinformatic pipelines, have given rise to the genomics era where traditional genetic approaches are being replaced by whole-genome approaches to conservation genetic research (Satam et al. 2023). While genomic data alone have no direct impact on conservation outcomes, they provide a foundational blueprint that that can be harnessed by geneticists and conservationists for a range of downstream applications (Hogg et al. 2022). These can include; aiding in the identification of genetic variants for population genetic analyses (Brandies et al. 2019); investigations into functionally important genetic variation such as immune genes (Peel et al. 2022), development of PCR primers and recently *in silico* extraction of complete mitochondrial genomes (hereafter 'mitogenomes') (Meng et al. 2019; Uliano-Silva, Nunes, and Krasheninnikova 2021).

The Bellinger River turtle (*Myuchelys georgesi*) is a species of short-necked turtle (Family Chelidae) and is one of two turtle species that is listed as Critically Endangered in Australia under the Environmental Protection and Biodiversity Conservation Act 1999 (Commonwealth of Australia 1999). The species has a current known distribution that is restricted to 60 km of the Bellinger River and a short section of its main tributary, the Kalang River, in north-eastern New South Wales (NSW), Australia (Cann et al. 2015). However, the species has not been recorded in the Kalang since 2007 (Georges et al. 2011). *Myuchelys georgesi* is a rare and cryptic species that has adapted to up-stream regions of the Bellinger River, preferring deeper waterholes surrounded by bedrock making them difficult to survey using conventional diving and trapping methods (Spencer et al. 2014). In 2015, a novel nidovirus outbreak resulted in the estimated death of more than 90% of individuals, further contributing to the species' rarity (Zhang et al. 2018; Chessman et al. 2020). The species also faces threats from competition with another locally occurring species, the Murray River turtle (*Emydura macquarii*). Implementation of eDNA analyses in both known and data deficient areas of the catchment (e.g., Kalang River) is currently listed in the species Conservation Action Plan to inform survey site selection (R. Jakob-Hoff et al., unpublished), yet no such tool currently exists.

Given the growing application of both eDNA and genomic data in conservation management, we used a PacBio HiFi reference genome to develop species-specific eDNA markers for *M. georgesi*. We also provide comprehensive methodologies and visual workflow for other threatened species, with reference genomes or genomic data, which would benefit from an eDNA assay.

2 | Methods

2.1 | Mitogenome Assembly

We previously assembled a chromosome-level reference genome for *M. georgesi* using PacBio High Fidelity (HiFi) (CA, USA) sequencing (Nelson et al. 2024). HiFi sequencing is a type of long-read data that is generated by circular consensus sequencing (CCS). Raw CCS reads can be as long as 15,000–20,000 base pairs, allowing full-length mitogenome sequences to be captured within a single read. To generate HiFi sequence data, high molecular weight DNA was extracted from the heart tissue of a male *M. georgesi* using the Nanobind Tissue Big DNA kit following the manufacturer's protocol (Circulomics, Pacific Biosciences, CA, USA). PacBio HiFi Single-Molecule Real-Time (SMRT) bell libraries were sequenced at the Australian Genome Research Facility (Brisbane, Australia). The HiFi genome was assembled using Hifiasm v.0.16.0 (Cheng et al. 2021). To obtain a complete mitogenome (i.e., the entire mitochondrial DNA), we bioinformatically extracted the mtDNA sequence from the HiFi genome fasta file (a text-based file format containing nucleotide sequences) using MitoHifi v2 (Uliano-Silva, Nunes, and Krasheninnikova 2021). The -c flag was used to identify and annotate the mitogenome from genome scaffolds, rather than assembling it from raw reads with the -r flag. MitoHiFi also requires a mitochondrial reference sequence as input in either fasta or GenBank format (e.g., https://www.ncbi.nlm.nih.gov/genbank/samplerecord/). MitoHiFi provides an internal script (findMitoReference.py) that can be used to find and download the most closely related mitogenome for the species of interest. For this study, we manually obtained reference sequences from the NCBI for the Green Sea Turtle (*Chelonia mydas*) (NC_000886) (Kumazawa and Nishida 1999), Murray River Turtle (*Emydura macquarii*) (NC_041302.1) (unpublished) and a previously Sanger sequenced *M. georgesi* mitogenome (NC_042474.1) (unpublished). These sequences were used to evaluate whether levels of divergence between reference and target species affected assembly quality. The mitogenome was visualised using Proksee (Grant et al. 2023) (Figure 1). To confirm efficacy of the bioinformatically extracted mitogenome, we used MEGA11 (Tamura, Stecher, and Kumar 2021) to align the assembly to a Sanger sequenced *M. georgesi* mitogenome for structural comparison and mismatches between sequences (NC_042474.1).

2.2 | Species-Specific Primer Development and Validation

Using the annotated fasta file output by MitoHiFi, we located genetic sequences labelled 'CO1' and 'CytB' and used the complete sequence (Figures S1 and S2) as input into Primer3Plus v3.3.0 (Untergasser et al. 2012) to design forward and reverse primer sequences with 0 base pair mismatches with the CO1 and CytB gene sequences. These genes were used as they are known to be highly variable among closely related

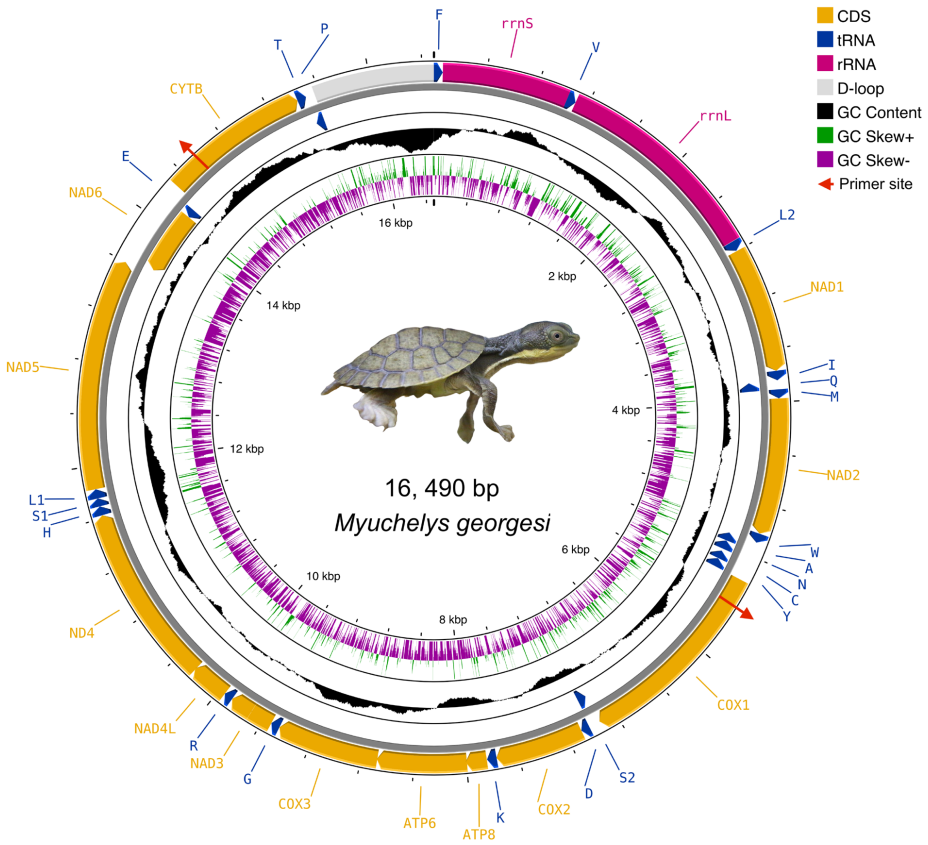


FIGURE 1 | The mitochondrial genome of *M. georgesi* extracted using MitoHiFi (Uliano-Silva, Nunes, and Krasheninnikova 2021). tRNAs are labelled according to their single-letter abbreviation. Arrows indicate direction of gene transcription. Protein coding genes are shown in yellow, rRNA genes in pink, tRNA genes in blue and the 920 bp non-coding region between P and F in white. The GC-skew depicting the deviation from the average in the complete mitogenome is depicted in green (positive) and maroon (negative), and the GC content is depicted in black. Figure generated using the Proksee (https://proksee.ca). *M. georgesi* juvenile image credit of Paul Fahy.

TABLE 1 | Primers designed (CytB and CO1) and used (12S) in this study for amplification of *M. georgesi* mitochondrial eDNA. T_m melting temperature.

Gene	Name	Forward/ Reverse	Nucleotide sequence	Primer length (bp)	Amplicon size (bp)	T_m (°C)
CytB	MG_CB	Forward	AATCTCCCACATCCAACGAG	20	186	59.9
		Reverse	ATGCGGTGGCTATGACTAGG			60.1
CO1	MG_C1	Forward	ACATTGGCACCCTCTACCTG	20	152	60
		Reverse	AATTAAGGCGTGGGCTGTAA			59.6
12S	12Sv5	Forward	TAGAACAGGCTCCTCTAG	18	~100	Riaz et al. (2011)
		Reverse	TTAGATACCCCACTATGC			

species providing greater specificity for species-specific eDNA assays compared to mitochondrial genes with lower inter-specific variation (Moritz, Dowling, and Brown 1987; Meyer 1994; Johns and Avise 1998; Hebert, Ratnasingham,

and de Waard 2003). Forward and reverse primers output by Primer3Plus were individually input into OligoAnalyzer (https://sg.idtdna.com/calc/analyzer) for quality checks using the hairpin and homodimer options to ensure efficiency and sensitivity of primer binding. To ensure primer stability and minimise the likelihood of hairpin structure formation (when complementary base-pair sequences create a loop), we used a ΔG (Gibbs free energy change) threshold of −4.5 kcal/mol. For homodimers (annealing of identical primer sequences) we ensured primers had no more than three complementary bases. The melting temperature (T_m) for all primer sequences fell between 59.6°C and 60.1°C. Final primer pairs are provided in Table 1 and Table S2. The specificity and sensitivity of primer sets were evaluated at three stages: *in silico*, *in vitro*, and *in situ*.

2.3 | *In Silico* Validation

To confirm specificity *in silico*, the alignment search tool Basic Local Alignment Search Tool (BLAST) was used to confirm percent of sequence similarity with other species (https://blast.ncbi.nlm.nih.gov/). To visually confirm specificity and optimal primer design against another locally occurring species (*E. macquarii*) we used MEGA v11 to align both our assembled and *E. macquarii* mitogenomes (NC_041302.1) and ensured 2–3 mismatches between primer design sequences (de Brauwer et al. 2022b). Custom primer sets were ordered using ThermoFisher Scientific (MA, USA) custom DNA oligos synthesis service.

2.4 | *In Vitro* Validation

To evaluate specificity and amplification efficiency *in vitro*, we carried out tests using tissue-derived DNA from *M. georgesi* and *E. macquarii* using conventional PCR (Figure S2). Heart tissue belonging to a female *M. georgesi* that required medical euthanasia in 2021 (C10031) was flash frozen at −80°C at Taronga Zoo and stored at −80°C at the University of Sydney. *E. macquarii* skin tissue was acquired from the trailing webbing of the hindfoot of a wild individual in 2015 (UC<Aus>AA063724) and stored at −80°C in the University of Canberra Wildlife Tissue Collection (GenBank UC<Aus>). To prevent contamination during lab procedures, equipment was sterilised in an autoclave and benchtops cleaned with 80% ethanol. DNA (Table S3) was extracted using the Qiagen DNeasy Mini Kit (Qiagen, Germany) following the manufacturer's protocol, except for a final elution in 100 μL buffer AE (Qiagen). Quality (fragmentation) and concentration of DNA were assessed using a combination of a Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific) and 1.5% agarose/TBE gel electrophoresis stained with SYBR safe (Life Technologies), alongside a 1 kb size standard (Bioline) and run for 55 min at 100 V. Samples yielding high concentrations of DNA were used for subsequent PCR amplification assays.

For PCR set-up, 0.25 μM of CytB and CO1 forward and reverse primers were used. A quantity of 0.25 μM of 12Sv5F/12Sv5R universal vertebrate primers was used as a positive control by amplifying a ~100-bp fragment of the V5 loop of the 12S mitochondrial gene (Riaz et al. 2011). The final PCR reaction consisted of 3 μL of *M. georgesi* DNA template or negative extraction control (*E. macquarii* DNA template, ddH_2O), 25 μL of Bioline MyTaq Mix (Bioline, UK), 2.5 μM of forward and reverse primers (either 12Sv5, CytB or CO1), and 17 μL of nuclease free water to make a total volume of 50 μL.

Real-time PCR cycling was carried out on a T100 Thermal Cycler (BioRad). Cycling conditions were 10 min for enzyme activation at 95°C, 35 cycles of denaturation at 95°C for 30 s, annealing at 50°C for 30 s, extension at 72°C for 30 s and a final extension at 72°C for 10 min. Amplification was confirmed using 1.5% agarose/TBE gel electrophoresis stained with SYBR safe (Life Technologies), alongside a 1 kb size standard (Bioline) and run for 55 min at 100 V. Bands were visualised under ultraviolet light using a ChemiDoc XRS + system (BioRad) and images were analysed with ImageLab (BioRad).

2.5 | *In Situ* Validation

M. georgesi eDNA water samples (positive controls) were obtained from three 4000 L, closed-system tanks at Symbio Wildlife Park in Helensburgh, Australia, each housing four or five animals. From each tank, two 500 mL water samples were collected. For negative controls, we collected two 500 mL water samples from a 2000 L pond containing four *E. macquarii* and two Eastern long-necked turtles (*Chelondina longicollis*). We transported the water samples on ice and stored them briefly at −2°C before filtering within 1–2 h of collection. Negative control samples were handled and stored separately to prevent contamination.

A 47 mm Whatman membrane filter paper with a pore size of 0.45 μm was dampened with deionised water before the 500 mL water samples were filtered through. The filtration system included a 50 mm Büchner funnel, adaptor, 500 mL Büchner flask, rubber tubing and a diaphragm pump (KNF, CA, USA). The filter papers were then placed in individual resealable bags and frozen at −80°C prior to DNA extraction the following day.

eDNA extractions were conducted on samples (Table S4) using Qiagen's DNeasy Blood & Tissue Kit (Qiagen, Germany). DNA extraction followed the protocol of Renshaw et al. (2015) with minor adjustments. Briefly, each filter paper was halved and finely cut before being placed in separate 2 mL screw-cap tubes. Volumes of 540 μL buffer ATL and 60 μL (rather than the recommended 180 and 20 μL, respectively) of Proteinase K were added to submerse each half filter and incubated at 65°C for 1 h. Following lysis, the paper was tightly pressed to the bottom of the tube, and supernatant transferred to a new 2 mL screw-cap tube. Volumes of 630 μL Buffer AL and 630 μL of ethanol were added and mixed thoroughly with a vortex. The lysates from each half were then combined by passing the mixtures through the same DNeasy Mini spin column, resulting in six rounds of centrifugation and discarded flow-through. Total eDNA was rinsed with 500 μL of AW1 and AW2 solutions respectively and eluted in 100 μL buffer AE (Qiagen). eDNA concentration was quantified using a Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific). All eDNA extractions were placed in a freezer (−20°C) for 12 h until PCR analysis.

Following our *in vitro* validation protocol, 0.25 μM of CytB and CO1 forward and reverse primers were used for PCR set-up. A quantity of 0.25 μM of 12Sv5F/12Sv5R universal vertebrate primers was used as a positive control. PCR mixes consisted of 3 μL of *M. georgesi* eDNA template or a negative extraction control (*E. macquarii* and *C. longicollis* eDNA template or ddH_2O), 25 μL of Bioline MyTaq Mix (Bioline, UK), 2.5 μM of forward and reverse primers, and 17 μL of nuclease free water to make a total volume of 50 μL. Real-time PCR cycling, agarose gel electrophoresis, and image analysis was conducted using the same methods described in *in vitro* validation section above.

3 | Results

3.1 | *Myuchelys georgesi* Mitogenome

The complete mitochondrial sequence was extracted from scaffold 9 of our reference assembly and yielded a complete length of 16,490 bp (Figure 1). The same mitogenome was assembled when the Green Sea Turtle, Murray River Turtle and Bellinger River Turtle mitogenomes were used as reference sequences, confirming that varying levels of divergence between reference input and target species does not affect final assembly quality. The size and structure of the mitochondrial genome is comparable to other chelid turtles (Fielder et al. 2012; Zhang et al. 2017), which includes 37 genes consisting of 22 transfer RNA (tRNA) genes, 13 protein coding genes, 2 ribosomal RNA (rRNA) genes, plus a non-coding region (CR). Additional details can be found in the Supporting Information.

Visual alignment of the bioinformatically assembled mitogenome to the Sanger sequenced *M. georgesi* mitogenome (NC_042474.1) using MEGA v11 showed a 100% sequence identity match, confirming efficacy of the *in silico*-based mitogenome.

3.2 | Primer Design and Validation

In silico primer assessment found greater species-specificity of the CO1 primers compared to CytB as BLAST results returned lower percentage identity with other species. Both CO1 and CytB primers successfully amplified *M. georgesi* tissue samples (Figure 2A; lanes 1–2, 4–5, 7–8). Both sets of primers showed no amplification on *E. macquarii* tissue (Figure 2A; lanes 10–11 and 13–14), confirming the species-specificity of the primers against the other locally occurring species. The 12Sv5 positive control amplified across both species, indicating the presence of mitochondrial DNA in the tissues (Figure 2A; lanes 3, 6, 9, 12 and 15), while no amplification was observed for the ddH2O negative controls (Figure 2A; lanes 16–17). *In situ* evaluation showed PCR products for both primers successfully amplified *M. georgesi* eDNA collected on cellulose ester filters from tank water (Figure 2B; lanes 1–2, 4–5, 7–8). Primer sets did not amplify eDNA from tank water containing *E. macquarii* or *C. longicollis* (Figure 2B; lanes 10–11), confirming species-specificity of primers against other locally occurring species. Amplification of the positive 12Sv5 control across tank water confirmed the presence of mtDNA in all samples (Figure 2B; lanes 3, 6, 9 and 12) while no amplification was observed for the ddH2O negative control (Figure 2B; lanes 13–14).

4 | Discussion

We developed the first eDNA markers for detection of *M. georgesi* using an existing long-read PacBio HiFi reference genome. The 100% sequence identity match between the Sanger sequenced and bioinformatically assembled mitogenome (NC_042474.1) and successful amplification of mtDNA across *in silico*, *in vitro* and *in situ* validations highlights the efficacy of genomic data-derived mitogenome assemblies, without the need for targeted mitochondrial DNA tissue extraction and sequencing.

We provide comprehensive methodologies of our workflow for other taxa that may benefit from this approach (Figure 3). Conventional approaches rely on the availability of mitochondrial sequence data from online databases or de novo extraction, sequencing and assembly when sequence data is not available (Schmidt et al. 2016; Zhang et al. 2017; Kundu et al. 2019, 2020; Frandsen, Figueroa, and George 2020; Chen et al. 2021). For conservation programs with genomic resources but lacking mitochondrial sequence data, this approach offers an avenue for developing a widely used conservation genetic tool.

When developing a species-specific eDNA assay, it is essential to have DNA sequence information unique to your target organism. The most efficient approach is identifying if relevant sequence data is available in online repositories (Figure 3: step 1) such as the National Centre for Biotechnology Information (NCBI), the Barcode of Life Data System (BOLD), and the European Molecular Biology Laboratory online repositories. The Sanger sequenced *M. georgesi* mitogenome used as our positive control for the bioinformatic mitogenome extraction was obtained from the NCBI database using '*Myuchelys georgesi* mitochondrion' as search terms. When relevant mitochondrial sequence data are not available in online repositories or if gene regions are missing for species or taxa (Freeland 2017; Nordstrom et al. 2022), genomic data (reference genome or raw HTS) provides an *in silico* alternative (Figure 3: step 2). For example, the availability of a reference genome has allowed for bioinformatic extraction of the mitogenomes for several cryptic and threatened species lacking mitochondrial sequence data including the Kroombit tinker frog (*Taudactylus pleione*) (Farquharson et al. 2023), blue-tailed skink (*Cryptoblepharus egeriae*) (Dodge et al. 2023), Lister's gecko (*Lepidodactylus listeri*) (Dodge et al. 2023), and southern stuttering barred frog (*Mixophyes balbus*) (Tang et al. 2024), providing capacity for the development of species-specific eDNA assays in the future. Although our approach leverages PacBio HiFi sequencing data and MitoHiFi mitochondrial genome assembly program for bioinformatic extraction (Figure 3: Step 3a), a suite of bioinformatic tools are available for extraction and assembly of mitogenomes from a range of HTS data types (Table 2). Additionally, some of these tools can take raw HTS sequencing data as input and do not require a reference genome.

If mitochondrial or genomic sequence data does not exist (Figure 3: Step 1 and 2), conventional approaches involving acquisition of genetic material; DNA extraction; sequencing; and assembly are needed to undertake species' assay design (Figure 3: Step 3b). Although targeted mitochondrial sequencing may be effective when programs have limited funds available (Schmidt, Thia, and Hoffmann 2024), these approaches often require substantial time and resources to undertake so

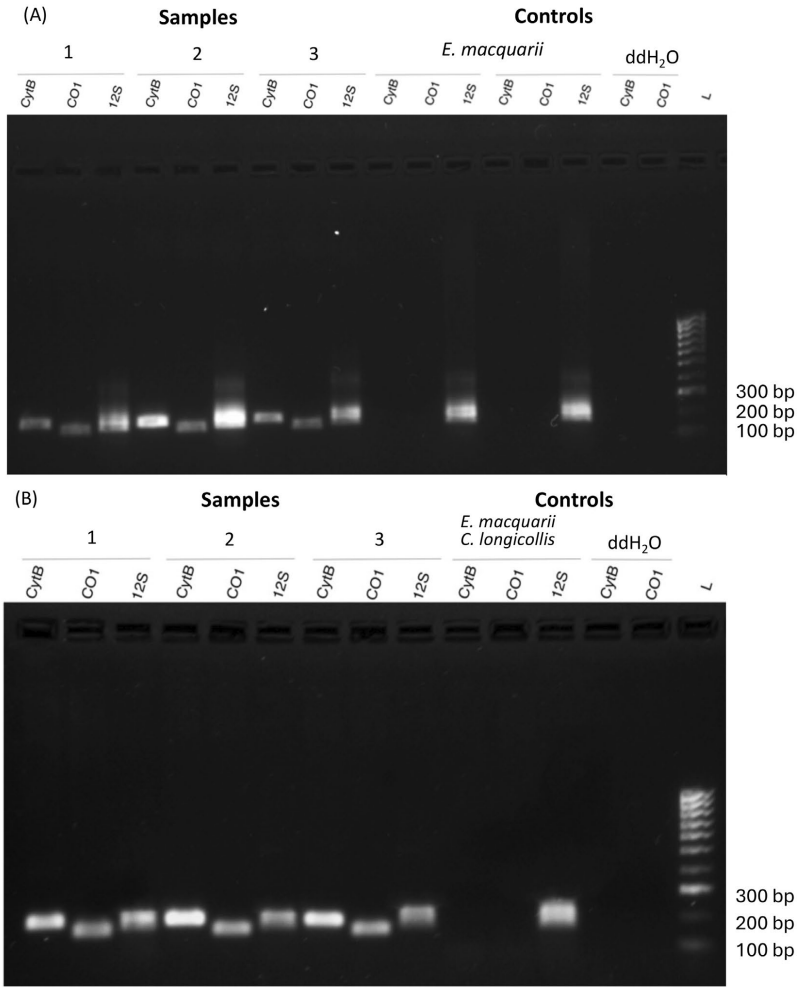


FIGURE 2 | 1.5% agarose gel and TBE stained with SYBR safe, showing (A) *in vitro* amplicon products of tissue derived DNA for Bellinger River Turtle (*Myuchelys georgesi*) with CytB, CO1 and 12Sv5 control (lanes 1–9), Murray River Turtle (*Emydura macquarii*) with CytB, CO1 and 12Sv5 control (lanes 10–15), and ddH_2O with CytB and CO1 (16–17). (B) *In situ* amplicon products of tank water derived eDNA for Bellinger River Turtle (*Myuchelys georgesi*) with CytB, CO1 and 12Sv5 control (lanes 1–9), Murray River Turtle (*Emydura macquarii*) and Eastern long-necked Turtle (*Chelondina longicollis*) with CytB, CO1 and 12Sv5 control (lanes 10–12), and ddH_2O with CytB and CO1 (lanes 13–14).

likely cost the same as whole genome sequencing when labour costs are accounted for. For example, completion of the existing *M. georgesi* mitogenome following the methods of Zhang et al. (2017) used Sanger sequencing and long-range PCR, took 12 weeks to complete, costing $15,000 AUD in labour and $1500 in lab consumables (Arthur Georges pers. comm., 2024). By-passing these steps, when genomic data is available, can save conservation programs time and money that can be invested elsewhere. For example, bioinformatic extraction of the mitogenome from a 1.9GB genome required 30 min, 1 CPU and 5.3GB of memory, offering a high cost-effectiveness in terms of labour, data acquisition and analysis. As the costs associated with genome assembly decrease, a 3GB long-read genome can cost ~$5005 in sequencing, ~$600 in labour and ~$200 in consumables (Elspeth McLennan pers. comm., 2024). Additionally, completion of a reference genome can only require 2 days of laboratory work, 6 weeks of sequencing, and 2 days for bioinformatic assembly. Although costs are not directly comparable, investment in genomic data provides a resource for a plethora of downstream applications beyond mitochondrial and eDNA (Formenti et al. 2022; de León et al. 2023; Schneider 2023; Brandies et al. 2019).

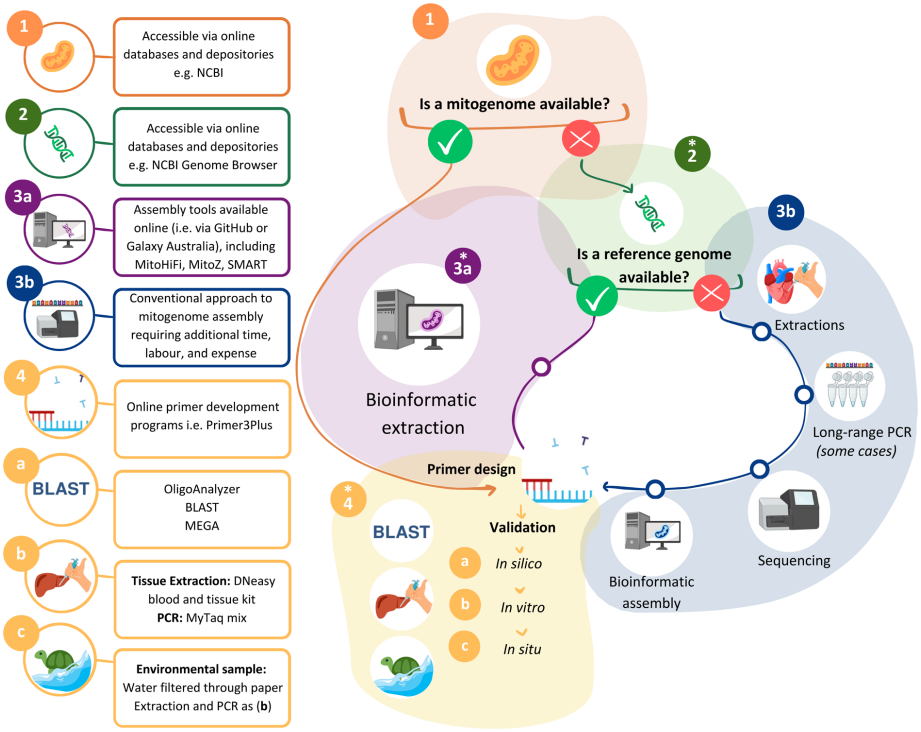


FIGURE 3 | Reference genome derived eDNA assay workflow used for species-specific primer development in *Myuchelys georgesi*. NCBI National Center for Biotechnology Information, BLAST Basic Local Alignment Search Tool, MEGA Molecular Evolutionary Genetic Analysis, PCR Polymerase Chain Reaction. Steps undertaken in this study are denoted by *. Image created using Biorender.com and Canva.com.

TABLE 2 | Bioinformatic tools for complete mitochondrial genome assembly using next generation sequencing (NGS) data (including whole genomes) for downstream eDNA assay development.

Name	Data input described	Reference
MitoHiFi	PacBio high fidelity (HiFi) (CCS) WGS data	Uliano-Silva, Nunes, and Krasheninnikova (2021)
PMAT	PacBio high fidelity (HiFi) (CCS) WGS data	Bi et al. (2024)
SMART	Low-coverage WGS	Alqahtani and Măndoiu (2020)
MitoZ	Short WGS raw reads	Meng et al. (2019)
Norgal	Short WGS raw reads	Al-Nakeeb, Petersen, and Sicheritz-Pontén (2017)
MITObim	Short WGS raw reads	Hahn, Bachmann, and Chevreux (2013)

Abbreviations: CCS, close consensus sequencing; CLR, continuous long reads; WGS, whole genome sequencing.

The key aspect of an eDNA assay is primer design (Figure 3: Step 4). As mentioned in step 1, primers are often developed using available reference sequences in online databases, however, regions may be missing for species or taxa (Freeland 2017; Nordstrom et al. 2022). For example, 12S, 16S, 18S sequence data is less often available compared to CO1 and CytB sequence information (Lacoursière-Roussel et al. 2016). An advantage of a reference genome-derived approach is that it provides researchers and managers with a complete or close to complete mitochondrial sequence. This enables the design of molecular markers for any gene in the mitogenome and provides the option to expand into nuclear marker design (Mccauley et al. 2024). *In silico*, *in vitro* and *in situ* validation methodologies (Figure 3: Step 4A-C) should follow a standardised approach (Nordstrom et al. 2022). Since the rapid uptake of eDNA analysis, comprehensive eDNA guidelines for assay development and validation have been

developed to assist researchers and managers in developing eDNA across a range of taxa and ecosystems that can be adapted to the habitat and biology of the target species (Goldberg and Strickler 2017; de Brauwer et al. 2022a, 2022b).

In summary, our methodologies and workflow for *M. georgesi* consist of four stages; (i) identifying availability of a mitogenome (Figure 3: step 1); (ii) identifying availability of a reference genome or genomic data when mitochondrial sequence does not exist in online repositories (Figure 3: step 2); (iii) bioinformatic assembly of a mitogenome from a reference genome (Figure 3: step 3a); and (iv) primer design and *in silico*, *in vitro* and *in situ* validation (Figure 3: step 4).

Our results provide *M. georgesi* managers with an eDNA assay that can be implemented into species monitoring. The assay can assist managers in resolving questions around distribution within the Bellinger River catchment, including reaches in the upper catchment and the Kalang River, and inform survey site selection through identification of occupancy hotspots. Future work is needed to evaluate efficacy of primers on Bellinger River water samples as environmental barriers such as water flow, sediment composition and microbial and enzyme activity (Barnes et al. 2014; Stoeckle et al. 2017; Stewart 2019) may influence detection. The technique will be useful for initially be used to identify areas to perform more intensive diving and trapping surveys, providing the species with a multifaceted detection and survey approach (Villacorta-Rath et al. 2022; Lam, Sung, and Fong 2022; Nordstrom et al. 2022; Carvalho et al. 2022).

As conservation genetics moves into the genomics-era, genomic data is becoming increasingly available for non-model organisms, making it important to leverage and apply the information genomic resources provide. We use a reference genome-based approach to develop an eDNA assay for *M. georgesi*. The development of species-specific eDNA primers provides a valuable tool for managers in assessing population dynamics of this rare species, supporting informed management decisions and guiding future conservation efforts.

Author Contributions

Holly V. Nelson: conceptualization (supporting), formal analysis (lead), investigation (lead), writing – original draft (lead). **Arthur Georges:** supervision (supporting), writing – review and editing (equal). **Katherine A. Farquharson:** investigation (supporting), writing – review and editing (supporting). **Elspeth A. McLennan:** investigation (supporting), methodology (supporting), supervision (supporting), writing – review and editing (supporting). **Jane L. DeGabriel:** funding acquisition (supporting), supervision (supporting), writing – review and editing (supporting). **Katherine Belov:** resources (equal), supervision (equal), writing – review and editing (supporting). **Carolyn J. Hogg:** conceptualization (lead), funding acquisition (lead), project administration (lead), resources (equal), supervision (equal), writing – review and editing (supporting).

Acknowledgements

We acknowledge Gumbaynggirr people and their ongoing connection to Country where the turtles live and pay respects to their Elders past and present. This research was made possible by the NSW DCCEEW Saving our Species program, Symbio Wildlife Park, and the NCRIS funded Bioplatforms Australia Threatened Species Initiative. Computing resources were provided by the Pawsey Supercomputing Research Centre; Bioplatforms Australia; Galaxy Australia; and the University of Sydney's High Performance Computing cluster Artemis. Thanks to Jarrad Prangell for providing access and assistance with sample acquisition and Kimberley Batley for guidance with lab work. Xiuwen Zhang generated the whole mitogenome for *Myuchelys georgesi* (NC_042474.1). Funding for this work was provided by the Australian Government's Bushfire Recovery Scheme (ERF-WRR2-020) to the Threatened Species Initiative, and HVN scholarship funding was provided by the NSW DCCEEW and the University of Sydney (CJH).

Ethics Statement

Collection of samples was conducted in accordance with the conditions of NSW DCCEEW Animal Ethics Committee (AEC151201-3, AEC160503-01 and AEC180904-5) and Scientific Licences (MWL00102467, SL101672 and SL10255), and the Taronga Conservation Society Australia Scientific Licence SL101204.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The male *M. georgesi* reference genome assembly, from which the mitogenome is derived, and all raw sequencing reads including the 3-tissue transcriptome RNA-seq reads are available from NCBI under BioProject PRJNA1003540.

References

Al-Nakeeb, K., T. N. Petersen, and T. Sicheritz-Pontén. 2017. "Norgal: Extraction and de Novo Assembly of Mitochondrial DNA From Whole-Genome Sequencing Data." *BMC Bioinformatics* 18: 1–7.

Alqahtani, F., and I. Măndoiu. 2020. "SMART2: Multi-Library Statistical Mitogenome Assembly With Repeats." In *Computational Advances in Bio and Medical Sciences: 9th International Conference*, ICCABS 2019, Miami, FL, USA, November 15–17, 2019, Revised Selected Papers 9, 2020, 184–198. New York, NY: Springer International Publishing.

Ardura, A., A. Zaiko, J. L. Martinez, A. Samulioviene, A. Semenova, and E. Garcia-Vazquez. 2015. "eDNA and Specific Primers for Early Detection of Invasive Species–A Case Study on the Bivalve *Rangia cuneata*, Currently Spreading in Europe." *Marine Environmental Research* 112: 48–55.

Barnes, M. A., C. R. Turner, C. L. Jerde, M. A. Renshaw, W. L. Chadderton, and D. M. Lodge. 2014. "Environmental Conditions Influence eDNA Persistence in Aquatic Systems." *Environmental Science & Technology* 48: 1819–1827.

Bi, C., F. Shen, F. Han, et al. 2024. "PMAT: An Efficient Plant Mitogenome Assembly Toolkit Using Low-Coverage HiFi Sequencing Data." *Horticulture Research* 11: uhae023.

Brandies, P., E. Peel, C. J. Hogg, and K. Belov. 2019. "The Value of Reference Genomes in the Conservation of Threatened Species." *Genes* 10: 846.

de Brauwer, M., A. Chariton, L. Clarke, et al. 2022a. *Environmental DNA protocol development guide for biomonitoring*, 49. Canberra, Australia: National eDNA Reference Centre. https://dx.doi.org/10.25607/OBP-1853.

de Brauwer, M., A. Chariton, L. Clarke, et al. 2022b. *Environmental DNA test validation guidelines*, 33. Canberra, Australia: National eDNA Reference Centre. https://dx.doi.org/10.25607/OBP-1886.

Bylemans, J., E. M. Furlan, D. M. Gleeson, C. M. Hardy, and R. P. Duncan. 2018. "Does Size Matter? An Experimental Evaluation of the Relative Abundance and Decay Rates of Aquatic Environmental DNA." *Environmental Science & Technology* 52: 6408–6416.

Cann, J., R.-J. Spencer, M. Welsh, and A. Georges. 2015. "*Myuchelys georgesi* (Cann 1997) – Bellinger River Turtle." *Chelonian Research Monographs*: 1–9. https://doi.org/10.3854/crm.5.091.georgesi.v1.2015.

Carvalho, C. S., M. E. de Oliveira, K. G. Rodriguez-Castro, B. H. Saranholi, and P. M. Galetti Jr. 2022. "Efficiency of eDNA and iDNA in Assessing Vertebrate Diversity and Its Abundance." *Molecular Ecology Resources* 22: 1262–1273.

Chen, C., H. Xiaoyou, L. Wei, Y. Lingyun, C. Haigang, and Z. Xinping. 2021. "Complete Mitochondrial Genome and the Phylogenetic Position of the Burmese Narrow-Headed Softshell Turtle *Chitra vandijki* (Testudines: Trionychidae)." *Mitochondrial DNA Part B Resources* 6: 1216–1218.

Cheng, H., G. T. Concepcion, X. Feng, H. Zhang, and H. Li. 2021. "Haplotype-Resolved de Novo Assembly Using Phased Assembly Graphs With Hifiasm." *Nature Methods* 18: 170–175.

Chessman, B. C., G. Mcgilvray, S. Ruming, et al. 2020. "On a Razor's Edge: Status and Prospects of the Critically Endangered Bellinger River Snapping Turtle, *Myuchelys georgesi*." *Aquatic Conservation: Marine and Freshwater Ecosystems* 30: 586–600.

Chevrinais, M., and G. J. Parent. 2023. "A Validated Targeted Assay for Environmental DNA Detections of the Atlantic Wolffish (*Anarhichas lupus*)." *Conservation Genetics Resources* 15: 59–66.

Dodge, T. O., K. A. Farquharson, C. Ford, et al. 2023. "Genomes of Two Extinct-in-the-Wild Reptiles From Christmas Island Reveal Distinct Evolutionary Histories and Conservation Insights." *Molecular Ecology Resources*: 1–17. https://doi.org/10.1111/1755-0998.13780.

Dunn, N., D. J. Curnick, C. Carbone, et al. 2023. "Environmental DNA Helps Reveal Reef Shark Distribution Across a Remote Archipelago." *Ecological Indicators* 154: 110718.

Farquharson, K. A., E. A. Mclennan, K. Belov, and C. J. Hogg. 2023. "The Genome Sequence of the Critically Endangered Kroombit Tinkerfrog (*Taudactylus pleione*)." *F1000Research* 12: 845. https://doi.org/10.12688/f1000research.138571.1.

Fielder, D., K. Vernes, E. Alacs, and A. Georges. 2012. "Mitochondrial Variation Among Australian Freshwater Turtles (Genus Myuchelys), With Special Reference to the Endangered *M. bellii*." *Endangered Species Research* 17: 63–71.

Fields, A. T., E. P. Hunt, K. Lanoue, et al. 2024. "Development, Verification, and Implementation of an eDNA Detection Assay for Emydids With a Case Study on Diamondback Terrapins, Malaclemys Terrapin, and Red Eared Slider, Trachemys Script Elegans." *Conservation Genetics Resources* 16: 147–157.

Formenti, G., K. Theissinger, C. Fernandes, et al. 2022. "The Era of Reference Genomes in Conservation Genomics." *Trends in Ecology & Evolution* 37: 197–202.

Frandsen, H. R., D. F. Figueroa, and J. A. George. 2020. "Mitochondrial Genomes and Genetic Structure of the Kemp's Ridley Sea Turtle (*Lepidochelys kempii*)." *Ecology and Evolution* 10: 249–262.

Freeland, J. R. 2017. "The Importance of Molecular Markers and Primer Design When Characterizing Biodiversity From Environmental DNA." *Genome* 60: 358–374.

Georges, A., R.-J. Spencer, M. Welsh, H. B. Shaffer, R. Walsh, and X. Zhang. 2011. "Application of the Precautionary Principle to Taxa of Uncertain Status: The Case of the Bellinger River Turtle." *Endangered Species Research* 14: 127–134.

Goldberg, C., and K. Strickler. 2017. eDNA protocol sample collection. Washington State University & University of Idaho Collaborative Report.

Grant, J. R., E. Enns, E. Marinier, et al. 2023. "Proksee: In-Depth Characterization and Visualization of Bacterial Genomes." *Nucleic Acids Research* 51: 484–492.

Hahn, C., L. Bachmann, and B. Chevreux. 2013. "Reconstructing Mitochondrial Genomes Directly From Genomic Next-Generation Sequencing Reads—A Baiting and Iterative Mapping Approach." *Nucleic Acids Research* 41: e129.

Hebert, P. D., S. Ratnasingham, and J. R. de Waard. 2003. "Barcoding Animal Life: Cytochrome c Oxidase Subunit 1 Divergences Among Closely Related Species." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270: S96–S99.

Hogg, C. J., K. Ottewell, P. Latch, et al. 2022. "Threatened Species Initiative: Empowering Conservation Action Using Genomic Resources." *Proceedings of the National Academy of Sciences* 119: e2115643118.

Hohenlohe, P. A., W. C. Funk, and O. P. Rajora. 2021. "Population Genomics for Wildlife Conservation and Management." *Molecular Ecology* 30: 62–82.

Johns, G. C., and J. C. Avise. 1998. "A Comparative Summary of Genetic Distances in the Vertebrates From the Mitochondrial Cytochrome b Gene." *Molecular Biology and Evolution* 15: 1481–1490.

Kumazawa, Y., and M. Nishida. 1999. "Complete Mitochondrial DNA Sequences of the Green Turtle and Blue-Tailed Mole Skink: Statistical Evidence for Archosaurian Affinity of Turtles." *Molecular Biology and Evolution* 16: 784–792.

Kundu, S., V. Kumar, K. Tyagi, R. Chakraborty, and K. Chandra. 2019. "The First Complete Mitochondrial Genome of the Indian Tent Turtle, *Pangshura tentoria* (Testudines: Geoemydidae): Characterization and Comparative Analysis." *Ecology and Evolution* 9: 10854–10868.

Kundu, S., V. Kumar, K. Tyagi, and K. Chandra. 2020. "The Complete Mitochondrial Genome of the Endangered Assam Roofed Turtle, *Pangshura sylhetensis* (Testudines: Geoemydidae): Genomic Features and Phylogeny." *PLoS One* 15: e0225233.

Lacoursière-Roussel, A., M. Rosabal, and L. Bernatchez. 2016. "Estimating Fish Abundance and Biomass from eDNA Concentrations: Variability Among Capture Methods and Environmental Conditions." *Molecular Ecology Resources* 16, no. 6: 1401–1414.

Lam, I. P., Y.-H. Sung, and J. J. Fong. 2022. "Using eDNA Techniques to Find the Endangered Big-Headed Turtle (*Platysternon megacephalum*)." *PLoS One* 17: e0262015.

de León, L. F., B. Silva, K. J. Avilés-Rodríguez, and D. Buitrago-Rosas. 2023. "Harnessing the Omics Revolution to Address the Global Biodiversity Crisis." *Current Opinion in Biotechnology* 80: 102901.

Lopes, C. M., D. Baêta, A. Valentini, et al. 2021. "Lost and Found: Frogs in a Biodiversity Hotspot Rediscovered With Environmental DNA." *Molecular Ecology* 30: 3289–3298.

Mauvisseau, Q., A. Burian, C. Gibson, R. Brys, A. Ramsey, and M. Sweet. 2019. "Influence of Accuracy, Repeatability and Detection Probability in the Reliability of Species-Specific eDNA Based Approaches." *Scientific Reports* 9: 580.

Mccauley, M., S. A. Koda, S. Loesgen, and D. J. Duffy. 2024. "Multicellular Species Environmental DNA (eDNA) Research Constrained by Overfocus on Mitochondrial DNA." *Science of the Total Environment* 912: 169550.

Meng, G., Y. Li, C. Yang, and S. Liu. 2019. "MitoZ: A Toolkit for Animal Mitochondrial Genome Assembly, Annotation and Visualization." *Nucleic Acids Research* 47: e63.

Meyer, A. 1994. "Shortcomings of the Cytochrome b Gene as a Molecular Marker." *Trends in Ecology & Evolution* 9: 278–280.

Moritz, C., T. Dowling, and W. Brown. 1987. "Evolution of Animal Mitochondrial DNA: Relevance for Population Biology and Systematics." *Annual Review of Ecology and Systematics* 18: 269–292.

Nelson, H. V., K. A. Farquharson, A. Georges, et al. 2024. "A Genomic Framework to Assist Conservation Breeding and Translocation Success: A Case Study of a Critically Endangered Turtle." *Conservation Science and Practice* 6, no. 10: e13204.

Nordstrom, B., N. Mitchell, M. Byrne, and S. Jarman. 2022. "A Review of Applications of Environmental DNA for Reptile Conservation and Management." *Ecology and Evolution* 12: e8995.

Peel, E., L. Silver, P. Brandies, et al. 2022. "Best Genome Sequencing Strategies for Annotation of Complex Immune Gene Families in Wildlife." *GigaScience* 11: 1–15.

Rees, H. C., B. C. Maddison, D. J. Middleditch, J. R. Patmore, and K. C. Gough. 2014. "The Detection of Aquatic Animal Species Using Environmental DNA-a Review of eDNA as a Survey Tool in Ecology." *Journal of Applied Ecology* 51: 1450–1459.

Renshaw, M. A., B. P. Olds, C. L. Jerde, M. M. McVeigh, and D. M. Lodge. 2015. "The Room Temperature Preservation of Filtered Environmental DNA Samples and Assimilation into a Phenol–Chloroform–Isoamyl Alcohol DNA Extraction." *Molecular Ecology Resources* 15, no. 1: 168–176.

Riaz, T., W. Shehzad, A. Viari, F. Pompanon, P. Taberlet, and E. Coissac. 2011. "ecoPrimers: Inference of New DNA Barcode Markers From Whole Genome Sequence Analysis." *Nucleic Acids Research* 39: e145.

Ruppert, K. M., R. J. Kline, and M. S. Rahman. 2019. "Past, Present, and Future Perspectives of Environmental DNA (eDNA) Metabarcoding: A Systematic Review in Methods, Monitoring, and Applications of Global eDNA." *Global Ecology and Conservation* 17: e00547.

Satam, H., K. Joshi, U. Mangrolia, et al. 2023. "Next-Generation Sequencing Technology: Current Trends and Advancements." *Biology* 12: 997.

Schmidt, D. J., B. Brockett, T. Espinoza, M. Connell, and J. M. Hughes. 2016. "Complete Mitochondrial Genome of the Endangered Mary River Turtle (*Elusor macrurus*) and Low mtDNA Variation Across the Species' Range." *Australian Journal of Zoology* 64: 117–121.

Schmidt, T. L., J. A. Thia, and A. A. Hoffmann. 2024. "How Can Genomics Help or Hinder Wildlife Conservation?" *Annual Review of Animal Biosciences* 12: 45–68.

Schneider, H. 2023. "Integrating Genomics and Conservation to Safeguard Plant Diversity." *Integrative Conservation* 2: 10–18.

Spencer, R.-J., A. Georges, D. Lim, M. Welsh, and A. M. Reid. 2014. "The Risk of Inter-Specific Competition in Australian Short-Necked Turtles." *Ecological Research* 29: 767–777.

Stewart, K. A. 2019. "Understanding the Effects of Biotic and Abiotic Factors on Sources of Aquatic Environmental DNA." *Biodiversity and Conservation* 28: 983–1001.

Stoeckle, B. C., S. Beggel, A. F. Cerwenka, E. Motivans, R. Kuehn, and J. Geist. 2017. "A Systematic Approach to Evaluate the Influence of Environmental Conditions on eDNA Detection Success in Aquatic Ecosystems." *PLoS One* 12: e0189119.

Tamura, K., G. Stecher, and S. Kumar. 2021. "MEGA11: Molecular Evolutionary Genetics Analysis Version 11." *Molecular Biology and Evolution* 38: 3022–3027.

Tang, S., E. Peel, K. Belov, C. J. Hogg, and K. A. Farquharson. 2024. "Multi-Omics Resources for the Australian Southern Stuttering Frog (*Mixophyes Australis*) Reveal Assorted Antimicrobial Peptides." *Scientific Reports* 14: 3991.

Tessler, M., S. W. Cunningham, M. R. Ingala, S. D. Warring, and M. R. Brugler. 2023. "An Environmental DNA Primer for Microbial and Restoration Ecology." *Microbial Ecology* 85: 796–808.

Uliano-Silva, M., J. Nunes, and K. Krasheninnikova. 2021. marcelauliano/MitoHiFi: Mitohifi_v2.0.

Untergasser, A., I. Cutcutache, T. Koressaar, et al. 2012. "Primer3—New Capabilities and Interfaces." *Nucleic Acids Research* 40: e115.

Valdivia-Carrillo, T., A. Rocha-Olivares, H. Reyes-Bonilla, J. F. Domínguez-Contreras, and A. Munguia-Vega. 2021. "Integrating eDNA Metabarcoding and Simultaneous Underwater Visual Surveys to Describe Complex Fish Communities in a Marine Biodiversity Hotspot." *Molecular Ecology Resources* 21: 1558–1574.

Villacorta-Rath, C., T. Espinoza, B. Cockayne, J. Schaffer, and D. Burrows. 2022. "Environmental DNA Analysis Confirms Extant Populations of the Cryptic Irwin's Turtle Within Its Historical Range." *BMC Ecology and Evolution* 22: 1–14.

Wilcox, T. M., K. S. Mckelvey, M. K. Young, et al. 2016. "Understanding Environmental DNA Detection Probabilities: A Case Study Using a Stream-Dwelling Char *Salvelinus fontinalis*." *Biological Conservation* 194: 209–216.

Zhang, J., D. S. Finlaison, M. J. Frost, et al. 2018. "Identification of a Novel Nidovirus as a Potential Cause of Large Scale Mortalities in the Endangered Bellinger River Snapping Turtle (*Myuchelys georgesi*)." *PLoS One* 13: e0205209.

Zhang, X., P. J. Unmack, G. Kuchling, Y. Wang, and A. Georges. 2017. "Resolution of the Enigmatic Phylogenetic Relationship of the Critically Endangered Western Swamp Tortoise *Pseudemydura umbrina* (Pleurodira: Chelidae) Using a Complete Mitochondrial Genome." *Molecular Phylogenetics and Evolution* 115: 58–61.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

A4.3 Genome-wide diversity and MHC characterisation in a critically endangered freshwater turtle susceptible to disease.

The published PDF version of Chapter 4 of this thesis is presented on the following pages.

Immunogenetics (2025) 77:21
https://doi.org/10.1007/s00251-025-01378-8

ORIGINAL ARTICLE

Genome-wide diversity and MHC characterisation in a critically endangered freshwater turtle susceptible to disease

Holly V. Nelson[1] · Luke Silver[1,2] · Toby G. L. Kovacs[1] · Elspeth A. McLennan[1] · Arthur Georges[3] · Jane L. DeGabriel[4] · Carolyn J. Hogg[1,2] · Katherine Belov[1,2]

Received: 18 November 2024 / Accepted: 23 April 2025

Abstract
Small, isolated populations are often vulnerable to increased inbreeding and genetic drift, both of which elevate the risk of extinction. The Bellinger River turtle (*Myuchelys georgesi*) is a critically endangered species endemic to a single river catchment in New South Wales, Australia. The only extant wild population, along with the breeding program, face significant threats from viral outbreaks, most notably a nidovirus outbreak in 2015 that led to a 90% population decline. To enhance our understanding of genomic characteristics in the species, including genome-wide and functional gene diversity, we re-sequenced, assembled, and analysed 31 re-sequenced genomes for pure *M. georgesi* ($N=31$). We manually annotated the major histocompatibility complex (MHC), identifying five MHC class I and ten MHC class II genes and investigated genetic diversity across both classes in *M. georgesi*. Our results showed that genome-wide diversity is critically low in pure *M. georgesi*, contexualised through comparison with opportunistically sampled backcross animals—offspring of F1 hybrids (*M. georgesi* × *Emydura macquarii*) backcrossed to pure *M. georgesi* ($N=4$). However, the variation observed within the core MHC region of pure *M. georgesi*, extending across scaffold 10, exceeded that of all other macrochromosomes. Additionally, no significant short-term changes in either genome-wide or immunogenetic diversity were detected following the 2015 nidovirus outbreak (before; $N=19$, after; $N=12$). Demographic history reconstructions indicated a sustained, long-term decline in effective population size since the last interglacial period, accompanied by more recent steep declines. These patterns suggested that prolonged isolation and reduced population size have significantly influenced the dynamics of genome-wide diversity. It is likely that contemporary stressors, including the recent nidovirus outbreak, are acting on an already genetically depleted population. This study offers new insights into genome-wide and immune gene diversity, including immune gene annotation data with broader implications for testudines. These findings provide crucial information to support future management strategies for the species.

Keywords Conservation genomics · Immune genes · Bellinger River turtle · Major histocompatibility complex · Whole-genome re-sequencing

Introduction

The emergence of novel infectious diseases is currently a major threat faced by species of conservation concern (Daszak et al. 2000; Smith et al. 2009). Warming climates, pollution, and introduced species that harbour invasive pathogens are facilitating the spread of disease across wildlife populations (Anderson et al. 2004). Managing declining populations in the presence of infectious disease, in addition to other anthropogenic threats, is of growing conservation concern. Understanding the mechanisms of resilience and resistance, and the degree to which populations can adapt to change is a key step towards preventing extinctions (Auteri

✉ Carolyn J. Hogg
carolyn.hogg@sydney.edu.au

1 School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia

2 Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science Science, The University of Sydney, Sydney, NSW 2006, Australia

3 Institute for Applied Ecology, University of Canberra, Bruce, ACT 2617, Australia

4 NSW Department of Climate Change, the Environment,, Energy and Water, Parramatta, NSW 2150, Australia

Published online: 06 May 2025

and Knowles 2020). Another important aspect of conservation management is maintaining the adaptive potential of a species. Adaptive potential is the ability of a population to adapt to changing environmental conditions, such as disease, habitat modifications, and climate change (Hoffmann et al. 2017; Holderegger et al. 2006). Species' recovery programs are increasingly using genomic data to understand the adaptive potential of populations through functional gene analyses (McLennan et al. 2024), with growing emphasis on how managers can maintain or potentially increase adaptive potential in wild populations and conservation breeding programs (Farquharson et al. 2022). Advances in sequencing technologies and bioinformatic tools, and reductions in sequencing costs, have made whole-genome data more accessible for threatened species recovery programs. The combination of high-quality reference genomes with whole-genome sequencing data facilitates high-resolution analyses of neutral genetic variation, while also enabling the reconstruction of historical demographic trends and characterisation of functional gene families (Theissinger et al., 2023). For example, Magid et al. (2022) used 66 re-sequenced genomes to investigate toll-like receptors (TLR) immune gene diversity in shore plovers (*Thinornis novaeseelandiae*) and found low levels of diversity.

The major histocompatibility complex (MHC) is a crucial gene family within the adaptive immune system, primarily responsible for antigen presentation (Piertney and Oliver 2006). MHC molecules bind and present pathogen-derived peptides to T cells, facilitating the recognition and initiation of immune responses to pathogens and ensuring the specificity of adaptive immunity (Neefjes et al. 2011). MHC genes are broadly classified into classical and non-classical categories. Classical MHC genes, for example MHC class Ia and class II, are highly polymorphic and play a direct role in antigen presentation, influencing immune recognition and pathogen resistance (Le Bouteiller and Lenfant 1996). MHC class Ia and class II molecules are widely expressed across various tissues but are most prominent in antigen-presenting cells (Neefjes et al. 2011). In contrast, non-classical MHC genes (e.g., MHC class Ib, such as HLA-E, HLA-F, and HLA-G in humans) tend to be less polymorphic, typically function in immune regulation, and exhibit more tissue-specific expression (LeMaoult et al. 2003).

Variation within MHC genes has frequently been linked with species' susceptibility to disease, where higher levels of variation at MHC loci is thought to enhance a hosts ability to respond to a broader range of pathogen-derived antigens (Hughes and Yeager 1998; Penn et al. 2002). The association between MHC genes and disease resilience has been observed across multiple wildlife species including the northern leopard frog (*Rana pipiens*) (Trujillo et al., 2021), guignas (*Leopardus guigna*) (Napolitano et al. 2023) and desert bighorn sheep (*Ovis canadensis nelsoni*) (Dugovich et al. 2023), with disease-resilient populations exhibiting greater heterozygosity at MHC loci. The importance of the MHC in the immune cascade makes them excellent candidates to start to understand a populations' adaptive genetic diversity. For many threatened species, the lack of a high-quality reference genome and high-quality immune gene annotations has limited our ability to investigate the repertoire and variation present within the MHC region (Peel et al. 2022). Research on the MHC region in non-avian reptiles has included studies on the komodo dragon (*Varanus komodoensis*) (Reed and Settlage 2021) and the tuatara (*Sphenodon punctatus*) (Gemmell et al. 2020; Miller et al. 2005). More recently, genomic analyses in two anole species (*Anolis carolinensis* and *Anolis sagrei*) (Card et al. 2022) and the Chinese alligator (*Alligator sinensis*) (He et al. 2022b), have provided a more comprehensive characterisation of reptile MHC.

The critically endangered Bellinger River turtle (*Myuchelys georgesi*) is a medium-sized omnivorous turtle that is restricted to a 60 km stretch of Bellinger River, on New South Wales mid-north coast (Fig. 1). The Bellinger catchment, along with several other freshwater catchments in NSW, remains isolated and largely undisturbed. Cut-off by the ocean to the east and the Great Dividing Range to the west, these geographic boundaries have promoted unique habitat specialisation and catchment-specific speciation (Spencer et al. 2014). The species consists of a single wild population making it highly susceptible to stochastic events. In recent years, the species is estimated to have undergone a population decline of over 90% due to nidovirus outbreaks of unknown origin (Chessman et al. 2020; Zhang et al. 2018).

In 2015, a species-specific nidovirus outbreak (the Bellinger River virus; BRV) caused the population to crash from *ca.* 4000 individuals to *ca.* 200 and resulted in a significant decline in genetic diversity, as measured by reduced representation sequence data (Chessman et al. 2020; Nelson et al. 2024). The near extinction of the species led to urgent conservation efforts including the development of a *M. georgesi* conservation action plan, a range of community engagement initiatives, habitat restoration projects, and the establishment of a conservation breeding program comprised of two populations founded from 16 and 19 wild individuals. Since 2015, two lesser-documented viral outbreaks have occurred in the river, one in January of 2022 (Parrish et al. 2024) and a second in May of 2024 (NSW DCCEEW, *pers. comm.*, 2024). The ongoing disease outbreaks puts the species at high risk of extinction and requires investigation into potential genetic mechanisms that have led to this susceptibility.

Interestingly, one of the species' closest relatives, the Murray River turtle (*Emydura macquarii*), demonstrates resilience to the viruses. *E. macquarii* is widespread along the east coast of Australia, and it is hypothesised that the species has been introduced to the Bellinger River

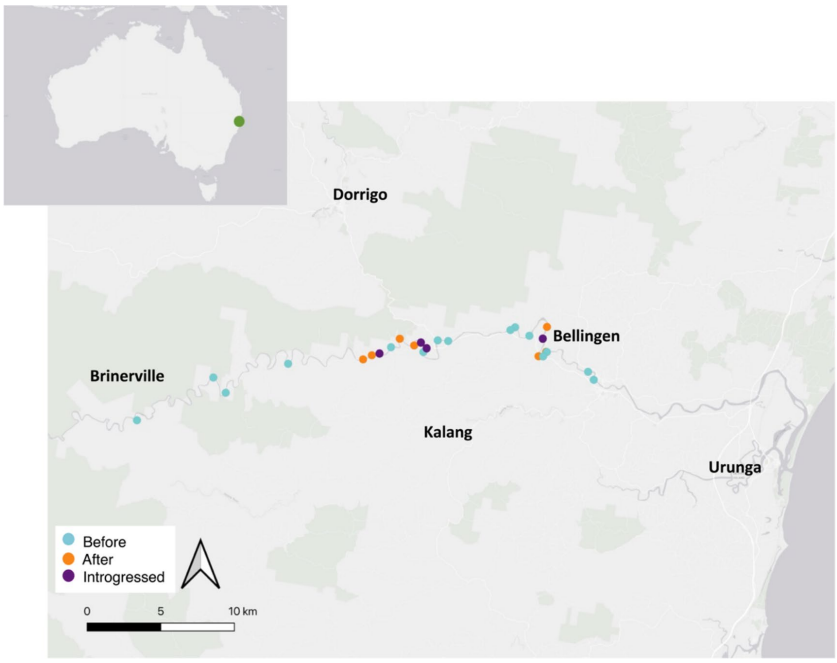


Fig. 1 An inset of Australia with a map of the Bellinger River basin showing the locations of historic (before: 2007, blue), contemporary (after: 2019, orange), and backcross (introgressed: 2019, purple) samples collected across the species' range (NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW) (unpublished data)

via human-mediated dispersal over the past two decades (Spencer et al. 2018). *E. macquarii* is known to hybridise with *M. georgesi* (Georges et al. 2018). Pure *E. macquarii*, F1 hybrids (*M. georgesi* × *E. macquarii*), and wild-caught F2 backcrosses to *M. georgesi*, previously identified by Georges et al. (2018), have tested positive to the recent viruses, but do not exhibit symptoms or succumb to the disease (Zhang et al. 2018). Given the high susceptibility of *M. georgesi* to nidovirus infections, conservation managers are looking for long-term viable options without the need for continual intensive management. A common strategy to enhance genetic diversity and potentially mitigate disease susceptibility is the introduction of individuals from genetically diverse or distinct populations (Frankham 2015; Frankham et al. 2017). Since this is not feasible for *M. georgesi*, alternative strategies may be required to enhance genetic diversity for long-term population viability such as interspecific hybridisation (Baack and Rieseberg, 2007), or the reintroduction of historical genetic variation lost over time through methods such as back-breeding, cloning, or genome editing (Shapiro 2017).

Previous work has shown that this species exhibits low genome-wide diversity (Nelson et al. 2024), characterised by limited intra-population differentiation and low allelic richness compared to other testudines populations (Buchanan et al. 2019; Fay et al. 2023). These findings, derived from reduced representation sequencing (RRS) data across 460 putatively neutral markers (Nelson et al. 2024), likely reflect the species' restricted geographic range, limited dispersal capacity, and long generation time. Additional analyses hypothesise that the species has undergone a number of historical bottlenecks, which have likely contributed to its low genetic diversity and ongoing disease susceptibility (Georges 2020).

In this study, we undertook manual MHC gene annotation by characterising the genetic features of MHC I and MHC II genes in *M. georgesi*. Building on previous research limited to 460 putatively neutral single nucleotide polymorphisms (SNPs), we use whole-genome re-sequencing (WGR) to analyse 12 contemporary turtles sampled after the disease outbreak, 19 historic turtles sampled before the outbreak, and four opportunistically sampled backcross individuals,

to undertake higher-resolution, genomic investigations. Here we aimed to (1) assess the effect of deep historical events on genome wide-diversity relative to backcross animals and effective population size, and (2) assess the effect of a recent bottleneck in the past 15 years due to a viral outbreak on genome-wide diversity and diversity within the MHC genes by comparing the before and after outbreak groups.

Methods

Immune gene annotation

We used a homology-based approach via BLAST v2.3.30 (Camacho et al. 2009) to manually characterise all MHC class I and II genes in the *M. georgesi* genome generated by Nelson et al. (2024) (rMyuGeo1.pri.20230808; GCA_040894355.1). The genome is 2 Gb, consists of 128 scaffolds and has an N50 of 123.4 Mb (Nelson et al. 2024). The high contiguity of the genome suggests that the MHC region is likely well assembled, providing a strong foundation for reliable MHC gene characterisation and downstream analyses (Peel et al. 2022). To annotate class I genes we acquired query sequences from the National Centre for Biotechnology Information (NCBI) including the tawny dragon (*Ctenophorus decresii*) (KY905241.1), caiman (*Caiman crocodilus*) (KF769542.1), marine iguana (*Amblyrhynchus cristatus*) (EU839663.1), Galapagos land iguana (*Conolophus subcristatus*) (EU604313.1), tuatara (DQ145788.1), and green sea turtle (*Chelonia mydas*) (OK135213.1) (Table S1). To annotate class II, we acquired query sequences from NCBI including the Chinese softshell turtle (*Pelodiscus sinensis*) (MT834970.1), marine iguana (FJ623752.1), and green-rumped parrotlet (*Forpus passerines*) (EF710746.1) (Table S1). We used query sequences from multiple organisms for manual annotation to improve gene predictions by increasing the likelihood of matching conserved MHC elements. Query sequences were used to search the *M. georgesi* genome and a global transcriptome generated by Nelson et al. (2024), using BLASTn and tBLASTn, respectively, with an *e*-value threshold of 1e-10. Exon splicing sites were manually checked by visualising against the reference genome, global transcriptome, and automated annotation in IGV v.2.16.0 (Robinson et al. 2011). Nucleotide sequences for each gene were then extracted from the reference genome using BEDtools v2.29.2 (Quinlan and Hall 2010) and input into MEGA v11 (Tamura et al. 2021). Nucleotide sequences were converted to protein sequences to ensure no stop codons were present within exons. Pairwise nucleotide similarity between genes was calculated using EMBL-EBI Clustal Omega (Madeira et al. 2019). To classify MHC genes as classical or non-classical, we analysed gene expression by aligning transcripts from brain, liver, and spleen tissues to MHC genes using IGV v.2.16.0. Phylogenetic relationships between genes within each class were inferred using sequences acquired from a maximum likelihood (ML) analysis performed in IQ- TREE2 (Minh et al. 2020). When possible, complete coding sequences for reptiles and amphibians were acquired from NCBI; however, predicted sequences were used for the inclusion of testudines data (Table S2, Table S3). The ModelFinder option (-m MFP) within IQTREE2 was used to select the best-fitting substitution model according to the Bayesian information criterion (Kalyaanamoorthy et al. 2017). Node support was assessed using the ultrafast bootstrap (-bb 1000) approximation and the like approximate likelihood-ratio test (-alrt 1000) (Guindon et al. 2010; Hoang et al. 2018).

Re-sequenced genome sampling, extraction, and sequencing

Re-sequenced genomes were generated using samples collected by the NSW Department of Climate Change, Energy, the Environment, and Water (DCCEEW) during surveys before the disease outbreak in April of 2007, after the outbreak in November 2019, and from F2 backcrosses to *M. georgesi* in November 2019 (Table S4). DNA samples were collected by removing part of the trailing webbing of the clawless toe on the hindfoot or by extracting blood from the jugular vein (Georges et al. 2018) (Table S4). Samples were stored in 75% ethanol at − 20 °C in the University of Canberra Wildlife Tissue Collection (GenBank UC < Aus >). For DNA extractions we performed a high salt method on 19 skin tissue biopsies and 16 dried bloods on Whatman card samples following a modified protocol from Aljanabi and Martinez (1997) that included an additional round of spin column centrifugation. DNA concentration and quality were assessed using a Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific) and 0.95% agarose gel electrophoresis for 30 min at 90 V. To maximise DNA quality for WGR, we undertook an additional DNA repair step using a FFPE DNA repair protocol (New England Biosciences). Repaired DNA concentration and quality were assessed using a Nanodrop 2000 Spectrophotometer (ThermoFisher Scientific) and 0.95% agarose gel electrophoresis for 30 min at 90 V. Samples were sent to Ramaciotti Centre for Genomics (University of New South Wales, Australia) for WGR. To minimise batch effects, all samples were processed simultaneously on an Illumina NovaSeq 6000, using a TruSeq DNA PCR free library prep kit across six lanes.

Re-sequenced genome alignment

Raw 150 bp paired-end FASTQ reads from 35 individuals were quality checked and trimmed using fastQC v0.11.8

(Andrews 2010) and trimmomatic v0.39 (Bolger and Giorgi 2014) with the parameters ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 SLIDINGWINDOW:4:5 LEADING:5 TRAILING:5 MINLEN:25. Reads were aligned to the reference genome using Burrows-Wheeler aligner (BWA) v0.7.17 (Li and Durbin 2009) 'mem' function with default parameters. The resulting alignment files were sorted into BAM format using SAMtools (Li et al. 2009) sort v1.6 and alignment rates and coverage calculated using SAMtools v1.6 'flagstat' and 'depth', respectively (Li et al. 2009) (Table S4). As individuals were sequenced across multiple lanes, BAM files pertaining to a single individual were merged using SAMtools merge and duplicates marked and removed using picard v2.21.9 MarkDuplicates (http://broadinstitute.github.io/picard/). For genome-wide investigations, we partitioned our data into three putative groups (two temporal, reflective of groups investigated by Nelson et al. (2024)); wild individuals sampled before the disease outbreak in 2007, $N = 19$ (hereafter 'before'); wild individuals sampled after the disease outbreak between 2015–2020, $N = 12$ (hereafter 'after'); and F2 backcrosses to *M. georgesi*, $N = 4$ (hereafter 'backcross'). For demographic history and MHC characterisation, the backcross group was excluded to avoid confounding effects on historical demography. Additionally, within the scope of this study, the absence of *E. macquarii* annotations, combined with the limitations of short-read data, posed a risk to the reliability of accurate variant calls (Figure S1).

Genome-wide diversity

To investigate genome-wide diversity, autosomal heterozygosity and runs of homozygosity (ROHs) were calculated across each genome ($N = 35$) using ROHan (Renaud et al. 2019). ROHan combines a local Bayesian model and hidden Markov model (HMM) to identify autosomal heterozygosity and ROHs from BAM files of individually mapped genomes (Renaud et al. 2019). Analyses were run on nine macrochromosomes, excluding the putative sex chromosome (scaffold 4) which was identified through genome coverage analyses (Nelson, unpublished data) and inferred from the *E. macquarii* karyotype (Martinez et al. 2008). We ran ROHan with three parameters; –rohmu 5e-5, –TsTv 1.965, -t 4 and –size 100,000 (100 kb windows) on an Amazon Web Services ubuntu 20.04 LTS cloud machine (r5.8 × large, 32 vCPU, 256 Gb RAM, 1 TB attached storage). Inbreeding coefficient based on ROH (F_{ROH}) (McQuillan et al. 2008) and average individual-level heterozygosity was calculated for each group using the hmmrohl and summary texts output by ROHan, respectively. Results were visualised in R v4.3.0 using dplyr() v1.1.2. and ggplot2() v3.5.0. We also calculated the number of genome-wide variants using BCFtools v1.3.1 'call' (Danecek et al. 2021).

Demographic reconstructions

Changes in ancient effective population size (N_e) over time were assessed with the pairwise sequentially Markovian coalescent (PSMC) model (Li and Durbin 2011) using sorted BAM files for pure *M. georgesi* ($N = 31$). For demographic reconstructions, sites with coverage below one-third or above twice the sample's average coverage were removed following the approach of Bi et al. (2020). Additionally, we included only scaffolds > 50 kb and excluded the sex chromosome. Individual consensus genome sequences containing SNP variants were generated for 31 M. *georgesi* using SAMtools v1.6 with the 'mpileup' command. 'mpileup' was used to directly represent sequencing variation without additional filtering, minimising potential biases in demographic inference (Li et al. 2009). For PSMC analyses, the files were processed using 'vcfutils.pl' and 'vcf2fq' to call SNP variants, following the recommended protocol (https://github.com/lh3/psmc). PSMC v0.6.5-r67 was run with the following default parameters, -N25 -t15 -r5 -p '4 + 25 × 2 + 4 + 6'. Results were scaled by a mutation rate of 4.61×10^{-9} substitutions per site per generation based on estimates for *Chrysemys picta* (Bergeron et al. 2023). As precise details of the life history of *M. georgesi* are not available, we used a generation time of 14 years based on manager recommendations. To investigate recent historical patterns of demography we calculated N_e using GoNe (Santiago et al. 2020). Following the SAMtools v1.6 'mpileup' step mentioned above, we used BCFtools 'call' for multi-sample variant calling. For stringent filtering, we retained variants with a MAF of > 0.04, where the alternate allele must be seen twice in the population, which could be homozygous in a single individual or heterozygous in two individuals. As a result, 2,979 variants were removed from the total of 3,681,055 (Table S8), as they were more likely to represent genome-wide sequencing errors than true variants in this genetically homogeneous species. We conducted multiple iterations of GoNe with default parameters to determine the optimal values for *hc* (0.01, 0.05, 0.1) and various recombination rate estimates (1, 2, 3). The optimal values were identified as a recombination rate of 3 cM/Mb and an *hc* of 0.01. N_e outputs were visualised using ggplot2() v3.5.0 in R v4.3.0.

MHC gene variant calling and analyses

To characterise MHC diversity in the before and after samples ($N = 31$), we generated a joint-genotyped dataset. This was done using GATK GenomicsDBImport to create a sample database, followed by GATK GenotypeGVCF to produce a whole-genome joint-genotyped multi-sample VCF file containing both temporal groups. As previously mentioned, backcross animals were excluded from these analyses due to the absence of manually annotated MHC

genes in *E. macquarii* and concerns about the reliability and confidence of short-read alignment and variant calls in the MHC core region. This is because uncharacterised gene duplication and structural factors can introduce artifacts when interspecific short reads are mapped to the genome (Jaegle et al. 2023) (Table S4), resulting incomplete or inaccurate variant identification and interpretation. We filtered the joint genotyped multi sample VCF using VCFtools v0.1.14 (Danecek et al. 2011), retaining SNPs, multiple nucleotide polymorphisms (MNPs), non-biallelic SNPs, and indels, to obtain raw variant counts within our manually annotated MHC exons using BCFtools (Table 2). Among these variant types, SNPs were the most abundant, providing the highest level of variation and making them the most informative for downstream analyses. For SNP analyses, we filtered the VCF to include only biallelic SNPs found within the MHC exons. We applied GATK VariantFiltration and SelectVariants v4.2.0.0 (McKenna et al. 2010) using GATK-recommended thresholds to remove variants with MQ < 40, − 12.5 < MQRankSum < 12.5, − 8 < ReadPosRankSum < 8, and a stringent QUAL score of < 80. BCFtools was then used to filter out sites with an average depth across all samples of < 10 and those with an allelic balance > 0.9. Finally, we used VCFtools to retain SNPs with a minor allele frequency (MAF) > 0.01, to retain rare variants within the MHC region. Using the filtered biallelic SNP set, we then calculated the number of SNPs in each gene and determined if they resulted in synonymous or non-synonymous amino acid substitutions using Geneious Prime 2020 (https://www.geneious.com/). To identify full length allele sequences for each MHC gene in each individual, we undertook phasing. First, we used GATK FastaAlternateReferenceMaker (v4.2.0.0) (McKenna et al. 2010) to generate single sample FASTA sequences for each gene. SeqPHASE (Flot 2010) was used to convert the FASTA sequences into phase format then PHASE (v.2.1.1) (Stephens et al. 2001; Stephens and Scheet 2005) was run to generate alleles and SeqPHASE used to convert phase output into phases FASTA sequences to give 70 sequences for each gene (two per individual). We calculated observed (H_o) and expected heterozygosity (H_E) across the complete MHC I and MHC II exon sequences in temporal (before and after) using GenAlEx v6.5 (Peakall and Smouse 2006) and visualised average individual H_o by group using the boxplot() functions in R. To assess genetic differentiation among individuals, we generated principal component analysis (PCoA) plots using our genome-wide and MHC joint genotyped VCFs, using the adegenet package (Jombart 2008) in R (v4.1.1) (R Core Team 2023). Comprehensive details of sample metadata and a breakdown of the methodological approaches used in this study are shown in Table S4 and Figure S1, respectively.

Results

MHC annotation

Through manual annotation of immune-related genes, we identified five MHC class I loci (*Myge-UA*, *Myge-UB*, *Myge-UC*, *Myge-UD*, *Myge-UE;* labelled from most to least number of detected variants), and ten MHC class II loci (*Myge-DAA1*, *Myge-DAB1*, *Myge-DAA2*, *Myge-DAB2*, *Myge-DAA3*, *Myge-DAB3*, *Myge-DAA4*, *Myge-DAB4*, *Myge-DAA5*, *Myge-DAB5*), all located within a 272,213 bp region on scaffold 10. The gene nomenclature adheres to the conventions established by Miller et al. (2005) and Miller et al. (2015) for a reptilian species; however, limited annotations and published sequences in other reptile species made comparative interpretation challenging. The *M. georgesi* MHC class II genes are densely clustered, consisting of alpha (here labelled *DAA*) and beta (here labelled *DAB*) chains encoded on the 3′ to 5′ and 5′ to 3′ strands, respectively (Table 1). Structurally, the class II genes exhibit high conservation, except for *Myge-DAB5* which shows a putative loss of exon 5 with no transcriptional evidence across three tissue types (Fig. 2). Similarly, MHC class I genes display a largely conserved architecture, except for *Myge-UA*, which has notably larger introns 1 and 2 compared to other MHC class I genes and contains an additional exon (Fig. 2; Table S5). Differentiation analyses revealed high sequence similarity between *Myge-UB* and *Myge-UE*, whereas *Myge-UD*, *Myge-UA*, and *Myge-UC* exhibited closer sequence similarity (Table S6 A). As expected, lower levels of differentiation were observed within the MHC II alpha (*DAA*) and beta (*DAB*) gene groups compared to between these gene groups (Table S6B). Transcriptome data did not provide a clear distinction between classical and non-classical Class I genes, as transcript levels across the three tissue types (brain, liver, and spleen) were consistent for all genes, suggesting that additional tissue types may be needed to identify if any of the identified class I genes may have non-classical functions.

To contextualise the genetic relationships among MHC homologs in *M. georgesi* and extend these comparisons across non-avian and avian reptiles, we constructed phylogenetic trees using complete coding sequences of MHC I and MHC IIB sequences. The resulting phylogenies showed that *M. georgesi* forms monophyletic groups with other testudines across both classes, with certain *M. georgesi* genes forming unique clades alongside other testudines (*Myge*-UA, *Myge*-UB, *Myge*-UE, *Myge*-DAB4) (Figure S2; Figure S3). Furthermore, testudines and members of the Crocodilia were shown to form a monophyletic group, separate from Squamata, Aves, and Rhynchocephalia in the MHC class II tree (Figure S3).

Table 1 Manually annotated MHC I and MHC II genes on scaffold 10 used for downstream analyses. Start and end coordinates refer to whole-genome coordinates on the reference genome (GCA_040894355.1). Coordinates for each exon are presented in table S5

Class	Gene	Scaffold	Exons	Start	End	Strand
MHC I	*Myge_UB*	10	8	30,471,574	30,493,484	-
	Myge_UE	10	7	30,500,310	30,508,525	-
	Myge_UC	10	7	30,533,405	30,538,940	-
	Myge_UD	10	7	30,567,953	30,577,665	-
	Myge_UA	10	7	30,583,497	30,591,368	-
MHC II	*Myge_DAA1*	10	4	30,660,982	30,663,487	-
	Myge_DAB1	10	6	30,665,109	30,669,452	+
	Myge_DAA2	10	4	30,677,698	30,680,223	-
	Myge_DAB2	10	6	30,681,654	30,686,193	+
	Myge_DAA3	10	4	30,698,489	30,700,988	-
	Myge_DAB3	10	6	30,702,397	30,706,348	+
	Myge_DAA4	10	4	30,717,695	30,720,191	-
	Myge_DAB4	10	6	30,721,446	30,725,542	+
	Myge_DAA5	10	4	30,735,202	30,737,782	-
	Myge_DAB5	10	5	30,739,701	30,743,969	+

MHC diversity

Following the identification and filtering of MHC variants, we identified 257 biallelic SNPs across all MHC exons in the before group, including 127 SNPs in class I and 136 SNPs in class II genes (Table 2). In the after group, 232 biallelic SNPs were identified across all MHC exons, comprising 108 SNPs in class I and 124 SNPs in class II genes (Table 2). The SNPs were unevenly distributed among MHC genes, with highest number of variants contained in *Myge-UA* ($N = 66$) and the lowest number of variants contained in *Myge-DAB3* which was monomorphic (Table 2). Overall, the 14 genes with variants had an average of 18 exonic SNPs (range: 3–66), most of which were in exons 2, 3, and 4 across both classes. Among the 257 SNPs identified in the 'before' animals, 171 (67%) were predicted to result in non-synonymous substitutions. In the 'after' animals, 154 of the 232 (59%) detected variants were also predicted to be non-synonymous. Non-synonymous variation was observed across 14 genes in the 'before' group and 13 genes in the 'after' group, respectively. A total of 15 and 13 alleles were observed in MHC I and 35 and 27 alleles were observed in MHC II in before and after groups, respectively (Table 2). The before population had one private allele in *Myge-UC* and *Myge-UE,* and 11 private alleles across *Myge-DAA1*, *Myge-DAA2, Myge-DAB4, Myge-DAA5, Myge-DAB5* that were unique to that population. We found no significant differences in allelic diversity between temporal groups (MHC I: $t = 1.8711$, df $= 8$, p-value $= 0.09823$; MHC II: $t = 1.688$, df $= 16$, p-value $= 0.1108$) (Table 2). On average, class I genes had higher numbers of SNPs, non-synonymous SNPs, and fewer alleles compared to class II. SNP H_O and H_E was similar across temporal groups (Table S9) with slightly lower observed than expected heterozygosity at MHC loci. Higher number of variants corresponded with higher levels of heterozygosity across both classes in pure *M. georgesi*. PCoA analyses revealed no clear temporal differentiation in MHC genes, although clustering of individuals was observed, potentially reflecting the presence of shared alleles (Figure S6). Slightly greater differentiation was observed in the 'before' group for MHC I genes (Figure S6 A), indicating a minor degree of temporal structuring. In contrast, genome-wide analyses showed no discernible clustering patterns. The inclusion of backcross individuals resulted in tighter clustering of purebred animals, reflecting the nearly identical genome-wide diversity within the species and the absence of distinct genetic structuring (Figure S7 A). This observation was further validated by the exclusion of backcross individuals, which revealed no significant clustering or differentiation among pure groups, suggesting minimal genetic divergence over time (Figure S7B).

Genome-wide diversity and demographic history

All whole-genome samples had an average alignment rate of 99.6% and 99.5% for pure and backcross animals, respectively, across ten putative macrochromosomes (scaffolds 1–10). Raw SNP calls across the genomes were substantially higher in backcross animals (~ 15,000,000) compared to pure *M. georgesi* (~ 650,000) (Table S7). ROH were distributed across all macrochromosomes (Fig. 3A). The longest Run of Homozygosity (ROH) was identified on scaffold 5 in purebred individuals and covered 36% of the scaffold's total length (Fig. 3A). In contrast, scaffold 10 exhibited the lowest density of ROH relative to its size (Fig. 3A). The pure samples had 90% of their genome in ROH, characterised by short (< 2 Mb) and long (> 2 Mb) ROH as defined by Ceballos et al. (2018), while only a

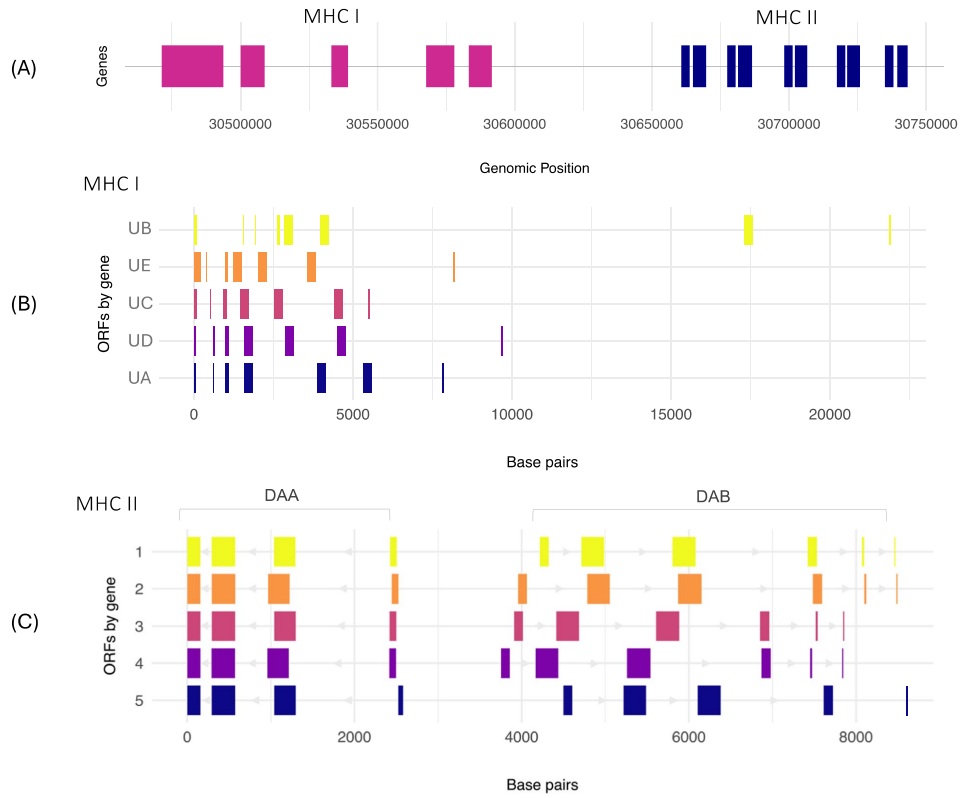


Fig. 2 Genomic architecture of major histocompatibility complex (MHC) genes. **A** Genomic architecture of the core major histocompatibility complex (MHC) region, including class I and class II genes, on scaffold 10. **B** and **C** Genomic coordinates and structural organisation of manually annotated *Myuchelys georgesi* (*Myge*) MHC genes presented as open reading frames (ORFs), in order of genomic position along scaffold 10

small number of short ROH were detected in backcross individuals (Fig. 3A, Table S7, Figure S4, Figure S5). Small ROHs measuring 0–500 kb, 500–1000 kb, and 1000–2000 kb had similar frequencies across groups (Figure S5 A). Although counts of small and long ROH were relatively similar, the proportion of the genome in long runs was notably higher than small runs (Figure S5B). This is consistent with the expectation that larger ROHs occupy a greater proportion of the genome. The before and after samples had an identical proportion of the genome in ROH and comparable ROH lengths. The average number of ROH per individual was also consistent between the two temporal groups. No significant differences in nucleotide diversity across the genome were observed between the before and after individuals, regardless of whether ROH were included or excluded. Both temporal groups displayed significantly lower nucleotide diversity compared to the backcross individuals (Table S7).

We identified significantly lower levels of genome-wide SNPs (Table S7) and autosomal heterozygosity in pure versus backcross animals (pure: 1.18×10^{-4}; backcross: 6.8×10^{-3}, $t = -16.01$, $p = 0.001$) (Fig. 4A). No significant difference was observed between temporal groups (before: 1.17×10^{-4}; after: 1.21×10^{-4}, $t = -1.40$, $p = 0.1745$) (Fig. 4B).

Our demographic history reconstructions using PSMC analysis indicate a gradual, long-term decline in effective population size beginning approximately 110,000 years ago, coinciding with the last glacial period (Fig. 5A). The largest effective population size estimates correspond to a period before the last interglacial, around 110,000 years ago. These estimates were consistent across all 31 individuals analysed. Estimates predating this timeframe

Table 2 Temporal variant statistics for each MHC gene investigated including number of indels; non-biallelic single nucleotide polymorphisms (SNPs); multiple nucleotide polymorphisms (MNPs); filtered SNPS; non-synonymous SNPs (ns); ratio of synonymous to non-synonymous SNPs (dN/dS); number of alleles; and allelic diversity. Before $N = 19$, after $N = 12$. Zero denoted by '-'

			Before								After							
Class	Gene (*Myge*)	ORF length (bp)	Indels	Non-biallelic SNPs	MNPs	SNPs	SNPs (ns)	SNP dN/dS	No. alleles	Allelic diversity	Indels	Non-biallelic	MNPs	SNPs	SNPs (ns)	SNP dN/dS	No. alleles	Allelic diversity
MHC I	*UA*	1063	-	-	-	66	40	0.606	4	0.610	-	-	-	60	40	0.606	3	0.513
	UB	1124	4	-	-	23	16	0.696	3	0.4–92	4	-	-	23	16	0.696	3	0.444
	UC	1096	-	-	-	18	15	0.833	4	0.495	-	-	-	12	9	0.750	3	0.513
	UD	1096	-	-	-	8	6	0.750	2	0.499	-	-	-	8	6	0.750	2	0.555
	UE	1196	1	2	-	6	1	0.167	2	0.512	1	2	-	5	-	0.000	2	0.555
MHC II	*DAA1*	774	-	-	1	16	13	0.813	7	0.697	-	-	1	13	9	0.563	5	0.506
	DAB1	805	-	-	1	12	1	0.083	2	0.493	-	-	1	12	1	0.083	2	0.444
	DAA2	774	-	-	-	3	3	1.000	3	0.370	-	-	-	3	3	1.000	2	0.180
	DAB2	807	-	-	-	16	12	0.750	3	0.553	-	-	-	11	9	0.563	3	0.537
	DAA3	774	-	-	-	4	4	1.000	2	0.347	-	-	-	4	4	1.000	2	0.297
	DAB3	804	-	-	-	-	N/A	N/A	N/A	N/A	-	-	-	-	N/A	N/A	N/A	N/A
	DAA4	775	-	-	-	29	22	0.759	3	0.590	-	-	-	27	21	0.724	3	0.509
	DAB4	807	-	-	-	17	13	0.765	5	0.674	-	-	-	15	11	0.647	4	0.675
	DAA5	751	-	-	-	27	17	0.630	7	0.763	-	-	-	27	17	0.630	4	0.555
	DAB5	777	1	-	-	12	8	0.667	3	0.562	1	-	-	12	8	0.667	2	0.486

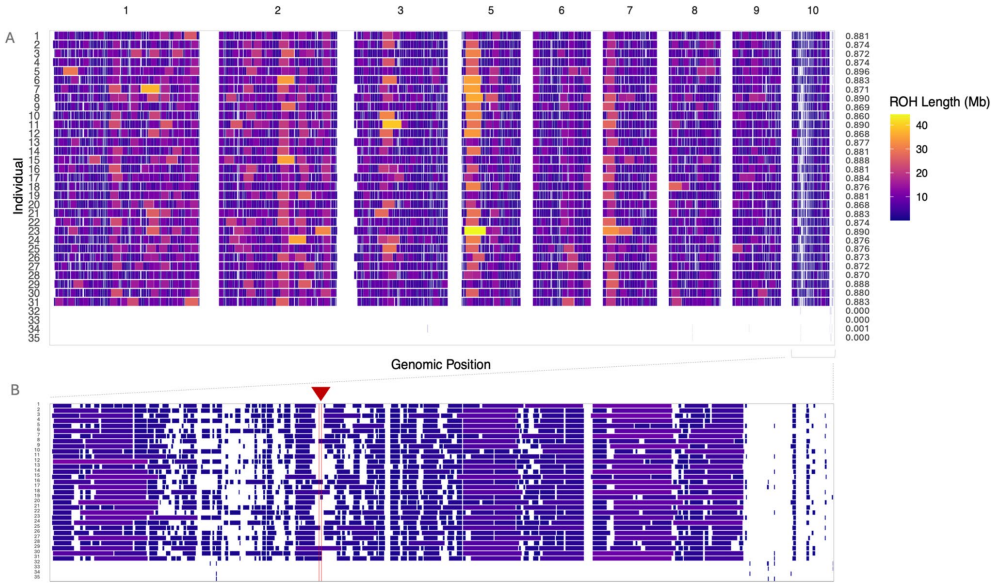


Fig. 3 Runs of homozygosity (ROH) heatmap illustrating **A** the chromosomal distribution, length, and the proportion of the genome in ROH across nine macrochromosomes in before (individuals 1–19) and after (individuals 20–31), relative to F2 backcrosses (individuals 32–35). **B** Enlarged view of scaffold 10. Major histocompatibility complex (MHC) core region marked by a red arrow and parallel lines

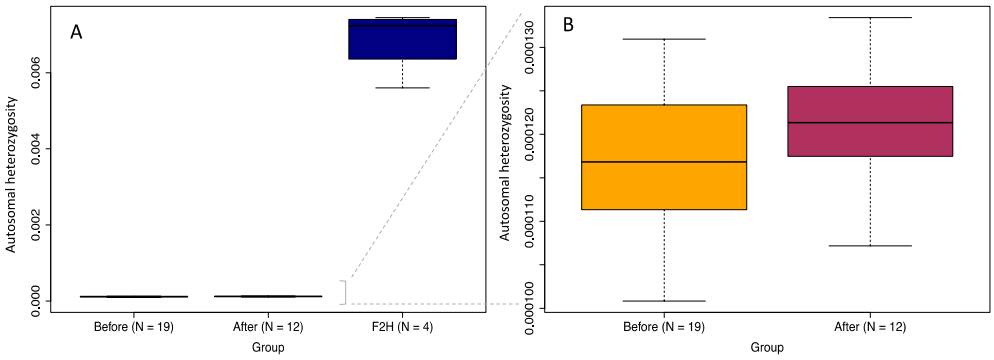


Fig. 4 Genome-wide heterozygosity estimates for *Myuchelys georgesi* calculated using ROHan. Boxes represent the interquartile range (IQR) with whiskers extending to the upper and lower ranges and the bold line representing the group mean. **A** Autosomal heterozygosity estimates for all groups including before, after, and backcross. **B** Rescaled autosomal heterozygosity estimates of pure *M. georgesi* before and after groups plotted in A

are likely less reliable due to including the accumulation of errors in coalescent events in deep time and limited resolution in ancient periods; therefore, they should be approached with caution. GoNe analyses showed there has been a decline in effective population size from an

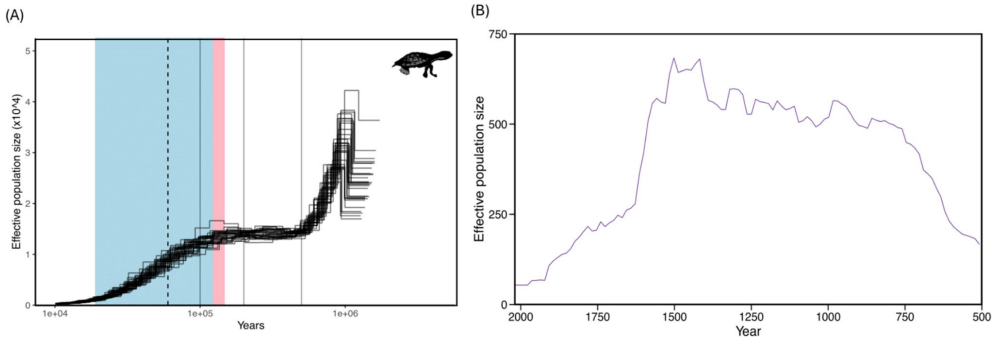


Fig. 5 Changes in effective population size (N_e) over time estimated using **A** PSMC and **B** GoNe for *Myuchelys georgesi*. The x axes indicate time before present in years, and the y axis indicates the effective population size. A generation time of 14 years was used for both analyses. For PSMC analyses, axes were scaled by mutation rate of 4.61 $\times 10^{-9}$ substitutions per site per generation. PSMC rectangles correspond to the last inter-glacial (red, warm) and glacial period (blue, cold). Vertical dotted line represents arrival of humans to Australia

estimate N_e of 700, 50 generations ago to 100 in the current generation (Fig. 5B).

Discussion

Here we provide a critical first step towards understanding the genome-wide and immune gene diversity in the Bellinger River turtle. We found low levels of genome-wide diversity in pure *M. georgesi* compared to backcross animals and evidence of long-term, continual declines in effective population size. We present the first comprehensive overview of the location and structure of the MHC region in a freshwater turtle species. Based on our manual annotations of the MHC genes, we observed relatively low levels of genetic variation within this gene region in pure *M. georgesi* relative to disease-resilient hybrids. However, scaffold 10, which encompasses the MHC region, exhibited higher levels of variation compared to other macrochromosomes. Consistent across both genome-wide and immune gene analyses, no significant changes in genetic diversity were detected before and after the disease outbreak.

Leveraging a scaffold-level genome assembly and RNAseq data, we identified MHC genes localised within a single core region, consistent with observations in many amniote species (Fig. 2A). Like *M. georgesi,* these studies revealed a single-core MHC region with linked class I and class II subregions. While the relocation of MHC genes to other genomic regions has been observed in reptiles—for instance, the displacement of class I and class II genes from the core MHC region in tuatara and anole (Card et al. 2022; Gemmell et al. 2020; Miller et al. 2015) —we did not find evidence for extensive duplication of functional MHC genes outside of the core region, aside from incomplete pseudogenes lacking major exons. This suggests a highly conserved MHC region with potential constraints on duplication to other regions of the genome in *M. georgesi*. The genomic organisation of each class forms two distinct clusters observed within the single-core MHC region as observed in the Chinese alligator (He et al. 2022b) and komodo dragon (Reed and Settlage 2021), with no interspersal of classes as observed in the tuatara (Gemmell et al. 2020), and some amphibian and mammalian species (He et al. 2023; Peel et al. 2019). These topologies align with the known evolutionary relationships within Reptilia (Zardoya and Meyer 2001), supporting both the conservation of MHC loci and their divergence along lineage-specific trajectories. Phylogenetic analyses of immune genes similarly reflect these established relationships, with both MHC classes showing the closest divergence from Crocodilia (Figure S2, Figure S3). This raises the question of whether these clades represent classical and non-classical genes or suggest specialised roles within the *M. georgesi* immune system. In humans, for instance, classical HLA (human leukocyte antigen) class Ia genes, such as HLA-A, are involved in antigen presentation to immune cells, are highly polymorphic, and play a critical role in adaptive immunity (Le Bouteiller and Lenfant 1996). In contrast, non-classical HLA class Ib genes such as HLA-G and HLA-E exhibit specialised immunomodulatory functions, are highly conserved and characterised by limited polymorphism, and are predominantly expressed in immunologically significant tissues (LeMaoult et al. 2003). Consistent with the genomic architecture observed in amphibians, MHC class I introns in *M. georgesi* are notably longer than those of MHC class II, where MHC II DAA and MHC II DAB are tightly linked (He et al. 2023). These

keys features suggest that the MHC region in *M. georgesi* is relatively simple as the core MHC I and MHC II regions tightly clustered and adjacent in the genome. This simplified genomic arrangement appears to have been largely conserved throughout the evolutionary history of tetrapods, including birds (He et al. 2022a), reptiles (Card et al. 2022; He et al. 2022b; Miller et al. 2015), and mammals (Silver et al. 2024), particularly in MHC II genes.

Our demographic reconstructions indicate that the effective population size of *M. georgesi* has been declining since the last interglacial period (Fig. 5), potentially due to limited gene flow and accumulation of runs of homozygosity. Mitochondrial data estimate that *M. georgesi* diverged from its close relative, *E. macquarii*, approximately 6.1 million years ago during a period of aridification (Le et al. 2013). Geographic isolation and habitat fragmentation likely contributed to the demographic trends identified in our PSMC analyses. The low N_e estimates (< 500) from approximately 10,000 years ago (Fig. 5A) suggest that climatic changes during the last glacial period may have influenced the genetic structure of *M. georgesi,* similar to patterns observed in other freshwater turtle species (Hilgers et al. 2024). However, unlike these common turtle species, which eventually reach an equilibrium in N_e (Hilgers et al. 2024), *M. georgesi* has continued to experience a gradual, long-term decline. The glacial period likely reduced habitat availability and fragmented freshwater populations, leading to isolation, speciation, and long-term declines in effective population size. The species isolation, combined with life-history traits, have likely intensified the impact of accumulated inbreeding, resulting in reduced fitness and an elevated risk of extinction over time (Frankham et al. 2017). The long-term isolation and gradual decline in effective population size represent key findings for the species, likely exacerbating the critically low levels of contemporary genome-wide diversity observed in heterozygosity and ROH analyses. Consequently, this accumulation of inbreeding effects, compounded by geographic barriers to dispersal, may have driven the species into an extinction vortex well before the onset of the Anthropocene. This is supported by the large proportion of the genome found in ROH (Fig. 3A), as long-term inbreeding and gene flow can directly influence ROH abundance (Ceballos et al. 2018; Foote et al. 2021; Hewett et al. 2023; Mooney et al. 2021). Short ROH are indicative of background relatedness or inbreeding resulting from distant common ancestry, while long ROH reflect recent parental relatedness or occur in genomic regions with low recombination rates (McQuillan et al. 2008; Pemberton et al. 2012). In *M. georgesi*, the ROH distribution includes both short (< 2 Mb) and long (> 2 Mb) segments (Fig. 3A, Figure S4, Figure S5). The elevated levels of homozygosity observed in the contemporary population are likely the consequence of both historical inbreeding among distant ancestors and recent background relatedness. This is further supported by the recent declines identified through GoNe analyses, which have likely contributed to the significant proportion of the genome comprising of long ROH (Fig. 5B, Figure S5B, Table S5) (Kardos et al. 2021; Pemberton et al. 2012). ROHs, which are typically regions identical by descent, are likely absent in backcross animals due to increased genetic diversity and reduced inbreeding in parental *E. macquarii* in addition to heterozygote advantage and levels of genetic divergence (7.8%) between the two species (Fielder et al. 2012). The absence of significant differences in both the number and length of ROH and in autosomal heterozygosity across temporal groups suggests that genomic signatures associated with the recent disease outbreak are not yet detectable at the genome-wide level in a species with long generation times. These findings conflict with previous analyses using 460 SNPs and a sample set of before = 92 and after = 38 (Nelson et al. 2024). As WGR leverages a dense array of markers spread across an entire genome, compared to the RRS data that surveyed 0.000023% of the genome, the increased statistical power and even genomic coverage of the genome-wide dataset yields results that are more robust and reliable (Jeon et al. 2024). Although these signatures are not apparent in the genome-wide data yet, these signatures may emerge in subsequent generations, potentially increasing the species' trajectory towards extinction, emphasising the importance of ongoing genetic monitoring.

Interestingly, the terminal region of scaffold 10 shows a notable absence of ROH, with only a few ROH observed in certain individuals, including backcrosses (Fig. 3). This pattern could potentially result from various factors, such as telomeric recombination disrupting long ROH tracts at the chromosomal ends (Bosse et al. 2012), or the presence of a functionally important gene region as observed in our characterisation of the MHC core region. However, further investigation would be necessary to confirm this.

Here we found higher levels of expected (H_E) than observed (H_O) at MHC loci. We previously found higher H_O than H_E heterozygosity in the species using 460 putatively neutral SNP markers (Nelson et al. 2024). Although neutral genetic variation is frequently employed as a proxy for adaptive potential, our findings support the argument that neutral diversity may not fully capture adaptive potential, particularly when only a small set of RRS variants is considered. Instead, genetic diversity at specific functional loci provides a more accurate measure of a population's adaptive potential (Holderegger et al. 2006; Teixeira and Huber 2021). MHC variants play a critical role in a wide range of biological traits. For instance (i) immune recognition where a single amino acid change in the antigen-binding region of the DRB*1302 allele in humans abolishes malaria recognition (Frank 2002; Summers et al. 2003); (ii) susceptibility to infectious and autoimmune diseases, as observed

in MHC IIB heterozygotes, which exhibit lower Ranavirus infection intensity compared to homozygotes in larval wood frogs (*Rana sylvatica*) (Savage et al. 2019); (iii) individual odours and mating preferences seen in song sparrows (*Melospiza melodia*) (Grieves et al. 2019) and other aves (Leclaire et al. 2014), where preen oil odour is used to discern MHC similarity and diversity of potential kin or mates and; (iv) pregnancy outcomes observed in giant pandas (*Ailuropoda melanoleuca*), where mating pairs with MHC dissimilarity exhibit higher reproductive success (Sommer 2005; Zhu et al. 2019). High levels of inbreeding often lead to an overall decrease in MHC variants, loss of rare and potentially advantageous alleles, and a decreased ability to adapt to novel or rapidly evolving pathogens (Altizer et al., 2003; Spielman et al., 2004). Despite the low genome-wide diversity observed in *M. georgesi*, MHC genes appear relatively conserved. The lack of ROH seen in the core MHC region, and scaffold 10 in general, relative to other macrochromosomes suggests that SNP diversity alone, may not be a primary driver of the species' disease susceptibility as initially hypothesised. Instead, the interspecific diversity introduced by *E. macquarii* across the genome more broadly, including other variant types not explored in this study, may contribute to the disease resilience of backcross individuals compared to their pure *M. georgesi* counterparts; however, this would need further investigation (Parrish et al. 2024; Yang et al. 2024). Individuals from the after disease group exhibit a reduced number of SNPs, non-synonymous mutations, and alleles at six out of the 15 MHC genes when compared to the before group. These observations may reflect a possible sampling artifact or could represent early evidence of allele loss in the wild, with some of the remaining alleles potentially conferring resistance to nidovirus infection, as seen in amphibians with chytridiomycosis (Fu et al. 2023). However, longer-term monitoring of subsequent generations in addition to metadata of resilient and animals that succumb to the disease will be necessary to validate these findings.

The results of this study should be interpreted with consideration of its limitations. The first limitation is the scarcity of complete coding sequences for testudines and other reptilian species, which hinders the ability to draw high-resolution conclusions about the evolutionary histories and classification of MHC homologs in *M. georgesi* and testudines more broadly. Compared to the extensive genomic and immunological resources available for other tetrapod clades, such as marsupials (Belov et al. 2013), reptiles—particularly testudines—still lack comprehensive immune gene resources. This deficiency in data restricts the ability to resolve evolutionary relationships of ancient lineages effectively. As next-generation sequencing (NGS) becomes more common for non-model organisms, the increasing availability of genomic data and resources will help address the current deficiency in MHC nomenclature and classification. Secondly, are comparisons between pure and backcross animals. Determining whether discrepancies between pure and backcross animals arise from misalignments due to structural variations or interference from paralogous sequences affecting accurate sequence alignment across backcross genomes can be difficult with short-read data. Synteny analyses between pure species could offer valuable insights into chromosomal rearrangements and regions with complex architecture, including the terminal region of scaffold 10 (Li et al. 2022). Additionally, this study uses short-read data to focus on biallelic SNPs as proxies for genetic diversity, similar to the approaches undertaken by Stroupe et al. (2022) and Askelson et al. (2023) investigating hybridisation in North American bison (*Bison bison*) and white-breasted nuthatches (*Sitta carolinensis*), respectively. Exploring structural variants and indels represents another promising approach to gain deeper insights into genomic diversity within and between species. These elements can span multiple alleles, genes, and gene regions simultaneously, potentially exerting a more substantial influence on fitness (Wold et al. 2021). For future functional studies, annotating MHC genes in *E. macquarii* and employing targeted amplicon sequencing would facilitate thorough and precise comparisons of the genomic impacts of hybridisation within MHC loci, including duplications, copy number variations, and expansions in gene families (Horton et al. 2004). Additionally, greater sequencing effort of first-generation hybrids and additional backcross individuals will offer a comprehensive view of genomic dynamics, facilitating a deeper understanding of the extent of introgression and its potential implications for genome-wide and immune function, and overall fitness across a larger sample set.

Conclusion

Our study demonstrates a high level of conservation in the functionally important MHC core region, in contrast to low, putatively neutral, genome-wide diversity of this critically endangered turtle. Using short-read Illumina re-sequenced genomes, we reconstructed the demographic history of the population, revealing a pronounced and ongoing gradual decline in effective population size and a persistent trajectory towards genetic depletion over time. Genome-wide diversity shows evidence of having diminished over time and is expected to decline further, particularly following the population crash triggered by the 2015 nidovirus outbreak. Lastly, by characterising the MHC region in a freshwater turtle species, this study provides important insights to support ongoing research within this species and lays the groundwork for future investigations into MHC diversity across testudines more broadly.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00251-025-01378-8.

Acknowledgements We acknowledge Gumbaynggirr people and their ongoing connection to Country where the turtles live and pay respects to their Elders past and present. This research was made possible by the NSW DCCEEW Saving our Species program and the NCRIS funded Bioplatforms Australia Threatened Species Initiative. Computing resources were provided by the University of Sydney's High Performance Computing cluster Artemis and Amazon Web Services.

Author contribution H.V.N., K.B., and C.J.H. conceptualised this study. H.V.N. conducted all lab work and analyses and wrote the first draft. C.J.H. obtained funding. A.G. provided samples for WGR and LS and T.G.L.K. assisted with bioinformatic analyses. E.A.M. and J.G. helped revise the manuscript with all authors.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. Funding for this work was provided by the Australian Government's Bushfire Recovery Scheme (ERF-WRR2-020) to the Threatened Species Initiative, and HVN scholarship funding was provided by the NSW DCCEEW and the University of Sydney.

Data availability The male *M. georgesi* reference genome assembly, all raw sequencing reads including the 3-tissue transcriptome RNA-seq reads, whole-genome re-sequencing data and annotated sequences are available from NCBI under BioProject PRJNA1003540. All scripts used to process data are publicly available at https://github.com/awgg-lab/australasiangenomes

Declarations

Ethics approval Collection of samples was conducted in accordance with the conditions of NSW DCCEEW Animal Ethics Committee (AEC151201-3, AEC160503-01 and AEC180904-5) and Scientific Licences (MWL00102467, SL101672, and SL10255).

Conflict of interest The authors declare no competing interests.

References

Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Res 25:4692–4693. https://doi.org/10.1093/nar/25.22.4692

Altizer S, Harvell D, Friedle E (2003) Rapid evolutionary dynamics and disease threats to biodiversity. Trends Ecol Evol 18:589–596

Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends Ecol Evol 19:535–544

Andrews S (2010) FastQC: A quality control tool for high throughput sequence data. Available online: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 20 Sept 2024

Askelson KK, Spellman GM, Irwin D (2023) Genomic divergence and introgression between cryptic species of a widespread North American songbird. Mol Ecol 32:6839–6853. https://doi.org/10.1111/mec.17169

Auteri GG, Knowles LL (2020) Decimated little brown bats show potential for adaptive change. Sci Rep 10:3023

Baack EJ, Rieseberg LH (2007) A genomic view of introgression and hybrid speciation. Curr Opin Genet Dev 17:513–518

Belov K, Miller RD, Old JM, Young LJ (2013) Marsupial immunology bounding ahead. Aust J Zool 61:24–40

Bergeron LA, Besenbacher S, Zheng J, Li P, Bertelsen MF, Quintard B, Hoffman JI, Li Z, St. Leger J, Shao C (2023) Evolution of the germline mutation rate across vertebrates. Nature 615:285–291

Bi C, Lu N, Huang Z, Chen J, He C, Lu Z (2020) Whole-genome resequencing reveals the pleistocene temporal dynamics of Branchiostoma belcheri and Branchiostoma floridae populations. Ecol Evol 10:8210–8224. https://doi.org/10.1002/ece3.6527

Bolger A, Giorgi F (2014) Trimmomatic: a flexible read trimming tool for illumina NGS data. Bioinformatics 30:2114–2120

Bosse M, Megens H-J, Madsen O, Paudel Y, Frantz LAF, Schook LB, Crooijmans RPMA, Groenen MAM (2012) Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. PLoS Genet 8:e1003100. https://doi.org/10.1371/journal.pgen.1003100

Buchanan SW, Kolbe JJ, Wegener JE, Atutubo JR, Karraker NE (2019) A comparison of the population genetic structure and diversity between a common (Chrysemys p. picta) and an endangered (Clemmys guttata) freshwater turtle. Diversity 11:99. https://doi.org/10.3390/d11070099

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinform 10:1–9

Card DC, Van Camp AG, Santonastaso T, Jensen-Seaman MI, Anthony NM, Edwards SV (2022) Structure and evolution of the squamate major histocompatibility complex as revealed by two Anolis lizard genomes. Front Genet 13:979746

Ceballos FC, Joshi PK, Clark DW, Ramsay M, Wilson JF (2018) Runs of homozygosity: windows into population history and trait architecture. Nat Rev Genet 19:220–234

Chessman BC, McGilvray G, Ruming S, Jones HA, Petrov K, Fielder DP, Spencer R, Georges A (2020) On a razor's edge: status and prospects of the critically endangered Bellinger River snapping turtle, Myuchelys georgesi. Aquat Conserv Mar Freshw Ecosyst 30:586–600. https://doi.org/10.1002/aqc.3258

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, Genomes Project Analysis G (2011) The variant call format and VCFtools. Bioinformatics 27:2156–8. https://doi.org/10.1093/bioinformatics/btr330

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. GigaScience 10:giab008. https://doi.org/10.1093/gigascience/giab008

Daszak P, Cunningham AA, Hyatt AD (2000) Emerging infectious diseases of wildlife-threats to biodiversity and human health. Science 287:443–449

Dugovich BS, Beechler BR, Dolan BP, Crowhurst RS, Gonzales BJ, Powers JG, Hughson DL, Vu RK, Epps CW, Jolles AE (2023) Population connectivity patterns of genetic diversity, immune

responses and exposure to infectious pneumonia in a metapopulation of desert bighorn sheep. J Anim Ecol 92:1456–1469
Farquharson KA, McLennan EA, Cheng Y, Alexander L, Fox S, Lee AV, Belov K, Hogg CJ (2022) Restoring faith in conservation action: maintaining wild genetic diversity through the Tasmanian devil insurance program. Iscience 25:104474
Fay R, Ficheux S, Béchet A, Besnard A, Crochet P-A, Leblois R, Crivelli A, Wattier R, Olivier A (2023) Direct and indirect estimates of dispersal support strong juvenile philopatry and male-biased dispersal in a freshwater turtle species (Emys orbicularis). Freshw Biol 68:2042–2053. https://doi.org/10.1111/fwb.14171
Fielder D, Vernes K, Alacs E, Georges A (2012) Mitochondrial variation among Australian freshwater turtles (genus Myuchelys), with special reference to the endangered M. bellii. Endanger Species Res 17:63–71
Flot J (2010) SeqPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. Mol Ecol Resour 10:162–166
Foote AD, Hooper R, Alexander A, Baird RW, Baker CS, Ballance L, Barlow J, Brownlow A, Collins T, Constantine R (2021) Runs of homozygosity in killer whale genomes provide a global record of demographic histories. Mol Ecol 30:6162–6177
Frank SA (2002) Immunology and evolution of infectious disease. Princeton University Press
Frankham R (2015) Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. Mol Ecol 24:2610–2618
Frankham R, Ballou JD, Ralls K, Eldridge M, Dudash MR, Fenster CB, Lacy RC, Sunnucks P (2017) Genetic management of fragmented animal and plant populations. Oxford University Press
Fu M, Eimes JA, Kong S, Lamichhaney S, Waldman B (2023) Identification of major histocompatibility complex genotypes associated with resistance to an amphibian emerging infectious disease. Infect Genet Evol 113:105470
Gemmell NJ, Rutherford K, Prost S, Tollis M, Winter D, Macey JR, Adelson DL, Suh A, Bertozzi T, Grau JH, Organ C, Gardner PP, Muffato M, Patricio M, Billis K, Martin FJ, Flicek P, Petersen B, Kang L, Michalak P, Buckley TR, Wilson M, Cheng Y, Miller H, Schott RK, Jordan MD, Newcomb RD, Arroyo JI, Valenzuela N, Hore TA, Renart J, Peona V, Peart CR, Warmuth VM, Zeng L, Kortschak RD, Raison JM, Zapata VV, Wu Z, Santesmasses D, Mariotti M, Guigó R, Rupp SM, Twort VG, Dussex N, Taylor H, Abe H, Bond DM, Paterson JM, Mulcahy DG, Gonzalez VL, Barbieri CG, DeMeo DP, Pabinger S, Van Stijn T, Clarke S, Ryder O, Edwards SV, Salzberg SL, Anderson L, Nelson N, Stone C, Stone C, Smillie J, Edmonds H (2020) The tuatara genome reveals ancient features of amniote evolution. Nature 584:403–409. https://doi.org/10.1038/s41586-020-2561-9
Georges A (2020) Assessment of hybridization and introgression in Bellinger River turtles, and assessment of parentage of hatchlings destined for release in the recovery program of Myuchelys georgesi. Biodiversity and conservation, NSW Department of Planning. Industry and Environment. http://georges.biomatix.org/storage/app/uploads/public/5f4/f7a/bf4/5f4f7abf48b9f904504354.pdf
Georges A, Spencer R-J, Kilian A, Welsh M, Zhang X (2018) Assault from all sides: hybridization and introgression threaten the already critically endangered Myuchelys georgesi (Chelonia: Chelidae). Endanger Species Res 37:239–247. https://doi.org/10.3354/esr00928
Grieves L, Gloor G, Bernards M, MacDougall-Shackleton E (2019) Songbirds show odour-based discrimination of similarity and diversity at the major histocompatibility complex. Anim Behav 158:131–138
Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321
He K, Babik W, Majda M, Minias P (2023) MHC architecture in amphibians—ancestral reconstruction, gene rearrangements, and duplication patterns. Genome Biol Evol 15:evad079. https://doi.org/10.1093/gbe/evad079
He K, Liang C, Zhu Y, Dunn P, Zhao A, Minias P (2022a) Reconstructing macroevolutionary patterns in avian MHC architecture with genomic data. Front Genet 13:823686
He K, Zhu Y, Yang S-C, Ye Q, Fang S-G, Wan Q-H (2022b) Major histocompatibility complex genomic investigation of endangered Chinese alligator provides insights into the evolution of tetrapod major histocompatibility complex and survival of critically bottlenecked species. Front Ecol Evol 10:1078058
Hewett AM, Stoffel MA, Peters L, Johnston SE, Pemberton JM (2023) Selection, recombination and population history effects on runs of homozygosity (ROH) in wild red deer (Cervus elaphus). Heredity 130:242–250
Hilgers L, Liu S, Jensen A, Brown T, Cousins T, Schweiger R, Guschanski K, Hiller M (2024) Avoidable false PSMC population size peaks occur across numerous studies. bioRxiv 2024.06.17.599025. https://doi.org/10.1101/2024.06.17.599025
Hoang DT, Vinh LS, Flouri T, Stamatakis A, von Haeseler A, Minh BQ (2018) MPBoot: fast phylogenetic maximum parsimony tree inference and bootstrap approximation. BMC Evol Biol 18:1–11
Hoffmann AA, Sgrò CM, Kristensen TN (2017) Revisiting adaptive potential, population size, and conservation. Trends Ecol Evol 32:506–517
Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landsc Ecol 21:797–807. https://doi.org/10.1007/s10980-005-5245-9
Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC Jr, Wright MW, Wain HM (2004) Gene map of the extended human MHC. Nat Rev Genet 5:889–899
Hughes AL, Yeager M (1998) Natural selection at major histocompatibility complex loci of vertebrates. Annu Rev Genet 32:415–435
Jaegle B, Pisupati R, Soto-Jiménez LM, Burns R, Rabanal FA, Nordborg M (2023) Extensive sequence duplication in Arabidopsis revealed by pseudo-heterozygosity. Genome Biol 24:44. https://doi.org/10.1186/s13059-023-02875-3
Jeon JY, Shin Y, Mularo AJ, Feng X, DeWoody JA (2024) The integration of whole-genome resequencing and ecological niche modelling to conserve profiles of local adaptation. Divers Distrib 30:e13847. https://doi.org/10.1111/ddi.13847
Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405
Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589
Kardos M, Armstrong EE, Fitzpatrick SW, Hauser S, Hedrick PW, Miller JM, Tallmon DA, Funk WC (2021) The crucial role of genome-wide genetic variation in conservation. Proc Natl Acad Sci U A 118. https://doi.org/10.1073/pnas.2104642118
Le M, Reid BN, McCord WP, Naro-Maciel E, Raxworthy CJ, Georges A (2013) Resolving the phylogenetic history of the short-necked turtles, genera *Elseya* and *Myuchelys* (Testudines: Chelidae) from Australia and New Guinea. Mol Phylogenet Evol 68:251–258
Le Bouteiller P, Lenfant F (1996) Antigen-presenting function(s) of the non-classical HLA-E, -F and -G class I molecules: the beginning of a story. Res Immunol 147:301–313. https://doi.org/10.1016/0923-2494(96)89643-X
Leclaire S, van Dongen WF, Voccia S, Merkling T, Ducamp C, Hatch SA, Blanchard P, Danchin É, Wagner RH (2014) Preen secretions

encode information on MHC similarity in certain sex-dyads in a monogamous seabird. Sci Rep 4:6920
LeMaoult J, Le Discorde M, Rouas-Freiss N, Moreau P, Menier C, McCluskey J, Carosella E (2003) Biology and functions of human leukocyte antigen-G in health and sickness. Tissue Antigens 62:273–284
Li H, Durbin R (2011) Inference of human population history from individual whole-genome sequences. Nature 475:493–496
Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079
Li X, He S-G, Li W-R, Luo L-Y, Yan Z, Mo D-X, Wan X, Lv F-H, Yang J, Xu Y-X, Deng J, Zhu Q-H, Xie X-L, Xu S-S, Liu C-X, Peng X-R, Han B, Li Z-H, Chen L, Han J-L, Ding X-Z, Dingkao R, Chu Y-F, Wu J-Y, Wang L-M, Zhou P, Liu M-J, Li M-H (2022) Genomic analyses of wild argali, domestic sheep, and their hybrids provide insights into chromosome evolution, phenotypic variation, and germplasm innovation. Genome Res 32:1669–1684. https://doi.org/10.1101/gr.276769.122
Madeira F, Park Y mi, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47:W636–W641. https://doi.org/10.1093/nar/gkz268
Magid M, Wold JR, Moraga R, Cubrinovska I, Houston DM, Gartrell BD, Steeves TE (2022) Leveraging an existing whole-genome resequencing population data set to characterize toll-like receptor gene diversity in a threatened bird. Mol Ecol Resour 22:2810–2825
Martinez PA, Ezaz T, Valenzuela N, Georges A, Marshall Graves JA (2008) An XX/XY heteromorphic sex chromosome system in the Australian chelid turtle Emydura macquarii: a new piece in the puzzle of sex chromosome evolution in turtles. Chromosome Res 16:815–825
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303. https://doi.org/10.1101/gr.107524.110
McLennan E, Cheng Y, Farquharson K, Grueber C, Elmer J, Alexander L, Fox S, Belov K, Hogg C (2024) Reinforcements in the face of ongoing threats: a case study from a critically small carnivore population. Anim Conserv. https://doi.org/10.1111/acv.12945C
McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A (2008) Runs of homozygosity in European populations. Am J Hum Genet 83:359–372
Miller HC, Belov K, Daugherty CH (2005) Characterization of MHC class II genes from an ancient reptile lineage, Sphenodon (tuatara). Immunogenetics 57:883–891
Miller HC, O'Meally D, Ezaz T, Amemiya C, Marshall-Graves JA, Edwards S (2015) Major histocompatibility complex genes map to two chromosomes in an evolutionarily ancient reptile, the Tuatara Sphenodon punctatus. G3 Genes Genomes Genet 5:1439–1451. https://doi.org/10.1534/g3.115.017467
Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534
Mooney JA, Yohannes A, Lohmueller KE (2021) The impact of identity by descent on fitness and disease in dogs. Proc Natl Acad Sci 118:e2019116118
Napolitano C, Sacristán I, Acuña F, Aguilar E, García S, López-Jara MJ, Cabello J, Hidalgo-Hermoso E, Poulin E, Grueber CE (2023) Assessing micro-macroparasite selective pressures and anthropogenic disturbance as drivers of immune gene diversity in a Neotropical wild cat. Sci Total Environ 897:166289
Neefjes J, Jongsma ML, Paul P, Bakke O (2011) Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat Rev Immunol 11:823–836
Nelson HV, Farquharson KA, Georges A, McLennan EA, DeGabriel JL, Giese M, Ormond C, McFadden M, Skidmore A, Prangell J (2024) A genomic framework to assist conservation breeding and translocation success: a case study of a critically endangered turtle. Conserv Sci Pract 6:e13204
Parrish K, Kirkland P, Horwood P, Chessman B, Ruming S, McGilvray G, Rose K, Hall J, Skerratt L (2024) Delving into the aftermath of a disease-associated near-extinction event: a five-year study of a serpentovirus (nidovirus) in a critically endangered turtle population. Viruses 16:653
Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295
Peel E, Ng HJJ, Belov K (2019) Monotreme, marsupial and bat immunology, Current therapy in medicine of Australian mammals. Collingwood, Victoria, Australia
Peel E, Silver L, Brandies P, Zhu Y, Cheng Y, Hogg CJ, 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
Pemberton TJ, Absher D, Feldman MW, Myers RM, Rosenberg NA, Li JZ (2012) Genomic patterns of homozygosity in worldwide human populations. Am J Hum Genet 91:275–292
Penn DJ, Damjanovich K, Potts WK (2002) MHC heterozygosity confers a selective advantage against multiple-strain infections. Proc Natl Acad Sci 99:11260–11264
Piertney S, Oliver M (2006) The evolutionary ecology of the major histocompatibility complex. Heredity 96:7–21
Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842
R Core Team (2023) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/
Reed KM, Settlage RE (2021) Major histocompatibility complex genes and locus organization in the Komodo dragon (Varanus komodoensis). Immunogenetics 73:405–417
Renaud G, Hanghøj K, Korneliussen TS, Willerslev E, Orlando L (2019) Joint estimates of heterozygosity and runs of homozygosity for modern and ancient samples. Genetics 212:587–614. https://doi.org/10.1534/genetics.119.302057
Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. Nat Biotechnol 29:24–26
Santiago E, Novo I, Pardiñas AF, Saura M, Wang J, Caballero A (2020) Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. Mol Biol Evol 37:3642–3653. https://doi.org/10.1093/molbev/msaa169
Savage AE, Muletz-Wolz CR, Campbell Grant EH, Fleischer RC, Mulder KP (2019) Functional variation at an expressed MHC class IIβ locus associates with Ranavirus infection intensity in larval anuran populations. Immunogenetics 71:335–346
Shapiro B (2017) Pathways to de-extinction: how close can we get to resurrection of an extinct species? Funct Ecol 31:996–1002. https://doi.org/10.1111/1365-2435.12705
Silver LW, Hogg CJ, Belov K (2024) Plethora of new marsupial genomes informs our knowledge of marsupial MHC class II. Genome Biol Evol 16. https://doi.org/10.1093/gbe/evae156
Smith KF, Acevedo-Whitehouse K, Pedersen AB (2009) The role of infectious diseases in biological conservation. Anim Conserv 12:1–12

Sommer S (2005) The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Front Zool 2:1–18
Spencer R-J, Georges A, Lim D, Welsh M, Reid AM (2014) The risk of inter-specific competition in Australian short-necked turtles. Ecol Res 29:767–777. https://doi.org/10.1007/s11284-014-1169-7
Spencer R-J, Van Dyke J, Petrov K, Ferronato B, McDougall F, Austin M, Keitel C, Georges A (2018) Profiling a possible rapid extinction event in a long-lived species. Biol Conserv 221:190–197. https://doi.org/10.1016/j.biocon.2018.03.009
Spielman D, Brook BW, Briscoe DA, Frankham R (2004) Does inbreeding and loss of genetic diversity decrease disease resistance? Conserv Genet 5:439–448
Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet 76:449–462. https://doi.org/10.1086/428594
Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989. https://doi.org/10.1086/319501
Stroupe S, Forgacs D, Harris A, Derr JN, Davis BW (2022) Genomic evaluation of hybridization in historic and modern North American Bison (Bison bison). Sci Rep 12:6397. https://doi.org/10.1038/s41598-022-09828-z
Summers K, McKeon S, Sellars J, Keusenkothen M, Morris J, Gloeckner D, Pressley C, Price B, Snow H (2003) Parasitic exploitation as an engine of diversity. Biol Rev 78:639–675
Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol 38:3022–3027
Teixeira JC, Huber CD (2021) The inflated significance of neutral genetic diversity in conservation genetics. Proc Natl Acad Sci 118:e2015096118
Theissinger K, Fernandes C, Formenti G, Bista I, Berg PR, Bleidorn C, Bombarely A, Crottini A, Gallo GR, Godoy JA, Jentoft S (2023) How genomics can help biodiversity conservation. Trends in Genetics 39:545–559
Trujillo AL, Hoffman EA, Becker CG, Savage AE (2021) Spatiotemporal adaptive evolution of an MHC immune gene in a frog-fungus disease system. Heredity 126:640–55
Wold J, Koepfli K-P, Galla SJ, Eccles D, Hogg CJ, Le Lec MF, Guhlin J, Santure AW, Steeves TE (2021) Expanding the conservation genomics toolbox: incorporating structural variants to enhance genomic studies for species of conservation concern. Mol Ecol 30:5949–5965. https://doi.org/10.1111/mec.16141
Yang Z, Zhang Z, Wang G, Huang Z, You W (2024) Exploiting hybrid vigor to enhance survival and growth for whiteleg shrimp (Litopenaeus vannamei): a complete diallel cross experiment. Aquac Fish. https://doi.org/10.1016/j.aaf.2024.04.004
Zardoya R, Meyer A (2001) The evolutionary position of turtles revised. Naturwissenschaften 88:193–200
Zhang J, Finlaison DS, Frost MJ, Gestier S, Gu X, Hall J, Jenkins C, Parrish K, Read AJ, Srivastava M (2018) Identification of a novel nidovirus as a potential cause of large scale mortalities in the endangered Bellinger River snapping turtle (Myuchelys georgesi). PLoS One 13:e0205209
Zhu Y, Wan Q-H, Zhang H-M, Fang S-G (2019) Reproductive strategy inferred from major histocompatibility complex-based inter-individual, sperm-egg, and mother-fetus recognitions in giant pandas (Ailuropoda melanoleuca). Cells 8:257

A4.4 Opportunities for research and conservation of freshwater turtles in Australia

The PDF version of publications contributed to during candidature are presented on the following pages.

Received: 29 June 2023 | Accepted: 18 July 2023

DOI: 10.1111/aec.13403

EDITORIAL

Opportunities for research and conservation of freshwater turtles in Australia

Abstract

Australia's freshwater turtles have high endemicity and many are threatened by extinction. Following a symposium held at the 2022 conference of the Australian Society of Herpetologists, we summarized the current status of research and conservation for Australian freshwater turtles and identified opportunities for future research. Eight species (32%) of Australia's 25 native freshwater turtles are listed as threatened by Australia's Federal Government. Symposium discussions on the primary gaps in research identified the lack of baseline data to inform population modelling as a key deficiency. Knowledge of the most effective conservation actions, the effectiveness of attempts to aid population recovery, and whether these actions are required at all, remains lacking for many species. A heavy bias exists between some well-studied species compared with others for which little or no information is published. Community science, engagement with First Nations people, advances in technology, and recognition of the importance of turtles are contributing to better knowledge.

INTRODUCTION

Globally, over half of our ~360 turtle species are threatened or endangered (Lovich et al., 2018; Stanford et al., 2020) resulting in a significant loss of ecosystem services. Turtles can reach high biomasses and influence the environment (Glorioso et al., 2010; Iverson, 1982) through their roles in food webs as both consumers (Blamires & Spencer, 2013; Petrov et al., 2018) and prey (Lettoof et al., 2021). Turtles influence water quality (Santori et al., 2020), soil bioturbation (Kaczor & Hartnett, 1990), seed dispersal (Falcón et al. 2020), and nutrient cycling (Wenger et al., 2019). Thus, ecosystem health is affected by declines in turtle populations. Socially, turtles were part of cosmic symbolism of ancient cultures (Rappengliick & Gilching, 2006) and remain important in many communities and traditional cultures (Cann & Sadlier, 2017; Fordham et al., 2006). They also play a role in storytelling as popular characters in books and movies, often as the underdog or hero (Hicks & Kanevsky, 1992). Thus, conservation of turtles is valuable environmentally and socially.

Freshwater turtle conservation in Australia is crucial, because the majority of species are endemic (88%), which places an imperative on Australian society to preserve their existence. Australian freshwater turtles, with the exception of *Carettochelys*, represent the distinct sub-family Chelodininae within the side-necked turtles (Pleurodira). Chelodininae is restricted to Australasia, where they are adapted to the dynamic freshwater environment. Australia is also a largely arid island continent with highly irregular rainfall in many places, which makes its freshwater systems more susceptible to climate change. Australian freshwater turtles are currently threatened by a suite of factors affecting the environment. Proximate causes for their declines include viruses (Chessman et al., 2020), habitat loss and degradation (Chessman, 2011), and invasive species (Fordham et al., 2006; Spencer & Thompson, 2005).

To address declines in Australian freshwater turtles and ensure their persistence for future generations, research (Chessman, Dillon, et al., 2023; Chessman, Fielder, et al., 2023; Coleman, 2023; Kidman et al., 2023; McKnight, 2023; McKnight et al., This issue; Nordberg & McKnight, 2023; Van Dyke et al., This issue) and conservation actions (Campbell et al., 2023; Streeting et al., 2023; Terry et al., 2023) are ongoing. However, compared to many other taxa, freshwater turtles have received less focus and funding in Australia (Gawne et al., 2020). Recognition of turtle declines, their potential to garner community support, and the role they play in facilitating ecosystem functioning, has led to increased interest from community groups and government in recent years (Santori et al., 2021).

Our special issue follows a symposium held to discuss the key priorities for the conservation and research of Australia's freshwater turtles at the conference of the Australian Society of Herpetologists at Mylor in South Australia 11th–14th July 2022. The symposium resulted in a series of ecological papers (this special issue) to increase our knowledge of Australian freshwater turtles and aid in their conservation management. Here, we summarize the current status of research and conservation for Australian freshwater turtles, identify opportunities for future research, and discuss important

factors for prioritizing future research, drawing on both new research published in this special issue and the literature more broadly.

STATUS AND DISTRIBUTION OF AUSTRALIAN TURTLES

Currently, 25 native species of freshwater turtles inhabit Australia, comprising 24 chelid species and *Carettochelys insculpta*. The geographic distributions of Australian freshwater turtle species vary considerably, from exceptionally narrow ranges that may be as small as a single catchment (e.g., *Myuchelys georgesi*), to widespread species occupying multiple catchments and states (e.g., *Emydura macquarii*, *Chelodina longicollis*, and *C. expansa*) (Figure 1). *Carettochelys insculpta*, *Emydura subglobosa*, and *Chelodina rugosa* also occur in New Guinea, while the remaining species are endemic to Australia. All of Australia's endemic turtles are aquatic, and thus they are largely absent from the arid centre (Figure 1). Centres of diversity occur in the Northern Territory and mid-eastern Australia where up to nine species are broadly sympatric.

Eight species (32%) of Australian freshwater turtles are listed as threatened by the Federal Government (EPBC Act, 1999; EPBC Act List of Threatened Fauna (environment.gov.au) accessed 10 May 2022). Three species are listed nationally as Critically Endangered. Two of these three species (*Elseya albagula* and *Pseudemydura umbrina*) are from monotypic genera, while the third (*Myuchelys georgesi*) is from a genus with three other species (two of which are listed as Endangered). Four species are listed as Endangered (*Elseya lavarackorum*, *Elusor macrurus*, *Myuchelys bellii*, and *M. purvisi*) and a further species is listed as Vulnerable (*Rheodytes leukops*). These species are all range-restricted (Figure 2; Table 1).

RESEARCH KNOWLEDGE AND GAPS

Australian freshwater turtles are unique in their habits and exhibit many remarkable adaptations to the Australian environment. For example, bizarre habits in reproduction include the underwater nesting of *Chelodina rugosa* (Kennett et al., 1993), the extensive embryonic diapause of *Chelodina expansa* (Booth, 2000), and the explosive hatching of *Carettochelys insculpta* (Webb et al., 1986). In oxygen-rich environments we see cloacal breathing as the primary form of respiration (FitzGibbon & Franklin, 2010), allowing extensive dive times in species adapted to

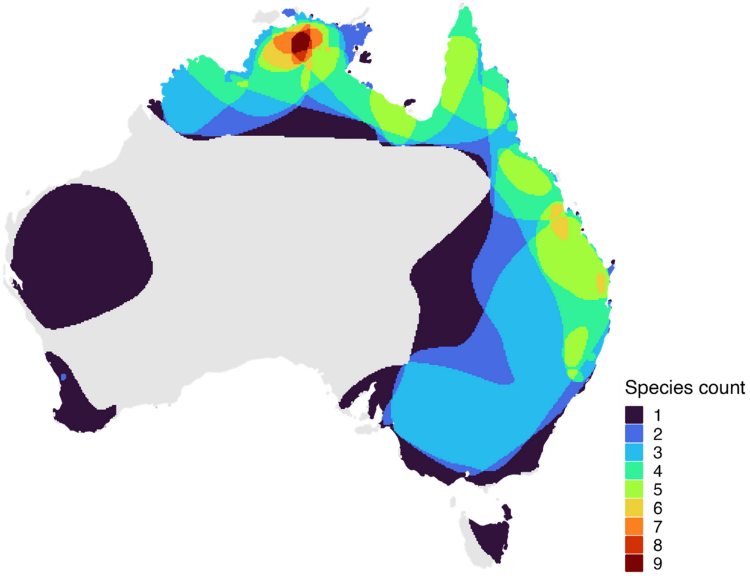


FIGURE 1 Map of turtle species distributions in Australia showing hotspots of richness in northern and eastern Australia and an absence of turtles in the arid centre. Notes: the one species found in Tasmania (*Chelodina longicollis*) has been introduced there; *Emydura macquarii* has also spread beyond its native range.

lotic environments with clean water (Clark et al., 2008; Gordos et al., 2004). In other species (e.g., *Chelodina longicollis*), a long coevolution with salt has driven the evolution of salt tolerance which approaches that of estuarine turtle species in other countries (Bower et al., 2016). Australian freshwater turtles were also among the first to be documented vocalizing, with an extensive repertoire of underwater vocalizations in *Chelodina oblonga* (Giles, 2005). It was also only recently discovered that some turtles in northern Australia bask nocturnally (Kidman et al., 2023; McKnight et al., 2023; Nordberg & McKnight, 2020, 2023). Some species have complex patterns of habitat use that include permanent, semi-permanent, and ephemeral wetlands (Santoro et al., 2020) and extensive terrestrial activity (Kennett et al., 2009). Additionally, life in large riverine environments enables long-distance movements by male turtles, the extent of which we are only just beginning to understand (Bower et al., 2012; Van Dyke et al., This issue). Future research will almost certainly continue to document many more weird and wonderful habits of Australian freshwater turtles.

While ecology is the primary research discipline in publications on Australian freshwater turtles, we still lack basic knowledge of home range sizes, movement cues, social structures, cognitive abilities, and environmental factors influencing growth and reproductive output. These knowledge gaps are true even for species for which there are substantial publication outputs (e.g., *Emydura macquarii*). With the advent of new technologies, such as camera trapping (McKnight et al., 2023), tissue stable isotope

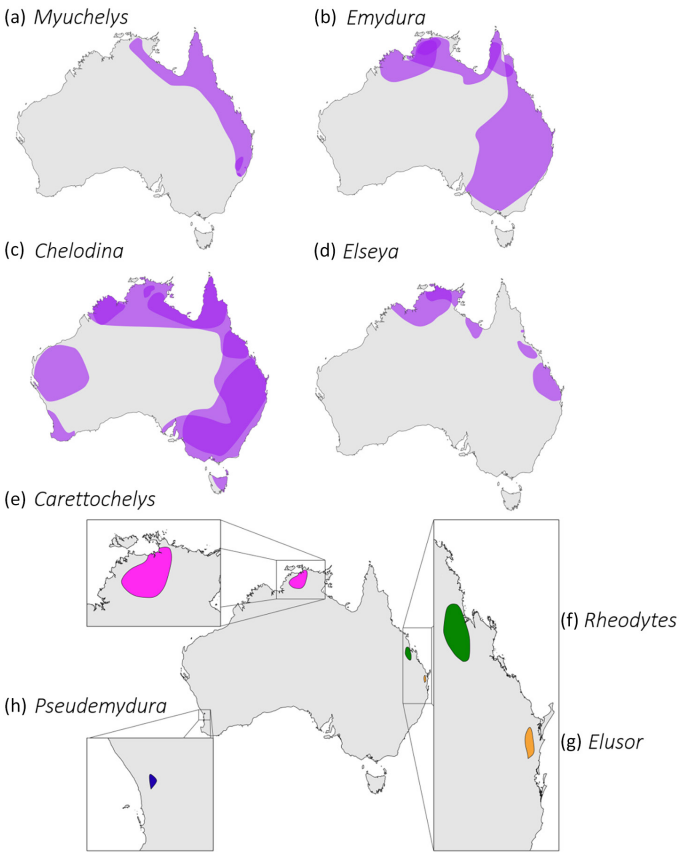


FIGURE 2 Species ranges overlaid to show variation in the Australian freshwater turtle genera distribution for (a) *Myuchelys* (b) *Emydura* (c) *Chelodina* (d) *Elseya* and monotypic (e) *Carettochelys insculpta* (f) *Rheodytes leukops* (g) *Elusor macrurus* (h) *Pseudemydura umbrina*.

TABLE 1 Current and interim proposed listing of Australian freshwater turtles under the *Environment Protection and Biodiversity Act 1999* (subject to change as the consultation process proceeds).

Species	Current Commonwealth listing	2022 SEAP proposed Commonwealth listing
Carettochelys insculpta	Not listed	Endangered
Chelodina burrungandjii	Not listed	
Chelodina canni	Not listed	Insufficient data
Chelodina expansa	Not listed	
Chelodina kuchlingi	Not listed	
Chelodina longicollis	Not listed	Vulnerable
Chelodina oblonga	Not listed	Insufficient data
Chelodina rugosa	Not listed	
Chelodina steindachneri	Not listed	
Elseya albagula	Critically endangered	Critically endangered
Elseya dentata	Not listed	
Elseya flaviventralis	Not listed	
Elseya irwini	Not listed	Endangered
Elseya lavarackorum	Endangered	Endangered
Elusor macrurus	Endangered	Critically endangered
Emydura macquarii gunabarra	Not listed	Endangered
Emydura macquarii	Not listed	Vulnerable
Emydura subglobosa subglobosa	Not listed	Critically endangered
Emydura tanybaraga	Not listed	
Emydura victoriae	Not listed	
Myuchelys bellii	Endangered	Endangered
Myuchelys georgesi	Critically endangered	Critically endangered
Myuchelys latisternum	Not listed	
Myuchelys purvisi	Endangered	Endangered
Pseudemydura umbrina	Critically endangered	
Rheodytes leukops	Vulnerable	Endangered

analysis (Micheli-Campbell et al., 2017), passive acoustic tracking (Micheli-Campbell et al., 2017; Van Dyke et al., This issue), environmental DNA (Villacorta-Rath et al., 2022), and underwater baited video cameras (Coleman, This issue), we are beginning to unravel the mysteries of turtle behaviour. In addition, the reduction in cost and increase in the capability of genetic techniques have enabled better resolution of phylogenetic relationships among and within turtle species, providing clarity around systematics and nomenclature (Georges et al., 1999, 2018; Thomson et al., 2021).

Symposium discussions on the primary gaps in research identified the lack of baseline data to inform population modelling as a key deficiency (Petrov et al., This issue). This lack of data includes relationships for fertility, hatching rates, sex ratios, age of maturity, maximum age, and age-dependent survival rates. Repeated population estimates are not available for most species, which limits our ability to track population changes through time (Howell et al., 2019). This sort of information is doubly important for long-lived species such as turtles because it is harder to detect population declines initially, and because they can take longer to respond to conservation efforts since they have delayed sexual maturity and high rates of juvenile mortality. For many northern species, there is little published information on basic ecology, such as nesting location and timing, reproductive output, or habitat use. Excluding subspecies, six of the seven species excluded from the Special Expert Assessment Panel owing to data deficiency were species from northern Australia (Petrov et al., This issue). Many species live in remote parts of northern Australia (e.g., *Elseya flaviventralis*, *Emydura tanybaraga*), are recently described (*Myuchelys purvisi*), or are rare or sympatric with crocodiles, and therefore difficult to study (e.g., *Emydura subglobosa subglobosa*). This lack of fundamental information may be an opportunity for reptile breeders to contribute collectively to aid in establishing baseline information for little-known species and their demographics.

PRIORITIES AND CHALLENGES FOR THE CONSERVATION OF AUSTRALIAN TURTLES

Conservation biologists continue to refine the best approaches to counteract population declines and causes of mortality in Australian freshwater turtles. Actions include attempts to mitigate threats of feral species such as pigs and foxes through predator-proof fencing (Cochrane, 2006; Streeting et al., 2023), nest protection measures (Campbell et al., 2020; Connell, 2018;Streeting et al., 2023; Terry et al., 2023), and artificial incubation of eggs and release of hatchlings (Streeting et al., 2022). Likewise, conservation translocations of Critically Endangered species such as reintroductions (*Myuchelys georgesi*) and assisted colonization (*Pseudemydura umbrina*) have attempted to mitigate critically low population sizes due to disease (Chessman et al., 2020) and habitat loss and degradation (Bouma et al., 2020; Kuchling et al., 2018). Improved river management, such as managed flow delivery in regulated systems, has attempted to improve wetland habitats and the condition of rivers (Espinoza et al., 2022; Francis et al., 2022). More broadly, technological advances such as biobanking gametes, controlling diseases, and advances in genetic and bioinformatic tools should assist future conservation measures (Clulow et al., 2022). However, knowledge of the most effective conservation actions, the effectiveness of attempts to aid population recovery, and whether these actions are required at all, remains lacking for many species.

Community Science is leading the way in the broad-scale reporting and mitigation of major threats to freshwater turtles, such as road mortality and nest predation (e.g., TurtleSAT.org.au – (Santori et al., 2021). The 1 Million Turtles (1MT) Community Conservation Program (1MillionTurtles.com) is engaging the public in turtle conservation and connects them with nature through a hands-on approach to conservation. Individuals and community groups can participate in activities such as nest monitoring, turtle rescues, and habitat restoration/modification (e.g., turtle islands). Similarly, community science programs to counteract road mortality and predation of nesting females and their nests have been established for *Chelodina oblonga* in Western Australia (Santoro, 2022).

Discussions on the priorities required in freshwater turtle conservation noted that science is hindered by a lack of baseline information. While research has been completed to counteract feral species as a source of mortality (Streeting et al., 2023; Terry et al., 2023), we still do not know how this mortality compares to the predation pressures that were evident prior to the decline of many medium-sized mammals (e.g., bandicoots, quolls) which were likely nest predators (Chessman, 2022). Although Australian turtles can withstand relatively high levels of nest predation, when coupled with unnaturally high levels of adult mortality, their populations will inevitably decrease (Spencer, 2018). Additionally, the decreased productivity created by regulating rivers and the associated habitat loss has likely reduced the density of freshwater turtles in areas they still occupy. This suggests that turtles were much more abundant previously (Thompson, 1993), but historical densities are difficult to accurately quantify. The lack of baseline information has placed an increasing importance on the need to determine current trends in turtle population growth over time, so that action can be taken where population trajectories suggest ongoing declines. In addition, the lack of any ecological information for some species of Australian freshwater turtle (e.g., *Elseya flaviventralis*, *Emydura tanybaraga*) highlights the need to better understand those species to inform conservation.

An ongoing challenge for Australian freshwater turtle conservation has been the lack of awareness of their importance to freshwater systems. Consequently, they have received less attention than sea turtles, many high-profile mammals (e.g. koalas), and economically or recreationally important taxa (e.g., fish). Despite being historically undervalued, the increasing understanding of the importance of turtles to river health (Santori et al., 2020) is helping garner support for increased protection. In addition, turtles are popular among Australians, which as previously noted, has led to successful community science campaigns, including individual participation, community group conservation, school participation (Santori et al., 2021), and the production of conservation guidelines for practitioners and land managers. The significance of turtles to First Nations Australians has also facilitated conservation efforts by Indigenous ranger groups (Cochrane, 2006), and the Working on Country program provides an ongoing opportunity for conservation and land and river management led by First Nations People. The increasing recognition of the value of engaging First Nations People when setting priorities around conservation and the benefit of collaboration on country (Ward-Fear et al., 2019) is leading to improved conservation outcomes.

CONCLUSION

While research has unravelled much of the biology and conservation needs of Australian freshwater turtles, there remains much to be learned. With many species recommended for an upgrade to more severe listings in the Special Assessment Expert Panel process in 2022 (Petrov et al., This issue), the trend in declines for Australian freshwater turtles highlights the need for increasing

conservation efforts. A heavy bias exists between some well-studied species compared with others for which no information is published. Our management for almost all species suffers from a lack of long-term population level data to assess trends in population size and demography, critical to inform conservation measures. Technological advances, alongside the increased participation of community members and investment in freshwater turtle conservation by governments, will likely catalyse our capacity to respond with appropriate long-term conservation action for these remarkable long-lived creatures. Our special issue is dedicated to further unravelling the mystery of Australian freshwater turtles, and improving access to our knowledge of these unique animals.

AUTHOR CONTRIBUTIONS

Deborah Bower: Conceptualization (equal); project administration (equal); writing – original draft (equal). **Donald McKnight:** Conceptualization (equal); methodology (equal); writing – original draft (equal). **Kyra Sullivan:** Data curation (equal); project administration (equal); writing – review and editing (equal). **Stewart Macdonald:** Data curation (equal); formal analysis (equal); investigation (equal); visualization (lead). **Arthur Georges:** Conceptualization (supporting); writing – review and editing (supporting). **Simon Clulow:** Conceptualization (supporting); writing – review and editing (supporting). **Rupert Mathwin:** Conceptualization (supporting); writing – review and editing (supporting). **Marilyn Joy Connell:** Conceptualization (supporting); writing – review and editing (supporting). **Holly Nelson:** Conceptualization (supporting); writing – review and editing (supporting). **Anthony Santoro:** Conceptualization (supporting); writing – review and editing (supporting). **Bethany Nordstrom:** Conceptualization (supporting); writing – review and editing (supporting). **James U Van Dyke:** Conceptualization (supporting); writing – review and editing (supporting). **Rosie Kidman:** Conceptualization (supporting); writing – review and editing (supporting). **Louise Streeting:** Conceptualization (supporting); writing – review and editing (supporting). **Martin Dillon:** Conceptualization (supporting); writing – review and editing (supporting). **Ricky-John Spencer:** Conceptualization (supporting); writing – review and editing (supporting). **Michael Thompson:** Conceptualization (equal); writing – review and editing (supporting). **Eric Nordberg:** Conceptualization (equal); methodology (equal); writing – original draft (equal).

KEYWORDS

chelonian, freshwater turtle, tortoise

ACKNOWLEDGEMENTS

We acknowledge the Peramangk people as the traditional custodians of the land on which we held the 55th Annual General Meeting of the Australian Society of Herpetologists at Mylor in South Australia.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Deborah S. Bower[1]
Donald T. McKnight[2]
Kyra Sullivan[1]
Stewart L. Macdonald[3]
Arthur Georges[4]
Simon Clulow[5]
Rupert Mathwin[6]
Marilyn J. Connell[7]
Holly V. Nelson[8]
Anthony Santoro[9]
Bethany Nordstrom[10]
James U. Van Dyke[2]
Rosie A. Kidman[1]
Louise M. Streeting[1]
Martin L. Dillon[11]
Ricky-John Spencer[12]
Michael B. Thompson[8]
Eric J. Nordberg[1]

[1]*School of Environmental and Rural Science, University of New England, Armidale, New South Wales, Australia*
[2]*Centre for Freshwater Ecosystems, Department of Environment and Genetics, School of Agriculture, Biomedicine and Environment, La Trobe University Wodonga Campus, Wodonga, Victoria, Australia*
[3]*CSIRO Health & Biosecurity, Townsville, Queensland, Australia*
[4]*Institute for Applied Ecology, University of Canberra, Canberra, Australian Capital Territory, Australia*
[5]*Centre for Conservation Ecology and Genomics, Institute for Applied Ecology, University of Canberra, Bruce, Australian Capital Territory, Australia*
[6]*Global Ecology (Partuyarta Ngadluku Wardli Kuu), College of Science and Engineering, Flinders University, Adelaide, South Australia, Australia*
[7]*Tiaro & District Landcare Group Inc, Tiaro, Queensland, Australia*
[8]*School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales, Australia*
[9]*Centre for Sustainable Aquatic Ecosystems, Harry Butler Institute, Murdoch University, Perth, Western Australia, Australia*
[10]*School of Biological Sciences, The University of Western Australia, Crawley, Western Australia, Australia*
[11]*Northern Tablelands Local Land Services, Armidale, New South Wales, Australia*

[12]*School of Science, Western Sydney University, Penrith, New South Wales, Australia*

Correspondence
Deborah S. Bower, School of Environmental and Rural Science, University of New England, Armidale, NSW 2351, Australia.
Email: dbower3@une.edu.au

ORCID

Deborah S. Bower https://orcid.org/0000-0003-0188-3290
Donald T. McKnight https://orcid.org/0000-0001-8543-098X
Arthur Georges https://orcid.org/0000-0003-2428-0361
Rupert Mathwin https://orcid.org/0000-0001-9316-9451
James U. Van Dyke https://orcid.org/0000-0002-3933-111X
Rosie A. Kidman https://orcid.org/0000-0001-7115-2691
Louise M. Streeting https://orcid.org/0000-0002-1663-0010
Eric J. Nordberg https://orcid.org/0000-0002-1333-622X

TWITTER

Deborah S. Bower @deborah_bower

REFERENCES

Blamires, S.J. & Spencer, R.-J. (2013) Influence of habitat and predation on population dynamics of the freshwater turtle *Myuchelys georgesi*. *Herpetologica*, 69, 46–57.

Booth, D.T. (2000) Incubation of eggs of the Australian broad-shelled turtle, *Chelodina expansa* (Testudinata: Chelidae), at different temperatures: effects on pattern of oxygen consumption and hatchling morphology. *Australian Journal of Zoology*, 48, 369–378.

Bouma, A., Kuchling, G., Zhai, S.Y. & Mitchell, N. (2020) Assisted colonisation trials for the western swamp turtle show that juveniles can grow in cooler and wetter climates. *Endangered Species Research*, 43, 75–88.

Bower, D., Hutchinson, M. & Georges, A. (2012) Movement and habitat use of Australia's largest snake-necked turtle: implications for water management. *Journal of Zoology*, 287, 76–80.

Bower, D.S., Scheltinga, D.M., Clulow, S., Clulow, J., Franklin, C.E. & Georges, A. (2016) Salinity tolerances of two Australian freshwater turtles, *Chelodina expansa* and *Emydura macquarii* (Testudinata: Chelidae). *Conservation Physiology*, 4, cow042.

Campbell, M.A., Connell, M.J., Clark, N.J., Espinoza, T., Flakus, S.P., Collett, S.J. et al. (2023) Saving the endangered Mary River turtle: how community engagement delivered conservation outcomes. *Austral Ecology*, 1–14.

Campbell, M.A., Connell, M.J., Collett, S.J., Udyawer, V., Crewe, T.L., McDougall, A. et al. (2020) The efficacy of protecting turtle nests as a conservation strategy to reverse population decline. *Biological Conservation*, 251, 108769.

Cann, J. & Sadlier, R.A. (2017) *Freshwater turtles of Australia*. Melbourne, Vic: CSIRO Publishing.

Chessman, B.C. (2011) Declines of freshwater turtles associated with climatic drying in Australia's Murray–Darling basin. *Wildlife Research*, 38, 664–671.

Chessman, B.C. (2022) Introduced red foxes (*Vulpes vulpes*) driving Australian freshwater turtles to extinction? A critical evaluation of the evidence. *Pacific Conservation Biology*, 28, 462–471.

Chessman, B.C., Dillon, M.L., Fielder, D.P., Spark, P. & Streeting, L.M. (2023) Juvenile proportion as a predictor of freshwater turtle population change. *Austral Ecology*, 1–15.

Chessman, B.C., Fielder, D.P., Spark, P. & Steed, A.A. (2023) Distribution, habitat and population structure of the threatened Manning River helmeted turtle, Myuchelys purvisi. *Austral Ecology*, 1–12. Available from: https://doi.org/10.1111/aec.13366

Chessman, B.C., McGilvray, G., Ruming, S., Jones, H.A., Petrov, K., Fielder, D.P. et al. (2020) On a razor's edge: status and prospects of the critically endangered Bellinger River snapping turtle, *Myuchelys georgesi*. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 30, 586–600.

Clark, N., Gordos, M. & Franklin, C. (2008) Diving behaviour, aquatic respiration and blood respiratory properties: a comparison of hatchling and juvenile Australian turtles. *Journal of Zoology*, 275, 399–406.

Clulow, S., Clulow, J., Marcec-Greaves, R., Della Togna, G. & Calatayud, N.E. (2022) Common goals, different stages: the state of the ARTs for reptile and amphibian conservation. *Reproduction, Fertility and Development*, 34, i–ix.

Cochrane, M. (2006) *The Djelk ranger program: an outsider's perspective*. Canberra, ACT: Centre for Aboriginal Economic Policy Research (CAEPR), The Australian National University.

Coleman, D. (2023) Using baited remote underwater videos to survey freshwater turtles. *Austral Ecology*, 1–10. Available from: https://doi.org/10.1111/aec.13365

Connell, M.J. (2018) Freshwater turtle assemblages of the Mary River (Queensland, Australia), with a focus on the population of the endangered *Elusor macrurus*. Thesis, Degree of Masters by Research, Charles Darwin University.

Espinoza, T., Marshall, S.M., Limpus, D.J., Limpus, C.J. & McDougall, A.J. (2022) Adaptive management to reduce nest inundation of a critically endangered freshwater turtle: confirming the win-win. *Environmental Management*, 69, 972–981.

FitzGibbon, S. & Franklin, C. (2010) The importance of the cloacal bursae as the primary site of aquatic respiration in the freshwater turtle, *Elseya albagula*. *Australian Zoologist*, 35, 276–282.

Fordham, D., Georges, A., Corey, B. & Brook, B.W. (2006) Feral pig predation threatens the indigenous harvest and local persistence of snake-necked turtles in northern Australia. *Biological Conservation*, 133, 379–388.

Francis, R.J., Brandis, K.J., Chessman, B.C., Slavich, E. & Kingsford, R.T. (2022) Abundance and nutritional status of freshwater turtles to guide delivery of environmental flows in the Murray-Darling basin, South-Eastern Australia. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 32, 282–293.

Gawne, B., Hale, J., Stewardson, M.J., Webb, J.A., Ryder, D.S., Brooks, S.S. et al. (2020) Monitoring of environmental flow outcomes in a large river basin: the commonwealth environmental water Holder's long-term intervention in the Murray–Darling basin, Australia. *River Research and Applications*, 36, 630–644.

Georges, A., Birrell, J., Saint, K., McCord, W. & Donnellan, S. (1999) A phylogeny for side-necked turtles (Chelonia: Pleurodira) based on mitochondrial and nuclear gene sequence variation. *Biological Journal of the Linnean Society*, 67, 213–246.

Georges, A., Gruber, B., Pauly, G.B., White, D., Adams, M., Young, M.J. et al. (2018) Genomewide SNP markers breathe new life into phylogeography and species delimitation for the problematic short-necked turtles (Chelidae: *Emydura*) of eastern Australia. *Molecular Ecology*, 27, 5195–5213.

Giles, J. (2005) *The underwater acoustic repertoire of the long-necked, freshwater turtle Chelodina oblonga*. Perth: Murdoch University.

Glorioso, B.M., Vaughn, A.J. & Waddle, J.H. (2010) The aquatic turtle assemblage inhabiting a highly altered landscape in Southeast Missouri. *Journal of Fish and Wildlife Management*, 1, 161–168.

Gordos, M.A., Franklin, C.E., Limpus, C.J. & Wilson, G. (2004) Blood-respiratory and acid–base changes during extended diving in the bimodally respiring freshwater turtle *Rheodytes leukops*. *Journal of Comparative Physiology B*, 174, 347–354.

Hicks, D. & Kanevsky, R. (1992) Ninja turtles and other superheros: a case study of one literacy learner. *Linguistics and Education*, 4, 59–105.

Howell, H.J., Legere, R.H., Jr., Holland, D.S. & Seigel, R.A. (2019) Long-term turtle declines: protected is a verb, not an outcome. *Copeia*, 107, 493–501.

Iverson, J.B. (1982) Biomass in turtle populations: a neglected subject. *Oecologia*, 55, 69–76.

Kaczor, S.A. & Hartnett, D.C. (1990) Gopher tortoise (*Gopherus polyphemus*) effects on soils and vegetation in a Florida sandhill community. *American Midland Naturalist*, 123, 100–111.

Kennett, R., Christian, K. & Pritchard, D. (1993) Underwater nesting by the tropical fresh-water turtle, *Chelodina rugosa* (Testudinata, Chelidae). *Australian Journal of Zoology*, 41, 47–52.

Kennett, R., Roe, J., Hodges, K. & Georges, A. (2009) *Chelodina longicollis* (Shaw 1794)–eastern long-necked turtle, common long-necked turtle, common snake-necked turtle. *Conservation Biology of Freshwater Turtles and Tortoises*, 31, 1–8.

Kidman, R.A., McKnight, D.T., Schwarzkopf, L. & Nordberg, E. (2023) Too cold is better than too hot: preferred temperatures and basking behaviour in a tropical freshwater turtle. *Austral Ecology*, 1–10. Available from: https://doi.org/10.1111/aec.13335

Kuchling, G., Burbidge, A., Page, M. & Olejnik, C. (2018) Western swamp tortoise: slow and steady wins the race. In: Garnett, S., Latch, P., Lindenmayer, D. & Woinarski, J. (Eds.) *Recovering Australian threatened species: a book of hope*. Clayton: CSIRO Publishing, pp. 217–226.

Lettoof, D., Santoro, A., Swinstead, C.V. & Cornelis, J. (2021) First record of predation of a hatchling turtle by the Western tiger snake (*Notechis scutatus occidentalis*). *Australian Zoologist*, 41, 773–777.

Lovich, J.E., Ennen, J.R., Agha, M. & Gibbons, J.W. (2018) Where have all the turtles gone, and why does it matter? *Bioscience*, 68, 771–781.

McKnight, D.T. (2023) Courtship and homosexual behavior in saw-shelled turtles (*Myuchelys latisternum*). *Austral Ecology*, 1–5. Available from: https://doi.org/10.1111/aec.13364

McKnight, D.T., Ard, K., Auguste, R.J., Barhadiya, G., Benard, M.F., Boban, P. et al. (2023) Nocturnal basking in freshwater turtles: a global assessment. *Global Ecology and Conservation*, 43, e02444.

McKnight, D.T., Georges, A., Guarino, F. & Bower, D.S. (This issue) The ecology and morphology of Australia's desert turtle (*Emydura macquarii emmotti*). *Austral Ecology*.

Micheli-Campbell, M.A., Connell, M.J., Dwyer, R.G., Franklin, C.E., Fry, B., Kennard, M.J. et al. (2017) Identifying critical habitat for freshwater turtles: integrating long-term monitoring tools to enhance conservation and management. *Biodiversity and Conservation*, 26, 1675–1688.

Nordberg, E. & McKnight, D.T. (2023) Seasonal, environmental, and anthropogenic influences on nocturnal basking in turtles and crocodiles from North-Eastern Australia. *Austral Ecology*, 1–16. Available from: https://doi.org/10.1111/aec.13320

Nordberg, E.J. & McKnight, D.T. (2020) Nocturnal basking behavior in a freshwater turtle. *Ecology*, 101, 1–3.

Petrov, K., Lewis, J., Malkiewicz, N., Van Dyke, J.U. & Spencer, R.-J. (2018) Food abundance and diet variation in freshwater turtles from the mid-Murray River, Australia. *Australian Journal of Zoology*, 66, 67–76.

Petrov, K., Sutcliffe, S., Helen, T., Kutay, C., Eisemberg, C., Spencer, J.A. et al. (This issue) Turtles in trouble. Conservation status and priorities for Australian freshwater turtles. *Austral Ecology*.

Rappengliick, M.A. & Gilching, V. (2006) The whole world put between to shells: the cosmic symbolism of tortoises and turtles. *Mediterranean Archaeology and Archaeometry*, 4, 223–230.

Santori, C., Keith, R.J., Whittington, C.M., Thompson, M.B., Van Dyke, J.U. & Spencer, R.J. (2021) Changes in participant behaviour and attitudes are associated with knowledge and skills gained by using a turtle conservation citizen science app. *People and Nature*, 3, 66–76.

Santori, C., Spencer, R.-J., Thompson, M.B., Whittington, C.M., Burd, T.H., Currie, S.B. et al. (2020) Scavenging by threatened turtles regulates freshwater ecosystem health during fish kills. *Scientific Reports*, 10, 1–7.

Santoro, A. (2022) *The implications of climate change and urbanisation throughout the lifecycle of a freshwater turtle*. Perth: Murdoch University.

Santoro, A., Chambers, J.M., Robson, B.J. & Beatty, S.J. (2020) Land use surrounding wetlands influences urban populations of a freshwater turtle. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 30, 1050–1060.

Spencer, R.-J. (2018) How much long-term data are required to effectively manage a wide-spread freshwater turtle? *Australian Zoologist*, 39, 568–575.

Spencer, R.J. & Thompson, M.B. (2005) Experimental analysis of the impact of foxes on freshwater turtle populations. *Conservation Biology*, 19, 845–854.

Stanford, C.B., Iverson, J.B., Rhodin, A.G., van Dijk, P.P., Mittermeier, R.A., Kuchling, G. et al. (2020) Turtles and tortoises are in trouble. *Current Biology*, 30, R721–R735.

Streeting, L.M., Bower, D.S., Dillon, M.L., Spark, P., Gough, M., Skidmore, A. et al. (2022) Optimising the hatching success of artificially incubated eggs for use in a conservation program for the western saw-shelled turtle (*Myuchelys bellii*). *Australian Journal of Zoology*, 70, 74–82.

Streeting, L.M., Dillon, M.L., Nesbitt, B., Nesbitt, L., Baker, L., Spark, P. et al. (2023) A shocking result – electric fences protect Western saw-shelled turtle (*Myuchelys bellii*) nests from depredation by foxes. *Austral Ecology*, 1–17.

Terry, R., Robert, K., Simms, A., Stockfeld, G. & Van Dyke, J.U. (2023) Ineffectiveness of plastic mesh for protecting artificial freshwater turtle nests from red fox (*Vulpes vulpes*) predation. *Austral Ecology*, 1–17. Available from: https://doi.org/10.1111/aec.13362

Thompson, M.B. (1993) Hypothetical considerations of the biomass of chelid tortoises in the river Murray and the possible influences of predation by introduced foxes. In: Lunney, D. & Ayers, D. (Eds.) *Herpetology in Australia: a diverse discipline*. Sydney: Royal Zoological Society of New South Wales, pp. 219–224.

Thomson, R.C., Spinks, P.Q. & Shaffer, H.B. (2021) A global phylogeny of turtles reveals a burst of climate-associated diversification on continental margins. *Proceedings of the National Academy of Sciences of the United States of America*, 118, e2012215118.

Van Dyke, J.U., McPhan, L., Howard, K., Hamilton, T., Ross, L. & Watts, R. (This issue) Effects of wetland connectivity on overwintering and movement behaviours of Australian freshwater turtles. *Austral Ecology*.

Villacorta-Rath, C., Espinoza, T., Cockayne, B., Schaffer, J. & Burrows, D. (2022) Environmental DNA analysis confirms extant populations of the cryptic Irwin's turtle within its historical range. *BMC Ecology and Evolution*, 22, 1–14.

Ward-Fear, G., Rangers, B., Pearson, D., Bruton, M. & Shine, R. (2019) Sharper eyes see shyer lizards: collaboration with indigenous peoples can alter the outcomes of conservation research. *Conservation Letters*, 12, e12643.

Webb, G., Choquenot, D. & Whitehead, P.J. (1986) Nests, eggs, and embryonic development of *Carettochelys insculpta* (Chelonia: Carettochelidae) from northern Australia. *Journal of Zoology*, 1, 521–550.

Wenger, S.J., Subalusky, A.L. & Freeman, M.C. (2019) The missing dead: the lost role of animal remains in nutrient cycling in north American Rivers. *Food Webs*, 18, e00106.

A4.5 Bellinger River Snapping Turtle Genetic Management Final Report for Symbio Wildlife Park

The PDF version of reports contributed to during candidature are presented on the following pages.

Bellinger River Snapping Turtle Genetic Management

Final Report for Symbio Wildlife Park

November 2021

Holly Nelson, Dr Elspeth McLennan and Dr Carolyn Hogg

Australasian Wildlife Genomics Group, SOLES, University of Sydney, Sydney

Executive Summary

The Bellinger River Snapping Turtle (*M. georgesi*) is a critically endangered species currently listed and managed under the NSW Governments Saving our Species program. Included in the species management plan is regular genetic monitoring of the wild and captive populations (ProjectReport (nsw.gov.au), 2016). There are currently two captive populations maintained at Taronga Zoo and Symbio Wildlife Park, NSW. The Taronga population was founded from emergency intakes in 2015 (n = 17), with animals collected from the upper reaches of the Bellinger River where the virus had not yet reached. The Symbio population was founded post-virus in 2016 from surviving juveniles collected from the lower reaches (n = 19). To assess genetic diversity and relatedness in the captive populations, as well as place this population into context with the wild population, samples were collected from live founders from both captive and the wild population between 2015-2020. Samples from all populations were sequenced using the DArTseq™ platform and the data analysed using a custom pipeline from the University of Sydney. A total of 65 *M. georgesi* samples were genotyped using a set of 406SNP markers to assess genetic diversity, relatedness, and differentiation within and across sites.

Levels of genetic diversity were not statistically different between the captive and wild population. However, there is some degree of genetic differentiation between the captive and wild populations which is primarily being driven by captive individuals descended from the original intake of founders from Mt Buller and Mt Hotham, which is consistent with previous surveys that showed wild populations at Mt Buller and Mt Loch to be genetically distinct (Weeks et al. 2017). Although inbreeding is accumulating within the captive and wild populations, the genetic differentiation means these populations are suitable as source populations for genetic augmentation actions if required.

Australasian Wildlife Genomics Group
Faculty of Science
School of Environmental and Life Sciences
Room 219, Gunn Building (B19)
The University of Sydney NSW 2006 Australia

T +61 2 8627 4896
E Carolyn.hogg@sydney.edu.au

ABN 15 211 513 464
CRICOS 00026A

Recommendations

- While genetic diversity was not statistically different between the captive and wild populations of *M. georgesi*, there is some degree of genetic differentiation (F_{ST} = 0.033) between the captive and wild populations which is primarily being driven by the descendants of the original Mt Buller & Mt Hotham founders (Figure 1). This means these populations are suitable as source populations for genetic augmentation actions if required
- If considering a release of captive individuals to Mt Loch, then care should be taken to choose individuals with the lowest possible degree of relationship to avoid potential inbreeding between released individuals. Further, individuals descended from the Mt Buller and Mt Hotham lineages (separated out in the PCoA; Figure 1), should be selected, as these individuals represent different diversity to that currently present at Mt Loch.
- Sourcing contemporary genetic samples from the remaining wild populations in Victoria, Mt Buller and Mt Hotham, should be a priority to 1) place these populations in context with Mt Loch and the captive population, and 2) identify aspects of wild genetic diversity not currently represented in the captive population for future management efforts. Interbreeding between Mt Loch and Mt Buller has previously resulted in genetic rescue with increased reproductive success (Weeks et al. 2017). Such a rescue effort may also be achieved with the captive population to increase diversity and reduce inbreeding.
- Mean kinship values can be incorporated into the captive population studbook to reduce incidences of mating between close relatives.

Methods and Results

Samples

Skin tissue was collected during NSW Office of Environment and Heritage surveys (2015 – 2020). DNA was extracted from 69 unique *M. georgesi* samples and was sent to Diversity Arrays Technology for high-density DArTseq™ genotyping. DArTseq is a restriction enzyme approach that targets non-repetitive regions of the genome followed by sequencing of the resulting fragments.

Generating reliable genetic markers

Filtered sequences were aligned to the draft *M. georgesi* reference genome and Single nucleotide polymorphisms (SNPs) were called. The resulting dataset was filtered using a custom version of the R script detailed in Wright et al. (2019) to obtain a high-quality set of XX SNPs. Four samples, all from the Symbio population, had insufficient data quality. The remaining 65 samples had sufficient quality to be included in the analysis with an average of XX% missing data across the dataset.

Diversity analysis

We examined genetic diversity (standardised heterozygosity [H_S]) using the GENHET package for R (Coulon 2010; Table 1). Standardised heterozygosity is calculated by dividing individual observed heterozygosity by the average observed heterozygosity of all the samples included. This metric is used to compare populations because it places each individual into context with the rest of the population, as these values are relative to which samples are included. Genetic diversity was slightly higher in the wild population but not significantly different to the captive population (t-test statistic = -0.771, p = 0.446). Observed (H_O) and expected heterozygosity (H_E) were calculated across loci (Table 1), with no significant difference between observed and expected heterozygosity (Bartlett's test statistics = 0.00016, p = 0.990). Population inbreeding (F_{IS}) was calculated using the diveRsity package for R (Keenan et al. 2013). A negative F_{IS} value denotes that the observed average heterozygosity is greater than the expected average (negative = outbreeding, positive = inbreeding). Both populations exhibited significant inbreeding with a positive F_{IS} value and the 95% confidence intervals not including 0; however, neither of these values are currently of concern but it should be noted that inbreeding is accumulating in both populations.

Population structure analysis

F_{ST} is a common metric used in population genetics to assess population genetic differentiation. These values should be used in conjunction with all other metrics presented here as values can be skewed by the samples included in the analysis. There was significant population structure observed between the captive and wild MPP population (F_{ST} = 0.033 [0.031 – 0.035 95% CI, p <0.001]). In the instance of MPP, the pattern appears to be driven by captive individuals born in 2016 to parents from the original founder intake from Mt Buller and Mt Hotham, prior to supplementation with Mt Loch females in 2017. These individuals appear to represent different diversity to that currently present at Mt Loch (see samples inside the dashed ellipse in Figure 1).

The data presented in Table 1, the PCoA and the F_{ST} value should be used in tandem when making the most informed translocation decision about how representative of the wild the current captive population is and whether it would benefit from genetic augmentation.

Parentage analysis

Parentage reconstruction for the captive population of MPPs at Zoos Victoria (N = 34) was performed using the Sequoia package for R (Huisman 2017). The Sequoia package is particularly useful when sampling efficiency is not 100% as it includes sibship clustering and grandparent assignment based on likelihood analysis (Huisman 2017). To supplement gaps in the pedigree, we also performed a kinship analysis using COANCESTRY v1.0.1.9 which provided a relatedness estimate for each dyad of individuals. Marissa Parrot (Zoos Victoria) provided detailed housing information which further refined

the pedigree resulting in a final total of 24 individuals assigned both parents, one individual assigned one parent and three individuals identified only as siblings, with the remainder having wild parentage (Table 2). This pedigree can be used for the current studbook. If required, an empirical kinship file of all individuals included in this study can be provided to Zoos Victoria to be used in studbook analysis to account for founder relationships within the captive population.

Table 1 Summary of genetic diversity statistics for the captive and Mt Loch populations of MPP based on 3,986 SNPs. *N* = sample size, H_O = mean observed heterozygosity, H_E = expected heterozygosity, H_S = standardised heterozygosity (relative to each individual included in the analysis), SD = standard deviation, CI = 95% lower and upper confidence intervals, MK = mean kinship based on estimates of molecular relatedness.

	Symbio population	Taronga population	Wild population
N	15	17	33
H_O (± SD)	0.121 (0.006)	0.121 (0.006)	0.129 (0.005)
H_E (± SD)	0.206 (0.005)	0.106 (0.005)	0.126 (0.005)
H_S (± SD)	0.878 (0.018)	0.880 (0.066)	0.982 (0.081)
F_{IS} (95% CI)			
MK	0.05	0.01	

Table 2 Pedigree for the captive population of MPPs at Healesville Sanctuary. The coloured blocks denote half sibling relationships.

YOB	ID	Name	Sire ID	Sire name	Dam ID	Dam name
2015	B70102	Plum	Wild		Wild	
2015	B70103	Mrs Loch	Wild		Wild	
2015	B70104	Bev	Wild		Wild	
2016	B70106	Dibbin	Wild		Wild	
2016	B70113	Spargo	Wild		Wild	
2016	B70105	Derrick	Wild		Wild	
2016	B70112	Lachlan	Wild		Wild	
2016	B70100	Harriet	Wild		Wild	
2016	B70099	Nessie	Wild		Wild	
2016	B70101	Pod	Wild		Wild	
2012	B20889	Snug	B10551	Casanova	B01409	Macadamia
2016	B60568	Vera	B20889	Snug	B20745	Autumn
2016	B60569	Nancy	B20889	Snug	B20745	Autumn
2016	B60442	Eugene	B10485	Fang	B20745	Autumn
2016	B60437	Cheryl	B20889	Snug	B01664	Gracie
2016	B60436	Beryl	B20889	Snug	B01664	Gracie
2016	B60439	Pearl	B20886	Winter	B20804	Floss
2016	B60440	Trike	B20886	Winter	B20804	Floss
2016	B60476	Larry	B20886	Winter	B20804	Floss
2016	B60477	Scooter	B20886	Winter	B20804	Floss
2018	B81898	Bobble	B60442	Eugene	B70104	Bev
2018	B81899	Buddha	B60442	Eugene	B70104	Bev
2018	B81407	Bamboozle	B60442	Eugene	B70104	Bev
2018	B81408	Bumblebee	B60442	Eugene	B70104	Bev
2019	B92506	Noodle	B70105	Derrick	B70099	Nessie
2019	B92505	NugNug	B70105	Derrick	B70099	Nessie
2019	B92502	Missy H	B70105	Derrick	B70103	Mrs Loch
2019	B92503	Mog Gog	B70105	Derrick	B70103	Mrs Loch
2019	B92504	Mr Magoo	B70105	Derrick	B70103	Mrs Loch
2019	B92508	Flora	B70106	Dibbin	B70100	Harriet
2019	B92507	The Hoth			B70100	Harriet
2020	CO1531	Claf	B70112	Lachlan	B70102	Plum
2020	CO1532	Compote	B70112	Lachlan	B70102	Plum
2020	CO1533	Crumble	B70112	Lachlan	B70102	Plum
2020	CO1935	Cobber	B70112	Lachlan	B70102	Plum

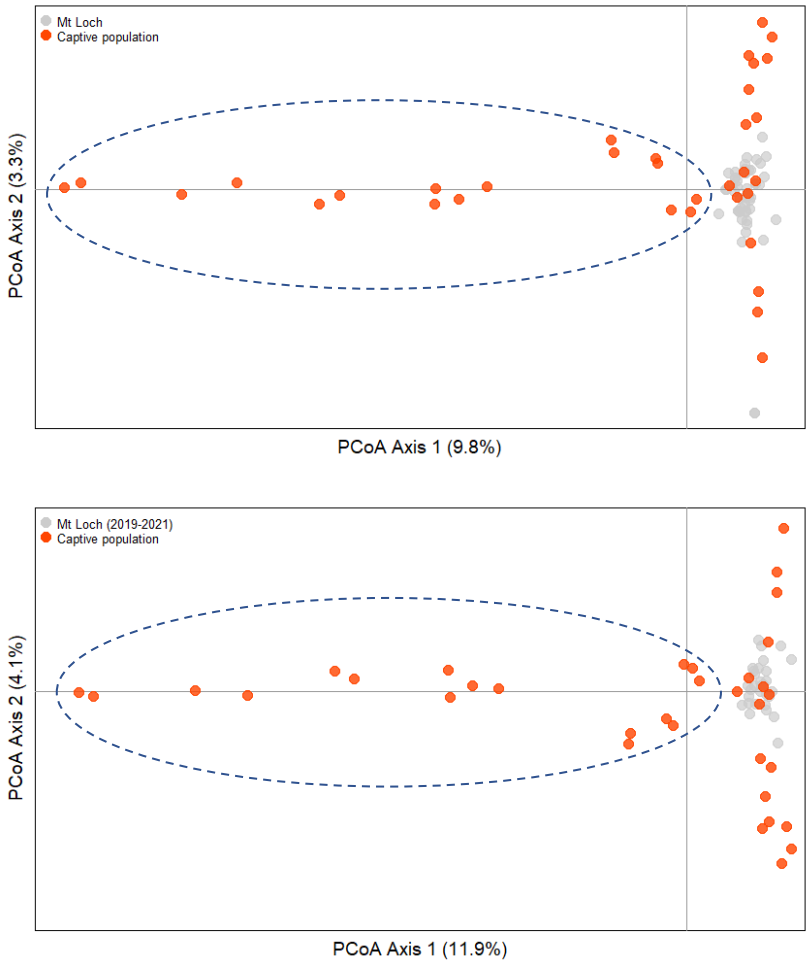


Figure 1 Principal coordinates analysis of a) 34 individuals from the captive MPP population at Zoos Victoria and 49 wild individuals from Mt Loch, Victoria, and b) 34 individuals from the captive MPP population at Zoos Victoria and 32 wild individuals from Mt Loch captured between 2019 and 2021 to represent the current diversity in the wild. Dots inside the dashed ellipse are the descendants of the original Mt Hotham and Mt Buller founders.

Acknowledgements

Thank you to Zoos Victoria, in particular Marissa Parrot, for providing samples for this work as well as sample metadata and historical information regarding the establishment of the captive breeding program. Thanks also to Andrew Weeks and Anthony van Rooyen (The University of Melbourne) for providing contemporary wild samples and metadata. Thank you to Kim Heasman for performing the DNA extractions and quality control assessments. This work was supported by Zoos Victoria and the Australian Government's Bushfire Recovery for Wildlife and their Habitats program under ERF-WRR2-020.

References

Coulon, A. 2010. GENHET: an easy-to-use R function to estimate individual heterozygosity. Molecular Ecology Resources **10**:167-169.

Huisman, J. 2017. Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. Molecular Ecology Resources **17**:1009-1024.

Keenan, K., P. McGinnity, T. F. Cross, W. W. Crozier, P. A. Prodöhl, and R. B. O'Hara. 2013. diveRsity: AnRpackage for the estimation and exploration of population genetics parameters and their associated errors. Methods in Ecology and Evolution **4**:782-788.

Weeks, A. R., D. Heinze, L. Perrin, J. Stoklosa, A. A. Hoffmann, A. van Rooyen, T. Kelly, and I. Mansergh. 2017. Genetic rescue increases fitness and aids rapid recovery of an endangered marsupial population. Nature Communications **8**:1071.

Wright, B., K. A. Farquharson, E. A. McLennan, K. Belov, C. J. Hogg, and C. E. Grueber. 2019. From reference genomes to population genomics: comparing three reference-aligned reduced-representation sequencing pipelines in two wildlife species. BMC Genomics **20**:453.

References

ADAMACK, A. T. & GRUBER, B. 2014. PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology & Evolution,* 5**,** 384-387.

ADAMS, C. I., HEPBURN, C., JEUNEN, G. J., CROSS, H., TAYLOR, H. R., GEMMELL, N. J., BUNCE, M. & KNAPP, M. 2023. Environmental DNA reflects common haplotypic variation. *Environmental DNA,* 5**,** 906-919.

AINSWORTH, G. B., REDPATH, S. M., WILSON, M., WERNHAM, C. & YOUNG, J. C. 2020. Integrating scientific and local knowledge to address conservation conflicts: Towards a practical framework based on lessons learned from a Scottish case study. *Environmental Science & Policy,* 107**,** 46-55.

AL-NAKEEB, K., PETERSEN, T. N. & SICHERITZ-PONTÉN, T. 2017. Norgal: extraction and de novo assembly of mitochondrial DNA from whole-genome sequencing data. *BMC Bioinformatics,* 18**,** 1-7.

ALFARO-NÚÑEZ, A., FROST BERTELSEN, M., BOJESEN, A. M., RASMUSSEN, I., ZEPEDA-MENDOZA, L., TANGE OLSEN, M. & GILBERT, M. T. P. 2014. Global distribution of Chelonid fibropapilloma-associated herpesvirus among clinically healthy sea turtles. *BMC Evolutionary Biology,* 14**,** 1-12.

ALLENDORF, F. W. 1983. Isolation, gene flow and genetic differentiation among populations. *Genetics & Conservation*.

ALLENDORF, F. W., LUIKART, G. H. & AITKEN, S. N. 2012. *Conservation and the genetics of populations*, John Wiley & Sons.

ALQAHTANI, F. & MĂNDOIU, I. SMART2: multi-library statistical mitogenome assembly with repeats. Computational Advances in Bio and Medical Sciences: 9th International Conference, ICCABS 2019, Miami, FL, USA, November 15–17, 2019, Revised Selected Papers 9, 2020. Springer International Publishing, 184-198.

ANDERSON, P. K., CUNNINGHAM, A. A., PATEL, N. G., MORALES, F. J., EPSTEIN, P. R. & DASZAK, P. 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution,* 19**,** 535-544.

ANDRES, K. J., LODGE, D. M., SETHI, S. A. & ANDRÉS, J. 2023. Detecting and analysing intraspecific genetic variation with eDNA: From population genetics to species abundance. *Molecular Ecology,* 32**,** 4118-4132.

ANDREW, P., COGGER, H., DRISCOLL, D., FLAKUS, S., HARLOW, P., MAPLE, D., MISSO, M., PINK, C., RETALLICK, K. & ROSE, K. 2018. Somewhat saved: a captive breeding programme for two endemic Christmas Island lizard species, now extinct in the wild. *Oryx,* 52**,** 171-174.

ANDREWS, S. 2010. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

ARDURA, A., ZAIKO, A., MARTINEZ, J. L., SAMULIOVIENE, A., SEMENOVA, A. & GARCIA-VAZQUEZ, E. 2015. eDNA and specific primers for early detection of invasive species–A case study on the bivalve *Rangia cuneata*, currently spreading in Europe. *Marine Environmental Research,* 112**,** 48-55.

AUTERI, G. G. & KNOWLES, L. L. 2020. Decimated little brown bats show potential for adaptive change. *Scientific Reports,* 10**,** 3023.

BACKUS, V. L., BRYANT, E. H., HUGHES, C. R. & MEFFERT, L. M. 1995. Effect of Migration or Inbreeding Followed by Selection on Low-Founder-Number Populations: Implications for Captive Breeding Programs. *Conservation Biology,* 9**,** 1216-1224.

BAKER, R. E., MAHMUD, A. S., MILLER, I. F., RAJEEV, M., RASAMBAINARIVO, F., RICE, B. L., TAKAHASHI, S., TATEM, A. J., WAGNER, C. E. & WANG, L.-F. 2022. Infectious disease in an era of global change. *Nature Reviews Microbiology,* 20**,** 193-205.

BALLOU, J. D. 1995. Identifying genetically important individuals for management of genetic variation in pedigreed populations. *Population Management for Survival & Recovery*.

BALLOU, J. D., LEES, C., FAUST, L. J., LONG, S., LYNCH, C., BINGAMAN LACKEY, L. & FOOSE, T. J. 2010. Demographic and genetic management of captive populations. *Wild Mammals in Captivity: Principles & Techniques for Zoo Management,* 219.

BALMFORD, A., LEADER-WILLIAMS, N. & GREEN, M. 1995. Parks or arks: where to conserve threatened mammals? *Biodiversity & Conservation,* 4**,** 595-607.

BARNES, M. A. & TURNER, C. R. 2016. The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics,* 17**,** 1-17.

BARNES, M. A., TURNER, C. R., JERDE, C. L., RENSHAW, M. A., CHADDERTON, W. L. & LODGE, D. M. 2014. Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science & Technology,* 48**,** 1819-1827.

BARRETT, K. G., AMARAL, G., ELPHINSTONE, M., MCADIE, M. L., DAVIS, C. S., JANES, J. K., CARNIO, J., MOEHRENSCHLAGER, A. & GORRELL, J. C. 2022. Genetic management on the brink of extinction: sequencing microsatellites does not improve estimates of inbreeding in wild and captive Vancouver Island marmots (*Marmota vancouverensis*). *Conservation Genetics,* 23**,** 417-428.

BENG, K. C. & CORLETT, R. T. 2020. Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. *Biodiversity & Conservation,* 29**,** 2089-2121.

BENTLEY, B. P., CARRASCO-VALENZUELA, T., RAMOS, E. K., PAWAR, H., ARANTES, L. S., ALEXANDER, A., BANERJEE, S. M., MASTERSON, P., KUHLWILM, M. & PIPPEL, M. 2022. Divergent sensory and immune gene evolution in sea turtles with contrasting demographic and life histories. *BioRxiv***,** 2022.01. 10.475373.

BERGERON, L. A., BESENBACHER, S., ZHENG, J., LI, P., BERTELSEN, M. F., QUINTARD, B., HOFFMAN, J. I., LI, Z., ST. LEGER, J. & SHAO, C. 2023. Evolution of the germline mutation rate across vertebrates. *Nature,* 615**,** 285-291.

BI, C., SHEN, F., HAN, F., QU, Y., HOU, J., XU, K., XU, L.-A., HE, W., WU, Z. & YIN, T. 2024. PMAT: an efficient plant mitogenome assembly toolkit using low-coverage HiFi sequencing data. *Horticulture Research,* 11**,** uhae023.

BLAMIRES, S. J., SPENCER, R.-J., KING, P. & THOMPSON, M. B. 2005. Population parameters and life-table analysis of two coexisting freshwater turtles: are the Bellinger River turtle populations threatened? *Wildlife Research,* 32**,** 339-347.

BLEHERT, D. S., HICKS, A. C., BEHR, M., METEYER, C. U., BERLOWSKI-ZIER, B. M., BUCKLES, E. L., COLEMAN, J. T., DARLING, S. R., GARGAS, A. & NIVER, R. 2009. Bat white-nose syndrome: an emerging fungal pathogen? *Science,* 323**,** 227-227.

BOLGER, A. & GIORGI, F. 2014. Trimmomatic: a flexible read trimming tool for illumina NGS data. *Bioinformatics,* 30**,** 2114-2120.

BOURQUE, J. R. 2016. Side-necked turtles (Testudines, Pleurodira) from the ancient Gulf coastal plain of Florida during middle Cenozoic megathermals. *Chelonian Conservation & Biology,* 15**,** 23-35.

BRANDIES, PEEL, HOGG & BELOV 2019. The Value of Reference Genomes in the Conservation of Threatened Species. *Genes,* 10.

BREARLEY, G., RHODES, J., BRADLEY, A., BAXTER, G., SEABROOK, L., LUNNEY, D., LIU, Y. & MCALPINE, C. 2013. Wildlife disease prevalence in human-modified landscapes. *Biological Reviews,* 88**,** 427-442.

BUSHNELL, B. 2014. BBMap: a fast, accurate, splice-aware aligner. Berkeley, California, United States: Lawrence Berkeley National Lab.

BUZAN, E., POTUŠEK, S., DUNIŠ, L. & POKORNY, B. 2022. Neutral and selective processes shape MHC diversity in roe deer in Slovenia. *Animals,* 12**,** 723.

BYLEMANS, J., FURLAN, E. M., GLEESON, D. M., HARDY, C. M. & DUNCAN, R. P. 2018. Does size matter? An experimental evaluation of the relative abundance and decay rates of aquatic environmental DNA. *Environmental Science & Technology,* 52**,** 6408-6416.

CABODEVILLA, X., GÓMEZ-MOLINER, B. J., ABAD, N. & MADEIRA, M. J. 2023. Simultaneous analysis of the intestinal parasites and diet through eDNA metabarcoding. *Integrative Zoology,* 18**,** 399-413.

CAMACHO, C., COULOURIS, G., AVAGYAN, V., MA, N., PAPADOPOULOS, J., BEALER, K. & MADDEN, T. L. 2009. BLAST+: architecture and applications. *BMC Bioinformatics,* 10**,** 421.

CANN, J., SPENCER, R.-J., WELSH, M. & GEORGES, A. 2015. *Myuchelys georgesi* (Cann 1997) – Bellinger River Turtle. *Chelonian Research Monographs***,** 1-9.

CARD, D. C., VAN CAMP, A. G., SANTONASTASO, T., JENSEN-SEAMAN, M. I., ANTHONY, N. M. & EDWARDS, S. V. 2022. Structure and evolution of the squamate major histocompatibility complex as revealed by two Anolis lizard genomes. *Frontiers in Genetics,* 13**,** 979746.

CARROLL, R. A., RICE, E. S., MURPHY, W. J., LYONS, L. A., THIBAUD-NISSEN, F., COGHILL, L. M., SWANSON, W. F., TERIO, K. A., BOYD, T. & WARREN, W. C. 2024. A chromosome-scale fishing cat reference genome for the evaluation of potential germline risk variants. *Scientific Reports,* 14**,** 8073.

CARVALHO, C. S., DE OLIVEIRA, M. E., RODRIGUEZ-CASTRO, K. G., SARANHOLI, B. H. & GALETTI JR, P. M. 2022. Efficiency of eDNA and iDNA in assessing vertebrate diversity and its abundance. *Molecular Ecology Resources,* 22**,** 1262-1273.

CATCHEN, J., HOHENLOHE, P. A., BASSHAM, S., AMORES, A. & CRESKO, W. A. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology,* 22**,** 3124-3140.

CEBALLOS, F. C., JOSHI, P. K., CLARK, D. W., RAMSAY, M. & WILSON, J. F. 2018. Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics,* 19**,** 220-234.

CHAFIN, T. K., ZBINDEN, Z. D., DOUGLAS, M. R., MARTIN, B. T., MIDDAUGH, C. R., GRAY, M. C., BALLARD, J. R. & DOUGLAS, M. E. 2021. Spatial population genetics in heavily managed species: Separating patterns of historical translocation from contemporary gene flow in white-tailed deer. *Evolutionary Applications,* 14**,** 1673-1689.

CHAPMAN, A. D. 2009. Numbers of living species in Australia and the world. *Australian Biological Resources Study.*

CHARLESWORTH, D., BARTON, N. H. & CHARLESWORTH, B. 2017. The sources of adaptive variation. *Proceedings of the Royal Society B: Biological Sciences,* 284**,** 20162864.

CHEN, C., XIAOYOU, H., WEI, L., LINGYUN, Y., HAIGANG, C. & XINPING, Z. 2021. Complete mitochondrial genome and the phylogenetic position of the Burmese narrow-headed softshell turtle *Chitra vandijki* (Testudines: Trionychidae). *Mitochondrial DNA Part B,* 6**,** 1216-1218.

CHEN, Z.-X., ZHENG, J.-C. & JIANG, Y.-L. 1999. A new iridovirus isolated from soft-shelled turtle. *Virus Research,* 63**,** 147-151.

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**,** 170-175.

CHESSMAN, B. C., MCGILVRAY, G., RUMING, S., JONES, H. A., PETROV, K., FIELDER, D. P., SPENCER, R. J. & GEORGES, A. 2020. On a razor's edge: Status and prospects of the critically endangered Bellinger River snapping turtle, *Myuchelys georgesi*. *Aquatic Conservation: Marine & Freshwater Ecosystems,* 30**,** 586-600.

CHHINA, A. K., ABHARI, N., MOOERS, A. & LEWTHWAITE, J. M. 2024. Linking the spatial and genomic structure of adaptive potential for conservation management: a review. *Genome,* 67**,** 403-423.

CHOJNACKA-PUCHTA, L. & SAWICKA, D. 2020. CRISPR/Cas9 gene editing in a chicken model: current approaches and applications. *Journal of Applied Genetics,* 61**,** 221-229.

CHRISTIE, M. R., MARINE, M. L., FRENCH, R. A. & BLOUIN, M. S. 2012. Genetic adaptation to captivity can occur in a single generation. *Proceedings of the National Academy of Sciences,* 109**,** 238-242.

COHEN, J. M., SAUER, E. L., SANTIAGO, O., SPENCER, S. & ROHR, J. R. 2020. Divergent impacts of warming weather on wildlife disease risk across climates. *Science,* 370**,** eabb1702.

COLLAR, N. 2020. Preparing captive-bred birds for reintroduction: the case of the Vietnam Pheasant *Lophura edwardsi*. *Bird Conservation International,* 30**,** 559-574.

COLLEN, B., WHITTON, F., DYER, E. E., BAILLIE, J. E., CUMBERLIDGE, N., DARWALL, W. R., POLLOCK, C., RICHMAN, N. I., SOULSBY, A. M. & BÖHM, M. 2014. Global patterns of freshwater species diversity, threat and endemism. *Global Ecology & Biogeography,* 23**,** 40-51.

COMMON, M. S. & NORTON, T. W. 1992. Biodiversity: its conservation in Australia. *Ambio***,** 258-265.

CONDE, D. A., COLCHERO, F., GUSSET, M., PEARCE-KELLY, P., BYERS, O., FLESNESS, N., BROWNE, R. K. & JONES, O. R. 2013. Zoos through the lens of the IUCN Red List: a global metapopulation approach to support conservation breeding programs. *PLoS One,* 8**,** e80311.

CONDE, D. A., FLESNESS, N., COLCHERO, F., JONES, O. R. & SCHEUERLEIN, A. 2011. An emerging role of zoos to conserve biodiversity. *Science,* 331**,** 1390-1391.

CONSORTIUM, I. H. G. S. 2004. Finishing the euchromatic sequence of the human genome. *Nature,* 431**,** 931-945.

CONVENTION ON BIOLOGICAL DIVERSITY 2020. Zero Draft of the Post-2020 Global Biodiversity Framework. (Convention on Biological Diversity).

CONWAY, W. G. 2011. Buying time for wild animals with zoos. Wiley Online Library.

COULON, A. 2010. GENHET: an easy-to-use R function to estimate individual heterozygosity. *Molecular Ecology Resources,* 10**,** 167-169.

COX, N., YOUNG, B. E., BOWLES, P., FERNANDEZ, M., MARIN, J., RAPACCIUOLO, G., BÖHM, M., BROOKS, T. M., HEDGES, S. B. & HILTON-TAYLOR, C. 2022. A global reptile assessment highlights shared conservation needs of tetrapods. *Nature,* 605**,** 285-290.

CRATES, R., STOJANOVIC, D. & HEINSOHN, R. 2023. The phenotypic costs of captivity. *Biological Reviews,* 98**,** 434-449.

CROWL, T. A., CRIST, T. O., PARMENTER, R. R., BELOVSKY, G. & LUGO, A. E. 2008. The spread of invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology & the Environment,* 6**,** 238-246.

DANECEK, P., AUTON, A., ABECASIS, G., ALBERS, C. A., BANKS, E., DEPRISTO, M. A., HANDSAKER, R. E., LUNTER, G., MARTH, G. T., SHERRY, S. T., MCVEAN, G., DURBIN, R. & GENOMES PROJECT ANALYSIS, G. 2011. The variant call format and VCFtools. *Bioinformatics,* 27**,** 2156-8.

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.

DARWALL, W., SMITH, K., ALLEN, D., SEDDON, M., REID, G. M., CLAUSNITZER, V. & KALKMAN, V. J. 2009. *Freshwater biodiversity: a hidden resource under threat*.

DASZAK, P., CUNNINGHAM, A. A. & HYATT, A. D. 2000. Emerging infectious diseases of wildlife-threats to biodiversity and human health. *Science,* 287**,** 443-449.

DAVEY, J. W., HOHENLOHE, P. A., ETTER, P. D., BOONE, J. Q., CATCHEN, J. M. & BLAXTER, M. L. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics,* 12**,** 499-510.

DE BRAUWER, M., CHARITON, A., CLARKE, L., COOPER, M., DIBATTISTA, J., FURLAN, E., GIBLOT-DUCRAY, D., GLEESON, D., HARFORD, A. & HERBERT, S. 2022a. Environmental DNA protocol development guide for biomonitoring.

DE BRAUWER, M., CHARITON, A., CLARKE, L., COOPER, M., DIBATTISTA, J., FURLAN, E., GIBLOT-DUCRAY, D., GLEESON, D., HARFORD, A. & HERBERT, S. 2022b. Environmental DNA test validation guidelines.

DE LEÓN, L. F., SILVA, B., AVILÉS-RODRÍGUEZ, K. J. & BUITRAGO-ROSAS, D. 2023. Harnessing the omics revolution to address the global biodiversity crisis. *Current Opinion in Biotechnology,* 80**,** 102901.

DE MIRANDA, E. B. 2017. The plight of reptiles as ecological actors in the tropics. *Frontiers in Ecology & Evolution,* 5**,** 159.

DEL MAR ORTEGA-VILLAIZAN, M., NOGUCHI, D. & TANIGUCHI, N. 2011. Minimization of genetic diversity loss of endangered fish species captive broodstocks by means of minimal kinship selective crossbreeding. *Aquaculture,* 318**,** 239-243.

DELORD, C., LASSALLE, G., OGER, A., BARLOY, D., COUTELLEC, M. A., DELCAMP, A., EVANNO, G., GENTHON, C., GUICHOUX, E. & LE BAIL, P. Y. 2018. A cost-and-time effective procedure to develop SNP markers for multiple species: A support for community genetics. *Methods in Ecology & Evolution,* 9**,** 1959-1974.

DÍAZ, S., SETTELE, J., BRONDÍZIO, E. S., NGO, H. T., AGARD, J., ARNETH, A., BALVANERA, P., BRAUMAN, K. A., BUTCHART, S. H. & CHAN, K. M. 2019. Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science,* 366**,** eaax3100.

DÍAZ-ABAD, L., BACCO-MANNINA, N., MADEIRA, F. M., NEIVA, J., AIRES, T., SERRAO, E. A., REGALLA, A., PATRÍCIO, A. R. & FRADE, P. R. 2022. eDNA metabarcoding for diet analyses of green sea turtles (*Chelonia mydas*). *Marine Biology,* 169**,** 18.

DO, C., WAPLES, R. S., PEEL, D., MACBETH, G., TILLETT, B. J. & OVENDEN, J. R. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources,* 14**,** 209-214.

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*.

DUDGEON, D., ARTHINGTON, A. H., GESSNER, M. O., KAWABATA, Z.-I., KNOWLER, D. J., LÉVÊQUE, C., NAIMAN, R. J., PRIEUR-RICHARD, A.-H., SOTO, D. & STIASSNY, M. L. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews,* 81**,** 163-182.

DUGOVICH, B. S., BEECHLER, B. R., DOLAN, B. P., CROWHURST, R. S., GONZALES, B. J., POWERS, J. G., HUGHSON, D. L., VU, R. K., EPPS, C. W. & JOLLES, A. E. 2023. Population connectivity patterns of genetic diversity, immune responses and exposure to infectious pneumonia in a metapopulation of desert bighorn sheep. *Journal of Animal Ecology,* 92**,** 1456-1469.

DURAND, N. C., ROBINSON, J. T., SHAMIM, M. S., MACHOL, I., MESIROV, J. P., LANDER, E. S. & AIDEN, E. L. 2016a. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Systems,* 3**,** 99-101.

DURAND, N. C., SHAMIM, M. S., MACHOL, I., RAO, S. S., HUNTLEY, M. H., LANDER, E. S. & AIDEN, E. L. 2016b. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Systems,* 3**,** 95-98.

ELDRIDGE, M. D., DEAKIN, J. E., MACDONALD, A. J., BYRNE, M., FITZGERALD, A., JOHNSON, R. N., MORITZ, C., PALMER, S. & YOUNG, A. 2020. The Oz Mammals Genomics (OMG) initiative: developing genomic resources for mammal conservation at a continental scale. *Australian Zoologist,* 40**,** 505-509.

ENQUIST, C. A., JACKSON, S. T., GARFIN, G. M., DAVIS, F. W., GERBER, L. R., LITTELL, J. A., TANK, J. L., TERANDO, A. J., WALL, T. U. & HALPERN, B. 2017. Foundations of translational ecology. *Frontiers in Ecology & the Environment,* 15**,** 541-550.

EXPOSITO-ALONSO, M., BOOKER, T. R., CZECH, L., GILLESPIE, L., HATELEY, S., KYRIAZIS, C. C., LANG, P. L., LEVENTHAL, L., NOGUES-BRAVO, D. & PAGOWSKI, V. 2022. Genetic diversity loss in the Anthropocene. *Science,* 377**,** 1431-1435.

FARQUHARSON, K. A., MCLENNAN, E. A., BELOV, K. & HOGG, C. J. 2023. The genome sequence of the critically endangered Kroombit tinkerfrog (*Taudactylus pleione*). *F1000Research,* 12.

FARQUHARSON, K. A., MCLENNAN, E. A., CHENG, Y., ALEXANDER, L., FOX, S., LEE, A. V., BELOV, K. & HOGG, C. J. 2022. Restoring faith in conservation action: Maintaining wild genetic diversity through the Tasmanian devil insurance program. *Iscience,* 25.

FEUERSTEIN, C. A., KOVACH, R. P., KRUSE, C. G., JAEGER, M. E., BELL, D. A., ROBINSON, Z. L. & WHITELEY, A. R. 2024. Genetic variation and hybridization determine the outcomes of conservation reintroductions. *Conservation Letters***,** e13049.

FIELDER, D., VERNES, K., ALACS, E. & GEORGES, A. 2012. Mitochondrial variation among Australian freshwater turtles (genus Myuchelys), with special reference to the Endangered M. bellii *Endangered Species Research,* 17**,** 63-71.

FLOT, J. F. 2010. seqphase: a web tool for interconverting phase input/output files and fasta sequence alignments. *Mol Ecol Resour,* 10**,** 162-6.

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,* 117**,** 9451-9457.

FOOTE, A. D., HOOPER, R., ALEXANDER, A., BAIRD, R. W., BAKER, C. S., BALLANCE, L., BARLOW, J., BROWNLOW, A., COLLINS, T. & CONSTANTINE, R. 2021. Runs of homozygosity in killer whale genomes provide a global record of demographic histories. *Molecular Ecology,* 30**,** 6162-6177.

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., DE VRIES, J., WALDVOGEL, A.-M., ZHANG, G., MAZZONI, C. J., JARVIS, E. D., BÁLINT, M., FORMENTI, G., THEISSINGER, K., FERNANDES, C., BISTA, I., BOMBARELY, A., BLEIDORN, C., ČIAMPOR, F., CIOFI, C., CROTTINI, A., GODOY, J. A., HOGLUND, J., MALUKIEWICZ, J., MOUTON, A., OOMEN, R. A., PAEZ, S., PALSBØLL, P., PAMPOULIE, C., RUIZ-LÓPEZ, M. J., SVARDAL, H., THEOFANOPOULOU, C., DE VRIES, J., WALDVOGEL, A.-M., ZHANG, G., MAZZONI, C. J., JARVIS, E., BÁLINT, M., AGHAYAN, S. A., ALIOTO, T. S., ALMUDI, I., ALVAREZ, N., ALVES, P. C., AMORIM, I. R., ANTUNES, A., ARRIBAS, P., BALDRIAN, P., BERG, P. R., BERTORELLE, G., BÖHNE, A., BONISOLI-ALQUATI, A., BOŠTJANČIĆ, L. L., BOUSSAU, B., BRETON, C. M., BUZAN, E., CAMPOS, P. F., CARRERAS, C., CASTRO, L. F., CHUECA, L. J., CONTI, E., COOK-DEEGAN, R., CROLL, D., CUNHA, M. V., DELSUC, F., DENNIS, A. B., DIMITROV, D., FARIA, R., FAVRE, A., FEDRIGO, O. D., FERNÁNDEZ, R., FICETOLA, G. F., FLOT, J.-F., GABALDÓN, T., GALEA AGIUS, D. R., GALLO, G. R., GIANI, A. M., GILBERT, M. T. P., GREBENC, T., GUSCHANSKI, K., GUYOT, R., HAUSDORF, B., HAWLITSCHEK, O., HEINTZMAN, P. D., HEINZE, B., HILLER, M., HUSEMANN, M., IANNUCCI, A., et al. 2022. The era of reference genomes in conservation genomics. *Trends in Ecology & Evolution,* 37**,** 197-202.

FRANDSEN, H. R., FIGUEROA, D. F. & GEORGE, J. A. 2020. Mitochondrial genomes and genetic structure of the Kemp's ridley sea turtle (*Lepidochelys kempii*). *Ecology & Evolution,* 10**,** 249-262.

FRANK, S. A. 2002. *Immunology and evolution of infectious disease*, Princeton University Press.

FRANKHAM, R. 1995. Conservation genetics. *Annual Review of Genetics,* 29**,** 305-327.

FRANKHAM, R. 1996. Relationship of genetic variation to population size in wildlife. *Conservation Biology,* 10**,** 1500-1508.

FRANKHAM, R. 2008. Genetic adaptation to captivity in species conservation programs. *Molecular Ecology,* 17**,** 325-333.

FRANKHAM, R. 2010. Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation,* 143**,** 1919-1927.

FRANKHAM, R. 2015. Genetic rescue of small inbred populations: Meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology,* 24**,** 2610-2618.

FRANKHAM, R., BALLOU, J. D., ELDRIDGE, M. D., LACY, R. C., RALLS, K., DUDASH, M. R. & FENSTER, C. B. 2011. Predicting the probability of outbreeding depression. *Conservation Biology,* 25**,** 465-475.

FRANKHAM, R., BALLOU, J. D., RALLS, K., ELDRIDGE, M., DUDASH, M. R., FENSTER, C. B., LACY, R. C. & SUNNUCKS, P. 2017. *Genetic management of fragmented animal and plant populations*, Oxford University Press.

FRANKHAM, R., BRISCOE, D. A. & BALLOU, J. D. 2010. *Introduction to conservation genetics*, Cambridge University Press.

FREELAND, J. R. 2017. The importance of molecular markers and primer design when characterizing biodiversity from environmental DNA. *Genome,* 60**,** 358-374.

FRYDENBERG, O. 1963. Population studies of a lethal mutant in *Drosophila melanogaster*: I. Behaviour in populations with discrete generations. *Hereditas,* 50**,** 89-116.

FU, M., EIMES, J. A., KONG, S., LAMICHHANEY, S. & WALDMAN, B. 2023. Identification of major histocompatibility complex genotypes associated with resistance to an amphibian emerging infectious disease. *Infection, Genetics & Evolution,* 113**,** 105470.

FUENTES-PARDO, A. P. & RUZZANTE, D. E. 2017. Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Molecular Ecology,* 26**,** 5369-5406.

GAERTNER, M., WILSON, J. R., CADOTTE, M. W., MACIVOR, J. S., ZENNI, R. D. & RICHARDSON, D. M. 2017. Non-native species in urban environments: patterns, processes, impacts and challenges. Springer.

GALLA, S. J., FORSDICK, N. J., BROWN, L., HOEPPNER, M. P., KNAPP, M., MALONEY, R. F., MORAGA, R., SANTURE, A. W. & STEEVES, T. E. 2018. Reference genomes from distantly related species can be used for discovery of single nucleotide polymorphisms to inform conservation management. *Genes,* 10**,** 9.

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., MICHALAK, P., BUCKLEY, T. R., WILSON, M., CHENG, Y., MILLER, H., SCHOTT, R. K., JORDAN, M. D., NEWCOMB, R. D., ARROYO, J. I., VALENZUELA, N., HORE, T. A., RENART, J., PEONA, V., PEART, C. R., WARMUTH, V. M., ZENG, L., KORTSCHAK, R. D., RAISON, J. M., ZAPATA, V. V., WU, Z., SANTESMASSES, D., MARIOTTI, M., GUIGÓ, R., RUPP, S. M., TWORT, V. G., DUSSEX, N., TAYLOR, H., ABE, H., BOND, D. M., PATERSON, J. M., MULCAHY, D. G., GONZALEZ, V. L., BARBIERI, C. G., DEMEO, D. P., PABINGER, S., VAN STIJN, T., CLARKE, S., RYDER, O., EDWARDS, S. V., SALZBERG, S. L., ANDERSON, L., NELSON, N., STONE, C., STONE, C., SMILLIE, J. & EDMONDS, H. 2020. The tuatara genome reveals ancient features of amniote evolution. *Nature,* 584**,** 403-409.

GEORGES, A. 2020. Assessment of hybridization and introgression in Bellinger River turtles, and assessment of parentage of hatchlings destined for release in the recovery program of *Myuchelys georgesi*. Parramatta, New South Wales, Australia: Biodiversity and Conservation, NSW Department of Planning, Industry and Environment.

GEORGES, A. & ADAMS, M. 1992. A phylogeny for Australian chelid turtles based on allozyme electrophoresis. *Australian Journal of Zoology,* 40**,** 453-476.

GEORGES, A., SPENCER, R.-J., KILIAN, A., WELSH, M. & ZHANG, X. 2018. Assault from all sides: hybridization and introgression threaten the already critically endangered *Myuchelys georgesi* (Chelonia: Chelidae). *Endangered Species Research,* 37**,** 239-247.

GEORGES, A., SPENCER, R.-J., WELSH, M., SHAFFER, H. B., WALSH, R. & ZHANG, X. 2011. Application of the precautionary principle to taxa of uncertain status: the case of the Bellinger River turtle. *Endangered Species Research,* 14**,** 127-134.

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. 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**,** 653-666.

GILLIGAN, D. M., BRISCOE, D. A. & FRANKHAM, R. 2005. Comparative losses of quantitative and molecular genetic variation in finite populations of *Drosophila melanogaster*. *Genetics Research,* 85**,** 47-55.

GILLIGAN, D. M. & FRANKHAM, R. 2003. Dynamics of genetic adaptation to captivity. *Conservation Genetics,* 4**,** 189-197.

GILPIN, M. E. 1986. Minimal viable populations: processes of species extinction. *Conservation Biology: The Science of Scarcity & Diversity*.

GIST, D. H. & JONES, J. M. 1987. Storage of sperm in the reptilian oviduct. *Scanning Microscopy,* 1**,** 31.

GOLDBERG, C. & STRICKLER, K. 2017. eDNA protocol sample collection. *Washington State University & University of Idaho collaborative report.*

GRANT, J. R., ENNS, E., MARINIER, E., MANDAL, A., HERMAN, E. K., CHEN, C.-Y., GRAHAM, M., VAN DOMSELAAR, G. & STOTHARD, P. 2023. Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Research***,** 484-492.

GRIEVES, L., GLOOR, G., BERNARDS, M. & MACDOUGALL-SHACKLETON, E. 2019. Songbirds show odour-based discrimination of similarity and diversity at the major histocompatibility complex. *Animal Behaviour,* 158**,** 131-138.

GRUBER, B., UNMACK, P. J., BERRY, O. F. & GEORGES, A. 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources,* 18**,** 691-699.

GRUEBER, C., PEEL, E., WRIGHT, B., HOGG, C. & BELOV, K. 2019. A Tasmanian devil breeding program to support wild recovery. *Reproduction, Fertility & Development,* 31**,** 1296-1304.

GRUEBER, C. E., FOX, S., MCLENNAN, E. A., GOOLEY, R. M., PEMBERTON, D., HOGG, C. J. & BELOV, K. 2018. Complex problems need detailed solutions: harnessing multiple data types to inform genetic management in the wild. *Evolutionary Applications*.

GUEDES, J. J., MOURA, M. R. & ALEXANDRE F. DINIZ-FILHO, J. 2023. Species out of sight: elucidating the determinants of research effort in global reptiles. *Ecography,* 2023**,** e06491.

GUÉGAN, J.-F., AYOUBA, A., CAPPELLE, J. & DE THOISY, B. 2020. Forests and emerging infectious diseases: unleashing the beast within. *Environmental Research Letters,* 15**,** 083007.

GUINDON, S., DUFAYARD, J.-F., LEFORT, V., ANISIMOVA, M., HORDIJK, W. & GASCUEL, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology,* 59**,** 307-321.

HAAS, B. & PAPANICOLAOU, A. 2016. TransDecoder (find coding regions within transcripts). *Google Scholar*.

HAHN, C., BACHMANN, L. & CHEVREUX, B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Research,* 41**,** e129-e129.

HARLEY, D., MAWSON, P. R., OLDS, L., MCFADDEN, M. & HOGG, C. 2018. The contribution of captive breeding in zoos to the conservation of Australia's threatened fauna. *Recovering Australian threatened species: A book of hope***,** 281-294.

HE, H., YANG, H., FOO, R., CHAN, W., ZHU, F., LIU, Y., ZHOU, X., MA, L., WANG, L.-F. & ZHAI, W. 2023a. Population genomic analysis reveals distinct demographics and recent adaptation in the black flying fox (*Pteropus alecto*). *Journal of Genetics & Genomics,* 50**,** 554-562.

HE, K., BABIK, W., MAJDA, M. & MINIAS, P. 2023b. MHC Architecture in Amphibians—Ancestral Reconstruction, Gene Rearrangements, and Duplication Patterns. *Genome Biology & Evolution,* 15**,** evad079.

HE, K., LIANG, C.-H., ZHU, Y., DUNN, P., ZHAO, A. & MINIAS, P. 2022a. Reconstructing macroevolutionary patterns in avian MHC architecture with genomic data. *Frontiers in Genetics,* 13**,** 823686.

HE, K., ZHU, Y., YANG, S.-C., YE, Q., FANG, S.-G. & WAN, Q.-H. 2022b. Major histocompatibility complex genomic investigation of endangered Chinese alligator provides insights into the evolution of

tetrapod major histocompatibility complex and survival of critically bottlenecked species. *Frontiers in Ecology & Evolution,* 10**,** 1078058.

HEBERT, P. D., RATNASINGHAM, S. & DE WAARD, J. R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences,* 270**,** S96-S99.

HEINSOHN, R., LACY, R., ELPHINSTONE, A., INGWERSEN, D., PITCHER, B. J., RODERICK, M., SCHMELITSCHEK, E., VAN SLUYS, M., STOJANOVIC, D. & TRIPOVICH, J. 2022. Population viability in data deficient nomadic species: what it will take to save regent honeyeaters from extinction. *Biological Conservation,* 266**,** 109430.

HEWETT, A. M., STOFFEL, M. A., PETERS, L., JOHNSTON, S. E. & PEMBERTON, J. M. 2023. Selection, recombination and population history effects on runs of homozygosity (ROH) in wild red deer (*Cervus elaphus*). *Heredity,* 130**,** 242-250.

HOANG, D. T., VINH, L. S., FLOURI, T., STAMATAKIS, A., VON HAESELER, A. & MINH, B. Q. 2018. MPBoot: fast phylogenetic maximum parsimony tree inference and bootstrap approximation. *BMC Evolutionary Biology,* 18**,** 1-11.

HOBAN, S., ARCHER, F. I., BERTOLA, L. D., BRAGG, J. G., BREED, M. F., BRUFORD, M. W., COLEMAN, M. A., EKBLOM, R., FUNK, W. C., GRUEBER, C. E., HAND, B. K., JAFFE, R., JENSEN, E., JOHNSON, J. S., KERSHAW, F., LIGGINS, L., MACDONALD, A. J., MERGEAY, J., MILLER, J. M., MULLER-KARGER, F., O'BRIEN, D., PAZ-VINAS, I., POTTER, K. M., RAZGOUR, O., VERNESI, C. & HUNTER, M. E. 2022. Global genetic diversity status and trends: towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition. *Biological Reviews*.

HOELZEL, A. R., BRUFORD, M. W. & FLEISCHER, R. C. 2019. Conservation of adaptive potential and functional diversity. *Conservation Genetics,* 20**,** 1-5.

HOFFMANN, A. A., SGRÒ, C. M. & KRISTENSEN, T. N. 2017. Revisiting adaptive potential, population size, and conservation. *Trends in Ecology & Evolution,* 32**,** 506-517.

HOFFMANN, M., HILTON-TAYLOR, C., ANGULO, A., BÖHM, M., BROOKS, T. M., BUTCHART, S. H., CARPENTER, K. E., CHANSON, J., COLLEN, B. & COX, N. A. 2010. The impact of conservation on the status of the world's vertebrates. *Science,* 330**,** 1503-1509.

HOGG, C., LEE, A., SRB, C. & HIBBARD, C. 2017a. Metapopulation management of an endangered species with limited genetic diversity in the presence of disease: the Tasmanian devil *Sarcophilus harrisii*. *International Zoo Yearbook,* 51**,** 137-153.

HOGG, C., MCLENNAN, E., WISE, P., LEE, A., PEMBERTON, D., FOX, S., BELOV, K. & GRUEBER, C. 2020. Preserving the demographic and genetic integrity of a single source population during multiple translocations. *Biological Conservation,* 241**,** 108318.

HOGG, C., WRIGHT, B., MORRIS, K., LEE, A., IVY, J., GRUEBER, C. & 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**,** 348-361.

HOGG, C. J. 2013. Preserving Australian native fauna: zoo-based breeding programs as part of a more unified strategic approach1. *Australian Journal of Zoology,* 61**,** 101-108.

HOGG, C. J., GRUEBER, C. E., PEMBERTON, D., FOX, S., LEE, A. V., IVY, J. A. & BELOV, K. 2017b. "Devil Tools & Tech": a synergy of conservation research and management practice. *Conservation Letters,* 10**,** 133-138.

HOGG, C. J., OTTEWELL, K., LATCH, P., ROSSETTO, M., BIGGS, J., GILBERT, A., RICHMOND, S. & BELOV, K. 2022. Threatened Species Initiative: Empowering conservation action using genomic resources. *Proceedings of the National Academy of Sciences,* 119**,** e2115643118.

HOHENLOHE, P. A., FUNK, W. C. & RAJORA, O. P. 2021. Population genomics for wildlife conservation and management. *Molecular Ecology,* 30**,** 62-82.

HOLDEREGGER, R., KAMM, U. & GUGERLI, F. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology,* 21**,** 797-807.

HORTON, R., WILMING, L., RAND, V., LOVERING, R. C., BRUFORD, E. A., KHODIYAR, V. K., LUSH, M. J., POVEY, S., TALBOT JR, C. C. & WRIGHT, M. W. 2004. Gene map of the extended human MHC. *Nature Reviews Genetics,* 5**,** 889-899.

HOWEY, C. A. & DINKELACKER, S. A. 2013. Characteristics of a historically harvested alligator snapping turtle (*Macrochelys temminckii*) population. *Copeia,* 2013**,** 58-63.

HU, H., NIGMATULINA, K. & ECKHOFF, P. 2013. The scaling of contact rates with population density for the infectious disease models. *Mathematical Biosciences,* 244**,** 125-134.

HU, T., CHITNIS, N., MONOS, D. & DINH, A. 2021. Next-generation sequencing technologies: An overview. *Human Immunology,* 82**,** 801-811.

HUDSON, P. J., RIZZOLI, A., GRENFELL, B. T., HEESTERBEEK, H. & DOBSON, A. P. 2002. *The ecology of wildlife diseases*, Oxford University Press Oxford.

HUGHES, A. L. & YEAGER, M. 1998. Natural selection at major histocompatibility complex loci of vertebrates. *Annual Review of Genetics,* 32**,** 415-435.

IUCN/SSC 2013. Guidelines for Reintroductions and Other Conservation Translocations. Version 1.0. Gland, Switzerland: IUCN Species Survival Commission.

IVY, J. A. & LACY, R. C. 2010. *Using molecular methods to improve the genetic management of captive breeding programs for threatened species*, Cambridge University Press.

IVY, J. A., MILLER, A., LACY, R. C. & DEWOODY, J. A. 2009. Methods and Prospects for Using Molecular Data in Captive Breeding Programs: An Empirical Example Using Parma Wallabies (*Macropus parma*). *Journal of Heredity,* 100**,** 441-454.

JAIN, M., KOREN, S., MIGA, K. H., QUICK, J., RAND, A. C., SASANI, T. A., TYSON, J. R., BEGGS, A. D., DILTHEY, A. T. & FIDDES, I. T. 2018. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature Biotechnology,* 36**,** 338-345.

JAKOB-HOFF, R., LEES, C. M., MCGILVRAY, G. & RUMING, S. 2017. Status review, disease risk analysis and conservation action plan for the Bellinger River snapping turtle. Apple Valley, Minnesota, United States: IUCN SSC Conservation Breeding Specialist Group, .

JARVIS, R. M., BORRELLE, S. B., FORSDICK, N. J., PÉREZ-HÄMMERLE, K. V., DUBOIS, N. S., GRIFFIN, S. R., RECALDE-SALAS, A., BUSCHKE, F., ROSE, D. C. & ARCHIBALD, C. L. 2020. Navigating spaces between conservation research and practice: Are we making progress? *Ecological Solutions & Evidence,* 1**,** e12028.

JAUREGUIBERRY, P., TITEUX, N., WIEMERS, M., BOWLER, D. E., COSCIEME, L., GOLDEN, A. S., GUERRA, C. A., JACOB, U., TAKAHASHI, Y. & SETTELE, J. 2022. The direct drivers of recent global anthropogenic biodiversity loss. *Science Advances,* 8**,** eabm9982.

JOHNS, G. C. & AVISE, J. C. 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene. *Molecular Biology & Evolution,* 15**,** 1481-1490.

JOHNSON, A. J., PESSIER, A. P., WELLEHAN, J. F., CHILDRESS, A., NORTON, T. M., STEDMAN, N. L., BLOOM, D. C., BELZER, W., TITUS, V. R. & WAGNER, R. 2008. Ranavirus infection of free-ranging and captive box turtles and tortoises in the United States. *Journal of Wildlife Diseases,* 44**,** 851-863.

JOMBART, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics,* 24**,** 1403-1405.

JONES, A., OVENDEN, J. & WANG, Y. 2016. Improved confidence intervals for the linkage disequilibrium method for estimating effective population size. *Heredity,* 117**,** 217-223.

JOYCE, W. G., PARHAM, J. F., LYSON, T. R., WARNOCK, R. C. & DONOGHUE, P. C. 2013. A divergence dating analysis of turtles using fossil calibrations: an example of best practices. *Journal of Paleontology,* 87**,** 612-634.

KALYAANAMOORTHY, S., MINH, B. Q., WONG, T. K., VON HAESELER, A. & JERMIIN, L. S. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods,* 14**,** 587-589.

KANG, Y.-J., YANG, D.-C., KONG, L., HOU, M., MENG, Y.-Q., WEI, L. & GAO, G. 2017. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Research,* 45**,** W12-W16.

KARDOS, M. & SHAFER, A. B. 2018. The peril of gene-targeted conservation. *Trends in Ecology & Evolution,* 33**,** 827-839.

KEENAN, K., MCGINNITY, P., CROSS, T. F., CROZIER, W. W. & PRODÖHL, P. A. 2013. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology & Evolution,* 4**,** 782-788.

KILIAN, A., WENZL, P., HUTTNER, E., CARLING, J., XIA, L., BLOIS, H., CAIG, V., HELLER-USZYNSKA, K., JACCOUD, D. & HOPPER, C. 2012. Diversity arrays technology: a generic genome profiling technology on open platforms. *Data Production and Analysis in Population Genomics: Methods & Protocols***,** 67-89.

KLEINMAN-RUIZ, D., MARTÍNEZ-CRUZ, B., SORIANO, L., LUCENA-PEREZ, M., CRUZ, F., VILLANUEVA, B., FERNÁNDEZ, J. & GODOY, J. A. 2017. Novel efficient genome-wide SNP panels for the conservation of the highly endangered Iberian lynx. *BMC Genomics,* 18**,** 1-12.

KLÜTSCH, C. F. & LAIKRE, L. 2021. Closing the conservation genetics gap: integrating genetic knowledge in conservation management to ensure evolutionary potential. *Closing the Knowledge-Implementation Gap in Conservation Science: Interdisciplinary Evidence Transfer Across Sectors and Spatiotemporal Scales***,** 51-82.

KNAUS, B. J. & GRÜNWALD, N. J. 2017. vcfr: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources,* 17**,** 44-53.

KOSCH, T. A., WADDLE, A. W., COOPER, C. A., ZENGER, K. R., GARRICK, D. J., BERGER, L. & SKERRATT, L. F. 2022. Genetic approaches for increasing fitness in endangered species. *Trends in Ecology & Evolution,* 37**,** 332-345.

KRAUS, R. H., VONHOLDT, B., COCCHIARARO, B., HARMS, V., BAYERL, H., KÜHN, R., FÖRSTER, D. W., FICKEL, J., ROOS, C. & NOWAK, C. 2015. A single-nucleotide polymorphism-based approach for rapid and cost-effective genetic wolf monitoring in Europe based on noninvasively collected samples. *Molecular Ecology Resources,* 15**,** 295-305.

KUMAZAWA, Y. & NISHIDA, M. 1999. Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles. *Molecular Biology & Evolution,* 16**,** 784-792.

KUNDU, S., KUMAR, V., TYAGI, K., CHAKRABORTY, R. & CHANDRA, K. 2019. The first complete mitochondrial genome of the Indian Tent Turtle, *Pangshura tentoria* (Testudines: Geoemydidae): Characterization and comparative analysis. *Ecology & Evolution,* 9**,** 10854-10868.

KUNDU, S., KUMAR, V., TYAGI, K. & CHANDRA, K. 2020. The complete mitochondrial genome of the endangered Assam Roofed Turtle, *Pangshura sylhetensis* (Testudines: Geoemydidae): Genomic features and phylogeny. *PLoS One,* 15**,** e0225233.

KUO, R. I., CHENG, Y., ZHANG, R., BROWN, J. W., 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**,** 1-22.

LACY, R. C. 1987. Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology,* 1**,** 143-158.

LACY, R. C., BALLOU, J. D., STARFIELD, A., THOMPSON, E. & THOMAS, A. 1995. Pedigree analysis for population management. *Population Management for Survival & Recovery*.

LAM, I. P., SUNG, Y.-H. & FONG, J. J. 2022. Using eDNA techniques to find the endangered big-headed turtle (*Platysternon megacephalum*). *PloS One,* 17**,** e0262015.

LANDE, R. & BARROWDOUGH, G. 1987. Effective population size, genetic variation, and their use in population. *Viable Populations for Conservation***,** 87.

LE BOUTEILLER, P. & LENFANT, F. 1996. Antigen-presenting function(s) of the non-classical HLA-E, -F and -G class I molecules: the beginning of a story. *Research in Immunology,* 147**,** 301-313.
LECLAIRE, S., VAN DONGEN, W. F., VOCCIA, S., MERKLING, T., DUCAMP, C., HATCH, S. A., BLANCHARD, P., DANCHIN, É. & WAGNER, R. H. 2014. Preen secretions encode information on MHC similarity in certain sex-dyads in a monogamous seabird. *Scientific Reports,* 4**,** 6920.
LEMAOULT, J., LE DISCORDE, M., ROUAS-FREISS, N., MOREAU, P., MENIER, C., MCCLUSKEY, J. & CAROSELLA, E. 2003. Biology and functions of human leukocyte antigen-G in health and sickness. *Tissue Antigens,* 62**,** 273-284.
LEUS, K. 2011. Captive breeding and conservation. *Zoology in the Middle East,* 54**,** 151-158.
LEWIN, H. A., ROBINSON, G. E., KRESS, W. J., BAKER, W. J., CODDINGTON, J., CRANDALL, K. A., DURBIN, R., EDWARDS, S. V., FOREST, F. & GILBERT, M. T. P. 2018. Earth BioGenome Project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences,* 115**,** 4325-4333.
LI, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics,* 27**,** 2987-93.
LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics,* 25**,** 1754-1760.
LI, H. & DURBIN, R. 2011. Inference of human population history from individual whole-genome sequences. *Nature,* 475**,** 493-496.
LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G. & DURBIN, R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics,* 25**,** 2078-2079.
LI, Q., SHAO, W., JIANG, Y., YAN, C. & LIAO, W. 2024. Assessing reptile conservation status under global climate change. *Biology,* 13**,** 436.
LOPES, C. M., BAÊTA, D., VALENTINI, A., LYRA, M. L., SABBAG, A. F., GASPARINI, J. L., DEJEAN, T., HADDAD, C. F. B. & ZAMUDIO, K. R. 2021. Lost and found: Frogs in a biodiversity hotspot rediscovered with environmental DNA. *Molecular Ecology,* 30**,** 3289-3298.
LYSON, T. R. & BEVER, G. S. 2020. Origin and evolution of the turtle body plan. *Annual Review of Ecology, Evolution, & Systematics,* 51**,** 143-166.
MADEIRA, F., PARK, Y. M., LEE, J., BUSO, N., GUR, T., MADHUSOODANAN, N., BASUTKAR, P., TIVEY, A. R. N., POTTER, S. C., FINN, R. D. & LOPEZ, R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research,* 47**,** W636-W641.
MAGID, M., WOLD, J. R., MORAGA, R., CUBRINOVSKA, I., HOUSTON, D. M., GARTRELL, B. D. & STEEVES, T. E. 2022. Leveraging an existing whole-genome resequencing population data set to characterize toll-like receptor gene diversity in a threatened bird. *Molecular Ecology Resources,* 22**,** 2810-2825.
MALLINSON, J. J. 1995. Conservation breeding programmes: an important ingredient for species survival. *Biodiversity & Conservation,* 4**,** 617-635.
MANLIK, O., KRÜTZEN, M., KOPPS, A. M., MANN, J., BEJDER, L., ALLEN, S. J., FRÈRE, C., CONNOR, R. C. & SHERWIN, W. B. 2019. Is MHC diversity a better marker for conservation than neutral genetic diversity? A case study of two contrasting dolphin populations. *Ecology & Evolution,* 9**,** 6986-6998.
MANNI, M., BERKELEY, M. R., SEPPEY, M., SIMÃO, F. A. & ZDOBNOV, E. M. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology & Evolution,* 38**,** 4647-4654.
MARSCHANG, R., BECHER, P., POSTHAUS, H., WILD, P., THIEL, H.-J., MÜLLER-DOBLIES, U., KALET, E. & BACCIARINI, L. 1999. Isolation and characterization of an iridovirus from Hermann's tortoises (*Testudo hermanni*). *Archives of virology,* 144**,** 1909-1922.
MARTINEZ, P. A., EZAZ, T., VALENZUELA, N., GEORGES, A. & MARSHALL GRAVES, J. A. 2008. An XX/XY heteromorphic sex chromosome system in the Australian chelid turtle *Emydura macquarii*: a

new piece in the puzzle of sex chromosome evolution in turtles. *Chromosome Research,* 16**,** 815-825.

MAUVISSEAU, Q., BURIAN, A., GIBSON, C., BRYS, R., RAMSEY, A. & SWEET, M. 2019. Influence of accuracy, repeatability and detection probability in the reliability of species-specific eDNA based approaches. *Scientific Reports,* 9**,** 580.

MCCALLUM, M. L. 2021. Turtle biodiversity losses suggest coming sixth mass extinction. *Biodiversity & Conservation,* 30**,** 1257-1275.

MCCAULEY, M., KODA, S. A., LOESGEN, S. & DUFFY, D. J. 2024. Multicellular species environmental DNA (eDNA) research constrained by overfocus on mitochondrial DNA. *Science of The Total Environment,* 912**,** 169550.

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 Res,* 20**,** 1297-303.

MCLENNAN, E., CHENG, Y., FARQUHARSON, K., GRUEBER, C., ELMER, J., ALEXANDER, L., FOX, S., BELOV, K. & HOGG, C. 2024. Reinforcements in the face of ongoing threats: a case study from a critically small carnivore population. *Animal Conservation*.

MCLENNAN, E. A., WRIGHT, B. R., BELOV, K., HOGG, C. J. & GRUEBER, C. E. 2019. Too much of a good thing? Finding the most informative genetic data set to answer conservation questions. *Molecular Ecology Resources,* 19**,** 659-671.

MCQUILLAN, R., LEUTENEGGER, A. L., ABDEL-RAHMAN, R., FRANKLIN, C. S., PERICIC, M., BARAC-LAUC, L., SMOLEJ-NARANCIC, N., JANICIJEVIC, B., POLASEK, O. & TENESA, A. 2008. Runs of homozygosity in European populations. *The American Journal of Human Genetics,* 83**,** 359-372.

MENAJOVSKY, M., ESPUNYES, J., CABEZON, O. & MAYOR, P. 2023. Infectious diseases of interest for the conservation of peccaries in the Amazon: A systematic quantitative review. *Biological Conservation,* 277**,** 109867.

MENG, G., LI, Y., YANG, C. & LIU, S. 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Research,* 47**,** e63-e63.

MENG, J., XU, F., YANG, H., LI, X. & ZHAO, P. 2024. Exploring microbiome and plankton responses and interactions in the mangrove ecosystem through eDNA and network analysis. *Science of The Total Environment,* 930**,** 172581.

MENNING, D. M., UHER-KOCH, B. D., FLAMME, M. J., SIMMONS, T., SCHMUTZ, J. A. & TALBOT, S. L. 2023. eDNA metabarcoding analyses of diet in Yellow-billed Loons of Northern Alaska. *Waterbirds,* 45**,** 159-166.

MEYER, A. 1994. Shortcomings of the cytochrome b gene as a molecular marker. *Trends in Ecology & Evolution***,** 278-280.

MILLER, H. C., BELOV, K. & DAUGHERTY, C. H. 2005. Characterization of MHC class II genes from an ancient reptile lineage, *Sphenodon* (tuatara). *Immunogenetics,* 57**,** 883-891.

MILLER, H. C., O'MEALLY, D., EZAZ, T., AMEMIYA, C., MARSHALL-GRAVES, J. A. & EDWARDS, S. 2015. Major Histocompatibility Complex Genes Map to Two Chromosomes in an Evolutionarily Ancient Reptile, the Tuatara *Sphenodon punctatus*. *G3: Genes, Genomes, Genetics,* 5**,** 1439-1451.

MINH, B. Q., SCHMIDT, H. A., CHERNOMOR, O., SCHREMPF, D., WOODHAMS, M. D., VON HAESELER, A. & LANFEAR, R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology & Evolution,* 37**,** 1530-1534.

MOONEY, J. A., YOHANNES, A. & LOHMUELLER, K. E. 2021. The impact of identity by descent on fitness and disease in dogs. *Proceedings of the National Academy of Sciences,* 118**,** e2019116118.

MORITZ, C., DOWLING, T. & BROWN, W. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology & Systematics,* 18**,** 269-292.

MULLIN, D. I., WHITE, R. C., LENTINI, A. M., BROOKS, R. J., BÉRIAULT, K. R. & LITZGUS, J. D. 2020. Predation and disease limit population recovery following 15 years of headstarting an endangered freshwater turtle. *Biological Conservation,* 245**,** 108496.

MYERS, S. S., GAFFIKIN, L., GOLDEN, C. D., OSTFELD, R. S., H. REDFORD, K., H. RICKETTS, T., TURNER, W. R. & OSOFSKY, S. A. 2013. Human health impacts of ecosystem alteration. *Proceedings of the National Academy of Sciences,* 110**,** 18753-18760.

NAPOLITANO, C., SACRISTÁN, I., ACUÑA, F., AGUILAR, E., GARCÍA, S., LÓPEZ-JARA, M. J., CABELLO, J., HIDALGO-HERMOSO, E., POULIN, E. & GRUEBER, C. E. 2023. Assessing micro-macroparasite selective pressures and anthropogenic disturbance as drivers of immune gene diversity in a Neotropical wild cat. *Science of the Total Environment,* 897**,** 166289.

NARDONI, S. & MANCIANTI, F. 2023. Mycotic Diseases in Chelonians. *Journal of Fungi,* 9**,** 518.

NEEFJES, J., JONGSMA, M. L., PAUL, P. & BAKKE, O. 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology,* 11**,** 823-836.

NELSON, H. V., FARQUHARSON, K. A., GEORGES, A., MCLENNAN, E. A., DEGABRIEL, J. L., GIESE, M., ORMOND, C., MCFADDEN, M., SKIDMORE, A. & PRANGELL, J. 2024. A genomic framework to assist conservation breeding and translocation success: A case study of a critically endangered turtle. *Conservation Science & Practice***,** e13204.

NICHOLS, S., EWEN, J. G., GOTTELLI, D., GRUEBER, C. E., SANTURE, A. W., TRASK, A. & BREKKE, P. 2024. Genetic rescue attempt in a small, inbred population of a wild endangered passerine. *Biological Conservation,* 290**,** 110430.

NISTELBERGER, H. M., ROYCROFT, E., MACDONALD, A. J., MCARTHUR, S., WHITE, L. C., GRADY, P. G., PIERSON, J., SIMS, C., COWEN, S. & MOSEBY, K. 2023. Genetic mixing in conservation translocations increases diversity of a keystone threatened species, *Bettongia lesueur*. *Molecular Ecology*.

NORDSTROM, B., MITCHELL, N., BYRNE, M. & JARMAN, S. 2022. A review of applications of environmental DNA for reptile conservation and management. *Ecology & Evolution,* 12**,** e8995.

OCHOA, A., WELLS, S. A., WEST, G., AL-SMADI, M. E., REDONDO, S. A., SEXTON, S. R. & CULVER, M. 2016. Can captive populations function as sources of genetic variation for reintroductions into the wild? A case study of the Arabian oryx from the Phoenix Zoo and the Shaumari Wildlife Reserve, Jordan. *Conservation Genetics,* 17**,** 1145-1155.

OLIVER, M. K. & PIERTNEY, S. B. 2012. Selection maintains MHC diversity through a natural population bottleneck. *Molecular Biology & Evolution,* 29**,** 1713-1720.

OLSEN, G. J., LANE, D. J., GIOVANNONI, S. J., PACE, N. R. & STAHL, D. A. 1986. Microbial ecology and evolution: a ribosomal RNA approach. *Annual Reviews in Microbiology,* 40**,** 337-365.

OLSON, D. H. & PILLIOD, D. S. 2022. Elevating human dimensions of amphibian and reptile conservation, a USA perspective. *Conservation Science & Practice,* 4**,** e12685.

ONLEY, I. R., MOSEBY, K. E. & AUSTIN, J. J. 2021. Genomic approaches for conservation management in Australia under climate change. *Life,* 11**,** 653.

OSSIBOFF, R. J., RAPHAEL, B. L., AMMAZZALORSO, A. D., SEIMON, T. A., NEWTON, A. L., CHANG, T. Y., ZARATE, B., WHITLOCK, A. L. & MCALOOSE, D. 2015. Three novel herpesviruses of endangered *Clemmys* and *Glyptemys* turtles. *PLoS One,* 10**,** e0122901.

PARRISH, K., KIRKLAND, P., HORWOOD, P., CHESSMAN, B., RUMING, S., MCGILVRAY, G., ROSE, K., HALL, J. & SKERRATT, L. 2024. Delving into the Aftermath of a Disease-Associated Near-Extinction Event: A Five-Year Study of a Serpentovirus (Nidovirus) in a Critically Endangered Turtle Population. *Viruses,* 16**,** 653.

PEAKALL, R. & SMOUSE, P. E. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* 6**,** 288-295.

PEEL, E., NG, H. J. J. & BELOV, K. 2019. Monotreme, marsupial and bat immunology. *Current therapy in medicine of Australian mammals.* Collingwood, Victoria, Australia.

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.

PELLEGRINO, I., BOATTI, L., CUCCO, M., MIGNONE, F., KRISTENSEN, T. N., MUCCI, N., RANDI, E., RUIZ-GONZALEZ, A. & PERTOLDI, C. 2016. Development of SNP markers for population structure and phylogeography characterization in little owl (*Athene noctua*) using a genotyping-by-sequencing approach. *Conservation Genetics Resources,* 8**,** 13-16.

PEMBERTON, T. J., ABSHER, D., FELDMAN, M. W., MYERS, R. M., ROSENBERG, N. A. & LI, J. Z. 2012. Genomic patterns of homozygosity in worldwide human populations. *The American Journal of Human Genetics,* 91**,** 275-292.

PENN, D. J., DAMJANOVICH, K. & POTTS, W. K. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences,* 99**,** 11260-11264.

PERTEA, M., PERTEA, G. M., ANTONESCU, C. M., CHANG, T.-C., MENDELL, J. T. & SALZBERG, S. L. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology,* 33**,** 290-295.

PIERTNEY, S. & OLIVER, M. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity,* 96**,** 7-21.

PIOTROWSKI, J. S., ANNIS, S. L. & LONGCORE, J. E. 2004. Physiology of Batrachochytrium dendrobatidis, a chytrid pathogen of amphibians. *Mycologia,* 96**,** 9-15.

PRICE, M. R. S. 1989. *Animal reintroductions: the Arabian oryx in Oman*, Cambridge University Press.

PRITCHARD, D. J., FA, J. E., OLDFIELD, S. & HARROP, S. R. 2012. Bring the captive closer to the wild: redefining the role of ex situ conservation. *Oryx,* 46**,** 18-23.

PRITCHARD, R. A., KELLY, E. L., BIGGS, J. R., EVERAARDT, A. N., LOYN, R., MAGRATH, M. J., MENKHORST, P., HOGG, C. J. & GEARY, W. L. 2022. Identifying cost-effective recovery actions for a critically endangered species. *Conservation Science & Practice,* 4**,** e546.

QUINLAN, A. R. & HALL, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* 26**,** 841-842.

R CORE TEAM 2023. R: A Language and Environment for Statistical Computing. 2023.06.0+421 ed. R Foundation for Statistical Computing, Vienna, Austria.

RAJ, A., STEPHENS, M. & PRITCHARD, J. K. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics,* 197**,** 573-589.

RALLS, K. & BALLOU, J. D. 2004. Genetic status and management of California condors. *The Condor,* 106**,** 215-228.

REED, D. H. & FRANKHAM, R. 2001. How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution,* 55**,** 1095-1103.

REED, K. M. & SETTLAGE, R. E. 2021. Major histocompatibility complex genes and locus organization in the Komodo dragon (*Varanus komodoensis*). *Immunogenetics,* 73**,** 405-417.

REES, H. C., MADDISON, B. C., MIDDLEDITCH, D. J., PATMORE, J. R. & GOUGH, K. C. 2014. The detection of aquatic animal species using environmental DNA-a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology,* 51**,** 1450-1459.

REID, A. J., CARLSON, A. K., CREED, I. F., ELIASON, E. J., GELL, P. A., JOHNSON, P. T., KIDD, K. A., MACCORMACK, T. J., OLDEN, J. D. & ORMEROD, S. J. 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews,* 94**,** 849-873.

RENAUD, G., HANGHØJ, K., KORNELIUSSEN, T. S., WILLERSLEV, E. & ORLANDO, L. 2019. Joint Estimates of Heterozygosity and Runs of Homozygosity for Modern and Ancient Samples. *Genetics,* 212**,** 587-614.

REVENGA, C., CAMPBELL, I., ABELL, R., DE VILLIERS, P. & BRYER, M. 2005. Prospects for monitoring freshwater ecosystems towards the 2010 targets. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 360**,** 397-413.

RIAZ, T., SHEHZAD, W., VIARI, A., POMPANON, F., TABERLET, P. & COISSAC, E. 2011. ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research,* 39**,** e145-e145.

ROBERTON, S., BELL, D., SMITH, G., NICHOLLS, J., CHAN, K., NGUYEN, D., TRAN, P., STREICHER, U., POON, L. & CHEN, H. 2006. Avian influenza H5N1 in viverrids: implications for wildlife health and conservation. *Proceedings of the Royal Society B: Biological Sciences,* 273**,** 1729-1732.

ROBINSON, J. T., THORVALDSDÓTTIR, H., WINCKLER, W., GUTTMAN, M., LANDER, E. S., GETZ, G. & MESIROV, J. P. 2011. Integrative genomics viewer. *Nature Biotechnology,* 29**,** 24-26.

ROCHETTE, N. C., RIVERA-COLÓN, A. G. & CATCHEN, J. M. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology,* 28**,** 4737-4754.

ROSENBERG, N. A. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes,* 4**,** 137-138.

RUMBALL, W., FRANKLIN, I., FRANKHAM, R. & SHELDON, B. 1994. Decline in heterozygosity under full-sib and double first-cousin inbreeding in Drosophila melanogaster. *Genetics,* 136**,** 1039-1049.

RUPPERT, K. M., KLINE, R. J. & RAHMAN, M. S. 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology & Conservation,* 17**,** e00547.

SAHU, A., KUMAR, N., SINGH, C. P. & SINGH, M. 2023. Environmental DNA (eDNA): Powerful technique for biodiversity conservation. *Journal for Nature Conservation,* 71**,** 126325.

SANTIAGO, E., NOVO, I., PARDIÑAS, A. F., SAURA, M., WANG, J. & CABALLERO, A. 2020. Recent Demographic History Inferred by High-Resolution Analysis of Linkage Disequilibrium. *Molecular Biology and Evolution,* 37**,** 3642-3653.

SATAM, H., JOSHI, K., MANGROLIA, U., WAGHOO, S., ZAIDI, G., RAWOOL, S., THAKARE, R. P., BANDAY, S., MISHRA, A. K. & DAS, G. 2023. Next-generation sequencing technology: current trends and advancements. *Biology,* 12**,** 997.

SAVAGE, A. E., MULETZ-WOLZ, C. R., CAMPBELL GRANT, E. H., FLEISCHER, R. C. & MULDER, K. P. 2019. Functional variation at an expressed MHC class IIβ locus associates with Ranavirus infection intensity in larval anuran populations. *Immunogenetics,* 71**,** 335-346.

SCHEELE, B. C., HOLLANDERS, M., HOFFMANN, E. P., NEWELL, D. A., LINDENMAYER, D. B., MCFADDEN, M., GILBERT, D. J. & GROGAN, L. F. 2021. Conservation translocations for amphibian species threatened by chytrid fungus: A review, conceptual framework, and recommendations. *Conservation Science & Practice,* 3**,** e524.

SCHMID, D. W., MEYER, M., WILHELM, K., TILLEY, T., LINK-HESSING, T., FLEISCHER, R., BADU, E. K., NKRUMAH, E. E., OPPONG, S. K., SCHWENSOW, N., TSCHAPKA, M., BALDWIN, H. J., VALLO, P., CORMAN, V. M., DROSTEN, C. & SOMMER, S. 2023. MHC class II genes mediate susceptibility and resistance to coronavirus infections in bats. *Molecular Ecology,* 32**,** 3989-4002.

SCHMIDT, D. J., BROCKETT, B., ESPINOZA, T., CONNELL, M. & HUGHES, J. M. 2016. Complete mitochondrial genome of the endangered Mary River turtle (*Elusor macrurus*) and low mtDNA variation across the species' range. *Australian Journal of Zoology,* 64**,** 117-121.

SCHMIDT, T. L., JASPER, M. E., WEEKS, A. R. & HOFFMANN, A. A. 2021. Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods in Ecology & Evolution,* 12**,** 1888-1898.

SCHMIDT, T. L., THIA, J. A. & HOFFMANN, A. A. 2024. How can genomics help or hinder wildlife conservation? *Annual Review of Animal Biosciences,* 12**,** 45-68.

SCHNEIDER, H. 2023. Integrating genomics and conservation to safeguard plant diversity. *Integrative Conservation,* 2**,** 10-18.

SCHOU, M. F., LOESCHCKE, V., BECHSGAARD, J., SCHLÖTTERER, C. & KRISTENSEN, T. N. 2017. Unexpected high genetic diversity in small populations suggests maintenance by associative overdominance. *Molecular Ecology,* 26**,** 6510-6523.

SCHWEIZER, R. M., SAARMAN, N., RAMSTAD, K. M., FORESTER, B. R., KELLEY, J. L., HAND, B. K., MALISON, R. L., ACKISS, A. S., WATSA, M., NELSON, T. C., BEJA-PEREIRA, A., WAPLES, R. S., FUNK, W. C. & LUIKART, G. 2021. Big Data in Conservation Genomics: Boosting Skills, Hedging Bets, and Staying Current in the Field. *Journal of Heredity,* 112**,** 313-327.

SENAPATI, D., BHATTACHARYA, M., KAR, A., CHINI, D. S., DAS, B. K. & PATRA, B. C. Environmental DNA (eDNA): A promising biological survey tool for aquatic species detection. Proceedings of the Zoological Society, 2019. Springer, 211-228.

SEVER, D. M. & HAMLETT, W. C. 2002. Female sperm storage in reptiles. *Journal of Experimental Zoology,* 292**,** 187-199.

SHAFER, A. B., PEART, C. R., TUSSO, S., MAAYAN, I., BRELSFORD, A., WHEAT, C. W. & WOLF, J. B. 2017. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology & Evolution,* 8**,** 907-917.

SHAFFER, H. B., MCCARTNEY-MELSTAD, E., NEAR, T. J., MOUNT, G. G. & SPINKS, P. Q. 2017. Phylogenomic analyses of 539 highly informative loci dates a fully resolved time tree for the major clades of living turtles (Testudines). *Molecular Phylogenetics & Evolution,* 115**,** 7-15.

SHORT, J. 2009. The characteristics and success of vertebrate translocations within Australia. Wildlife Research and Management Pty Ltd, Perth, & Australian Government Department of Agriculture, Fisheries and Forestry, Canberra.

SILVER, L. W., HOGG, C. J. & BELOV, K. 2024. Plethora of New Marsupial Genomes Informs Our Knowledge of Marsupial MHC Class II. *Genome Biology & Evolution,* 16.

SIM, R. R., NORTON, T. M., BRONSON, E., ALLENDER, M. C., STEDMAN, N., CHILDRESS, A. L. & WELLEHAN JR, J. F. 2015. Identification of a novel herpesvirus in captive Eastern box turtles (*Terrapene carolina carolina*). *Veterinary Microbiology,* 175**,** 218-223.

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**,** 1-7.

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**,** 3210-3212.

SMALL, E. 2012. The new Noah's Ark: beautiful and useful species only. Part 2. The chosen species. *Biodiversity,* 13**,** 37-53.

SMIT, A., HUBLEY, R. & GREEN, P. 2015. RepeatMasker Open-4.0. 2013–2015.

SMITH, D. A., LETHBRIDGE, M. R., ALLEN, B. L. & ANDREW, R. L. 2023. Inferring inter-colony movement within metapopulations of yellow-footed rock-wallabies using estimates of kinship. *Conservation Genetics,* 24**,** 265-278.

SMITH, K. F., ACEVEDO-WHITEHOUSE, K. & PEDERSEN, A. B. 2009. The role of infectious diseases in biological conservation. *Animal Conservation,* 12**,** 1-12.

SNYDER, N. F., DERRICKSON, S. R., BEISSINGER, S. R., WILEY, J. W., SMITH, T. B., TOONE, W. D. & MILLER, B. 1996. Limitations of captive breeding in endangered species recovery. *Conservation Biology,* 10**,** 338-348.

SOLOVYEV, V., KOSAREV, P., SELEDSOV, I. & VOROBYEV, D. 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biology,* 7**,** 1-12.

SOMMER, S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology,* 2**,** 1-18.

SPENCER, R.-J., GEORGES, A., LIM, D., WELSH, M. & REID, A. M. 2014. The risk of inter-specific competition in Australian short-necked turtles. *Ecological Research,* 29**,** 767-777.

SPENCER, R.-J., VAN DYKE, J., PETROV, K., FERRONATO, B., MCDOUGALL, F., AUSTIN, M., KEITEL, C. & GEORGES, A. 2018. Profiling a possible rapid extinction event in a long-lived species. *Biological Conservation,* 221**,** 190-197.

STEPHENS, M. & SCHEET, P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet,* 76**,** 449-62.

STEPHENS, M., SMITH, N. J. & DONNELLY, P. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet,* 68**,** 978-89.

STEVENTON, C., WICKER, L., DOBSON, E., LEGIONE, A., DEVLIN, J. & HARLEY, D. 2024. Management lessons from a long-term captive-breeding program for a critically endangered species, Leadbeater's possum, *Gymnobelideus leadbeateri*. *Conservation Science & Practice***,** e13215.

STEWART, K. A. 2019. Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. *Biodiversity & Conservation,* 28**,** 983-1001.

STOECKLE, B. C., BEGGEL, S., CERWENKA, A. F., MOTIVANS, E., KUEHN, R. & GEIST, J. 2017. A systematic approach to evaluate the influence of environmental conditions on eDNA detection success in aquatic ecosystems. *PLoS One,* 12**,** e0189119.

SUES, H.-D. 2019. *The rise of reptiles: 320 million years of evolution*, JHU Press.

SUMMERS, K., MCKEON, S., SELLARS, J., KEUSENKOTHEN, M., MORRIS, J., GLOECKNER, D., PRESSLEY, C., PRICE, B. & SNOW, H. 2003. Parasitic exploitation as an engine of diversity. *Biological Reviews,* 78**,** 639-675.

SUNDE, J., YILDIRIM, Y., TIBBLIN, P., BEKKEVOLD, D., SKOV, C., NORDAHL, O., LARSSON, P. & FORSMAN, A. 2022. Drivers of neutral and adaptive differentiation in pike (*Esox lucius*) populations from contrasting environments. *Molecular Ecology,* 31**,** 1093-1110.

TABERLET, P., COISSAC, E., POMPANON, F., BROCHMANN, C. & WILLERSLEV, E. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology,* 21**,** 2045-2050.

TAKAHASHI, M., SACCÒ, M., KESTEL, J. H., NESTER, G., CAMPBELL, M. A., VAN DER HEYDE, M., HEYDENRYCH, M. J., JUSZKIEWICZ, D. J., NEVILL, P. & DAWKINS, K. L. 2023. Aquatic environmental DNA: A review of the macro-organismal biomonitoring revolution. *Science of the Total Environment,* 873**,** 162322.

TAMURA, K., STECHER, G. & KUMAR, S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology & Evolution,* 38**,** 3022-3027.

TANG, S., PEEL, E., BELOV, K., HOGG, C. J. & FARQUHARSON, K. A. 2024. Multi-omics resources for the Australian southern stuttering frog (*Mixophyes australis*) reveal assorted antimicrobial peptides. *Scientific Reports,* 14**,** 3991.

TARONGA CONSERVATION SOCIETY AUSTRALIA 2023. Standard Operating Procedure – Bellinger River Turtle *Myuchelys georgesi*. Taronga Zoo, Sydney, NSW.

TAYLOR, H. R., DUSSEX, N. & VAN HEEZIK, Y. 2017. Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Global Ecology & Conservation,* 10**,** 231-242.

TEIXEIRA, J. C. & HUBER, C. D. 2021. The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences,* 118**,** e2015096118.

TENSEN, L. & FISCHER, K. 2024. Evaluating hybrid speciation and swamping in wild carnivores with a decision-tree approach. *Conservation Biology,* 38**,** e14197.

TESSLER, M., CUNNINGHAM, S. W., INGALA, M. R., WARRING, S. D. & BRUGLER, M. R. 2023. An environmental DNA primer for microbial and restoration ecology. *Microbial Ecology,* 85**,** 796-808.

THAVORNKANLAPACHAI, R., MILLS, H. R., OTTEWELL, K., FRIEND, J. A. & KENNINGTON, W. J. 2021. Temporal Variation in the Genetic Composition of an Endangered Marsupial Reflects Reintroduction History. *Diversity,* 13**,** 257.

THEISSINGER, K., FERNANDES, C., FORMENTI, G., BISTA, I., BERG, P. R., BLEIDORN, C., BOMBARELY, A., CROTTINI, A., GALLO, G. R. & GODOY, J. A. 2023. How genomics can help biodiversity conservation. *Trends in Genetics*.

THOMSON, R. C., SPINKS, P. Q. & SHAFFER, H. B. 2021. A global phylogeny of turtles reveals a burst of climate-associated diversification on continental margins. *Proceedings of the National Academy of Sciences,* 118**,** e2012215118.

TODD, B. D., JENKINSON, T. S., ESCALONA, M., BERAUT, E., NGUYEN, O., SAHASRABUDHE, R., SCOTT, P. A., TOFFELMIER, E., WANG, I. J. & SHAFFER, H. B. 2022. Reference genome of the northwestern pond turtle, *Actinemys marmorata*. *Journal of Heredity,* 113**,** 624-631.

TODD, B. D., WILLSON, J. D. & GIBBONS, J. W. 2010. The global status of reptiles and causes of their decline. *Ecotoxicology of Amphibians & Reptiles,* 47**,** 67.

TORKAMANEH, D., LAROCHE, J. & BELZILE, F. 2016. Genome-wide SNP calling from genotyping by sequencing (GBS) data: a comparison of seven pipelines and two sequencing technologies. *PloS One,* 11**,** e0161333.

TRUJILLO, A. L., HOFFMAN, E. A., BECKER, C. G. & SAVAGE, A. E. 2021. Spatiotemporal adaptive evolution of an MHC immune gene in a frog-fungus disease system. *Heredity,* 126**,** 640-655.

ULIANO-SILVA, M., NUNES, J. & KRASHENINNIKOVA, K. 2021. marcelauliano/MitoHiFi: mitohifi_v2. 0.

UNTERGASSER, A., CUTCUTACHE, I., KORESSAAR, T., YE, J., FAIRCLOTH, B. C., REMM, M. & ROZEN, S. G. 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Research,* 40**,** e115-e115.

VALDIVIA-CARRILLO, T., ROCHA-OLIVARES, A., REYES-BONILLA, H., DOMÍNGUEZ-CONTRERAS, J. F. & MUNGUIA-VEGA, A. 2021. Integrating eDNA metabarcoding and simultaneous underwater visual surveys to describe complex fish communities in a marine biodiversity hotspot. *Molecular Ecology Resources,* 21**,** 1558-1574.

VAN DYKE, J. U., FERRONATO, B. D. O. & SPENCER, R.-J. 2018. Current conservation status of Australian freshwater turtles. *Australian Journal of Zoology,* 66**,** 1-3.

VILLACORTA-RATH, C., ESPINOZA, T., COCKAYNE, B., SCHAFFER, J. & BURROWS, D. 2022. Environmental DNA analysis confirms extant populations of the cryptic Irwin's turtle within its historical range. *BMC Ecology & Evolution,* 22**,** 1-14.

WANG, J. 2011. COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources,* 11**,** 141-145.

WELDON, C., DU PREEZ, L. H., HYATT, A. D., MULLER, R. & SPEARE, R. 2004. Origin of the amphibian chytrid fungus. *Emerging Infectious Diseases,* 10**,** 2100.

WILCOX, T. M., MCKELVEY, K. S., YOUNG, M. K., SEPULVEDA, A. J., SHEPARD, B. B., JANE, S. F., WHITELEY, A. R., LOWE, W. H. & SCHWARTZ, M. K. 2016. Understanding environmental DNA detection probabilities: A case study using a stream-dwelling char *Salvelinus fontinalis*. *Biological Conservation,* 194**,** 209-216.

WILKINSON, D. A., MARSHALL, J. C., FRENCH, N. P. & HAYMAN, D. T. 2018. Habitat fragmentation, biodiversity loss and the risk of novel infectious disease emergence. *Journal of the Royal Society Interface,* 15**,** 20180403.

WILLIAMS, S. E. & HOFFMAN, E. A. 2009. Minimizing genetic adaptation in captive breeding programs: a review. *Biological Conservation,* 142**,** 2388-2400.

WISELY, S. M., MCDONALD, D. B. & BUSKIRK, S. W. 2003. Evaluation of the genetic management of the endangered black-footed ferret (*Mustela nigripes*). *Zoo Biology: Published in affiliation with the American Zoo & Aquarium Association,* 22**,** 287-298.

WOINARSKI, J. C., BURBIDGE, A. A. & HARRISON, P. L. 2015. Ongoing unraveling of a continental fauna: decline and extinction of Australian mammals since European settlement. *Proceedings of the National Academy of Sciences,* 112**,** 4531-4540.

WRIGHT, B., FARQUHARSON, K. A., MCLENNAN, E. A., BELOV, K., HOGG, C. J. & GRUEBER, C. E. 2019. From reference genomes to population genomics: comparing three reference-aligned reduced-representation sequencing pipelines in two wildlife species. *BMC Genomics,* 20.

WRIGHT, B., MORRIS, K., GRUEBER, C. E., WILLET, C. E., GOOLEY, R., HOGG, C. J., O'MEALLY, D., HAMEDE, R., JONES, M. & WADE, C. 2015. Development of a SNP-based assay for measuring genetic diversity in the Tasmanian devil insurance population. *BMC Genomics,* 16**,** 1-11.

WRIGHT, B. R., FARQUHARSON, K. A., MCLENNAN, E. A., BELOV, K., HOGG, C. J. & GRUEBER, C. E. 2020. A demonstration of conservation genomics for threatened species management. *Molecular Ecology Resources,* 20**,** 1526-1541.

WU, S., WANG, K., DOU, T., YUAN, S., WU, D.-D., GE, C., JIA, J. & SU, Z. 2024. High-quality genome assembly of a *C. crossoptilon* and related functional and genetics data resources. *Scientific Data,* 11**,** 247.

YANG, Z., ZHANG, Z., WANG, G., HUANG, Z. & YOU, W. 2024. Exploiting hybrid vigor to enhance survival and growth for whiteleg shrimp (*Litopenaeus vannamei*): A complete diallel cross experiment. *Aquaculture & Fisheries*.

YILDIRIM, Y., TINNERT, J. & FORSMAN, A. 2018. Contrasting patterns of neutral and functional genetic diversity in stable and disturbed environments. *Ecology & Evolution,* 8**,** 12073-12089.

ZARDOYA, R. & MEYER, A. 2001. The evolutionary position of turtles revised. *Naturwissenschaften,* 88**,** 193-200.

ZHANG, J., FINLAISON, D. S., FROST, M. J., GESTIER, S., GU, X., HALL, J., JENKINS, C., PARRISH, K., READ, A. J., SRIVASTAVA, M., ROSE, K. & KIRKLAND, P. D. 2018. Identification of a novel nidovirus as a potential cause of large scale mortalities in the endangered Bellinger River snapping turtle (*Myuchelys georgesi*). *PLoS One,* 13**,** e0205209.

ZHANG, J., KHAZALWA, E. M., ABKALLO, H. M., ZHOU, Y., NIE, X., RUAN, J., ZHAO, C., WANG, J., XU, J. & LI, X. 2021. The advancements, challenges, and future implications of the CRISPR/Cas9 system in swine research. *Journal of Genetics & Genomics,* 48**,** 347-360.

ZHANG, X., UNMACK, P. J., KUCHLING, G., WANG, Y. & GEORGES, A. 2017. Resolution of the enigmatic phylogenetic relationship of the critically endangered Western Swamp Tortoise *Pseudemydura umbrina* (Pleurodira: Chelidae) using a complete mitochondrial genome. *Molecular Phylogenetics & Evolution,* 115**,** 58-61.

ZHOU, C., MCCARTHY, S. A. & DURBIN, R. 2022. YaHS: yet another Hi-C scaffolding tool. *bioRxiv*.

ZHU, Y., WAN, Q.-H., ZHANG, H.-M. & FANG, S.-G. 2019. Reproductive strategy inferred from major histocompatibility complex-based inter-individual, sperm-egg, and mother-fetus recognitions in giant pandas (*Ailuropoda melanoleuca*). *Cells,* 8**,** 257.