

https://dspace.org/

Uncovering genetic changes underlying adaptation in southern African dwarf chameleons (Bradypodion)

Taft, Jody M.

2024
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/42166
Taft, Jody M.. (2024). Uncovering genetic changes underlying adaptation in southern African dwarf chameleons (Bradypodion) [PhD thesis, University of the Witwatersrand, Johannesburg]. WireDSpace.

Downloaded from a DSpace repository. https://dspace.org/

Uncovering genetic changes underlying adaptation in southern African dwarf chameleons (*Bradypodion*)

Jody M. Taft



Supervisor: Prof. Graham J. Alexander

University of the Witwatersrand

Co-supervisor: Prof. Krystal A. Tolley

South African National Biodiversity Institute

Co-supervisor: Prof. Anthony J. Geneva

Rutgers University-Camden

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa, in fulfilment of the requirements for the degree of Doctor of Philosophy

School of Animal, Plant and Environmental Sciences

March 2024

Declaration

I, Jody M. Taft (Student number: 2395479), declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg, South Africa. It has not been submitted before for any degree or examination at any other University. All the sources I have used or quoted have been indicated and acknowledged by complete reference.



Jody M. Taft

Date: 02 June 2024

Abstract

Natural selection acting on the available range of phenotypes in a population will favour alleles beneficial to an organism's fitness, resulting in adaptation. Characterizing patterns of adaptive genetic variation in wild populations is a fundamental first step toward understanding the potential for adaptation in response to selection. With recent advances in sequencing technology, the increased availability of high-quality genomic resources has made efforts to identify loci under selection across the genome more tractable. As a result, there is a growing body of research challenging the notion that adaption takes place over deep evolutionary timescales, with evidence of rapid adaptation occurring in a range of taxa. Examples of rapid adaptation are rife within reptiles, particularly in response to contemporary environmental changes such as urbanization. Despite this, insights into the genetic architecture of adaptation in reptiles remain limited as appropriate genomic resources are currently lacking.

The southern African dwarf chameleons, *Bradypodion*, are the most recently diverged lineage in Chamaeleonidae. These arboreal lizards are well known for their high capacity for adaptation, especially in response to environmental change. *Bradypodion* are known to have undergone rapid diversification linked to habitat specialization resulting in the emergence of few convergent phenotypes, or ecomorphs, across the genus. These lizards occupy a variety of habitat types, from closed-canopy forests to open-canopy habitats such as grasslands, but are also known to occur within urban environments. As ecomorphs of various *Bradypodion* species display phenotypic convergence in similar habitat types, populations may also converge phenotypically in urban habitats. However, the underlying genetic architecture of these phenotypes remains unknown for *Bradypodion*, further constrained given the absence of appropriate genomic resources for these lizards.

To facilitate insights into the genetic architecture underlying adaptation in *Bradypodion*, two *de novo* assembled whole genomes were produced using Pacific Biosciences long-read sequencing data. These assemblies are among the highest-quality squamate genomes published to date. In addition, coalescent analyses of these assemblies indicated that historical changes in effective population size correspond to notable shifts in the southern African environment. Furthermore, the high-quality annotations of both *Bradypodion* genome assemblies were generated and used to describe the genetic feature landscapes and evaluate the gene family evolution of *Bradypodion*. Findings indicated that expanded gene families

within this genus are likely the result of responses to changing environmental conditions, facilitating the diversification of *Bradypodion*. In addition, differences in gene family composition at the species level provide insights into underlying genetic pathways resulting in adaptive traits possibly promoting ecomorph divergence across the genus. Lastly, to detect loci under selection, population structure and genetic diversity were assessed in five species of *Bradypodion* known to have populations in natural and urban habitats using a pairwise comparative approach. While there is evidence of allele frequency differentiation between urban and natural populations, it is uncertain that this is due to selection pressures experienced by the urban population. While links between genes containing outlier loci and the phenotypic traits known to be associated with adaptation in urban habitats are made, it is recommended that these genes be used as candidates for targeted gene modification to evaluate the phenotypes generated by those modifications and contrasted with phenotypes present in urban populations.

This thesis provides insights into the mechanistic genomic basis of adaption in *Bradypodion*. The high-quality genomes and annotations produced here will hopefully serve as a resource for further assessments of genetic adaptation in chameleons and reptiles more broadly. Furthermore, these findings present a framework from which to formulate robust hypotheses to assess candidate genes presented here as likely underlying adaptive phenotypes in urban populations.

Acknowledgements

First and foremost, I would like to thank Prof. Krystal Tolley. Never would I have imagined that nearly a decade after attending your guest lecture at the University of the Western Cape in 2015, I would have experienced, explored and accomplished so much because of you. I will be forever grateful for your mentorship, guidance, and blasting *Stadium Arcadium* on a remote mountain pass in the middle of Karoo at 2 am. You have granted me the opportunity to develop my version of a herpetologist, encouraged me to share my thoughts and opinions, challenged me to use my voice in pursuit of new knowledge, and provided the platform for me to broaden my horizons, both professionally and personally. Thank you for all the support throughout these years.

To Prof. Graham Alexander, thank you for always reminding me to see the bigger picture. You've shown me that, even though things may not be clear from the outset, maintaining perspective on the "why" will always keep one true. I am appreciative of your dedication and the continued support you've provided throughout this process. Thank you for engaging in thoughtful discussions and challenging me to expand my thinking. You've shaped both my approach to science and the way in which I deploy temperature loggers in Afrotemperate forests.

Next, I thank my co-supervisor, Prof. Anthony Geneva. Your insight has been immeasurable throughout this project. You've introduced me to so much, in both a professional and personal capacity. Thank you for creating the opportunity to expand my capabilities as a scientist, for providing the scope to explore different ways to solve the same problem and guidance I needed to hone my thoughts and ideas. I am honoured to call you mentor, supervisor, and friend. Thank you for making the world much smaller, creating a home away from home, and always making me feel welcome.

I would like to acknowledge the funding support provided for this research. The work conducted in this study was jointly funded by the National Research Foundation of South Africa (grant number 136381) and the National Science Foundation (award numbers 1927194 and 2152059). Thank you to the National Research Foundation of South Africa for the extension support grant (PMDS22071138617). Samples for this study were collected under permits from the provinces of Eastern Cape, KwaZulu-Natal, Western Cape (018-CPM403-00001, 0092-CPM401-00006, AAA004-00107-0035, CPM-333-00002, CRO 3/19CR,CRO 35/15CR, CRO 36/15CR, CRO 32/20CR, CRo 33/20CR, FAUNA 110/2011, OP 4758/2010, OP 4596/2010, OP 2635, OP 1259/2014, KZN 1647/2009, MPB.5299,

MPB.5371, MPB.5544, MPB.5604, CN44-59-5795, RSH 24/2021, WRO 41/03WR; WRO 15/03WR), ethics clearances (South African National Biodiversity Institute: 003/2011, 001/2013, 001/2014, 0001/2015, University of Johannesburg: 2019-10-10, University of the Witwatersrand: 2019/10/56/B), subsamples exported to the USA under CITES permits (142667, 206199, 257511, 260568 and 260570) issued by the South African Department of Fisheries, Forestry and Environment.

To Prof. Jonathan Losos and Prof. Krystal Tolley, for facilitating, developing and procuring funding for the Dimensions of Biodiversity project. Through this project, I was given the opportunity to collaborate and interact with exceptional scientists globally. Special thanks to Prof. Bettine van Vuuren (University of Johannesburg), Prof. Graham Alexander, Dr Anthony Herrel (Muséum National d'Histoire Naturelle), Prof. Doug Menke (University of Georgia), and Prof. Anthony Geneva.

I would like to thank all the support staff involved in the project. You have been instrumental in managing the administrative-related processes. Special mentions to Mpho August (University of Johannesburg), Shilpa Parbhu (University of Johannesburg), Rowena Siebritz (SANBI), and Gail Van Aswegen (SANBI).

To one of my closest friends, Dr Jessica da Silva. You've played a pivotal role not only during this journey but also in my development as a research scientist. You've shown me how to remain curious, introspective, and forthright. Your unwavering belief in me, endless support and encouragement is immensely appreciated.

To my friend, colleague, former roommate, and travel companion, Devon Main. I will forever be grateful for our discussions, related to science and beyond. You've broadened my way of thinking, challenged my views on various topics, and supported so much of my outlandish ideas whether we were in a remote coastal forest or the middle of a foreign city 12,000 km away from home. It has been a privilege to have taken this journey with you.

To Cleo Falvey and Alyssa Vanerelli—the friends I regard as family. I am beyond appreciative of your friendship, support and valuable contributions you've made to my work. Thank you for being the support structure during my time in the US. You've made a considerable impact on my life and for this, and so much more, I am eternally grateful.

To the greatest team I've had the honour of conducting fieldwork with, Dr Melissa Petford, Devon Main, Azraa Ebrahim, Matthew Adair, Juan-Jacques Forgus, and Annie Basson. Thank you for the support, laughs, and all-around good times we've shared while searching for chameleons across South Africa. I will forever cherish the unforgettable memories we've made together, from the nudibranch to the midnight karaoke sessions belting out Ben E. King's *Stand by Me*.

I would like to make special mention to the following people for their contributions to this work. Thanks to Dr Kristin Winchell (NYU), for inspiring the approach used, particularly the code, in Chapter 4. For the use of their images, I thank Melissa Petford (habitat) and Tyrone Ping (*Bradypodion melanocephalum* portrait). To Gerrie Le Roux from Grootnek Guesthouse, in the Langkloof Valley, thank you so much for welcoming us into your home. I will forever remember your enthusiasm and interest in our work. I also remain eternally grateful for the cricket updates during our stay at Grootnek. Lastly, to Dr Ferozah Conrad (SANBI), thank you so much for the early morning wellness checks and words of motivation.

Thank you to Denise Foster, and the Foster family. I am immensely appreciative of the check-ins, consideration, and overwhelming support you have given me throughout this journey.

To my brother, Jade Taft. I'm grateful for the 2-hour phone calls in the middle of the night, whether I am home or in a different time zone. Thank you so much for the support.

To my parents, Joan and Deon Taft, your encouragement and support have been instrumental to me throughout these years. Thank you for always believing in me, reminding me why I started, and motivating me to keep going. I am grateful for the grace you've afforded me to pursue my passions.

Finally, I would like to thank Jodene Foster. You have been with me every step of the way, encouraging, supporting, and motivating me. In difficult times, you were there to keep me grounded. In moments of success, you celebrated with me but reminded me that we still have a ways to go. Thank you for your unwavering belief that I could achieve this. *We* did this together, and I am so grateful.

Thesis structure

This thesis is comprised of five chapters; a general introduction, three data chapters, and a concluding chapter. The three data chapters have been written as stand-alone manuscripts to facilitate the publication process. Consequently, this has resulted in some repetition in certain sections of the data chapters. Also, reference lists were provided at the end of each chapter because each chapter is treated as separate sections in this thesis. All in-text citations (including figure and table references) are clickable for easier navigation through the thesis. Below are the details of each chapter:

Chapter 1: General Introduction

This chapter serves as an introduction to the themes explored in this thesis. I detail mechanisms of divergence, outlining adaptive responses to contemporary habitat transformation. I present evidence for adaptive responses to transformed habitats examined in reptiles and explore genetic approaches to quantifying adaptive evolution. I also introduce my study group, *Bradypodion*, and provide the aims and objectives of this thesis.

Chapter 2: *De Novo* Whole Genome Assemblies for Two Southern African Dwarf Chameleons (*Bradypodion*, Chamaeleonidae)

In this chapter, I present two *de novo* assembled whole-genomes for *Bradypodion*. I also assess the quality of these assemblies relative to available vertebrate genomes, quantify the structural elements of the assemblies, compare synteny within *Bradypodion* and between *Anolis sagrei*, and estimate population size history using both *Bradypodion* assemblies. This chapter has subsequently been submitted to, and published in *Genomes, Biology and Evolution*. The formatting for this chapter follows the guidelines stipulated for *Genomes, Biology and Evolution*, therefore additional sections were included as per the requirements of the journal.

Taft JM, Tolley KA, Alexander GJ, Geneva AJ. 2023. De Novo Whole Genome Assemblies for Two Southern African Dwarf Chameleons (*Bradypodion*, Chamaeleonidae). *Genome Biology and Evolution. 15*(10):evad182.

Chapter 3: Exploring Gene Family Evolutionary Dynamics in *Bradypodion*

In this chapter, I generate the annotations for both *Bradypodion* genome assemblies. I report quality assessments of the annotations and estimate orthologous gene families across representative species in Iguania. I then evaluate changes in gene family composition for *Bradypodion* and explore the likelihood of gene family expansions promoting the diversification of *Bradypodion*.

Chapter 4: Detecting Signals of Selection within *Bradypodion*

In this chapter, I assess the population structure and genetic diversity of urban and natural populations of *Bradypodion*. In addition to traditional methods used to assess genetic structure and diversity, I use a pairwise comparison approach to evaluate allele frequency differences between populations. I also detect outlier loci at high frequencies and proceed to evaluate the traits underlain by genes containing outlier loci.

Chapter 5: Synthesis

Finally, I provide a thesis summary and contextualise the outcomes presented in each chapter. I give a broad context to the general findings, detail the limitations in my approaches used, and provide recommendations for using these results as a framework to develop testable hypotheses.

Contents

Declaration i

Abstract ii

Acknowledgements iv

Thesis structure vii

1 General Introduction **1**

1.1 Introduction 2

1.2 Facilitating Adaptation in Human-transformed Environments 3

1.3 Reptiles in Transformed Environments 4

1.4 Quantifying the extent of adaptive evolution using genetics 5

1.5 Study species: *Bradypodion* 6

1.6 Problem statement 12

1.7 Aims and Objectives 12

References 14

2 De Novo-Whole Genome Assemblies for Two Southern African Dwarf Chameleons (*Bradypodion*, Chamaeleonidae) **23**

Abstract 24

2.1 Introduction 25

2.2 Results and Discussion 26

2.2.1 Assembly contiguity and chromosome size 26

2.2.2 Assembly completeness and Phylogenetics 26

2.2.3 Hi-C and Inversion 28

2.2.4 Repetitive element content 30

2.2.5 Synteny analysis 30

2.2.6 Estimation of population size history 31

2.3 Methods . . . 32
2.3.1 Sampling and sequencing . . . 32
2.3.2 Assembly contiguity and chromosome size analysis . . . 32
2.3.3 Assembly completeness and Phylogenetics . . . 33
2.3.4 HiC chromatin conformation mapping . . . 33
2.3.5 Repetitive element content . . . 33
2.3.6 Synteny analysis . . . 34
2.3.7 Estimation of population size history . . . 34
2.4 Acknowledgements . . . 34
2.4.1 Data Availability . . . 35
References . . . 35
Supplementary Methods . . . 40
Supplementary References . . . 41
Supplementary Figures . . . 42
Supplementary Tables . . . 49

3 Exploring Gene Family Evolutionary Dynamics in *Bradypodion* 52
Abstract . . . 53
3.1 Introduction . . . 54
3.2 Methods . . . 57
3.2.1 Genome Annotation . . . 57
3.2.2 Genomic feature landscape . . . 58
3.2.3 Orthologous gene families . . . 58
3.2.4 Likelihood analysis of gene family gain and loss . . . 58
3.2.5 Functional enrichment analysis . . . 59
3.3 Results . . . 59
3.3.1 Genome annotation summary statistics . . . 59
3.3.2 Orthologous gene families . . . 61
3.3.3 Likelihood analysis of gene family gain and loss . . . 61
3.3.4 Functional enrichment analysis . . . 62
3.4 Discussion . . . 65
3.5 Future directions . . . 68
References . . . 69
Supplementary Figures . . . 81

4 Detecting signals of selection within *Bradypodion* 89

Abstract . . . 90

4.1 Introduction . . . 91

4.2 Methods and materials . . . 93

4.2.1 Sampling and tissue collection . . . 93

4.2.2 Sequencing . . . 94

4.2.3 Sequence-read processing . . . 95

4.2.4 Variant calling and filtering . . . 95

4.2.5 Species-level structure analysis . . . 96

4.2.6 Within-species population differentiation, divergence, and diversity . . . 97

4.2.7 Selection scan and functional term estimation analyses . . . 97

4.2.8 Isolation by Distance . . . 98

4.3 Results . . . 99

4.3.1 Species-level structure . . . 99

4.3.2 Within-species population differentiation, divergence, and diversity . . . 100

4.3.3 Selection scan and functional term estimation analyses . . . 103

4.3.4 Isolation by Distance . . . 105

4.4 Discussion . . . 105

4.5 Future Directions . . . 110

References . . . 111

Supplementary Figures . . . 125

Supplementary Tables . . . 140

5 Conclusion 163

5.1 Thesis summary . . . 163

5.2 Synthesis . . . 164

References . . . 168

Chapter 1

General Introduction

1.1 Introduction

Uncovering the mechanisms facilitating divergence within lineages has, for the most part, been addressed along geological timescales (Gillespie, 1991; Messer et al., 2016). In recent years, more attention has been paid to assessing lineage responses to environmental change during the contemporary period (Johnson and Munshi-South, 2017). Human-facilitated habitat change is recognised as a major threat to natural systems on a global scale. Growing evidence has also highlighted that habitat transformation has a direct negative effect on several species' survival (Templeton et al., 1990; McKinney, 2008; Jeltsch et al., 2011; Aronson et al., 2014; Akerman and Bürger, 2014). Human populations have also moved toward urban settlements rather than rural living, with an estimated 65% of the population projected to live in urban settings by 2050 (Kundu and Pandey, 2020). Further predictions indicate that urban expansion is expected to increase more rapidly in the coming years (Makido et al., 2020). To accommodate urban expansion, natural systems are unfortunately experiencing increased rates of transformation (Seto et al., 2010). As a result, transformed systems might result in strong and often novel selection pressures on native fauna and flora.

Selection pressures within transformed habitats have led to changes in animals' behavioural, physiological, and morphological responses (Johnson and Munshi-South, 2017; Szulkin et al., 2020). Urban dwelling fauna tends to be bolder in response to reduced natural predatory pressure and strong habituation to human presence (Ducatez et al., 2017; Lapiedra et al., 2017; Baxter-Gilbert et al., 2019; Huang et al., 2020). On the contrary, changes in behaviour can often involve shifts to reduce exposure to humans. Diurnal mammals have altered their activity times, becoming more nocturnal (Gaynor et al., 2018, but see Frey et al., 2020). In groups reliant on calls for communication, as in birds and frogs, urban populations have altered their sound frequencies to limit overlap with anthropogenic sounds (Halfwerk et al., 2019). Additionally, the thermal environment within transformed systems is also considerably different from natural areas (Forman, 2014; Winchell et al., 2016). Urban areas typically produce higher absolute temperatures relative to adjacent natural environments (coined as 'urban heat islands') and, as a result, there is evidence for several species developing physiological tolerances to higher temperatures in urban populations relative to their natural counterparts (Sears et al., 2007; Angilletta Jr et al., 2007; Campbell-Staton et al., 2020). Shifts in morphological traits are

[1]Current literature is predominantly saturated with works focused on adaptation in 'typical' urban systems (Szulkin et al., 2020). When referring to transformed habitats/environments, this will include all forms of human-facilitated habitat transformation (namely urbanisation, peri-urban habitats, transformation for agriculture and resource removal)

also commonly reported; with urban insects being smaller in body size (Merckx et al., 2018; Eggenberger et al., 2019), birds showing changes in beak depth and length relative to non-urban populations (Giraudeau et al., 2014; Miller et al., 2018), and the development of domesticated features in urban foxes (shorter snout, broader head: Parsons et al., 2020). Although numerous examples of urban adaptation do exist, many lineages are not able to adapt and are ultimately excluded (either by extinction or dispersal to remaining fragmented natural habitats: (Miles et al., 2021). In addition, rapid adaptation to urban systems has been demonstrated within several species, and it should be mentioned that phenotypic shifts are more likely a result of plasticity or phenotypic sorting across urban-rural-natural or urban-natural gradients (Brans et al., 2017; Evans et al., 2018; Caspi et al., 2022).

1.2 Facilitating Adaptation in Human-transformed Environments

Urbanisation is currently recognised as the most drastic form of habitat change resulting from human disturbance (McKinney, 2008; Luck and Smallbone, 2010; Seress and Liker, 2015). In these novel environments, organisms experience unnatural media that impose novel pressures that act as a filter on diversity capable of persisting within transformed habitats. Species that occur within transformed habitats often represent relic populations that predate the transformation, although now restricted to certain spaces, or have populations that have been established after habitat transformation (Miles et al., 2021). Within transformed systems, species distributions are often highly fragmented because of novel barriers to dispersal. As a result, these species experience reduced gene flow between populations, which further contributes to genetic differentiation (Jha, 2015; Johnson et al., 2015; Wilson et al., 2016). These differences have provided an opportunity to detect the effects of transformation on evolution by natural selection.

Since attention to evolution in response to urbanization is fairly novel (Johnson and Munshi-South, 2017, but see Cook and Saccheri, 2013), disentangling the factors driving adaptation in these environments is essential to clarifying adaptive and non-adaptive evolution in these contemporary environments. In urban environments that are considerably labile, consistent changes to biotic and abiotic environmental factors are more likely to promote adaptive responses given the shifting pressures on the fitness of urban-dwelling species (Miles et al., 2021). Where changes in urban environments alter physical landscapes,

non-adaptive responses are more likely since the strength of genetic drift and restricted gene flow reduces the evolutionary potential of persisting isolated populations (Beninde et al., 2016). At present, the most well-documented driving forces behind evolutionary responses to urbanization are urban heat islands (Diamond and Martin, 2020), pollution (Whitehead et al., 2017), and habitat fragmentation (Miles et al., 2019), with each underlying force having documented cases of parallel evolution in multiple cities (Reid et al., 2016; Johnson and Munshi-South, 2017; Diamond et al., 2018; Santangelo et al., 2020). However, these cases are mostly explored in species with short generation times and high reproductive capacity, ideal candidates to experience rapid evolution (Whitehead et al., 2017; Diamond et al., 2022).

1.3 Reptiles in Transformed Environments

In many fields of study, reptiles are often not considered candidate taxa in many systems and often take a backseat when larger paradigms are established (Card et al., 2023; but see Booth and Schuett, 2011; Booth et al., 2014; Booth and Schuett, 2016; Card et al., 2021). However, in the field of urban evolutionary ecology, reptiles are fast becoming model groups to study as populations are more resilient, particularly when they are able to use man-made structures effectively (Winchell et al., 2016, 2018; Campbell-Staton et al., 2020). In urban populations of the Puerto Rican crested anole, *Anolis cristatellus*, there is evidence of distinct morphological characteristics due to urban living relative to their natural counterparts. Populations of urban anoles have repeatedly evolved longer limbs and toe pads with more lamellae in cities of Puerto Rico (Winchell et al., 2018, 2023). Experiments show that these evolved traits increase the locomotor performance of urban lizards on flat, smooth, artificial surfaces commonly found in cities (Winchell et al., 2018). Together, these examples of adaptation to urban living demonstrate how traits may diverge in response to specific urban pressures. Such clear examples of reptiles responding to urban environments exist in very few urban areas (Beninde et al., 2016; Winchell et al., 2016; Littleford-Colquhoun et al., 2017; Winchell et al., 2023), and thus, this represents an important aspect missing from our understanding of evolution in urban environments (Rivkin et al., 2019).

While certain species are able to utilise man-made structures, those species unable to do so are mostly restricted to green spaces. Green spaces in cities are known to be heterogeneous in terms of habitat structure (Lepczyk et al., 2017; Callaghan et al., 2019). This is largely due to these spaces being dominated by non-native vegetation. However, this provides an

opportunity to host a greater diversity of species relative to spaces with impervious surfaces (Vergnes et al., 2013; Munshi-South and Richardson, 2020). While green spaces are known to facilitate movement, aiding in maintaining gene flow (Munshi-South, 2012), these spaces may also act as urban archipelagos as well (Faeth and Kane, 1978), facilitating genetic differentiation and diversification. In the case of the eastern water dragon, *Intellagama lesuerii*, a semi-aquatic arboreal lizard, there is evidence for both genetic and morphological differentiation between populations occurring in nearby isolated urban green spaces (Littleford-Colquhoun et al., 2017). Further investigation also revealed that selection likely led to phenotypic differences among populations in various city parks, with population divergence exceeding that of drift alone (Jackson et al., 2022). These differences between populations were estimated to occur within 175 years (roughly 35 water dragon generations), suggesting that species unable to use man-made structures effectively also undergo rapid adaptation while persisting in isolated urban green spaces.

1.4 Quantifying the extent of adaptive evolution using genetics

Comparing and contrasting populations within transformed and natural habitats enables the opportunity to test for adaptive responses in these environments, especially when adaptive solutions are repeated across taxa or geographic localities (Rivkin et al., 2019; Santangelo et al., 2020). Repeated occurrences of the same phenotypes or genotypes in response to similar selection pressures across populations highlights the importance of natural selection in adaptive evolution (Arendt and Reznick, 2008; Conte et al., 2012; Stuart et al., 2017; Langerhans, 2018). Being able to quantify the extent of these evolutionary responses provides insights into the mechanics of adaptation, with high levels of parallelism suggesting that successful adaptive outcomes can only come about through a select number of pathways (Gould et al., 2006; Losos, 2011). However, these adaptive parallels can occur at multiple levels, from phenotypes to amino acids and individual nucleotides (Reid et al., 2016; Campbell-Staton et al., 2020), and discovering the factors that facilitate or limit parallel responses at all levels is important for understanding how adaptation occurs in nature.

In recent years, breakthroughs in high-throughput, low-cost sequencing together with long-distance scaffolding libraries have brought about a wave in the number of high-quality genome assemblies (Geneva et al., 2022). These genomic resources have become a

fundamental tool used to study functional, comparative, and evolutionary genomics (Genome 10K Community of Scientists, 2009; Koepfli et al., 2015; Lewin et al., 2018; Blaxter et al., 2022; Ebenezer et al., 2022). A major advantage of these high-quality genomic resources is the ability to detect the effects of selection using genetic rather than phenotypic data (Allendorf et al., 2022). Furthermore, by taking a genome-wide approach, detecting signatures of selection and recent demographic events has become more accurate, especially when highlighting loci underlying adaptive traits (Winchell et al., 2023). For example, recent selective sweeps leave patterns of reduced genetic diversity in the genomic regions surrounding the selected loci. Methods to detect recent selection relying on these patterns ultimately depend on the quality and contiguity of the reference assembly as these signals are better detected in the context of completely assembled genomes (Hahn et al., 2007; Hahn, 2019; Gable et al., 2023). As more genomic assemblies become available the ability to detect the underpinnings of selection become ever more attainable, especially for non-model organisms, given the scope and motivation to generate these genomic resources (Card et al., 2023; Gable et al., 2023).

1.5 Study species: *Bradypodion*

Different lineages within the family Chameleonidae occurring in Africa are known to have responded differently to the reduction of forested habitats during the Quaternary period (Tolley et al., 2013). Since most chameleon species are typically associated with forested/woodland habitats, the reduction of forested habitats would have promoted vicariant events driving speciation. While certain groups do exhibit strong niche conservatism (Tolley et al., 2013), persisting in fragmented forest patches (e.g., *Kinyongia*), other groups have diversified through niche divergence (e.g., *Bradypodion*; Tolley et al., 2008; da Silva and Tolley, 2013; Tolley et al., 2013, 2022), and are able to colonise newly available habitats (i.e., grasslands) adjacent to remaining forest patches.

In *Bradypodion*, an African genus of dwarf chameleons, several lineages within the group are shown to have diversified in response to habitat shifts driven by global climate change (Tolley et al., 2008, 2022). Multiple convergence events have resulted in the emergence of a few morphological phenotypes, or ecomorphs, within the genus. These arboreal lizards typically inhabit trees, bushes, shrubs or grasses, depending on the species, however, they are also capable of traversing bare ground between vegetation. *Bradypodion* are also known to occur

within urban transformed habitats, usually inhabiting green spaces such as parks, gardens and roadside vegetation. Of the 20 described species of *Bradypodion*, roughly a third occur in both urban and natural habitats, five of which were considered in this thesis: *B. damaranum*, *B. melanocephalum*, *B. setaroi*, *B. thamnobates*, and *B. ventrale*.

Bradypodion damaranum

The Knysna Dwarf Chameleon, *Bradypodion damaranum* (Boulenger 1887), is one of the larger dwarf chameleons, brightly coloured (Fig. 1.1 A), and typically occurs in Afrotemperate forests (Fig. 1.1 C). In forests, *B. damaranum* is found using vegetation ranging from low understory to high canopy (Tolley and Burger, 2007) but is also fairly common in urban spaces (Fig. 1.1 D), especially gardens with dense, leafy cover.

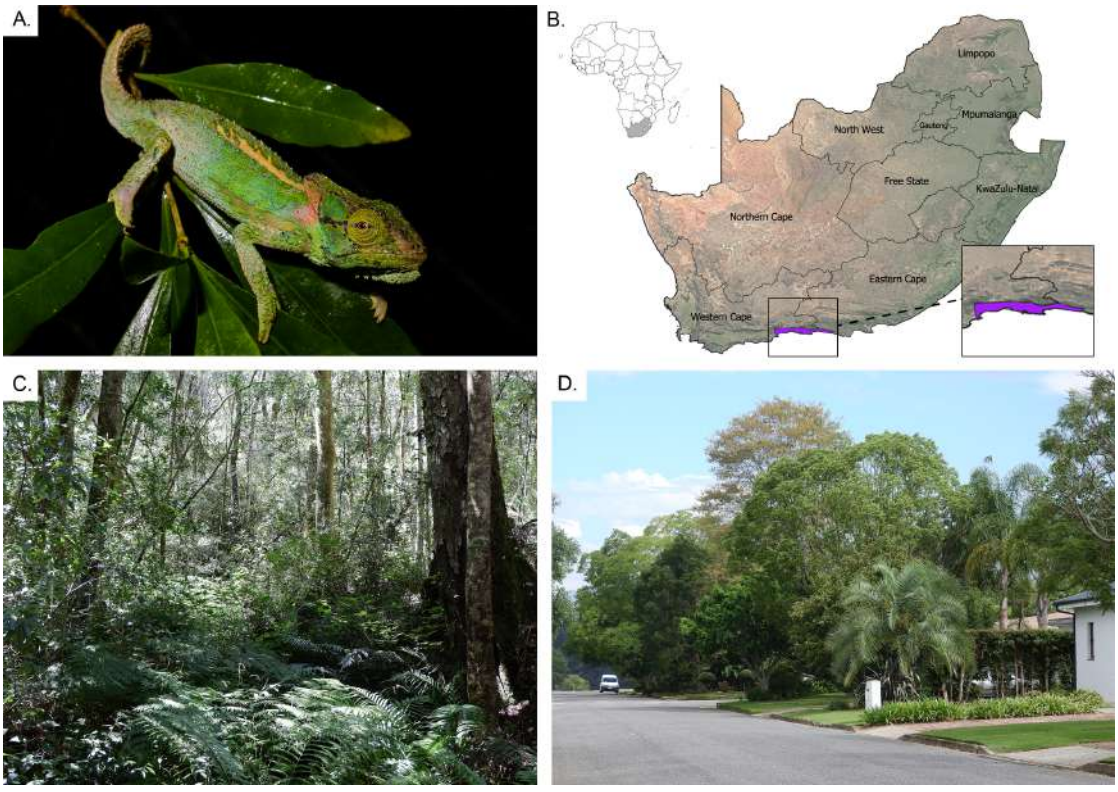


Figure 1.1: A) An adult *Bradypodion damaranum* from George, Western Cape. B) A map depicting the natural range (purple) of *B. damaranum* in South Africa, inset: Africa, with South Africa in grey. Images for C) natural and D) urban habitat, credit: M Petford.

Bradypodion melanocephalum

The KwaZulu Dwarf Chameleon, *Bradypodion melanocephalum* (Gray 1865), is a smaller, dull-coloured and inornate species (Fig. 1.2 A), with a total length of no more than 11 cm

(Tolley and Burger, 2007). *Bradypodion melanocephalum* is typically associated with open-canopy habitats (Fig. 1.2 C; Tolley and Burger, 2007), known to occupy a range of vegetation types, including high grasses, bushes, riparian thicket as well as urban habitats (Fig. 1.2 D).

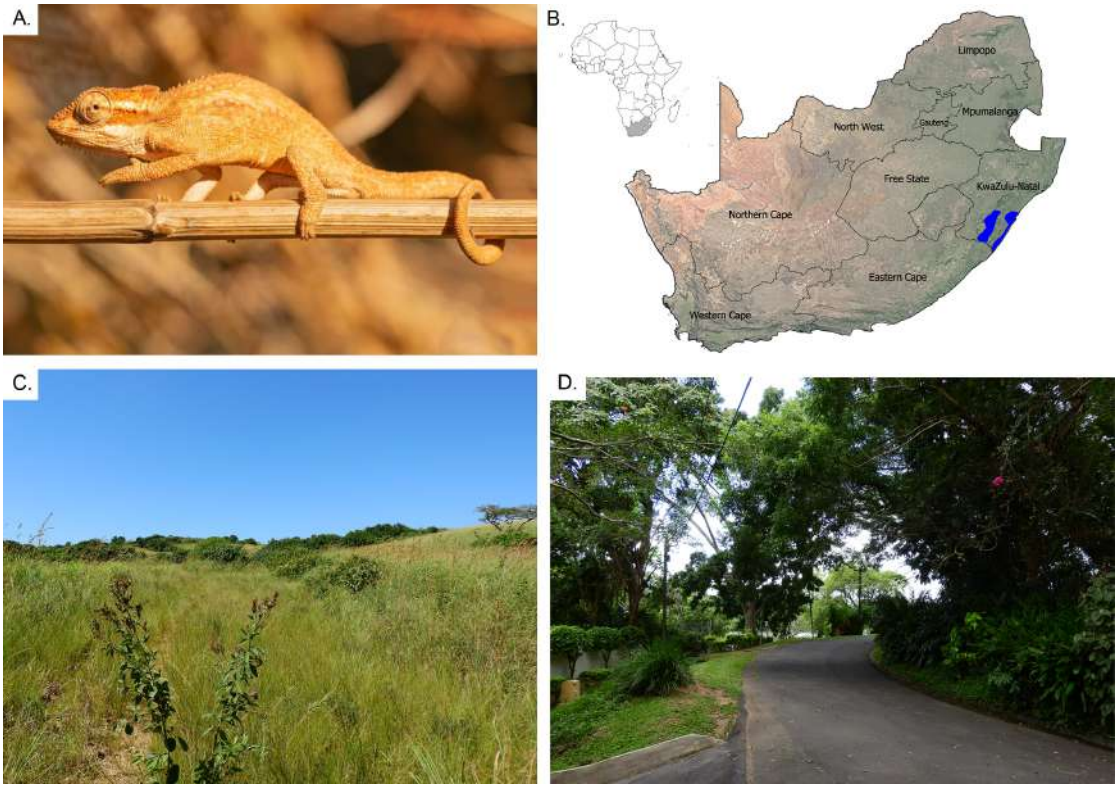


Figure 1.2: A) An adult *Bradypodion melanocephalum* from Durban, KwaZulu-Natal. credit: T Ping. B) A map depicting the natural range (blue) of *B. melanocephalum* in South Africa, inset: Africa, with South Africa in grey. Images for C) natural and D) urban habitat, credit: M Petford.

Bradypodion setaroi

The Setaro's Dwarf Chameleon, *Bradypodion setaroi* (Raw 1976), is a smaller species of *Bradypodion* not easily confused with others due to the distinct shape of the casque, pointed snout and lack of dorsal spines (Fig. 1.3 A; Tolley and Burger, 2007). *Bradypodion setaroi* naturally occurs in coastal dune forests (Fig. 1.3 C; Tolley and Burger, 2007), but are also regularly sighted in urban environments(Fig. 1.3 D, Tolley pers. comm.).

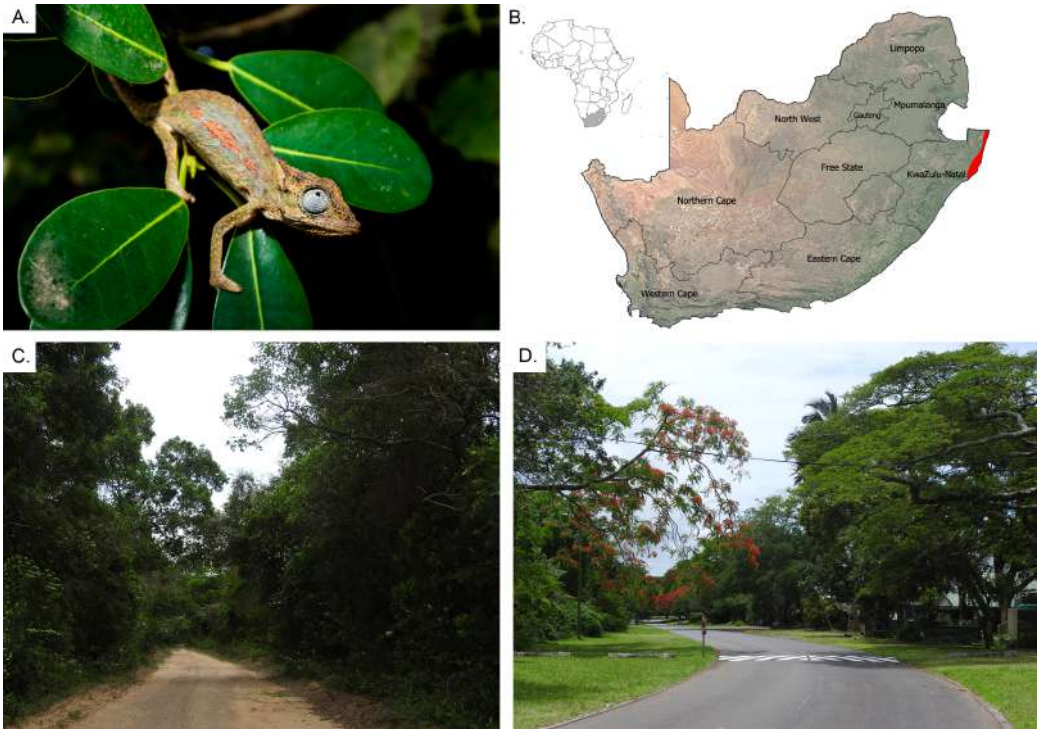


Figure 1.3: A) An adult *Bradypodion setaroi* from St. Lucia, KwaZulu-Natal. B) A map depicting the natural range (red) of *B. setaroi* in South Africa, inset: Africa, with South Africa in grey. Images for C) natural and D) urban habitat, credit: M Petford.

Bradypodion thamnobates

The Natal Midlands Dwarf Chameleon, *Bradypodion thamnobates* (Raw 1976), is the sister species to *B. melanocephalum* but is a considerably larger, more ornate and brightly coloured species (Fig. 1.4 A). *Bradypodion thamnobates* is typically associated with closed-canopy indigenous forests but also occupy urban gardens (Figs. 1.4 C-D, da Silva and Tolley, 2013).

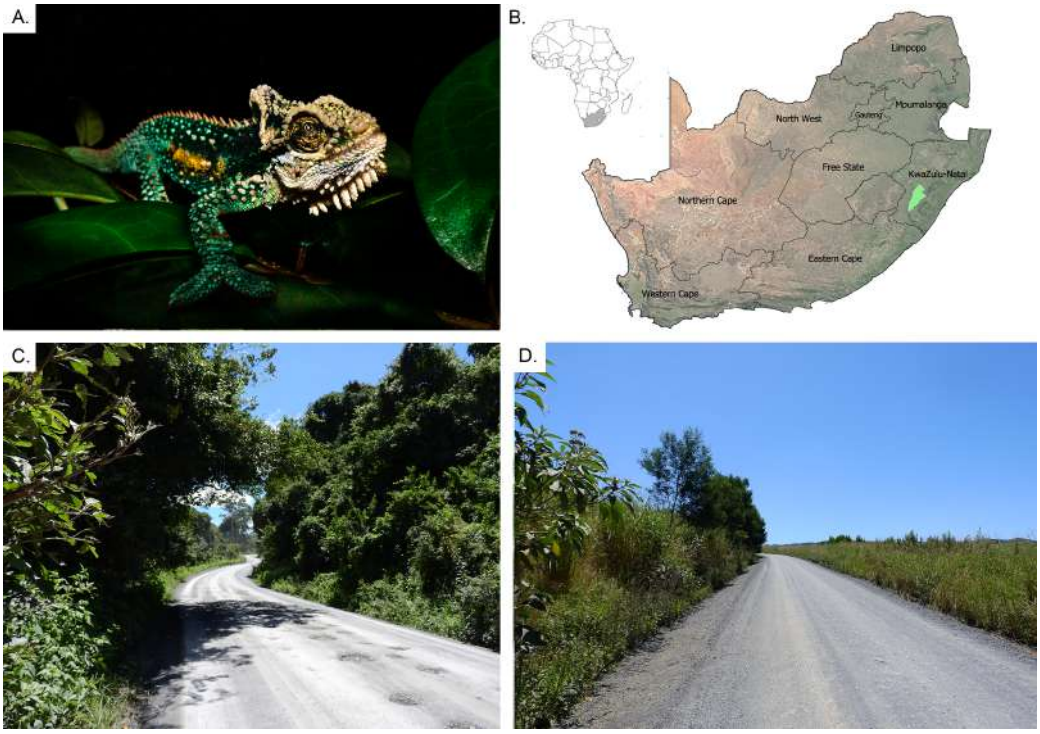


Figure 1.4: A) An adult *Bradypodion thamnobates* from the Natal Midlands, KwaZulu-Natal. B) A map depicting the natural range (green) of *B. thamnobates* in South Africa, inset: Africa, with South Africa in grey. Images for C) natural and D) transformed habitat, credit: M Petford.

Bradypodion ventrale

The Southern Dwarf Chameleon, *Bradypodion ventrale* (Gray 1845), are relatively dull in colouration, with shades of grey with olive mottling, and are one of the larger species (Tolley and Burger, 2007). *Bradypodion ventrale* has the largest distribution of any *Bradypodion* (Fig. 1.5 B). These lizards typically occur in more open canopy shrubby habitats (Fig. 1.5 C), but are also known to occur in urban environments (Fig. 1.5 D, Tolley and Burger, 2007; Herrel et al., 2024).

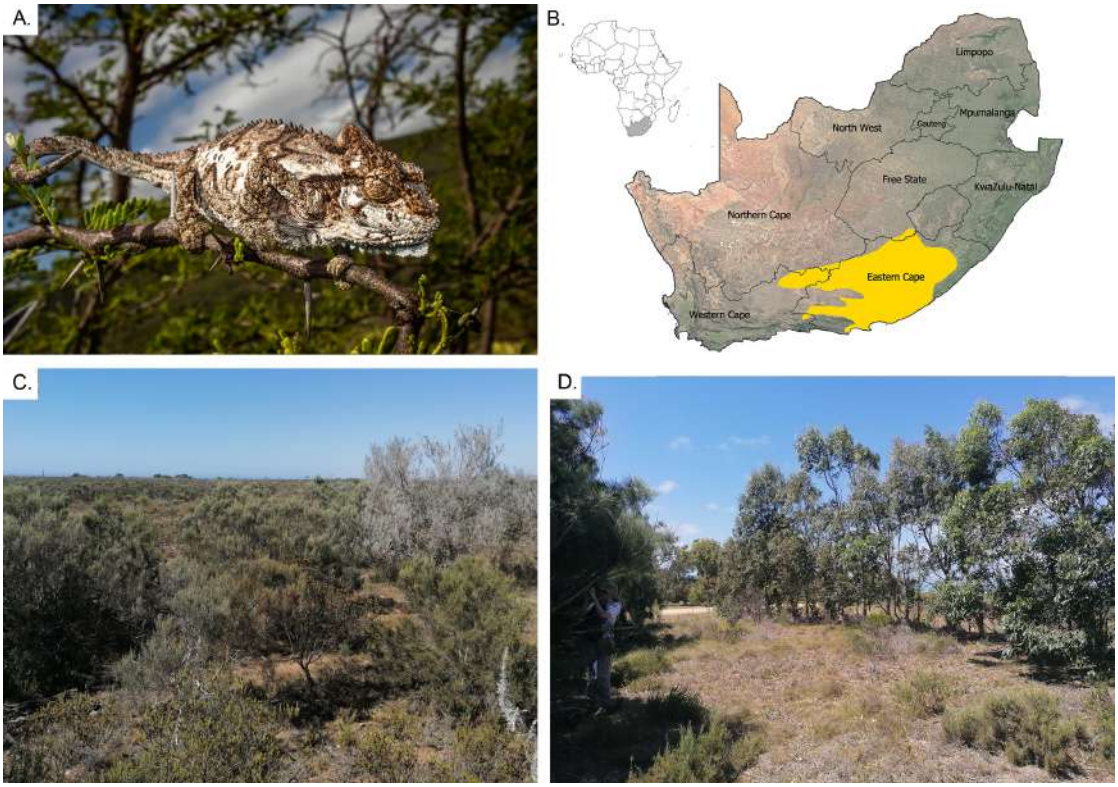


Figure 1.5: A) An adult *Bradypodion ventrale* from Jeffrey's Bay, Eastern Cape. B) A map depicting the natural range (gold) of *B. ventrale* in South Africa, inset: Africa, with South Africa in grey. Images for C) natural and D) urban habitat, credit: JM da Silva.

1.6 Problem statement

In order to reveal the genetic architecture underlying adaptive traits, particularly for complex, polygenic traits, integrating genomic and phenotypic data are essential (Gable et al., 2023). This approach demands significant resources, and as a result has been limited to reduced representation genomic data (Margres et al., 2017; Card et al., 2019; Winchell et al., 2023), or feasible in species with available genome and transcriptome data (Ollonen et al., 2018). Only with access to whole genome assemblies is identifying specific genes underlying traits of interest possible (Gable et al., 2023). However, due to finite resources, limited access to appropriate sequencing technologies, and lack of bioinformatic expertise, efforts to produce high-quality genome assemblies for African taxa are falling behind (Helmy et al., 2016; Ebenezer et al., 2022). In addition, available whole-genome assemblies for vertebrates are biased toward more charismatic groups. Even though reptiles are arguably one of the most evolutionary diverse taxonomic groups, reptile genomes are particularly underrepresented (Card et al., 2023), especially for acrodont lizards which include the chameleons (Gable et al., 2023).

Urbanisation and habitat transformation are the leading causes of the reduction and destruction of natural environments (United Nations, 2018). A consequence of this rapid change is the reduction of species diversity as novel pressures remove certain species from inhabiting modified environments (Johnson and Munshi-South, 2017; Winchell et al., 2018). Species that can persist in human-altered habitats tend to exhibit several common characteristics: habitat generalists and broad diet breadth, high vagility, high reproductive output and small body size (Hufbauer et al., 2012; Winchell et al., 2018; but see Croci et al., 2008). As *Bradypodion* ecomorphs display phenotypic convergence in similar natural habitat types, chameleon populations may also converge morphologically in urban habitats mirroring the patterns seen in other species (Littleford-Colquhoun et al., 2017; Winchell et al., 2018, 2023).

1.7 Aims and Objectives

The southern African dwarf chameleons (*Bradypodion*) are a highly adaptive group of lizards, known to have undergone rapid diversification linked to habitat specialization (Hopkins and Tolley, 2011; Herrel et al., 2013). As a result, there was an emergence of convergent phenotypes evident throughout the phylogeny. However, the underlying genetic architecture of these phenotypes remains unknown for *Bradypodion*. Since the appropriate genomic

resources are not currently available for these lizards, many evolutionary questions cannot be answered. The primary aim of this thesis was to uncover the genetic architecture underlying adaptation in *Bradypodion*. First, two genome assemblies for *Bradypodion* were assembled *de novo* using PacBio long-read sequence data. Second, both genomes were annotated to identify functional elements along the sequence of each genome. These annotations were then used to detect orthologous sequences needed to evaluate the changes in gene family composition within *Bradypodion*. Next, population structure and allele frequency differences were assessed for five species of *Bradypodion* with urban and natural populations. Genome scans were carried out to identify outlier loci along candidate genes, likely under selection. Overall, this thesis provides insights into how, with the appropriate genomic resources, the architecture and evolutionary underpinnings of adaptive traits are able to be assessed at two different evolutionary timescales—along longer evolutionary time (gene family evolution) and rapid adaptive responses to contemporary habitat change (urban environments). Specific aims, summarised hypotheses, and objectives are detailed below:

Produce de novo assembled genomes for *Bradypodion* (Chapter 2)

1. Assemble genomes *de novo* for *Bradypodion*, *B. pumilum* and *B. ventrale*, using long-read sequencing.

2. Evaluate the quality of the *Bradypodion* genome assemblies relative to other publicly available vertebrate genomes.

3. Compare and contrast the genomic structural elements between both *Bradypodion* assemblies.

Assess the gene family evolutionary dynamics in *Bradypodion* (Chapter 3)

Hypothesis: I tested the hypothesis that if there are major differences in the gene family composition within *Bradypodion*, it is likely associated with changes in environmental conditions experienced by this group.

1. Provide annotations for both de novo assembled genomes of *Bradypodion*.

2. Evaluate gene family expansion and contractions in *Bradypodion* relative to other members of the superfamily of Iguania (Chamaeleonidae, Agamidae, Iguanidae).

Detect sites of selection facilitating contemporary adaptation in *Bradypodion* (Chapter 4)

Hypothesis: I test the hypothesis that differences in allele frequencies between urban and natural populations of *Bradypodion* are recovered in genes underlying traits better suited to urban environments.

1. I compared and contrasted the genetic variation and diversity of urban and natural *Bradypodion* populations.

2. Conducted selection scans to identify outlier loci.

3. Evaluated gene function for genes containing outlier loci.

References

Akerman, A. and Bürger, R. (2014). The consequences of dominance and gene flow for local adaptation and differentiation at two linked loci. *Theoretical Population Biology*, 94:42–62.

Allendorf, F. W., Funk, W. C., Aitken, S. N., Byrne, M., and Luikart, G. H. (2022). *Conservation and the Genomics of Populations*. Oxford University Press.

Angilletta Jr, M. J., Wilson, R. S., Niehaus, A. C., Sears, M. W., Navas, C. A., and Ribeiro, P. L. (2007). Urban physiology: city ants possess high heat tolerance. *PLoS One*, 2(2):e258.

Arendt, J. and Reznick, D. (2008). Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in ecology & evolution*, 23(1):26–32.

Aronson, M. F., La Sorte, F. A., Nilon, C. H., Katti, M., Goddard, M. A., Lepczyk, C. A., Warren, P. S., Williams, N. S., Cilliers, S., Clarkson, B., et al. (2014). A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proceedings of the royal society B: biological sciences*, 281(1780):20133330.

Baxter-Gilbert, J., Riley, J. L., and Whiting, M. J. (2019). Bold new world: urbanization promotes an innate behavioral trait in a lizard. *Behavioral Ecology and Sociobiology*, 73:1–10.

Beninde, J., Feldmeier, S., Werner, M., Peroverde, D., Schulte, U., Hochkirch, A., and Veith, M. (2016). Cityscape genetics: structural vs. functional connectivity of an urban lizard population. *Molecular ecology*, 25(20):4984–5000.

Blaxter, M., Archibald, J. M., Childers, A. K., Coddington, J. A., Crandall, K. A., Di Palma, F., Durbin, R., Edwards, S. V., Graves, J. A., Hackett, K. J., et al. (2022). Why sequence all eukaryotes? *Proceedings of the National Academy of Sciences*, 119(4):e2115636118.

Booth, W. and Schuett, G. W. (2011). Molecular genetic evidence for alternative reproductive strategies in north american pitvipers (serpentes: Viperidae): long-term sperm storage and facultative parthenogenesis. *Biological Journal of the Linnean Society*, 104(4):934–942.

Booth, W. and Schuett, G. W. (2016). The emerging phylogenetic pattern of parthenogenesis in snakes. *Biological Journal of the Linnean Society*, 118(2):172–186.

Booth, W., Schuett, G. W., Ridgway, A., Buxton, D. W., Castoe, T. A., Bastone, G., Bennett, C., and McMahan, W. (2014). New insights on facultative parthenogenesis in pythons. *Biological Journal of the Linnean Society*, 112(3):461–468.

Brans, K. I., Govaert, L., Engelen, J. M., Gianuca, A. T., Souffreau, C., and De Meester, L. (2017). Eco-evolutionary dynamics in urbanized landscapes: evolution, species sorting and the change in zooplankton body size along urbanization gradients. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1712):20160030.

Callaghan, C. T., Bino, G., Major, R. E., Martin, J. M., Lyons, M. B., and Kingsford, R. T. (2019). Heterogeneous urban green areas are bird diversity hotspots: insights using continental-scale citizen science data. *Landscape Ecology*, 34:1231–1246.

Campbell-Staton, S. C., Winchell, K. M., Rochette, N. C., Fredette, J., Maayan, I., Schweizer, R. M., and Catchen, J. (2020). Parallel selection on thermal physiology facilitates repeated adaptation of city lizards to urban heat islands. *Nature Ecology & Evolution*, 4(4):652–658.

Card, D. C., Adams, R. H., Schield, D. R., Perry, B. W., Corbin, A. B., Pasquesi, G. I., Row, K., Van Kleeck, M. J., Daza, J. M., Booth, W., et al. (2019). Genomic basis of convergent island phenotypes in boa constrictors. *Genome biology and evolution*, 11(11):3123–3143.

Card, D. C., Jennings, W. B., and Edwards, S. V. (2023). Genome evolution and the future of phylogenomics of non-avian reptiles. *Animals*, 13(3):471.

Card, D. C., Vonk, F. J., Smalbrugge, S., Casewell, N. R., Wüster, W., Castoe, T. A., Schuett, G. W., and Booth, W. (2021). Genome-wide data implicate terminal fusion automixis in king cobra facultative parthenogenesis. *Scientific Reports*, 11(1):7271.

Caspi, T., Johnson, J. R., Lambert, M. R., Schell, C. J., and Sih, A. (2022). Behavioral plasticity can facilitate evolution in urban environments. *Trends in Ecology & Evolution*, 37(12):1092–1103.

Conte, G. L., Arnegard, M. E., Peichel, C. L., and Schluter, D. (2012). The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B: Biological Sciences*, 279(1749):5039–5047.

Cook, L. and Saccheri, I. J. (2013). The peppered moth and industrial melanism: evolution of a natural selection case study. *Heredity*, 110(3):207–212.

Croci, S., Butet, A., and Clergeau, P. (2008). Does urbanization filter birds on the basis of their biological traits. *The Condor*, 110(2):223–240.

da Silva, J. M. and Tolley, K. A. (2013). Ecomorphological variation and sexual dimorphism in a recent radiation of dwarf chameleons (*Bradypodion*). *Biological Journal of the Linnean Society*, 109(1):113–130.

Diamond, S. E., Chick, L. D., Perez, A., Strickler, S. A., and Martin, R. A. (2018). Evolution of thermal tolerance and its fitness consequences: parallel and non-parallel responses to urban heat islands across three cities. *Proceedings of the Royal Society B: Biological Sciences*, 285(1882):20180036.

Diamond, S. E. and Martin, R. A. (2020). Evolutionary consequences of the urban heat island. In Szulkin, M., Munshi-South, J., and Charmantier, A., editors, *Urban evolutionary biology*. Oxford University Press, USA.

Diamond, S. E., Martin, R. A., Bellino, G., Crown, K. N., and Prileson, E. G. (2022). Urban evolution of thermal physiology in a range-expanding, mycophagous fruit fly, drosophila tripunctata. *Biological Journal of the Linnean Society*, 137(3):409–420.

Ducatez, S., Audet, J.-N., Rodriguez, J. R., Kayello, L., and Lefebvre, L. (2017). Innovativeness and the effects of urbanization on risk-taking behaviors in wild barbados birds. *Animal Cognition*, 20:33–42.

Ebenezer, T. E., Muigai, A. W., Nouala, S., Badaoui, B., Blaxter, M., Buddie, A. G., Jarvis, E. D., Korlach, J., Kuja, J. O., Lewin, H. A., et al. (2022). Africa: sequence 100,000 species to safeguard biodiversity. *Nature*, 603(7901):388–392.

Eggenberger, H., Frey, D., Pellissier, L., Ghazoul, J., Fontana, S., and Moretti, M. (2019). Urban bumblebees are smaller and more phenotypically diverse than their rural counterparts. *Journal of Animal Ecology*, 88(10):1522–1533.

Evans, B. S., Reitsma, R., Hurlbert, A. H., and Marra, P. P. (2018). Environmental filtering of avian communities along a rural-to-urban gradient in Greater Washington, DC, USA. *Ecosphere*, 9(11):e02402.

Faeth, S. H. and Kane, T. C. (1978). Urban biogeography: city parks as islands for Diptera and Coleoptera. *Oecologia*, 32:127–133.

Forman, R. T. (2014). *Urban ecology: science of cities*. Cambridge University Press.

Frey, S., Volpe, J., Heim, N., Paczkowski, J., and Fisher, J. (2020). Move to nocturnality not a universal trend in carnivore species on disturbed landscapes. *Oikos*, 129(8):1128–1140.

Gable, S. M., Mendez, J. M., Bushroe, N. A., Wilson, A., Byars, M. I., and Tollis, M. (2023). The State of Squamate Genomics: Past, Present, and Future of Genome Research in the Most Speciose Terrestrial Vertebrate Order. *Genes*, 14(7):1387.

Gaynor, K. M., Hojnowski, C. E., Carter, N. H., and Brashares, J. S. (2018). The influence of human disturbance on wildlife nocturnality. *Science*, 360(6394):1232–1235.

Geneva, A. J., Park, S., Bock, D. G., de Mello, P. L., Sarigol, F., Tollis, M., Donihue, C. M., Reynolds, R. G., Feiner, N., Rasys, A. M., et al. (2022). Chromosome-scale genome assembly of the brown anole (*Anolis sagrei*), an emerging model species. *Communications Biology*, 5(1):1126.

Genome 10K Community of Scientists (2009). Genome 10k: a proposal to obtain whole-genome sequence for 10 000 vertebrate species. *Journal of Heredity*, 100(6):659–674.

Gillespie, J. H. (1991). *The causes of molecular evolution*, volume 2. Oxford University Press, USA.

Giraudeau, M., Nolan, P. M., Black, C. E., Earl, S. R., Hasegawa, M., and McGraw, K. J. (2014). Song characteristics track bill morphology along a gradient of urbanization in house finches (*Haemorhous mexicanus*). *Frontiers in Zoology*, 11:1–8.

Gould, J., Getz, G., Monti, S., Reich, M., and Mesirov, J. P. (2006). Comparative gene marker selection suite. *Bioinformatics*, 22(15):1924–1925.

Hahn, M. H. (2019). *Molecular Population Genetics*. Oxford University Press, USA.

Hahn, M. W., Han, M. V., and Han, S.-G. (2007). Gene Family Evolution across 12 *Drosophila* Genomes. *PLoS genetics*, 3(11):e197.

Halfwerk, W., Blaas, M., Kramer, L., Hijner, N., Trillo, P. A., Bernal, X. E., Page, R. A., Goutte, S., Ryan, M. J., and Ellers, J. (2019). Adaptive changes in sexual signalling in response to urbanization. *Nature ecology & evolution*, 3(3):374–380.

Helmy, M., Awad, M., and Mosa, K. A. (2016). Limited resources of genome sequencing in developing countries: challenges and solutions. *Applied & translational genomics*, 9:15–19.

Herrel, A., Taft, J. M., Main, D. C., da Silva, J. M., and Tolley, K. A. (2024). Branching out: perch diameter and orientation affect pull strength in chameleons (genus: *Bradypodion*). *Biological Journal of the Linnean Society*, page blae003.

Herrel, A., Tolley, K. A., Measey, G. J., da Silva, J. M., Potgieter, D. F., Boller, E., Boistel, R., and Vanhooydonck, B. (2013). Slow but tenacious: an analysis of running and gripping performance in chameleons. *Journal of Experimental Biology*, 216(6):1025–1030.

Hopkins, K. P. and Tolley, K. A. (2011). Morphological variation in the Cape Dwarf Chameleon (*Bradypodion pumilum*) as a consequence of spatially explicit habitat structure differences. *Biological Journal of the Linnean Society*, 102(4):878–888.

Huang, P., St. Mary, C. M., and Kimball, R. T. (2020). Habitat urbanization and stress response are primary predictors of personality variation in northern cardinals (*Cardinalis cardinalis*). *Journal of Urban Ecology*, 6(1):juaa015.

Hufbauer, R. A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C. E., Rey, O., and Estoup, A. (2012). Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary applications*, 5(1):89–101.

Jackson, N., Littleford-Colquhoun, B. L., Strickland, K., Class, B., and Frere, C. H. (2022). Selection in the city: Rapid and fine-scale evolution of urban eastern water dragons. *Evolution*, 76(10):2302–2314.

Jeltsch, F., Moloney, K. A., Schwager, M., Körner, K., and Blaum, N. (2011). Consequences of

correlations between habitat modifications and negative impact of climate change for regional species survival. *Agriculture, ecosystems & environment*, 145(1):49–58.

Jha, S. (2015). Contemporary human-altered landscapes and oceanic barriers reduce bumble bee gene flow. *Molecular Ecology*, 24(5):993–1006.

Johnson, M. T. and Munshi-South, J. (2017). Evolution of life in urban environments. *Science*, 358(6363):eaam8327.

Johnson, M. T., Thompson, K. A., and Saini, H. S. (2015). Plant evolution in the urban jungle. *American Journal of Botany*, 102(12):1951–1953.

Koepfli, K.-P., Paten, B., of Scientists, G. K. C., and O'Brien, S. J. (2015). The Genome 10K Project: a way forward. *Annual Review of Animal Biosciences*, 3(1):57–111.

Kundu, D. and Pandey, A. K. (2020). World urbanisation: trends and patterns. *Developing national urban policies: Ways forward to green and smart cities*, pages 13–49.

Langerhans, R. B. (2018). Predictability and parallelism of multitrait adaptation. *Journal of Heredity*, 109(1):59–70.

Lapiedra, O., Chejanovski, Z., and Kolbe, J. J. (2017). Urbanization and biological invasion shape animal personalities. *Global Change Biology*, 23(2):592–603.

Lepczyk, C. A., Aronson, M. F., Evans, K. L., Goddard, M. A., Lerman, S. B., and MacIvor, J. S. (2017). Biodiversity in the city: fundamental questions for understanding the ecology of urban green spaces for biodiversity conservation. *BioScience*, 67(9):799–807.

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., et al. (2018). Earth Biogenome Project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences*, 115(17):4325–4333.

Littleford-Colquhoun, B. L., Clemente, C., Whiting, M. J., Ortiz-Barrientos, D., and Frère, C. H. (2017). Archipelagos of the Anthropocene: rapid and extensive differentiation of native terrestrial vertebrates in a single metropolis. *Molecular Ecology*, 26(9):2466–2481.

Losos, J. B. (2011). Convergence, adaptation, and constraint. *Evolution*, 65(7):1827–1840.

Luck, G. W. and Smallbone, L. T. (2010). Species diversity and urbanisation: patterns, drivers

and implications. *Urban ecology*, pages 88–119.

Makido, Y., Shandas, V., and Ferwati, S. (2020). Predicting urban growth. *Urban adaptation to climate change: The role of urban form in mediating rising temperatures*, pages 75–92.

Margres, M. J., Wray, K. P., Hassinger, A. T., Ward, M. J., McGivern, J. J., Moriarty Lemmon, E., Lemmon, A. R., and Rokyta, D. R. (2017). Quantity, not quality: rapid adaptation in a polygenic trait proceeded exclusively through expression differentiation. *Molecular biology and evolution*, 34(12):3099–3110.

McKinney, M. L. (2008). Effects of urbanization on species richness: a review of plants and animals. *Urban ecosystems*, 11:161–176.

Merckx, T., Souffreau, C., Kaiser, A., Baardsen, L. F., Backeljau, T., Bonte, D., Brans, K. I., Cours, M., Dahirel, M., Debortoli, N., et al. (2018). Body-size shifts in aquatic and terrestrial urban communities. *Nature*, 558(7708):113–116.

Messer, P. W., Ellner, S. P., and Hairston, N. G. (2016). Can population genetics adapt to rapid evolution? *Trends in Genetics*, 32(7):408–418.

Miles, L. S., Carlen, E. J., Winchell, K. M., and Johnson, M. T. (2021). Urban evolution comes into its own: Emerging themes and future directions of a burgeoning field. *Evolutionary Applications*, 14(1):3–11.

Miles, L. S., Rivkin, L. R., Johnson, M. T., Munshi-South, J., and Verrelli, B. C. (2019). Gene flow and genetic drift in urban environments. *Molecular ecology*, 28(18):4138–4151.

Miller, C. R., Latimer, C. E., and Zuckerberg, B. (2018). Bill size variation in northern cardinals associated with anthropogenic drivers across North America. *Ecology and Evolution*, 8(10):4841–4851.

Munshi-South, J. (2012). Urban landscape genetics: canopy cover predicts gene flow between white-footed mouse (*Peromyscus leucopus*) populations in New York City. *Molecular Ecology*, 21(6):1360–1378.

Munshi-South, J. and Richardson, J. L. (2020). Landscape genetic approaches to understanding movement. In Szulkin, M., Munshi-South, J., and Charmantier, A., editors, *Urban evolutionary biology*. Oxford University Press, USA.

Ollonen, J., Da Silva, F. O., Mahlow, K., and Di-Poï, N. (2018). Skull development,

ossification pattern, and adult shape in the emerging lizard model organism *Pogona vitticeps*: a comparative analysis with other squamates. *Frontiers in Physiology*, 9:350637.

Parsons, K., Rigg, A., Conith, A., Kitchener, A. C., Harris, S., and Zhu, H. (2020). Skull morphology diverges between urban and rural populations of red foxes mirroring patterns of domestication and macroevolution. *Proceedings of the Royal Society B*, 287(1928):20200763.

Reid, N. M., Proestou, D. A., Clark, B. W., Warren, W. C., Colbourne, J. K., Shaw, J. R., Karchner, S. I., Hahn, M. E., Nacci, D., Oleksiak, M. F., et al. (2016). The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science*, 354(6317):1305–1308.

Rivkin, L. R., Santangelo, J. S., Alberti, M., Aronson, M. F., de Keyzer, C. W., Diamond, S. E., Fortin, M.-J., Frazee, L. J., Gorton, A. J., Hendry, A. P., et al. (2019). A roadmap for urban evolutionary ecology. *Evolutionary applications*, 12(3):384–398.

Santangelo, J. S., Miles, L. S., Breitbart, S. T., Murray-Stoker, D., Rivkin, L. R., Johnson, M. T., Ness, R. W., Szulkin, M., Munshi-South, J., and Charmantier, A. (2020). Urban environments as a framework to study parallel evolution. *Urban evolutionary biology*, 36.

Sears, M. W., Angilletta, M. J., Wilson, R. S., Niehaus, A. C., Ribeiro, P. L., and Navas, C. A. (2007). Urban physiology: City ants possess high heat tolerance. *PLOS ONE*.

Seress, G. and Liker, A. (2015). Habitat urbanization and its effects on birds. *Acta Zoologica Academiae Scientiarum Hungaricae*, 61(4):373–408.

Seto, K. C., Sánchez-Rodríguez, R., and Fragkias, M. (2010). The new geography of contemporary urbanization and the environment. *Annual review of environment and resources*, 35:167–194.

Stuart, Y. E., Veen, T., Weber, J. N., Hanson, D., Ravinet, M., Lohman, B. K., Thompson, C. J., Tasneem, T., Doggett, A., Izen, R., et al. (2017). Contrasting effects of environment and genetics generate a continuum of parallel evolution. *Nature ecology & evolution*, 1(6):0158.

Szulkin, M., Munshi-South, J., and Charmantier, A. (2020). *Urban Evolutionary Biology*. Oxford University Press, USA.

Templeton, A. R., Shaw, K., Routman, E., and Davis, S. K. (1990). The genetic consequences of habitat fragmentation. *Annals of the Missouri Botanical Garden*, pages 13–27.

Tolley, K. and Burger, M. (2007). *Chameleons of Southern Africa*. Struik Publishers.

Tolley, K. A., Chase, B. M., and Forest, F. (2008). Speciation and radiations track climate transitions since the Miocene Climatic Optimum: a case study of southern African chameleons. *Journal of Biogeography*, 35(8):1402–1414.

Tolley, K. A., Tilbury, C. R., and Burger, M. (2022). Convergence and vicariance: speciation of chameleons in the Cape Fold Mountains, South Africa, and the description of three new species of *Bradypodion* Fitzinger, 1843. *African Journal of Herpetology*, 71(1):14–38.

Tolley, K. A., Townsend, T. M., and Vences, M. (2013). Large-scale phylogeny of chameleons suggests African origins and Eocene diversification. *Proceedings of the Royal Society B: Biological Sciences*, 280(1759):20130184.

Vergnes, A., Kerbiriou, C., and Clergeau, P. (2013). Ecological corridors also operate in an urban matrix: a test case with garden shrews. *Urban Ecosystems*, 16:511–525.

Whitehead, A., Clark, B. W., Reid, N. M., Hahn, M. E., and Nacci, D. (2017). When evolution is the solution to pollution: Key principles, and lessons from rapid repeated adaptation of killifish (*Fundulus heteroclitus*) populations. *Evolutionary Applications*, 10(8):762–783.

Wilson, A., Fenton, B., Malloch, G., Boag, B., Hubbard, S., and Begg, G. (2016). Urbanisation versus agriculture: a comparison of local genetic diversity and gene flow between wood mouse apodemus sylvaticus populations in human-modified landscapes. *Ecography*, 39(1):87–97.

Winchell, K. M., Campbell-Staton, S. C., Losos, J. B., Revell, L. J., Verrelli, B. C., and Geneva, A. J. (2023). Genome-wide parallelism underlies contemporary adaptation in urban lizards. *Proceedings of the National Academy of Sciences*, 120(3):e2216789120.

Winchell, K. M., Maayan, I., Fredette, J. R., and Revell, L. J. (2018). Linking locomotor performance to morphological shifts in urban lizards. *Proceedings of the Royal Society B*, 285(1880):20180229.

Winchell, K. M., Reynolds, R. G., Prado-Irwin, S. R., Puente-Rolon, A. R., and Revell, L. J. (2016). Phenotypic shifts in urban areas in the tropical lizard *Anolis cristatellus*. *Evolution*, 70(5):1009–1022.

Chapter 2

De Novo-Whole Genome Assemblies for Two Southern African Dwarf Chameleons (*Bradypodion*, Chamaeleonidae)

Abstract

A complete and high-quality reference genome has become a fundamental tool for the study of functional, comparative, and evolutionary genomics. However, efforts to produce high-quality genomes for African taxa are lagging given the limited access to sufficient resources and technologies. The southern African dwarf chameleons (*Bradypodion*) are a relatively young lineage, with a large body of evidence demonstrating the highly adaptive capacity of these lizards. *Bradypodion* are known for their habitat specialization, with evidence of convergent phenotypes across the phylogeny. However, the underlying genetic architecture of these phenotypes remains unknown for *Bradypodion* and without adequate genomic resources, many evolutionary questions cannot be answered. We present de novo assembled whole-genomes for *B. pumilum* and *B. ventrale*, using Pacific Biosciences long-read sequencing data. BUSCO analysis revealed that 96.36 % of single copy orthologs were present in the *B. pumilum* genome and 94 % in *B. ventrale*. Moreover, these genomes boast scaffold N50 of 389.6 Mb and 374.9 Mb, respectively. Based on a whole genome alignment of both *Bradypodion* genomes, *B. pumilum* is highly syntenic with *B. ventrale*. Furthermore, *Bradypodion* is also syntenic with *Anolis* lizards, despite the divergence between these lineages estimated to be nearly 170 Mya. Coalescent analysis of the genomic data also suggests that historical changes in effective population size for these species corresponds to notable shifts in the southern African environment. These high-quality *Bradypodion* genome assemblies will support future research on the evolutionary history, diversification, and genetic underpinnings of adaptation in *Bradypodion*.

Significance We generated high-quality reference genome assemblies for two southern African dwarf chameleons, *Bradypodion pumilum* and *B. ventrale*. These genomes are currently among the most complete and contiguous reptile assemblies and are the second and third on record for Chamaeleonidae. Chameleons show unique adaptive traits and ecology, and these assemblies presented here provide a novel resource for evolutionary research into this group.

Keywords: Africa, chameleon genomics, HiFi data, Hi-C, reference genome, reptiles

2.1 Introduction

There has been a recent concerted effort to produce fully sequenced genomes for a range of vertebrate species. Initially centered around model organisms and charismatic groups, attention has now turned to include less studied groups (Genome 10K Community of Scientists, 2009; Koepfli et al., 2015; Lewin et al., 2018; Blaxter et al., 2022; Ebenezer et al., 2022). Reptiles are one of the most diverse clades of extant tetrapods, boasting nearly 12000 species. However, the number of reptile genome assemblies are few relative to other tetrapod taxonomic groups. Furthermore, over 2200 reptile species (18% of the total) occur in mainland Africa and associated landmasses (Uetz et al., 2023), yet only two full genome assemblies have yet been published (*Paroedura picta*, Hara et al., 2018; *Hemicordylus capensis*, Leitão et al., 2023). With finite resources, minimal access to sequencing technologies, and limited bioinformatic expertise, efforts to produce high-quality genome assemblies for African taxa are lagging (Ebenezer et al., 2022). However, given the efforts of genomic initiatives (Ebenezer et al., 2022; Lewin et al., 2018), the production of high-quality genomic resources for African taxa is gaining momentum (Ebenezer et al., 2022; Gupta, 2022; Lewin et al., 2022; Ishengoma, 2023).

An iconic group of African reptiles are the Chamaeleonidae - arguably one of the most specialized groups of reptiles (Bickel and Losos, 2002; Tolley and Herrel, 2013; Diaz et al., 2015). Chameleons have undergone evolutionary shifts from a terrestrial to an arboreal lifestyle (Tolley et al., 2013), exhibiting an array of adaptive functional traits presumably linked to habitat specialization (Hopkins and Tolley, 2011; Herrel et al., 2013). The southern African Dwarf Chameleons, *Bradypodion*, are the youngest chameleon lineage, with multiple instances of adaptive responses to habitat shifts since the Miocene. Some clades have undergone rapid diversification, largely associated with changes in natural environments through ecological speciation (Tolley et al., 2008, 2022). Multiple convergence events have resulted in a few phenotypic morphs, or ecomorphs, in the genus. However, it is unclear whether the convergence of ecomorphs is due to standing genetic variation, or if ecomorphs arise because of convergent novel mutations regardless of their phylogenetic history. Similarly, *Anolis*, a large clade of new world lizards occupying arboreal habitats broadly comparable to *Bradypodion*, display convergent ecomorphs across distantly related lineages. Given these parallels, anole and dwarf chameleon ecomorphs might employ similar genetic pathways to give rise to shared elements of their ecomorphological phenotypes. High-quality reference genomes from these clades are necessary to investigate this hypothesis.

Using long-read sequencing, we generated genome assemblies of high quality and completeness de novo for two *Bradypodion* species, the Cape Dwarf Chameleon (*B. pumilum*) and the Eastern Dwarf Chameleon (*B. ventrale*). We place these assemblies within a phylogenetic framework, identify syntenic relationships within *Bradypodion* and between *Anolis sagrei*, and explore the demographic history of *B. pumilum* and *B. ventrale*. These genomes provide an exciting opportunity to explore the genetic architecture that underlies the genotype-phenotype and genotype-environment associations in *Bradypodion*, which could extend to other chameleon genera.

2.2 Results and Discussion

2.2.1 Assembly contiguity and chromosome size

The assemblies for both dwarf chameleon species are the most contiguous reptile genomes produced to date, with a scaffold N50 of 389.6 Mb (*Bradypodion pumilum*, genome size: 2.43 Gb) and 374.9 Mb (*B. ventrale*, genome size: 2.40 Gb, fig. 2.1 *A*). The next most contiguous published assembly is *Hemicordylus capensis* (359.65 Mb, genome size: 2.29 Gb,Leitão et al., 2023), followed by *Shinisaurus crocodilurus* (296.95 Mb, genome size: 2.19 Gb, Xie et al., 2022). For each *Bradypodion* assembly, the three largest scaffolds together cover more than 50% of the genome. *Bradypodion pumilum* (BraPum1.0) compromises 90% of the genome within the seven largest scaffolds, while *B. ventrale* (BraVen1.1) holds 90% of the genome within the nine largest scaffolds similar to the contiguity of *Sceloporus undulatus* and *Phrynosoma platyrhinos* (table S1). Both *Bradypodion* assemblies are highly correlated (fig. S1; BraPum1.0: $r^2 = 0.991$, $p \leq 2.2 \times 10^{-16}$; BraVen1.1: $r^2 = 0.982$, $p \leq 9.7 \times 10^{-15}$) with chromosome sizes estimated based on the published karyotype for *Bradypodion* (Rovatsos et al., 2017).

2.2.2 Assembly completeness and Phylogenetics

To estimate assembly completeness, *Bradypodion* assemblies were tested for the occurrence of a curated set of 3354 protein-coding genes present in single-copy across vertebrate genomes (vertebrata_odb10; Manni et al., 2021). BraPum1.0 contains 3232 single-copy genes with 39 duplicated genes and BraVen1.1 has 3230 genes in single-copy and 49 duplicated (fig. 2.1*B*). BraVen1.1 has a total of 3279 genes present, behind *Hemicordylus capensis* with 3287 genes and *Furcifer pardalis* with 3282 of the possible 3354 coding genes that are assumed to be present in all vertebrates. Both BraPum1.0 and BraVen1.1 retain approximately 97% of the possible

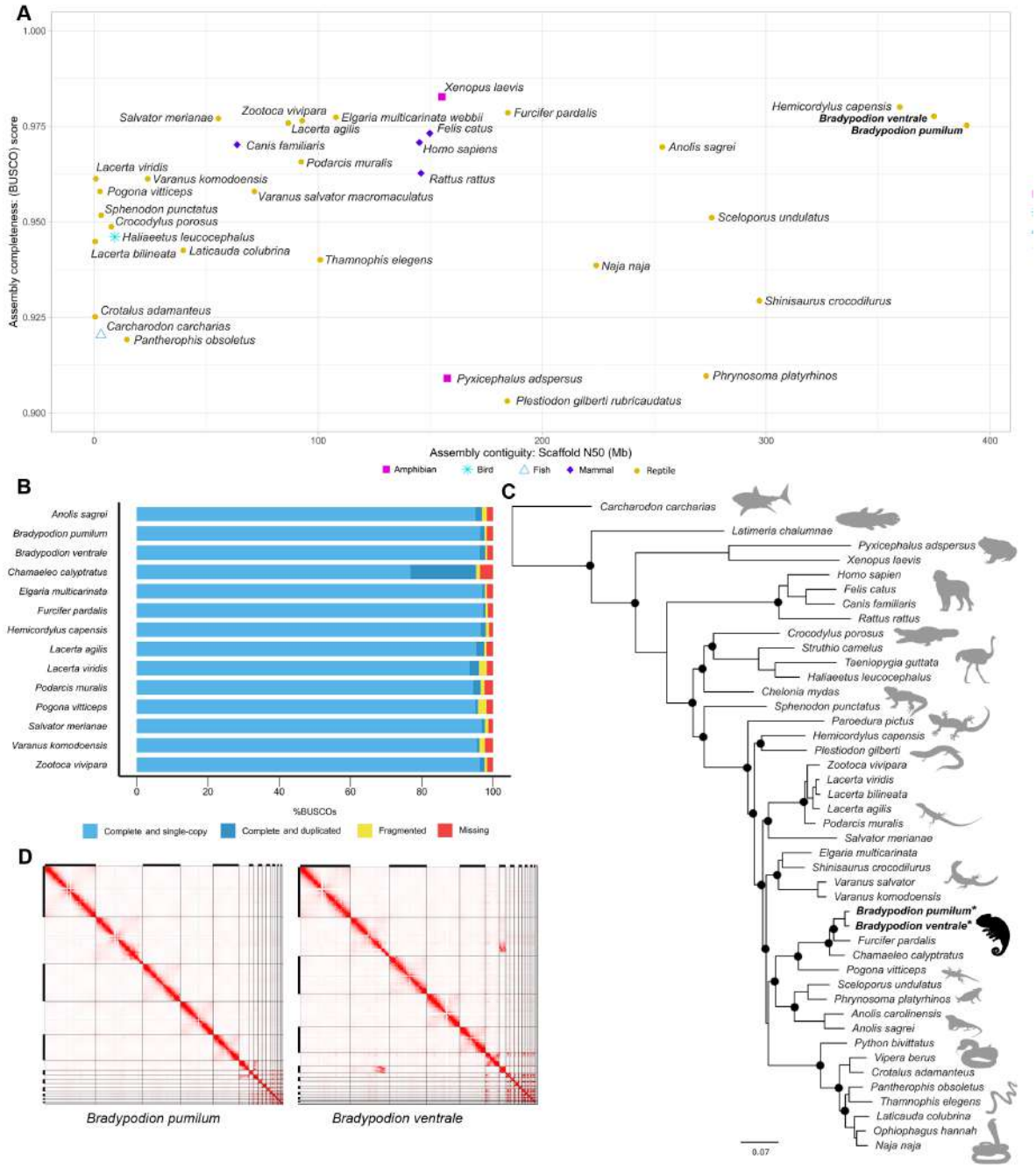


Figure 2.1: A. Scatterplot of recent high-quality vertebrate genomes comparing assembly completeness scores against scaffold N50. Study species are highlighted in bold. BUSCO scores presented only contain assemblies with scores above 0.9. B. BUSCO assessment of assembly completeness for *Bradypodion* and closest nine available reptile assemblies based on BUSCO scores. Colours indicate single-copy (light blue), duplicated (dark blue), fragmented (yellow), and missing (red) genes. C. Maximum-likelihood phylogenetic tree based on single-copy proteins from BUSCO analysis. *Bradypodion* assemblies are placed within the context of representative vertebrate taxa in all major groups. The assemblies presented here are indicated with *. Supported nodes ($\geq$ 70% bootstrap) are indicated with a black dot. D. Link density histograms depict associations between parts of the *Bradypodion* assemblies, *B. pumilum* and *B. ventrale*, depicting the mapping location of HiC read pairs. Increasing intensity of red indicated more read pairs map to a particular coordinate of the histogram. White shaded areas indicate no reads mapping to that coordinate in the histogram. Black bars on each axis indicate every other chromosome in the assembly.

coding genes present in all vertebrates.

Using 3354 single-copy orthologs from the vertebrata_odb10 BUSCO dataset, we generated a phylogenetic tree to incorporate the placement of *Bradypodion* relative to other vertebrates with available assemblies (fig. 2.1*C*). Tree estimates using the concatenated sequence alignment produced a topology that mirrors the most complete squamate phylogeny (Burbrink et al., 2020). We also included the transcriptome of *Chamaeleo calyptratus* and recovered the placement of Chamaeleonidae as sister to the family Agamidae, supporting the hypothesis that acrodonts are a monophyletic group. The phylogeny generated using representative genomes from a few groups contains a fraction of the power that could be harnessed using genomic data to address hypotheses relating to evolutionary histories. While the implementation of whole genomic data within the field of phylogenetics is still in its infancy (Simões and Pyron, 2021; Card et al., 2023), the addition of these *Bradypodion* assemblies will be valuable in resolving relationships between lineages while uncovering the origins of traits within squamates (Card et al., 2023).

2.2.3 Hi-C and Inversion

Each link density histogram for the *Bradypodion* assemblies indicated the presence of six larger chromosomes, consistent with the six macrochromosomes observed in a karyotype of *Bradypodion thamnobates* (fig. 2.1*D*, Rovatsos et al., 2017). For *B. ventrale*, scaffolds seven and eight were not colinear; however, these scaffolds are homologous to each other with large inversion along scaffold eight (fig. S2). Upon further investigation of the reduced coverage regions in both scaffolds, we detected breakpoints for an inversion on scaffold eight (position Sc8: 11460 kbp-11465 kbp; position Sc8: 56620 kbp-56630 kbp; fig. S3). Furthermore, along the inversion breakpoints, we found that PacBio long reads span these sections of the assembly indicating the inversion is not an assembly artefact (fig. S3). Subsequently, we retained scaffold seven while removing scaffold eight from the initial *B. ventrale* assembly. After removing this scaffold, HiC reads appropriately mapped along scaffold seven in the updated assembly (BraVen1.1; fig. 1*D*), with minor contacts corresponding to scaffold two indicative of homology in this region.

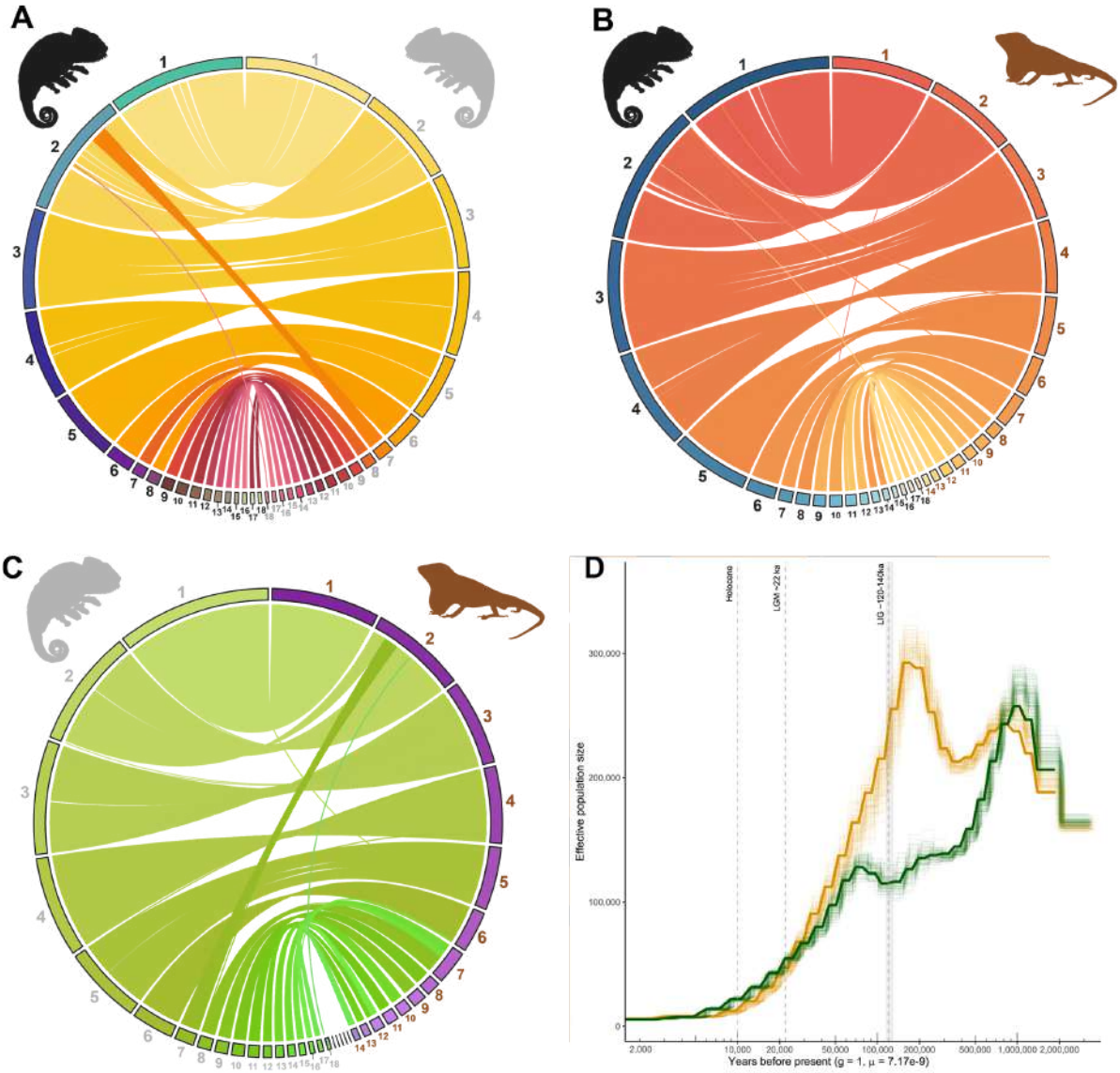


Figure 2.2: Synteny Circos plots for whole genome alignment between A. *B. pumilum* (black) and *B. ventrale* (gray), B. *B. pumilum* and *A. sagrei* (brown), and C. *B. ventrale* and *A. sagrei*. Each segment of the outer circle represents a unique scaffold in the genome alignment, with the largest scaffolds at the top of each plot. Links are drawn between scaffolds with similarity, where scaffolds with links of at least 1 kb in length were retained for each plot. D. Changes in effective population size for *B. pumilum* (green) and *B. ventrale* (gold) over longer time periods inferred using the PSMC method. Solid lines indicate the model results using the average mutation rate (7.17×10^{-9}). Clouded lines indicate unique bootstrapping results. The dotted lines indicate Holocene, the Last Glacial Maximum, and the Last Interglacial periods.

2.2.4 Repetitive element content

The total repetitive element content for *Bradypodion pumilum* was estimated to be 61% of the genome, with 52% comprised of known interspersed repeats, and 5.62% unclassified repeats (table S2). Comparatively, the *B. ventrale* assembly was shown to have a slightly lower repeat content estimated to be 59.54%, 50% of which were found to be interspersed repeats, and 6.65% unclassified repeats. Both assemblies held a similar diversity in the type of repeat classes, including short, interspersed elements (SINEs), long interspersed nuclear elements (LINEs), long terminal repeat retrotransposons (LTRs), and DNA transposons. However, *B. pumilum* contained a higher proportion of LTRs, whereas *B. ventrale* retained a higher proportion of LINEs. Overall, the repetitive element content appears to be relatively conserved within *Bradypodion* (fig. S4), given that these two species are roughly 7 Mya divergent (Tolley et al. 2008).

2.2.5 Synteny analysis

A synteny analysis of the two *Bradypodion* assemblies shows a conserved genome structure when comparing the 18 largest chromosomes (links $>$ 1kbp), with synteny conserved primarily between the five largest chromosomes (fig. 2.2*A*). Similarly, this pattern is repeated between the two *Bradypodion* assemblies with *Anolis sagrei*, with rearrangements only on a few smaller chromosomes. Within *B. pumilum*, chromosomes six and eight both align with chromosome six of *B. ventrale* and *A. sagrei*. This may be due to a chromosomal fission in the *Bradypodion pumilum* lineage after the divergence of *B. pumilum* and *B. ventrale*. However, this pattern might instead suggest a failure to fully assemble chromosome six in *B. pumilum*, and this alternative explanation is supported by the substantial number of HiC links between these scaffolds, more than either scaffolds 6 or 8 share with any other scaffold in the assembly (fig. 1*D*). Similarly, both chromosomes two and seven of *B. ventrale* are homologous to chromosome two in *B. pumilum* and *A. sagrei*. Chromosome seven in *B. pumilum* is syntenic with chromosome eight in *B. ventrale*, with both these chromosomes aligning partly to chromosome seven in *A. sagrei* (fig. 2*B*). We were unable to conclusively distinguish between the alternatives - either a lineage-specific fission of chromosomes 2 and 7 in *B. ventrale* or a failure to assemble these scaffolds into a single chromosome in *B. pumilum*. The latter, however, is more likely given the increased number of links between chromosomes 2 and 7 in *B. ventrale* but additional data are required to confidently resolve between the alternatives.

Anolis sagrei has an XY sex-determination system, and recently chromosome seven in *A. sagrei* was identified as the X chromosome (Geneva et al., 2022). Despite these highly complete *Bradypodion* assemblies, it remains unclear which sex-determination system is present in *Bradypodion*. Within the chameleon genus *Furcifer*, the sex determination system was found to be a highly differentiated ZW system; however, similar patterns were inconclusive when tested across the Chamaeleonidae family (Rovatsos et al., 2017). The discovery of XX/XY sex chromosomes in *Chamaeleo calyptratus* suggests that at least one transition between ZZ/ZW and XX/XY systems has occurred within the Chamaeleonidae (Nielsen et al., 2018). When comparing the synteny between *Bradypodion* and *A. sagrei*, chromosome seven in *A. sagrei* is related to chromosomes seven, 13 and 14 in *B. pumilum*, and chromosomes eight, 14, and 15 in *B. ventrale* (fig. $2C$). Given that there is no distinct pattern between *Bradypodion* and known X chromosome in *A. sagrei*, it is likely that the sex chromosomes for *Bradypodion* reside elsewhere.

2.2.6 Estimation of population size history

The trajectory of changes in the effective population size (N_e) of *Bradypodion pumilum* and *B. ventrale* were estimated using a pairwise sequentially Markovian coalescent (PSMC) approach (Li and Durbin, 2011). We found the greatest change in predicted population size using the maximum mutation rate of 3.76×10^{-9} per bp per year relative to the minimum and average in our models (fig. S5). When using the minimum and average rates, there were similar trends between both predicted models, resulting in a conservative estimate for changes in the population size of these species. An increase in N_e for both species of *Bradypodion* was estimated for the early Pleistocene (1 - 2 mya: fig. $2D$), followed by a sharp decline in N_e of *B. pumilum* from 1 mya - 100 kya. In contrast, there was a gradual reduction in N_e of *B. ventrale* between 1 mya - 500 kya, after which a sharp increase led to a peak in N_e at ∼200 kya. Furthermore, the results suggest that both species of *Bradypodion* experienced a decrease in N_e along similar trajectories over the last ∼70 kya; however, PSMC estimates of N_e near present should be interpreted with caution as these estimates can be unreliable (i.e., <10 kya BP; (Dussex et al., 2021). While the low estimates of N_e for the currently range-restricted *B. pumilum* are plausible, the predicted N_e over the last 70 kyr for *B. ventrale* may be underestimated given its widespread distribution at present (Tolley and Burger, 2007), and N_e is unlikely to be extremely low. Nevertheless, some initial inferences can be made regarding the estimated trends in N_e in the Pleistocene. Over the last 120 kya, the biomes

of South Africa underwent shifts in distribution before reaching their current extent roughly 10 kya (Verboom et al., 2009; Tolley et al., 2014). Similar to other species occurring in this region (Barlow et al., 2013; Péron and Altwegg, 2015; Taft et al., 2022), favourable habitats for *Bradypodion* likely contracted as other biomes expanded. These range contractions could have resulting in a decline in N_e. Given that *Bradypodion* is largely affected by drastic changes in habitat type over longer time periods (Tolley et al., 2008; da Silva and Tolley, 2013; Tolley et al., 2022), it is possible that the decline in population size of these two species is related to recent contractions in their natural available habitats and is not due to recent anthropogenic habitat transformation.

2.3 Methods

2.3.1 Sampling and sequencing

For long-read sequencing, a male *Bradypodion pumilum* from Cape Town (-34.03, 18.73) and a male *B. ventrale* from an introduced population in Johannesburg (-26.15, 28.07) were collected. High-molecular-weight DNA was extracted from muscle and liver tissues using the Qiagen Genomic-tip kit and sequenced on PacBio Sequel II 8M SMRT cells. Additional details regarding sampling, Omni-C library preparation and sequencing are in the supplementary methods.

2.3.2 Assembly contiguity and chromosome size analysis

Basic assembly statistics were generated using the stats tool in BBTools v38.9 to calculate contiguity statistics (Bushnell, 2021). To compare contiguity between representative assemblies across all vertebrate groups, we downloaded 38 additional assemblies from the NCBI database (https://www.ncbi.nlm.nih.gov/genome/) to compare contiguity between representative assemblies across all vertebrate groups, with 70% being reptiles (table S1).

We estimated whether chromosome sizes in *Bradypodion* correlate with the length of scaffolds in our assemblies (Geneva et al., 2022) by estimating the size of each chromosome in the published *Bradypodion* karyotype (fig. $1A$ in Rovatsos et al., 2017) by measuring the chromosomes sizes using ImageJ (Schneider et al., 2012). The fraction of the total karyotype occupied by each chromosome was calculated and multiplied by the total size of *each* *Bradypodion* assembly to generate an estimate of nucleotide content. We then calculated the

correlation between scaffold size and estimated chromosome size via linear regression using the lm function in R v4.2.2 (R Core Team, 2022).

2.3.3 Assembly completeness and Phylogenetics

A Benchmarking Universal Single-Copy Orthologs (BUSCO v5.4.4; Simão et al., 2015) analysis was used to evaluate the completeness of the assemblies using the vertebrata odb10 dataset. Subsequently, we constructed a phylogenetic tree using the single-copy BUSCOs for all assemblies downloaded from NCBI. Implementing the BUSCO phylogenomic pipeline (McGowan 2019), we generated multiple sequence alignments using MUSCLE for all protein sequences present in all assemblies. Alignments were trimmed with trimal v1.2rev59 (Capella-Gutiérrez et al., 2009). A maximum likelihood phylogenetic tree was generated using FastTree v2.1.11 (Price et al., 2010).

2.3.4 HiC chromatin conformation mapping

Link density histograms of paired reads were generated from the Omni-C Libraries using Juicer v1.6 ad visualized in Juicebox v2.16.00. This revealed reduced read mapping on scaffolds seven and eight in the *Bradypodion ventrale* initial assembly (fig. S6). With MUMmer v4.0.0 (Marçais et al., 2018), the initial assembly as aligned as a query and reference using nucmer to inspect whether scaffolds seven and eight are colinear duplicates of the same chromosome. Coverage across the genome assembly was estimated using SAMtools v1.7 (Li et al., 2009) to examine the points in the assembly with reduced mapping.

2.3.5 Repetitive element content

Estimating the repetitive landscape of the *Bradypodion* assemblies was performed by modelling repeats de novo on each assembly using RepeatModeler v v2.0.2a (Flynn et al., 2020) and annotated the repeat consensus sequences using RepeatMasker v4.1.2 (http://www.repeatmasker.org/). To maximize element identification, a custom bash script to specify the order of the four libraries was used as references for the masking process: (i) simple repeats, (ii) Tetrapoda DFam library v3.7, (iii) classified elements from the species-specific library, and (iv) unknown elements from the species-specific library. For the age distribution of transposable elements in each genome, the divergence of an insert from its family consensus was used as a proxy for its age. Alignments for each repeat family were

generated and the Kimura-2 parameter divergence from consensus (correcting for CpG sites) was calculated using the calcDivergenceFromAlign.pl RepeatMasker tool.

2.3.6 Synteny analysis

A whole genome alignment using SatsumaSynteny v3.1 was used to investigate synteny between the *Bradypodion* assemblies. The alignment was visualized in the R package circlize v0.4.15 to implement Circos. Both *Bradypodion* assemblies were aligned to *Anolis sagrei* v2.1 to investigate synteny between these 160-170 mya divergent taxa (Zheng and Wiens, 2016; Burbrink et al., 2020).

2.3.7 Estimation of population size history

We applied PSMC analysis to infer changes in effective population size history from the assembly sequence of *Bradypodion pumilum* and *B. ventrale* (Li and Durbin, 2011). For each genome, the associated PacBio long-read data were mapped using Minimap2 (Li, 2018). PSMC parameters were used as described in Bergeron et al. (2023), including -d 13 and -D 100. Next, we generated PSMC curves with 100 bootstraps using the parameters linked above. For the mutation rate, we used the minimum (9.45×10^{-9}), average (7.17×10^{-9}), and maximum (3.76×10^{-9}) per bp per year, estimated from the closest relative *Pogona vitticeps* (Bergeron et al., 2023) and a generation time of 1 year (Tolley and Burger, 2007).

2.4 Acknowledgements

This work was supported by the National Research Foundation of South Africa [grant number 136381] and the National Science Foundation [grant numbers 1927194, 2152059]. Samples exported to the USA under CITES permit [#206297] issued by the South African Department of Fisheries, Forestry and Environment, and provincial permits (CapeNature #CNN4459-5795 permit to KA Tolley and Gauteng Nature Conservation #000186 permit to GJ Alexander) and ethics clearance from the University of Johannesburg (2019-10-10). Thanks to Matthew Adair, Marius Burger, Devon C. Main and Bettine van Vuuren for their assistance with this project. To Cleo H. Falvey and Alyssa A. Vanerelli for their helpful comments on this manuscript.

2.4.1 Data Availability

Data presented in this article are available from NCBI under BioProject PRJNA986319, BioSample SAMN35825189 (*Bradypodion pumilum*), and SAMN35825190 (*Bradypodion ventrale*). Repetitive elements available at https://doi.org/ 10.7910/DVN/EPSA3O.

References

Barlow, A., Baker, K., Hendry, C. R., Peppin, L., Phelps, T., Tolley, K. A., Wüster, C. E., and Wüster, W. (2013). Phylogeography of the widespread African puff adder (*Bitis arietans*) reveals multiple Pleistocene refugia in southern Africa. *Molecular ecology*, 22(4):1134–1157.

Bergeron, L. A., Besenbacher, S., Zheng, J., Li, P., Bertelsen, M. F., Quintard, B., Hoffman, J. I., Li, Z., St. Leger, J., Shao, C., et al. (2023). Evolution of the germline mutation rate across vertebrates. *Nature*, pages 1–7.

Bickel, R. and Losos, J. B. (2002). Patterns of morphological variation and correlates of habitat use in chameleons. *Biological Journal of the Linnean Society*, 76(1):91–103.

Blaxter, M., Archibald, J. M., Childers, A. K., Coddington, J. A., Crandall, K. A., Di Palma, F., Durbin, R., Edwards, S. V., Graves, J. A., Hackett, K. J., et al. (2022). Why sequence all eukaryotes? *Proceedings of the National Academy of Sciences*, 119(4):e2115636118.

Burbrink, F. T., Grazziotin, F. G., Pyron, R. A., Cundall, D., Donnellan, S., Irish, F., Keogh, J. S., Kraus, F., Murphy, R. W., Noonan, B., et al. (2020). Interrogating genomic-scale data for Squamata (lizards, snakes, and amphisbaenians) shows no support for key traditional morphological relationships. *Systematic Biology*, 69(3):502–520.

Bushnell, B. (2021). Bbmap (version 38.90).

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15):1972–1973.

Card, D. C., Jennings, W. B., and Edwards, S. V. (2023). Genome evolution and the future of phylogenomics of non-avian reptiles. *Animals*, 13(3):471.

da Silva, J. M. and Tolley, K. A. (2013). Ecomorphological variation and sexual dimorphism

in a recent radiation of dwarf chameleons (*Bradypodion*). *Biological Journal of the Linnean Society*, 109(1):113–130.

Diaz, R. E., Anderson, C. V., Baumann, D. P., Kupronis, R., Jewell, D., Piraquive, C., Kupronis, J., Winter, K., Bertocchini, F., and Trainor, P. A. (2015). The veiled chameleon (chamaeleo calyptratus duméril and duméril 1851): a model for studying reptile body plan development and evolution. *Cold Spring Harbor Protocols*, 2015(10):pdb–emo087700.

Dussex, N., Van Der Valk, T., Morales, H. E., Wheat, C. W., Díez-del Molino, D., Von Seth, J., Foster, Y., Kutschera, V. E., Guschanski, K., Rhie, A., et al. (2021). Population genomics of the critically endangered kākāpō. *Cell Genomics*, 1(1):100002.

Ebenezer, T. E., Muigai, A. W., Nouala, S., Badaoui, B., Blaxter, M., Buddie, A. G., Jarvis, E. D., Korlach, J., Kuja, J. O., Lewin, H. A., et al. (2022). Africa: sequence 100,000 species to safeguard biodiversity. *Nature*, 603(7901):388–392.

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., and Smit, A. F. (2020). Repeatmodeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences*, 117(17):9451–9457.

Geneva, A. J., Park, S., Bock, D. G., de Mello, P. L., Sarigol, F., Tollis, M., Donihue, C. M., Reynolds, R. G., Feiner, N., Rasys, A. M., et al. (2022). Chromosome-scale genome assembly of the brown anole (*Anolis sagrei*), an emerging model species. *Communications Biology*, 5(1):1126.

Genome 10K Community of Scientists (2009). Genome 10k: a proposal to obtain whole-genome sequence for 10 000 vertebrate species. *Journal of Heredity*, 100(6):659–674.

Gupta, P. K. (2022). Earth biogenome project: present status and future plans. *Trends in Genetics*.

Hara, Y., Takeuchi, M., Kageyama, Y., Tatsumi, K., Hibi, M., Kiyonari, H., and Kuraku, S. (2018). Madagascar ground gecko genome analysis characterizes asymmetric fates of duplicated genes. *BMC biology*, 16(1):1–19.

Herrel, A., Tolley, K. A., Measey, G. J., da Silva, J. M., Potgieter, D. F., Boller, E., Boistel, R., and Vanhooydonck, B. (2013). Slow but tenacious: an analysis of running and gripping performance in chameleons. *Journal of Experimental Biology*, 216(6):1025–1030.

Hopkins, K. P. and Tolley, K. A. (2011). Morphological variation in the Cape Dwarf Chameleon (*Bradypodion pumilum*) as a consequence of spatially explicit habitat structure differences. *Biological Journal of the Linnean Society*, 102(4):878–888.

Ishengoma, E. (2023). Vertebrate genomics and adaptation–status and prospects in africa. *Molecular Ecology*.

Koepfli, K.-P., Paten, B., of Scientists, G. K. C., and O'Brien, S. J. (2015). The Genome 10K Project: a way forward. *Annual Review of Animal Biosciences*, 3(1):57–111.

Leitão, H. G., Diedericks, G., Broeckhoven, C., Baeckens, S., and Svardal, H. (2023). Chromosome-level genome assembly of the Cape cliff lizard (*Hemicordylus capensis*). *Genome Biology and Evolution*, 15(2):evad001.

Lewin, H. A., Richards, S., Lieberman Aiden, E., Allende, M. L., Archibald, J. M., Bálint, M., Barker, K. B., Baumgartner, B., Belov, K., Bertorelle, G., et al. (2022). The earth biogenome project 2020: Starting the clock.

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., et al. (2018). Earth BioGenome Project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences*, 115(17):4325–4333.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, 34(18):3094–3100.

Li, H. and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, 475(7357):493–496.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, . G. P. D. P. (2009). The sequence alignment/map format and samtools. *Bioinformatics*, 25(16):2078–2079.

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., and 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 and evolution*, 38(10):4647–4654.

Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., and Zimin, A. (2018).

MUMmer4: A fast and versatile genome alignment system. *PLoS Computational Biology*, 14(1):e1005944.

Nielsen, S. V., Banks, J., Diaz Jr, R., Trainor, P., and Gamble, T. (2018). Dynamic sex chromosomes in old world chameleons (Squamata: Chamaeleonidae). *Journal of evolutionary biology*, 31(4):484–490.

Péron, G. and Altwegg, R. (2015). Departures from the energy-biodiversity relationship in south african passerines: are the legacies of past climates mediated by behavioral constraints on dispersal? *PloS one*, 10(7):e0133992.

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS one*, 5(3):e9490.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rovatsos, M., Altmanová, M., Johnson Pokorná, M., Velenskỳ, P., Sanchez Baca, A., and Kratochvíl, L. (2017). Evolution of karyotypes in chameleons. *Genes*, 8(12):382.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods*, 9(7):671–675.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19):3210–3212.

Simões, T. R. and Pyron, R. A. (2021). The squamate tree of life. *Bulletin of the Museum of Comparative Zoology*, 163(2):47–95.

Taft, J. M., Maritz, B., and Tolley, K. A. (2022). Stable climate corridors promote gene flow in the Cape sand snake species complex (Psammophiidae). *Zoologica Scripta*, 51(1):58–75.

Tolley, K. and Burger, M. (2007). *Chameleons of Southern Africa*. Struik Publishers.

Tolley, K. A., Bowie, R. C., Measey, J. G., Price, B. W., and Forest, F. (2014). The shifting landscape of genes since the Pliocene: terrestrial phylogeography in the Greater Cape Floristic Region. In Allsopp, N., Colville, J. F., and Verboom, G. A., editors, *Fynbos: ecology, evolution, and conservation of a megadiverse region*. Oxford University Press, USA.

Tolley, K. A., Chase, B. M., and Forest, F. (2008). Speciation and radiations track climate transitions since the Miocene Climatic Optimum: a case study of southern African chameleons. *Journal of Biogeography*, 35(8):1402–1414.

Tolley, K. A. and Herrel, A. (2013). *The Biology of Chameleons*. Univ of California Press.

Tolley, K. A., Tilbury, C. R., and Burger, M. (2022). Convergence and vicariance: speciation of chameleons in the Cape Fold Mountains, South Africa, and the description of three new species of *Bradypodion* Fitzinger, 1843. *African Journal of Herpetology*, 71(1):14–38.

Tolley, K. A., Townsend, T. M., and Vences, M. (2013). Large-scale phylogeny of chameleons suggests African origins and Eocene diversification. *Proceedings of the Royal Society B: Biological Sciences*, 280(1759):20130184.

Uetz, P., Freed, P., Aguilar, R., Reyes, F., and Hošek, J. (2023). The Reptile Database.

Verboom, G. A., Archibald, J. K., Bakker, F. T., Bellstedt, D. U., Conrad, F., Dreyer, L. L., Forest, F., Galley, C., Goldblatt, P., Henning, J. F., et al. (2009). Origin and diversification of the Greater Cape flora: ancient species repository, hot-bed of recent radiation, or both? *Molecular Phylogenetics and Evolution*, 51(1):44–53.

Xie, H.-X., Liang, X.-X., Chen, Z.-Q., Li, W.-M., Mi, C.-R., Li, M., Wu, Z.-J., Zhou, X.-M., and Du, W.-G. (2022). Ancient demographics determine the effectiveness of genetic purging in endangered lizards. *Molecular Biology and Bvolution*, 39(1):msab359.

Zheng, Y. and Wiens, J. J. (2016). Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for Squamate reptiles (lizards and snakes) based on 52 genes and 4162 species. *Molecular Phylogenetics and Evolution*, 94:537–547.

Supplementary Methods

Sample collection and DNA isolation

For long-read sequencing, a male *Bradypodion pumilum* from Cape Town, Western Cape Province (-34.03, 18.73) and a male *B. ventrale* were collected from an introduced population in Johannesburg, Gauteng Province (-26.15, 28.07). Ethics clearances for the collection of these specimens were granted by a nationally accredited animal research ethics committee at the University of Johannesburg (Ref. 2019-10-10). The ethical guidelines used are in accordance with regulations set out in the South African National Standards (SANS 10386: 2008) and also follow the ARRIVE guidelines (https://arriveguidelines.org). After humane euthanasia using an oral application of 20% benzocaine gel, each animal was flash-frozen with liquid nitrogen and subsequently stored at -80 ^{o}C. High-molecular-weight DNA was extracted from muscle and liver tissues using the Qiagen Genomic-tip kit for isolation of high molecular-weight DNA and extracts were quantified using Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA).

PacBio Library, Sequencing, and Assembly

For each sample, a PacBio SMRTbell library (20kb) for PacBio Sequel was constructed using SMRTbell Express Template Prep Kit 2.0 (PacBio, Menlo Park, CA, USA) using the manufacturer-recommended protocol. The library was bound to polymerase using the Sequel II Binding Kit 2.0 (PacBio) and loaded onto PacBio Sequel II. Sequencing was run on PacBio Sequel II 8M SMRT cells. PacBio CCS reads were then used as input to Hifiasm v0.15.4-r347 with default parameters. BLAST (Basic Local Alignment Search Tool) results of the initial Hifiasm output assembly against the NCBI nucleotide database (https://www.ncbi.nlm.nih.gov/) were used as input for blobtools2 v1.1.1 and scaffolds identified as possible contamination were removed from the assembly. Finally, purge dups3 v1.2.5 was used to remove haplotigs and contig overlaps.

Dovetail Omni-C Library Preparation and Sequencing

To scaffold initial Hifiasm assemblies additional Dovetail Omni-C libraries were prepared. For each Dovetail Omni-C library, chromatin was fixed in place with formaldehyde in the nucleus and then extracted. Fixed chromatin was digested with DNAse I, and chromatin ends were repaired and ligated to a biotinylated bridge adapter followed by proximity ligation of the

adapter-containing ends. After proximity ligation, the cross-links were reversed, and the DNA was purified. Purified DNA was treated to remove biotin that was not internal to the ligated fragments. Sequencing libraries were generated using NEBNext Ultra enzymes and Illumina-compatible adapters. Biotin-containing fragments were isolated using streptavidin beads prior to PCR enrichment of each library. The library was sequenced on an Illumina HiSeqX platform to produce a sequence coverage of approximately 30x.

Scaffolding assemblies with HiRise

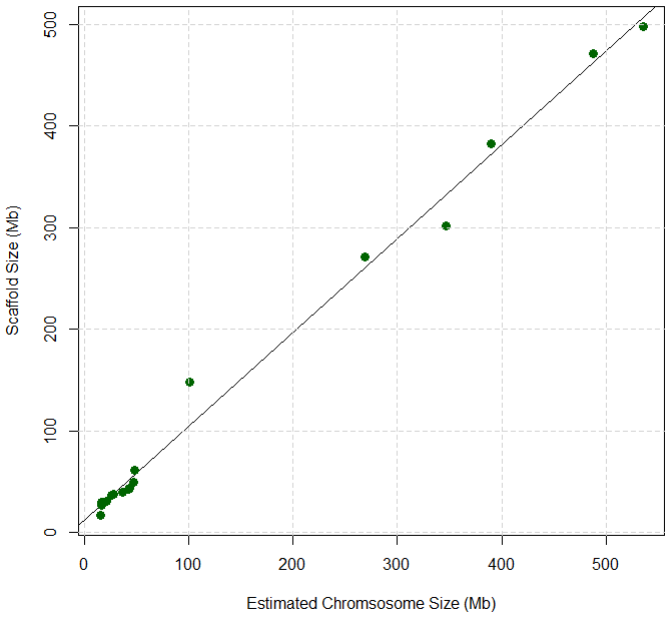
The input de novo assembly and Dovetail OmniC library reads were used as input data for HiRise, a software pipeline explicitly designed for the use of proximity ligation data for scaffold genome assemblies (Putnam et al., 2016). Dovetail OmniC library sequences were aligned to the draft input assembly using bwa (Li, 2013). The separations of Dovetail OmniC read pairs mapped within draft scaffolds were analysed by HiRise to produce a likelihood model for genomic distance between read pairs, and the model was used to identify and break putative misjoins, to score prospective joins, and make joins above a threshold.

References

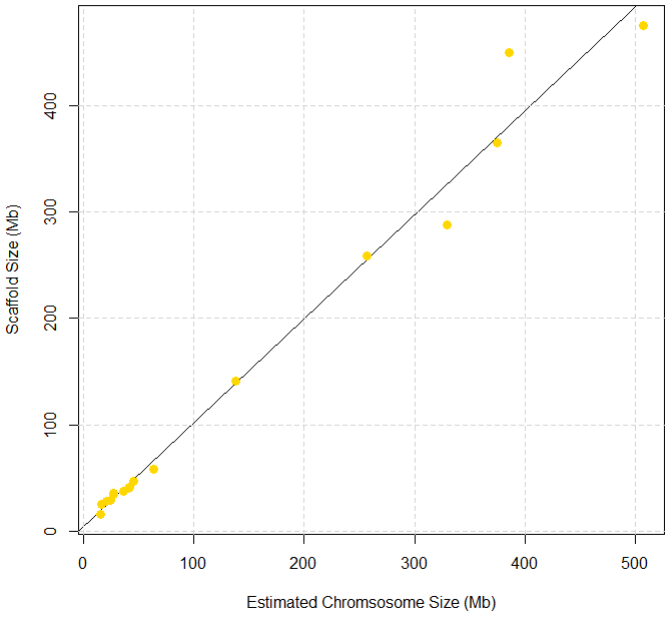
Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with bwa-mem. *arXiv preprint arXiv:1303.3997*.

Putnam, N. H., O'Connell, B. L., Stites, J. C., Rice, B. J., Blanchette, M., Calef, R., Troll, C. J., Fields, A., Hartley, P. D., Sugnet, C. W., et al. (2016). Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome research*, 26(3):342–350.

Supplementary Figures



(a) *Bradypodion pumilum* (BraPum1.0)



(b) *Bradypodion ventrale* (BraVen1.1)

Figure S1: Linear regression showing a) *Bradypodion pumilum* and b) *B. ventrale* scaffold sizes correlated with estimated chromosome sizes from published karyotype

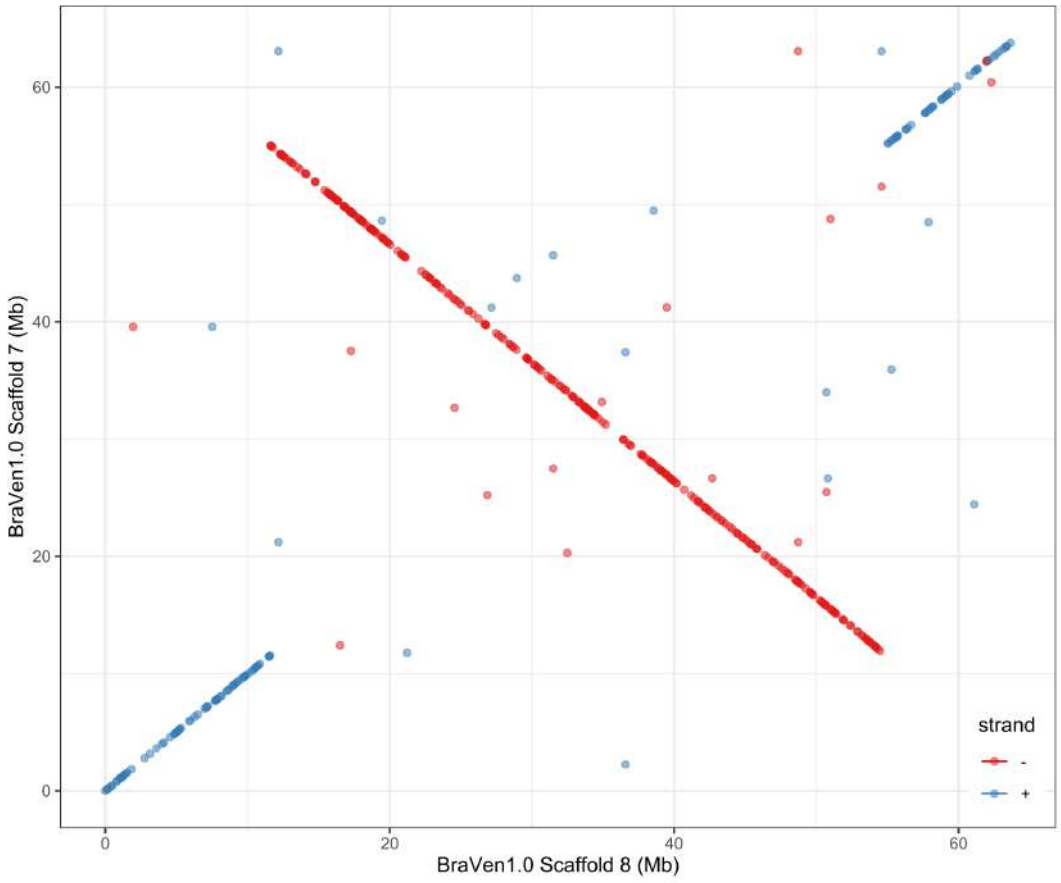
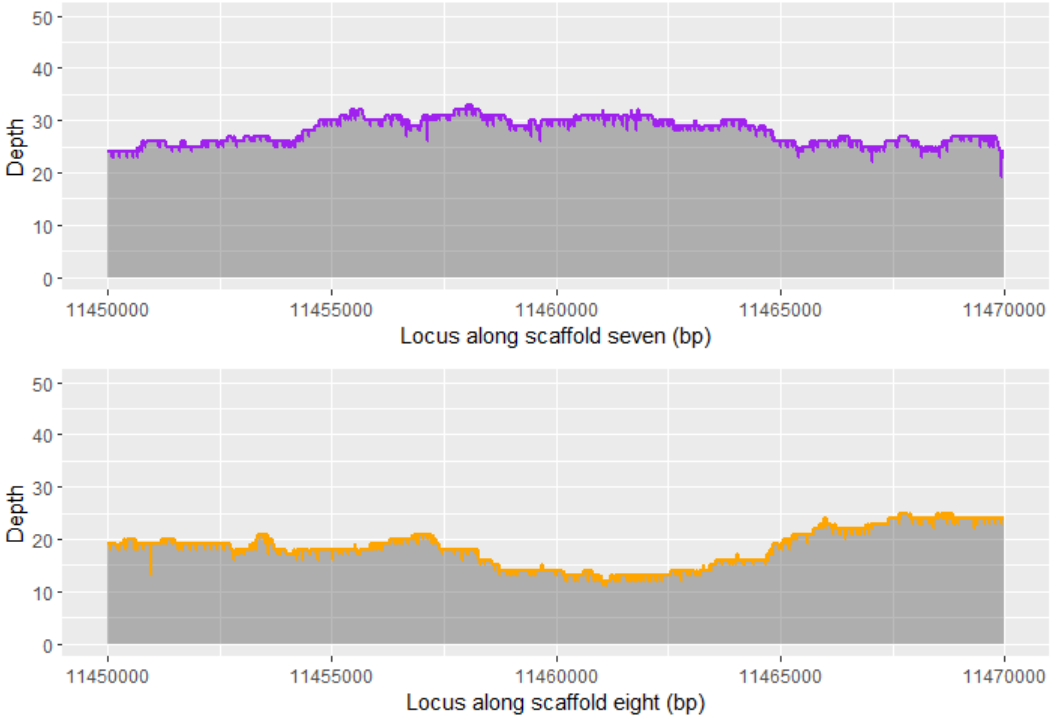
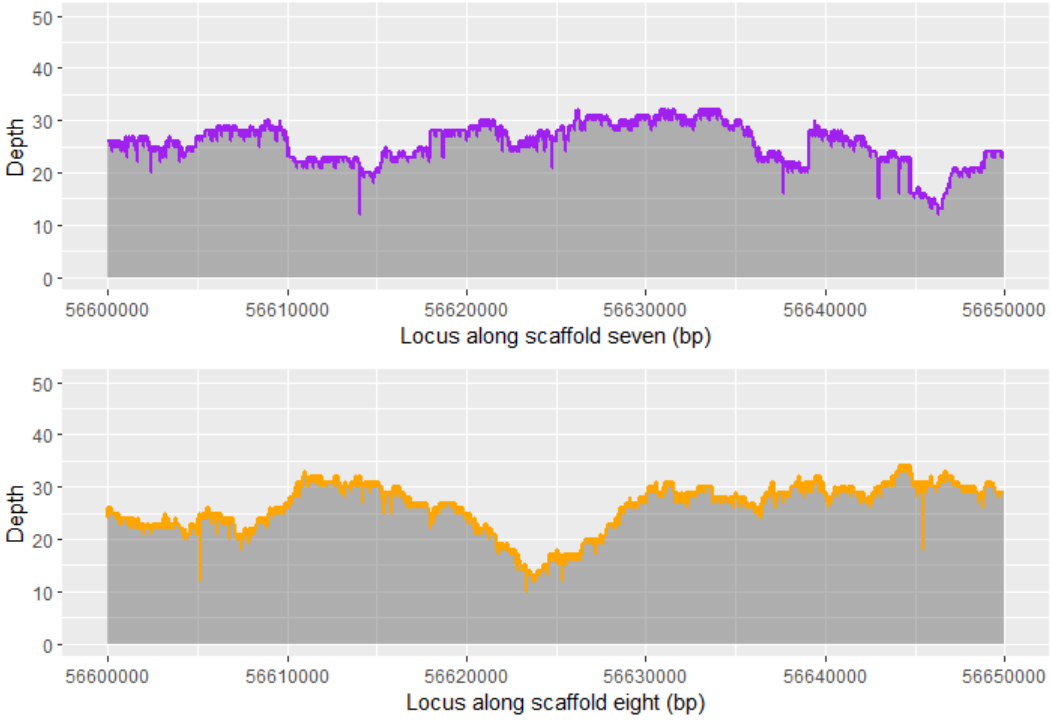


Figure S2: Inversion pattern detected in MUMmer dotplot when comparing the *Bradypodion ventrale* (BraVen1.0) scaffolds seven and eight. Blue dots indicate the same orientation of sites along the scaffold. Red dots indicate inverted orientation of sites along the scaffold



(a) Breakpoint at the front end of scaffold 8



(b) Breakpoint at the tail end of scaffold 8

Figure S3: PacBio long read coverage plots comparison plots for scaffolds seven and eight. Breakpoints along scaffold 8 in the initial *Bradypodion ventrale* (BraVen 1.0) assembly are evident by the reduced coverage "valley". Scaffold seven is indicated in purple, and scaffold eight is indicated in yellow.

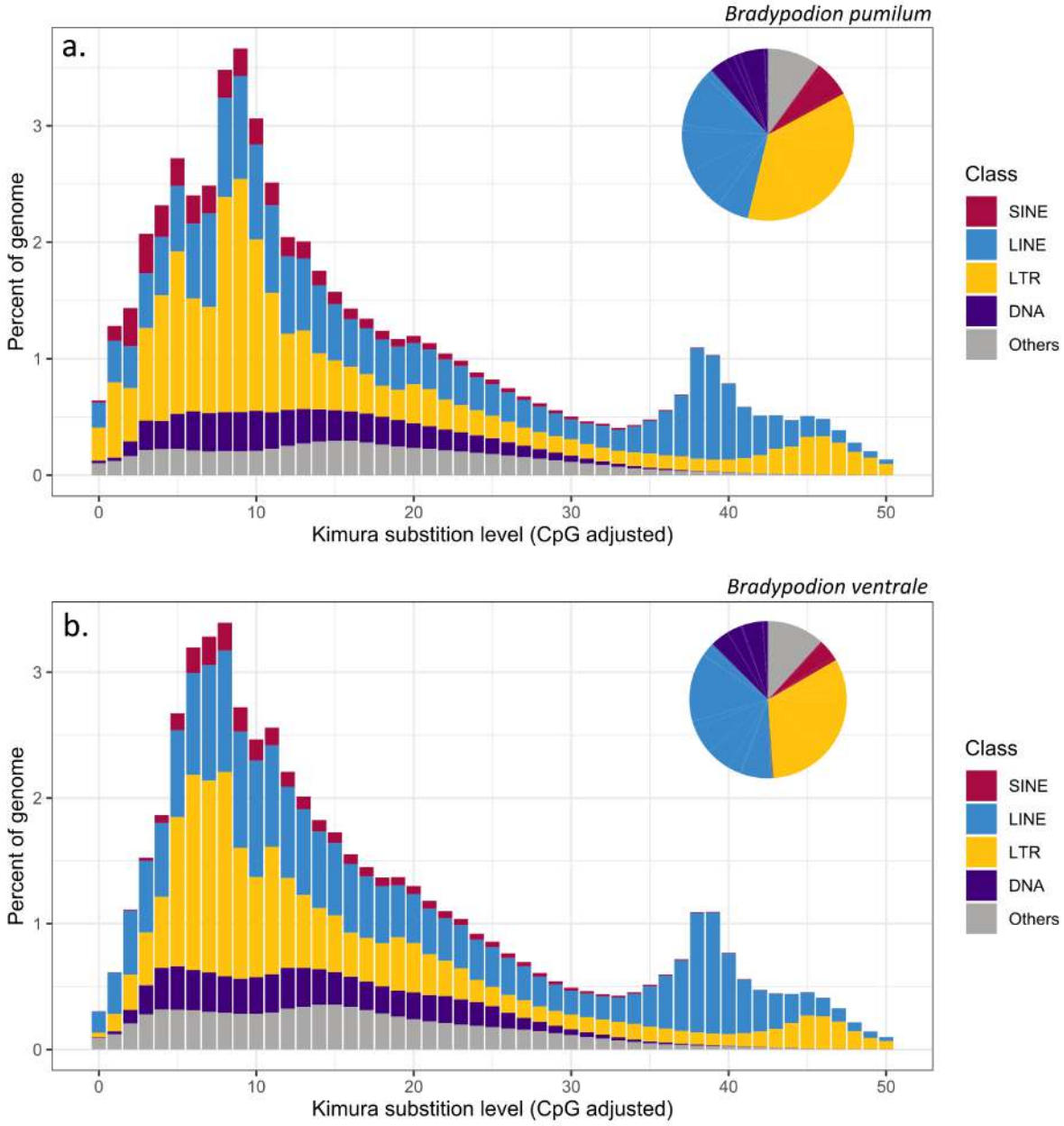
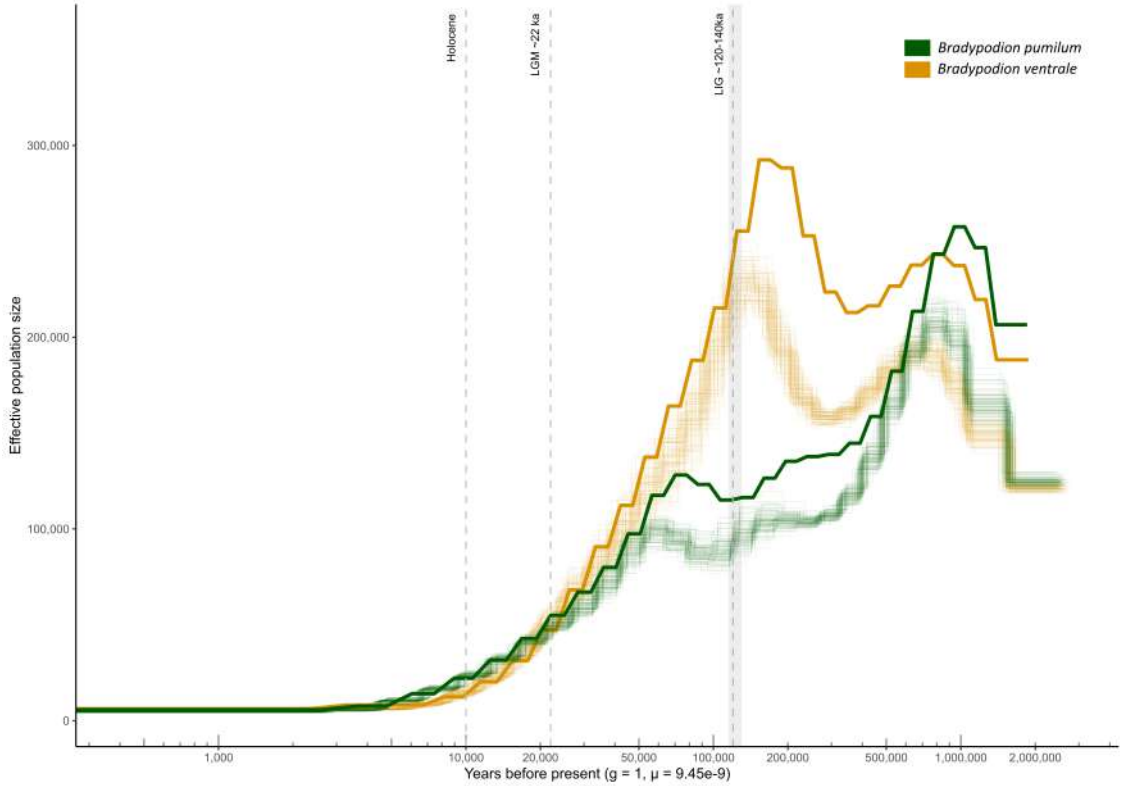
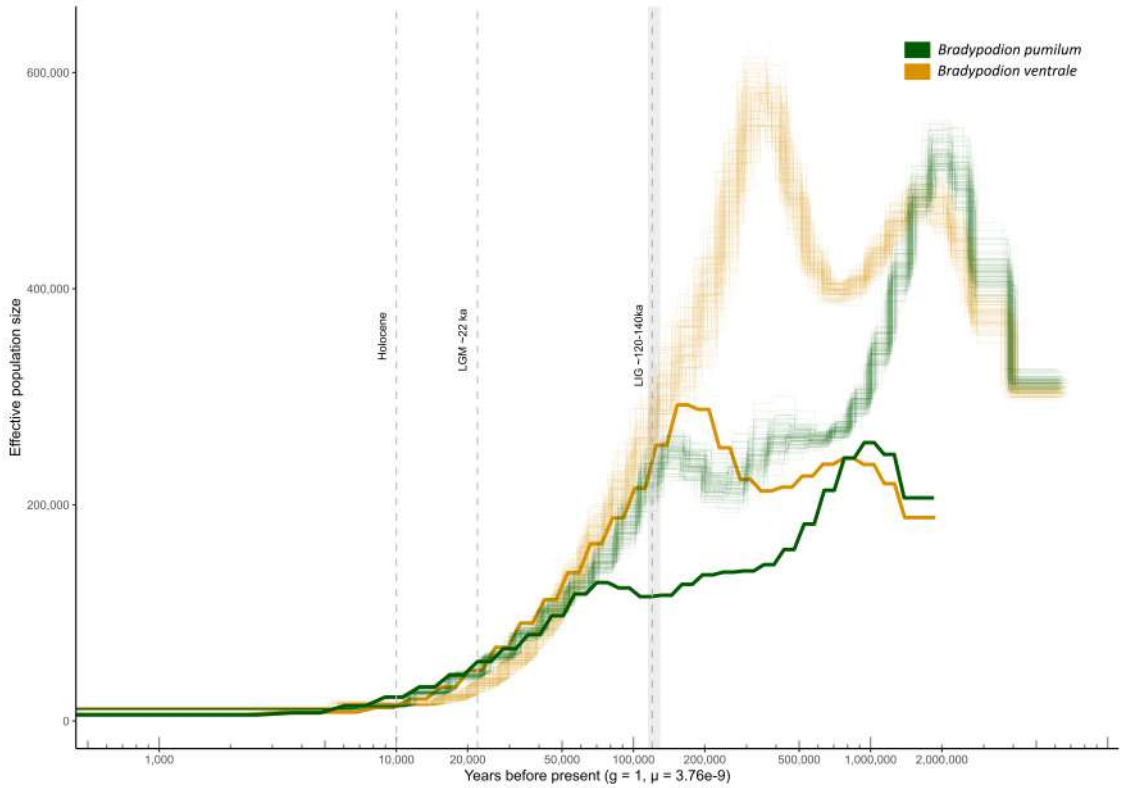


Figure S4: Comparison of repeat landscapes for the classes of transposable elements in a) *Bradypodion pumilum* and b) *Bradypodion ventrale*. The proportion of the genome consisting of transposable element insertions (short interspersed elements=SINE, long interspersed elements=LINE, long terminal repeat retrotransposons=LTR, and DNA transposons=DNA) of different ages according to their Kimura 2-parameter divergence (CpG adjusted) from consensus.



(a) minimum mutation rate: 9.45e-09



(b) maximum mutation rate: 3.76e-09

Figure S5: PSMC results using minimum and maximum mutation rates for *Bradypodion pumilum* (green) and *B. ventrale* (gold). Solid lines indicate the model results using the average mutation rate (7.17×10^{-9}). Clouded lines indicate unique bootstrapping results for a) minimum (9.45×10^{-9}) and b) maximum (3.76×10^{-9}) mutation rates

.

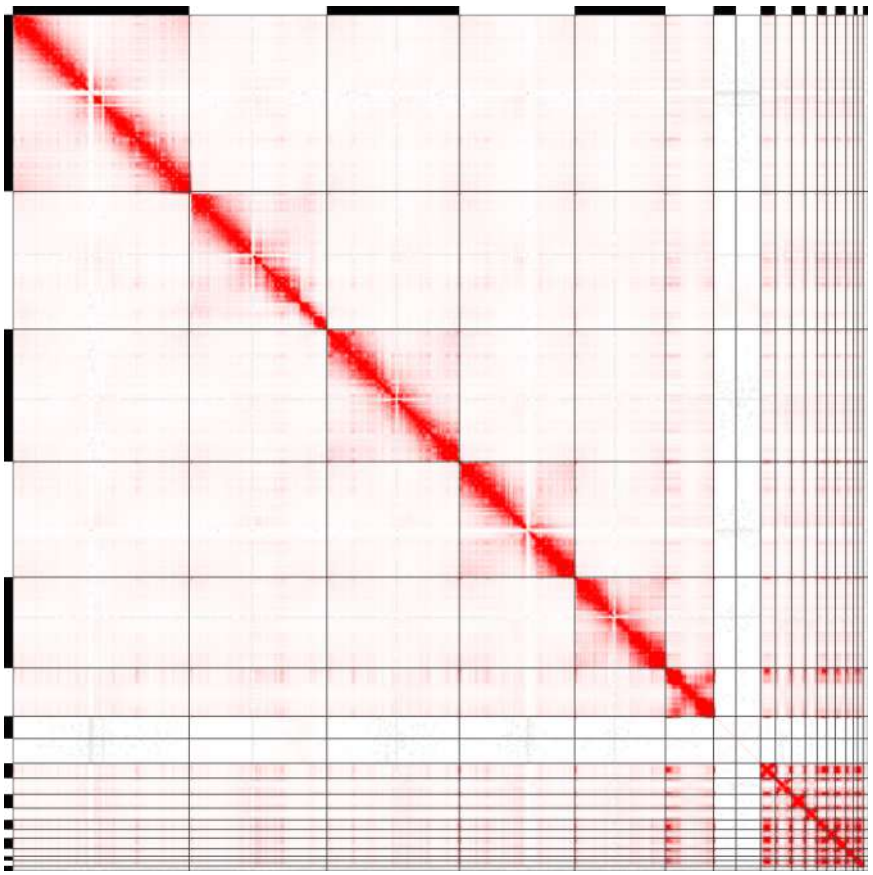


Figure S6: Link density histograms depict associations between parts of the *Bradypodion ventrale* 1.0 assembly. Increasing intensity of red indicated more read pairs map to a particular coordinate of the histogram. White shaded areas indicate no reads mapping to that coordinate in the histogram. All white areas indicate no reads mapping to scaffolds seven and eight.

Supplementary Tables

Table S1: Assembly statistics for vertebrate assemblies assessed for contiguity used in this study.

Group	Species	Accession	Sequence total (Mb)	Scaffold L50	Scaffold N50 (Mb)	Max scaffold length (Mb)	Genes Present
Reptile	*Anolis carolinensis*	GCF_000090745.1	1799.144	5	150.642	236.920	2971
Reptile	*Anolis sagrei*	GCA_025583915.1	1926.425	4	253.587	335.360	3252
Reptile	*Bradypodion pumilum*	JAWDJD000000000	2486.275	3	389.613	535.128	3271
Reptile	*Bradypodion ventrale*	JAWDJE000000000	2373.374	3	374.919	506.857	3279
Mammal	*Canis familiaris*	GCF_000002285.5	2312.802	14	63.739	122.014	3254
Fish	*Carcharodon carcharias*	GCF_017639515.1	3915.281	330	2.992	23.593	3087
Reptile	*Chamaeleo calyptratus*	PRJNA429753	-	-	-	-	3197
Reptile	*Chelonia mydas*	GCF_015237465.2	2208.394	158	3.864	22.971	3174
Reptile	*Crocodylus porosus*	GCA_001723895.1	2125.620	87	7.557	33.354	3182
Reptile	*Crotalus adamanteus*	GCA_018446365.1	1597.648	1275	0.338	3.371	3103
Reptile	*Elgaria multicarinata*	GCA_023053635.1	1790.509	7	107.859	174.634	3278
Mammal	*Felis catus*	GCF_018350175.1	2521.711	7	149.752	242.101	3264
Reptile	*Furcifer pardarlis*	-	1609.777	4.000	184.669	317.478	3282
Bird	*Haliaeetus leucocephalus*	GCF_000737465.1	1178.409	40	9.145	30.461	3173
Reptile	*Hemicordylus capensis*	GCA_027244095.1	2294.751	3	359.650	459.129	3287
Mammal	*Homo sapiens*	GCF_000001405.40	3298.431	9	145.139	248.956	3256
Reptile	*Lacerta agilis*	GCA_009819535.1	1391.404	7	86.566	133.751	3273
Reptile	*Lacerta bilineata*	GCA_900245895.1	1418.149	1150	0.368	2.880	3169
Reptile	*Lacerta viridis*	GCA_900245905.1	1439.843	661	0.663	3.630	3224
Reptile	*Laticauda colubrina*	GCA_004320045.1	2038.820	16	39.960	114.693	3163
Fish	*Latimeria chalumnae*	GCF_000225785.1	2860.575	830	0.925	10.737	3001
Reptile	*Naja naja*	GCA_009733165.1	1768.535	3	224.089	375.027	3148
Reptile	*Ophiophagus hannah*	GCA_000516915.1	1594.075	1750	0.242	2.845	2861
Reptile	*Pantherophis obsoletus*	GCA_012654085.1	1692.462	29	14.520	59.843	3083
Reptile	*Paroedura pictus*	GCA_003118565.2	1562.176	6	109.005	184.259	2943
Reptile	*Phrynosoma platyrhinoss*	GCA_020142125.1	1901.840	3	273.213	396.191	3051
Reptile	*Plestiodon gilbertis*	GCA_026170595.1	1447.230	4	184.425	267.325	3029
Reptile	*Podarcis muraliss*	GCA_004329235.1	1511.020	7	92.398	140.210	3239
Reptile	*Pogona vitticepss*	GCA_900067755.1	1716.675	199	2.478	14.681	3213
Reptile	*Python bivittatuss*	GCA_000186305.2	1435.035	1939	0.214	1.453	3005
Amphibian	*Pyxicephalus adspersuss*	GCA_004786255.1	1563.368	5	157.522	194.783	3049
Mammal	*Rattus rattuss*	GCF_011064425.1	2250.262	6	145.848	260.810	3229
Reptile	*Salvator merianaes*	GCA_003586115.2	2068.170	12	55.382	154.337	3277
Reptile	*Sceloporus undulatuss*	GCA_019175285.1	1905.484	3	275.632	382.616	3190
Reptile	*Shinisaurus crocodilurus*	GCA_021292165.1	2189.995	4	296.945	321.224	3117
Reptile	*Sphenodon punctatus*	GCA_003113815.1	4272.215	370	3.053	29.980	3192
Bird	*Struthio camelus*	GCA_000698965.1	1225.025	101	3.593	19.376	2955
Bird	*Taeniopygia guttata*	GCA_000151805.2	1232.119	7	62.375	156.413	2953
Reptile	*Thamnophis elegens*	GCA_009769535.1	1672.190	6	100.852	186.725	3153
Reptile	*Varanus komodoensis*	GCA_004798865.1	1507.946	17	23.832	138.280	3224
Reptile	*Varanus salvator*	GCA_023646645.1	1702.542	9	71.462	121.089	3213
Reptile	*Vipera berus*	GCA_000800605.1	1532.391	3408	0.126	2.584	2781
Amphibian	*Xenopus laevis*	GCA_017654675.1	2742.469	8	155.251	233.740	3296
Reptile	*Zootoca vivipara*	GCA_011800845.1	1464.301	7	92.810	131.769	3275

Table S2: Comparison of the interspersed repeat contents of *Bradypodion pumilum* and *Bradypodion ventrale*.

Repeat Class	***Bradypodion pumilum***		***Bradypodion ventrale***	
	Length occupied (bp)	**Percent of Genome**	**Length occupied (bp)**	**Percent of Genome**
SINEs	100 066 638	4.02	61 988 050	2.61
LINEs	487 041 094	19.59	513 854 029	21.65
LTR Elements	541 014 879	21.76	443 823 189	18.7
DNA transposons	164 429 416	6.61	158 813 066	6.69
Unclassified	139 767 023	5.62	157 935 903	6.65
Total Interspersed Repeats	1 432 319 050	57.61	1 336 414 237	56.31
Total Repeats	1 518 034 260	61.06	1 413 185 695	59.54

Chapter 3

Exploring Gene Family Evolutionary Dynamics in *Bradypodion*

Abstract

Adaptation via natural selection underpins the phenotypic diversity seen throughout the tree of life, yet our understanding of the genomic mechanisms behind adaptation remains limited. Recent advancements in technology have enabled the exploration of alternative, emerging genomic characteristics such as gene family expansion, which could potentially drive adaptive evolution. The southern African dwarf chameleons, *Bradypodion*, are known for their highly adaptive capacity in response to habitat change. However, the underlying genetic architecture of adaptive traits remains unknown. With the recently assembled genomes for *Bradypodion*, I assess the gene family dynamics of these lizards in order to uncover whether changes in gene family composition drive adaptive diversification in *Bradypodion*. First, I provide annotations for both Bradypodion assemblies, where gene model predictions yielded 36 266 protein-coding sequences in the *B. pumilum* annotation, with 37 828 sequences in *B. ventrale*. Comparisons of gene models to the BUSCO vertebrate dataset suggested that genome annotation completeness was 73.49% and 75.58% for *B. pumilum* and *B. ventrale* assemblies respectively. These annotations were utilized to assess gene family evolution within *Bradypodion* and among Iguania representative species. Notable gene family expansions detected at the *Bradypodion* ancestral node were associated with ontology terms relating to responses to light and ultra-violet radiation stimuli and developmental processes. In addition, gene family expansions were notably different between *B. pumilum* and *B. ventrale*. While *Bradypodion pumilum* contained similar expansions to the *Bradypodion* ancestral node, gene family expansions in *B. ventrale* were associated with visual perception of light stimuli and tissue development, particularly skin and bone development. Given the contrasts in gene family composition between *B. pumilum* and *B. ventrale*, diversification in *Bradypodion* in response to environmental change was likely facilitated by the adaptive traits associated with the gene family dynamics detected here.

Keywords: dwarf chameleons, genome annotation, CAFE, comparative genomics

3.1 Introduction

The diversity of forms present across the tree of life is underpinned by genetic changes inherited from ancestor to descendant. The process of natural selection favours those phenotypic changes that increase an individual's fitness. Mutations are at the core of phenotypic changes and these genetic changes come in many forms ranging from modifications of single nucleotides to the wholesale duplication of entire genomes (Kaessmann, 2010). With numerous genome projects contributing tens of thousands of nucleotide sequences to public databases (Genome 10K Community of Scientists, 2009; Ebenezer et al., 2022), uncovering genomic evolutionary histories of non-model organisms is more accessible than ever (Ellegren, 2014). This is due to the increasing accuracy and affordability of whole genome assembly and the development of high-throughput sequencing technology (Ellegren, 2014). Although, even with these recent innovations, identifying the underlying genes and mutations of putative adaptive phenotypes still requires considerable investigation (Benfey and Mitchell-Olds, 2008; Henry et al., 2011). Especially when candidate variants might be in genes that have not yet been characterized (Shefchek et al., 2020; Bomblies and Peichel, 2022).

While the number of available sequenced genomes continues to increase, there is still a need to fully assess the complex organization of genetic material (Jauhal and Newcomb, 2021; Rhie et al., 2021). The natural complexity of the genome structure makes it challenging to associate DNA sequences with particular phenotypes (Bradnam et al., 2013; Chaisson et al., 2015; Jauhal and Newcomb, 2021). The variation in the genomic landscape further contributes to this challenge, i.e., changes in the types and abundance of transposable elements or composition of repetitive elements (Ayarpadikannan and Kim, 2014; Bourque et al., 2018; Shao et al., 2019; Whibley et al., 2021). Additionally, complex phenotypes may often result from the combined interactions of multiple genes (Glazier et al., 2002; Cooper et al., 2005; Werth, 2021), as well as the impact of environmental factors (Yan et al., 2022), and the differential expression of genes (Shyamli et al., 2021; Valenza-Troubat et al., 2022). While exploring genes related to environmental factors might require an integrated experimental approach (Kawecki and Ebert, 2004; Hargreaves et al., 2020), having access to whole genome sequences allows for inference on the evolutionary processes affecting gene evolution. Given the rapid increase in sequenced genomes, efforts towards providing accurate annotations of genomic features facilitate steps toward uncovering the genetic architecture of lineage-specific traits and adaptations (Primmer et al., 2013; Ekblom and Wolf, 2014).

Many genome-sequencing initiatives involving non-model organisms have revealed lineage-specific genes likely underlying the evolutionary adaptations of phenotypic traits (Qiu et al., 2012; Perry et al., 2018; Lind et al., 2019; Gemmell et al., 2020). A common feature in discovering lineage-specific genes has been the observations of expanded gene families (Das et al., 2020; Ferrari et al., 2023). Gene families are those sets of genes that share a common evolutionary origin (Guo, 2013), the size and composition of which vary due to multiple factors (Hahn et al., 2007; Freeling, 2009; De Smet and Van de Peer, 2012; Ellegren, 2014). To begin with, gene duplication or whole-genome duplication augments gene family size, and contributes to genome evolution at a similar scale as the mutation rate per nucleotide site (Lynch and Conery, 2003; Flagel and Wendel, 2009; Van de Peer, 2011). Moreover, gene family size can be reduced either by the deletion of a single gene or several genes within a segmental fragment. Finally, the *de novo* formation of genes may create new gene families (McLysaght and Hurst, 2016). Given the continued growth in the number of genomes for non-model organisms, the power of estimating changes in gene family evolution between closely related species has increased substantially Trouern-Trend et al. (2020); Lin et al. (2021); Sumner et al. (2023).

A significant concern when investigating changes in gene family size is the quality of the underlying genome assembly and the accompanying annotation (Card et al., 2023; Demuth and Hahn, 2009). While long-read sequencing technologies have greatly exceeded the limitations of short-read sequencing approaches (formerly at $\tilde{0}.1\%$ error rate and only $\tilde{5}00$bp amplicon length; Weirather et al., 2017; Hu et al., 2021; Karst et al., 2021), there still exists concerns in terms of sequencing coverage, inaccurate assembly approaches, or sequencing platform related errors (Hahn et al., 2007; Han et al., 2013). These issues can lead to both the erroneous addition and subtraction of genes in datasets generated by this technology (e.g., Hubisz et al., 2011). A direct consequence of these types of errors occurs when trying to annotate genome assemblies (Ejigu and Jung, 2020). Generating a successful annotation is highly dependent on the quality of the underlying genome assembly (Denton et al., 2014; Stuart et al., 2022). Unfortunately, generating annotations has not kept pace with sequencing and assembly of new genome assemblies (Shumate and Salzberg, 2021).

There has been considerable effort put into producing whole genome sequences for a wide variety of vertebrate species (Genome 10K Community of Scientists, 2009; Koepfli et al., 2015; Lewin et al., 2018; Blaxter et al., 2022; Ebenezer et al., 2022). However, there is still a substantial scarcity of genome assemblies for reptiles (Card et al., 2023), even though this

diverse clade of extant tetrapods has over 12,000 recognized species (Uetz et al., 2023). Given the overall lack of genome annotations relative to all sequenced assemblies, it is no surprise that accurate annotations for reptile assemblies are also needed (Gable et al., 2023). The value of genome assemblies is greatly increased with access to highly accurate annotations, identifying biologically relevant information such as gene models and functionally relevant elements (Yandell and Ence, 2012; Ekblom and Wolf, 2014). This is especially true in non-model organisms, as annotations are typically restricted to protein-coding sequences or transcripts more generally (Ekblom and Wolf, 2014).

Chameleons are one of the least represented groups among reptiles with a publicly available genome assembly (Xie et al., 2022; Taft et al., 2023). Of the 220 species in the family, there are only three assemblies to date, none of which are annotated. Thus, the genetic basis of their unique physiological and morphological characteristics remains unexplored. Members of Chamaeleonidae are a phenotypically distinct family of lizards with specialized traits rarely seen in other reptiles, such as fused, opposable digits on their hands/feet for grasping, projectile tongue, independently moving eyes, and a prehensile tail (Boistel et al., 2010; Higham and Anderson, 2013; Diaz et al., 2015; Pinto et al., 2019). Ecologically, chameleons have undergone evolutionary shifts from inhabiting the forest floor to becoming highly adapted for an arboreal lifestyle (Bickel and Losos, 2002; Tolley et al., 2013), which has entailed several major adaptations in morphology and ecophysiology (Tolley and Herrel, 2013; Giles and Arbuckle, 2022), e.g., a range in body sizes (Diaz and Trainor, 2015) and diverse reproductive life histories (Measey et al., 2013). The southern African dwarf chameleons (*Bradypodion*) are the most recently diverged chameleon lineage known to have undergone rapid diversification linked to changes in natural environments through ecological speciation (Tolley et al., 2008; Hopkins and Tolley, 2011; da Silva and Tolley, 2013; Tolley et al., 2022). Additionally, phenotypic ecomorphs exist within *Bradypodion*, presumably the result of multiple convergence events. However, it remains unclear whether convergence for particular phenotypes is a result of standing genetic variation (i.e., plasticity - but see Miller and Alexander, 2009), or due to convergent novel mutations independent of their phylogenetic history.

Given that two of the three existing Chamaeleonidae genome assemblies are of *Bradypodion*, the opportunity to explore gene family dynamics may shed light on the evolutionary history underlying the diversification of this group. Here, I provide high-quality annotations of two *Bradypodion* genome assemblies, describe gene landscapes in both assemblies, and use these

annotations to evaluate the gene family evolution of *Bradypodion* relative to other members of the superfamily Iguania (Chamaeleonidae, Agamidae, Iguanidae). Should this comparative approach reveal significant changes in gene family evolution in *Bradypodion*, it would suggest unique genomic adaptations exist within dwarf chameleons, potentially linked to adaptations in response to changing environmental factors and ecological niches. Conversely, if no significant differences in gene families are found, it is likely that diversification within Iguania is driven by genomic features unique to each family.

3.2 Methods

3.2.1 Genome Annotation

The genome assemblies of *Bradypodion pumilum* and *B. ventrale* were annotated following the Dovetail Genomics annotation pipeline (Jarvis et al., 2017; Taft et al., 2023). Species-specific repeat models were first identified *de novo* and classified using RepeatModeler v2.0.3 (Flynn et al., 2020), RECON v1.08 (Bao and Eddy, 2002), and RepeatScout v1.0.6 (Price et al., 2005). Once the custom repeat library for each assembly was generated, these were used as input to identify and mask repeats in the original assembly using RepeatMasker v4.1.0. Coding sequences from publicly available squamate assemblies were used as input (*Anolis sagrei*, *Pogona vitticeps*, *Chamaeleo calyptratus*, *Salvator merianae*) to train the initial *ab initio* model for each *Bradypodion* annotation using AUGUSTUS v2.5.5, followed by six rounds of prediction optimization. The same additional coding sequences were used to train separate *ab initio* models for each *Bradypodion* assembly using SNAP v.2006-07-28. RNAseq reads were generated with tissue sampled from the eye, muscle, skin (flank, gular, and ventral), and testes of each genome animal were mapped onto the respective genome assembly using the STAR aligner v2.7 (Dobin et al., 2013), and intron hints were generated using the AUGUSTUS bam2hints tools. Gene models for each *Bradypodion* assembly were then estimated using MAKER v3.01.03, SNAP and AUGUSTUS (with intron-exon boundary hints provided from RNA-Seq) to predict genes in the repeat-masked reference assembly. The Swiss-Prot peptide sequences from the UniProt data were used in conjunction with the protein sequences of (list additional assemblies) to provide peptide evidence which assisted the prediction process in MAKER. Gene models predicted by both SNAP and AUGUSTUS were retained in the final gene sets. The quality of gene prediction was assessed by generating AED scores for each gene model in MAKER. Genome annotation summary statistics were generated using Genome Annotation Generator (GAG; Geib et al.,

2018). Gene annotations were evaluated for gene completeness with BUSCO (Benchmarking Universal Single-Copy Orthologs) v5.4.4 (Simão et al., 2015) using the vertebrata odb10 dataset.

3.2.2 Genomic feature landscape

Patterns in the predicted annotations for both *Bradypodion pumilum* and *B. ventrale* assemblies were assessed by sliding windows. Sliding windows were calculated using BEDTools v2.26.0 with 1 Mb window sizes to estimate gene density. The repetitive element annotations generated in Taft et al., (2023) (Chapter 2) were used as input to summarise the density of repetitive elements for both *Bradypodion* assemblies, with window sizes set at 10 Kb.

3.2.3 Orthologous gene families

Four additional genome assemblies and accompanying annotations were downloaded from the NCBI genome database (https://www.ncbi.nlm.nih.gov/genome/), namely *Furcifer pardalis* (Chamaeleonidae), *Pogona vitticeps* (Agamidae), *Laudakia sacra* (Agamidae), and *Anolis sagrei* (Dactyloidae), to estimate orthologous gene families. Amino acid sequences for all protein-coding genes for each assembly were extracted using GFFRead v0.12.8 and then filtered for the longest transcript variant per gene. Filtered sequences were used as input to phylogenetically predict orthologous proteins (orthogroups) with Orthofinder v2.5.4. Orthofinder categorizes gene families as sets of homologous genes that stem from a single gene in the last common ancestor of the input species and remain reliable even with varying gene lengths or larger phylogenetic distances (Emms and Kelly, 2015). Multiple sequence alignment was done using the '-M msa' function within Orthofinder aligning sequences using MAFFT (Katoh and Standley, 2013), before inferring both species and gene trees using IQTree v2.1.4-beta.

3.2.4 Likelihood analysis of gene family gain and loss

Gene family gain and loss were estimated using CAFE v 5.0 (Mendes et al., 2020), a tool to statistically analyse the evolution of the gene family size under a phylogenetic framework (De Bie et al., 2006). The species tree generated in the Orthofinder analysis was used as input after confirming the tree is rooted and ultrametric with the R package 'ape' (Paradis et al., 2004). The phylogenetic hierarchical orthogroups (gene families) from the Orthofinder analysis were used to retrieve gene counts per gene family per taxon as input for CAFE. These gene families were first filtered by removing gene families with gene counts exceeding

100 in any single taxon to prevent biases in gene family evolution estimation (Emms and Kelly, 2019). The lambda parameter was calculated uniformly across the phylogeny and gene families with a large size variance among taxa were selected using a significance threshold of $P \leq 0.01$. Estimated changes in gene families were subsequently plotted with CafePlotter v0.1.1 (https://github.com/moshi4/CafePlotter).

3.2.5 Functional enrichment analysis

To ascertain the functions associated with *Bradypodion* gene families, the underlying genes within each gene family were tested for functional Gene Ontology (GO) terms. Functional classes were estimated using g:GOSt in g:Profiler to perform the functional enrichment test for the gene list derived from significant gene families in the CAFE analysis. Significant gene families were included for further analysis where the estimated gene family change resulted in $\geq$ 80% increase in family size. The tools available within g:Profiler are frequently used in standard pipelines of biological entity (gene/protein) centered computational analysis (Raudvere et al., 2019; Kolberg et al., 2023), with g:GOSt able to conduct the functional enrichment analysis of individual or multiple gene lists. GO terms were assigned to the list of genes based on the Ensembl GO predictions for *Anolis carolinensis* used as the reference organism. P-values were adjusted using Benjamini and Hochberg's approach for controlling the false discovery rate (FDR) for a significance threshold of $P < 0.05$. Significantly enriched GO terms were returned unordered. Further filtering of significant terms within g:GOSt was done to ensure unique leading terms in each GO component (molecular function, biological process, cellular component) by removing bias for terms with the highest significant levels (Kolberg et al., 2023).

3.3 Results

3.3.1 Genome annotation summary statistics

Protein-coding sequences in the *Bradypodion pumilum* annotation predicted 36,266 sequences, for which 16,917 showed homology with genes from other species (Table 3.1). Comparatively, predictions for the *B. ventrale* yielded a greater number of protein-coding sequences (37,828) with 17,712 being homologous to previously annotated genes. Genes for *B. pumilum* were estimated at an average length of 12.22 Kb, with a mean of five exons and four introns per gene. Average gene length estimations for *B. ventrale* were 12.58 Kb, with the same mean number of exons and introns per gene at five and four respectively. Of the 3354 total

Table 3.1: Summary of the annotation features and BUSCO assessment statistics for each *Bradypodion genome* assembly.

	Bradypodion pumilum	*Bradypodion ventrale*
Number of genes	36528	37828
Number of annotated genes	16917	17712
Number of mRNAs	36528	37828
Number of exons	183198	193906
Number of introns	146671	156079
Number of CDS	36266	37447
mean gene length (bp)	12229	12588
mean mRNA length (bp)	12229	12588
mean exon length (bp)	219	218
mean intron length (bp)	2774	2782
mean CDS length (bp)	1103	1126
% of genome covered by genes	18	20
Complete BUSCOs	2465	2535
Single copy BUSCOs	2452	2512
Duplicated BUSCOs	13	23
Fragmented BUSCOs	288	275
Missing BUSCOs	601	544
Total BUSCOs searched	3354	3354

BUSCOs available in the vertebrate odb10 database, assessment of gene completeness recovered 2465 (73.49%) complete BUSCOs in the *B. pumilum* annotation of which 2452 are single-copy with 13 duplicates. Gene completeness was marginally greater in *B. ventrale*, with 2535 complete BUSCOs recovered (75.58%) with 2512 in single-copy and 23 duplicates (Table 3.1). Visual inspection of the gene density along chromosomes revealed similar distributions for the 6 largest chromosomes in both *Bradypodion* annotations (Figure S1-S2). However, inspection of the smaller chromosomes revealed an inverse distribution of genes on chromosome 9 as peak gene densities in *B. pumilum* occur toward the telomeres, but peak toward the centromere in *B. ventrale*. Repeat densities are similarly distributed along chromosomes in both *Bradypodion* (Figures S3-S4), however, excess repeats are more evident in B. ventrale, particularly on chromosomes two, three, eight - ten and 17.

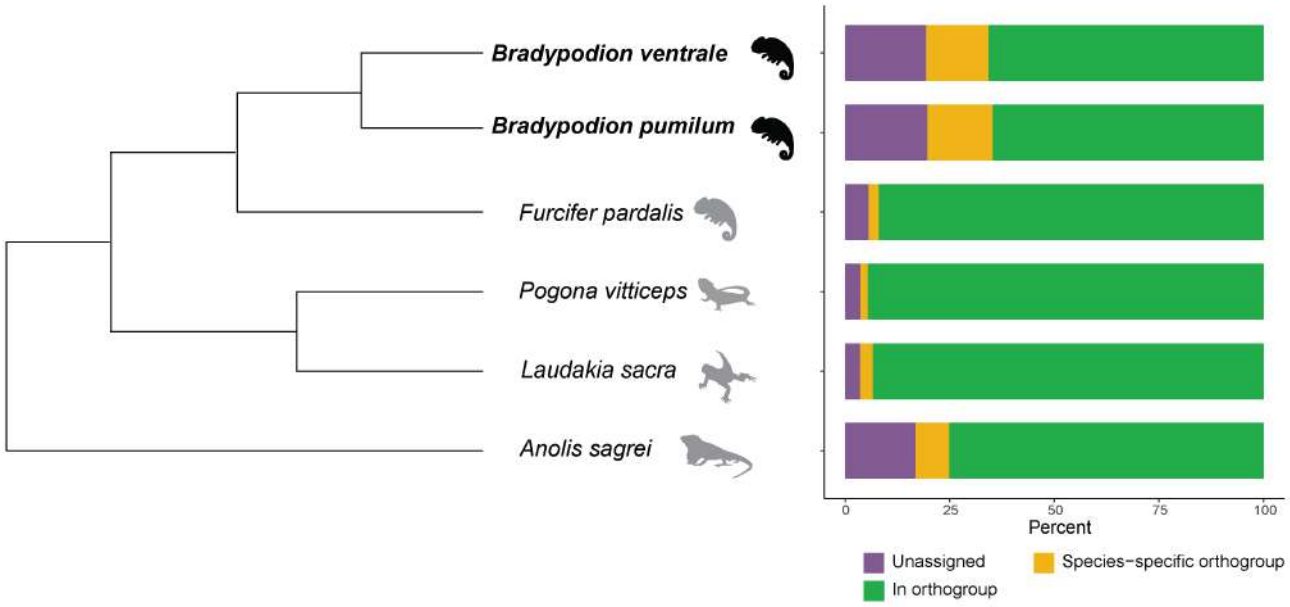


Figure 3.1: Percentage of protein-coding genes assigned to orthogroups in candidate species. The phylogenetic species tree is based on reconciled gene trees generated by Orthofinder. All organism silhouettes are from PhyloPic (www.phylopic.org)

3.3.2 Orthologous gene families

Designation of gene families based on the six input species resulted in 134,083 of 154,669 total genes (86.7%) being assigned to 22,445 orthogroups. A total of 14,469 genes (9.4%) are in 3765 species-specific orthogroups. Only 5731 orthogroups contain genes from all species included in the analysis. Orthogroups range from 2 - 486 genes per group, with 97.9% of orthogroups containing ≤ 20 genes from at least two species. While *Bradypodion pumilum* and *B. ventrale* had the greatest number of genes present, 36,266 and 37,447 genes respectively, roughly 80% of genes were assigned to orthogroups–the lowest gene assignment rate relative to other species (Figure 3.1). Both *Laudakia sacra* (96.4%) and *Pogona vitticeps* (96.3%) have the highest gene assignment to orthogroups, followed by *Furcifer pardalis* (94.4%).

3.3.3 Likelihood analysis of gene family gain and loss

The evaluation of gene family change among the representative species successfully detected significant changes between taxa (Figure 3.2). At the Acrodont ancestral node (Figure 3.2B), 107 gene families showed significant changes (19 families expanded, 88 families contracted). The comparison between Agamidae and Chamaeleonidae ancestral nodes revealed a greater number of significant gene family changes in Agamidae, with 71 expansions and 346 contractions, relative to Chameleonidae (38 expansions and 52 contractions). *Furcifer pardalis* had the greatest number of total gene family changes for chameleons relative to *Bradypodion*,

with 150 expansions and 516 contractions across 666 gene families. At the *Bradypodion* ancestral node, 233 gene families experienced significant changes. Both species of *Bradypodion* were estimated to have roughly double the number of significant changes, with *B. pumilum* exhibiting 452 significant changes (142 expansions and 310 contractions) relative to *B. ventrale* with 416 total changes (131 expansions and 285 contractions).

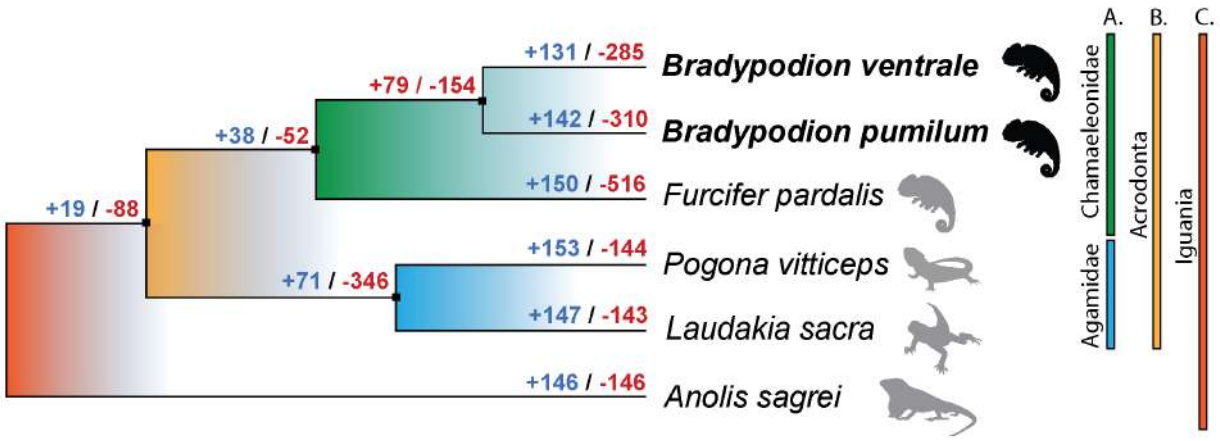


Figure 3.2: The number of significant gene family changes across the phylogenetic species constructed with Orthofinder, p-value $\leq$ 0.01. Significant gene expansion (blue) and contractions (red) are presented at each node and leaf. Coloured bars indicate branches belonging to A) Chamaeleonidae (green) and Agamidae (light blue), B) Acrodonta, and C) Iguania (orange). All species silhouettes are from PhyloPic (www.phylopic.org)

3.3.4 Functional enrichment analysis

Within significantly expanding gene families of *Bradypodion*, more than half of the annotated genes were functionally ambiguous with no identified function (*Bradypodion* ancestral node, 63 %; *B. pumilum*, 55 %; *B. ventrale*, 60 %). For functionally annotated genes, gene families at the *Bradypodion* ancestral node yielded 51 unique genes, with *B. pumilum* families holding 34 genes, and 30 genes across *B. ventrale* gene families. When assessing gene ontology terms, both molecular function (MF) and cellular component (CC) GO terms are largely conserved across *Bradypodion*, with MF terms involving molecular binding processes (GO:0005488) and CC terms relating to cellular anatomical features (GO:0110165) (Figure S5, S6, S7). In contrast, ontology terms linked to the biological process (BP) are the most dynamic within *Bradypodion* (Figure 3.3). Expanded gene families within *Bradypodion* contribute toward three broad BP ontology terms: response to stimulus (GO:0050896), cell signalling (GO:0023052), and the development process (GO:0032502)–specifically anatomical structure

(GO:0048856), system development (GO:0048731), and organ development (GO:0048513) (Figure 3.3 A). Specific pathways underlying response to stimuli are primarily involved with responses to light (GO:0009416) and radiation (GO:0009314), specifically UV-B (GO:0010224), but also responses to temperature stimuli by thermotaxis (GO:0043052). Expansion in gene families relating to anatomical structure development is the only broad BP term retained in *B. pumilum* similar to the ancestral node. Unique terms within *B. pumilum* are linked to cellular component organization (GO:0071840) and cell localization (GO:0051179) but also maintain pathways involving positive regulation of the target of rapamycin (TOR) signalling, specifically TORC1 (GO:1904263) which promotes nutrient and hormonal sensing. *Bradypodion ventrale* was shown to maintain the pathways relating to responses to light stimuli, but directly involve visual (GO:0007601) and sensory perception (GO:0050953) terms. Developmental processes are also defined by unique terms in *B. ventrale*, particularly tissue (GO:0009888) and organ development (GO:0048513), which includes muscle (GO:0061061), skin (GO:0043588), kidney (GO:0001822) and gonad (GO:0008406), as well as terms relating to blood vessels (GO:0001570, GO:0002040, GO:0001525). Anatomical structure development largely includes terms involving bone development such as osteoblast differentiation (GO:0001649) and ossification (GO:0001503).

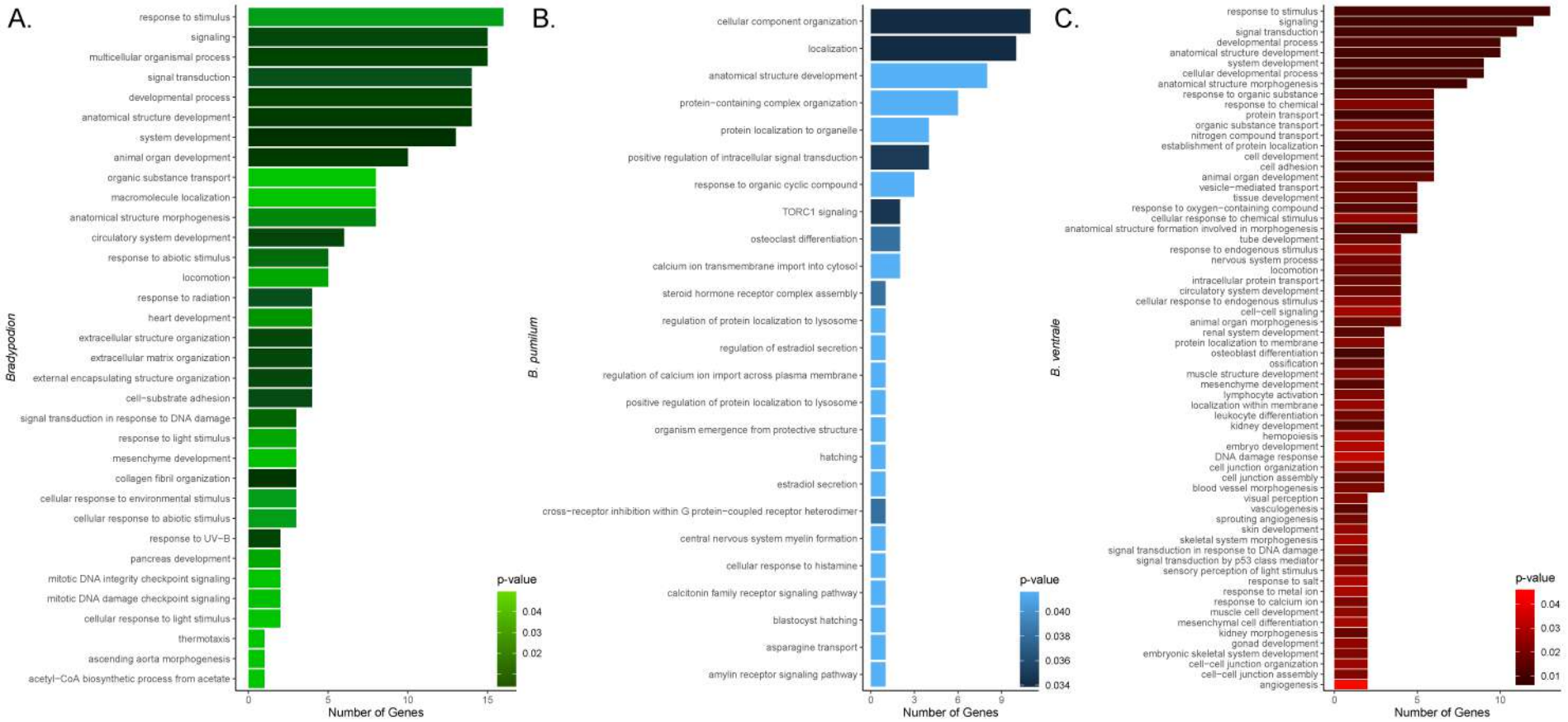


Figure 3.3: Functional gene ontology (GO) terms for biological processes in *Bradypodion*. GO terms are provided for A. *Bradypodion* ancestral node, B. *B. pumilum*, and C. *B. ventrale*

3.4 Discussion

Gene family expansions in *Bradypodion* appear to be associated with adaptations unique to the Chamaeleonidae relative to other members of Iguania. Expanded gene families within this genus are likely to have resulted from changing environmental conditions, facilitating the diversification of *Bradypodion*. In addition, changes in gene family composition indicate differences at the species level, as *B. pumilum* and *B. ventrale* exhibited expansions in gene families with distinct genetic pathways. The underlying genome annotation generated for each assembly yielded a similar structural composition, number of identified protein-coding genes, and gene and repetitive elements distribution. The overall completeness of these annotations was also comparable, considering the number of genes able to be identified by the annotation algorithms used. These quality annotations enabled the appropriate identification of orthologous groups as well as notable differences in species-specific orthogroups, the greatest of which were present in these dwarf chameleons. Therefore, these results set the stage for novel insights to be made into the genetic underpinnings of *Bradypodion* diversification and adaptive solutions unique to dwarf chameleons.

The expansion of gene families in *Bradypodion* underlying responses to light, radiation and thermotaxis could be an adaptive response driven by the different exposures of abiotic factors related to the differing structures of the habitats that they occupy. It is well documented that the diversification of *Bradypodion* correlates to adaptive responses to changes in the natural environment since the Miocene (Tolley et al., 2008). As a result, convergent phenotypic morphs are associated with at least two broad habitat types: closed-canopy forests and open-canopy grassy or fynbos habitats. While the adaptive traits unique to each morph are correlated to the challenges encountered within each habitat type (Measey et al., 2009; da Silva and Tolley, 2013), it should be noted that ancestral habitat reconstruction indicated that these dwarf chameleons originated within closed-canopy forests (Tolley et al., 2008). During the Late Miocene-Pliocene, southern Africa experienced dramatic shifts in climate which promoted the expansion of more open-canopy habitats (Chase and Meadows, 2007; Lee-Thorp et al., 2007). Consequently, closed-canopy forests contracted and became highly fragmented (Tolley et al., 2013). It is, therefore, possible that exposure to light and radiation differ between habitat types, with radiation near the forest floor generally lower compared to the outer canopy (Brown et al., 1994), and overall exposure is greater in open-canopy habitats relative to closed-canopy forest (Gao et al., 2002).

The gene family expansion in *Bradypodion* underlying response to radiation is specifically linked to ultraviolet-B (UV-B) radiation. In reptiles, exposure to ultraviolet light has both physiological benefits and costs (Watson and Mitchell, 2014). UV light is known to catalyse a reaction in the skin of reptiles allowing for the availability of active vitamin D_3 which promotes the uptake of calcium (Holick et al., 1995; Klaphake, 2010). However, of the two wavelengths of UV light, UV-A and UV-B, overexposure to UV-B is known to result in skin damage and skin-related diseases (Gardiner et al., 2009; Watson and Mitchell, 2014; Hedley et al., 2018). While reptiles are known to behaviourally self-regulate exposure to UV light to take advantage of the benefits of UV while minimizing the costs (Ferguson et al., 2010, 2013), regulating exposure is also achieved through morphological adaptations in the skin as well as temperature regulation (Chang et al., 2009; Gholamifard et al., 2015; Sakich and Tattersall, 2022). In *Chamaeleo calyptratus*, the Veiled Chameleon, the endogenous synthesis of vitamin D proved more efficient over dietary ingestion (Hoby et al., 2010). While the Panther Chameleon, *Furcifer pardalis*, was shown to behaviourally adjust UV-B exposure depending on the availability of dietary vitamin D_3 (Ferguson et al., 2003). In *Bradypodion*, the UV-related genetic pathways involved in the gene family expansions highlighted in this study suggest that these dwarf chameleons are more likely adept in efficiently regulating exposure imbalances depending on the habitat type, potentially leading to adaptive traits evolving in ecomorphs associated with UV radiation and temperature-related behaviour.

The dense forest canopy largely filters light and radiation penetration into the forest understory (Brown et al., 1994). Overexposure to UV radiation is therefore unlikely to be experienced by closed-canopy ecomorphs such as *Bradypodion pumilum*. There are still preventative measures available to mitigate potential UV radiation damage, even though it is not the primary concern when inhabiting closed-canopy habitats. Expanded gene families in *B. pumilum* are related to the serine/threonine protein kinase (TOR) signalling pathway which regulates cellular responses to stresses such as DNA damage in tissues (Wullschleger et al., 2006). Within structurally dense forests, these chameleons would naturally move along temperature gradients to regulate body temperatures and likely select exposed patches of direct sunlight to do so. In contrast, open-canopy ecomorphs such as *B. ventrale*, are doubtless exposed to higher UV irradiance across their open-canopy habitats and likely possess adaptive strategies to mitigate the increased radiation stress.

Protection measures against UV exposure are seemingly more available in *Bradypodion ventrale* based on gene family expansions, particularly underlying skin development. The

ability to rapidly alter colour in chameleons is driven by actively reorganizing nanocrystals in iridophores present in the skin to structurally adjust the colour (Teyssier et al., 2015), with colour change being essential for intra-specific communication and thermoregulation. Modulating skin brightness is also achieved by the dispersal/aggregation of dark chromatophores containing melanin, with darker skin tones typically indicative of stress. Melanin (a pigment within iridophores) can absorb UV radiation and thus prevents damage to more sensitive structures underneath the epidermis and dermis (Rouzaud et al., 2006). Given the exposed nature of open-canopy habitats (Gao et al., 2002), it is plausible that a higher density of melanin pigment in chromatophores is present in open-canopy species of *Bradypodion*. This likely presents as a trade-off in the distribution of pigments, favouring melanin over colour-mediating structures, resulting in the dull colouration typically associated with open-canopy ecomorphs (Stuart-Fox et al., 2008; da Silva and Tolley, 2013). However, differences in melanin concentration between *Bradypodion* ecomorphs remain to be tested.

Gene family expansions underlying the perception of light stimuli and bone development are also present in *Bradypodion ventrale* and not in *B. pumilum*. Most tetrapods can see in the UV light spectrum, particularly UV-A given the sensitivity of ocular photoreceptors (Cronin and Bok, 2016). This is particularly relevant in chameleons as many species are known to possess fluorescent tubercles, cranial bone protrusions through the skin, which optimally reflect within the UV-A wavelengths (Prötzel et al., 2018). However, using UV reflectance for communication in *Bradypodion* is more common in closed habitats than in open habitats (Stuart-Fox et al., 2007). Given this, expanded gene families underlying bone development in *B. ventrale* might not be specifically related to UV reflectance. Instead, these expansions might be associated with mediating appropriate amounts of calcium uptake driven by endogenous synthesis of vitamin D in the skin as a response to increased radiation in open-canopy habitats. It is also unclear to what degree various species are able to distinguish between UV-A, UV-B and visible light (Ferguson et al., 2003, 2013). A number of lizards in the sister family, Agamidae, have visual sensitivity to UV light through the pineal-gland-associated eye-like structure known as the 'parietal eye' (Barbour et al., 2002; Yewers et al., 2015). These structures are rudimentary in chameleons (Gundy and Wurst, 1976). Therefore, the expanded gene families underlying visual perception of light stimuli in *B. ventrale* are associated with the intensity of visible light in open-canopy habitats and may have similarly developed in other open-canopy species of *Bradypodion*.

Given the contrast in gene family expansion between *Bradypodion pumilum* and *B. ventrale*,

these results have provided insight into the genomic basis underlying adaptive solutions in response to challenges by the respective habitats. By extension, these findings contribute to the current understanding of ecomorph evolution in *Bradypodion* but were only possible through the generation of high-quality genomic resources (i.e., genome assembly and accompanying annotations). In recent years, assessments of gene family evolution in a number of species have revealed the genetic architecture underlying adaptive responses, particularly within harsh environments (Wang et al., 2019; Peng et al., 2021; Kim et al., 2022). With the rise in *de novo* genome assemblies and annotations of non-model organisms (Genome 10K Community of Scientists, 2009; Koepfli et al., 2015; Lewin et al., 2018; Blaxter et al., 2022; Ebenezer et al., 2022), evaluating the changes in gene family composition will likely provide a new lens from which to routinely assess genetic underpinnings of adaptive evolution.

3.5 Future directions

Bradypodion are the most recently diverged chameleon lineage, characterised by ecologically driven phenotypic specialisation resulting in ecomorphs associated with either closed-canopy forests or open-canopy grass and fynbos (Tolley et al., 2008, 2013; da Silva et al., 2014; da Silva and Tolley, 2017). These ecomorphs are often, but not always, phylogenetically paired and suggest that morphological similarities have arisen independently in response to similar ecological pressures. This presents a unique opportunity to assess the genetic basis underlying adaptation in these chameleons, by leveraging the replicated diversification of *Bradypodion* to investigate the role of gene family evolution in driving the adaptive convergent phenotypes. To truly take advantage of detecting changes in gene family evolution, *de novo* assemblies are required for the taxa in question (Han et al., 2013; Ellis et al., 2021; Yan et al., 2022). Given that whole genome sequencing has become more feasible in recent years (Rice and Green, 2019), targeting closely related ecomorph pairs within *Bradypodion* will be most appropriate, as well as contrasting ecomorphs that occur as different populations within a single species (Herrel et al., 2011; Hopkins and Tolley, 2011). This approach will then allow for detecting convergent gene family changes in a comparative framework. In addition, testing whether the underlying genes within expanded gene families are under positive selection within each ecomorph will provide additional evidence in uncovering the genetic architecture underlying convergent phenotypes in these dwarf chameleons. In conclusion, by investigating convergent gene family changes and assessing positive selection within ecomorphs, the genetic basis of adaptation in the evolutionary history of *Bradypodion* can be revealed.

References

Ayarpadikannan, S. and Kim, H. (2014). The impact of transposable elements in genome evolution and genetic instability and their implications in various diseases. *Genomics & informatics*, 12(3):98.

Bao, Z. and Eddy, S. R. (2002). Automated de novo identification of repeat sequence families in sequenced genomes. *Genome Research*, 12:1269–1276.

Barbour, H. R., Archer, M. A., Hart, N. S., Thomas, N., Dunlop, S. A., Beazley, L. D., and Shand, J. (2002). Retinal characteristics of the ornate dragon lizard, *Ctenophorus ornatus*. *Journal of Comparative Neurology*, 450(4):334–344.

Benfey, P. N. and Mitchell-Olds, T. (2008). From genotype to phenotype: systems biology meets natural variation. *Science*, 320(5875):495–497.

Bickel, R. and Losos, J. B. (2002). Patterns of morphological variation and correlates of habitat use in chameleons. *Biological Journal of the Linnean Society*, 76(1):91–103.

Blaxter, M., Archibald, J. M., Childers, A. K., Coddington, J. A., Crandall, K. A., Di Palma, F., Durbin, R., Edwards, S. V., Graves, J. A., Hackett, K. J., et al. (2022). Why sequence all eukaryotes? *Proceedings of the National Academy of Sciences*, 119(4):e2115636118.

Boistel, R., Herrel, A., Daghfous, G., Libourel, P.-A., Boller, E., Tafforeau, P., and Bels, V. (2010). Assisted walking in malagasy dwarf chamaeleons. *Biology Letters*, 6(6):740–743.

Bomblies, K. and Peichel, C. L. (2022). Genetics of adaptation. *Proceedings of the National Academy of Sciences*, 119(30).

Bourque, G., Burns, K. H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H. L., Macfarlan, T. S., et al. (2018). Ten things you should know about transposable elements. *Genome biology*, 19:1–12.

Bradnam, K. R., Fass, J. N., Alexandrov, A., Baranay, P., Bechner, M., Birol, I., Boisvert, S., Chapman, J. A., Chapuis, G., Chikhi, R., et al. (2013). Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. *Gigascience*, 2(1):2047–217.

Brown, M. J., Parker, G. G., and Posner, N. E. (1994). A survey of ultraviolet-b radiation in forests. *Journal of Ecology*, pages 843–854.

Card, D. C., Jennings, W. B., and Edwards, S. V. (2023). Genome evolution and the future of phylogenomics of non-avian reptiles. *Animals*, 13(3):471.

Chaisson, M. J., Wilson, R. K., and Eichler, E. E. (2015). Genetic variation and the de novo assembly of human genomes. *Nature Reviews Genetics*, 16(11):627–640.

Chang, C., Wu, P., Baker, R. E., Maini, P. K., Alibardi, L., and Chuong, C.-M. (2009). Reptile scale paradigm: Evo-devo, pattern formation and regeneration. *The International journal of developmental biology*, 53(5-6):813.

Chase, B. M. and Meadows, M. E. (2007). Late quaternary dynamics of southern africa's winter rainfall zone. *Earth-Science Reviews*, 84(3-4):103–138.

Cooper, M., Podlich, D. W., and Smith, O. S. (2005). Gene-to-phenotype models and complex trait genetics. *Australian Journal of Agricultural Research*, 56(9):895–918.

Cronin, T. W. and Bok, M. J. (2016). Photoreception and vision in the ultraviolet. *Journal of Experimental Biology*, 219(18):2790–2801.

da Silva, J. M., Herrel, A., Measey, G. J., Vanhooydonck, B., and Tolley, K. A. (2014). Linking microhabitat structure, morphology and locomotor performance traits in a recent radiation of dwarf chameleons. *Functional Ecology*, 28(3):702–713.

da Silva, J. M. and Tolley, K. A. (2013). Ecomorphological variation and sexual dimorphism in a recent radiation of dwarf chameleons (*Bradypodion*). *Biological Journal of the Linnean Society*, 109(1):113–130.

da Silva, J. M. and Tolley, K. A. (2017). Diversification through ecological opportunity in dwarf chameleons. *Journal of Biogeography*, 44(4):834–847.

Das, D., Singh, S. K., Bierstedt, J., Erickson, A., Galli, G. L., Crossley, D. A., and Rhen, T. (2020). Draft genome of the common snapping turtle, *Chelydra serpentina*, a model for phenotypic plasticity in reptiles. *G3: Genes, Genomes, Genetics*, 10(12):4299–4314.

De Bie, T., Cristianini, N., Demuth, J. P., and Hahn, M. W. (2006). CAFE: a computational tool for the study of gene family evolution. *Bioinformatics*, 22(10):1269–1271.

De Smet, R. and Van de Peer, Y. (2012). Redundancy and rewiring of genetic networks following genome-wide duplication events. *Current opinion in plant biology*, 15(2):168–176.

Demuth, J. P. and Hahn, M. W. (2009). The life and death of gene families. *Bioessays*, 31(1):29–39.

Denton, J. F., Lugo-Martinez, J., Tucker, A. E., Schrider, D. R., Warren, W. C., and Hahn, M. W. (2014). Extensive error in the number of genes inferred from draft genome assemblies. *PLoS computational biology*, 10(12):e1003998.

Diaz, R. E., Anderson, C. V., Baumann, D. P., Kupronis, R., Jewell, D., Piraquive, C., Kupronis, J., Winter, K., Bertocchini, F., and Trainor, P. A. (2015). The veiled chameleon (*Chamaeleo calyptratus* duméril and duméril 1851): A model for studying reptile body plan development and evolution. *Cold Spring Harbor Protocols*, 2015(10):pdb–emo087700.

Diaz, R. E. and Trainor, P. A. (2015). Hand/foot splitting and the 're-evolution' of mesopodial skeletal elements during the evolution and radiation of chameleons. *BMC Evolutionary Biology*, 15(1):1–25.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*, 29(1):15–21.

Ebenezer, T. E., Muigai, A. W., Nouala, S., Badaoui, B., Blaxter, M., Buddie, A. G., Jarvis, E. D., Korlach, J., Kuja, J. O., Lewin, H. A., et al. (2022). Africa: sequence 100,000 species to safeguard biodiversity. *Nature*, 603(7901):388–392.

Ejigu, G. F. and Jung, J. (2020). Review on the computational genome annotation of sequences obtained by next-generation sequencing. *Biology*, 9(9):295.

Ekblom, R. and Wolf, J. B. (2014). A field guide to whole-genome sequencing, assembly and annotation. *Evolutionary applications*, 7(9):1026–1042.

Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in ecology & evolution*, 29(1):51–63.

Ellis, E. A., Storer, C. G., and Kawahara, A. Y. (2021). De novo genome assemblies of butterflies. *GigaScience*, 10(6):giab041.

Emms, D. M. and Kelly, S. (2015). Orthofinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome biology*, 16(1):1–14.

Emms, D. M. and Kelly, S. (2019). Orthofinder: phylogenetic orthology inference for comparative genomics. *Genome biology*, 20:1–14.

Ferguson, G. W., Brinker, A. M., Gehrmann, W. H., Bucklin, S. E., Baines, F. M., and Mackin, S. J. (2010). Voluntary exposure of some western-hemisphere snake and lizard species to ultraviolet-b radiation in the field: how much ultraviolet-b should a lizard or snake receive in captivity? *Zoo biology*, 29(3):317–334.

Ferguson, G. W., Gehrmann, W. H., Karsten, K. B., Hammack, S. H., McRae, M., Chen, T. C., Lung, N. P., and Holick, M. F. (2003). Do panther chameleons bask to regulate endogenous vitamin D3 production? *Physiological and Biochemical Zoology*, 76(1):52–59.

Ferguson, G. W., Kingeter, A. J., and Gehrmann, W. H. (2013). Ultraviolet light exposure and response to dietary vitamin D3 in two jamaican anoles. *Journal of Herpetology*, 47(4):524–529.

Ferrari, R. R., Batista, T. M., Zhou, Q., Hilário, H. O., Orr, M. C., Luo, A., and Zhu, C. (2023). The whole genome of *Colletes collaris* (Hymenoptera: Colletidae): an important step in comparative genomics of cellophane bees. *Genome Biology and Evolution*, 15(5):evad062.

Flagel, L. E. and Wendel, J. F. (2009). Gene duplication and evolutionary novelty in plants. *New Phytologist*, 183(3):557–564.

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., and Smit, A. F. (2020). Repeatmodeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences*, 117(17):9451–9457.

Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annual review of plant biology*, 60:433–453.

Gable, S. M., Mendez, J. M., Bushroe, N. A., Wilson, A., Byars, M. I., and Tollis, M. (2023). The state of squamate genomics: Past, present, and future of genome research in the most speciose terrestrial vertebrate order. *Genes*, 14(7):1387.

Gao, W., Grant, R. H., Heisler, G. M., and Slusser, J. R. (2002). A geometric ultraviolet-b radiation transfer model applied to vegetation canopies. *Agronomy journal*, 94(3):475–482.

Gardiner, D. W., Baines, F. M., and Pandher, K. (2009). Photodermatitis and photokeratoconjunctivitis in a ball python (*Python regius*) and a blue-tongue skink (*Tiliqua* spp.). *Journal of Zoo and Wildlife Medicine*, 40(4):757–766.

Geib, S. M., Hall, B., Derego, T., Bremer, F. T., Cannoles, K., and Sim, S. B. (2018). Genome Annotation Generator: a simple tool for generating and correcting WGS annotation tables for NCBI submission. *Gigascience*, 7(4):giy018.

Gemmell, N. J., Rutherford, K., Prost, S., Tollis, M., Winter, D., Macey, J. R., Adelson, D. L., Suh, A., Bertozzi, T., Grau, J. H., et al. (2020). The tuatara genome reveals ancient features of amniote evolution. *Nature*, 584(7821):403–409.

Genome 10K Community of Scientists (2009). Genome 10K: a proposal to obtain whole-genome sequence for 10 000 vertebrate species. *Journal of Heredity*, 100(6):659–674.

Gholamifard, A., Rastegar-Pouyani, N., and Ostovari, H. (2015). Some aspects of the ecology and natural history of Keyserling's plate-tailed gecko, *Teratoscincus keyserlingii*, strauch 1863 (Sauria: Sphaerodactylidae) from South Khorasan Province, Eastern Iran. *Amphibian and Reptile Conservation*, 9(1):58–65.

Giles, S. and Arbuckle, K. (2022). Diversification dynamics of chameleons (Chamaeleonidae). *Journal of Zoology*, 318(4):241–252.

Glazier, A. M., Nadeau, J. H., and Aitman, T. J. (2002). Finding genes that underlie complex traits. *science*, 298(5602):2345–2349.

Gundy, G. C. and Wurst, G. Z. (1976). Parietal eye-pineal morphology in lizards and its physiological implications. *The Anatomical Record*, 185(4):419–431.

Guo, Y.-L. (2013). Gene family evolution in green plants with emphasis on the origination and evolution of *Arabidopsis thaliana* genes. *The Plant Journal*, 73(6):941–951.

Hahn, M. W., Han, M. V., and Han, S.-G. (2007). Gene Family Evolution across 12 *Drosophila* Genomes. *PLoS genetics*, 3(11):e197.

Han, M. V., Thomas, G. W., Lugo-Martinez, J., and Hahn, M. W. (2013). Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Molecular biology and evolution*, 30(8):1987–1997.

Hargreaves, A. L., Germain, R. M., Bontrager, M., Persi, J., and Angert, A. L. (2020). Local adaptation to biotic interactions: a meta-analysis across latitudes. *The American Naturalist*, 195(3):395–411.

Hedley, J., Johnson, R., and Yeates, J. (2018). Reptiles (Reptilia). *Companion Animal Care and Welfare: The UFAW Companion Animal Handbook*, pages 371–394.

Henry, C. S., Overbeek, R., Xia, F., Best, A. A., Glass, E., Gilbert, J., Larsen, P., Edwards, R., Disz, T., Meyer, F., et al. (2011). Connecting genotype to phenotype in the era of high-throughput sequencing. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1810(10):967–977.

Herrel, A., Measey, G. J., Vanhooydonck, B., and Tolley, K. A. (2011). Functional consequences of morphological differentiation between populations of the Cape Dwarf Chameleon (*Bradypodion pumilum*). *Biological Journal of the Linnean Society*, 104(3):692–700.

Higham, T. E. and Anderson, C. V. (2013). Function and Adaptation of Chameleons. In Tolley, K. A. and Herrel, A., editors, *The Biology of Chameleons*. University of California Press, USA.

Hoby, S., Wenker, C., Robert, N., Jermann, T., Hartnack, S., Segner, H., Aebischer, C.-P., and Liesegang, A. (2010). Nutritional metabolic bone disease in juvenile veiled chameleons (*Chamaeleo calyptratus*) and its prevention. *The Journal of nutrition*, 140(11):1923–1931.

Holick, M. F., Tian, X. Q., and Allen, M. (1995). Evolutionary importance for the membrane enhancement of the production of vitamin D3 in the skin of poikilothermic animals. *Proceedings of the National Academy of Sciences*, 92(8):3124–3126.

Hopkins, K. P. and Tolley, K. A. (2011). Morphological variation in the Cape Dwarf Chameleon (*Bradypodion pumilum*) as a consequence of spatially explicit habitat structure differences. *Biological Journal of the Linnean Society*, 102(4):878–888.

Hu, T., Chitnis, N., Monos, D., and Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human Immunology*, 82(11):801–811.

Hubisz, M. J., Lin, M. F., Kellis, M., and Siepel, A. (2011). Error and error mitigation in low-coverage genome assemblies. *PloS one*, 6(2):e17034.

Jarvis, D. E., Ho, Y. S., Lightfoot, D. J., Schmöckel, S. M., Li, B., Borm, T. J., Ohyanagi, H., Mineta, K., Michell, C. T., Saber, N., et al. (2017). The genome of *Chenopodium quinoa*. *Nature*, 542(7641):307–312.

Jauhal, A. A. and Newcomb, R. D. (2021). Assessing genome assembly quality prior to

downstream analysis: N50 versus busco. *Molecular Ecology Resources*, 21(5):1416–1421.

Kaessmann, H. (2010). Origins, evolution, and phenotypic impact of new genes. *Genome research*, 20(10):1313–1326.

Karst, S. M., Ziels, R. M., Kirkegaard, R. H., Sørensen, E. A., McDonald, D., Zhu, Q., Knight, R., and Albertsen, M. (2021). High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nature methods*, 18(2):165–169.

Katoh, K. and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4):772–780.

Kawecki, T. J. and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology letters*, 7(12):1225–1241.

Kim, H., Kim, H.-W., Lee, J. H., Park, J., Lee, H., Kim, S., and Shin, S. C. (2022). Gene family expansions in antarctic winged midge as a strategy for adaptation to cold environments. *Scientific Reports*, 12(1):18263.

Klaphake, E. (2010). A fresh look at metabolic bone diseases in reptiles and amphibians. *Veterinary Clinics: Exotic Animal Practice*, 13(3):375–392.

Koepfli, K.-P., Paten, B., of Scientists, G. K. C., and O'Brien, S. J. (2015). The genome 10K Project: a way forward. *Annual Review of Animal Biosciences*, 3(1):57–111.

Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., and Peterson, H. (2023). g: Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Research*, page gkad347.

Lee-Thorp, J. A., Sponheimer, M., and Luyt, J. (2007). Tracking changing environments using stable carbon isotopes in fossil tooth enamel: an example from the South African hominin sites. *Journal of Human Evolution*, 53(5):595–601.

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., et al. (2018). Earth biogenome project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences*, 115(17):4325–4333.

Lin, D., Lan, L., Zheng, T., Shi, P., Xu, J., and Li, J. (2021). Comparative genomics reveals

recent adaptive evolution in Himalayan giant honeybee *Apis laboriosa*. *Genome Biology and Evolution*, 13(10):evab227.

Lind, A. L., Lai, Y. Y., Mostovoy, Y., Holloway, A. K., Iannucci, A., Mak, A. C., Fondi, M., Orlandini, V., Eckalbar, W. L., Milan, M., et al. (2019). Genome of the Komodo dragon reveals adaptations in the cardiovascular and chemosensory systems of monitor lizards. *Nature ecology & evolution*, 3(8):1241–1252.

Lynch, M. and Conery, J. S. (2003). The evolutionary demography of duplicate genes. *Genome Evolution: Gene and Genome Duplications and the Origin of Novel Gene Functions*, pages 35–44.

McLysaght, A. and Hurst, L. D. (2016). Open questions in the study of de novo genes: what, how and why. *Nature Reviews Genetics*, 17(9):567–578.

Measey, G. J., Hopkins, K., and Tolley, K. A. (2009). Morphology, ornaments and performance in two chameleon ecomorphs: is the casque bigger than the bite? *Zoology*, 112(3):217–226.

Measey, J. G., Raselimanana, A., and Herrel, A. (2013). Ecology and life history of chameleons. In Tolley, K. A. and Herrel, A., editors, *The Biology of Chameleons*. University of California Press, USA.

Mendes, F. K., Vanderpool, D., Fulton, B., and Hahn, M. W. (2020). CAFE 5 models variation in evolutionary rates among gene families. *Bioinformatics*, 36(22-23):5516–5518.

Miller, A. K. and Alexander, G. J. (2009). Do dwarf chameleons (*Bradypodion*) show developmental plasticity? *African Zoology*, 44(1):45–54.

Paradis, E., Claude, J., and Strimmer, K. (2004). Ape: analyses of phylogenetics and evolution in r language. *Bioinformatics*, 20(2):289–290.

Peng, Y., Li, H., Liu, Z., Zhang, C., Li, K., Gong, Y., Geng, L., Su, J., Guan, X., Liu, L., et al. (2021). Chromosome-level genome assembly of the Arctic fox (*Vulpes lagopus*) using PacBio sequencing and Hi-C technology. *Molecular ecology resources*, 21(6):2093–2108.

Perry, B. W., Card, D. C., McGlothlin, J. W., Pasquesi, G. I., Adams, R. H., Schield, D. R., Hales, N. R., Corbin, A. B., Demuth, J. P., Hoffmann, F. G., et al. (2018). Molecular adaptations for sensing and securing prey and insight into amniote genome diversity from the garter snake genome. *Genome Biology and Evolution*, 10(8):2110–2129.

Pinto, B. J., Card, D. C., Castoe, T. A., Diaz Jr, R. E., Nielsen, S. V., Trainor, P. A., and Gamble, T. (2019). The transcriptome of the veiled chameleon (*Chamaeleo calyptratus*): a resource for studying the evolution and development of vertebrates. *Developmental Dynamics*, 248(8):702–708.

Price, A. L., Jones, N. C., and Pevzner, P. A. (2005). De novo identification of repeat families in large genomes. *Bioinformatics*, 21(1):i351–i358.

Primmer, C., Papakostas, S., Leder, E., Davis, M., and Ragan, M. (2013). Annotated genes and nonannotated genomes: cross-species use of gene ontology in ecology and evolution research. *Molecular ecology*, 22(12):3216–3241.

Prötzel, D., Heß, M., Scherz, M. D., Schwager, M., Padje, A. v., and Glaw, F. (2018). Widespread bone-based fluorescence in chameleons. *Scientific reports*, 8(1):698.

Qiu, Q., Zhang, G., Ma, T., Qian, W., Wang, J., Ye, Z., Cao, C., Hu, Q., Kim, J., Larkin, D. M., et al. (2012). The yak genome and adaptation to life at high altitude. *Nature genetics*, 44(8):946–949.

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic acids research*, 47(1):191–198.

Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., et al. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, 592(7856):737–746.

Rice, E. S. and Green, R. E. (2019). New approaches for genome assembly and scaffolding. *Annual review of animal biosciences*, 7:17–40.

Rouzaud, F., Costin, G.-E., Yamaguchi, Y., Valencia, J. C., Berens, W. F., Chen, K. G., Hoashi, T., Böhm, M., Abdel-Malek, Z. A., Hearing, V. J., et al. (2006). Regulation of constitutive and UVR-induced skin pigmentation by melanocortin 1-receptor isoforms. *The FASEB journal*, 20(11):1927–1929.

Sakich, N. B. and Tattersall, G. J. (2022). Regulation of exposure to ultraviolet light in bearded dragons (*Pogona vitticeps*) in relation to temperature and scalation phenotype. *Ichthyology & Herpetology*, 110(3):477–488.

Shao, F., Han, M., and Peng, Z. (2019). Evolution and diversity of transposable elements in fish genomes. *Scientific reports*, 9(1):15399.

Shefchek, K. A., Harris, N. L., Gargano, M., Matentzoglu, N., Unni, D., Brush, M., Keith, D., Conlin, T., Vasilevsky, N., Zhang, X. A., et al. (2020). The monarch initiative in 2019: an integrative data and analytic platform connecting phenotypes to genotypes across species. *Nucleic acids research*, 48(1):704–715.

Shumate, A. and Salzberg, S. L. (2021). Liftoff: accurate mapping of gene annotations. *Bioinformatics*, 37(12):1639–1643.

Shyamli, P. S., Pradhan, S., Panda, M., and Parida, A. (2021). De novo whole-genome assembly of *Moringa oleifera* helps identify genes regulating drought stress tolerance. *Frontiers in Plant Science*, 12:766999.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19):3210–3212.

Stuart, K. C., Edwards, R. J., Cheng, Y., Warren, W. C., Burt, D. W., Sherwin, W. B., Hofmeister, N. R., Werner, S. J., Ball, G. F., Bateson, M., et al. (2022). Transcript-and annotation-guided genome assembly of the European starling. *Molecular Ecology Resources*, 22(8):3141–3160.

Stuart-Fox, D., Moussalli, A., and Whiting, M. J. (2007). Natural selection on social signals: signal efficacy and the evolution of chameleon display coloration. *The American Naturalist*, 170(6):916–930.

Stuart-Fox, D., Moussalli, A., and Whiting, M. J. (2008). Predator-specific camouflage in chameleons. *Biology letters*, 4(4):326–329.

Sumner, J. T., Andrasz, C. L., Johnson, C. A., Wax, S., Anderson, P., Keeling, E. L., and Davidson, J. M. (2023). De novo genome assembly and comparative genomics for the colonial ascidian *Botrylloides violaceus*. *G3: Genes, Genomes, Genetics*, 13(10):jkad181.

Taft, J. M., Tolley, K. A., Alexander, G. J., and Geneva, A. J. (2023). De novo Whole Genome Assemblies for Two Southern African Dwarf Chameleons (*Bradypodion*, Chamaeleonidae). *Genome biology and evolution*, 15(10):evad182.

Teyssier, J., Saenko, S. V., Van Der Marel, D., and Milinkovitch, M. C. (2015). Photonic crystals cause active colour change in chameleons. *Nature communications*, 6(1):6368.

Tolley, K. A., Chase, B. M., and Forest, F. (2008). Speciation and radiations track climate transitions since the Miocene Climatic Optimum: a case study of southern African chameleons. *Journal of Biogeography*, 35(8):1402–1414.

Tolley, K. A. and Herrel, A. (2013). *The Biology of Chameleons*. Univ of California Press.

Tolley, K. A., Tilbury, C. R., and Burger, M. (2022). Convergence and vicariance: speciation of chameleons in the Cape Fold Mountains, South Africa, and the description of three new species of *Bradypodion* Fitzinger, 1843. *African Journal of Herpetology*, 71(1):14–38.

Tolley, K. A., Townsend, T. M., and Vences, M. (2013). Large-scale phylogeny of chameleons suggests African origins and Eocene diversification. *Proceedings of the Royal Society B: Biological Sciences*, 280(1759):20130184.

Trouern-Trend, A. J., Falk, T., Zaman, S., Caballero, M., Neale, D. B., Langley, C. H., Dandekar, A. M., Stevens, K. A., and Wegrzyn, J. L. (2020). Comparative genomics of six *Juglans* species reveals disease-associated gene family contractions. *The Plant Journal*, 102(2):410–423.

Uetz, P., Freed, P., Aguilar, R., Reyes, F., and Hošek, J. (2023). The Reptile Database.

Valenza-Troubat, N., Davy, M., Storey, R., Wylie, M. J., Hilario, E., Ritchie, P., and Wellenreuther, M. (2022). Differential expression analyses reveal extensive transcriptional plasticity induced by temperature in New Zealand silver trevally (*Pseudocaranx georgianus*). *Evolutionary Applications*, 15(2):237–248.

Van de Peer, Y. (2011). A mystery unveiled. *Genome Biology*, 12(5):1–2.

Wang, N., Yang, Y., Moore, M. J., Brockington, S. F., Walker, J. F., Brown, J. W., Liang, B., Feng, T., Edwards, C., Mikenas, J., et al. (2019). Evolution of Portulacineae marked by gene tree conflict and gene family expansion associated with adaptation to harsh environments. *Molecular Biology and Evolution*, 36(1):112–126.

Watson, M. K. and Mitchell, M. A. (2014). Vitamin D and ultraviolet B radiation considerations for exotic pets. *Journal of Exotic Pet Medicine*, 23(4):369–379.

Weirather, J. L., de Cesare, M., Wang, Y., Piazza, P., Sebastiano, V., Wang, X.-J., Buck,

D., and Au, K. F. (2017). Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. *F1000Research*, 6.

Werth, A. J. (2021). How do complex phenotypes evolve? Solving the "gene for x" problem with atavisms, homeosis, and other evo-devo surprises. *Open Science Journal*, 6(4).

Whibley, A., Kelley, J. L., and Narum, S. R. (2021). The changing face of genome assemblies: Guidance on achieving high-quality reference genomes. *Molecular Ecology Resources*, 21(3):641–652.

Wullschleger, S., Loewith, R., and Hall, M. N. (2006). TOR signaling in growth and metabolism. *Cell*, 124(3):471–484.

Xie, H.-X., Liang, X.-X., Chen, Z.-Q., Li, W.-M., Mi, C.-R., Li, M., Wu, Z.-J., Zhou, X.-M., and Du, W.-G. (2022). Ancient demographics determine the effectiveness of genetic purging in endangered lizards. *Molecular biology and evolution*, 39(1):msab359.

Yan, C., Zhang, Z.-Y., Lv, Y., Wang, Z., Jiang, K., and Li, J.-T. (2022). Genome of *Laudakia sacra* provides new insights into high-altitude adaptation of ectotherms. *International Journal of Molecular Sciences*, 23(17):10081.

Yandell, M. and Ence, D. (2012). A beginner's guide to eukaryotic genome annotation. *Nature Reviews Genetics*, 13(5):329–342.

Yewers, M. S., McLean, C. A., Moussalli, A., Stuart-Fox, D., Bennett, A. T., and Knott, B. (2015). Spectral sensitivity of cone photoreceptors and opsin expression in two colour-divergent lineages of the lizard *Ctenophorus decresii*. *The Journal of experimental biology*, 218(10):1556–1563.

Supplemental figures

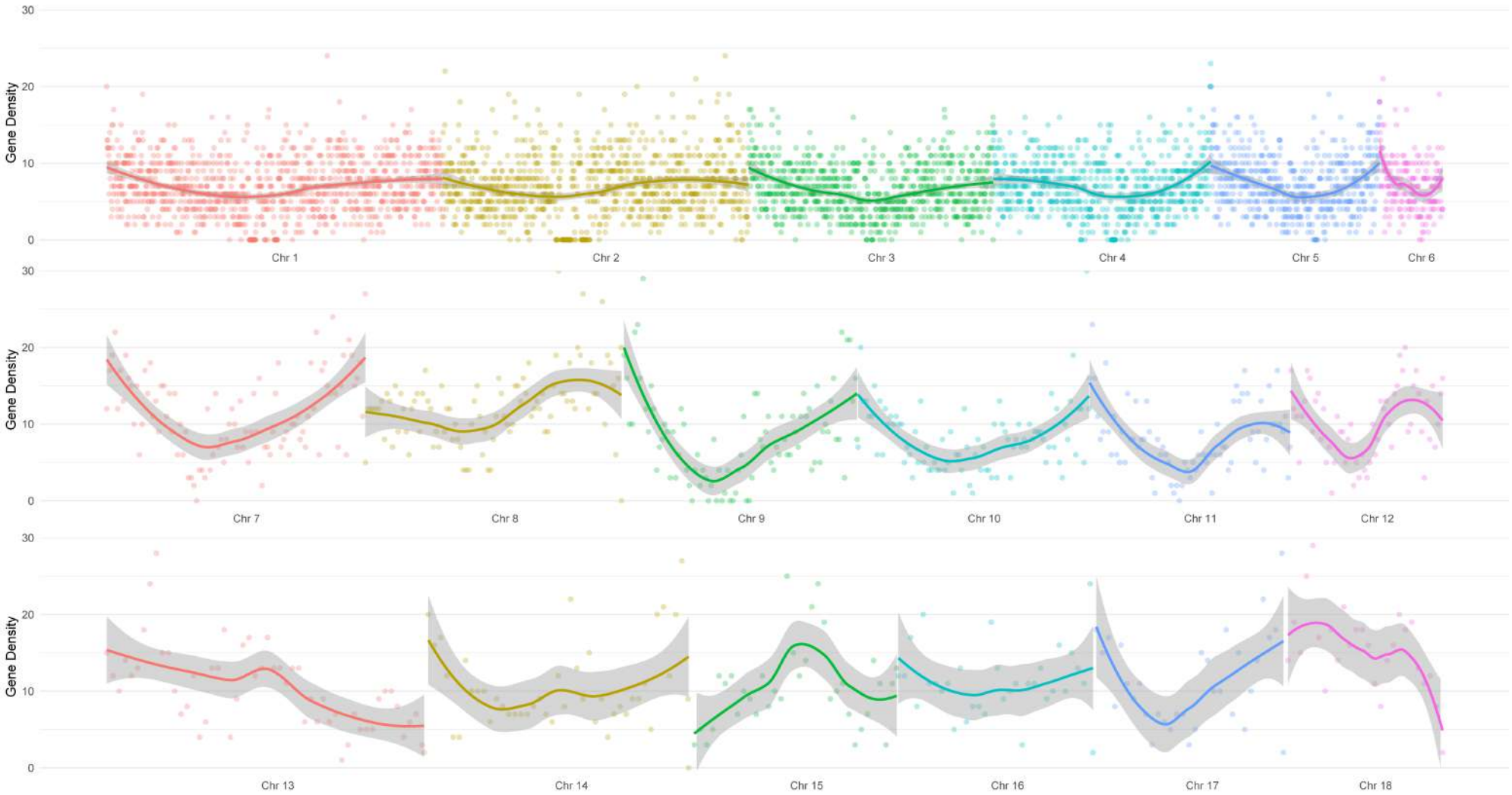


Figure S1: *B. pumilum* gene density landscape. The x-axis is not scaled to reflect the sizes of each chromosome



Figure S2: *B. ventrale* gene density landscape. The x-axis is not scaled to reflect the sizes of each chromosome

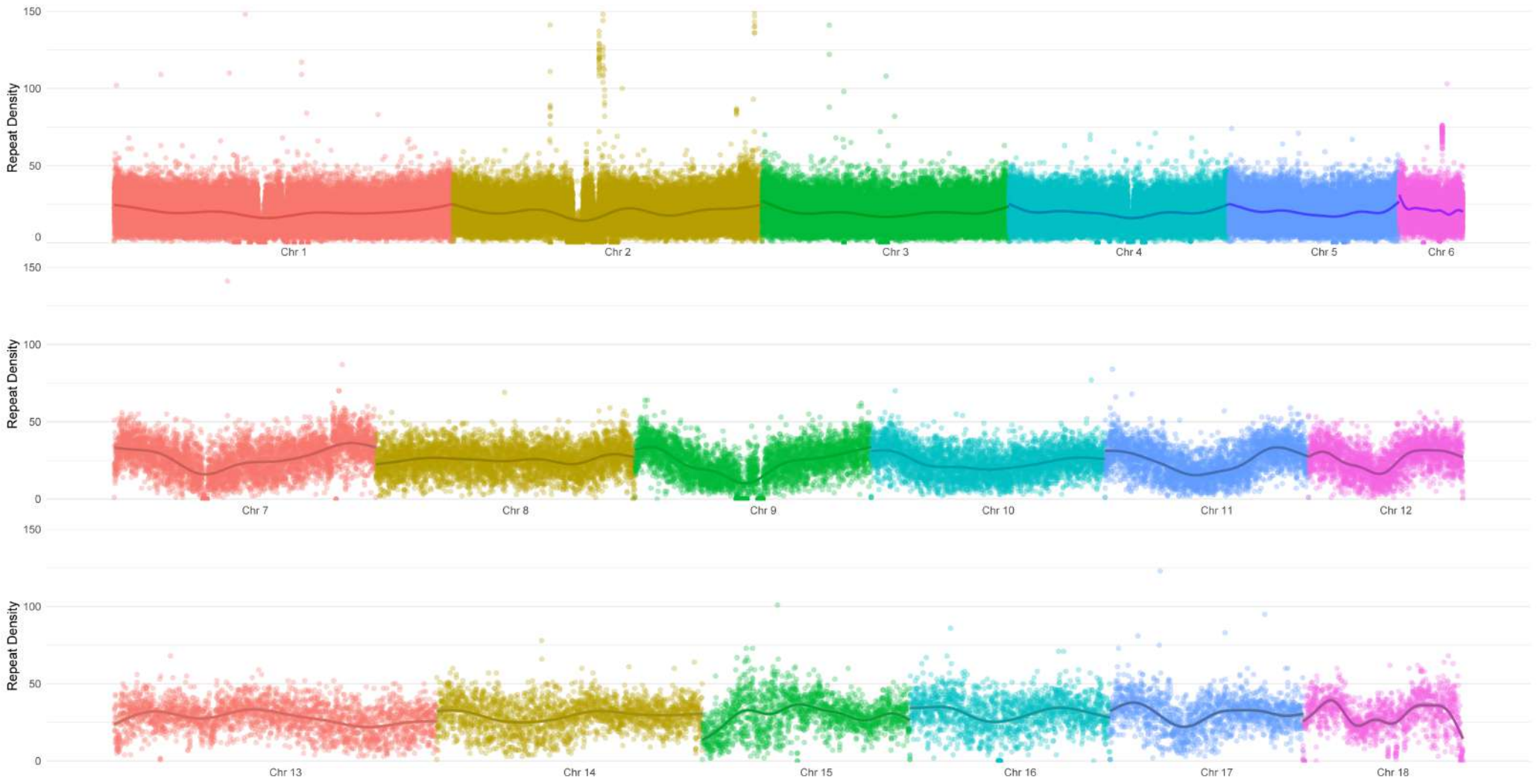


Figure S3: *B. pumilum* repeat density landscape. The x-axis is not scaled to reflect the sizes of each chromosome

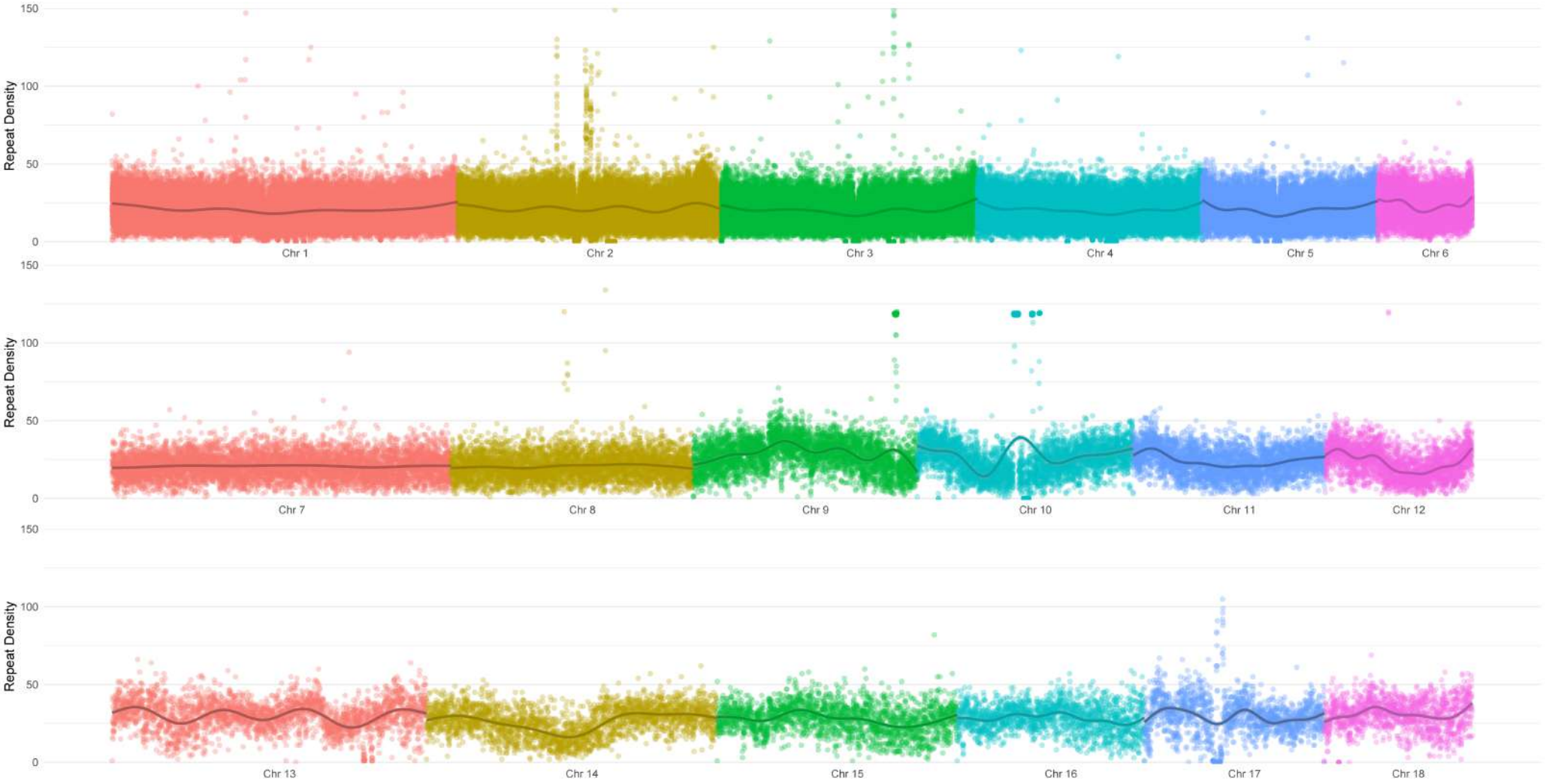


Figure S4: *B. ventrale* repeat density landscape. The x-axis is not scaled to reflect the sizes of each chromosome

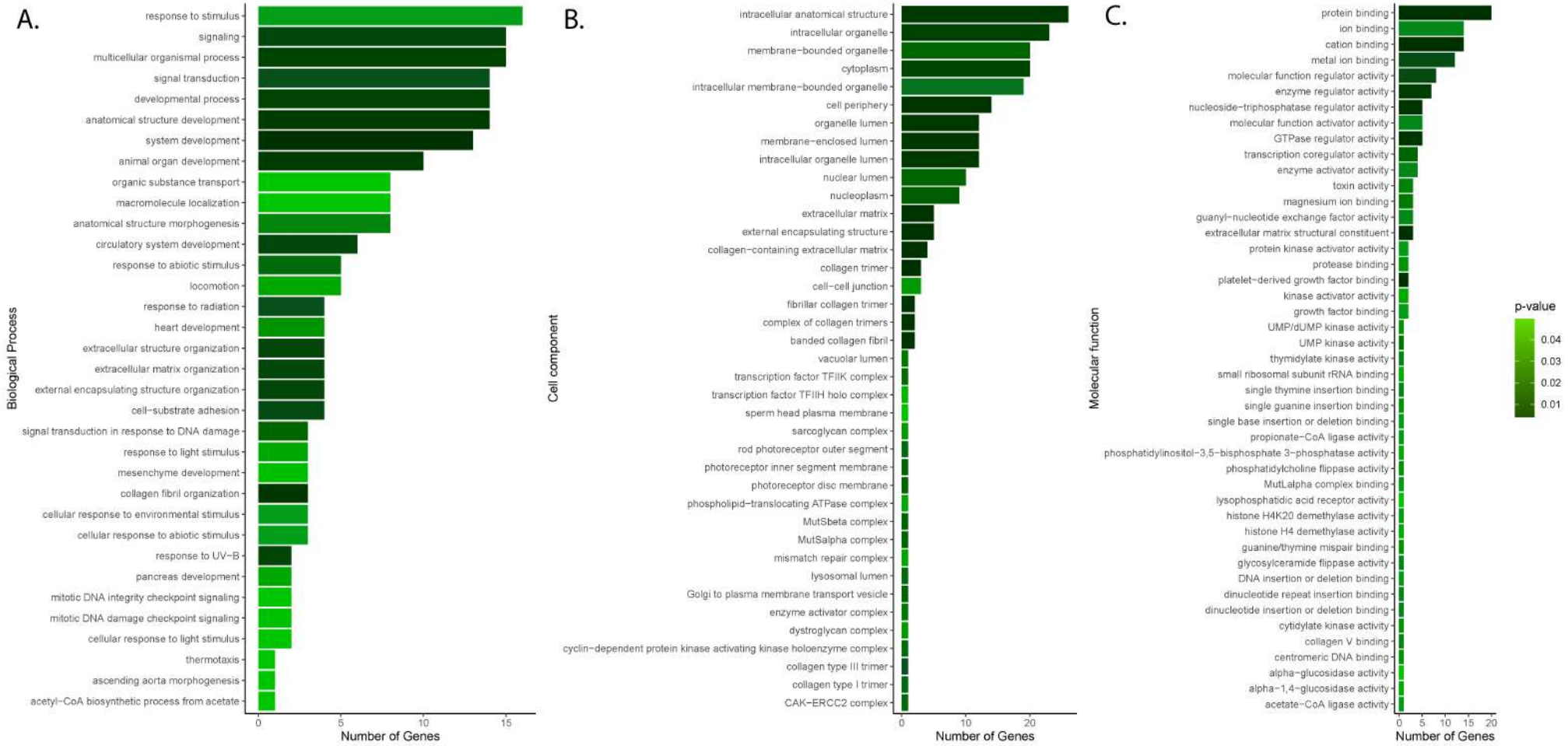


Figure S5: Functional GO terms for Bradypodion ancestral node for each gene ontology function: A. Biological process B. Cellular component C. Molecular function

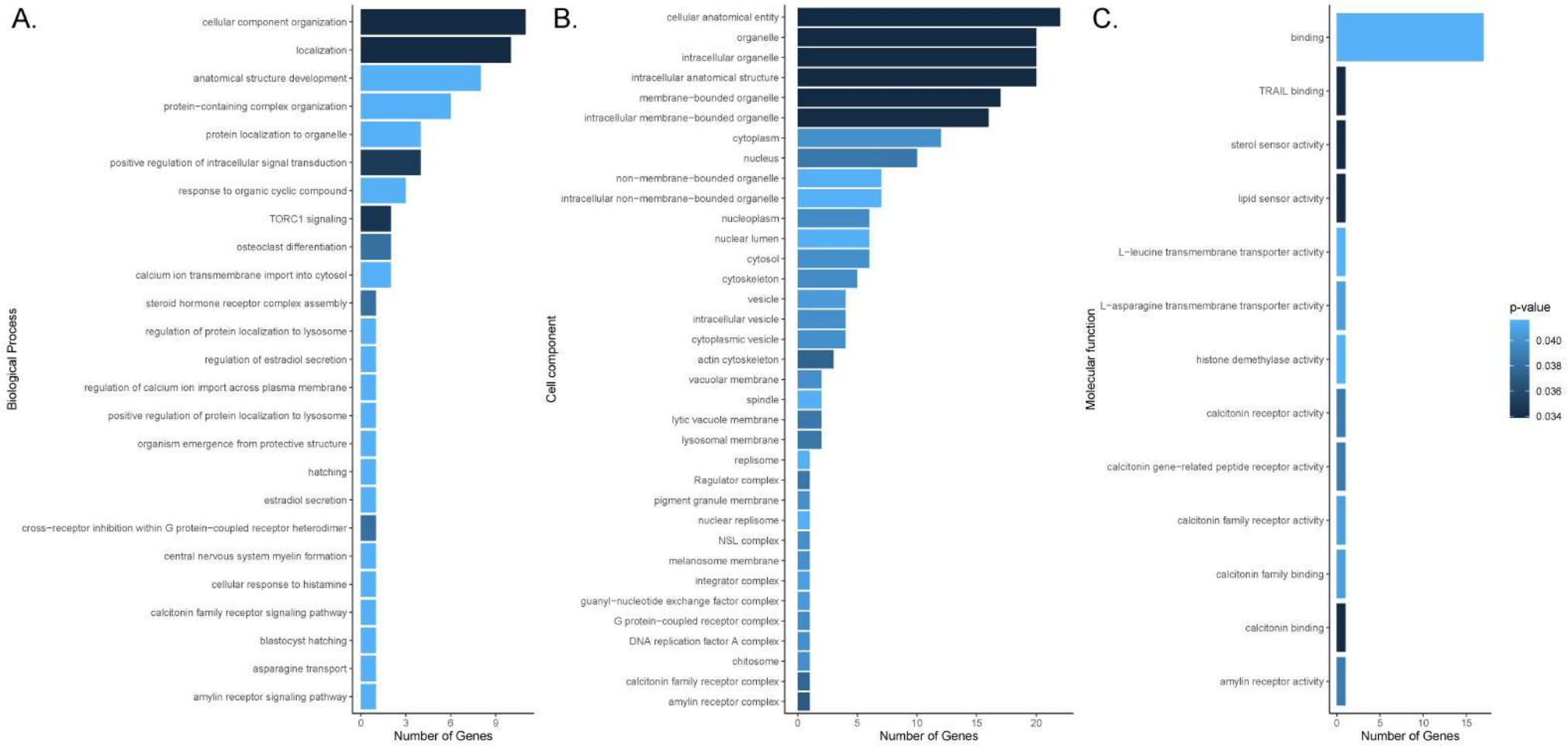


Figure S6: Functional GO terms for *Bradypodion pumilum* for each gene ontology function: A. Biological process B. Cellular component C. Molecular function

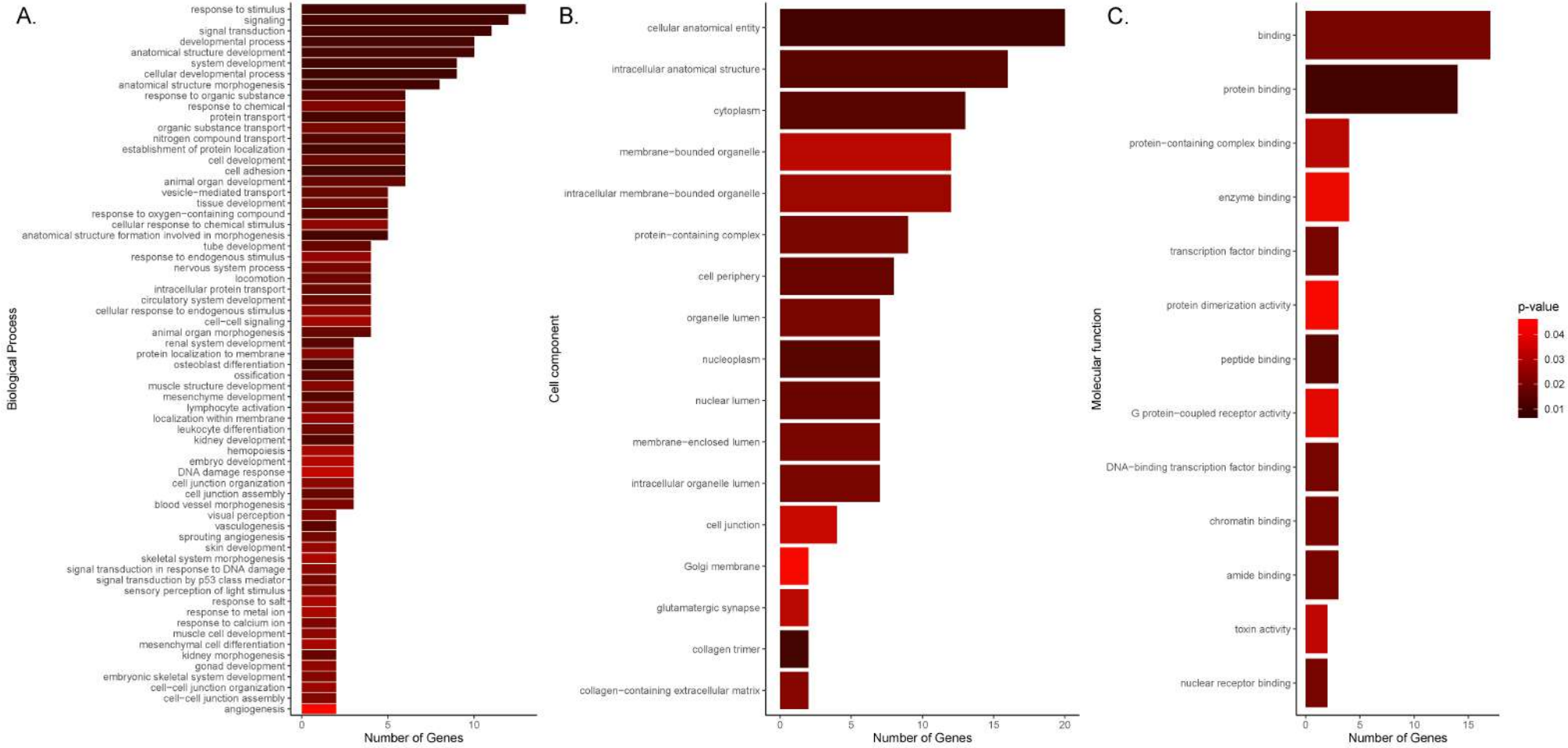


Figure S7: Functional GO terms for *Bradypodion ventrale* for each gene ontology function: A. Biological process B. Cellular component C. Molecular function

Chapter 4

Detecting signals of selection within *Bradypodion*

Abstract

In response to anthropogenic-driven environmental change, there is increasing evidence for rapid contemporary adaptation in several taxa which has garnered considerable interest in evolutionary biology and urban ecology. However, these remarkable examples of adaptation are largely explored in taxa that are able to utilize man-made substrates effectively. For those species inhabiting urban areas but are restricted to fragmented green spaces, the effects of urbanization remain relatively unknown. The southern African dwarf chameleons, *Bradypodion*, are one such group, where several species are known to occupy arboreal niches in urbanized areas. Here I assess population structure and the genetic variation within five species of *Bradypodion* with urban and natural populations. I took a population genomic approach and implement a pairwise comparison to evaluate allele frequency differences between these populations. After accounting for population differentiation, I was able to identify outlier loci exhibiting significantly divergent allele frequencies between urban and natural populations. Through linking genes containing outlier loci and the associated gene function, phenotypic traits were identified that are known to be associated with adaptation in urban habitats. Even though I was unable to identify which population experienced divergence, strong arguments are made in favour of urban populations possessing the genes containing outlier loci.

4.1 Introduction

The rapid transformation of the natural environment into human-dominated landscapes is a key feature of the Anthropocene (Halfwerk et al., 2019a). The transformation of habitat is especially evident in urbanised areas. As more people move to urban areas, the expansion of cities has become one of the biggest impacts on natural systems today. While urbanised areas around the world share overlapping characteristics in terms of infrastructure, they differ in other aspects such as the age since establishment, the historical context, urban planning, and local government policies. However, the commonalities between urban areas represent an unintended but highly replicated system for studying adaptive evolution. As more evidence highlights patterns of evolution within urban systems, we can now seek to understand both the repeatability and the pace of evolution in response to human activities (Johnson and Munshi-South, 2017).

Urban systems undergo habitat transformation in extreme ways and are likely to affect populations at the microevolutionary level if they are able to persist in these systems. Microevolution is the change in allele frequencies in a population from one generation to the next, with four main mechanisms likely to influence this change: mutation, genetic drift, gene flow, and selection. Evolution resulting from new mutations typically occurs over longer timescales than the habitat transformation process in urban areas. It is therefore unlikely that evolutionary change in these novel environments would be primarily due to new mutations (Barrett and Schluter, 2008; Johnson and Munshi-South, 2017, but see Yauk et al., 2000; Somers et al., 2004; Mousseau and Møller, 2014). Rather, selective pressures on standing genetic variation (including pre-existing mutations) within populations are more likely to result in evolutionary responses to transformed habitats and could be more drastic in urban systems which experience particularly rapid and large habitat shifts (Reid et al., 2016; Diamond and Martin, 2021; McCulloch and Waters, 2023).

Urbanisation is predicted to strongly influence genetic drift, which produces stochastic changes in allele frequencies between generations (Miles et al., 2019; Fusco et al., 2021). Genetic drift is most prominent in small, isolated populations, and the influence of drift within urban areas is expected to increase whenever population sizes are reduced due to urbanisation. Reduction in population sizes can be the result of 1) the loss of natural habitat by fragmentation, 2) the founder effects associated with newly established populations, and 3) severe bottlenecks due to population declines as a result of direct selection pressures from

humans. These scenarios will likely result in reduced genetic diversity within populations and increased differentiation between urban & natural populations.

As populations occupying different habitats become more divergent, natural selection can drive local adaption in response to selective pressures within each habitat (Schluter and Conte, 2009; Schluter et al., 2010; Nosil, 2012). However, gene flow, if strong between populations may inhibit local adaptation (Lenormand, 2002; Akerman and Bürger, 2014). The impact of gene flow on adaptation is well documented when uncovering drivers of speciation (Smadja and Butlin, 2011), as the spatial context of population divergence and the extent of gene flow can restrict the pace of speciation (Carvalho et al., 2023). This principle is also evident when detecting local adaptation of intraspecific populations while in the presence of gene flow (Butlin et al., 2014; Tigano and Friesen, 2016). Even with restricted gene flow in urbanized habitats (Yang et al., 2020), urban-associated phenotypic and genomic changes are evident within persisting populations (Johnson and Munshi-South, 2017). However, given that changes are only evident in a limited number of taxa, the underlying genetic elements subject to selection in urban environments remain largely unidentified (Perrier et al., 2020, but see Winchell et al., 2023), especially since the appropriate resources with which to identify genetic elements are lacking (e.g., high-quality genome assemblies and annotations; Card et al., 2023; Gable et al., 2023).

It has been demonstrated that taxa capable of using man-made structures (e.g. walls of buildings and fences) display strongly divergent traits to their natural counterparts (e.g., longer limb lengths and increased lamellae in anoles, Winchell et al., 2016; shorter limb lengths and digits in fence lizards, Putman et al., 2019). While these effects are mostly seen and tested within typical urban systems, evidence suggests that similar adaptation patterns may occur across urban-rural-natural gradients (Garcia et al., 2017; Samia et al., 2017; DeCandia et al., 2019; Parsons et al., 2020). Therefore, a response to increased habitat transformation likely results in rapid adaptation; however, evidence to support this is best contrasted at the urban-natural level. Moreover, taxa restricted to arboreal habitats are expected to display similar trends of rapid adaptation (Winchell et al., 2018, 2023), but the extent of population genetic and morphological divergence is largely untested for many arboreal groups (Winchell et al., 2020).

Bradypodion is a genus of endemic dwarf chameleons in southern Africa. These arboreal lizards occupy an array of habitats and have been shown to have undergone rapid diversification

driven by changes in habitat (reduction of forests and expansion of fynbos, grasslands, etc., Tolley et al., 2008). *Bradypodion* are also known to persist within several urban areas, occupying arboreal habitats within parks, gardens and roadside vegetation (Tolley and Burger, 2007; Petford et al., 2023). The available vegetation within urban habitats is less dense, and/or structurally different to vegetation in natural areas (e.g., Palamuleni et al., 2015; Stickler and Shackleton, 2015; Magidi and Ahmed, 2022), likely resulting in unique challenges and differences in selection pressures (Hendry et al., 2008; Winchell et al., 2018). In urban populations of the Knysna Dwarf Chameleon, *B. damaranum*, individuals exhibit different head morphology and higher bite forces than their natural counterparts (Petford et al., 2023). Given that the vegetation density in urban areas provides less cover than in natural habitats, urban populations of *Bradypodion* are likely more exposed and are at higher risk of predation(Petford et al., 2023). To remain less conspicuous in exposed urban habitats, urban populations of *B. damaranum* exhibit reduced ornamentation relative to their natural counterparts, but this comes at a cost of effective signalling leading to a breakdown in intraspecific communication (Petford et al., 2023). Whether the presence of these urban-associated traits exists in other species of *Bradypodion* occupying similar habitats remains untested. However, these chameleons provide a model for testing the effects of urban habitats on a predominantly arboreal group.

Here, I assess the population structure and genetic diversity of five species of *Bradypodion* known to have populations in natural and urban habitats and report genes that contain loci that may be under selection. I predicted that the greatest level of genetic divergence would be observed between different species, reflecting the broader evolutionary divergence. Furthermore, I anticipated that urban and natural populations of the same species would exhibit evidence of allele frequency differences between populations. I anticipated that these allele frequency differences will involve loci that occur along genes associated with traits likely under selection. If so, genes with known functions may underlie traits that are suited to urban environments, suggesting that the allele frequency differences are driven by unique loci at higher frequencies in urban populations.

4.2 Methods and materials

4.2.1 Sampling and tissue collection

Adult *Bradypodion* from five species that occur within both urban and natural habitats were collected across South Africa. Using a paired sampling approach, sites were selected based on

the proximity (e.g. within 15 km) of natural to urban habitats: *B. damaranum* from George, Western Cape; *B. melanocephalum* from Durban, KwaZulu-Natal; *B. setaroi* from St. Lucia, KwaZulu-Natal; *B. thamnobates* from Midlands, KwaZulu-Natal; *B. ventrale* from Jeffreys Bay, Eastern Cape (Fig. 4.1). The collection of lizards took place in a single habitat patch to ensure the comparison of distinct populations. Lizards were caught by hand as encountered or by encouraging them to climb onto a 7-meter long telescopic pole when lizards were out of reach. Tissue samples were then collected by clipping from the distal tail (< 3 mm) of each lizard and then preserved in nucleic acid preservation buffer and stored at −40 °C.

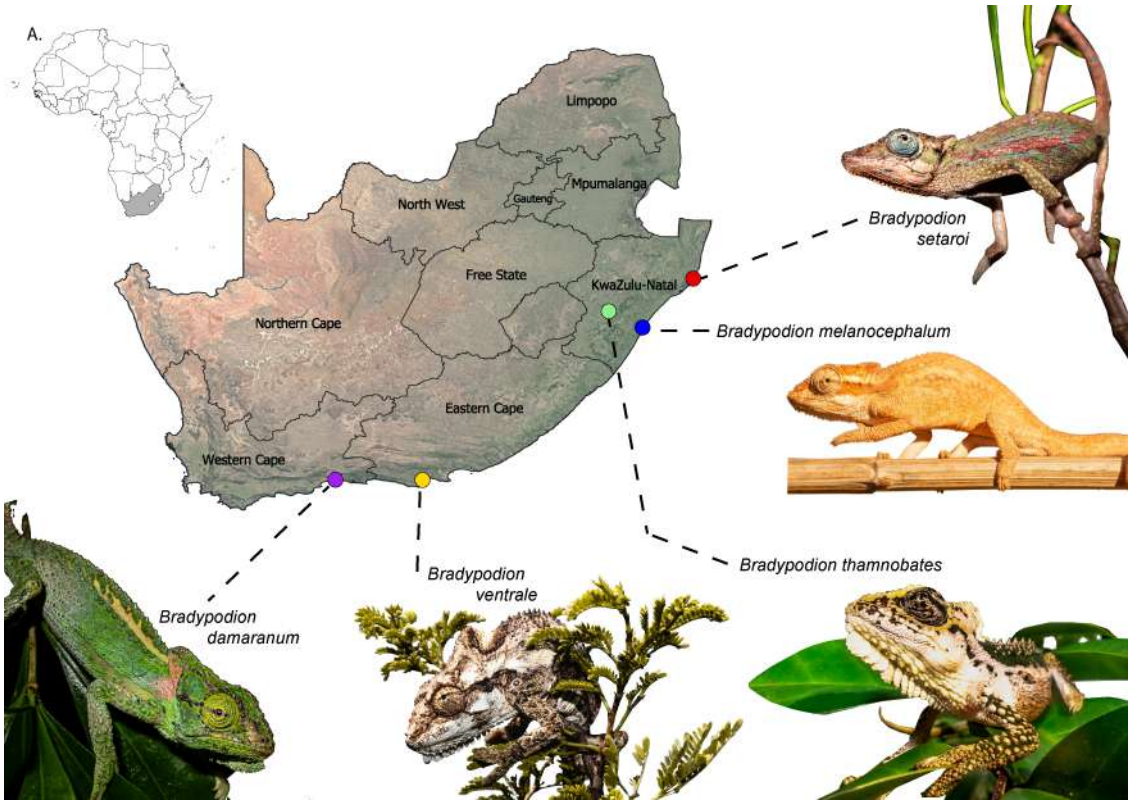


Figure 4.1: Sampling localities for each study species in South Africa. *Bradypodion damaranum* from George, Western Cape (purple); *B. melanocephalum* from Durban, KwaZulu-Natal (blue); *B. setaroi* from St. Lucia, KwaZulu-Natal (red); *B. thamnobates* from Midlands, KwaZulu-Natal (green); *B. ventrale* from Jeffreys Bay, Eastern Cape (yellow). Inset: Map of Africa, with South Africa in grey.

4.2.2 Sequencing

Total genomic DNA was extracted from a total of 88 samples (n = 17, *B. damaranum*; 19, *B. melanocephalum*; 15, *B. setaroi*; 17, *B. thamnobates*; 20, *B. ventrale*) using the OMEGA E.Z.N.A®Tissue Kit. The extraction protocol was followed as per the manufacturer's instructions, however, each sample was eluted twice at a volume of 100 µL to recover as much genetic material as possible. DNA concentrations were estimated fluorometrically with a Qubit 3 Fluorometer (Life Technologies) using the Qubit dsDNA BR Assay Kit, as well as

confirming the presence of genomic DNA with minimal degradation by gel electrophoresis. Total DNA concentrations were also screened using the Genomic DNA ScreenTape Analysis from Agilent Technologies Inc. prior to library preparation. Sequencing libraries were prepared using 10 ng of DNA per sample using the Illumina DNA Library Prep Kit which made use of Transposase-mediated Tagmentation. Each sample was then assigned a barcode using IDT Nextera for Illumina DNA Unique Dual 10 bp Indexes adaptors. After library preparation, re-sequencing was carried out on an Illumina NovaSeq 6000 S4 sequencer.

4.2.3 Sequence-read processing

The read quality of the initial sequence was assessed using FastQC v0.11.8 (Andrews, 2010) on each sample. Sequence adapters were removed, and reads were trimmed and filtered based on quality scores using Trimmomatic v0.32 (Bolger et al., 2014) with the following syntax: LEADING:20 TRAILING:20 MINLEN:23 SLIDINGWINDOW:13:20. All sequence reads were quality assessed in FastQC once again to ensure that filtering produced the highest-quality reads. Before quality-filtered reads were aligned with the *Bradypodion ventrale* v1.1 assembly (Taft et al., 2023), the genome was first assessed to be sorted and renamed by size using the BBMap short read toolkit (Bushnell, 2021). The quality reads were then aligned with *the Bradypodion ventrale* v1.1 assembly and sorted by scaffold size using the Burrows-Wheeler Aligner (BWA) v0.7.17-r1188 (Li, 2013). To retain as much sequence data as possible, paired-end and single-end reads were merged and then assessed using SAMtools (Li et al., 2009). The merged reads were evaluated based on mapping statistics and the reads were retained if the following criteria were met; the proportion of reads mapped to reference > 95; mean read depth > 3; breadth of coverage > 50.

4.2.4 Variant calling and filtering

Variants (single nucleotide polymorphisms, SNPs) were called and filtered using the Genome Analysis Toolkit (GATK; McKenna et al., 2010) following the GATK best practices (DePristo et al., 2011; Van der Auwera and O'Connor, 2020). All best-practice steps were followed apart from Base Quality Score Recalibration, which was not possible as a reference variant set for any *Bradypodion* species does not exist. ReadGroup information was manually added, followed by marking duplicate reads in each bam file and then merging files generated from unique lanes to call haplotypes (which included syntax $-$ERC GVCF) for each sample individually. gVCFs were combined for each population (urban and natural individuals) and jointly called genotypes for

each species separately. All sites were retained with a standard minimum confidence threshold of 20 for calling used. Both variant and invariant files were isolated using GATK SelectVariants and pushed through GATK VariantsToTable to appropriately estimate filtering parameters. Filtered SNPs using GATK VariantFiltration were done using the following filtering expression "QUAL < 0.00 || MQ < 40.00 || SOR > 10.00 || QD < 2.000 || FS > 60.000 || MQRankSum < -12.50 || ReadPosRankSum < -8.00 || ReadPosRankSum > 8.00" and then jointly filtered both variant and invariant sites to remove sites with read depths less than 3 and greater than 70. Further filtering on SNPs was done using VCFtools v0.1.15 for a minimum quality of 25 and a maximum of 25% missing samples per site.

4.2.5 Species-level structure analysis

SNPs datasets were filtered using bcftools v1.15.1 to contain no missing sites, a minimum minor allele frequency of 0.01, and to remove sites with linkage greater than $r^2 = 0.2$ within 10-kb windows (retaining the site in an LD pair with the greater allele frequency). Species SNP datasets were then converted from vcf to bed using Plink v1.9 (Purcell et al., 2007). Population structure was analysed using fastSTRUCTURE v1.0 (Raj et al., 2014) using the combined SNP dataset in bed format. To estimate the appropriate number of model components that explain structure in the combined dataset, fastSTRUCTURE was run with multiple choices of K (number of ancestral populations), from 2 to 10 with fivefold cross-validation. Stucture_threader, a wrapper tool for structure analyses (Pina-Martins et al., 2017), was used to automate and parallelize the running of fastSTRUCTURE. The fastSTRUCTURE utility tool was used to parse through all outputs of these runs to estimate the model complexity appropriate for the best K value. To visualize admixture proportions, the structure result was plotted using the R package, Pophelper v2.3.0 (Francis, 2017). An exploratory principal components analysis (PCA) was also implemented to distil the genetic variation in the dataset into two component axes to assess population structure in a model-free approach. The PCA was conducted using the *pcadapt* package conducted in R (Luu et al., 2017), with varying estimates of K (2-10) to enable estimating structure underlain by the greatest amount of genetic variation in the dataset.

4.2.6 Within-species population differentiation, divergence, and diversity

To compare urban and natural populations for each species, two metrics were estimated: F_{ST} (the proportion of genetic variance contained in each population) and D_{XY} (sequence divergence between populations) using the Python scripts parseVCF.py and popgenWindows.py (https://github.com/ simonhmartin/genomics_general). Nucleotide diversity (π) was estimated separately for each population (i.e., urban and natural). For this analysis, variant and invariant sites were used while indels were excluded. Summary statistics were calculated across 10kb windows, and all windows retaining fewer than 100 called sites weren't considered. Additionally, heterozygosity was estimated for each population using VCFtools v0.1.15 (–het and –relatedness). To assess whether each population is at neutrality (mutations are random) or instead showing signs of selection (non-random patterns of mutation), Tajima's D (–TajimaD 100000) was estimated.

4.2.7 Selection scan and functional term estimation analyses

Selection scan analyses were conducted using the dataset of genomic SNPs for each species after being filtered to remove SNPs with more than 0.25 missing sites and a minor allele frequency threshold of 0.06. Following the approach outlined in Winchell et al. (2023), two complementary methods were used to identify outlier loci: a differentiation outlier method (a modified principle component analysis: pcadapt) and a genetic–environment association method (GEA). First, the *pcadapt* method is implemented in R carries out a principal component analysis and computes p-values to test for outliers (Duforet-Frebourg et al., 2016; Martins et al., 2016; Luu et al., 2017). The test for outliers is based on the correlations between genetic variation and the principal components retained (Luu et al., 2017). For this analysis, principal components were retained that accounted for variation that distinguished urban and natural populations. The Mahalanobis distance was then calculated as the test statistic for detecting outlier SNPs, which is a multi-dimensional approach that measures how distant a point is from the mean (McLachlan, 1999; Ghorbani, 2019). P-values were then transformed into q-values with the 'qvalue' function in R (Dabney and Storey, 2011), with the q-values used to estimate outliers based on a threshold ≤ 0.001 while accounting for a false-discovery rate of 1% (Luu et al., 2017). Second, a genotype association test was done using a logistic linear mixed effects model (binomial family) applied with the GENESIS (Gogarten et al., 2019) functions fitNullModel and assocTestSingle. The model was built using a 50-kb sliding window size with a 10-kb step.

A genetic relatedness matrix was incorporated to account for population structure which was estimated with functions pcair and pcrelate in GENESIS. Outlier SNPs from the GEA were identified as the smallest 1% of the distribution of P-values for the association test. Outlier genes were identified with at least one outlier SNP present when mapped against the *Bradypodion ventrale* genome annotation. Using the intersection of outlier SNPs in the PCA analysis and the GEA analysis, a core set of genes was identified that are associated with urbanisation.

To ascertain the biological functions linked with the identified outlier genes, genes were tested for associated functional Gene Ontology (GO) terms. Functional classes were estimated using g:GOSt in g:Profiler (Raudvere et al., 2019), following the approach outlined in Chapter 3, with *Anolis carolinensis* set as the reference organism and P-values to be adjusted using Benjamini and Hochberg's approach, controlling the false discovery rate for a significance threshold of $P \leq 0.05$. Significantly enriched GO terms were returned unordered. Further filtering of significant terms within g:GOSt was done to ensure unique leading terms in each GO component (molecular function, biological process, cellular component) by removing bias for terms with the highest significant levels (Kolberg et al., 2023).

4.2.8 Isolation by Distance

Isolation by distance as a factor in explaining the differences between the urban and natural populations was assessed by examining the relationship if there was spatial autocorrelation between the genetic and geographic distance of all sampled individuals, separately for each species sampling sites and the genetic distances of samples per species. The latitude and longitude were logged for each sample using a GarminMAP 64s or GarminMAP 65 (Garmin International Incorporated, Olathe, Kansas, USA) handheld GPS ($\pm$ 3 m precision; projection: WGS84). Geographic distances were then computed as the shortest distance between samples using the disHaversine function from the R package geosphere (Hijmans et al., 2017). Using the PC scores generated in the PCA from signatures of selection, the genetic distance matrix was estimated using the Euclidean distances generated from the mean of PC1 to PC3 per sample, accounting for the majority of the variation in the PCA analysis. Finally, Isolation by distance was assessed using a Mantel's test (Legendre et al., 2015, R package vegan; Oksanen et al., 2019) was run using the pairwise geographic distances and genetic distance matrix as input, using running the analysis for 9999 permutations.

4.3 Results

4.3.1 Species-level structure

When assessing population structure, the fastSTRUCTURE analysis produced the greatest support for K = 5 genetic clusters which best described the data (Fig. 4.2). Model complexity that maximized marginal likelihood reported the same number of clusters (K = 5). These genetic clusters grouped samples by species, with both urban and natural populations clustering within their respective species. Notably, however, a single *Bradypodion melanocephalum* (sampled from the urban population within that species range) was 40% admixed with *B. setaroi* and is therefore assumed to be a hybrid and was removed from downstream analyses. *Bradypodion melanocephalum* and *B. setaroi* do not naturally occur in sympathy (Tolley and Burger, 2007). This particular sample was collected well within the natural range of *B. melanocephalum* and was likely the result of a translocation given the number of observations of *B. setaroi* in Durban, KwaZulu-Natal (Tolley, pers. comm.).

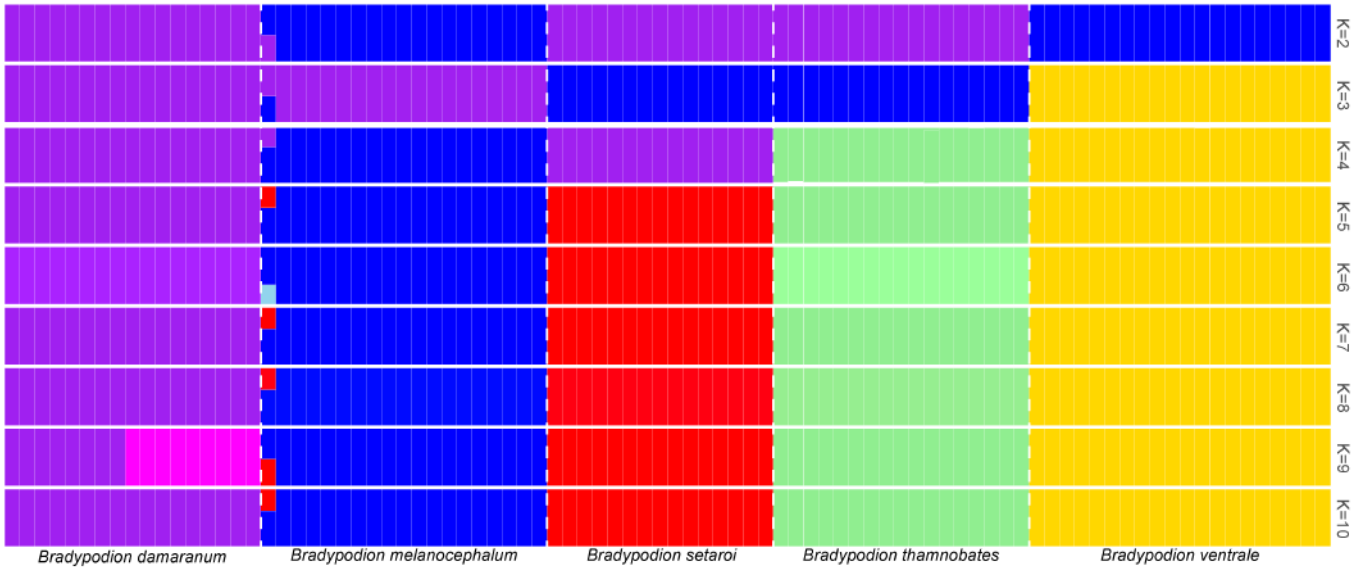


Figure 4.2: fastSTRUCTURE plot showing admixture proportions for five species of *Bradypodion* (n = 88 samples). Estimates of K are shown for values of 2 - 10. K = 5 best explains the data with species groups delineated by the vertical dotted line in white.

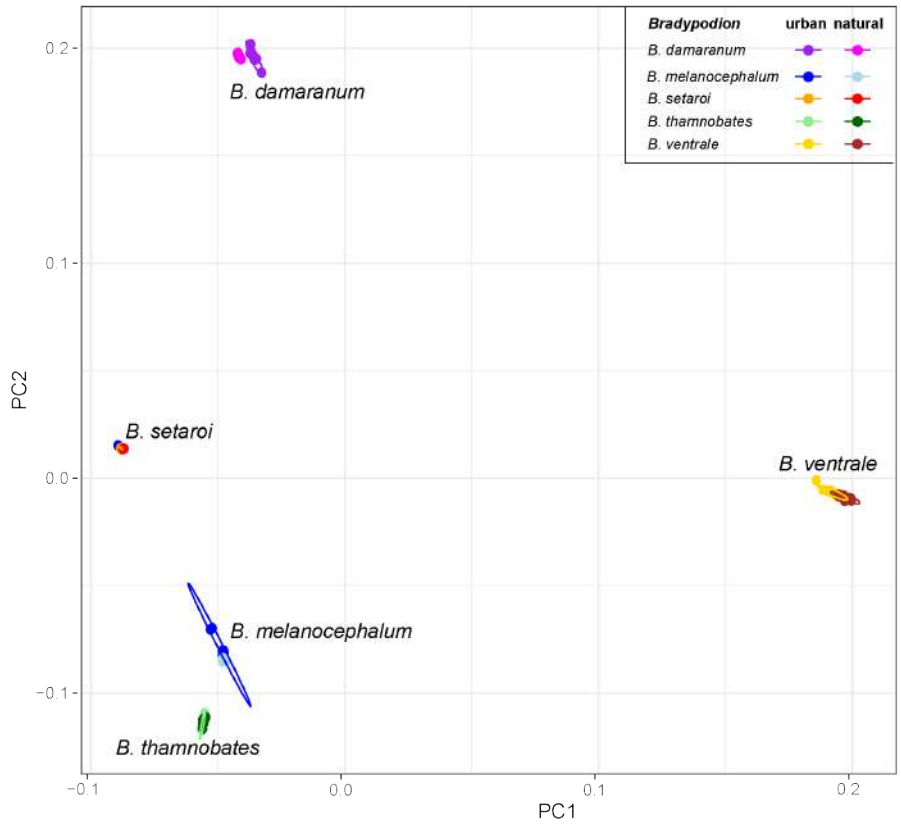


Figure 4.3: Scatterplot of principal components (PC1 and PC2) for five species of *Bradypodion*, colour-coded by species and urban and natural populations within each species with 95% confidence ellipses.

The principal components analysis for all five species showed that principal components 1 - 3 contained 63% of the variation (Supplemental Fig. S1 A). Each species occupies unique areas of multivariate space (Fig. 4.3). Additionally, while urban and natural populations share genetic diversity, differentiation between populations was detected, for *B. damaranum*, *B. melanocephalum*, and *B. ventrale*, each population occupying fairly distinct multivariate space (Fig. 4.3, Supplemental Fig. S1 B).

4.3.2 Within-species population differentiation, divergence, and diversity

The assessment of average pairwise F_{ST} indicated little differentiation between urban-non-urban populations. These average pairwise F_{ST} (and standard deviation) for each species population pairs are as follows: *B. damaranum*, 0.049 ± 0.042; *B. melanocephalum*, 0.035 ± 0.053; *B. setaroi*, 0.007 ± 0.034; *B. thamnobates*, 0.006 ± 0.027; *B. ventrale*, 0.042 ± 0.04. When assessing patterns of F_{ST} and D_{XY} across the genome, in regions with elevated F_{ST} there is often a reduction in D_{XY}, particularly evident in *B. damaranum* (Supplementary

Fig. S2), *B. melanocephalum* (Supplementary Fig. S3), and *B. ventrale* (Supplementary Fig. S4). This contrast of F_{ST} and D_{XY} indicates that the genetic differences between urban and natural populations in these species are not due to the accumulation of new mutations in the underlying regions, but a result of differences in allele frequencies (Cruickshank and Hahn, 2014). While this pattern between F_{ST} and D_{XY} is not as clear in *B. setaroi* and *B. thamnobates*, regions of population differentiation are still evident in these species with no correlation in D_{XY} (Supplementary Figs S5 and S6).

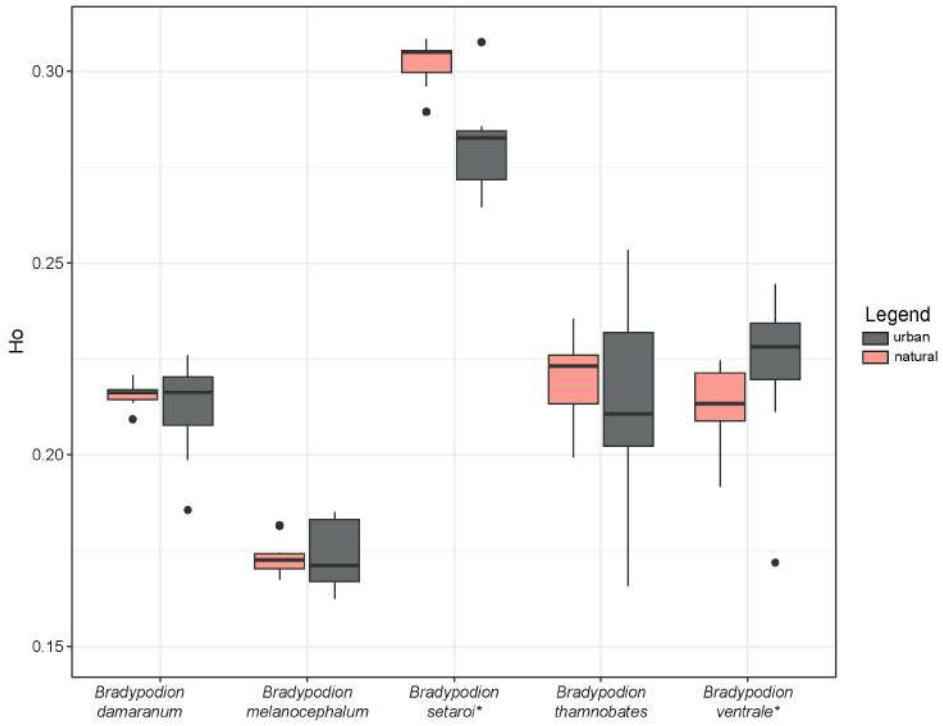


Figure 4.4: Genome-wide heterozygosity for each urban and natural population pair for five species of *Bradypodion*. Significant differences in heterozygosity between populations denoted by * ($p \leq 0.05$)

When assessing within-population nucleotide diversity (π), there are no clear differences between urban and natural populations of *Bradypodion* (Supplementary Figs S2-S6). Only two species had populations with significant differences in observed heterozygosity between populations (Fig. 4.4), *B. setaroi* and *B. ventrale*. While the urban population of *B. setaroi* had lower levels of heterozygosity than its natural counterpart, higher levels of heterozygosity were present in the urban population of *B. ventrale*.

Typically, Tajima's D is used to assess whether a segment of DNA sequence is evolving randomly or evolving under a non-random process, such as directional or balancing selection, demographic processes, hitchhiking, or introgression (Tajima 1989). In general, the

Table 4.1: Genome-wide genetic diversity for *Bradypodion* from urban and natural habitats. Given are the allele diversity (π), observed heterozygosity (H_o), and Tajima's D.

Species	**Habitat**	π	$\mathbf{H_o}$	**Tajima's D**
Bradypodion damaranum	urban	0.0071	0.211	0.049
	natural	0.0072	0.215	-0.143
Bradypodion melanocephalum	urban	0.0116	0.173	-0.510
	natural	0.0099	0.172	-0.087
Bradypodion setaroi	urban	0.0043	0.277	0.216
	natural	0.0046	0.301	0.368
Bradypodion thamnobates	urban	0.0067	0.214	-0.141
	natural	0.0070	0.221	-0.232
Bradypodion ventrale	urban	0.0068	0.229	-0.074
	natural	0.0062	0.213	0.249

expectation is that true signals for departures from neutrality yield Tajima's D $\geq |2|$ (Tajima, 1989; Hahn, 2019), with no expectation to recover signals genome-wide. Even so, urban and natural populations had different departure trajectories of Tajima's D when assessed genome-wide (Table 4.1). Specifically, both *B. melanocephalum* and *B. thamnobates* populations have negative Tajima's D statistics (Table 4.1). Negative values can result from different demographic processes including a recent selective sweep or population expansion after a recent bottleneck (Table 4.1). In contrast, the urban and natural populations of *B. setaroi* both exhibit positive Tajima's D indicating that both populations are likely experiencing balancing selection and maintaining genetic diversity. Populations in *B. damaranum* and *B. ventrale* display alternate patterns of Tajima's D, with the natural population of *B. damaranum* shown as negative and the urban population being positive, while the opposite is shown for populations of *B. ventrale* (Table 4.1). Fine-scale evaluation of Tajima's D across 100 Kb windows indicates similar distributions of Tajima's D between both urban and natural populations in all species. However, it is noted that a greater number of windows in urban populations tend toward negative Tajima's D $\leq$ -2 indicative of selective sweeps or expansion after a recent bottleneck, particularly in *B. damaranum* (Supplementary Fig. S7), *B. setaroi* (Supplementary Fig. S8), and *B. ventrale* (Supplementary Fig. S9). Although this pattern is not as clear in *B. melanocephalum* (Supplementary Fig. S10) or *B. thamnobates* (Supplementary Fig. S11), there is still an overall negative trend of Tajima's D in urban populations.

Table 4.2: Summary of SNPs and genes detected in outlier analyses for *Bradypodion*. These include the species sample sizes (n), total filtered SNPs, number of outlier SNPs detected in the principal components analysis (PCA), and genotype-environment association test (GEA). Outlier genes are summarised by the number of genes identified and total genes with known symbols

			Outlier SNPs			Outlier genes	
Species	***n***	**Filtered SNPs**	**PCA**	**GEA**	**Overlap**	**Identified genes**	**Gene symbols**
B. damaranum	17	1 556 562	553	15 775	487	430	48
B. melanocephalum	18	1 519 773	211	15 399	170	157	18
B. setaroi	15	1 194 306	128	12 052	49	41	5
B. thamnobates	17	1 227 109	622	12 450	15	15	2
B. ventrale	20	1 304 659	277	13 242	441	207	21

4.3.3 Selection scan and functional term estimation analyses

After filtering, the number of total SNPs retained for outlier analyses ranged between 1 194 306 - 1 556 562 sites per species (Table 4.2). For the PCA conducted on each species, at least one principal component accounted for the variation that distinguished urban and natural populations (Supplementary Figs. S12-S15), with the exception of *Bradypodion thamnobates* (Supplementary Fig. S16).

When assessing outliers, the greatest number of outlier SNPs, across both the PCA and GEA analyses, were identified in *B. damaranum* with 487 SNPs resulting in 430 unique genes, 48 of which have known gene symbols with associated functional information. Following in the number of outliers identified, 441 outlier SNPs were detected in *B. ventrale* with 207 associated genes, 21 of which have known gene symbols. From the 170 outlier SNPs highlighted in *B. setaroi*, only 18 outlier genes had known symbols. The lowest number of outlier SNPs were recovered in *B. melanocephalum* and *B. thamnobates* which resulted in only five and two genes with known symbols, respectively (Table 4.2).

When assessing biologically relevant Gene Ontology (GO) terms, functional terms were broadly assigned across three pathways, namely immune system function, organism development, and responses to environmental stimuli. Outlier genes detected in *B. damaranum* had functional associations in all three described pathways (Table S1). Pathways associated with the immune system directly involved responses to chemokine (GO:1990869) and cytokine stimuli (GO:0071345). Chemokine proteins facilitate the migration of white

blood cells within tissues as part of an immune response and cytokines are involved in mediating the inflammatory response within the organism. Additional GO terms for *B. damaranum* outliers were also associated with antiviral defence (GO:0051607) and endogenous growth factor stimulus linked to healing (GO:0071363). Outlier genes underlying organism development encompassed both ectoderm development (GO:0007398), responsible for the formation of the epidermis and the nervous system, and osteoblast differentiation (GO:0001649), whereby unspecialized cells gain the specialized features of an osteoblast, a mesodermal or neural crest cell that gives rise to bone. Functional GO terms linked to stress response (GO:0080134) were also detected in genes underlying the population differences of *B. damaranum*. In particular, *CTNNB1* was shown to downregulate the programmed cell death of neurons in response to oxidative stress (GO:1903377).

In the sister species *B. melanocephalum* and *B. thamnobates*, outlier genes were similarly associated with underlying genetic pathways which directly influenced brain development (Table S2-S3). In *B. melanocephalum*, outlier genes were shown to be involved in the development of the thalamus (GO:0021794), responsible for all incoming motor and sensory signals. Whereas in *B. thamnobates*, genetic pathways were associated with the development of the pallium (GO:0021543) and cerebral cortex (GO:0021987), responsible for higher-level processing, reasoning, learning and decision-making. Genes underlying sensory perception in response to environmental stimuli appeared to have genetic pathway differences in urban and natural populations of *Bradypodion*. In *B. setaroi*, *Cfap69* contained allele differences between these populations and was mainly associated with chemosensory behaviour (GO:0007635), particularly in the olfactory system (GO:0042048) and in response to odorants (GO:1990834, Table S4). In *B. ventrale*, population differences in genes associated with the sensory perception of sound were detected (GO:0007605, Table S5). The biological process involves the detection of mechanical stimulus relating to sound (GO:0050910), with the underlying pathway involved during the reception of auditory stimuli and converting this to a molecular signal for recognition and characterization of the signal. These genes are also associated with phenotypes in humans linked to sensory hypersensitivity (HP:5200058). Additional genes underlying population differences in *B. ventrale* are also underlying processes linked to the immune system, particularly mast cell activation (GO:0002279).

4.3.4 Isolation by Distance

There was no significant signal of isolation by distance (IBD) for any of the species (Table 4.3). This suggests that the population structure observed between the urban and natural populations for each species is the result of reduced gene flow between these habitats, or of local adaptation. Significant IBD would have suggested that the genetic differences observed are the result of a proportional reduction in relatedness as geographic distance increases due to the natural resistance over the landscape that restricts gene flow.

Table 4.3: Results of the Mantel test for the relationship between genetic and geographic distance between urban and natural population sample sites of *Bradypodion*. Pearson's correlation coefficient (ρ) and significance value (P) are given. ns - not significant

Species	ρ	$\boldsymbol{P}$
Bradypodion damaranum	-0.016	0.464 ns
Bradypodion melanocephalum	0.044	0.367 ns
Bradypodion setaroi	-0.191	0.955 ns
Bradypodion thamnobates	0.122	0.119 ns
Bradypodion ventrale	0.074	0.108 ns

4.4 Discussion

The findings presented here provide compelling support for urban and natural populations exhibiting distinct patterns of genetic variation. After accounting for population differentiation through the assessment of traditional metrics, contrasting population-level differences led to the identification of genomic outlier loci exhibiting significantly divergent allele frequencies between urban and natural populations. I next identified genes containing those outlier loci and used public gene ontology databases to connect outlier genes with their organismal functions and detect categories of gene function that were over-represented among our outlier genes. This then allowed me to identify connections between outlier genomic loci I detected, and the phenotypic traits known to be associated with adaptation in urban habitats.

While clear evidence of allele frequency differentiation between urban and natural populations was found, it cannot be said for certain that this is due to directional or disruptive selection experienced by these populations. These analyses are unable to indicate

which of the two tested populations experienced divergence, but rather point out those loci for which substantial divergence has occurred in one population. However, there are several reasons to suspect the bulk of outlier loci are the result of adaptive evolution operating in urban populations. First, the outlier loci appear to be driven by natural selection rather than simple population divergence. If there was evidence of reduced gene flow between urban and natural populations we would expect these populations to exhibit divergence genome-wide due to founder effects or genetic drift. However, given that F_{ST} is relatively low across the genome, the outlier peaks in F_{ST} observed between urban and natural populations suggest these regions of the genome are more likely impacted by recent selection. Second, since these lizards have only occupied urban habitats fairly recently in terms of evolutionary time, with the modern establishment of these urban settlements beginning roughly 200 years ago (Western Cape Government, 2021), urban populations more likely experienced novel selection pressures relative to their natural counterparts. Finally, the gene functions over-represented among outlier loci are tied to phenotypic traits that appear to be associated with urban evolution in *Bradypodion* chameleons. This presents the opportunity to generate new hypotheses about the genomic basis of phenotypic traits shown to be associated with urban adaptation. Subsequent functional genomic work could be performed to test the hypotheses generated here via targeted gene modification techniques such as CRISPR-Cas9-mediated gene editing and assessment of the phenotypes generated by those modifications (Rasys et al., 2019).

In urban environments, species able to persist within these habitats overcome urban pressures through behavioural, physiological, and morphological responses (Johnson and Munshi-South, 2017). In birds, biological predictors of successful urban living include lack of plumage dichromatism (Iglesias-Carrasco et al., 2019), higher body mass (Croci et al., 2008), large brain size (Maklakov et al., 2011), and broad environmental tolerance (Bonier et al., 2007; Lowry et al., 2013). In reptiles, adaptive phenotypes associated with urban living are mostly identified as morphological solutions, especially evident in taxa able to effectively use man-made substrates with traits including differing limb lengths and body sizes (Littleford-Colquhoun et al., 2017; Winchell et al., 2018, but see Batabyal and Thaker, 2019). In *Bradypodion*, urban populations were shown to have head dimension and ornamentation differences relative to their natural counterparts (Petford et al., 2023). The findings presented in this chapter provide evidence that there may be an optimal urban phenotype for *Bradypodion*, and that through the functional term analysis of outlier loci, it appears that

traits linked to three broad categories might differ between populations, organism development, immune system function, and responses to external stimuli.

Disentangling how certain species can adapt to urban environments while others decline and become locally extinct remains a complex issue (Sih et al., 2011; Sol et al., 2013; Lee and Thornton, 2021). In addition to ecological and life history factors, behavioural regulation has been shown to be a crucial factor in facilitating the exploitation of urban habitats, enabling species to capitalize on new opportunities and effectively respond to novel threats (Wright et al., 2010; Sol et al., 2013, 2020; Ducatez et al., 2020). Cognitive abilities that determine animals' ability to process information from the urban environment are likely to influence the potential for behavioural adjustments. A growing number of studies compare urban-rural differences in cognitive performance (Sol et al., 2013; Griffin et al., 2017), with many cases of greater cognitive performance shown in urban populations over rural counterparts (Federspiel et al., 2017; Kozlovsky et al., 2017; Preiszner et al., 2017; Batabyal and Thaker, 2019, but see Papp et al., 2015; Kang et al., 2018). Outlier genes detected in this study highlighted two pathways involved with brain development in *Bradypodion*, particularly the development of the thalamus in *B. melanocephalum*, and the pallium and cerebral cortex in *B. thamnobates*. These structures of the brain are directly responsible for interpreting motor and sensory signals as well as higher-level processing, reasoning, and decision-making (Javed et al., 2019). Given the scale and intensity of environmental stimuli in urban systems (Johnson and Munshi-South, 2017), having the cognitive capacity to deal with intense stimuli is likely an adaptive trait in these dwarf chameleons. This aligns with the expectation that species possessing cognitive abilities that allow them to track and respond flexibly to new information may be more successful in novel environments (Dridi and Lehmann, 2016; Lee and Thornton, 2021).

In natural environments, habitat quality is related to habitat complexity i.e., the structural complexity of vegetation (Tews et al., 2004; McElhinny et al., 2005). In contrast, urban environments are structurally homogeneous (paved surfaces such as roads and buildings, Groffman et al., 2014), with green spaces, such as parks or gardens, often being the most structurally complex (Haddad et al., 2015; Dubois and Cheptou, 2017). Green spaces are often highly fragmented in urban spaces, which poses numerous challenges encountered by species restricted to these habitats (Sandström et al., 2006; Vergnes et al., 2013; Munshi-South and Richardson, 2020). The Knysna Dwarf Chameleon, *Bradypodion damaranum*, occurs naturally in Afrotemperate forests (Tolley and Burger, 2007), but recent evidence has shown differences for some traits in populations occurring in urban habitats

(Petford et al., 2023). These include a reduction in ornamentation, a distinguishing feature of closed-canopy forest chameleons (Tolley and Burger, 2007; Tolley et al., 2008; Measey et al., 2009; da Silva and Tolley, 2013), as well as higher and wider heads (Petford et al., 2023). Outlier genes involved with the development of skin and bone detected in this chapter could relate to the development for these traits in urban populations, but this requires further assessment. It was also shown that urban populations of *B. damaranum* had significantly more injuries than their natural counterparts (Petford et al., 2023), potentially promoting selection for appropriate healing strategies in urban populations. While an over-representation of genes associated with immune and inflammatory responses (e.g., facilitated by chemokine and cytokine proteins; Abdulkhaleq et al., 2018; Harvanová et al., 2023), it is unclear if this was driven by directional selection in urban populations. Additional outlier genes were also associated with antiviral defence and endogenous growth factor stimulus linked to healing as well as responses to oxidative stress. However, this may suggest healing strategies may not exclusively be related to combat-related injuries, but also be a response to novel predators or pathogens in urban environments (Poljsak, 2011; Giraudeau et al., 2014; Isaksson, 2015).

Urban environments are contaminated with different forms of pollution, including air and soil pollution, as well as light and noise pollution. Recently, increasing efforts have been devoted to understanding the impact of toxicity of inhalable particulate matter on immune function in wild living animals (Li et al., 2021). In addition, there is evidence to suggest that air pollution not only inhibits airway immunity and induces inflammation, but also generates oxidative stress (Ling and van Eeden, 2009; Arias-Pérez et al., 2020; Minias, 2023). While over-represented genes relating to inflammation and oxidative stress inhibition were found in *Bradypodion damaranum*, responses to olfactory stimuli were found in *B. setaroi*, likely in response to air pollution. A recent study has shown that, in mammals, olfactory dysfunctions related to air pollution drove over-representation in similar genes to those detected in this study (Kim et al., 2020). *Bradypodion setaroi* inhabits the coastal forests of St. Lucia, Kwa-Zulu Natal(Tolley and Burger, 2007) , which is characterised by its estuarine system (Perissinotto et al., 2013). A recent study found that, while South African estuaries face a growing threat of pollution, pollution in these environments has mostly been associated with pollutants within the water systems (Govender et al., 2020; Rajkaran, 2022). Given this, it is unlikely that the over-representation of genes associated with olfactory stimuli in *B. setaroi* is in response to air pollutants, especially since pollution from residential, tourism and industrial activities in St Lucia are relatively minimal (Rajkaran, 2022). It should be mentioned that, while no direct

links between olfactory function and air pollution were found in *Bradypodion*, assessments of the impacts affecting reptiles from such pollutants are lacking considering the known effects identified in other groups (Llacuna et al., 1993; Lovett et al., 2009; Sanderfoot and Holloway, 2017, but see Read, 1998).

Noise pollution has been recognized as an additional major pollutant in urban environments and a potential evolutionary force (Swaddle et al., 2015). Assessing the impact of noise disturbance on urban wildlife has largely been the focus of sensory disturbance research efforts, with findings routinely reporting negative effects on species diversity and richness (Stone, 2000; Rheindt, 2003; Francis et al., 2011). In urban populations of many acoustically communicating species, a prominent response to noise pollution has been to avoid any acoustic masking of their signals. This is achieved by altering or adjusting their acoustic frequency of signals, often to a higher frequency than the background noise (Slabbekoorn and den Boer-Visser, 2006; Nemeth and Brumm, 2009), or behaviourally shifting activity times when calling (Halfwerk et al., 2019b; Lukanov and Naumov, 2019, but see Mitchell et al., 2020). When the effect of anthropogenic noise at different decibels and frequencies on behaviour was assessed in blue-tongued skinks, *Tiliqua scincoides*, findings indicated that loud and high-frequency noises resulted in these lizards spending more time stuck in position, indicative of a typical stress response in reptiles (Alarcon and Fabiola, 2016). While the ears in chameleons are greatly reduced (Anderson and Higham, 2013), they are still able to detect airborne sound at lower frequencies, roughly $\tilde{2}$00-600 Hz (Wever, 1968, 1978; Olroyd, 2022). In *Bradypodion ventrale*, over-represented genes were associated with the sensory perception of sound were detected, particularly regarding the mechanical recognition of sound frequencies. Low-frequency noises generated by industry and transport are ubiquitous in urban environments (Ravinder et al., 2014; Okokon et al., 2017; Bąkowski et al., 2019; Soni et al., 2022), and likely mask the lower frequency sounds detectable by these chameleons. It should be noted that the same genes identified in *B. ventrale* are also associated with sensory hypersensitivity in mammals (Ruan et al., 2005; Driessen et al., 2017). While the association may not be the same in these chameleons, there is likely still an adaptive response to excessive noise stimuli, however, uncovering the sources of these stimuli warrants further investigation.

The findings presented in this chapter provide further evidence that alleles underlying adaptation can rapidly emerge in response to selection, contrary to prior belief (Gillespie, 1991; Messer et al., 2016). In addition, the diversity in traits detected by outlier genes would suggest that there may not be a universal solution to urbanisation in these chameleons. Given that

these species differ in life history strategies, it may be that the differences in standing variation result in different solutions to similar selection pressures. However, whether these loci under selection are a result of pressures in urban environments remains to be tested. Therefore, assessing whether these over-represented genes are present to the same degree in additional urban populations, using a replicated approach, would further indicate if these responses are due to urban pressures.

4.5 Future Directions

Being able to detect signatures of natural selection is crucial to understanding local adaption in any environment (Kawecki and Ebert, 2004). In a genetic context, uncovering the architecture underlying adaptive traits is inherently difficult given the complex structural nature of the genome (Rhie et al., 2021). In addition, disentangling selection acting on single versus multiple correlated traits can be particularly difficult (Winchell et al., 2022). Identifying genes associated with many complex traits can be challenging due to polygenic adaptation's suspected prevalence and significance (Rockman, 2012). Moreover, rapid adaptation may involve soft rather than hard selective sweeps, further complicating the process (Messer and Petrov, 2013). The findings in this chapter are likely a result of hard selective sweeps given the genetic variation unique to urban populations. However, if this is the case, assessment of soft selective sweeps may uncover additional sites of selection undetected by the approach used here. With the recent success and surge of machine learning and deep learning-based methods (Zhang et al., 2020; Xue et al., 2021; Kumar et al., 2022), leveraging these tool's sensitivity to detect selective sweeps may be most appropriate in uncovering signals of soft sweeps. Many of the genes and associated traits highlighted in this chapter remain unverified in reptiles, with links to appropriate phenotypes based on the known phenotypes in other taxa. Given the successes in gene editing techniques applicable to reptiles (CRISPR-based gene editing, Rasys et al., 2019), functionally validating candidate genes targeted by selection is now an approach worth consideration.

References

Abdulkhaleq, L., Assi, M., Abdullah, R., Zamri-Saad, M., Taufiq-Yap, Y., and Hezmee, M. (2018). The crucial roles of inflammatory mediators in inflammation: A review. *Veterinary world*, 11(5):627.

Akerman, A. and Bürger, R. (2014). The consequences of gene flow for local adaptation and differentiation: a two-locus two-deme model. *Journal of mathematical biology*, 68(5):1135–1198.

Alarcon, M. and Fabiola, K. (2016). *Effects of anthropogenic noise on the behaviour, physiological traits and welfare of two animal models: wild mice (Mus musculus) and Eastern blue tongued lizard (Tiliqua scincoides)*. PhD thesis, The University of Queensland.

Anderson, C. A. and Higham, T. E. (2013). Chameleon Anatomy. In Tolley, K. A. and Herrel, A., editors, *The Biology of Chameleons*. University of California Press, USA.

Andrews, S. (2010). Fastqc: a quality control tool for high throughput sequence data.

Arias-Pérez, R. D., Taborda, N. A., Gómez, D. M., Narvaez, J. F., Porras, J., and Hernandez, J. C. (2020). Inflammatory effects of particulate matter air pollution. *Environmental Science and Pollution Research*, 27(34):42390–42404.

Bąkowski, A., Radziszewski, L., Dekỳš, V., and Świetlik, P. (2019). Frequency analysis of urban traffic noise. In *2019 20th International Carpathian Control Conference (ICCC)*, pages 1–6. IEEE.

Barrett, R. D. and Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in ecology & evolution*, 23(1):38–44.

Batabyal, A. and Thaker, M. (2019). Lizards from suburban areas learn faster to stay safe. *Biology Letters*, 15(2):20190009.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*, 30(15):2114–2120.

Bonier, F., Martin, P. R., and Wingfield, J. C. (2007). Urban birds have broader environmental tolerance. *Biology letters*, 3(6):670–673.

Bushnell, B. (2021). Bbmap (version 38.90).

Butlin, R. K., Saura, M., Charrier, G., Jackson, B., André, C., Caballero, A., Coyne, J. A., Galindo, J., Grahame, J. W., Hollander, J., et al. (2014). Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. *Evolution*, 68(4):935–949.

Card, D. C., Jennings, W. B., and Edwards, S. V. (2023). Genome evolution and the future of phylogenomics of non-avian reptiles. *Animals*, 13(3):471.

Carvalho, J., Faria, R., Butlin, R. K., and Sousa, V. C. (2023). Parallel evolution of local adaptation in *Littorina saxatilis* inferred with whole-genome pool-seq data. *BioRxiv*, pages 2023–09.

Croci, S., Butet, A., and Clergeau, P. (2008). Does urbanization filter birds on the basis of their biological traits. *The Condor*, 110(2):223–240.

Cruickshank, T. E. and Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular ecology*, 23(13):3133–3157.

da Silva, J. M. and Tolley, K. A. (2013). Ecomorphological variation and sexual dimorphism in a recent radiation of dwarf chameleons (*Bradypodion*). *Biological Journal of the Linnean Society*, 109(1):113–130.

Dabney, A. and Storey, J. D. (2011). Package 'qvalue'. *Medicine*, 344:539–548.

DeCandia, A. L., Brzeski, K. E., Heppenheimer, E., Caro, C. V., Camenisch, G., Wandeler, P., Driscoll, C., and vonHoldt, B. M. (2019). Urban colonization through multiple genetic lenses: The city-fox phenomenon revisited. *Ecology and evolution*, 9(4):2046–2060.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., Del Angel, G., Rivas, M. A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation dna sequencing data. *Nature genetics*, 43(5):491–498.

Diamond, S. E. and Martin, R. A. (2021). Evolution in cities. *Annual Review of Ecology, Evolution, and Systematics*, 52:519–540.

Dridi, S. and Lehmann, L. (2016). Environmental complexity favors the evolution of learning. *Behavioral Ecology*, 27(3):842–850.

Driessen, A. K., McGovern, A. E., Narula, M., Yang, S.-K., Keller, J. A., Farrell, M. J., and

Mazzone, S. B. (2017). Central mechanisms of airway sensation and cough hypersensitivity. *Pulmonary pharmacology & therapeutics*, 47:9–15.

Dubois, J. and Cheptou, P.-O. (2017). Effects of fragmentation on plant adaptation to urban environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1712):20160038.

Ducatez, S., Sol, D., Sayol, F., and Lefebvre, L. (2020). Behavioural plasticity is associated with reduced extinction risk in birds. *Nature Ecology & Evolution*, 4(6):788–793.

Duforet-Frebourg, N., Luu, K., Laval, G., Bazin, E., and Blum, M. G. (2016). Detecting genomic signatures of natural selection with principal component analysis: application to the 1000 genomes data. *Molecular biology and evolution*, 33(4):1082–1093.

Federspiel, I. G., Garland, A., Guez, D., Bugnyar, T., Healy, S. D., Güntürkün, O., and Griffin, A. S. (2017). Adjusting foraging strategies: a comparison of rural and urban common mynas (*Acridotheres tristis*). *Animal cognition*, 20:65–74.

Francis, C. D., Ortega, C. P., and Cruz, A. (2011). Vocal frequency change reflects different responses to anthropogenic noise in two suboscine tyrant flycatchers. *Proceedings of the Royal Society B: Biological Sciences*, 278(1714):2025–2031.

Francis, R. M. (2017). Pophelper: An R package and web app to analyse and visualize population structure. *Molecular ecology resources*, 17(1):27–32.

Fusco, N. A., Pehek, E., and Munshi-South, J. (2021). Urbanization reduces gene flow but not genetic diversity of stream salamander populations in the new york city metropolitan area. *Evolutionary Applications*, 14(1):99–116.

Gable, S. M., Mendez, J. M., Bushroe, N. A., Wilson, A., Byars, M. I., and Tollis, M. (2023). The State of Squamate Genomics: Past, Present, and Future of Genome Research in the Most Speciose Terrestrial Vertebrate Order. *Genes*, 14(7):1387.

Garcia, C. M., Suárez-Rodríguez, M., and López-Rull, I. (2017). Becoming citizens: avian adaptations to urban life. *Ecology and conservation of birds in urban environments*, pages 91–112.

Ghorbani, H. (2019). Mahalanobis distance and its application for detecting multivariate outliers. *Facta Universitatis, Series: Mathematics and Informatics*, pages 583–595.

Gillespie, J. H. (1991). *The causes of molecular evolution*, volume 2. Oxford University Press, USA.

Giraudeau, M., Mousel, M., Earl, S., and McGraw, K. (2014). Parasites in the city: degree of urbanization predicts poxvirus and coccidian infections in house finches (*Haemorhous mexicanus*). *PloS one*, 9(2):e86747.

Gogarten, S. M., Sofer, T., Chen, H., Yu, C., Brody, J. A., Thornton, T. A., Rice, K. M., and Conomos, M. P. (2019). Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics*, 35(24):5346–5348.

Govender, J., Naidoo, T., Rajkaran, A., Cebekhulu, S., Bhugeloo, A., and Sershen (2020). Towards characterising microplastic abundance, typology and retention in mangrove-dominated estuaries. *Water*, 12(10):2802.

Griffin, A. S., Netto, K., and Peneaux, C. (2017). Neophilia, innovation and learning in an urbanized world: a critical evaluation of mixed findings. *Current Opinion in Behavioral Sciences*, 16:15–22.

Groffman, P. M., Cavender-Bares, J., Bettez, N. D., Grove, J. M., Hall, S. J., Heffernan, J. B., Hobbie, S. E., Larson, K. L., Morse, J. L., Neill, C., et al. (2014). Ecological homogenization of urban USA. *Frontiers in Ecology and the Environment*, 12(1):74–81.

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., Lovejoy, T. E., Sexton, J. O., Austin, M. P., Collins, C. D., et al. (2015). Habitat fragmentation and its lasting impact on earth's ecosystems. *Science advances*, 1(2):e1500052.

Hahn, M. H. (2019). *Molecular Population Genetics*. Oxford University Press, USA.

Halfwerk, W., Blaas, M., Kramer, L., Hijner, N., Trillo, P. A., Bernal, X. E., Page, R. A., Goutte, S., Ryan, M. J., and Ellers, J. (2019a). Adaptive changes in sexual signalling in response to urbanization. *Nature ecology & evolution*, 3(3):374–380.

Halfwerk, W., Blaas, M., Kramer, L., Hijner, N., Trillo, P. A., Bernal, X. E., Page, R. A., Goutte, S., Ryan, M. J., and Ellers, J. (2019b). Adaptive changes in sexual signalling in response to urbanization. *Nature ecology & evolution*, 3(3):374–380.

Harvanová, G., Duranková, S., and Bernasovská, J. (2023). The role of cytokines and

chemokines in the inflammatory response. *Alergologia Polska-Polish Journal of Allergology*, 10(3):210–219.

Hendry, A. P., Farrugia, T. J., and Kinnison, M. T. (2008). Human influences on rates of phenotypic change in wild animal populations. *Molecular ecology*, 17(1):20–29.

Hijmans, R. J., Williams, E., Vennes, C., and Hijmans, M. R. J. (2017). Package 'geosphere'. *Spherical trigonometry*, 1(7):1–45.

Iglesias-Carrasco, M., Duchêne, D. A., Head, M. L., Møller, A. P., and Cain, K. (2019). Sex in the city: Sexual selection and urban colonization in passerines. *Biology letters*, 15(9):20190257.

Isaksson, C. (2015). Urbanization, oxidative stress and inflammation: a question of evolving, acclimatizing or coping with urban environmental stress. *Functional Ecology*, 29(7):913–923.

Javed, K., Reddy, V., and Lui, F. (2019). *Neuroanatomy, Cerebral Cortex*. StatPearls Publishing LLC.

Johnson, M. T. and Munshi-South, J. (2017). Evolution of life in urban environments. *Science*, 358(6363):eaam8327.

Kang, F., Goulet, C. T., and Chapple, D. G. (2018). The impact of urbanization on learning ability in an invasive lizard. *Biological Journal of the Linnean Society*, 123(1):55–62.

Kawecki, T. J. and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology letters*, 7(12):1225–1241.

Kim, S. J., Kim, N., Park, S. H., Kim, H. S., Song, J.-J., Son, B.-S., Jang, A.-S., Park, M. K., and Seo, Y. R. (2020). Genomic approach to explore altered signaling networks of olfaction in response to diesel exhaust particles in mice. *Scientific Reports*, 10(1):16972.

Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., and Peterson, H. (2023). g: Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping. *Nucleic Acids Research*, page gkad347.

Kozlovsky, D. Y., Weissgerber, E. A., and Pravosudov, V. V. (2017). What makes specialized food-caching mountain chickadees successful city slickers? *Proceedings of the Royal Society B: Biological Sciences*, 284(1855):20162613.

Kumar, H., Panigrahi, M., Panwar, A., Rajawat, D., Nayak, S. S., Saravanan, K., Kaisa, K., Parida, S., Bhushan, B., and Dutt, T. (2022). Machine-learning prospects for detecting selection signatures using population genomics data. *Journal of Computational Biology*, 29(9):943–960.

Lee, V. E. and Thornton, A. (2021). Animal cognition in an urbanised world. *Frontiers in ecology and evolution*, 9:633947.

Legendre, P., Fortin, M.-J., and Borcard, D. (2015). Should the mantel test be used in spatial analysis? *Methods in Ecology and Evolution*, 6(11):1239–1247.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in ecology & evolution*, 17(4):183–189.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with bwa-mem. *arXiv preprint arXiv:1303.3997*.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, . G. P. D. P. (2009). The sequence alignment/map format and samtools. *bioinformatics*, 25(16):2078–2079.

Li, M., Nabi, G., Sun, Y., Wang, Y., Wang, L., Jiang, C., Cao, P., Wu, Y., and Li, D. (2021). The effect of air pollution on immunological, antioxidative and hematological parameters, and body condition of eurasian tree sparrows. *Ecotoxicology and environmental safety*, 208:111755.

Ling, S. H. and van Eeden, S. F. (2009). Particulate matter air pollution exposure: role in the development and exacerbation of chronic obstructive pulmonary disease. *International journal of chronic obstructive pulmonary disease*, pages 233–243.

Littleford-Colquhoun, B. L., Clemente, C., Whiting, M. J., Ortiz-Barrientos, D., and Frère, C. H. (2017). Archipelagos of the Anthropocene: rapid and extensive differentiation of native terrestrial vertebrates in a single metropolis. *Molecular Ecology*, 26(9):2466–2481.

Llacuna, S., Gorriz, A., Durfort, M., and Nadal, J. (1993). Effects of air pollution on passerine birds and small mammals. *Archives of Environmental Contamination and Toxicology*, 24:59–66.

Lovett, G. M., Tear, T. H., Evers, D. C., Findlay, S. E., Cosby, B. J., Dunscomb, J. K.,

Driscoll, C. T., and Weathers, K. C. (2009). Effects of air pollution on ecosystems and biological diversity in the eastern United States. *Annals of the New York Academy of Sciences*, 1162(1):99–135.

Lowry, H., Lill, A., and Wong, B. B. (2013). Behavioural responses of wildlife to urban environments. *Biological reviews*, 88(3):537–549.

Lukanov, S. and Naumov, B. (2019). Effect of anthropogenic noise on call parameters of *Hyla arborea* (Anura: Hylidae). *Ecological Questions*, 30(1):55–60.

Luu, K., Bazin, E., and Blum, M. G. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular ecology resources*, 17(1):67–77.

Magidi, J. and Ahmed, F. (2022). Monitoring vegetation phenology using MODIS NDVI 250m in the City of Tshwane, South Africa. *South African Journal of Geomatics*, 11(2):176–189.

Maklakov, A. A., Immler, S., Gonzalez-Voyer, A., Rönn, J., and Kolm, N. (2011). Brains and the city: big-brained passerine birds succeed in urban environments. *Biology letters*, 7(5):730–732.

Martins, H., Caye, K., Luu, K., Blum, M. G., and François, O. (2016). Identifying outlier loci in admixed and in continuous populations using ancestral population differentiation statistics. *Molecular ecology*, 25(20):5029–5042.

McCulloch, G. A. and Waters, J. M. (2023). Rapid adaptation in a fast-changing world: Emerging insights from insect genomics. *Global Change Biology*, 29(4):943–954.

McElhinny, C., Gibbons, P., Brack, C., and Bauhus, J. (2005). Forest and woodland stand structural complexity: its definition and measurement. *Forest Ecology and Management*, 218(1-3):1–24.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9):1297–1303.

McLachlan, G. J. (1999). Mahalanobis distance. *Resonance*, 4(6):20–26.

Measey, G. J., Hopkins, K., and Tolley, K. A. (2009). Morphology, ornaments and performance in two chameleon ecomorphs: is the casque bigger than the bite? *Zoology*, 112(3):217–226.

Messer, P. W., Ellner, S. P., and Hairston, N. G. (2016). Can population genetics adapt to rapid evolution? *Trends in Genetics*, 32(7):408–418.

Messer, P. W. and Petrov, D. A. (2013). Population genomics of rapid adaptation by soft selective sweeps. *Trends in ecology & evolution*, 28(11):659–669.

Miles, L. S., Rivkin, L. R., Johnson, M. T., Munshi-South, J., and Verrelli, B. C. (2019). Gene flow and genetic drift in urban environments. *Molecular ecology*, 28(18):4138–4151.

Minias, P. (2023). The effects of urban life on animal immunity: Adaptations and constraints. *Science of The Total Environment*, page 165085.

Mitchell, B. A., Callaghan, C. T., and Rowley, J. J. (2020). Continental-scale citizen science data reveal no changes in acoustic responses of a widespread tree frog to an urbanisation gradient. *Journal of Urban Ecology*, 6(1):juaa002.

Mousseau, T. A. and Møller, A. P. (2014). Genetic and ecological studies of animals in Chernobyl and Fukushima. *Journal of Heredity*, 105(5):704–709.

Munshi-South, J. and Richardson, J. L. (2020). Landscape Genetic Approaches to Understanding Movement. In Szulkin, M., Munshi-South, J., and Charmantier, A., editors, *Urban Evolutionary Biology*. Oxford University Press, USA.

Nemeth, E. and Brumm, H. (2009). Blackbirds sing higher-pitched songs in cities: adaptation to habitat acoustics or side-effect of urbanization? *Animal behaviour*, 78(3):637–641.

Nosil, P. (2012). *Ecological speciation*. Oxford Ecology and Evolution.

Okokon, E. O., Yli-Tuomi, T., Turunen, A. W., Taimisto, P., Pennanen, A., Vouitsis, I., Samaras, Z., Voogt, M., Keuken, M., and Lanki, T. (2017). Particulates and noise exposure during bicycle, bus and car commuting: A study in three European cities. *Environmental research*, 154:181–189.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'hara, R., Simpson, G. L., Solymos, P., et al. (2019). Package 'vegan'. *Community ecology package, version*, 2(9).

Olroyd, S. L. (2022). Independent origins of a novel atympanic middle ear system within Chamaeleonidae. *The Anatomical Record*, 305(12):3371–3384.

Palamuleni, L. G., Turyahikayo, A. K., et al. (2015). Vegetation density assessment in the Upper Molopo River Catchment, South Africa. *Journal of Food, Agriculture & Environment*, 13(2):190–196.

Papp, S., Vincze, E., Preiszner, B., Liker, A., and Bókony, V. (2015). A comparison of problem-solving success between urban and rural house sparrows. *Behavioral ecology and sociobiology*, 69:471–480.

Parsons, K., Rigg, A., Conith, A., Kitchener, A. C., Harris, S., and Zhu, H. (2020). Skull morphology diverges between urban and rural populations of red foxes mirroring patterns of domestication and macroevolution. *Proceedings of the Royal Society B*, 287(1928):20200763.

Perissinotto, R., Stretch, D. D., and Taylor, R. H. (2013). *Ecology and conservation of estuarine ecosystems: Lake St Lucia as a global model*. Cambridge University Press.

Perrier, C., Caizergues, A., and Charmantier, A. (2020). Adaptation Genomics in Urban Environments. In Szulkin, M., Munshi-South, J., and Charmantier, A., editors, *Urban evolutionary biology*. Oxford University Press, USA.

Petford, M. A., Herrel, A., Alexander, G. J., and Tolley, K. A. (2023). Differences between urban and natural populations of dwarf chameleons (*Bradypodion damaranum*): a case of urban warfare? *Urban Ecosystems*, pages 1–13.

Pina-Martins, F., Silva, D. N., Fino, J., and Paulo, O. S. (2017). Structure_threader: an improved method for automation and parallelization of programs structure, faststructure and maverick on multicore cpu systems. *Molecular ecology resources*, 17(6):e268–e274.

Poljsak, B. (2011). Strategies for reducing or preventing the generation of oxidative stress. *Oxidative medicine and cellular longevity*, 2011.

Preiszner, B., Papp, S., Pipoly, I., Seress, G., Vincze, E., Liker, A., and Bókony, V. (2017). Problem-solving performance and reproductive success of great tits in urban and forest habitats. *Animal Cognition*, 20:53–63.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J., et al. (2007). Plink: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3):559–575.

Putman, B. J., Gasca, M., Blumstein, D. T., and Pauly, G. B. (2019). Downsizing for downtown: limb lengths, toe lengths, and scale counts decrease with urbanization in western fence lizards (*Sceloporus occidentalis*). *Urban Ecosystems*, 22:1071–1081.

Raj, A., Stephens, M., and Pritchard, J. K. (2014). faststructure: variational inference of population structure in large snp data sets. *Genetics*, 197(2):573–589.

Rajkaran, A. (2022). South Africa's estuaries face a growing threat from pollution. *Water&Sanitation Africa*, 17(4):26–27.

Rasys, A. M., Park, S., Ball, R. E., Alcala, A. J., Lauderdale, J. D., and Menke, D. B. (2019). CRISPR-Cas9 gene editing in lizards through microinjection of unfertilized oocytes. *Cell reports*, 28(9):2288–2292.

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g: Profiler: a web server for functional enrichment analysis and conversions of gene lists. *Nucleic acids research*, 47(1):191–198.

Ravinder, L., Belachew, M. G., et al. (2014). Urban noise in a metropolitan towns. *Open Journal of Acoustics*, 4(04):163.

Read, J. (1998). Are geckos useful bioindicators of air pollution? *Oecologia*, 114(2):180–187.

Reid, N. M., Proestou, D. A., Clark, B. W., Warren, W. C., Colbourne, J. K., Shaw, J. R., Karchner, S. I., Hahn, M. E., Nacci, D., Oleksiak, M. F., et al. (2016). The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science*, 354(6317):1305–1308.

Rheindt, F. E. (2003). The impact of roads on birds: does song frequency play a role in determining susceptibility to noise pollution? *Journal für Ornithologie*, 144(3):295–306.

Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., et al. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, 592(7856):737–746.

Rockman, M. V. (2012). The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution*, 66(1):1–17.

Ruan, H.-B., Zhang, N., and Gao, X. (2005). Identification of a novel point mutation of mouse proto-oncogene c-kit through n-ethyl-n-nitrosourea mutagenesis. *Genetics*, 169(2):819–831.

Samia, D. S., Blumstein, D. T., Díaz, M., Grim, T., Ibáñez-Álamo, J. D., Jokimäki, J., Tätte, K., Markó, G., Tryjanowski, P., and Møller, A. P. (2017). Rural-urban differences in escape behavior of European birds across a latitudinal gradient. *Frontiers in Ecology and Evolution*, 5:66.

Sanderfoot, O. V. and Holloway, T. (2017). Air pollution impacts on avian species via inhalation exposure and associated outcomes. *Environmental Research Letters*, 12(8):083002.

Sandström, U., Angelstam, P., and Mikusiński, G. (2006). Ecological diversity of birds in relation to the structure of urban green space. *Landscape and urban planning*, 77(1-2):39–53.

Schluter, D. and Conte, G. L. (2009). Genetics and ecological speciation. *Proceedings of the National Academy of Sciences*, 106(1):9955–9962.

Schluter, D., Marchinko, K. B., Barrett, R., and Rogers, S. M. (2010). Natural selection and the genetics of adaptation in threespine stickleback. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1552):2479–2486.

Sih, A., Ferrari, M. C., and Harris, D. J. (2011). Evolution and behavioural responses to human-induced rapid environmental change. *Evolutionary applications*, 4(2):367–387.

Slabbekoorn, H. and den Boer-Visser, A. (2006). Cities change the songs of birds. *Current biology*, 16(23):2326–2331.

Smadja, C. M. and Butlin, R. K. (2011). A framework for comparing processes of speciation in the presence of gene flow. *Molecular ecology*, 20(24):5123–5140.

Sol, D., Lapiedra, O., and Ducatez, S. (2020). Cognition and adaptation to urban environments. In Szulkin, M., Munshi-South, J., and Charmantier, A., editors, *Urban evolutionary biology*. Oxford University Press, USA.

Sol, D., Lapiedra, O., and González-Lagos, C. (2013). Behavioural adjustments for a life in the city. *Animal behaviour*, 85(5):1101–1112.

Somers, C. M., McCarry, B. E., Malek, F., and Quinn, J. S. (2004). Reduction of particulate air pollution lowers the risk of heritable mutations in mice. *Science*, 304(5673):1008–1010.

Soni, A. R., Makde, K., Amrit, K., Vijay, R., and Kumar, R. (2022). Noise prediction and environmental noise capacity for urban traffic of Mumbai. *Applied Acoustics*, 188:108516.

Stickler, M. and Shackleton, C. (2015). Local wood demand, land cover change and the state of Albany thicket on an urban commonage in the Eastern Cape, South Africa. *Environmental management*, 55:411–422.

Stone, E. (2000). Separating the noise from the noise: a finding in support of the "niche hypothesis," that birds are influenced by human-induced noise in natural habitats. *Anthrozoös*, 13(4):225–231.

Swaddle, J. P., Francis, C. D., Barber, J. R., Cooper, C. B., Kyba, C. C., Dominoni, D. M., Shannon, G., Aschehoug, E., Goodwin, S. E., Kawahara, A. Y., et al. (2015). A framework to assess evolutionary responses to anthropogenic light and sound. *Trends in ecology & evolution*, 30(9):550–560.

Taft, J. M., Tolley, K. A., Alexander, G. J., and Geneva, A. J. (2023). De Novo Whole Genome Assemblies for Two Southern African Dwarf Chameleons (*Bradypodion*, Chamaeleonidae). *Genome biology and evolution*, 15(10):evad182.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by dna polymorphism. *Genetics*, 123(3):585–595.

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M. C., Schwager, M., and Jeltsch, F. (2004). Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. *Journal of biogeography*, 31(1):79–92.

Tigano, A. and Friesen, V. L. (2016). Genomics of local adaptation with gene flow. *Molecular ecology*, 25(10):2144–2164.

Tolley, K. and Burger, M. (2007). *Chameleons of Southern Africa*. Struik Publishers.

Tolley, K. A., Chase, B. M., and Forest, F. (2008). Speciation and radiations track climate transitions since the Miocene Climatic Optimum: a case study of southern African chameleons. *Journal of Biogeography*, 35(8):1402–1414.

Van der Auwera, G. A. and O'Connor, B. D. (2020). *Genomics in the cloud: using Docker, GATK, and WDL in Terra*. O'Reilly Media.

Vergnes, A., Kerbiriou, C., and Clergeau, P. (2013). Ecological corridors also operate in an urban matrix: a test case with garden shrews. *Urban Ecosystems*, 16:511–525.

Western Cape Government (2021). Socio-economic profile: George municipality.

Wever, E. (1978). *The reptile ear Princeton*. Princeton University Press New Jersey.

Wever, E. G. (1968). The ear of the chameleon: *Chamaeleo senegalensis* and *Chamaeleo quilensis*. *Journal of Experimental Zoology*, 168(4):423–436.

Winchell, K. M., Aviles-Rodriguez, K. J., Carlen, E. J., Miles, L. S., Charmantier, A., De León, L. F., Gotanda, K. M., Rivkin, L. R., Szulkin, M., and Verrelli, B. C. (2022). Moving past the challenges and misconceptions in urban adaptation research. *Ecology and Evolution*, 12(11):e9552.

Winchell, K. M., Battles, A. C., and Moore, T. Y. (2020). Terrestrial Locomotor Evolution in Urban Environments. In Szulkin, M., Munshi-South, J., and Charmantier, A., editors, *Urban evolutionary biology*. Oxford University Press, USA.

Winchell, K. M., Campbell-Staton, S. C., Losos, J. B., Revell, L. J., Verrelli, B. C., and Geneva, A. J. (2023). Genome-wide parallelism underlies contemporary adaptation in urban lizards. *Proceedings of the National Academy of Sciences*, 120(3):e2216789120.

Winchell, K. M., Maayan, I., Fredette, J. R., and Revell, L. J. (2018). Linking locomotor performance to morphological shifts in urban lizards. *Proceedings of the Royal Society B*, 285(1880):20180229.

Winchell, K. M., Reynolds, R. G., Prado-Irwin, S. R., Puente-Rolon, A. R., and Revell, L. J. (2016). Phenotypic shifts in urban areas in the tropical lizard *Anolis cristatellus*. *Evolution*, 70(5):1009–1022.

Wright, T. F., Eberhard, J. R., Hobson, E. A., Avery, M. L., and Russello, M. A. (2010). Behavioral flexibility and species invasions: the adaptive flexibility hypothesis. *Ethology Ecology & Evolution*, 22(4):393–404.

Xue, A. T., Schrider, D. R., and Kern, A. D. (2021). Discovery of ongoing selective sweeps within *Anopheles* mosquito populations using deep learning. *Molecular biology and evolution*, 38(3):1168–1183.

Yang, W., Feiner, N., Laakkonen, H., Sacchi, R., Zuffi, M. A., Scali, S., While, G. M., and Uller, T. (2020). Spatial variation in gene flow across a hybrid zone reveals causes of reproductive isolation and asymmetric introgression in wall lizards. *Evolution*, 74(7):1289–1300.

Yauk, C. L., Fox, G. A., McCarry, B. E., and Quinn, J. S. (2000). Induced minisatellite

germline mutations in herring gulls (*Larus argentatus*) living near steel mills. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 452(2):211–218.

Zhang, Y., Yan, J., Chen, S., Gong, M., Gao, D., Zhu, M., and Gan, W. (2020). Review of the applications of deep learning in bioinformatics. *Current Bioinformatics*, 15(8):898–911.

Supplemental figures

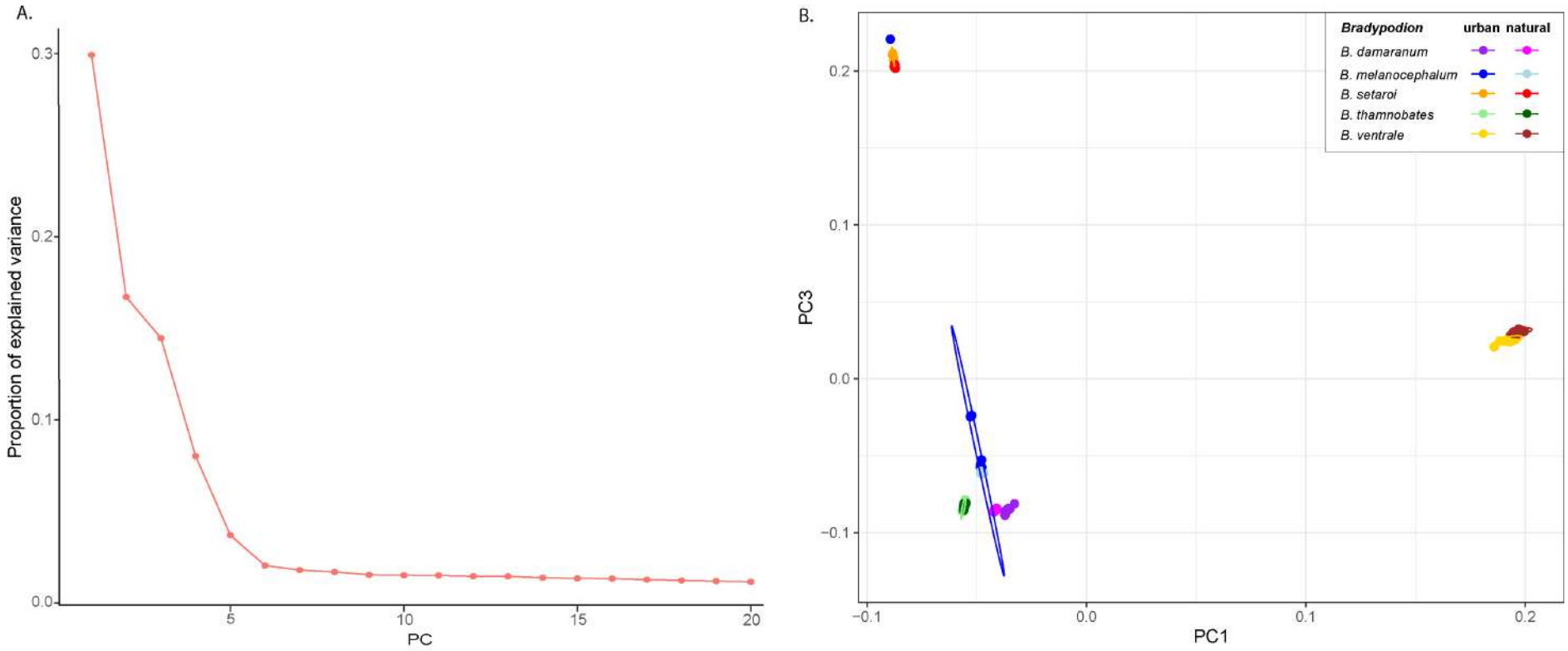


Figure S1: A. Scree plot indicating the proportion of explained variance. B. Scatterplot of principal components (PC1 and PC3) for five species of *Bradypodion*.

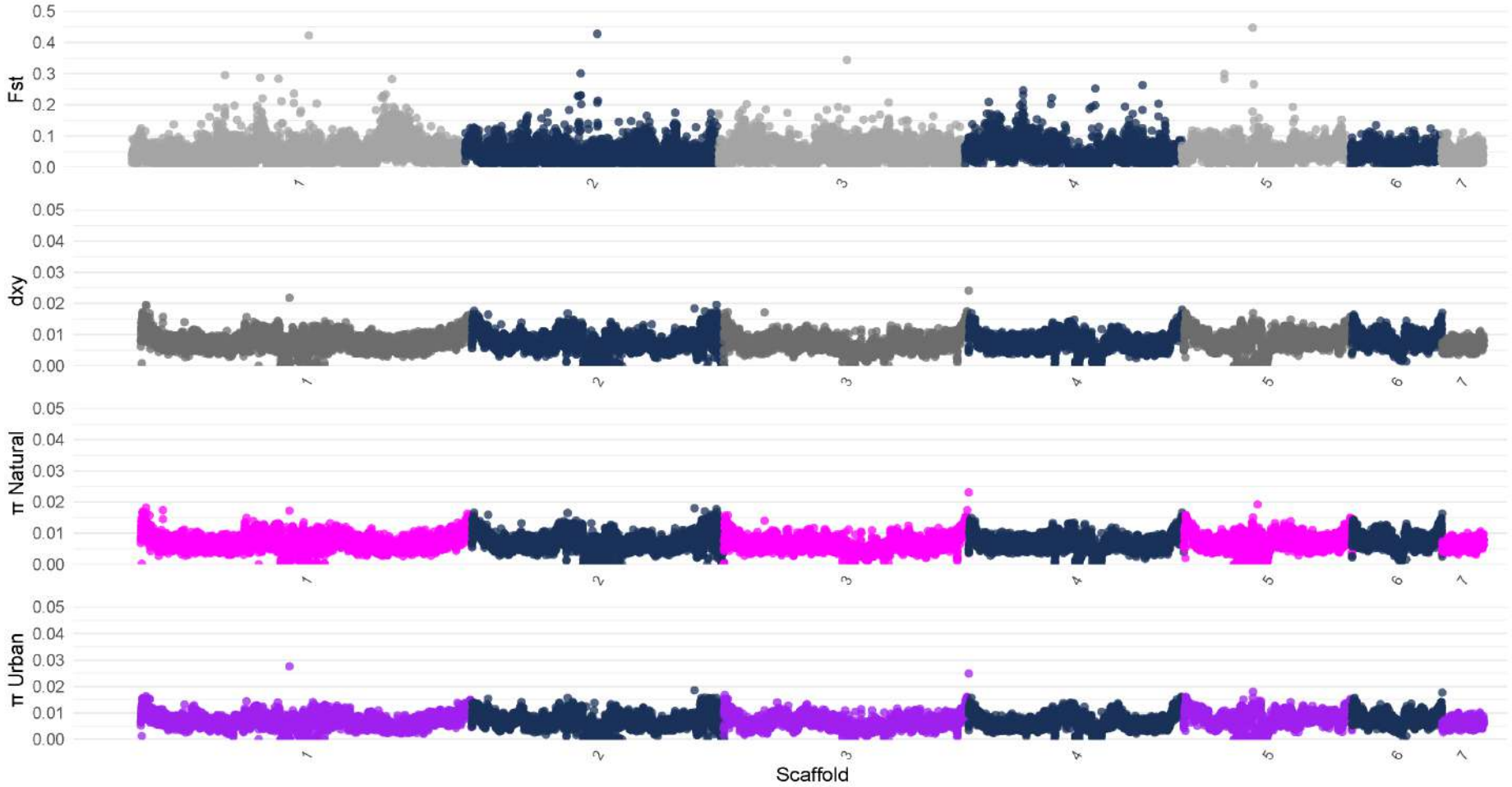


Figure S2: Stacked Manhattan plots estimated for *Bradypodion damaranum* depicting F_{ST} and genetic divergence (D_{XY}), as well as nucleotide diversity (π) was estimated for urban (pink) and natural (purple) populations in 10kb windows across the seven largest chromosomes.

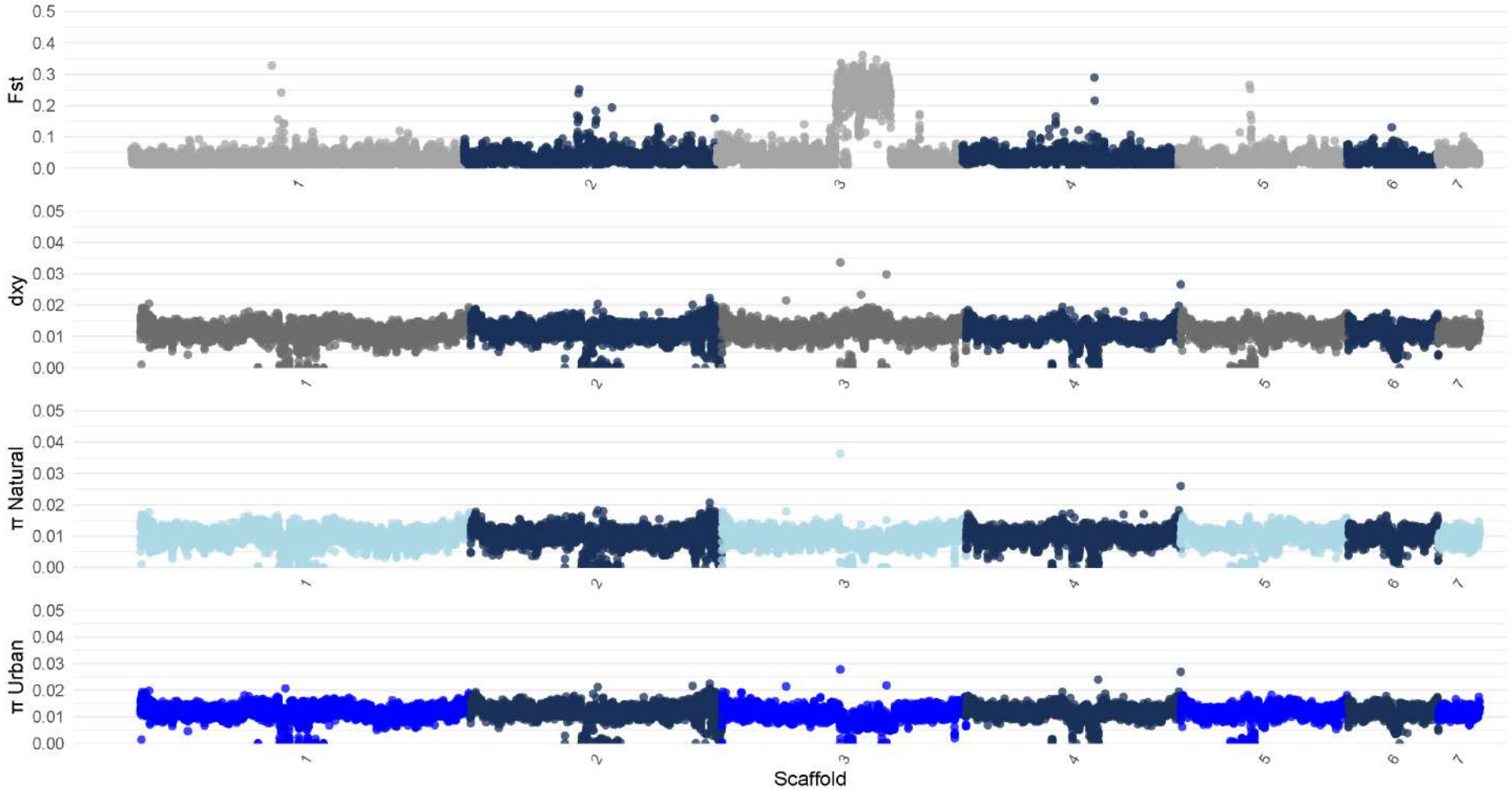


Figure S3: Stacked Manhattan plots estimated for *Bradypodion melanocephalum* depicting F_{ST} and genetic divergence (D_{XY}), as well as nucleotide diversity (π) was estimated for natural (teal) and urban (blue) populations in 10kb windows across the seven largest chromosomes.

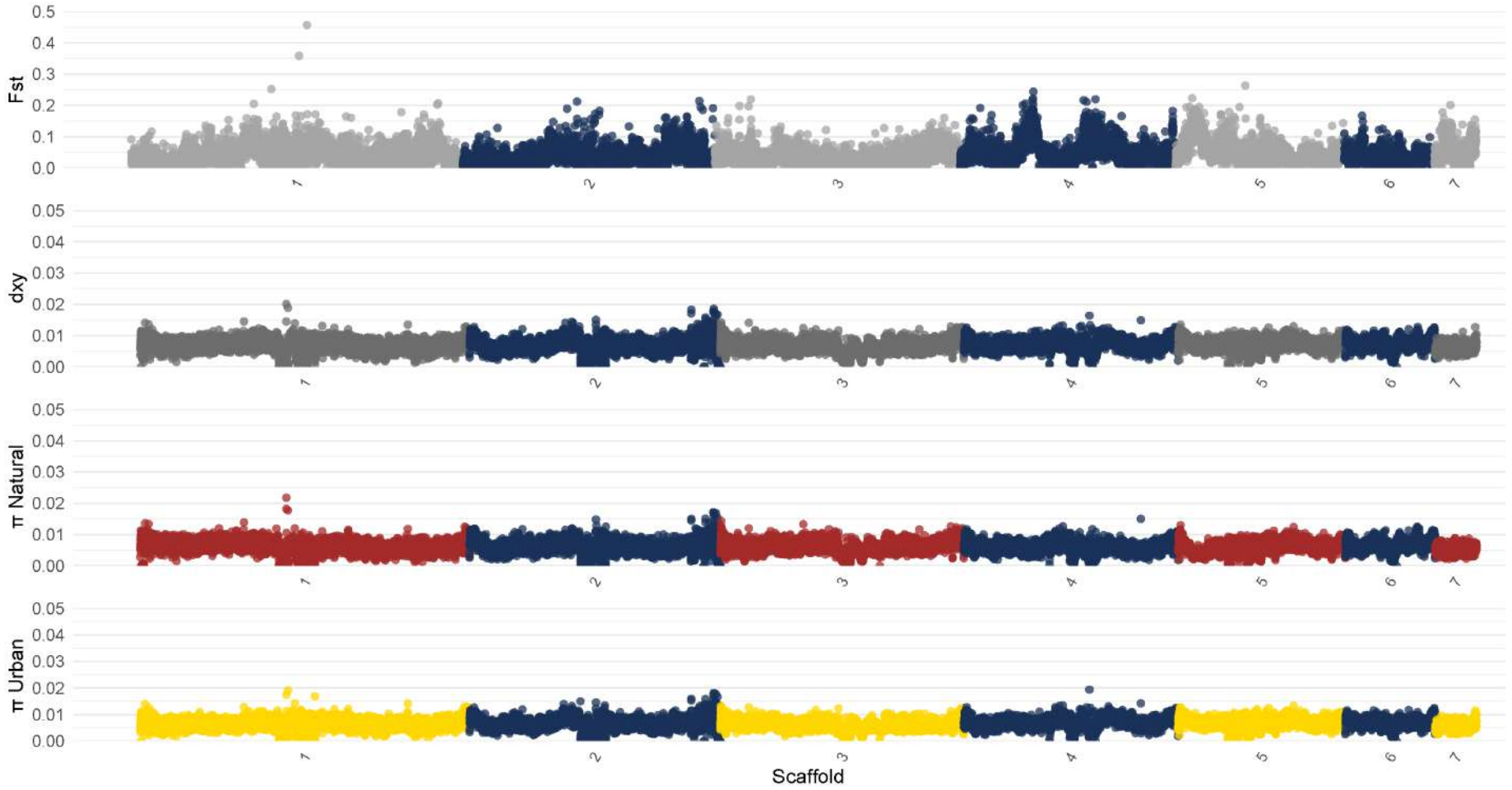


Figure S4: Stacked Manhattan plots estimated for *Bradypodion ventrale* depicting F_{ST} and genetic divergence (D_{XY}), as well as nucleotide diversity (π) was estimated for natural (brown) and urban (gold) populations in 10kb windows across the seven largest chromosomes.

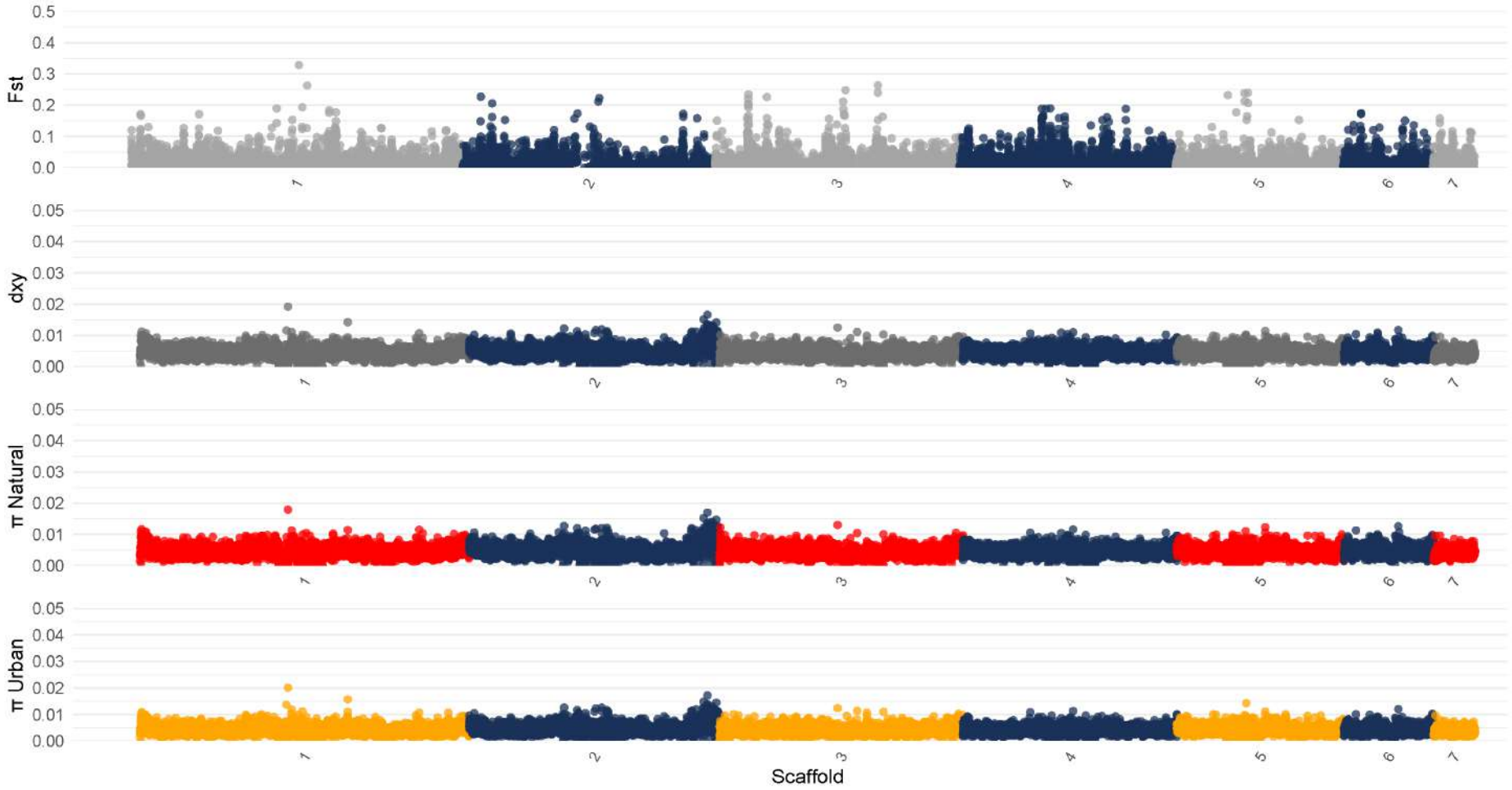


Figure S5: Stacked Manhattan plots estimated for *Bradypodion setaroi* depicting F_{ST} and genetic divergence (D_{XY}), as well as nucleotide diversity (π) was estimated for natural (red) and urban (gold) populations in 10kb windows across the seven largest chromosomes.

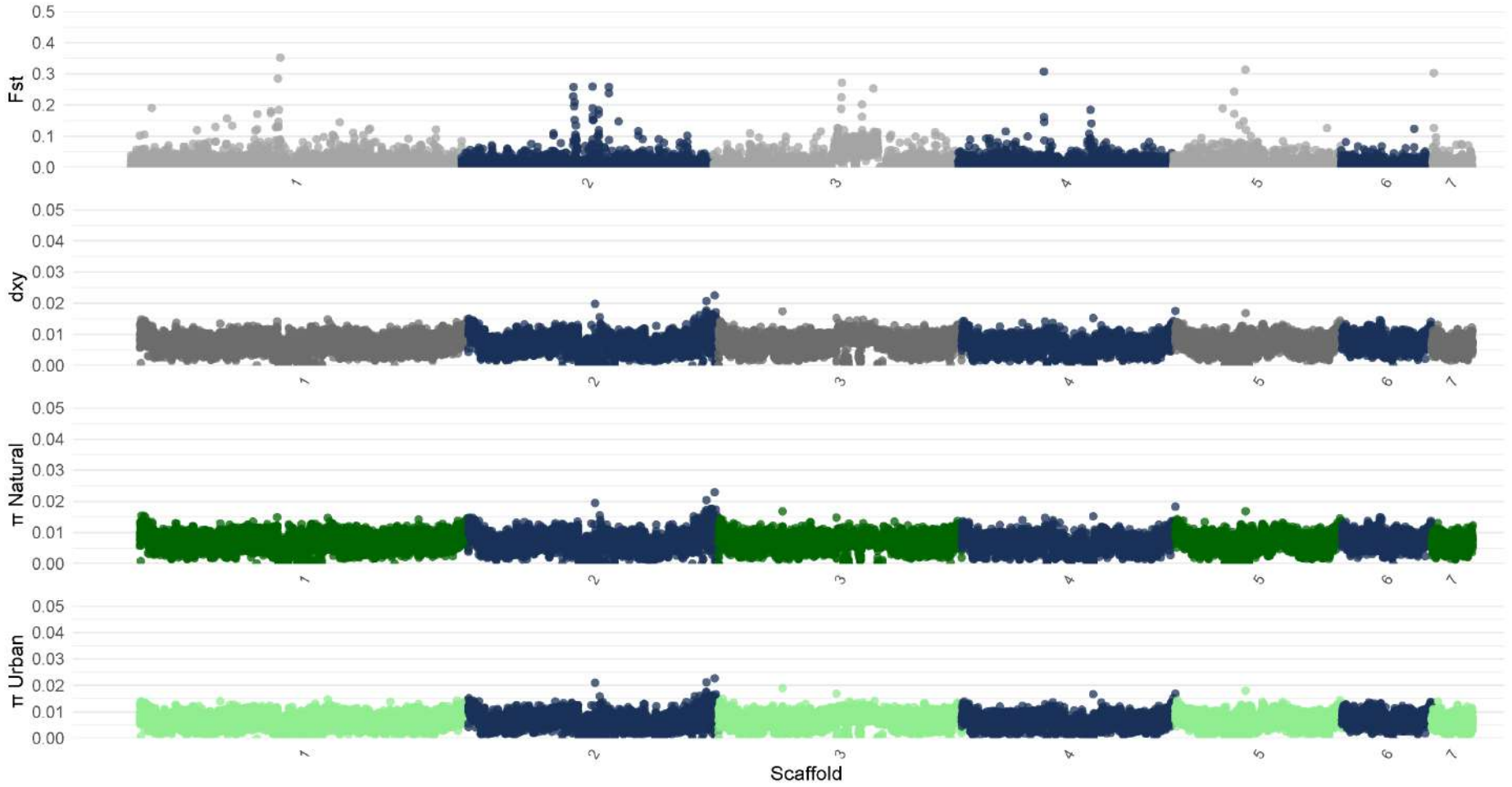


Figure S6: Stacked Manhattan plots estimated for *Bradypodion thamnobates* depicting F_{ST} and genetic divergence (D_{XY}), as well as nucleotide diversity (π) was estimated for natural (green) and urban (lime) populations in 10kb windows across the seven largest chromosomes.

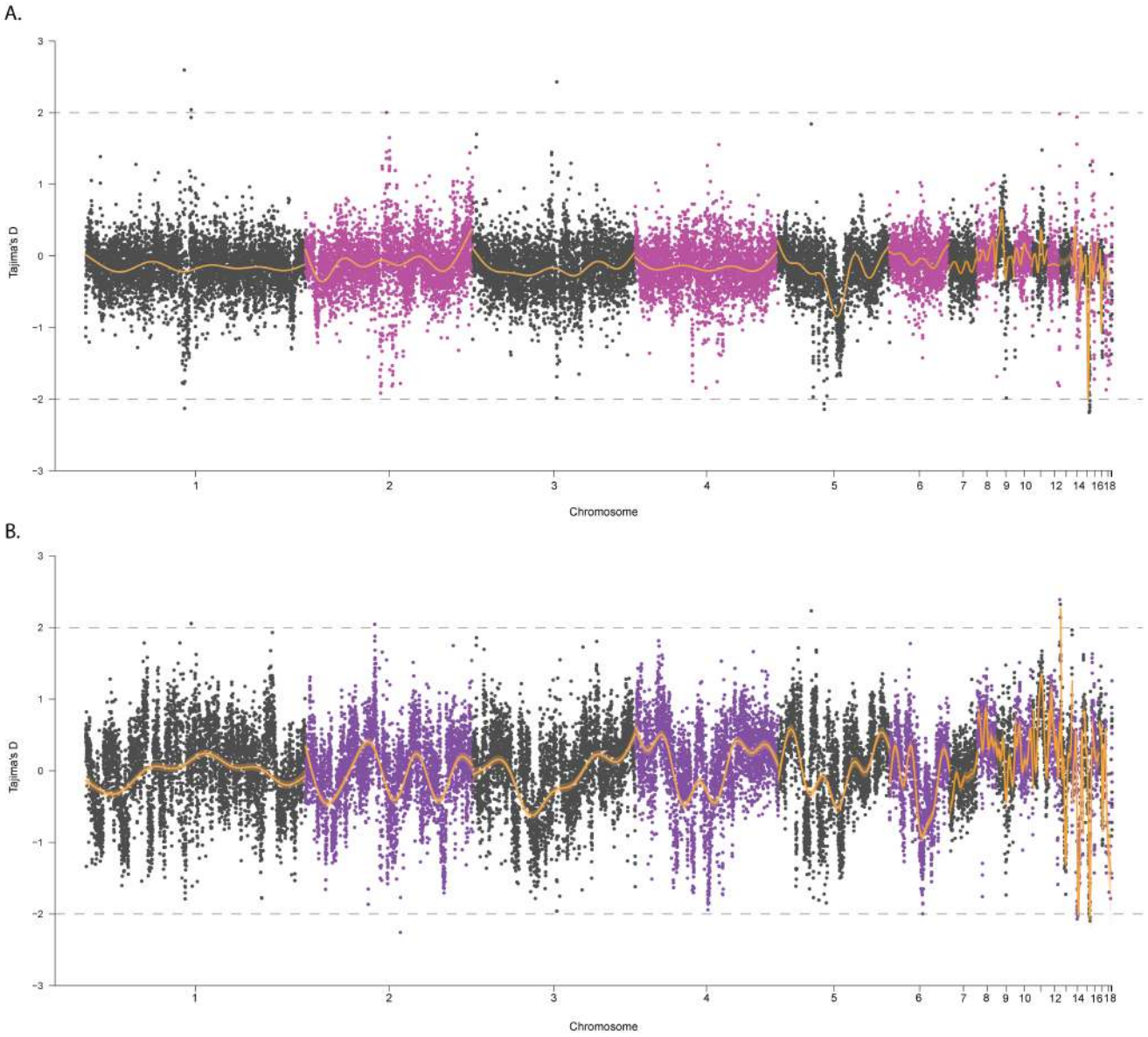


Figure S7: Stacked Manhattan plots depicting Tajima's D in 100kb windows for A) natural (pink) and B) urban (purple) for *Bradypodion thamnobates*. Locally estimated scatterplot smoothing (LOESS) is depicted by the gold line in both Manhattan plots.

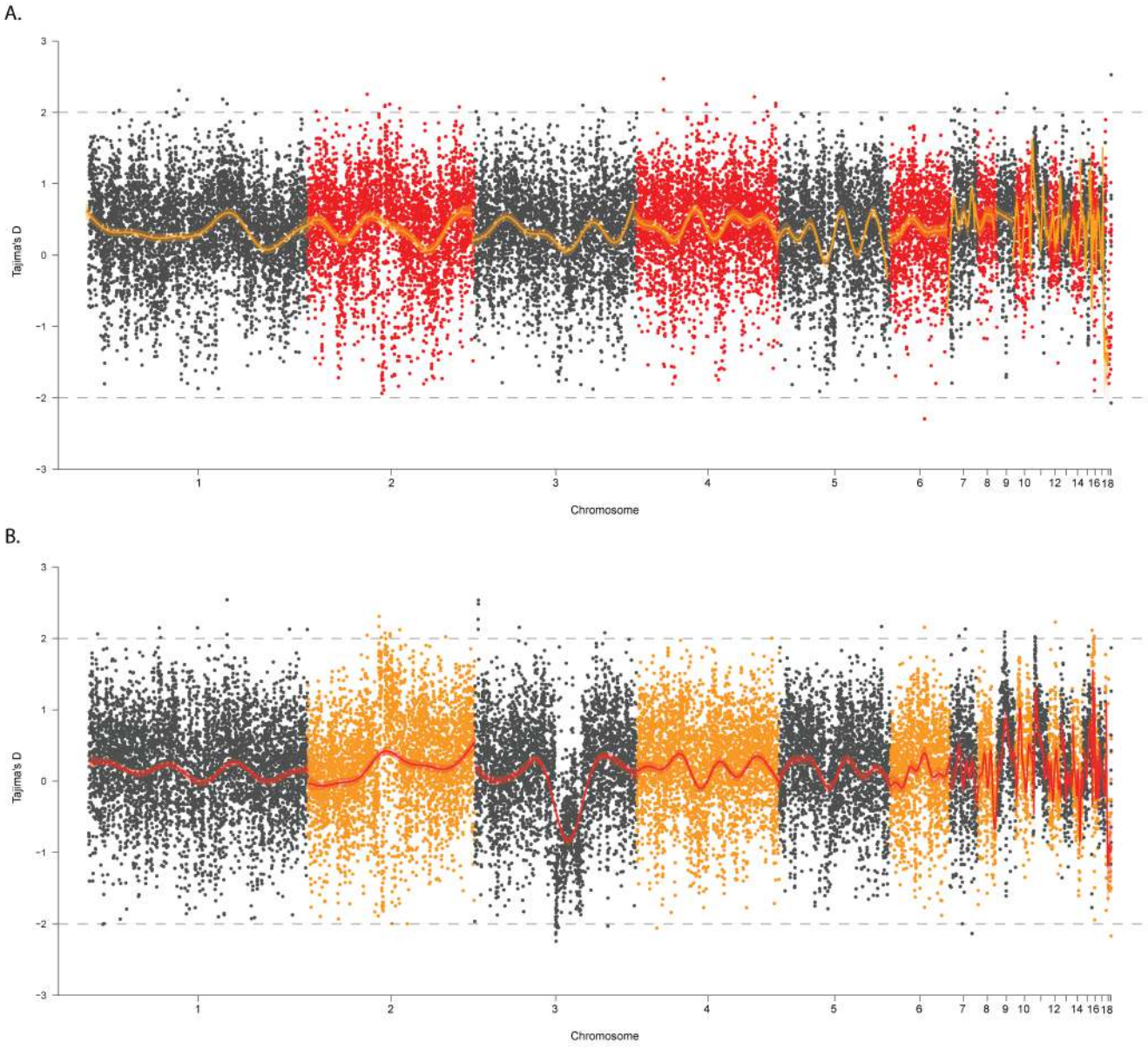


Figure S8: Stacked Manhattan plots depicting Tajima's D across 100kb windows for A) natural (red) and B) urban (gold) for *Bradypodion setaroi*. Locally estimated scatterplot smoothing (LOESS) is depicted by the gold line in the natural plot, and the red line in the urban plot.

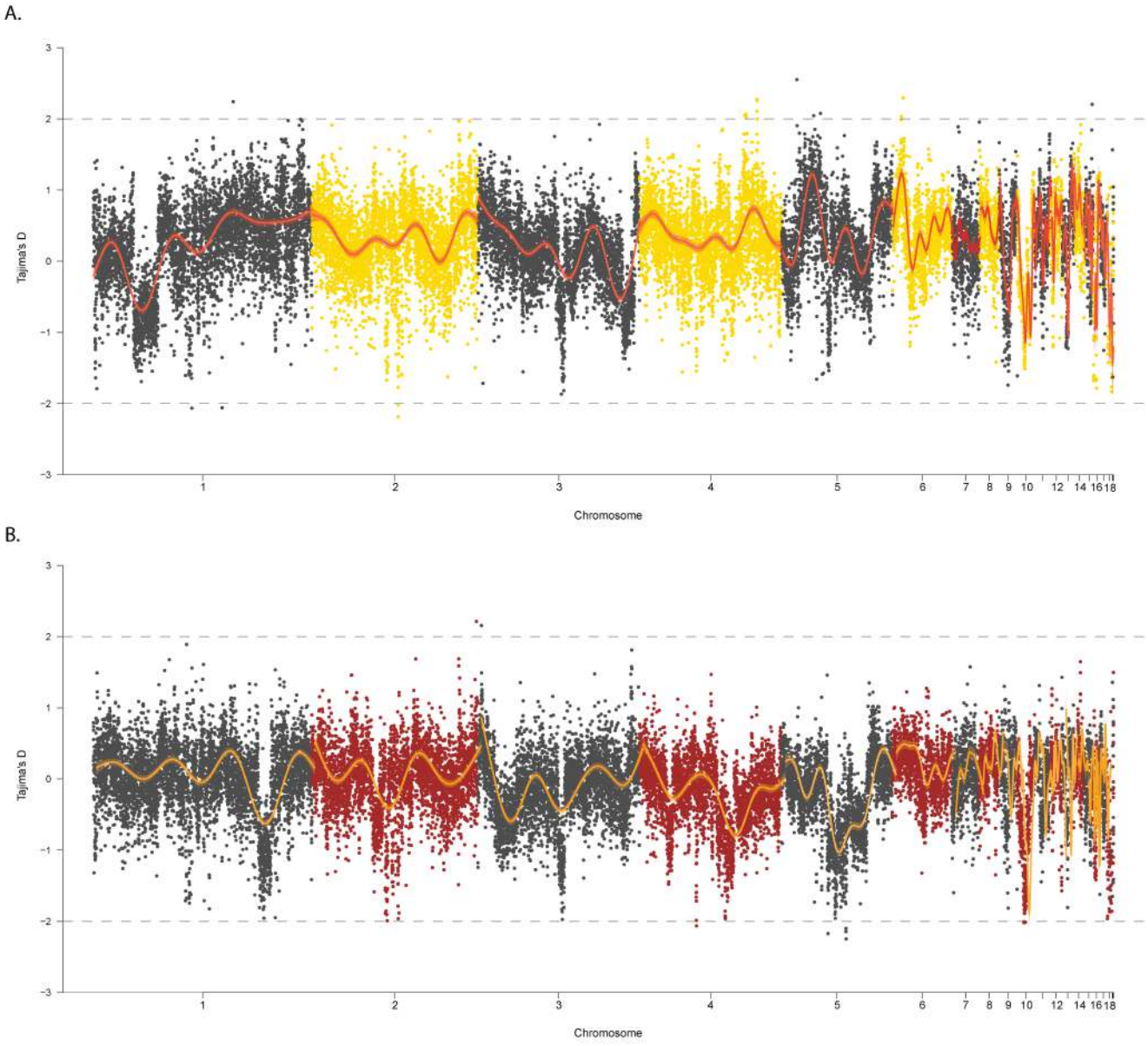


Figure S9: Stacked Manhattan plots depicting Tajima's D across 100kb windows for A) natural (gold) and B) urban (brown) for *Bradypodion ventrale*. Locally estimated scatterplot smoothing (LOESS) is depicted by the brown line in the natural plot, and the gold line in the urban plot.

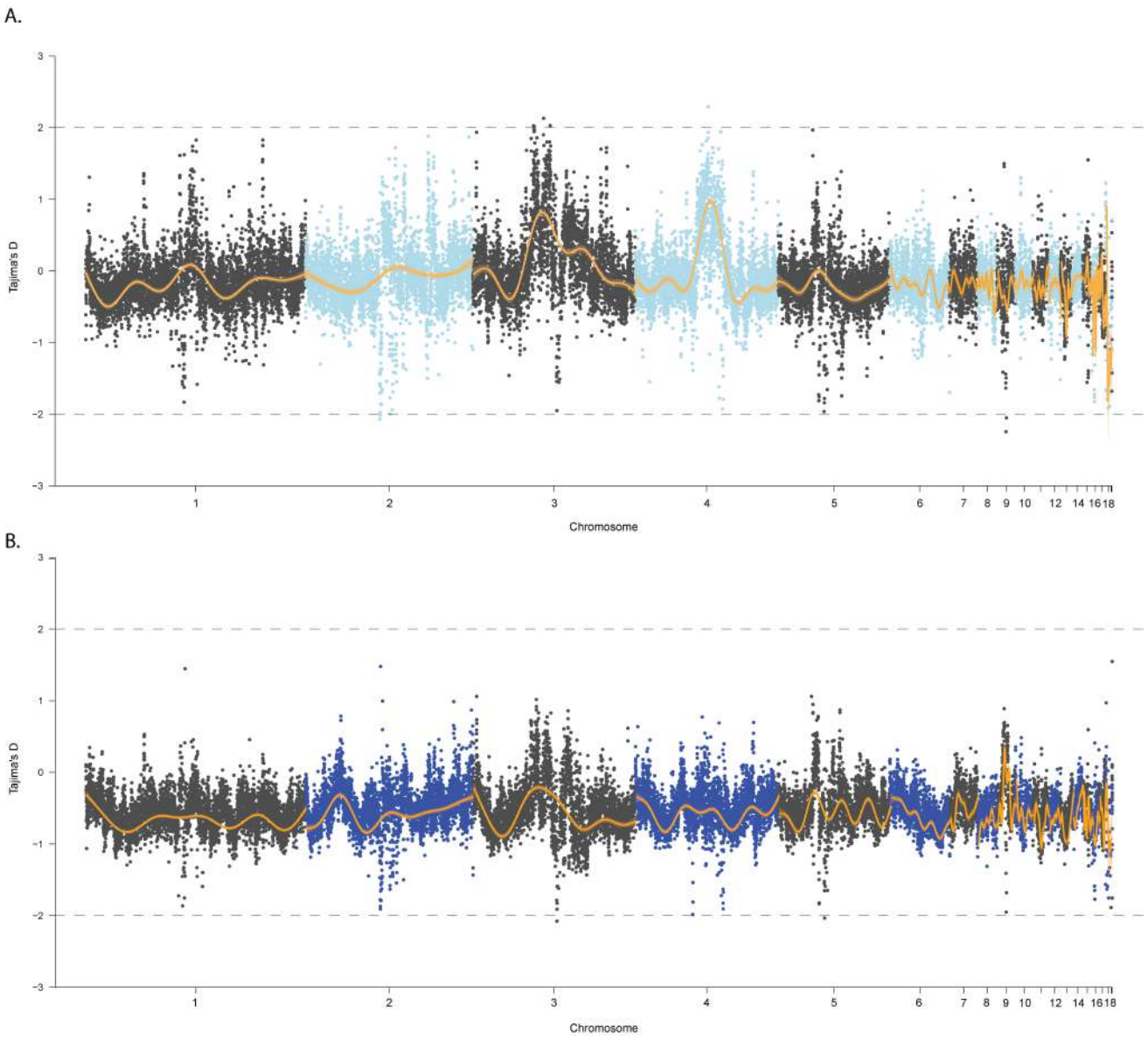


Figure S10: Stacked Manhattan plots depicting Tajima's D in 100kb windows for A) natural (teal) and B) urban (blue) for *Bradypodion melanocephalum*. Locally estimated scatterplot smoothing (LOESS) is depicted by the gold line in both Manhattan plots.

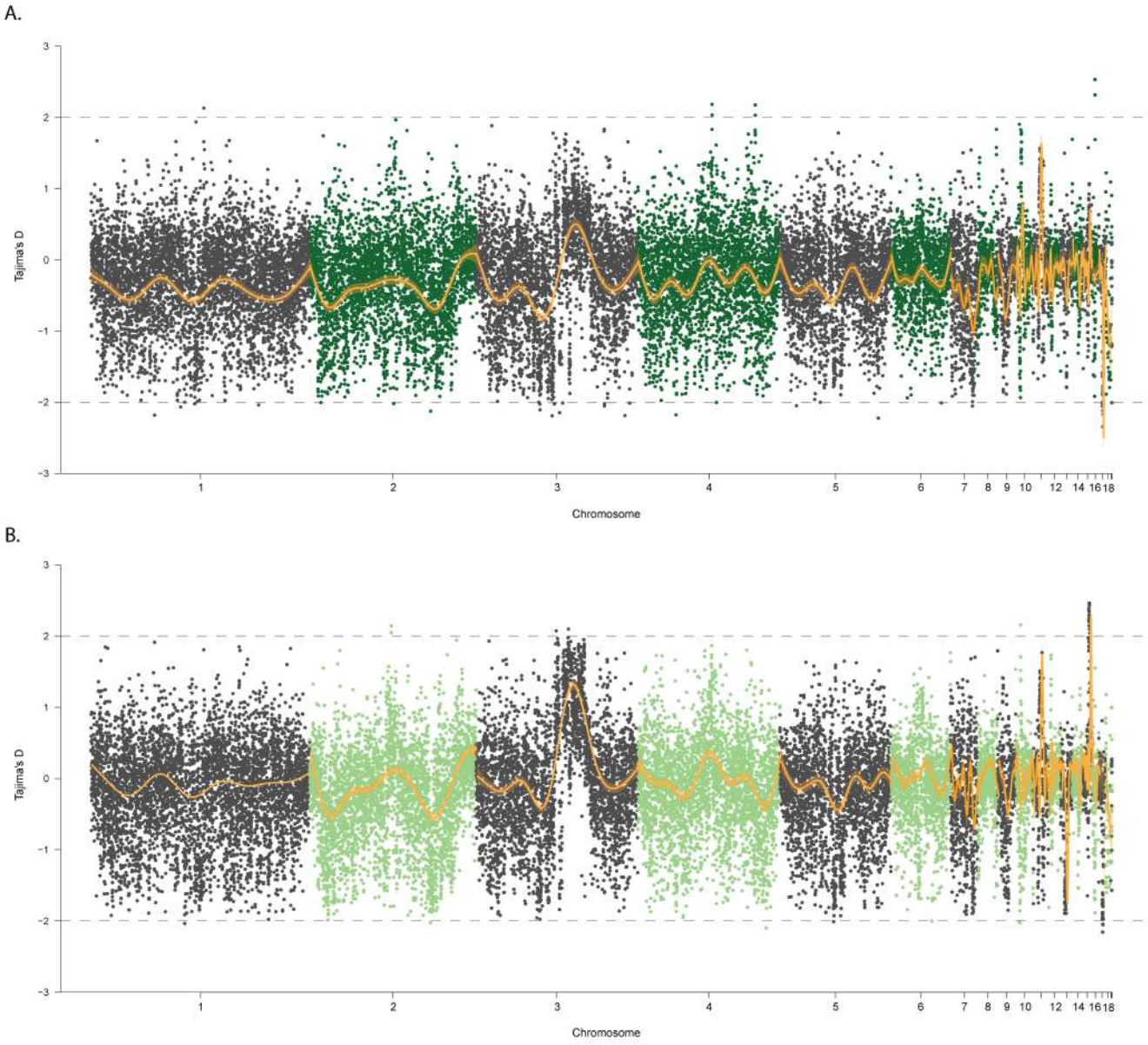


Figure S11: Stacked Manhattan plots depicting Tajima's D in 100kb windows for A) natural (green) and B) urban (lime) for *Bradypodion thamnobates*. Locally estimated scatterplot smoothing (LOESS) is depicted by the gold line in both Manhattan plots.

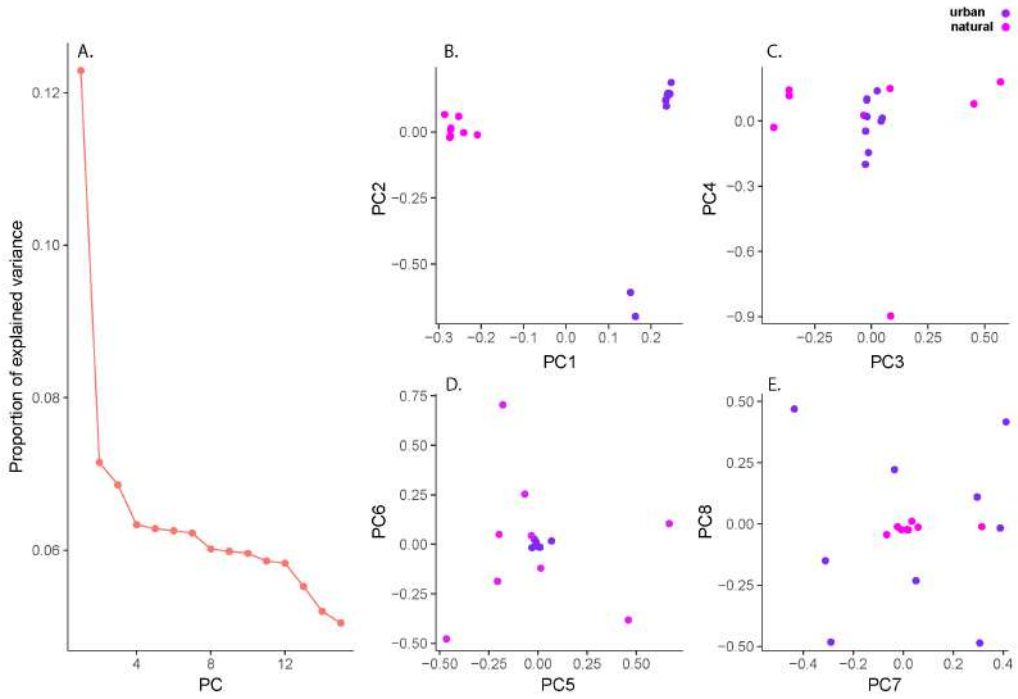


Figure S12: A. Scree plot indicating the proportion of estimated variance. Scatterplots of principal components B) PC1 and PC2, C) PC3 and PC4, and D) PC5 and PC6 for *Bradypodion damaranum*. Samples are coloured based on habitat type, urban (purple) and natural (pink).

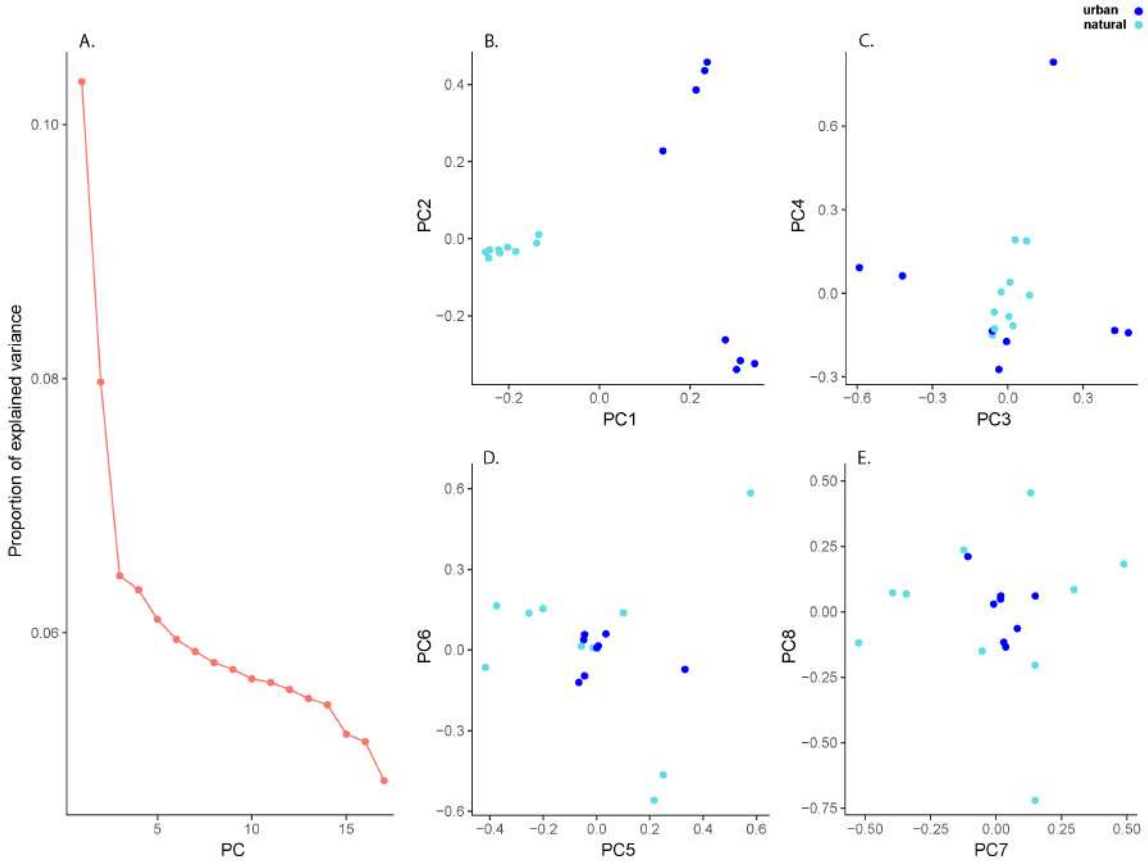


Figure S13: A. Scree plot indicating the proportion of estimated variance. Scatterplots of principal components B) PC1 and PC2, C) PC3 and PC4, and D) PC5 and PC6 for *Bradypodion melanocephalum*. Samples are coloured based on habitat type, urban (blue) and natural (teal).

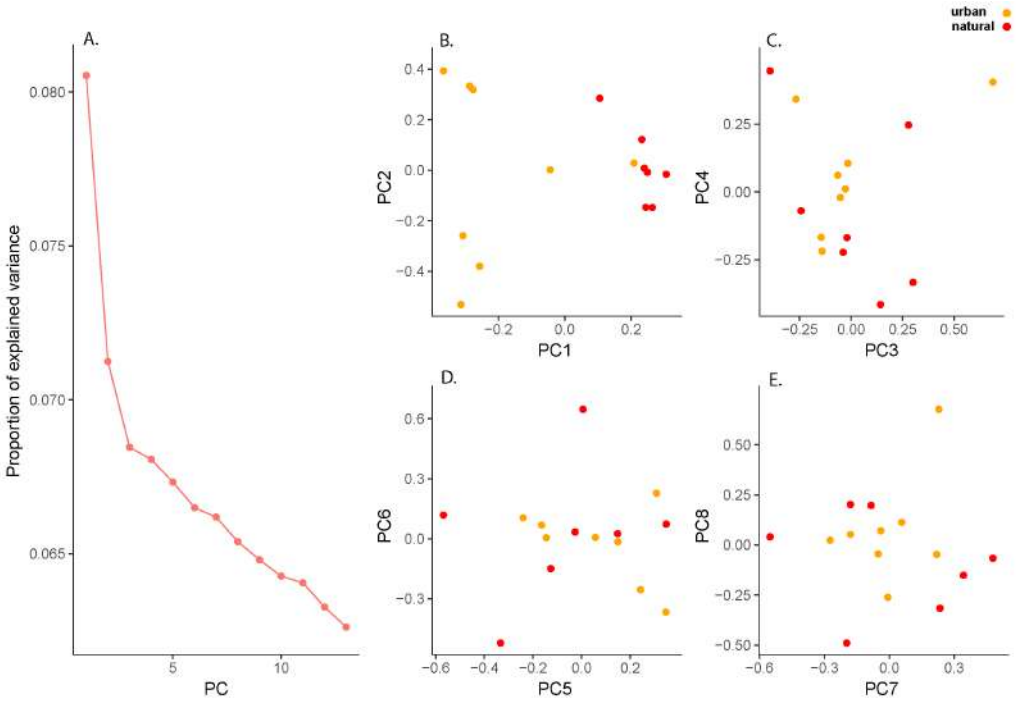


Figure S14: A. Scree plot indicating the proportion of estimated variance. Scatterplots of principal components B) PC1 and PC2, C) PC3 and PC4, and D) PC5 and PC6 for *Bradypodion setaroi*. Samples are coloured based on habitat type, urban (gold) and natural (red).

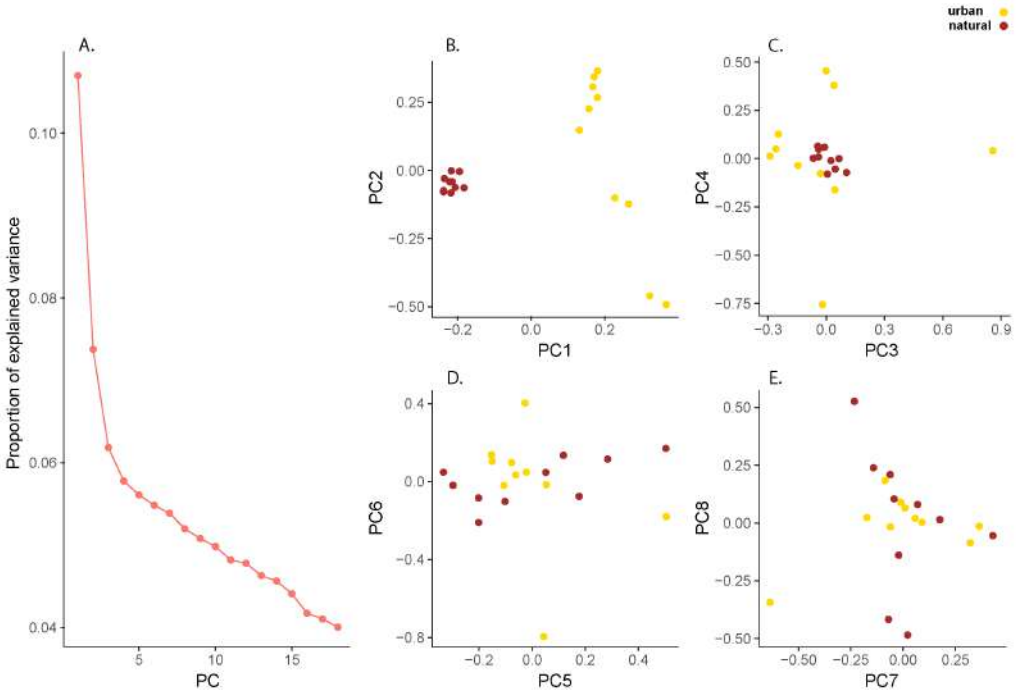


Figure S15: A. Scree plot indicating the proportion of estimated variance. Scatterplots of principal components B) PC1 and PC2, C) PC3 and PC4, and D) PC5 and PC6 for *Bradypodion ventrale*. Samples are coloured based on habitat type, urban (gold) and natural (brown).

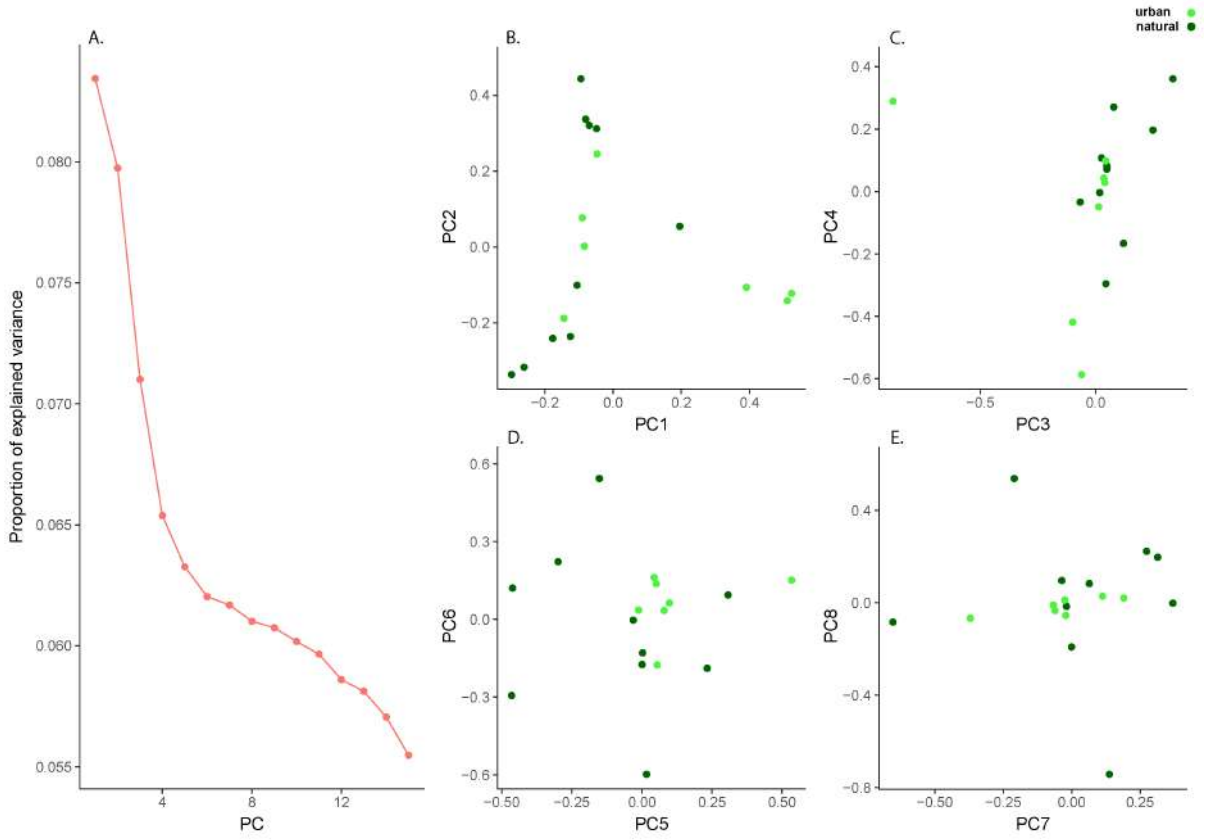


Figure S16: A. Scree plot indicating the proportion of estimated variance. Scatterplots of principal components B) PC1 and PC2, C) PC3 and PC4, and D) PC5 and PC6 for *Bradypodion thamnobates*. Samples are coloured based on habitat type, urban (lime) and natural (green).

Supplementary Tables

Table S1: Full list of Gene Ontology (GO) terms for genes containing outlier loci detected in *Bradypodion damaranum* populations. Provided are the GO Source, GO term name and ID, Adjusted p-value, Term size (total genes associated with the GO term), and Intersection size (number of genes from this study matching genes in the term).

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	nervous system development	GO:0007399	0.005	1250	9
GO:BP	positive regulation of immune system process	GO:0002684	0.005	450	6
GO:BP	regulation of immune system process	GO:0002682	0.005	675	7
GO:BP	biological regulation	GO:0065007	0.005	7729	23
GO:BP	multicellular organism development	GO:0007275	0.005	2307	12
GO:BP	microtubule polymerization	GO:0046785	0.005	45	3
GO:BP	regulation of protein modification process	GO:0031399	0.005	430	6
GO:BP	positive regulation of leukocyte migration	GO:0002687	0.006	54	3
GO:BP	protein modification process	GO:0036211	0.006	1705	10
GO:BP	regulation of cellular process	GO:0050794	0.006	7075	21
GO:BP	protein metabolic process	GO:0019538	0.006	3272	14
GO:BP	mitotic cell cycle	GO:0000278	0.006	520	6
GO:BP	embryonic morphogenesis	GO:0048598	0.007	330	5
GO:BP	tube development	GO:0035295	0.007	539	6
GO:BP	anatomical structure development	GO:0048856	0.007	2974	13
GO:BP	synaptic vesicle clustering	GO:0097091	0.007	11	2
GO:BP	regulation of RIG-I signaling pathway	GO:0039535	0.007	11	2
GO:BP	regulation of protein metabolic process	GO:0051246	0.008	861	7
GO:BP	developmental process	GO:0032502	0.008	3196	13
GO:BP	regulation of protein deacetylation	GO:0090311	0.008	16	2
GO:BP	protein deacetylation	GO:0006476	0.008	16	2
GO:BP	cellular component organization	GO:0016043	0.008	4146	15
GO:BP	ensheathment of neurons	GO:0007272	0.008	78	3
GO:BP	macromolecule modification	GO:0043412	0.008	1882	10
GO:BP	protein deacylation	GO:0035601	0.008	16	2
GO:BP	microtubule polymerization or depolymerization	GO:0031109	0.008	74	3
GO:BP	morphogenesis of embryonic epithelium	GO:0016331	0.008	83	3
GO:BP	axon ensheathment	GO:0008366	0.008	78	3
GO:BP	myelination	GO:0042552	0.008	76	3
GO:BP	multicellular organismal process	GO:0032501	0.008	3623	14
GO:BP	RIG-I signaling pathway	GO:0039529	0.008	14	2
GO:BP	negative regulation of signal transduction by p53 class mediator	GO:1901797	0.008	13	2
GO:BP	regulation of myelination	GO:0031641	0.008	16	2
GO:BP	positive regulation of T cell migration	GO:2000406	0.008	14	2
GO:BP	regulation of leukocyte migration	GO:0002685	0.008	85	3
GO:BP	regulation of T cell migration	GO:2000404	0.008	16	2
GO:BP	organonitrogen compound metabolic process	GO:1901564	0.008	4018	15
GO:BP	positive regulation of lymphocyte migration	GO:2000403	0.008	16	2
GO:BP	cell cycle	GO:0007049	0.008	931	7
GO:BP	cellular component organization or biogenesis	GO:0071840	0.009	4295	15
GO:BP	tube morphogenesis	GO:0035239	0.010	432	5
GO:BP	T cell migration	GO:0072678	0.011	20	2

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	microtubule nucleation	GO:0007020	0.011	20	2
GO:BP	positive regulation of cytokinesis	GO:0032467	0.011	20	2
GO:BP	positive regulation of biological process	GO:0048518	0.011	3406	13
GO:BP	positive regulation of multicellular organismal process	GO:0051240	0.011	707	6
GO:BP	angiogenesis	GO:0001525	0.011	249	4
GO:BP	cell cycle process	GO:0022402	0.011	719	6
GO:BP	non-membrane-bounded organelle assembly	GO:0140694	0.011	255	4
GO:BP	regulation of lymphocyte migration	GO:2000401	0.013	23	2
GO:BP	positive regulation of response to external stimulus	GO:0032103	0.013	270	4
GO:BP	anatomical structure morphogenesis	GO:0009653	0.013	1407	8
GO:BP	vesicle localization	GO:0051648	0.013	114	3
GO:BP	macromolecule deacylation	GO:0098732	0.014	25	2
GO:BP	embryo development	GO:0009790	0.014	499	5
GO:BP	lymphocyte proliferation	GO:0046651	0.014	118	3
GO:BP	regulation of macromolecule metabolic process	GO:0060255	0.014	3592	13
GO:BP	mononuclear cell proliferation	GO:0032943	0.015	121	3
GO:BP	positive regulation of macromolecule metabolic process	GO:0010604	0.015	1826	9
GO:BP	positive regulation of cell division	GO:0051781	0.015	27	2
GO:BP	positive regulation of mononuclear cell migration	GO:0071677	0.015	27	2
GO:BP	regulation of angiogenesis	GO:0045765	0.015	125	3
GO:BP	synaptic vesicle localization	GO:0097479	0.015	28	2
GO:BP	regulation of vasculature development	GO:1901342	0.015	127	3
GO:BP	metabolic process	GO:0008152	0.016	9632	23
GO:BP	immune system process	GO:0002376	0.016	1130	7
GO:BP	negative regulation of protein modification process	GO:0031400	0.016	130	3
GO:BP	blood vessel morphogenesis	GO:0048514	0.016	306	4
GO:BP	leukocyte proliferation	GO:0070661	0.017	134	3
GO:BP	positive regulation of stem cell proliferation	GO:2000648	0.017	31	2
GO:BP	regulation of cell cycle	GO:0051726	0.017	551	5
GO:BP	regulation of signal transduction by p53 class mediator	GO:1901796	0.018	32	2
GO:BP	positive regulation of immune response	GO:0050778	0.018	322	4
GO:BP	positive regulation of metabolic process	GO:0009893	0.018	1978	9
GO:BP	leukocyte migration	GO:0050900	0.018	146	3
GO:BP	regulation of metabolic process	GO:0019222	0.018	3877	13
GO:BP	positive regulation of nitrogen compound metabolic process	GO:0051173	0.018	1635	8
GO:BP	response to chemokine	GO:1990868	0.018	39	2
GO:BP	cranial skeletal system development	GO:1904888	0.018	39	2
GO:BP	organelle assembly	GO:0070925	0.018	586	5
GO:BP	anatomical structure formation involved in morphogenesis	GO:0048646	0.018	564	5
GO:BP	positive regulation of protein deacetylation	GO:0090312	0.018	1	1
GO:BP	central nervous system vasculogenesis	GO:0022009	0.018	1	1
GO:BP	bone remodeling	GO:0046849	0.018	39	2
GO:BP	negative regulation of mitotic cell cycle, embryonic	GO:0045976	0.018	1	1
GO:BP	blood vessel development	GO:0001568	0.018	351	4

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	cellular component assembly	GO:0022607	0.018	1982	9
GO:BP	gliogenesis	GO:0042063	0.018	151	3
GO:BP	cellular response to chemokine	GO:1990869	0.018	39	2
GO:BP	positive regulation of leukocyte chemotaxis	GO:0002690	0.018	40	2
GO:BP	centrosome-templated microtubule nucleation	GO:0090222	0.018	1	1
GO:BP	glial cell fate determination	GO:0007403	0.018	1	1
GO:BP	positive regulation of branching involved in lung morphogenesis	GO:0061047	0.018	1	1
GO:BP	endomitotic cell cycle	GO:0007113	0.018	1	1
GO:BP	positive regulation of termination of DNA-templated transcription	GO:0060566	0.018	1	1
GO:BP	renal inner medulla development	GO:0072053	0.018	1	1
GO:BP	renal outer medulla development	GO:0072054	0.018	1	1
GO:BP	memory T cell proliferation	GO:0061485	0.018	1	1
GO:BP	organic substance metabolic process	GO:0071704	0.018	9245	22
GO:BP	regulation of gliogenesis	GO:0014013	0.018	34	2
GO:BP	regulation of establishment of T cell polarity	GO:1903903	0.018	1	1
GO:BP	positive regulation of establishment of T cell polarity	GO:1903905	0.018	1	1
GO:BP	lymphocyte migration	GO:0072676	0.018	34	2
GO:BP	positive regulation of termination of RNA polymerase II transcription	GO:1904595	0.018	1	1
GO:BP	lipoate metabolic process	GO:0009106	0.018	1	1
GO:BP	lipoate biosynthetic process	GO:0009107	0.018	1	1
GO:BP	positive regulation of MDA-5 signaling pathway	GO:1900245	0.018	1	1
GO:BP	negative regulation of post-translational protein modification	GO:1901874	0.020	42	2
GO:BP	regulation of mononuclear cell migration	GO:0071675	0.020	42	2
GO:BP	protein polymerization	GO:0051258	0.021	169	3
GO:BP	microtubule-based process	GO:0007017	0.021	635	5
GO:BP	vasculature development	GO:0001944	0.021	372	4
GO:BP	cell junction organization	GO:0034330	0.021	374	4
GO:BP	regulation of cell cycle process	GO:0010564	0.021	376	4
GO:BP	signal transduction	GO:0007165	0.021	3511	12
GO:BP	regulation of immune response	GO:0050776	0.023	387	4
GO:BP	oligodendrocyte differentiation	GO:0048709	0.023	47	2
GO:BP	cellular component biogenesis	GO:0044085	0.024	2158	9
GO:BP	positive regulation of cellular process	GO:0048522	0.024	3088	11
GO:BP	astrocyte-dopaminergic neuron signaling	GO:0036520	0.025	2	1
GO:BP	negative regulation of cell-cell adhesion mediated by integrin	GO:0033633	0.025	2	1
GO:BP	positive regulation of cellular metabolic process	GO:0031325	0.025	1800	8
GO:BP	central nervous system myelin maintenance	GO:0032286	0.025	2	1
GO:BP	regulation of cytokinesis	GO:0032465	0.025	54	2
GO:BP	microtubule cytoskeleton organization	GO:0000226	0.025	431	4
GO:BP	regulation of cytoplasmic pattern recognition receptor signaling pathway	GO:0039531	0.025	50	2
GO:BP	positive regulation of determination of dorsal identity	GO:2000017	0.025	2	1
GO:BP	establishment of blood-retinal barrier	GO:1990963	0.025	2	1

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	positive regulation of mitotic cytokinesis	GO:1903490	0.025	2	1
GO:BP	dense core granule transport	GO:1901950	0.025	2	1
GO:BP	positive regulation of lymphangiogenesis	GO:1901492	0.025	2	1
GO:BP	glial cell-neuron signaling	GO:0150098	0.025	2	1
GO:BP	presynapse to nucleus signaling pathway	GO:0099526	0.025	2	1
GO:BP	dense core granule cytoskeletal transport	GO:0099519	0.025	2	1
GO:BP	presynaptic signal transduction	GO:0098928	0.025	2	1
GO:BP	lymphocyte migration into lymph node	GO:0097022	0.025	2	1
GO:BP	chemokine (C-C motif) ligand 21 signaling pathway	GO:0038116	0.025	2	1
GO:BP	DNA clamp unloading	GO:0090618	0.025	2	1
GO:BP	regulation of primary metabolic process	GO:0080090	0.025	3192	11
GO:BP	regulation of stem cell proliferation	GO:0072091	0.025	51	2
GO:BP	fungiform papilla formation	GO:0061198	0.025	2	1
GO:BP	epithelial cell differentiation involved in prostate gland development	GO:0060742	0.025	2	1
GO:BP	lung induction	GO:0060492	0.025	2	1
GO:BP	lung field specification	GO:0060424	0.025	2	1
GO:BP	foregut regionalization	GO:0060423	0.025	2	1
GO:BP	regulation of lymphocyte activation	GO:0051249	0.025	190	3
GO:BP	regulation of nitrogen compound metabolic process	GO:0051171	0.025	3116	11
GO:BP	positive regulation of endothelial cell differentiation	GO:0045603	0.025	2	1
GO:BP	mitotic cell cycle, embryonic	GO:0045448	0.025	2	1
GO:BP	negative regulation of DNA methylation-dependent heterochromatin formation	GO:0090310	0.025	2	1
GO:BP	regulation of centriole-centriole cohesion	GO:0030997	0.025	2	1
GO:BP	negative regulation of intrinsic apoptotic signaling pathway	GO:2001243	0.025	50	2
GO:BP	regulation of leukocyte chemotaxis	GO:0002688	0.025	51	2
GO:BP	dorsal/ventral pattern formation	GO:0009953	0.025	53	2
GO:BP	regulation of anatomical structure morphogenesis	GO:0022603	0.025	410	4
GO:BP	organelle organization	GO:0006996	0.025	2209	9
GO:BP	positive regulation of heparan sulfate proteoglycan biosynthetic process	GO:0010909	0.025	2	1
GO:BP	regulation of mitotic cell cycle, embryonic	GO:0009794	0.025	2	1
GO:BP	embryonic skeletal system morphogenesis	GO:0048704	0.027	59	2
GO:BP	cell-cell adhesion	GO:0098609	0.027	441	4
GO:BP	regulation of leukocyte activation	GO:0002694	0.027	211	3
GO:BP	mononuclear cell migration	GO:0071674	0.027	60	2
GO:BP	immune response-activating signaling pathway	GO:0002757	0.027	214	3
GO:BP	positive regulation of macromolecule biosynthetic process	GO:0010557	0.027	1455	7
GO:BP	positive regulation of response to stimulus	GO:0048584	0.029	1089	6

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	immune response-regulating signaling pathway	GO:0002764	0.029	219	3
GO:BP	cell junction assembly	GO:0034329	0.029	220	3
GO:BP	positive regulation of chemotaxis	GO:0050921	0.029	63	2
GO:BP	signaling	GO:0023052	0.029	3820	12
GO:BP	synapse organization	GO:0050808	0.029	223	3
GO:BP	macromolecule metabolic process	GO:0043170	0.029	7919	19
GO:BP	embryonic skeletal limb joint morphogenesis	GO:0036023	0.030	3	1
GO:BP	acinar cell differentiation	GO:0090425	0.030	3	1
GO:BP	regulation of MDA-5 signaling pathway	GO:0039533	0.030	3	1
GO:BP	stem cell proliferation	GO:0072089	0.030	65	2
GO:BP	positive regulation of systemic arterial blood pressure	GO:0003084	0.030	3	1
GO:BP	regulation of nervous system development	GO:0051960	0.030	229	3
GO:BP	cell communication	GO:0007154	0.030	3878	12
GO:BP	regulation of branching involved in lung morphogenesis	GO:0061046	0.030	3	1
GO:BP	mesenchymal cell proliferation involved in lung development	GO:0060916	0.030	3	1
GO:BP	primary lung bud formation	GO:0060431	0.030	3	1
GO:BP	digestive system development	GO:0055123	0.030	68	2
GO:BP	microtubule nucleation by microtubule organizing center	GO:0051418	0.030	3	1
GO:BP	glycoprotein biosynthetic process	GO:0009101	0.030	227	3
GO:BP	positive regulation of protein metabolic process	GO:0051247	0.030	480	4
GO:BP	positive regulation of biosynthetic process	GO:0009891	0.030	1502	7
GO:BP	negative regulation of biological process	GO:0048519	0.030	2842	10
GO:BP	regulation of cell activation	GO:0050865	0.030	234	3
GO:BP	tissue remodeling	GO:0048771	0.030	70	2
GO:BP	sympathetic ganglion development	GO:0061549	0.030	3	1
GO:BP	primary metabolic process	GO:0044238	0.030	8663	20
GO:BP	regulation of heparan sulfate proteoglycan biosynthetic process	GO:0010908	0.030	3	1
GO:BP	regulation of mesenchymal stem cell proliferation	GO:1902460	0.030	3	1
GO:BP	positive regulation of mesenchymal stem cell proliferation	GO:1902462	0.030	3	1
GO:BP	presynaptic active zone assembly	GO:1904071	0.030	3	1
GO:BP	regulation of termination of RNA polymerase II transcription	GO:1904594	0.030	3	1
GO:BP	establishment of lymphocyte polarity	GO:0001767	0.030	3	1
GO:BP	regulation of determination of dorsal identity	GO:2000015	0.030	3	1
GO:BP	establishment of T cell polarity	GO:0001768	0.030	3	1
GO:BP	regulation of termination of DNA-templated transcription	GO:0031554	0.030	3	1
GO:BP	positive regulation of cellular biosynthetic process	GO:0031328	0.030	1488	7
GO:BP	embryonic epithelial tube formation	GO:0001838	0.030	70	2
GO:BP	regulation of response to external stimulus	GO:0032101	0.030	478	4

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	signal transduction by p53 class mediator	GO:0072331	0.030	71	2
GO:BP	morphogenesis of an epithelium	GO:0002009	0.030	240	3
GO:BP	chordate embryonic development	GO:0043009	0.030	240	3
GO:BP	biosynthetic process	GO:0009058	0.030	5583	15
GO:BP	response to external stimulus	GO:0009605	0.030	1149	6
GO:BP	embryonic appendage morphogenesis	GO:0035113	0.030	72	2
GO:BP	embryonic limb morphogenesis	GO:0030326	0.030	72	2
GO:BP	cytoplasmic pattern recognition receptor signaling pathway	GO:0002753	0.031	75	2
GO:BP	regulation of osteoblast differentiation	GO:0045667	0.031	79	2
GO:BP	positive regulation of mast cell chemotaxis	GO:0060754	0.031	4	1
GO:BP	regulation of epithelial cell proliferation involved in prostate gland development	GO:0060768	0.031	4	1
GO:BP	smoothened signaling pathway involved in dorsal/ventral neural tube patterning	GO:0060831	0.031	4	1
GO:BP	myotome development	GO:0061055	0.031	4	1
GO:BP	negative regulation of tubulin deacetylation	GO:1904428	0.031	4	1
GO:BP	trachea formation	GO:0060440	0.031	4	1
GO:BP	positive regulation of angiogenesis	GO:0045766	0.031	75	2
GO:BP	dorsal root ganglion development	GO:1990791	0.031	4	1
GO:BP	negative regulation of chromatin organization	GO:1905268	0.031	4	1
GO:BP	regulation of timing of anagen	GO:0051884	0.031	4	1
GO:BP	glycoprotein metabolic process	GO:0009100	0.031	259	3
GO:BP	spindle assembly	GO:0051225	0.031	77	2
GO:BP	positive regulation of sulfur metabolic process	GO:0051176	0.031	4	1
GO:BP	embryo development ending in birth or egg hatching	GO:0009792	0.031	246	3
GO:BP	negative regulation of heterochromatin formation	GO:0031452	0.031	4	1
GO:BP	positive regulation of cell migration	GO:0030335	0.031	260	3
GO:BP	response to stimulus	GO:0050896	0.031	5063	14
GO:BP	regulation of developmental process	GO:0050793	0.031	1199	6
GO:BP	maintenance of presynaptic active zone structure	GO:0048790	0.031	4	1
GO:BP	negative regulation of oligodendrocyte differentiation	GO:0048715	0.031	4	1
GO:BP	oviduct development	GO:0060066	0.031	4	1
GO:BP	fungiform papilla morphogenesis	GO:0061197	0.031	4	1
GO:BP	embryonic skeletal system development	GO:0048706	0.031	81	2
GO:BP	positive regulation of vasculature development	GO:1904018	0.031	76	2
GO:BP	epithelial tube formation	GO:0072175	0.031	76	2
GO:BP	activation of immune response	GO:0002253	0.031	260	3
GO:BP	nitrogen compound metabolic process	GO:0006807	0.031	8196	19
GO:BP	T cell proliferation	GO:0042098	0.031	79	2
GO:BP	mesenchymal stem cell proliferation	GO:0097168	0.031	4	1
GO:BP	anagen	GO:0042640	0.031	4	1
GO:BP	supramolecular fiber organization	GO:0097435	0.031	508	4
GO:BP	regulation of lymphangiogenesis	GO:1901490	0.031	4	1
GO:BP	cellular response to indole-3-methanol	GO:0071681	0.031	4	1

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	positive regulation of odontoblast differentiation	GO:1901331	0.031	4	1
GO:BP	response to indole-3-methanol	GO:0071680	0.031	4	1
GO:BP	proteasome-mediated ubiquitin-dependent protein catabolic process	GO:0043161	0.031	255	3
GO:BP	positive regulation of proteoglycan biosynthetic process	GO:1902730	0.031	4	1
GO:BP	regulation of odontoblast differentiation	GO:1901329	0.031	4	1
GO:BP	positive regulation of RIG-I signaling pathway	GO:1900246	0.031	4	1
GO:BP	regulation of centromeric sister chromatid cohesion	GO:0070602	0.031	4	1
GO:BP	negative regulation of heterochromatin organization	GO:0120262	0.031	4	1
GO:BP	regulation of cell-cell adhesion involved in gastrulation	GO:0070587	0.031	4	1
GO:BP	regulation of pattern recognition receptor signaling pathway	GO:0062207	0.031	73	2
GO:BP	tube formation	GO:0035148	0.031	81	2
GO:BP	genitalia morphogenesis	GO:0035112	0.031	4	1
GO:BP	cranial ganglion development	GO:0061550	0.031	4	1
GO:BP	negative regulation of heterotypic cell-cell adhesion	GO:0034115	0.031	4	1
GO:BP	MDA-5 signaling pathway	GO:0039530	0.031	4	1
GO:BP	positive regulation of gene expression	GO:0010628	0.031	526	4
GO:BP	positive regulation of cell motility	GO:2000147	0.032	266	3
GO:BP	centrosome cycle	GO:0007098	0.032	83	2
GO:BP	cellular metabolic process	GO:0044237	0.033	8281	19
GO:BP	regulation of lymphocyte proliferation	GO:0050670	0.033	85	2
GO:BP	regulation of cell division	GO:0051302	0.033	85	2
GO:BP	regulation of intrinsic apoptotic signaling pathway	GO:2001242	0.033	85	2
GO:BP	negative regulation of pancreatic juice secretion	GO:0090188	0.033	5	1
GO:BP	fungiform papilla development	GO:0061196	0.033	5	1
GO:BP	cell morphogenesis	GO:0000902	0.033	543	4
GO:BP	positive regulation of immune effector process	GO:0002699	0.033	87	2
GO:BP	cell-cell adhesion involved in gastrulation	GO:0070586	0.033	5	1
GO:BP	neural plate development	GO:0001840	0.033	5	1
GO:BP	midbrain dopaminergic neuron differentiation	GO:1904948	0.033	5	1
GO:BP	cell-cell junction assembly	GO:0007043	0.033	91	2
GO:BP	epithelial cell proliferation involved in prostate gland development	GO:0060767	0.033	5	1
GO:BP	positive regulation of neuroepithelial cell differentiation	GO:1902913	0.033	5	1
GO:BP	morphogenesis of an endothelium	GO:0003159	0.033	5	1
GO:BP	synapse assembly	GO:0007416	0.033	89	2
GO:BP	foregut morphogenesis	GO:0007440	0.033	5	1
GO:BP	negative regulation of intracellular signal transduction	GO:1902532	0.033	282	3
GO:BP	lymphocyte migration into lymphoid organs	GO:0097021	0.033	5	1
GO:BP	presynaptic active zone organization	GO:1990709	0.033	5	1

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	termination of RNA polymerase II transcription	GO:0006369	0.033	5	1
GO:BP	epithelium development	GO:0060429	0.033	555	4
GO:BP	endothelial tube morphogenesis	GO:0061154	0.033	5	1
GO:BP	intracellular chloride ion homeostasis	GO:0030644	0.033	5	1
GO:BP	determination of dorsal identity	GO:0048263	0.033	5	1
GO:BP	positive regulation of fibroblast growth factor receptor signaling pathway	GO:0045743	0.033	5	1
GO:BP	regulation of mRNA export from nucleus	GO:0010793	0.033	5	1
GO:BP	regulation of phosphoprotein phosphatase activity	GO:0043666	0.033	5	1
GO:BP	positive regulation of locomotion	GO:0040017	0.033	276	3
GO:BP	positive regulation of nucleobase-containing compound metabolic process	GO:0045935	0.033	1265	6
GO:BP	appendage morphogenesis	GO:0035107	0.033	89	2
GO:BP	neural tube development	GO:0021915	0.033	90	2
GO:BP	negative regulation of protein sumoylation	GO:0033234	0.033	5	1
GO:BP	embryonic foregut morphogenesis	GO:0048617	0.033	5	1
GO:BP	regulation of mononuclear cell proliferation	GO:0032944	0.033	86	2
GO:BP	limb morphogenesis	GO:0035108	0.033	89	2
GO:BP	negative regulation of microtubule polymerization	GO:0031115	0.033	5	1
GO:BP	microtubule organizing center organization	GO:0031023	0.033	91	2
GO:BP	leukocyte chemotaxis	GO:0030595	0.033	89	2
GO:BP	intracellular monoatomic anion homeostasis	GO:0030002	0.033	5	1
GO:BP	positive regulation of myelination	GO:0031643	0.033	5	1
GO:BP	regulation of cellular metabolic process	GO:0031323	0.033	3615	11
GO:BP	circulatory system development	GO:0072359	0.035	568	4
GO:BP	regulation of protein ubiquitination	GO:0031396	0.035	94	2
GO:BP	positive regulation of epithelial cell proliferation	GO:0050679	0.036	95	2
GO:BP	regulation of cell-cell adhesion mediated by integrin	GO:0033632	0.036	6	1
GO:BP	neurogenesis	GO:0022008	0.036	932	5
GO:BP	trachea morphogenesis	GO:0060439	0.036	6	1
GO:BP	lung-associated mesenchyme development	GO:0060484	0.036	6	1
GO:BP	gastrulation	GO:0007369	0.036	96	2
GO:BP	regulation of mast cell chemotaxis	GO:0060753	0.036	6	1
GO:BP	regulation of ribonucleoprotein complex localization	GO:2000197	0.036	6	1
GO:BP	hair cycle phase	GO:0044851	0.036	6	1
GO:BP	biological phase	GO:0044848	0.036	6	1
GO:BP	positive regulation of potassium ion import across plasma membrane	GO:1903288	0.036	6	1
GO:BP	centromeric sister chromatid cohesion	GO:0070601	0.036	6	1
GO:BP	regulation of leukocyte proliferation	GO:0070663	0.036	96	2
GO:BP	response to external biotic stimulus	GO:0043207	0.036	581	4
GO:BP	regulation of potassium ion import	GO:1903286	0.036	6	1
GO:BP	regulation of mitotic cytokinesis	GO:1902412	0.036	6	1
GO:BP	regulation of response to stress	GO:0080134	0.036	578	4
GO:BP	regulation of pancreatic juice secretion	GO:0090186	0.036	6	1

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	regulation of secondary heart field cardioblast proliferation	GO:0003266	0.036	6	1
GO:BP	limb joint morphogenesis	GO:0036022	0.036	6	1
GO:BP	immunological synapse formation	GO:0001771	0.036	6	1
GO:BP	response to other organism	GO:0051707	0.036	580	4
GO:BP	renal vesicle formation	GO:0072033	0.036	6	1
GO:BP	neuron fate determination	GO:0048664	0.036	6	1
GO:BP	tissue morphogenesis	GO:0048729	0.036	296	3
GO:BP	negative regulation of protein metabolic process	GO:0051248	0.036	301	3
GO:BP	centriole-centriole cohesion	GO:0010457	0.036	6	1
GO:BP	determination of dorsal/ventral asymmetry	GO:0048262	0.036	6	1
GO:BP	regulation of hair follicle maturation	GO:0048819	0.036	6	1
GO:BP	response to biotic stimulus	GO:0009607	0.037	596	4
GO:BP	establishment of vesicle localization	GO:0051650	0.037	102	2
GO:BP	mesenchymal to epithelial transition involved in metanephros morphogenesis	GO:0003337	0.037	7	1
GO:BP	regulation of sodium ion transmembrane transport	GO:1902305	0.037	7	1
GO:BP	regulation of T cell chemotaxis	GO:0010819	0.037	7	1
GO:BP	dendritic cell migration	GO:0036336	0.037	7	1
GO:BP	pancreatic juice secretion	GO:0030157	0.037	7	1
GO:BP	regulation of cardioblast proliferation	GO:0003264	0.037	7	1
GO:BP	sulfur compound biosynthetic process	GO:0044272	0.037	103	2
GO:BP	cellular biosynthetic process	GO:0044249	0.037	5463	14
GO:BP	cardioblast proliferation	GO:0003263	0.037	7	1
GO:BP	ganglion development	GO:0061548	0.037	7	1
GO:BP	hair follicle maturation	GO:0048820	0.037	7	1
GO:BP	tongue morphogenesis	GO:0043587	0.037	7	1
GO:BP	regulation of signaling	GO:0023051	0.037	1789	7
GO:BP	negative regulation of cellular process	GO:0048523	0.037	2699	9
GO:BP	regulation of nephron tubule epithelial cell differentiation	GO:0072182	0.037	7	1
GO:BP	nephron tubule epithelial cell differentiation	GO:0072160	0.037	7	1
GO:BP	dense core granule localization	GO:0032253	0.037	7	1
GO:BP	negative regulation of RIG-I signaling pathway	GO:0039536	0.037	7	1
GO:BP	positive regulation of T cell chemotaxis	GO:0010820	0.037	7	1
GO:BP	positive regulation of glycoprotein biosynthetic process	GO:0010560	0.037	7	1
GO:BP	ribonucleoprotein complex localization	GO:0071166	0.037	7	1
GO:BP	protein localization to ciliary membrane	GO:1903441	0.037	7	1
GO:BP	signal transduction in response to DNA damage	GO:0042770	0.037	106	2
GO:BP	appendage development	GO:0048736	0.037	105	2
GO:BP	regulation of chemotaxis	GO:0050920	0.037	106	2
GO:BP	proteasomal protein catabolic process	GO:0010498	0.037	316	3
GO:BP	lymphocyte activation	GO:0046649	0.037	318	3
GO:BP	dorsal/ventral axis specification	GO:0009950	0.037	7	1
GO:BP	limb development	GO:0060173	0.037	105	2
GO:BP	regulation of cell communication	GO:0010646	0.037	1784	7
GO:BP	regulation of synapse organization	GO:0050807	0.037	107	2
GO:BP	positive regulation of glial cell proliferation	GO:0060252	0.037	7	1

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	layer formation in cerebral cortex	GO:0021819	0.037	7	1
GO:BP	cytokinesis	GO:0000910	0.037	108	2
GO:BP	immune system development	GO:0002520	0.037	105	2
GO:BP	positive regulation of lymphocyte activation	GO:0051251	0.037	104	2
GO:BP	negative regulation of cell adhesion mediated by integrin	GO:0033629	0.037	7	1
GO:BP	glial cell fate commitment	GO:0021781	0.037	7	1
GO:BP	cytoskeleton organization	GO:0007010	0.037	968	5
GO:BP	organelle localization	GO:0051640	0.038	324	3
GO:BP	osteoblast differentiation	GO:0001649	0.038	109	2
GO:BP	positive regulation of developmental process	GO:0051094	0.038	620	4
GO:BP	positive regulation of neurogenesis	GO:0050769	0.038	110	2
GO:BP	regulation of synapse structure or activity	GO:0050803	0.038	110	2
GO:BP	regulation of multicellular organismal process	GO:0051239	0.038	1373	6
GO:BP	cellular response to stimulus	GO:0051716	0.039	4365	12
GO:BP	positive regulation of leukocyte activation	GO:0002696	0.039	112	2
GO:BP	regulation of protein modification by small protein conjugation or removal	GO:1903320	0.039	112	2
GO:BP	odontoblast differentiation	GO:0071895	0.040	8	1
GO:BP	regulation of phosphatase activity	GO:0010921	0.040	8	1
GO:BP	mast cell chemotaxis	GO:0002551	0.040	8	1
GO:BP	regulation of post-translational protein modification	GO:1901873	0.040	115	2
GO:BP	embryonic brain development	GO:1990403	0.040	8	1
GO:BP	positive regulation of nervous system process	GO:0031646	0.040	8	1
GO:BP	induction of positive chemotaxis	GO:0050930	0.040	8	1
GO:BP	regulation of glial cell proliferation	GO:0060251	0.040	8	1
GO:BP	embryonic skeletal joint morphogenesis	GO:0060272	0.040	8	1
GO:BP	positive regulation of lymphocyte chemotaxis	GO:0140131	0.040	8	1
GO:BP	negative regulation of sodium ion transport	GO:0010766	0.040	8	1
GO:BP	T cell chemotaxis	GO:0010818	0.040	8	1
GO:BP	negative regulation of glial cell differentiation	GO:0045686	0.040	8	1
GO:BP	negative regulation of digestive system process	GO:0060457	0.040	8	1
GO:BP	morphogenesis of an epithelial bud	GO:0060572	0.040	8	1
GO:BP	endodermal cell fate commitment	GO:0001711	0.040	8	1
GO:BP	organic substance biosynthetic process	GO:1901576	0.040	5546	14
GO:BP	cellular response to growth factor stimulus	GO:0071363	0.041	342	3
GO:BP	response to stress	GO:0006950	0.041	1852	7
GO:BP	positive regulation of cell activation	GO:0050867	0.041	118	2
GO:BP	glial cell differentiation	GO:0010001	0.042	119	2
GO:BP	cell-cell junction organization	GO:0045216	0.042	120	2
GO:BP	regulation of protein sumoylation	GO:0033233	0.043	9	1
GO:BP	adherens junction assembly	GO:0034333	0.043	9	1
GO:BP	hypothalamus development	GO:0021854	0.043	9	1
GO:BP	ectoderm development	GO:0007398	0.043	9	1
GO:BP	mast cell migration	GO:0097531	0.043	9	1
GO:BP	sialylation	GO:0097503	0.043	9	1

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	positive regulation of isotype switching to IgG isotypes	GO:0048304	0.043	9	1
GO:BP	cellular hyperosmotic response	GO:0071474	0.043	9	1
GO:BP	choline transport	GO:0015871	0.043	9	1
GO:BP	negative regulation of apoptotic signaling pathway	GO:2001234	0.043	121	2
GO:BP	lymphangiogenesis	GO:0001946	0.043	9	1
GO:BP	negative regulation of gliogenesis	GO:0014014	0.043	9	1
GO:BP	secretory granule localization	GO:0032252	0.043	9	1
GO:BP	positive regulation of glycoprotein metabolic process	GO:1903020	0.043	9	1
GO:BP	regulation of lymphocyte chemotaxis	GO:1901623	0.043	9	1
GO:BP	response to growth factor	GO:0070848	0.043	350	3
GO:BP	negative regulation of oxidative stress-induced neuron intrinsic apoptotic signaling	GO:1903377	0.043	9	1
GO:BP	negative regulation of signal transduction	GO:0009968	0.044	669	4
GO:BP	cellular response to cytokine stimulus	GO:0071345	0.044	357	3
GO:BP	regulation of leukocyte cell-cell adhesion	GO:1903037	0.044	125	2
GO:BP	regulation of RNA export from nucleus	GO:0046831	0.044	10	1
GO:BP	negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage	GO:1902230	0.044	10	1
GO:BP	negative regulation of macromolecule metabolic process	GO:0010605	0.044	1457	6
GO:BP	regulation of oxidative stress-induced neuron intrinsic apoptotic signaling pathway	GO:1903376	0.044	10	1
GO:BP	tetrahydrofolate interconversion	GO:0035999	0.044	10	1
GO:BP	negative regulation of T cell apoptotic process	GO:0070233	0.044	10	1
GO:BP	synaptic vesicle cycle	GO:0099504	0.044	126	2
GO:BP	regulation of epithelial cell differentiation involved in kidney development	GO:2000696	0.044	10	1
GO:BP	cell chemotaxis	GO:0060326	0.044	130	2
GO:BP	mesenchymal to epithelial transition	GO:0060231	0.044	10	1
GO:BP	positive regulation of myoblast proliferation	GO:2000288	0.044	10	1
GO:BP	cytoskeleton-dependent intracellular transport	GO:0030705	0.044	130	2
GO:BP	hyperosmotic response	GO:0006972	0.044	10	1
GO:BP	cell development	GO:0048468	0.044	1461	6
GO:BP	regulation of multicellular organismal development	GO:2000026	0.044	675	4
GO:BP	sympathetic nervous system development	GO:0048485	0.044	10	1
GO:BP	negative regulation of intrinsic apoptotic signaling pathway by p53 class mediator	GO:1902254	0.044	10	1
GO:BP	metanephric renal vesicle morphogenesis	GO:0072283	0.044	10	1
GO:BP	regulation of metal ion transport	GO:0010959	0.044	126	2
GO:BP	spindle organization	GO:0007051	0.044	128	2
GO:BP	mesenchymal stem cell differentiation	GO:0072497	0.044	10	1
GO:BP	embryonic skeletal joint development	GO:0072498	0.044	10	1

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	regulation of nucleobase-containing compound transport	GO:0032239	0.044	10	1
GO:BP	protein dephosphorylation	GO:0006470	0.044	127	2
GO:BP	regulation of T cell activation	GO:0050863	0.044	129	2
GO:BP	axon ensheathment in central nervous system	GO:0032291	0.044	10	1
GO:BP	regulation of heterotypic cell-cell adhesion	GO:0034114	0.044	10	1
GO:BP	positive regulation of skeletal muscle tissue development	GO:0048643	0.044	10	1
GO:BP	establishment of blood-brain barrier	GO:0060856	0.044	10	1
GO:BP	dorsal/ventral neural tube patterning	GO:0021904	0.044	10	1
GO:BP	nephron tubule formation	GO:0072079	0.044	10	1
GO:BP	regulation of DNA methylation-dependent heterochromatin formation	GO:0090308	0.044	10	1
GO:BP	myelin maintenance	GO:0043217	0.044	10	1
GO:BP	leukocyte activation	GO:0045321	0.044	368	3
GO:BP	central nervous system myelination	GO:0022010	0.044	10	1
GO:BP	pattern recognition receptor signaling pathway	GO:0002221	0.044	131	2
GO:BP	innate immune response-activating signaling pathway	GO:0002758	0.045	132	2
GO:BP	ubiquitin-dependent protein catabolic process	GO:0006511	0.045	372	3
GO:BP	vesicle-mediated transport in synapse	GO:0099003	0.045	133	2
GO:BP	modification-dependent protein catabolic process	GO:0019941	0.046	376	3
GO:BP	lens morphogenesis in camera-type eye	GO:0002089	0.046	11	1
GO:BP	regulation of cell proliferation involved in heart morphogenesis	GO:2000136	0.046	11	1
GO:BP	male genitalia development	GO:0030539	0.046	11	1
GO:BP	positive regulation of RNA biosynthetic process	GO:1902680	0.046	1080	5
GO:BP	cell-cell adhesion mediated by integrin	GO:0033631	0.046	11	1
GO:BP	negative regulation of mRNA splicing, via spliceosome	GO:0048025	0.046	11	1
GO:BP	positive regulation of cell cycle process	GO:0090068	0.046	135	2
GO:BP	positive regulation of DNA-templated transcription	GO:0045893	0.046	1078	5
GO:BP	regulation of sulfur metabolic process	GO:0042762	0.046	11	1
GO:BP	trachea development	GO:0060438	0.046	11	1
GO:BP	positive regulation of blood pressure	GO:0045777	0.046	11	1
GO:BP	neuron intrinsic apoptotic signaling pathway in response to oxidative stress	GO:0036480	0.046	11	1
GO:BP	lymph vessel morphogenesis	GO:0036303	0.046	11	1
GO:BP	modification-dependent macromolecule catabolic process	GO:0043632	0.046	379	3
GO:BP	maintenance of synapse structure	GO:0099558	0.046	11	1
GO:BP	lipid phosphorylation	GO:0046834	0.046	11	1
GO:BP	skeletal system morphogenesis	GO:0048705	0.046	137	2
GO:BP	positive regulation of nervous system development	GO:0051962	0.047	138	2
GO:BP	regulation of immune effector process	GO:0002697	0.047	139	2
GO:BP	negative regulation of cell communication	GO:0010648	0.048	713	4
GO:BP	sex differentiation	GO:0007548	0.048	140	2
GO:BP	immune response	GO:0006955	0.048	714	4

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	negative regulation of signaling	GO:0023057	0.048	715	4
GO:BP	positive regulation of cell population proliferation	GO:0008284	0.048	390	3
GO:BP	positive regulation of neuron apoptotic process	GO:0043525	0.048	12	1
GO:BP	positive regulation of positive chemotaxis	GO:0050927	0.048	12	1
GO:BP	negative regulation of chondrocyte differentiation	GO:0032331	0.048	12	1
GO:BP	cell proliferation involved in heart morphogenesis	GO:0061323	0.048	12	1
GO:BP	positive regulation of DNA metabolic process	GO:0051054	0.048	141	2
GO:BP	protein neddylation	GO:0045116	0.048	12	1
GO:BP	regulation of isotype switching to IgG isotypes	GO:0048302	0.048	12	1
GO:BP	renal vesicle morphogenesis	GO:0072077	0.048	12	1
GO:BP	intraciliary retrograde transport	GO:0035721	0.048	12	1
GO:BP	negative regulation of gene expression	GO:0010629	0.048	720	4
GO:BP	negative regulation of lymphocyte apoptotic process	GO:0070229	0.048	12	1
GO:BP	regulation of myoblast proliferation	GO:2000291	0.048	12	1
GO:BP	negative regulation of mRNA processing	GO:0050686	0.048	12	1
GO:BP	DNA-templated transcription termination	GO:0006353	0.048	12	1
GO:BP	regulation of positive chemotaxis	GO:0050926	0.048	12	1
GO:BP	response to cytokine	GO:0034097	0.048	394	3
GO:BP	regulation of response to stimulus	GO:0048583	0.048	1991	7
GO:BP	reproductive structure development	GO:0048608	0.048	144	2
GO:BP	regulation of signal transduction	GO:0009966	0.048	1534	6
GO:BP	phosphate-containing compound metabolic process	GO:0006796	0.048	1535	6
GO:BP	negative regulation of metabolic process	GO:0009892	0.048	1537	6
GO:BP	reproductive system development	GO:0061458	0.048	145	2
GO:BP	leukocyte cell-cell adhesion	GO:0007159	0.048	145	2
GO:BP	regulation of sister chromatid cohesion	GO:0007063	0.049	13	1
GO:BP	positive regulation of muscle tissue development	GO:1901863	0.049	13	1
GO:BP	regulation of oligodendrocyte differentiation	GO:0048713	0.049	13	1
GO:BP	positive regulation of potassium ion transmembrane transport	GO:1901381	0.049	13	1
GO:BP	ganglioside biosynthetic process	GO:0001574	0.049	13	1
GO:BP	regulation of intrinsic apoptotic signaling pathway by p53 class mediator	GO:1902253	0.049	13	1
GO:BP	cell migration	GO:0016477	0.049	734	4
GO:BP	organ induction	GO:0001759	0.049	13	1
GO:BP	regulation of tubulin deacetylation	GO:0090043	0.049	13	1
GO:BP	tubulin deacetylation	GO:0090042	0.049	13	1
GO:BP	isotype switching to IgG isotypes	GO:0048291	0.049	13	1
GO:BP	N-glycan processing	GO:0006491	0.049	13	1
GO:BP	phosphorus metabolic process	GO:0006793	0.049	1554	6
GO:BP	regulation of establishment of cell polarity	GO:2000114	0.049	13	1
GO:BP	renal vesicle development	GO:0072087	0.049	13	1

Continued on next page

Table S1 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	regulation of skeletal muscle tissue development	GO:0048641	0.049	13	1
GO:BP	protein kinase C-activating G protein-coupled receptor signaling pathway	GO:0007205	0.049	13	1
GO:BP	cell adhesion	GO:0007155	0.049	738	4
GO:BP	activation of innate immune response	GO:0002218	0.049	147	2
GO:BP	negative regulation of apoptotic process	GO:0043066	0.049	403	3
GO:BP	positive regulation of mesenchymal cell proliferation	GO:0002053	0.049	13	1
GO:BP	tongue development	GO:0043586	0.049	13	1
GO:BP	response to organic substance	GO:0010033	0.049	1127	5
GO:BP	intrinsic apoptotic signaling pathway	GO:0097193	0.049	150	2

Table S2: Full list of Gene Ontology (GO) terms for genes containing outlier loci detected in *Bradypodion melanocephalum* populations. Provided are the GO Source, GO term name and ID, Adjusted p-value, Term size (total genes associated with the GO term), and Intersection size (number of genes from this study matching genes in the term).

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	protein transport	GO:0015031	0.0038	753	5
GO:BP	protein localization	GO:0008104	0.0038	1401	6
GO:BP	nitrogen compound transport	GO:0071705	0.0038	1081	6
GO:BP	organic substance transport	GO:0071702	0.0038	1395	6
GO:BP	cellular macromolecule localization	GO:0070727	0.0038	1405	6
GO:BP	localization	GO:0051179	0.0054	3427	8
GO:BP	establishment of protein localization	GO:0045184	0.0054	907	5
GO:BP	macromolecule localization	GO:0033036	0.0072	1713	6
GO:BP	transport	GO:0006810	0.0119	2843	7
GO:BP	cellular localization	GO:0051641	0.0135	1994	6
GO:BP	establishment of localization	GO:0051234	0.0143	3017	7
GO:BP	epithelial-mesenchymal cell signaling	GO:0060684	0.0199	1	1
GO:BP	entry receptor-mediated virion attachment to host cell	GO:0098670	0.0199	1	1
GO:BP	mesenchymal to epithelial transition involved in metanephric renal vesicle formation	GO:0072285	0.0199	1	1
GO:BP	mesenchymal to epithelial transition involved in renal vesicle formation	GO:0072036	0.0199	1	1
GO:BP	ventral midline determination	GO:0007371	0.0199	1	1
GO:BP	angiotensin-mediated drinking behavior	GO:0003051	0.0199	1	1
GO:BP	positive regulation of cardiac muscle contraction	GO:0060452	0.0254	3	1
GO:BP	regulation of proline transport	GO:0070881	0.0254	2	1
GO:BP	metanephric renal vesicle formation	GO:0072093	0.0254	3	1
GO:BP	receptor-mediated virion attachment to host cell	GO:0046813	0.0254	3	1
GO:BP	renal protein absorption	GO:0097017	0.0254	3	1
GO:BP	pancreas morphogenesis	GO:0061113	0.0254	3	1
GO:BP	tryptophan transport	GO:0015827	0.0254	3	1
GO:BP	ascorbate homeostasis	GO:0140576	0.0254	2	1
GO:BP	regulation of heart morphogenesis	GO:2000826	0.0254	3	1
GO:BP	multicellular organismal-level homeostasis	GO:0048871	0.0254	387	3

Continued on next page

Table S2 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	positive regulation of proline import across plasma membrane	GO:1902836	0.0254	2	1
GO:BP	L-proline transmembrane transport	GO:1904555	0.0254	3	1
GO:BP	S-adenosylmethionine biosynthetic process	GO:0006556	0.0254	3	1
GO:BP	determination of left/right asymmetry in lateral mesoderm	GO:0003140	0.0254	3	1
GO:BP	ventral midline development	GO:0007418	0.0254	3	1
GO:BP	L-proline import across plasma membrane	GO:1904271	0.0254	3	1
GO:BP	proline import across plasma membrane	GO:1905647	0.0254	3	1
GO:BP	midgut development	GO:0007494	0.0254	3	1
GO:BP	regulation of L-proline import across plasma membrane	GO:1905735	0.0254	2	1
GO:BP	brain renin-angiotensin system	GO:0002035	0.0254	2	1
GO:BP	regulation of blood volume by renin-angiotensin	GO:0002016	0.0254	3	1
GO:BP	positive regulation of L-proline import across plasma membrane	GO:1905737	0.0254	2	1
GO:BP	regulation of proline import across plasma membrane	GO:1902834	0.0254	2	1
GO:BP	adhesion of symbiont to host cell	GO:0044650	0.0301	4	1
GO:BP	positive regulation of striated muscle contraction	GO:0045989	0.0301	4	1
GO:BP	regulation of somatic stem cell population maintenance	GO:1904672	0.0301	4	1
GO:BP	negative regulation of hair follicle development	GO:0051799	0.0301	4	1
GO:BP	virion attachment to host cell	GO:0019062	0.0301	4	1
GO:BP	protein localization to organelle	GO:0033365	0.0319	545	3
GO:BP	positive regulation of gap junction assembly	GO:1903598	0.0325	5	1
GO:BP	drinking behavior	GO:0042756	0.0325	5	1
GO:BP	adhesion of symbiont to host	GO:0044406	0.0325	5	1
GO:BP	thalamus development	GO:0021794	0.0325	5	1
GO:BP	cobalamin transport	GO:0015889	0.0325	5	1
GO:BP	regulation of gap junction assembly	GO:1903596	0.0325	5	1
GO:BP	left/right axis specification	GO:0070986	0.0363	6	1
GO:BP	commissural neuron axon guidance	GO:0071679	0.0363	6	1
GO:BP	renal vesicle formation	GO:0072033	0.0363	6	1
GO:BP	astrocyte activation	GO:0048143	0.0363	6	1
GO:BP	lateral mesoderm development	GO:0048368	0.0370	7	1
GO:BP	regulation of amino acid import across plasma membrane	GO:0010958	0.0370	7	1
GO:BP	gap junction assembly	GO:0016264	0.0370	7	1
GO:BP	regulation of amino acid transmembrane transport	GO:1903789	0.0370	7	1
GO:BP	response to iron ion	GO:0010039	0.0370	7	1
GO:BP	mesenchymal to epithelial transition involved in metanephros morphogenesis	GO:0003337	0.0370	7	1
GO:BP	cobalamin metabolic process	GO:0009235	0.0370	7	1
GO:BP	homeostasis of number of cells	GO:0048872	0.0370	171	2
GO:BP	cellular response to cholesterol	GO:0071397	0.0387	8	1
GO:BP	positive regulation of amino acid transport	GO:0051957	0.0387	8	1
GO:BP	Golgi vesicle transport	GO:0048193	0.0387	192	2

Continued on next page

Table S2 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	forebrain morphogenesis	GO:0048853	0.0387	8	1
GO:BP	nervous system process involved in regulation of systemic arterial blood pressure	GO:0001976	0.0387	8	1
GO:BP	proline transport	GO:0015824	0.0387	8	1
GO:BP	regulation of cardiac conduction	GO:1903779	0.0400	10	1
GO:BP	aromatic amino acid transport	GO:0015801	0.0400	9	1
GO:BP	cell migration	GO:0016477	0.0400	734	3
GO:BP	spinal cord dorsal/ventral patterning	GO:0021513	0.0400	10	1
GO:BP	metanephric renal vesicle morphogenesis	GO:0072283	0.0400	10	1
GO:BP	multicellular organismal process	GO:0032501	0.0400	3623	6
GO:BP	proline transmembrane transport	GO:0035524	0.0400	9	1
GO:BP	response to cholesterol	GO:0070723	0.0400	10	1
GO:BP	hematopoietic stem cell homeostasis	GO:0061484	0.0400	10	1
GO:BP	dorsal/ventral neural tube patterning	GO:0021904	0.0400	10	1
GO:BP	S-adenosylmethionine metabolic process	GO:0046500	0.0400	10	1
GO:BP	myoblast migration	GO:0051451	0.0400	9	1
GO:BP	mesenchymal to epithelial transition	GO:0060231	0.0400	10	1
GO:BP	positive regulation of amine transport	GO:0051954	0.0400	9	1
GO:BP	regulation of amino acid transport	GO:0051955	0.0400	10	1
GO:BP	cellular response to sterol	GO:0036315	0.0422	11	1
GO:BP	atrial septum morphogenesis	GO:0060413	0.0422	11	1
GO:BP	spinal cord patterning	GO:0021511	0.0422	11	1
GO:BP	type B pancreatic cell development	GO:0003323	0.0427	12	1
GO:BP	positive regulation of branching involved in ureteric bud morphogenesis	GO:0090190	0.0427	12	1
GO:BP	mitotic chromosome condensation	GO:0007076	0.0427	12	1
GO:BP	renal vesicle morphogenesis	GO:0072077	0.0427	12	1
GO:BP	protein import into mitochondrial matrix	GO:0030150	0.0427	12	1
GO:BP	ameboidal-type cell migration	GO:0001667	0.0427	233	2
GO:BP	glial cell activation	GO:0061900	0.0427	12	1
GO:BP	renal vesicle development	GO:0072087	0.0431	13	1
GO:BP	response to sterol	GO:0036314	0.0431	13	1
GO:BP	positive regulation of blood circulation	GO:1903524	0.0431	13	1
GO:BP	positive regulation of heart contraction	GO:0045823	0.0431	13	1
GO:BP	regulation of systemic arterial blood pressure by renin-angiotensin	GO:0003081	0.0431	13	1
GO:BP	positive regulation of mesenchymal cell proliferation	GO:0002053	0.0431	13	1
GO:BP	regulation of branching involved in ureteric bud morphogenesis	GO:0090189	0.0431	13	1
GO:BP	establishment of protein localization to organelle	GO:0072594	0.0440	252	2
GO:BP	astrocyte development	GO:0014002	0.0440	14	1
GO:BP	positive regulation of muscle contraction	GO:0045933	0.0440	15	1
GO:BP	regulation of hair follicle development	GO:0051797	0.0440	14	1
GO:BP	renal absorption	GO:0070293	0.0440	14	1
GO:BP	type B pancreatic cell differentiation	GO:0003309	0.0440	15	1
GO:BP	metanephric nephron morphogenesis	GO:0072273	0.0440	15	1
GO:BP	cell differentiation involved in metanephros development	GO:0072202	0.0440	15	1
GO:BP	dentate gyrus development	GO:0021542	0.0440	14	1
GO:BP	positive regulation of transcription elongation by RNA polymerase II	GO:0032968	0.0440	15	1

Continued on next page

Table S2 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	positive regulation of organic acid transport	GO:0032892	0.0440	15	1
GO:BP	cell motility	GO:0048870	0.0440	843	3
GO:BP	enteroendocrine cell differentiation	GO:0035883	0.0440	15	1
GO:BP	regulation of stem cell population maintenance	GO:2000036	0.0450	16	1
GO:BP	multicellular organismal-level iron ion homeostasis	GO:0060586	0.0450	16	1
GO:BP	regulation of kidney development	GO:0090183	0.0450	16	1
GO:BP	ARF protein signal transduction	GO:0032011	0.0450	16	1
GO:BP	regulation of ARF protein signal transduction	GO:0032012	0.0450	16	1
GO:BP	atrial septum development	GO:0003283	0.0470	17	1
GO:BP	cardiac atrium morphogenesis	GO:0003209	0.0470	17	1
GO:BP	homeostatic process	GO:0042592	0.0475	892	3
GO:BP	regulation of mesenchymal cell proliferation	GO:0010464	0.0486	18	1
GO:BP	positive regulation of smoothened signaling pathway	GO:0045880	0.0486	18	1

Table S3: Full list of Gene Ontology (GO) terms for genes containing outlier loci detected in *Bradypodion thamnobates* populations. Provided are the GO Source, GO term name and ID, Adjusted p-value, Term size (total genes associated with the GO term), and Intersection size (number of genes from this study matching genes in the term).

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	cerebral cortex development	GO:0021987	0.049	55	1
GO:BP	forebrain cell migration	GO:0021885	0.049	34	1
GO:BP	regulation of cell division	GO:0051302	0.049	83	1
GO:BP	erythrocyte homeostasis	GO:0034101	0.049	76	1
GO:BP	erythrocyte differentiation	GO:0030218	0.049	73	1
GO:BP	telencephalon cell migration	GO:0022029	0.049	31	1
GO:BP	mitotic cytokinesis	GO:0000281	0.049	50	1
GO:BP	cilium or flagellum-dependent cell motility	GO:0001539	0.049	81	1
GO:BP	mitotic spindle organization	GO:0007052	0.049	79	1
GO:BP	pallium development	GO:0021543	0.049	80	1
GO:BP	cerebral cortex cell migration	GO:0021795	0.049	24	1
GO:BP	cilium-dependent cell motility	GO:0060285	0.049	81	1
GO:BP	cytoskeleton-dependent cytokinesis	GO:0061640	0.049	68	1

Table S4: Full list of Gene Ontology (GO) terms for genes containing outlier loci detected in *Bradypodion setaroi* populations. Provided are the GO Source, GO term name and ID, Adjusted p-value, Term size (total genes associated with the GO term), and Intersection size (number of genes from this study matching genes in the term).

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	regulation of cell motility	GO:2000145	0.004366382	462	2
GO:BP	positive regulation of flagellated sperm motility	GO:1902093	0.004366382	5	1
GO:BP	positive regulation of fertilization	GO:1905516	0.004366382	2	1
GO:BP	response to odorant	GO:1990834	0.004366382	2	1
GO:BP	locomotion	GO:0040011	0.004366382	574	2
GO:BP	chemosensory behavior	GO:0007635	0.004366382	6	1
GO:BP	positive regulation of cilium-dependent cell motility	GO:2000155	0.004366382	5	1

Continued on next page

Table S4 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	regulation of locomotion	GO:0040012	0.004366382	489	2
GO:BP	olfactory behavior	GO:0042048	0.004366382	6	1
GO:BP	positive regulation of cilium movement	GO:0003353	0.004366382	6	1
GO:BP	regulation of flagellated sperm motility	GO:1901317	0.004366382	6	1
GO:BP	regulation of fertilization	GO:0080154	0.004366382	4	1
GO:BP	microtubule cytoskeleton organization	GO:0000226	0.004366382	424	2
GO:BP	microtubule-based process	GO:0007017	0.004807676	625	2
GO:BP	regulation of cilium movement involved in cell motility	GO:0060295	0.005711103	10	1
GO:BP	regulation of cilium-dependent cell motility	GO:1902019	0.005711103	10	1
GO:BP	cell motility	GO:0048870	0.006751456	816	2
GO:BP	cytoskeleton organization	GO:0007010	0.00815941	923	2
GO:BP	regulation of cilium movement	GO:0003352	0.008655522	18	1
GO:BP	sperm axoneme assembly	GO:0007288	0.009550772	21	1
GO:BP	regulation of microtubule-based movement	GO:0060632	0.009550772	23	1
GO:BP	sperm flagellum assembly	GO:0120316	0.009550772	23	1
GO:BP	regulation of cell-substrate junction assembly	GO:0090109	0.011798265	31	1
GO:BP	regulation of focal adhesion assembly	GO:0051893	0.011798265	31	1
GO:BP	regulation of cell-substrate junction organization	GO:0150116	0.012421727	34	1
GO:BP	focal adhesion assembly	GO:0048041	0.013226006	40	1
GO:BP	motile cilium assembly	GO:0044458	0.013226006	42	1
GO:BP	cytoplasmic microtubule organization	GO:0031122	0.013226006	41	1
GO:BP	positive regulation of reproductive process	GO:2000243	0.013226006	42	1
GO:BP	cell-substrate junction assembly	GO:0007044	0.013393449	44	1
GO:BP	cell-substrate junction organization	GO:0150115	0.014138642	48	1
GO:BP	regulation of cell-matrix adhesion	GO:0001952	0.015122132	53	1
GO:BP	axoneme assembly	GO:0035082	0.016322096	59	1
GO:BP	fertilization	GO:0009566	0.017816137	71	1
GO:BP	sperm motility	GO:0097722	0.017816137	69	1
GO:BP	microtubule bundle formation	GO:0001578	0.017816137	77	1
GO:BP	cellular component assembly	GO:0022607	0.017816137	1959	2
GO:BP	cilium or flagellum-dependent cell motility	GO:0001539	0.017816137	81	1
GO:BP	cilium movement involved in cell motility	GO:0060294	0.017816137	78	1
GO:BP	cilium-dependent cell motility	GO:0060285	0.017816137	81	1
GO:BP	regulation of reproductive process	GO:2000241	0.017816137	82	1
GO:BP	flagellated sperm motility	GO:0030317	0.017816137	69	1
GO:BP	cellular component biogenesis	GO:0044085	0.018207022	2132	2
GO:BP	organelle organization	GO:0006996	0.018207022	2155	2
GO:BP	regulation of cell junction assembly	GO:1901888	0.019435056	96	1
GO:BP	spermatid development	GO:0007286	0.019435056	98	1
GO:BP	spermatid differentiation	GO:0048515	0.019545791	102	1
GO:BP	regulation of cell-substrate adhesion	GO:0010810	0.019545791	103	1
GO:BP	cilium movement	GO:0003341	0.019545791	105	1
GO:BP	cell-matrix adhesion	GO:0007160	0.022067031	121	1
GO:BP	protein autophosphorylation	GO:0046777	0.022526194	126	1
GO:BP	regulation of microtubule-based process	GO:0032886	0.024890917	142	1
GO:BP	germ cell development	GO:0007281	0.032304235	188	1
GO:BP	cell-substrate adhesion	GO:0031589	0.033890094	201	1
GO:BP	cell junction assembly	GO:0034329	0.036076493	218	1

Continued on next page

Table S4 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	cellular process involved in reproduction in multicellular organism	GO:0022412	0.036710187	227	1
GO:BP	cilium assembly	GO:0060271	0.036710187	234	1
GO:BP	spermatogenesis	GO:0007283	0.036710187	230	1
GO:BP	male gamete generation	GO:0048232	0.036855595	239	1
GO:BP	cilium organization	GO:0044782	0.039060651	258	1
GO:BP	positive regulation of cell motility	GO:2000147	0.039060651	262	1
GO:BP	microtubule-based movement	GO:0007018	0.039160334	267	1
GO:BP	positive regulation of locomotion	GO:0040017	0.039256715	272	1
GO:BP	behavior	GO:0007610	0.042175083	297	1
GO:BP	cellular component organization	GO:0016043	0.044491274	4094	2
GO:BP	regulation of cell adhesion	GO:0030155	0.045269014	338	1
GO:BP	plasma membrane bounded cell projection assembly	GO:0120031	0.045269014	344	1
GO:BP	cellular component organization or biogenesis	GO:0071840	0.045269014	4241	2
GO:BP	gamete generation	GO:0007276	0.045269014	341	1
GO:BP	cell projection assembly	GO:0030031	0.045653099	352	1
GO:BP	cell junction organization	GO:0034330	0.046263166	367	1
GO:BP	multicellular organismal reproductive process	GO:0048609	0.046263166	366	1
GO:BP	multicellular organism reproduction	GO:0032504	0.046740402	376	1

Table S5: Full list of Gene Ontology (GO) terms for genes containing outlier loci detected in *Bradypodion ventrale* populations. Provided are the GO Source, GO term name and ID, Adjusted p-value, Term size (total genes associated with the GO term), and Intersection size (number of genes from this study matching genes in the term).

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	protein metabolic process	GO:0019538	0.032	3817	9
GO:BP	regulation of male gonad development	GO:2000018	0.032	4	1
GO:BP	positive regulation of male gonad development	GO:2000020	0.032	4	1
GO:BP	regulation of reproductive process	GO:2000241	0.032	82	2
GO:BP	Kit signaling pathway	GO:0038109	0.032	3	1
GO:BP	regulation of nitrogen compound metabolic process	GO:0051171	0.032	3390	8
GO:BP	response to stem cell factor	GO:0036215	0.032	3	1
GO:BP	protein modification process	GO:0036211	0.032	2283	6
GO:BP	cellular response to erythropoietin	GO:0036018	0.032	2	1
GO:BP	gonad development	GO:0008406	0.032	107	2
GO:BP	response to erythropoietin	GO:0036017	0.032	2	1
GO:BP	regulation of DNA replication	GO:0006275	0.032	71	2
GO:BP	protein acetylation	GO:0006473	0.032	98	2
GO:BP	regulation of macrophage fusion	GO:0034239	0.032	3	1
GO:BP	cell projection assembly	GO:0030031	0.032	352	3
GO:BP	macrophage fusion	GO:0034238	0.032	3	1
GO:BP	nitrogen compound metabolic process	GO:0006807	0.032	8426	12
GO:BP	positive regulation of mast cell cytokine production	GO:0032765	0.032	2	1
GO:BP	male gonad development	GO:0008584	0.032	64	2
GO:BP	regulation of mast cell cytokine production	GO:0032763	0.032	2	1
GO:BP	response to manganese ion	GO:0010042	0.032	3	1
GO:BP	positive regulation of gonad development	GO:1905941	0.032	4	1
GO:BP	regulation of cell cycle process	GO:0010564	0.032	373	3

Continued on next page

Table S5 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	cellular component organization	GO:0016043	0.032	4094	8
GO:BP	membrane protein intracellular domain proteolysis	GO:0031293	0.032	4	1
GO:BP	regulation of metabolic process	GO:0019222	0.032	4106	8
GO:BP	positive regulation of macrophage fusion	GO:0034241	0.032	3	1
GO:BP	mast cell cytokine production	GO:0032762	0.032	2	1
GO:BP	cellular response to stem cell factor stimulus	GO:0036216	0.032	3	1
GO:BP	erythropoietin-mediated signaling pathway	GO:0038162	0.032	2	1
GO:BP	regulation of protein metabolic process	GO:0051246	0.032	1224	5
GO:BP	regulation of transferase activity	GO:0051338	0.032	421	3
GO:BP	male sex differentiation	GO:0046661	0.032	80	2
GO:BP	development of primary male sexual characteristics	GO:0046546	0.032	65	2
GO:BP	clustering of voltage-gated sodium channels	GO:0045162	0.032	4	1
GO:BP	regulation of macromolecule metabolic process	GO:0060255	0.032	3827	8
GO:BP	mast cell differentiation	GO:0060374	0.032	3	1
GO:BP	development of primary sexual characteristics	GO:0045137	0.032	110	2
GO:BP	positive regulation of mast cell proliferation	GO:0070668	0.032	4	1
GO:BP	primary metabolic process	GO:0044238	0.032	8916	12
GO:BP	positive regulation of protein serine/threonine kinase activity	GO:0071902	0.032	86	2
GO:BP	regulation of gonad development	GO:1905939	0.032	4	1
GO:BP	melanocyte migration	GO:0097324	0.032	2	1
GO:BP	melanocyte adhesion	GO:0097326	0.032	1	1
GO:BP	regulation of primary metabolic process	GO:0080090	0.032	3479	8
GO:BP	de novo centriole assembly	GO:0097742	0.032	3	1
GO:BP	positive regulation of vascular associated smooth muscle cell differentiation	GO:1905065	0.032	4	1
GO:BP	negative regulation of mitotic cell cycle phase transition	GO:1901991	0.032	99	2
GO:BP	positive regulation of MAP kinase activity	GO:0043406	0.032	56	2
GO:BP	myeloid leukocyte activation	GO:0002274	0.032	75	2
GO:BP	organonitrogen compound metabolic process	GO:1901564	0.032	4495	9
GO:BP	glutathione import into mitochondrion	GO:0160007	0.032	1	1
GO:BP	macromolecule metabolic process	GO:0043170	0.032	8169	12
GO:BP	regulation of MAP kinase activity	GO:0043405	0.032	92	2
GO:BP	plasma membrane bounded cell projection assembly	GO:0120031	0.032	344	3
GO:BP	de novo centriole assembly involved in multi-ciliated epithelial cell differentiation	GO:0098535	0.032	3	1
GO:BP	cellular component organization or biogenesis	GO:0071840	0.033	4241	8
GO:BP	myeloid progenitor cell differentiation	GO:0002318	0.034	6	1
GO:BP	immature B cell differentiation	GO:0002327	0.034	5	1
GO:BP	positive regulation of smooth muscle cell differentiation	GO:0051152	0.034	5	1

Continued on next page

Table S5 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	phospholipid efflux	GO:0033700	0.034	6	1
GO:BP	reproductive structure development	GO:0048608	0.034	146	2
GO:BP	positive regulation of centriole replication	GO:0046601	0.034	5	1
GO:BP	positive regulation of keratinocyte migration	GO:0051549	0.034	6	1
GO:BP	negative regulation of mitotic cell cycle	GO:0045930	0.034	130	2
GO:BP	neuronal ion channel clustering	GO:0045161	0.034	6	1
GO:BP	mast cell proliferation	GO:0070662	0.034	6	1
GO:BP	regulation of mast cell proliferation	GO:0070666	0.034	6	1
GO:BP	organic substance metabolic process	GO:0071704	0.034	9451	12
GO:BP	protein acylation	GO:0043543	0.034	134	2
GO:BP	macromolecule modification	GO:0043412	0.034	2440	6
GO:BP	response to thyroid hormone	GO:0097066	0.034	6	1
GO:BP	cellular response to thyroid hormone stimulus	GO:0097067	0.034	6	1
GO:BP	negative regulation of cell cycle phase transition	GO:1901988	0.034	146	2
GO:BP	monocyte activation	GO:0042117	0.034	5	1
GO:BP	multi-ciliated epithelial cell differentiation	GO:1903251	0.034	5	1
GO:BP	regulation of protein catabolic process in the vacuole	GO:1904350	0.034	5	1
GO:BP	regulation of lysosomal protein catabolic process	GO:1905165	0.034	5	1
GO:BP	glutathione transmembrane transport	GO:0034775	0.034	5	1
GO:BP	glutathione transport	GO:0034635	0.034	6	1
GO:BP	sex differentiation	GO:0007548	0.034	141	2
GO:BP	histone H4-K12 acetylation	GO:0043983	0.034	6	1
GO:BP	reproductive system development	GO:0061458	0.034	147	2
GO:BP	hematopoietic stem cell migration	GO:0035701	0.036	7	1
GO:BP	N-terminal peptidyl-methionine acetylation	GO:0017196	0.036	7	1
GO:BP	Fc receptor signaling pathway	GO:0038093	0.036	7	1
GO:BP	positive regulation of protein kinase activity	GO:0045860	0.036	158	2
GO:BP	tissue migration	GO:0090130	0.036	158	2
GO:BP	epithelium migration	GO:0090132	0.036	153	2
GO:BP	epithelial cell migration	GO:0010631	0.036	153	2
GO:BP	regulation of keratinocyte migration	GO:0051547	0.036	7	1
GO:BP	positive regulation of syncytium formation by plasma membrane fusion	GO:0060143	0.038	8	1
GO:BP	metabolic process	GO:0008152	0.038	9775	12
GO:BP	negative regulation of cell cycle process	GO:0010948	0.038	167	2
GO:BP	positive regulation of dendritic cell cytokine production	GO:0002732	0.038	8	1
GO:BP	lysosomal protein catabolic process	GO:1905146	0.038	8	1
GO:BP	protein catabolic process in the vacuole	GO:0007039	0.038	8	1
GO:BP	mast cell chemotaxis	GO:0002551	0.038	8	1
GO:BP	regulation of cell cycle	GO:0051726	0.038	565	3
GO:BP	ectopic germ cell programmed cell death	GO:0035234	0.038	9	1
GO:BP	regulation of protein serine/threonine kinase activity	GO:0071900	0.038	174	2
GO:BP	peptidyl-methionine modification	GO:0018206	0.038	9	1
GO:BP	mast cell migration	GO:0097531	0.038	9	1
GO:BP	organelle assembly	GO:0070925	0.038	571	3
GO:BP	dendritic cell cytokine production	GO:0002371	0.038	9	1

Continued on next page

Table S5 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	DNA replication	GO:0006260	0.038	173	2
GO:BP	keratinocyte migration	GO:0051546	0.038	9	1
GO:BP	regulation of dendritic cell cytokine production	GO:0002730	0.038	9	1
GO:BP	regulation of mitotic cell cycle phase transition	GO:1901990	0.039	182	2
GO:BP	regulation of syncytium formation by plasma membrane fusion	GO:0060142	0.040	10	1
GO:BP	epithelial cell-cell adhesion	GO:0090136	0.040	10	1
GO:BP	regulation of cellular metabolic process	GO:0031323	0.040	3796	7
GO:BP	regulation of vascular associated smooth muscle cell differentiation	GO:1905063	0.040	10	1
GO:BP	regulation of cellular component organization	GO:0051128	0.040	1198	4
GO:BP	reproduction	GO:0000003	0.040	607	3
GO:BP	detection of mechanical stimulus involved in sensory perception of sound	GO:0050910	0.040	10	1
GO:BP	reproductive process	GO:0022414	0.040	601	3
GO:BP	positive regulation of developmental process	GO:0051094	0.040	610	3
GO:BP	negative regulation of ubiquitin-protein transferase activity	GO:0051444	0.042	11	1
GO:BP	positive regulation of cell adhesion mediated by integrin	GO:0033630	0.042	11	1
GO:BP	peptidyl-amino acid modification	GO:0018193	0.042	630	3
GO:BP	negative regulation of cell cycle	GO:0045786	0.043	201	2
GO:BP	lymphoid progenitor cell differentiation	GO:0002320	0.043	12	1
GO:BP	histone H4-K5 acetylation	GO:0043981	0.043	12	1
GO:BP	histone H4-K8 acetylation	GO:0043982	0.043	12	1
GO:BP	protein localization to kinetochore	GO:0034501	0.043	12	1
GO:BP	positive regulation of kinase activity	GO:0033674	0.043	207	2
GO:BP	protein localization to condensed chromosome	GO:1903083	0.043	12	1
GO:BP	cell-cell adhesion mediated by integrin	GO:0033631	0.043	12	1
GO:BP	anatomical structure development	GO:0048856	0.044	2946	6
GO:BP	cellular developmental process	GO:0048869	0.044	2079	5
GO:BP	programmed cell death involved in cell development	GO:0010623	0.044	13	1
GO:BP	regulation of centriole replication	GO:0046599	0.044	13	1
GO:BP	cell differentiation	GO:0030154	0.044	2079	5
GO:BP	positive regulation of protein metabolic process	GO:0051247	0.044	667	3
GO:BP	N-terminal protein amino acid acetylation	GO:0006474	0.044	13	1
GO:BP	regulation of biological process	GO:0050789	0.044	7524	10
GO:BP	positive regulation of MAPK cascade	GO:0043410	0.045	219	2
GO:BP	regulation of protein modification process	GO:0031399	0.045	678	3
GO:BP	mast cell degranulation	GO:0043303	0.046	14	1
GO:BP	embryonic hemopoiesis	GO:0035162	0.046	14	1
GO:BP	positive regulation of receptor signaling pathway via JAK-STAT	GO:0046427	0.047	15	1
GO:BP	regulation of cell cycle phase transition	GO:1901987	0.047	233	2
GO:BP	organelle organization	GO:0006996	0.047	2155	5
GO:BP	establishment of organelle localization	GO:0051656	0.047	230	2
GO:BP	megakaryocyte development	GO:0035855	0.047	15	1
GO:BP	establishment of localization	GO:0051234	0.047	3069	6

Continued on next page

Table S5 – continued from previous page

Source	Gene Ontology (GO) term name	Term ID	Adjusted p-value	Term size	Intersection size
GO:BP	detection of mechanical stimulus involved in sensory perception	GO:0050974	0.047	15	1
GO:BP	ameboidal-type cell migration	GO:0001667	0.047	232	2
GO:BP	cilium assembly	GO:0060271	0.047	234	2
GO:BP	mast cell activation involved in immune response	GO:0002279	0.048	16	1
GO:BP	positive regulation of receptor signaling pathway via STAT	GO:1904894	0.048	16	1
GO:BP	protein localization to chromosome, centromeric region	GO:0071459	0.048	16	1
GO:BP	regulation of smooth muscle cell differentiation	GO:0051150	0.048	16	1
GO:BP	cell cycle process	GO:0022402	0.048	719	3
GO:BP	mast cell mediated immunity	GO:0002448	0.048	16	1
GO:BP	non-membrane-bounded organelle assembly	GO:0140694	0.049	245	2
GO:BP	melanocyte differentiation	GO:0030318	0.050	17	1
GO:BP	positive regulation of tyrosine phosphorylation of STAT protein	GO:0042531	0.050	17	1
GO:BP	mitotic cell cycle phase transition	GO:0044772	0.050	250	2
GO:BP	regulation of DNA metabolic process	GO:0051052	0.050	250	2

Chapter 5

Conclusion

5.1 Thesis summary

The southern African dwarf chameleons (*Bradypodion*) are a relatively young lineage, with a large body of evidence demonstrating the highly adaptive capacity of these lizards. *Bradypodion* are well known for their habitat specialization, with evidence of convergent phenotypes throughout the phylogeny. However, the underlying genetic architecture of these phenotypes remains unknown for *Bradypodion*, and without adequate genomic resources, many evolutionary questions cannot be answered. In this thesis, I aimed to uncover the genetic architecture underlying adaptation in *Bradypodion*. In pursuit of this, I first assembled two genomes *de novo* for *Bradypodion* using long-read sequencing data, with completeness and contiguity metrics demonstrating the high quality of these assembled genomes (Chapter 2). I then proceeded to annotate these assemblies, identifying functional elements along the sequence of each genome. Complete, well-annotated genome assemblies were required for the detection of changes in gene family composition. By extracting the complete set of annotated protein-coding genes from each species, I implemented a phylogenetic comparative approach to identify both gene family change and the extant or ancestral lineage where the changes occurred. This allowed for insights into the genetic evolution of adaptive traits within *Bradypodion* which likely contributed to the diversification of ecomorphs in these lizards (Chapter 3).

As *Bradypodion* ecomorphs are known to display phenotypic convergence in similar natural habitat types, chameleon populations may also converge phenotypically in urban

habitats. A pairwise comparison of urban and natural populations was used to assess the genetic variation and diversity for five species of *Bradypodion*. This assessment allowed for the detection of genes containing outlier loci reflecting the divergence of allele frequencies between urban and natural populations. The verification of gene function was done using public gene ontology databases, which enabled the connection of outlier genes with their organismal functions. Having verified the gene functions, I detected possible phenotypic traits presumed to be associated with adaptation in urban habitats (Chapter 4). Overall, this thesis provides support that, with the appropriate genomic resources, the architecture and evolutionary underpinnings of adaptive traits may be uncovered in remarkable, non-model organisms such as species of *Bradypodion*.

5.2 Synthesis

Disentangling the mechanisms of adaptation across the tree of life remains an underlying thread in evolutionary biology (Munson, 1971; Powers et al., 1991; Campbell-Staton et al., 2020; Szulkin et al., 2020). In the pre-genomic era, identifying adaptive alleles was often an integrative process which leveraged genetics, physiology, and ecology (Lasky et al., 2023). With the advancement in sequencing technologies, uncovering the genetic architecture underlying adaptive traits became more attainable, but has remained fairly complex (Hansen, 2006; Hudson, 2008; Stapley et al., 2010). Furthermore, the process of linking genotypes to phenotypes still required significant resources, often limited to reduced representation genomic data (Margres et al., 2017; Card et al., 2023). Major advances in sequencing technology have significantly improved in recent years, reducing previous sequencing limitations (see Karst et al., 2021), and facilitating the rise in whole-genome assemblies. As a tool in evolutionary biology, access to whole genome assemblies enables the identification of specific genes underlying traits of interest (Gable et al., 2023). However, assembly efforts predominantly focused on model or charismatic organisms (Blaxter et al., 2022; Ebenezer et al., 2022). Efforts to correct these biases are underway, but reptile genome assemblies in particular are still lagging (Card et al., 2023), with a many key squamate lineages still lacking high-quality genome sequences, including Chamaeleonidae (Gable et al., 2023).

By generating whole genome assemblies, I have contributed toward bridging the gap in genomic resources available for Chamaeleonidae, and reptiles in general. By leveraging the advantages of long-read sequencing technology, spanning difficult regions to sequence, the

quality of these assemblies is comparatively high compared to other available genomic resources (Chapter 2). It is well documented that the quality of the reference genomes directly impacts the outcome of evolutionary inferences (Rhie et al., 2021; Bentley and Armstrong, 2022). While the intention for generating whole-genome assemblies for *Bradypodion* was to uncover genes associated with adaptation, these genomes also allowed for direct comparisons of structural elements within the genomes of *Bradypodion*. The findings in Chapter 2 provided the opportunity to explore synteny, the arrangement of genes or genetic sequences on chromosomes, within *Bradypodion* as well as between the distantly related *Anolis sagrei*. Research studies have routinely used *Anolis* lizards as a biological model system, often challenging long-standing paradigms in the fields of ecology and evolution (Losos, 2009; Muñoz et al., 2023). *Bradypodion* is ecologically similar to some anoles, with similarities also extended to genome structural organisation. Furthermore, both *Bradypodion* assemblies generated for this thesis are fairly consistent with the expected size and structure of other squamate genomes (Card et al., 2023). While these higher taxonomic comparisons still provide insight into the evolution of genomes over deep timescales, having access to multiple assemblies within a single genus allows for the evaluation of genome structures underlying diversification at a much finer scale.

Comparative analyses of whole genome assemblies have revealed that new genes often originate over evolutionary timescales, and the expansion of gene families contributes to genome evolution at a rate exceeding the nucleotide mutation rate (Rogers et al., 2009, 2021; Demuth and Hahn, 2009; Van de Peer, 2011). Changes in gene family composition have been associated with adaptive evolution in plants (Caputi et al., 2012) and mammals (Demuth et al., 2006). Similarly, the contrast in gene family expansions between *Bradypodion pumilum* and *B. ventrale* revealed insights into the genomic basis underlying adaptive solutions in response to different habitat types (Chapter 3). However, given the limited number of available acrodont genome assemblies (Gable et al., 2023), the power to assess adaptive traits driven by gene family changes developed over deeper timescales was greatly reduced, minimising additional evolutionary inferences that could be made in Chapter 3, unlike in other more saturated groups (Lind et al., 2019; Trouern-Trend et al., 2020; Bentley et al., 2023). It is clear there is still a need for additional reptile assemblies, however, additional assemblies for *Bradypodion* will enable leveraging the repeated emergence of ecomorphs to uncover the role of gene family evolution driving the diversification of adaptive phenotypes in this group, particularly targeting species with populations exhibiting divergent ecomorphs.

The findings presented in Chapter 4 support the current view that, in the right conditions, adaptation of a population to novel conditions can occur fairly rapidly (Catullo et al., 2019). It is well documented that selection pressures in urban areas have resulted in rapid adaptive responses in a number of taxa (Donihue and Lambert, 2015; Johnson and Munshi-South, 2017; Szulkin et al., 2020; Miles et al., 2021; Lewis et al., 2023; Tang et al., 2024). Even though it is unclear whether directional or disruptive selection has driven the allele frequency differences between *Bradypodion* populations, the arguments made in Chapter 4 explore why the outlier loci detected are likely a result of adaptive evolution operating in urban populations. There is an expectation that urban-adapted purely arboreal species tend toward morphological traits that enhance stability, clinging and inter-perch aerial movement (Winchell et al., 2020). Evidence for morphological divergence in urban populations of *Bradypodion* has not yet been rigorously interrogated. The few studies so far, however, show there are some differing traits such as ornamentation and head dimensions (Petford et al., 2023) and body size and condition (Barends and Tolley, in review).

Given that the vegetation density in urban areas provides less cover than in natural habitats, urban populations of *Bradypodion* are likely more exposed and are at higher risk of predation (Petford et al., 2023). To remain less conspicuous in exposed urban habitats, urban populations of *B. damaranum* exhibit reduced ornamentation relative to their natural counterparts, but this comes at a cost of effective signalling leading to a breakdown in intraspecific communication (Petford et al., 2023). Given this, urban populations in *Bradypodion* are likely to tend toward inornate phenotypes similar to open-canopy ecomorphs, discussed in Chapter 3. Thus, rapid adaptive solutions in *Bradypodion* occupying urban habitats are likely constrained by their evolutionary history and habitat use which may lead to repeated morphological outcomes, while also favouring traits unique to each species likely to enable urban living detected in Chapter 4.

In conclusion, the findings presented in this thesis show there is a suite of genetic pathways that most likely contribute to adaptation in *Bradypodion*. In addition, these findings provide a framework from which to formulate robust hypotheses to assess candidate genes underlying adaptive solutions in urban populations. The current knowledge gaps relate to implementing a replicated experimental design, targeting multiple urban populations in more than one urban habitat (Perrier et al., 2020). However, this may be logistically challenging, especially when collecting phenotypic, genotypic and fitness-related data simultaneously (Winchell et al., 2020). As an alternative, the pairwise approach I have used in this thesis still led to the detection of candidate genes likely to underlie adaptation.

However, these candidate genes do require validation through functional genomic work performed via targeted gene modification techniques to directly evaluate gene function in response to urban adaptive responses (Rasys et al., 2019). Nevertheless, this body of work provides evidence for *Bradypodion* being a useful model to assess the genetic underpinnings of adaptation, as well as extending the opportunity to further delve into the genetic architecture of adaptation in chameleons, and reptiles more broadly, given the genomic resources and candidate loci presented here.

References

Barends, J. M. and Tolley, K. A. (2024). Big cities, big bodies: urbanisation correlates with large body sizes and enhanced body condition in African dwarf chameleons (genus: *Bradypodion*). *African Zoology*. In review.

Bentley, B. P. and Armstrong, E. E. (2022). Good from far, but far from good: The impact of a reference genome on evolutionary inference. *Molecular Ecology Resources*.

Bentley, B. P., Carrasco-Valenzuela, T., Ramos, E. K., Pawar, H., Souza Arantes, L., Alexander, A., Banerjee, S. M., Masterson, P., Kuhlwilm, M., Pippel, M., et al. (2023). Divergent sensory and immune gene evolution in sea turtles with contrasting demographic and life histories. *Proceedings of the National Academy of Sciences*, 120(7):e2201076120.

Blaxter, M., Archibald, J. M., Childers, A. K., Coddington, J. A., Crandall, K. A., Di Palma, F., Durbin, R., Edwards, S. V., Graves, J. A., Hackett, K. J., et al. (2022). Why sequence all eukaryotes? *Proceedings of the National Academy of Sciences*, 119(4):e2115636118.

Campbell-Staton, S. C., Winchell, K. M., Rochette, N. C., Fredette, J., Maayan, I., Schweizer, R. M., and Catchen, J. (2020). Parallel selection on thermal physiology facilitates repeated adaptation of city lizards to urban heat islands. *Nature Ecology & Evolution*, 4(4):652–658.

Caputi, L., Malnoy, M., Goremykin, V., Nikiforova, S., and Martens, S. (2012). A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. *The Plant Journal*, 69(6):1030–1042.

Card, D. C., Jennings, W. B., and Edwards, S. V. (2023). Genome evolution and the future of phylogenomics of non-avian reptiles. *Animals*, 13(3):471.

Catullo, R. A., Llewelyn, J., Phillips, B. L., and Moritz, C. C. (2019). The potential for rapid evolution under anthropogenic climate change. *Current Biology*, 29(19):R996–R1007.

Demuth, J. P., Bie, T. D., Stajich, J. E., Cristianini, N., and Hahn, M. W. (2006). The evolution of mammalian gene families. *PloS one*, 1(1):e85.

Demuth, J. P. and Hahn, M. W. (2009). The life and death of gene families. *BioEssays*, 31(1):29–39.

Donihue, C. M. and Lambert, M. R. (2015). Adaptive evolution in urban ecosystems. *Ambio*, 44:194–203.

Ebenezer, T. E., Muigai, A. W., Nouala, S., Badaoui, B., Blaxter, M., Buddie, A. G., Jarvis, E. D., Korlach, J., Kuja, J. O., Lewin, H. A., et al. (2022). Africa: sequence 100,000 species to safeguard biodiversity. *Nature*, 603(7901):388–392.

Gable, S. M., Mendez, J. M., Bushroe, N. A., Wilson, A., Byars, M. I., and Tollis, M. (2023). The State of Squamate Genomics: Past, Present, and Future of Genome Research in the Most Speciose Terrestrial Vertebrate Order. *Genes*, 14(7):1387.

Hansen, T. F. (2006). The evolution of genetic architecture. *Annual Review of Ecology, Evolution, and Systematics*, 37:123–157.

Hudson, M. E. (2008). Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular ecology resources*, 8(1):3–17.

Johnson, M. T. and Munshi-South, J. (2017). Evolution of life in urban environments. *Science*, 358(6363):eaam8327.

Karst, S. M., Ziels, R. M., Kirkegaard, R. H., Sørensen, E. A., McDonald, D., Zhu, Q., Knight, R., and Albertsen, M. (2021). High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nature methods*, 18(2):165–169.

Lasky, J. R., Josephs, E. B., and Morris, G. P. (2023). Genotype–environment associations to reveal the molecular basis of environmental adaptation. *The Plant Cell*, 35(1):125–138.

Lewis, C. D., Levine, B. A., Schal, C., Vargo, E. L., and Booth, W. (2023). Decade long upsurge in mutations associated with pyrethroid resistance in bed bug populations in the usa. *Journal of Pest Science*, 96(1):415–423.

Lind, A. L., Lai, Y. Y., Mostovoy, Y., Holloway, A. K., Iannucci, A., Mak, A. C., Fondi, M., Orlandini, V., Eckalbar, W. L., Milan, M., et al. (2019). Genome of the Komodo dragon reveals adaptations in the cardiovascular and chemosensory systems of monitor lizards. *Nature ecology & evolution*, 3(8):1241–1252.

Losos, J. B. (2009). *Lizards in an Evolutionary Tree*, volume 60. University of California Press.

Margres, M. J., Wray, K. P., Hassinger, A. T., Ward, M. J., McGivern, J. J., Moriarty Lemmon, E., Lemmon, A. R., and Rokyta, D. R. (2017). Quantity, not quality: rapid adaptation in a polygenic trait proceeded exclusively through expression differentiation. *Molecular biology and evolution*, 34(12):3099–3110.

Miles, L. S., Carlen, E. J., Winchell, K. M., and Johnson, M. T. (2021). Urban evolution comes into its own: Emerging themes and future directions of a burgeoning field. *Evolutionary Applications*, 14(1):3–11.

Muñoz, M. M., Frishkoff, L. O., Pruett, J., and Mahler, D. L. (2023). Evolution of a model system: New insights from the study of *Anolis* lizards. *Annual Review of Ecology, Evolution, and Systematics*, 54:475–503.

Munson, R. (1971). Biological adaptation. *Philosophy of Science*, 38(2):200–215.

Perrier, C., Caizergues, A., and Charmantier, A. (2020). Adaptation Genomics in Urban Environments. In Szulkin, M., Munshi-South, J., and Charmantier, A., editors, *Urban Evolutionary Biology*. Oxford University Press, USA.

Petford, M. A., Herrel, A., Alexander, G. J., and Tolley, K. A. (2023). Differences between urban and natural populations of dwarf chameleons (*Bradypodion damaranum*): a case of urban warfare? *Urban Ecosystems*, pages 1–13.

Powers, D. A., Lauerman, T., Crawford, D., and DiMichele, L. (1991). Genetic mechanisms for adapting to a changing environment. *Annual Review of Genetics*, 25(1):629–660.

Rasys, A. M., Park, S., Ball, R. E., Alcala, A. J., Lauderdale, J. D., and Menke, D. B. (2019). CRISPR-Cas9 gene editing in lizards through microinjection of unfertilized oocytes. *Cell reports*, 28(9):2288–2292.

Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., et al. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, 592(7856):737–746.

Rogers, R. L., Bedford, T., and Hartl, D. L. (2009). Formation and Longevity of Chimeric and Duplicate Genes in *Drosophila melanogaster*. *Genetics*, 181(1):313–322.

Rogers, R. L., Grizzard, S. L., Titus-McQuillan, J. E., Bockrath, K., Patel, S., Wares, J. P., Garner, J. T., and Moore, C. C. (2021). Gene family amplification facilitates adaptation in freshwater unionid bivalve *Megalonaias nervosa*. *Molecular Ecology*, 30(5):1155–1173.

Stapley, J., Reger, J., Feulner, P. G., Smadja, C., Galindo, J., Ekblom, R., Bennison, C., Ball, A. D., Beckerman, A. P., and Slate, J. (2010). Adaptation genomics: the next generation. *Trends in ecology & evolution*, 25(12):705–712.

Szulkin, M., Munshi-South, J., and Charmantier, A. (2020). *Urban Evolutionary Biology*. Oxford University Press, USA.

Tang, Q., Vargo, E. L., Ahmad, I., Jiang, H., Varadínová, Z. K., Dovih, P., Kim, D., Bourguignon, T., Booth, W., Schal, C., et al. (2024). Solving the 250-year-old mystery of the origin and global spread of the german cockroach, blattella germanica. *Proceedings of the National Academy of Sciences*, 121(22):e2401185121.

Trouern-Trend, A. J., Falk, T., Zaman, S., Caballero, M., Neale, D. B., Langley, C. H., Dandekar, A. M., Stevens, K. A., and Wegrzyn, J. L. (2020). Comparative genomics of six *Juglans* species reveals disease-associated gene family contractions. *The Plant Journal*, 102(2):410–423.

Van de Peer, Y. (2011). A mystery unveiled. *Genome biology*, 12(5):1–2.

Winchell, K. M., Battles, A. C., and Moore, T. Y. (2020). Terrestrial Locomotor Evolution in Urban Environments. In Szulkin, M., Munshi-South, J., and Charmantier, A., editors, *Urban Evolutionary Biology*. Oxford University Press, USA.