

Using squamate systems to understand molecular underpinnings of evolutionary processes

by

Randy Luke Klabacka

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 10, 2022

Keywords: biogeography; phylogenetics, mitochondrial physiology, parthenogenesis, population genetics, senescence

Copyright 2022 by Randy Luke Klabacka

Approved by

Tonia S. Schwartz, Co-chair, Associate Professor, Department of Biological Sciences
Jamie R. Oaks, Co-chair, Associate Professor, Department of Biological Sciences
Geoffrey E. Hill, Professor, Department of Biological Sciences
Andreas N. Kavazis, Professor, School of Kinesiology

Abstract

The solidification of evolutionary biology as a scientific theory provided a foundation for understanding the source of life's variation, an objective that has since become a central aim in biology. While scientific acceptance of evolution answered some biological questions, it created and continues to create more questions than it has answered. Several questions being studied across the globe include (1) "How are species formed?" (2) "What factors influence trait evolution?", and (3) "How do changes in genetics and environment determine how phenotypes respond to selection?" Squamate reptiles, the most species-rich group of tetrapods, are a diverse natural resource for empirical approaches to understanding biological questions. Within this dissertation, I utilize three squamate reptile systems (the species complex of spotted flying lizard *Draco maculatus*, the north American whiptail lizards *Aspidoscelis*, and the western terrestrial garter snake *Thamnophis elegans*) to answer questions regarding causes of lineage diversification, consequences of asexual reproduction, and genomics of life history evolution. I integrate results from molecular phylogenetics, whole-organism performance, mitochondrial physiology, and population genomics to test the riverine barrier hypothesis, the association of asexual reproduction with mitochondrial respiration, and the genetic underpinnings of senescence. Finally, I discuss the role of the mitochondrion in shaping evolutionary patterns, examine findings from this dissertation in broader biological and societal contexts, and provide recommendations for future endeavors to further tease out answers to these complex questions.

Acknowledgements

I produced this work with the help of many contributors, and here I both recognize and thank them. In addition to recognition here, each data chapter contains an acknowledgement section.

I thank my advisors, Jamie Oaks and Tonia Schwartz, for providing me with resources, counsel, and a healthy environment to do science. Their focus, relentlessness, honesty, inclusivity, patience, encouragement, and vision helped me discover the type of scientist I want to be. I once heard success defined as “lifting others as you rise”, and every time I left a meeting with them I felt uplifted and motivated. I am grateful to them for providing me with a true image of success in science.

I thank the additional two members of my committee, Geoff Hill and Andreas Kavazis, for providing their expertise during committee meetings. Geoff inspired me on multiple levels– as a scientist, a writer, a teacher, and a mentor. Andreas likewise provided crucial mentorship, and he also provided lab resources and training.

I thank additional mentors and collaborators Anne Bronikowski, Jim McGuire, and Matt Fujita. We had many productive, encouraging, and enjoyable conversations, including in the field, over email, in collaborative meetings, and in casual settings. Being able to work with leading experts in life history evolution, *Draco* diversification, and parthenogenesis, who all happen to be of exemplary character, was a true honor.

I thank my office mate and close friend, Kerry Cobb, for always being up for a distracting, though insightful, conversation of any subject matter. I also thank him and fellow Phyletician (Aundrea Westfall, Breanna Siple, Claire Tracey, Matt Buehler, Saman Jahangiri, Tanner Myers, Tashitso Anamza) and Schwartz (Abby Beatty, Amanda Clark, Anet Filipova, Dasia Simpson, Ryan Hardin, Taylor McKibben, Sam Weaver) lab members for many fun social gatherings and helpful lab meetings. Our adventures on and around the waterways of Alabama supplied much needed breaks from the computer.

I thank additional academic collaborators, including Haley Parry, Jesse Grismer, Jessica Judson, Jeff Yap, Laurie Stevison, and Matthew Wolak. I enjoyed all our lab work, white board discussions, and zoom meetings. I thank other mentors, teachers, and friends I had while at Auburn that helped me develop the knowledgebase and skillset necessary for this work, including Brian Counterman, Brian Folt, Dan Warner, Haruka Wada, Josh Hall, Rory Telemeco, Stephen Sefick, Todd Steury, and Wendy Hood.

I thank those who helped me in the field, including Corey Roelke, Danielle Rivera, Guillermo Álvarez, Jose Maldonado, Kathleen Currie, Kyle White, Miles Horne (especially Miles), Océane Da Cunha, Ryan Cook, Tori Herron, and many undergraduate students attending the UT El Paso field biology course at Indio Mountains Research Center.

I thank additional individuals who provided resources and assistance, including Dave Larson (Big Bend National Park), Gary Gruer (Mesilla Valley Bosque State Park), Jay Cole, Jerry

Johnson (Indio Mountains Research Station), Mike Roberts (treadmill), reviewers on two manuscripts, and an anonymous hiker in the Chisos Basin.

I thank the two people who helped me rediscover my childhood love for reptiles, Jack Sites and J.R. Wood. I thank Chad Hancock for introducing me to the mitochondrion. As an undergraduate these three helped me learn the science by providing opportunities for field work, lab work, writing, and presenting at international conferences. Their support, mentorship, and friendship helped me feel a sense of belonging in herpetology and physiology as a young scientist.

I thank the earth, all the forces that created it and its beauty, and the organisms I've had the privilege to study over the past few years.

I thank my family members, beginning with my parents (Matt and Angie Klabacka). They noticed my love for science as a young boy and supplied continual kindling in the form of books, dinosaur toys, field guides, and fossil-hunting trips. Since then, they have shown undaunting support in innumerable ways. Thank you, Mom and Dad. I also thank my brothers, Bryce and Ty, whose continual and undeserved admiration could always be felt from the other side of the country. Agradezco a mis otros papas, Adrián and Nancy Ochoa, who provided much support and were happy to listen to a practice talk or entertain a convoluted conversation.

I thank my two huiquitas, Lily and Rosie, for being beaming balls of joy and sharing my love for the natural world. Lastly, I thank Paulina Klabacka, my wife, field tech, animal caretaker, figure editor, and best friend. She's made this adventure one we will always cherish.

Table of Contents

| | |
|--|----|
| Abstract..... | 2 |
| Acknowledgments | 3 |
| List of Tables and Appendices..... | 9 |
| List of Figures..... | 11 |
| | |
| Chapter 1. Introduction to Questions and Study Systems | 14 |
| Questions | 14 |
| What Drives Lineage Diversification in a Biodiversity Hotspot? | 16 |
| What Are the Physiological Consequences of Asexual Reproduction? | 19 |
| How Do Gene Networks Shape Divergence in Life History Strategies? | 22 |
| Scientific Techniques and Implications | 25 |
| References..... | 25 |
| Figures..... | 35 |
| | |
| Chapter 2. Rivers of Indochina as Potential Drivers of Lineage Diversification | 38 |
| Abstract..... | 39 |
| Introduction..... | 40 |
| Methods | 43 |
| Results..... | 50 |
| Discussion | 54 |
| Conclusion | 61 |
| Acknowledgements..... | 62 |

| | |
|--|-----|
| References | 64 |
| Figures and Tables | 79 |
| Supplementary Materials | 102 |
| | |
| Chapter 3. Reduced Mitochondrial Respiration in Hybrid Asexual Lizards | 112 |
| Abstract | 113 |
| Introduction..... | 114 |
| Methods | 116 |
| Results..... | 120 |
| Discussion..... | 121 |
| Acknowledgements..... | 125 |
| References..... | 128 |
| Figures and Tables | 138 |
| Supplementary Materials | 142 |
| | |
| Chapter 4. Divergence in Genetic Variation and Expression in Molecular Networks that Underly Aging between Fast- and Slow-aging Garter Snakes | 159 |
| Abstract..... | 160 |
| Introduction..... | 161 |
| Methods | 164 |
| Results..... | 173 |
| Discussion..... | 179 |
| Acknowledgements..... | 194 |

| | |
|--|-----|
| References..... | 196 |
| Figures and Tables | 220 |
| Supplementary Materials | 228 |
| | |
| Chapter 5. Relevance, Recommendations, and Concluding Remarks..... | 253 |
| History and Current Understanding of Mitochondriology..... | 253 |
| Riverine Barriers as Potential Drivers of Lineage Diversification in Indochina | 257 |
| Reduced Mitochondrial Respiration in Hybrid Asexual..... | 261 |
| Divergence in Molecular Networks that Underly Aging..... | 269 |
| Dissertation Contributions and Implications | 276 |
| Concluding Remarks..... | 277 |
| References..... | 279 |
| Figures..... | 291 |

List of Tables and Appendices

Chapter 2

| | |
|--|----|
| Table 2.1 Genetic Summary Statistics | 85 |
| Table 2.2 Marginal-likelihood Estimations | 86 |
| Table 2.3 Lineages from River-8 Hypothesis | 87 |
| Appendix 2.A Subspecies Names | 88 |
| Appendix 2.B List of Samples | 89 |
| Appendix 2.C PCR Conditions | 93 |
| Appendix 2.D Genetic Summary Statistics for Mitochondrial Loci | 94 |
| Appendix 2.E Sister Pairs and <i>ND2</i> Phylogenetic Distances for <i>Draco</i> genus | 95 |
| Appendix 2.F Replicate Runs for Each BFD Hypothesis | 96 |
| Appendix 2.G Statistical Values Used for <i>Paralaudakia ND2</i> Molecular Clock | 98 |

Chapter 3

| | |
|--|-----|
| Table 3.1 Results from PhyloNetwork and Mixed-effects Linear Models | 141 |
| Table 3.S1 Sample Information | 149 |
| Table 3.S2 Individuals Used for Phylogenetic Inference | 151 |
| Table 3.S3 Summary Statistics for Species and Reproductive Modes | 152 |
| Table 3.S4 Within-group Linear Models | 153 |
| Table 3.S5 Differences in Variance Between Sexual and Hybrid Asexual Species | 154 |

Chapter 4

| | |
|--|-----|
| Table 4.1 Ecotype Divergence at Hallmarks of Aging | 227 |
|--|-----|

| | |
|---|-----|
| Table 4.S1 Statistics for Targeted Genes | 240 |
| Table 4.S2 Genes with Significant Interpopulation-Ecotype Divergence | 251 |
| Table 4.S3 Nonsynonymous SNPs with Significant Interpopulation-Ecotype Divergence | 252 |

List of Figures

Chapter 1

| | |
|---|----|
| Figure 1.1 Questions and Study Systems | 36 |
| Figure 1.1 The Costs of Asexual Reproduction..... | 37 |

Chapter 2

| | |
|---|-----|
| Figure 2.1 Map of Indochina with Sample Localities and Dewlap Illustrations | 80 |
| Figure 2.2 Mitochondrial and Nuclear Maximum Likelihood Estimated Phylogenies | 81 |
| Figure 2.3 Genus-wide <i>ND2</i> Analysis of Estimated Phylogenetic Distance | 82 |
| Figure 2.4 Experimental Design and Results from Bayes Factor Delimitation..... | 83 |
| Figure 2.5 Time-calibrated Maximum Clade Credibility Tree for River-8 Model..... | 84 |
| Figure 2.S1 Regression for Vicariance-based <i>Paralaudakia</i> Molecular Clock Calibration | 106 |
| Figure 2.S2 <i>ND2</i> Gene Tree for All Collected Samples | 107 |
| Figure 2.S3 Six-locus (12S, 16S, <i>ND2</i> , <i>BDNF</i> , <i>CMOS</i> , <i>PNN</i>) Concatenated Gene Tree | 108 |
| Figure 2.S4 12S Gene Tree..... | 108 |
| Figure 2.S5 16S Gene Tree..... | 109 |
| Figure 2.S6 <i>ND2</i> Gene Tree | 109 |
| Figure 2.S7 <i>BDNF</i> Gene Tree | 110 |
| Figure 2.S8 <i>CMOS</i> Gene Tree | 110 |
| Figure 2.S9 <i>PNN</i> Gene Tree | 111 |
| Figure 2.S10 <i>ND2</i> Gene Tree for <i>Draco</i> genus | 111 |

Chapter 3

| | |
|---|-----|
| Figure 3.1 Reduced Endurance and Mitochondrial Respiration in Hybrid Parthenogens | 139 |
| Figure 3.2 Positive Relationship Between Mitochondrial Respiration States and Endurance . | 140 |
| Figure 3.S1 Mitochondrial Gene Tree and Phylogenetic Network | 155 |
| Figure 3.S2 Effect of Hybrid Asexuality on Endurance and Mitochondrial Respiration..... | 156 |
| Figure 3.S3 Approximate Posterior Distribution for Standard Deviations of Repro Modes.... | 157 |
| Figure 3.S4 Difference in Coefficient of Variation Between Reproductive Modes..... | 158 |

Chapter 4

| | |
|---|-----|
| Figure 4.1 Map with Sample Geographic and Dataset Distribution | 221 |
| Figure 4.2 Differentially Expressed Genes from Meta Analysis | 222 |
| Figure 4.3 Networks of Enriched Pathways | 223 |
| Figure 4.4 Average Fst and dxy for “within” vs “between” ecotype population | 224 |
| Figure 4.5 Targeted Genes and Significant Loci | 225 |
| Figure 4.6 Summary of Differential Gene Expression and Sequence Variation Results | 226 |
| Figure 4.S1 Bioinformatics Pipeline– Gene Expression | 229 |
| Figure 4.S2 Bioinformatics Pipeline– Sequence Variation | 230 |
| Figure 4.S3 Gene Expression Upset Plots | 231 |
| Figure 4.S4 Cytoscape Networks of Enriched Pathways | 232 |
| Figure 4.S5 Insulin Signaling Gene Expression Enplot | 233 |
| Figure 4.S6 Oxidative Phosphorylation FA Heat Treatment Gene Expression Enplot | 234 |
| Figure 4.S7 Oxidative Phosphorylation SA Heat Treatment Gene Expression Enplot | 235 |
| Figure 4.S8 Nonsynonymous SNP Functional Impact Scores | 236 |
| Figure 4.S9 SNP Frequency Distributions | 237 |

| | |
|--|-----|
| Figure 4.S10 Comparison of Dxy between ecotypes for each category | 238 |
| Figure 4.S11 Electron Transport Chain Significant Subunits and Nonsynonymous SNPs..... | 239 |

Chapter 5

| | |
|---|-----|
| Figure 5.1 Publications Involving Mitochondria Through Time | 292 |
| Figure 5.2 SNP Paleo-river and Paleo-niche Hypotheses | 293 |
| Figure 5.3 Ideal Experimental Sampling of Parthenoforms and Potential Results | 294 |
| Figure 5.4 Potential Complex Activity Results | 295 |

Chapter 1: Introduction to evolutionary questions and study systems

1.1 Questions

“There is grandeur in this view of life... from so simple a beginning endless forms most beautiful and most wonderful have been, and are being evolved.” (Darwin 1859)

The solidification of evolutionary biology as a scientific theory provided a foundation for understanding the source of life’s variation, an objective that has since become a central aim in biology. While scientific acceptance of evolution has answered some biological questions, it continues to create more questions than it has answered. Scientists today still seek to understand the forces that are shaping these “endless forms most beautiful”, and often these forces are complex. For example, Darwin, the same individual who referred to an evolutionary mindset as a view with “grandeur”, was so troubled by understanding how evolution could shape all life that he once wrote, “The sight of a feather in a peacock’s tail, whenever I gaze at it, it makes me sick!” (Burkhardt et al. 1993).

The challenge of answering difficult questions has been mitigated by a concerted effort focused on several highly-accessible species. Studies of these organisms, deemed “model organisms”, make up the vast majority of all biological studies. Examining the genetics, behavior, physiology, and evolution of taxa such as *Drosophila* flies, *Caenorhabditis* worms, *Saccharomyces* yeast, and *Mus* mice (just to name a few) has resulted in major breakthroughs regarding our understanding of the natural world. However, if a primary goal of evolutionary biology is to understand the forces shaping “endless forms”, then findings derived from a

handful of organisms will necessarily fall short of this goal. Several major questions examined by evolutionary biologists that cannot be answered without a broader taxonomic scope include (1) How are new species formed? (2) What factors influence trait evolution? And (3) How do changes in genetics and environment shape phenotypes?

Squamate reptiles (snakes and lizards) provide a wealth of natural systems teeming with information that can help us answer these and other questions in evolutionary biology. With approximately 11,500 known species (Uetz et al. 2022), squamates are the most species-rich group of tetrapods on earth. These species occupy a diverse array of habitats and niches, including those of low and high elevation, humid and xeric climates, cold and warm temperatures, and multiple levels in their respective food webs from primary consumers to apex predators. They can be found in habitats viewed as mild and those considered extreme. Multiple, divergent clades rich with species with diverse morphology and life history make squamates a great system for understanding how natural processes shape life on Earth. For example, squamates contain multiple, independent transitions from viviparity-oviparity, gonochorism-parthenogenesis, carnivory-herbivory, and limbs-no limbs, among other traits, enabling the study of the general mechanisms that underlie these evolutionary transitions.

As a doctoral student at Auburn University interested in evolutionary mechanisms that shape biodiversity, I focused my research on three questions:

- 1) What drives lineage diversification?
- 2) What are the physiological consequences of asexual reproduction?
- 3) How does variation in gene networks shape divergence in life history strategies?

These questions, along with their respective squamate reptile study systems used for this dissertation work, are shown in Fig. 1.1. While every chapter is set within a unique field in evolutionary biology, each contains a component linked to the evolution of the mitochondrion and its contribution to the observed patterns. In the final chapter (Chapter 5), I discuss the history of mitochondriology, its connection with evolutionary biology, and how this dissertation contributes to our understanding of the influence mitochondria have on shaping biodiversity. In the remainder of this introductory chapter, I will provide background information for each question, study, and squamate system.

1.2 What Drives Lineage Diversification in a Biodiversity Hotspot?

Evolution implies shared ancestry, and provokes two major questions: (1) what forces create the diverse traits in organisms, and (2) what forces make one species become two? Natural selection provided an overarching answer for the first question, described by contemporary scientists Charles Darwin and Alfred Wallace (Darwin and Wallace 1958). Reproductive isolation of populations and subsequent independent evolution provided the answer to the second question. However, the agents responsible for the restriction of gene flow can be complex and are dependent on whether the mode of speciation is allopatric (or a variation of allopatry, such as peripatry) or sympatric. Allopatric speciation is intrinsically placed within the field of biogeography— another field in biology originally described by Alfred Wallace (Wallace 1855). Biogeography posits that the evolutionary history of organisms is connected to their geographic distributions. While the linkage of speciation and geography is widely accepted, much remains to be explained regarding the specific biogeographic mechanisms that are directly responsible for

restricting gene flow. This is especially true in Indochina- the continental outreach of Southeast Asia that has engendered significantly more species than any of the other highly diverse regions in the area (De Bruyn et al. 2014).

Using Indochina as a natural laboratory that has formed a remarkable number of lineages *in situ*, we can apply tools in molecular evolution to test hypotheses of lineage divergence in this megadiverse region. One hypothesis that may explain the source for some of the biodiversity in Indochina is the riverine barrier hypothesis (Wallace 1852), an idea based in biogeography that predates the description of evolution by natural selection. The riverine barrier hypothesis suggests that rivers dividing a population create a barrier to gene flow. With time, the population disconnectivity results in the independent evolution of populations on either side of the river resulting in two new species. Alfred Wallace originally described this idea while traversing the rivers of the Amazon Basin and observing the distribution of monkey species (Wallace 1852).

While riverine barriers may not play critical roles in reducing gene flow for all species (Naka and Pil 2020), empirical evidence of rivers as barriers to gene flow exists for many. A database search on Web of Science using the search criteria “Riverine Barrier Hypothesis” recovered 32 studies with results that implicate rivers as barriers to gene flow (out of 49 total studies involving the hypothesis; range of studies 1994–2022). Because the effect of rivers on biodiversification across Indochina as a whole had not been examined, I assessed the evolutionary history of a widespread Indochinese species complex (*Draco maculatus*) and statistically tested whether historical taxonomy, mitochondrial phylogeny, or riverine barriers best explained the patterns in sequence data I collected (Chapter 2). My results indicate that riverine barriers best explained the

patterns in the data, and I suggest that the high levels of in situ diversification in this hotspot may be due in-part to riverine vicariance followed by lineage divergence.

Study System

Because of their remarkable species richness, diverse morphology, and ability to adapt to many (including extreme) niches, squamate reptiles have become great model systems for understanding lineage diversification in the context of biogeography (Camargo et al. 2010; Marshall et al. 2018). Flying lizards (genus *Draco*) are among the squamates that have been useful in understanding how the geological history of Southeast Asia has contributed to the high levels of biodiversity within the region (Honda et al. 1999; McGuire and Kiew 2001; McGuire et al. 2007; Reilly et al. 2021). The expansive range of *Draco maculatus* extending across Indochina, along with the presence of morphological polymorphisms, led us to use this putative species complex to test if the geography of Indochina shapes the lineage diversity of its terrestrial inhabitants.

Soon after its description by John Edward Gray in 1845, *Draco maculatus* was recognized by scientists for its variation in dewlap color (Gray 1845; Musters 1983). At various points throughout its taxonomic history, its subspecies (*D. maculatus divergens*, *D. maculatus haseii*, *D. maculatus maculatus*, and *D. maculatus whiteheadi*) were considered independent species (Boettger 1893; Boulenger 1899; Taylor 1934) only to be re-categorized within the *D. maculatus* as subspecies. I followed the suggestion of McGuire and Kiew (2001) by conducting a phylogeographic study and implementing tests for several hypotheses that may explain the potentially discrete morphological variation in this species complex.

1.3 What Are the Physiological Consequences of Asexual Reproduction?

In 1930 renowned evolutionary biologist Sir Ronald A. Fisher wrote that asexual reproduction could not “be ascribed with certainty to any known group” (Fisher 1930). Two years later, Carl and Laura Hubbs described the first all-female asexual animal species, the Amazon Molly (*Poecilia formosa*), whose common name originates in the mythological tribe of warrior women (Hubbs and Hubbs 1932). Three years after that, British herpetologist Malcom Arthur Smith collected over 100 samples of fox geckos (*Hemidactylus garnotii*) from Southeast Asia without finding a single male (Smith 1935), marking the dawn of the first described parthenogenetic¹ organism. Since then, numerous asexual taxa have been described, and our understanding of reproduction without sex has gone from “virtually nonexistent” to “fairly common” in the animal kingdom (Suomalainen 1962). However, within vertebrates, parthenogenesis (Greek “partheno” : virgin, “genesis” : origin) remains a rare phenomenon (Dawley and Bogart 1989). While several types of asexual reproduction exist throughout vertebrates, transgenerational asexual reproduction devoid of male contribution (parthenogenesis) has only been described in several instances of squamate reptiles. Essentially all instances of these parthenoforms are a result of hybridization (but see Sinclair et al. 2010), and a large degree of heterozygosity is maintained due to “pseudo-recombination” after an endoreplication event before Meiosis I (Cuellar 1971; Lutes et al. 2010). For this reason, these organisms are frequently referred to as “biotypes” or “parthenoforms” rather than species. However, because many asexual organisms (including the

¹ “Parthenogenetic” organisms reproduce without sperm as their primary mode of reproduction, which separates them from organisms that use other modes of reproduction requiring sperm involvement (e.g., “gynogenetic”, “hybridogenetic”, and “kleptogenetic”) and those that typically reproduce sexually yet have the ability to reproduce asexually (e.g., “tychoparthenogenetic”).

focal group of this dissertation) are morphologically diagnosable and evolving along independent lineages, I will refer to them as species for the sake of simplicity.

John Maynard-Smith first coined the “two-fold cost of sex”, which describes the sexual disadvantages of (1) wasting half of one’s reproductive effort on producing males [cost of males] and (2) only passing on half of one’s genetic material [cost of meiosis] (Maynard Smith 1958; Williams 1975). The fact that the process of fertilization is the exact opposite of cell division indicates that an asexually-reproducing organism will have a two-fold reproductive advantage over a sexually-reproducing organism. In other words, if two female individuals (one sexual and one asexual) have the same number of offspring with the same fitness level, the asexual female will have passed on double the amount of genetic material compared to the sexual female (Fig. 1.2). This cost is reduced, or even eliminated, when fitness differs between the offspring of the reproductive modes (sexual vs asexual), thus providing a potential avenue for explaining the scarcity of asexual vertebrates.

Organismal and intracellular physiology constitute a collection of functional phenotypes critical for organismal survival and reproduction, making the factors contributing to variation in physiological phenotypes worthy targets of selection. Several past studies found reduced performance in aerobic activities for asexual species of whiptail lizards (Cullum 1997), dace fish (Mee et al. 2011), and mole salamanders (Denton et al. 2017). I aimed to elucidate the source of this reduced performance by measuring aerobic performance at the organismal level and intracellular traits in whiptail lizards (genus *Aspidoscelis*): endurance running and mitochondrial respiration (Chapter 3). My results in endurance running show reduced endurance capacity in the

asexual hybrid species, reflecting those found previously (Cullum 1997). These are refined by my results in isolated mitochondrial respiration, which show a reduced rate of oxygen consumption by the asexual hybrid species. Additionally, I found that endurance capacity and mitochondrial respiration are positively associated with one another. At the end of Chapter 3 and within the conclusion (Chapter 5) I discuss potential underpinnings of these findings.

Study System

Whiptail lizards (genus *Aspidoscelis*; “whiptails”) have become the model system for understanding the biology of parthenogenesis. With one-third of this genus reproduce via parthenogenesis, this system allows for examination of (1) the cytological mechanisms involved in parthenogenesis (Lutes et al. 2010), (2) the evolutionary conditions that result in parthenogenesis (Moritz et al. 1992; Barley et al. 2019, 2021, 2022a), (3) the genomic consequences of parthenogenesis (Fujita et al. 2020), and (4) the physiological consequences of parthenogenesis (Cullum 1997; Mata-Silva et al. 2008). *Aspidoscelis* puzzled taxonomists for over a century, but advances in technology for estimating evolutionary history have aided significantly in resolving relationships among the species in this genus (Reeder et al. 2002; Barley et al. 2022b), and genomic resources for this genus are being actively developed by multiple laboratories.

Within the same year that the first parthenogenetic lizard was officially described (Darevsky 1958), Minton (1958) and Tinkle (1959) reported independently that they had never seen a male *Aspidoscelis tessellatus*. Within a few days, reviews of specimens and field notes found that there were no males in *A. exsanguis*, *A. neomexicanus*, *A. velox*, *A. flagellicaudus*, *A. sonorae*, and *A.*

uniparens (Lowe 1993). The first publications confirming parthenogenesis in *Aspidoscelis* came just a few years later (Maslin 1962, 1967, 1971), and others described the mechanism of parthenogenetic species formation by hybridization of divergent gonochoristic species (Lowe and Wright 1966*a*, 1966*b*; Neaves and Gerald 1968, 1969; Neaves 1969; Cuellar 1974; Cuellar and McKinney 1976).

The endurance capacity of whiptails in the context of evolutionary ecology has been of interest to biologists since the 1980s. The connection of whiptail aerobic performance with observed patterns in behavior and physiology include notes on their high energy expenditure (Anderson and Karasov 1981), large home range sizes (Garland Jr. 1993), and wide foraging strategy (Garland Jr. 1993). However, it wasn't until 1997 that a difference in endurance capacity was seen between gonochoristic and parthenogenetic species (Cullum 1997). With the historical precedence of running whiptails on treadmills (Garland 1994; Cullum 1998, 2000), testing for a connection between aerobic performance and mitochondrial function was an appropriate next step to understand whether mitochondria may be underpinning the observed reduction in performance.

1.4 How Do Gene Networks Shape Divergence in Life History Strategies?

Uncovering the underpinnings responsible for variation in traits such as lifespan, growth rate, fecundity, and senescence across taxa will help us understand the targets of selection that cause divergence in these traits. Senescence (hereafter “aging”), defined as the process of decline in the probability of survival and/or reproduction with age, is a widespread phenomenon nested within the life history strategy of many organisms. Despite this phenomenon being relevant on a broad

scale taxonomically, the vast majority of studies on aging are focused on the few model organisms mentioned earlier in section 1.1.

Identifying the genetic mechanisms that contribute to aging is a necessary step in understanding the targets of natural selection that facilitate divergence in life history strategies. Examination of genetic dissimilarity in natural populations with variation in aging provides a window to genes and gene networks involved with aging. In addition to most studies of aging being focused on the heretofore mentioned model organisms, these studies were conducted in controlled laboratory settings with inbred lines where inference was drawn based on experimental manipulation. While these studies are highly informative and benefit from exclusion of natural noise experienced by populations outside of the lab, they lack the factors present in native environments that put the “nature” in “natural selection.” Several studies have examined evolutionary alterations in aging of other taxa, but these have focused primarily on differences between species, genera, families, or classes that have diverged in numerous other traits in addition to those related to aging (Remot et al. 2021; Opazo et al. 2022; Reinke et al. 2022). An accurate understanding of the genes behind aging requires examination of sequence variation in populations with variation in aging.

Two divergent ecotypes of terrestrial garter snakes (*Thamnophis elegans*) are present in populations around Eagle Lake in Northeastern California (USA). These ecotypes are at different positions on the “pace of life” spectrum, with one ecotype having a higher rate of aging, shorter lifespan, and higher metabolic rate compared to the other. Targeting gene networks associated with aging in model organisms (e.g., genes of the insulin signaling network and genes of the electron transport system in the mitochondrion), I sought to identify patterns in gene expression

and sequence variation associated with aging. Between ecotypes, I identified networks with enriched expression and genes with significant sequence variation. I also examined the functional implications of sequence variation (e.g., effect of mutation on peptide sequence and protein function). My results reflect many of those found in model systems, with variation in the same gene networks (and some of the same genes) associated with variation in aging, yet I also found highly divergent responses to heat in gene expression between the ecotypes and identify novel genes of interest with functional implications.

Study System

The garter snakes of Eagle Lake have been studied in-detail over the past half century. Stevan Arnold used populations around Eagle Lake to compare their ecology with those along the California coastline (Arnold 1977, 1981, 1992; Arnold and Wassersug 1978; Kephart and Arnold 1982). After twenty years of researchers collecting data at the site, Anne Bronikowski, a new graduate student at that time, found that the variation between populations around Eagle Lake was largely partitioned between individual ecotypes with divergent life history strategies (Bronikowski and Arnold 1999). Twenty more years of research of the ecotypes revealed that the lakeshore populations grow larger (Bronikowski and Arnold 1999), produce larger litter sizes (Sparkman et al. 2007), possess a higher metabolic rate (Gangloff et al. 2015), experience more oxidative stress (Schwartz and Bronikowski 2013), and die younger (Bronikowski and Vleck 2010) than the higher-elevation meadow populations. The ecotypes can also be differentiated based on general pattern and coloration, with the lakeshore populations being more gray and checkered with a pale dorsal stripe, whereas the meadow populations are more solid black with a yellow dorsal stripe (Manier and Arnold 2005).

Examination of the mitochondrial genome by Tonia Schwartz found that a nonsynonymous mutation in CYTB, a core subunit of Cytochrome C Reductase, segregated between the ecotypes in her sample set (Schwartz et al. 2015). Examining this further revealed that the individuals with the derived CYTB SNP were almost entirely from meadow populations, and these made up a monophyletic group within our estimated phylogeny (Gangloff et al. 2020). With multiple lines of evidence supporting the differences in ecological and physiological traits between these ecotypes, I set out to examine the potential genomic underpinnings of this divergence.

1.5 Scientific Techniques and Implications

Using the previously described squamate systems within the context of their respective research questions, I apply techniques in evolutionary biology to better understand the processes driving observed patterns. Techniques applied in chapters two, three, and four include molecular phylogenetics, Bayesian hypothesis testing, linear statistical modeling, high-resolution mitochondrial respirometry, phylogenetic network statistical modeling, processing of high-throughput targeted sequencing, and population genomics analyses. By examining data using these techniques, I shed light on potential (1) processes driving lineage diversification in a biodiversity hotspot, (2) physiological consequences of asexual reproduction in vertebrates, and (3) genomic underpinnings of variation in aging. As a conclusion, in chapter five, I discuss some of the contributions of this dissertation to science and society, including an overview of historical perspectives on the connection between the mitochondrion and general patterns in evolution.

1.6 References

- Anderson, R. A., and W. H. Karasov. 1981. Contrasts in energy intake and expenditure in sit-and-wait and widely foraging lizards. *Oecologia* 49:67–72.
- Arnold, S. J. 1977. Polymorphism and geographic variation in the feeding behavior of the garter snake *Thamnophis elegans*. *Science* 197:676–678.
- . 1981. Behavioral Variation in Natural Populations. II. The Inheritance of a Feeding Response in Crosses Between Geographic Races of the Garter Snake, *Thamnophis elegans*. *Evolution* 35:510.
- . 1992. Behavioural variation in natural populations. VI. Prey responses by two species of garter snakes in three regions of sympatry. *Animal Behaviour* 44:705–719.
- Arnold, S. J., and R. J. Wassersug. 1978. Differential Predation on Metamorphic Anurans by Garter Snakes (*Thamnophis*): Social Behavior as a Possible Defense. *Ecology* 59:1014–1022.
- Barley, A. J., J. E. Cordes, J. M. Walker, and R. C. Thomson. 2022a. Genetic diversity and the origins of parthenogenesis in the teiid lizard *Aspidoscelis laredoensis*. *Molecular Ecology* 31:266–278.
- Barley, A. J., A. Nieto-Montes de Oca, N. L. Manríquez-Morán, and R. C. Thomson. 2022b. The evolutionary network of whiptail lizards reveals predictable outcomes of hybridization. *Science (New York, N.Y.)* 377:773–777.
- Barley, A. J., A. Nieto-Montes de Oca, T. W. Reeder, N. L. Manríquez-Morán, J. C. Arenas Monroy, O. Hernández-Gallegos, and R. C. Thomson. 2019. Complex patterns of hybridization and introgression across evolutionary timescales in Mexican whiptail

- lizards (*Aspidoscelis*). *Molecular Phylogenetics and Evolution* 132:284–295.
- Barley, A. J., T. W. Reeder, A. N. M. de Oca, C. J. Cole, and R. C. Thomson. 2021. A New Diploid Parthenogenetic Whiptail Lizard from Sonora, Mexico, Is the “Missing Link” in the Evolutionary Transition to Polyploidy. <https://doi.org/10.1086/715056> 198:295–309.
- Boettger, O. 1893. Ein neuer Drache (*Draco*) aus Siam. *Zool. Anz.* 16:429–430.
- Boulenger, G. A. 1899. On the reptiles, batrachians (and fishes) collected by the late Mr. John Whitehead in the interior of Hainan. *Proc. Zoolog. Soc. London* 956–959.
- Bronikowski, A. M., and S. J. Arnold. 1999. The evolutionary ecology of life history variation in the garter snake *Thamnophis elegans*. *Ecology* 80:2314–2325.
- Bronikowski, A., and D. Vleck. 2010. Metabolism, Body Size and Life Span: A Case Study in Evolutionarily Divergent Populations of the Garter Snake (*Thamnophis elegans*). *Integrative and Comparative Biology* 50:880–887.
- Burkhardt, F., D. M. Porter, J. Browne, and M. Richmond, eds. 1993. *The Correspondence of Charles Darwin, Volume 8, 1860. History and Philosophy of the Life Sciences.* Cambridge University Press, Cambridge.
- Camargo, A., B. Sinervo, and J. W. Sites. 2010. Lizards as model organisms for linking phylogeographic and speciation studies. *Molecular ecology* 19:3250–3270.
- Cuellar, O. 1971. Reproduction and the mechanism of meiotic restitution in the parthenogenetic lizard *Cnemidophorus uniparens*. *Journal of Morphology* 133:139–165.
- . 1974. On the Origin of Parthenogenesis in Vertebrates: The Cytogenetic Factors. <https://doi.org/10.1086/282940> 108:625–648.
- Cuellar, O., and C. O. McKinney. 1976. Natural hybridization between parthenogenetic and bisexual lizards: detection of uniparental source of skin grafting. *The Journal of*

- Experimental Zoology 196:341–350.
- Cullum, A. J. 1997. Comparisons of physiological performance in sexual and asexual whiptail lizards (genus *Cnemidophorus*): implications for the role of heterozygosity. *The American naturalist* 150:24–47.
- Cullum, A. J. 1998. Sexual dimorphism in physiological performance of whiptail lizards (Genus *Cnemidophorus*). *Physiological Zoology* 71:541–552.
- Cullum, A. J. 2000. *Phenotypic variability of physiological traits in populations of sexual and asexual whiptail lizards (genus Cnemidophorus)*. *Evolutionary Ecology Research* (Vol. 2).
- Darevsky, I. 1958. Natural parthenogenesis in certain subspecies of rock lizard, *Lacerta asaxicola*. Eversmann. *Doklady Akad. Nauk SSR, Biol. Sci. Sect* 122:877–879.
- Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection, Or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.
- Darwin, C., and A. Wallace. 1958. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Journal of the Proceedings of the Linnean Society of London* 45–50.
- Dawley, R., and J. Bogart. 1989. *Evolution and Ecology of Unisexual Vertebrates*. New York State Museum, Albany.
- De Bruyn, M., B. Stelbrink, R. J. Morley, R. Hall, G. R. Carvalho, C. H. Cannon, G. Van Den Bergh, et al. 2014. Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity. *Systematic Biology* 63:879–901.
- Denton, R. D., K. R. Greenwald, and H. L. Gibbs. 2017. Locomotor endurance predicts differences in realized dispersal between sympatric sexual and unisexual salamanders.

- Functional Ecology 31:915–926.
- Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. Oxford at the Clarendon Press.
- Fujita, M. K., S. Singhal, T. O. Brunes, and J. A. Maldonado. 2020. Evolutionary Dynamics and Consequences of Parthenogenesis in Vertebrates. *Annual Review of Ecology, Evolution, and Systematics* 51.
- Gangloff, E. J., T. S. Schwartz, R. Klabacka, N. Huebschman, A. Y. Liu, and A. M. Bronikowski. 2020. Mitochondria as central characters in a complex narrative: Linking genomics, energetics, pace-of-life, and aging in natural populations of garter snakes. *Experimental Gerontology* 137.
- Gangloff, E. J., D. Vleck, and A. M. Bronikowski. 2015. Developmental and Immediate Thermal Environments Shape Energetic Trade-Offs, Growth Efficiency, and Metabolic Rate in Divergent Life-History Ecotypes of the Garter Snake *Thamnophis elegans*. *Physiological and Biochemical Zoology* 88:550–563.
- Garland Jr., T. 1993. Locomotor performance and activity metabolism of *Cnemidophorus tigris* in relation to natural behaviors. Pages 163–206 in J. W. Wright and L. J. Vitt, eds. *Biology of Whiptail Lizards (genus Cnemidophorus)*. The Oklahoma Museum of Natural History, Norman, OK.
- Garland, T. 1994. Phylogenetic analyses of lizard endurance capacity in relation to body size and body temperature. Pages 237–259 in *Lizard Ecology: Historical and Experimental Perspectives*.
- Gray, J. E. 1845. *Catalogue of the Specimens of lizards in the collection of the British museum*. Catalogue of the Specimens of lizards in the collection of the British museum. Printed by order of the Trustees, London,.

- Honda, M., H. Ota, M. Kobayashi, J. Nabhitabhata, H.-S. Yong, and T. Hikida. 1999. Phylogenetic Relationships of the Flying Lizards, Genus *Draco* (Reptilia, Agamidae). *Zoological Science* 16:535–549.
- Hubbs, C. L., and L. C. Hubbs. 1932. APPARENT PARTHENOGENESIS IN NATURE, IN A FORM OF FISH OF HYBRID ORIGIN. *Science* 76:628–630.
- Kephart, D. G., and S. J. Arnold. 1982. Garter Snake Diets in a Fluctuating Environment: A Seven-Year Study 63:1232–1236.
- Lowe, C. H., and J. W. Wright. 1966a. Evolution of parthenogenetic species of *Cnemidophorus* (whiptail lizards) in western North America. *Journal of the Arizona Academy of Sciences* 4:81–87.
- Lowe, C. H., and J. W. Wright. 1966b. Chromosomes and karyotypes of *Cnemidophorine* teiid lizards. *Mammalina Chromosomes Newsletter* 22:199–200.
- Lowe, C. harle. H. 1993. Introduction to Biology of *Cnemidophorus*. Page 5 in J. W. Wright and L. J. Vitt, eds. *Biology of Whiptail Lizards (genus Cnemidophorus)*. The Oklahoma Museum of Natural History, Norman, OK.
- Lutes, A. a, W. B. Neaves, D. P. Baumann, W. Wiegraebe, and P. Baumann. 2010. Sister chromosome pairing maintains heterozygosity in parthenogenetic lizards. *Nature* 464:283–286.
- Manier, M. K., and S. J. Arnold. 2005. Population genetic analysis identifies source–sink dynamics for two sympatric garter snake species (*Thamnophis elegans* and *Thamnophis sirtalis*). *Molecular Ecology* 14:3965–3976.
- Marshall, J. C., E. Bastiaans, A. Caccone, A. Camargo, M. Morando, M. L. Niemiller, M. Pabijan, et al. 2018. Mechanisms of speciation in reptiles and amphibians: a synopsis.

- Maslin, P. T. 1962. All-female species of the lizard genus *Cnemidophorus*, Teiidae. *Science* 135:212–213.
- Maslin, T. P. 1967. Skin grafting in the bisexual teiid lizard *Cnemidophorus sexlineatus* and in the unisexual *C. tessellatus*. *Journal of Experimental Zoology* 166:137–149.
- . 1971. Conclusive evidence of parthenogenesis in three species of *Cnemidophorus* (Teiidae). *Copeia* 156–158.
- Mata-Silva, V., C. R. Bursey, and J. D. Johnson. 2008. Gut parasites of two syntopic species of whiptail lizards, *Aspidoscelis marmorata* and *Aspidoscelis tessellata* from the Northern Chihuahuan Desert. *Sociedad Herpetológica Mexicana* 16:1–4.
- Maynard Smith, J. 1958. *The Theory of Evolution*. Cambridge University Press.
- McGuire, J. a., R. M. Brown, Mumpuni, A. Riyanto, and N. Andayani. 2007. the Flying Lizards of the *Draco Lineatus* Group (Squamata: Iguania: Agamidae): a Taxonomic Revision With Descriptions of Two New Species. *Herpetological Monographs* 21:179.
- Mcguire, J. a, and B. H. Kiew. 2001. Phylogenetic systematics of Southeast Asian flying lizards (Iguania : Agamidae : *Draco*) as inferred from mitochondrial DNA sequence data. *Biological Journal of the Linnean Society* 72:203–229.
- Mee, J. A., C. J. Brauner, and E. B. Taylor. 2011. Repeat Swimming Performance and Its Implications for Inferring the Relative Fitness of Asexual Hybrid Dace (Pisces: *Phoxinus*) and Their Sexually Reproducing Parental Species. *Physiological and Biochemical Zoology* 84:306–315.
- Minton, S. A. 1958. Observations on Amphibians and Reptiles of the Big Bend Region of Texas. *The Southwestern Naturalist* 3:28.
- Moritz, C. C., J. W. Wright, and W. M. Brown. 1992. Mitochondrial DNA Analyses and the

- Origin and Relative Age of Parthenogenetic *Cnemidophorus* : Phylogenetic Constraints on Hybrid Origins Author (s): C . Moritz , J . W . Wright and W . M . Brown Reviewed work (s): Published by : Society for the Study 46:184–192.
- Musters, C. J. . M. 1983. Taxonomy of the genus *Draco* (Agamidae, Lacertilia, Reptilia). *Zoologische Verhandelingen* 199.
- Naka, L. N., and M. W. Pil. 2020. Moving beyond the riverine barrier vicariant paradigm. *Molecular Ecology* 29:2129–2132.
- Neaves, W. B. 1969. Adenosine deaminase phenotypes among sexual and parthenogenetic lizards in the genus *Cnemidophorus* (teiidae). *Journal of Experimental Zoology* 171:175–183.
- Neaves, W. B., and P. S. Gerald. 1968. Lactate Dehydrogenase Isozymes in Parthenogenetic Teiid Lizards (*Cnemidophorus*). *Science* 160:1004–1005.
- . 1969. Gene Dosage at the Lactate Dehydrogenase b Locus in Triploid and Diploid Teiid Lizards. *Science* 164:557–559.
- Opazo, J. C., M. W. Vandewege, F. G. Hoffmann, K. Zavala, C. Meléndez, C. Luchsinger, V. A. Cavieres, et al. 2022. How many sirtuin genes are out there? evolution of sirtuin genes in vertebrates with a description of a new family member. *bioRxiv* 2020.07.17.209510.
- Reeder, T. W., C. J. Cole, and H. C. Dessauer. 2002. Phylogenetic relationships of whiptail lizards of the genus *Cnemidophorus* (Squamata: Teiidae): a test of monophyly, reevaluation of karyotypic evolution, and review of hybrid origins. *American Museum Novitates* 3365:1–61.
- Reilly, S. B., A. L. Stubbs, E. Arida, B. R. Karin, U. Arifin, H. Kaiser, K. Bi, et al. 2021. *Phylogenomic Analysis Reveals Dispersal-Driven Speciation and Divergence with Gene*

- Flow in Lesser Sunda Flying Lizards (Genus Draco). *Systematic Biology* 71:221–241.
- Reinke, B. A., H. Cayuela, F. J. Janzen, J. F. Lemaître, J. M. Gaillard, A. M. Lawing, J. B. Iverson, et al. 2022. Diverse aging rates in ectothermic tetrapods provide insights for the evolution of aging and longevity. *Science (New York, N.Y.)* 376:1459–1466.
- Remot, F., V. Ronget, H. Froy, B. Rey, J. M. Gaillard, D. H. Nussey, and J. F. Lemaitre. 2021. Decline in telomere length with increasing age across nonhuman vertebrates: A meta-analysis. *Molecular Ecology*.
- Schwartz, T. S., Z. W. Arendsee, and A. M. Bronikowski. 2015. Mitochondrial divergence between slow- and fast-aging garter snakes. *Experimental Gerontology* 71:135–146.
- Schwartz, T. S., and A. M. Bronikowski. 2013. Dissecting molecular stress networks: Identifying nodes of divergence between life-history phenotypes. *Molecular Ecology* 22:739–756.
- Sinclair, E. A., J. B. Pramuk, R. L. Bezy, K. A. Crandall, and J. W. Sites. 2010. Dna evidence for nonhybrid origins of parthenogenesis in natural populations of vertebrates. *Evolution* 64:1346–1357.
- Smith, M. A. 1935. *Reptilia and Amphibia: Sauria. The fauna of British India including Ceylon and Burma (2nd ed.)*. Taylor & Francis, London.
- Sparkman, A. M., S. J. Arnold, and A. M. Bronikowski. 2007. An empirical test of evolutionary theories for reproductive senescence and reproductive effort in the garter snake *Thamnophis elegans*. *Proceedings of the Royal Society* 274:943–950.
- Suomalainen, E. 1962. *Significance of parthenogenesis in the evolution of insects*.
- Taylor, E. H. 1934. Zoological results of othe third De Schauensee Siamese Expedition, part III: Amphibians and reptiles. *Proceedings of the Academy of Natural Sciences of Philadelphia* 86:281–310.

Tinkle, D. W. 1959. Observations on the Lizards *Cnemidophorus Tigris*, *Cnemidophorus Tessellatus* and *Crotaphytus Wislizeni*. *The Southwestern Naturalist* 4:195.

Uetz, P., P. Freed, R. Aguilar, and J. Hosek. 2022. *The Reptile Database*.

Wallace, A. 1852. On the monkeys of the Amazon. *Proceedings of the Zoological Society, London* 20:107–110.

Wallace, A. R. 1855. On the law which has regulated the introduction of new species. *Annals and magazine of natural history* 16:184–196.

Williams, C. G. 1975. *Sex and Evolution*. Princeton University Press, Princeton.

Chapter 1: Figures

Figure 1.1 Introduction to Questions and Study Systems of This Dissertation

Figure 1.2 The Costs of Sexual Reproduction

Ch 1: Figures

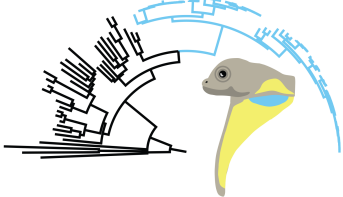
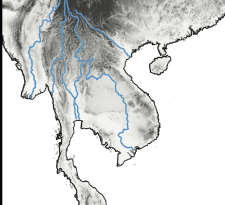

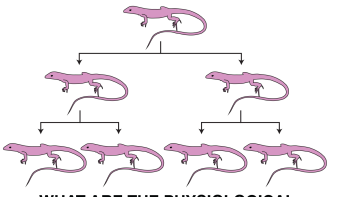
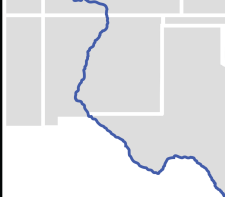




| Question | Location | Squamate System |
|--|--|--|
|  <p>WHAT DRIVES LINEAGE DIVERSIFICATION?</p> |  <p>Indochina</p> |  <p><i>Draco maculatus</i></p> |
|  <p>WHAT ARE THE PHYSIOLOGICAL CONSEQUENCES OF ASEX REPRO?</p> |  <p>Rio Grande</p> |  <p><i>Aspidoscelis</i></p> |
|  <p>pace of life</p> <p>HOW DO GENE NETWORKS SHAPE DIVERGENCE IN LIFE HISTORY STRATEGIES?</p> |  <p>Eagle Lake</p> |  <p><i>Thamnophis elegans</i></p> |

Figure 1.1: **Introduction to Questions and Study Systems of This Dissertation**

Focal questions of dissertation, along with the geographic locations and images of the focal species. Photo credit for *Draco maculatus* image: Matthieu Berroneau (www.matthieu-berroneau.fr). Photo credit for *Aspidoscelis* image: L.

Miles Horne.

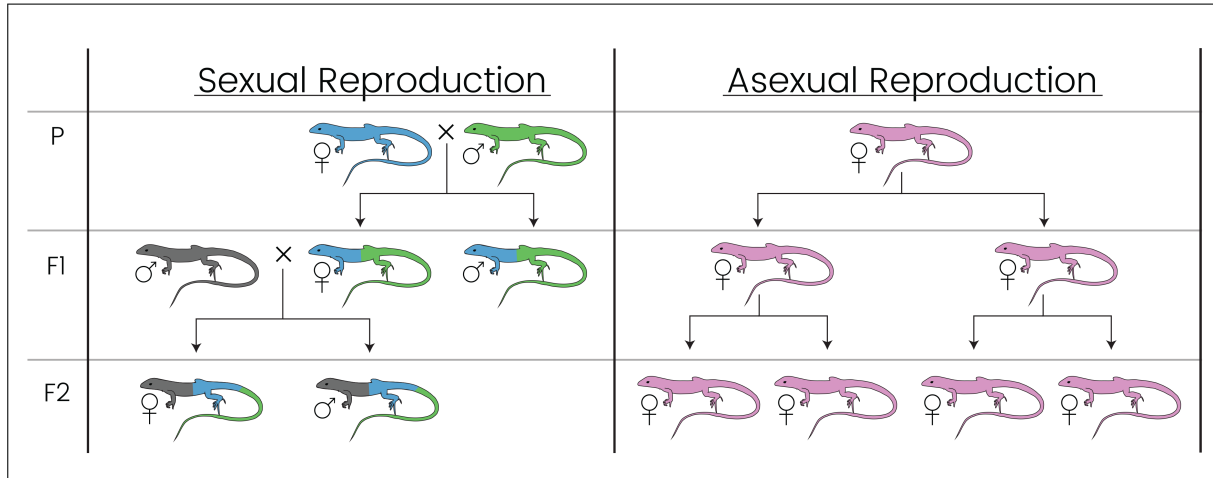


Figure 1.2: **The Costs of Sexual Reproduction**

The costs of sex are depicted here in number and genetic composition of offspring. In this example we will assume a fixed clutch size of two, a fixed sex ratio of 1:1, and no differential fitness between reproductive modes. The colors are indicative of genetic composition. Two females are shown in the P generation: one individual of a sexually reproducing species (far left; blue) and one individual of an asexually reproducing species (right; pink). Each of these individuals produce F1 offspring. The sexually reproducing female only passes on 50 percent of her genetic material to each F1 progeny, whereas the asexually reproducing individual passes on 100 percent of her genetic material. Further, the sexually-reproducing species has to waste half of her reproductive investment in males, whereas the asexually reproducing individual creates only female offspring.

Chapter 2: Rivers of Indochina as potential drivers of lineage diversification in the spotted flying lizard (*Draco maculatus*) species complex

Published in Molecular Phylogenetics and Evolution (Klabacka et al., 2020)

Randy L. Klabacka ^{a,b}, Perry L. Wood Jr. ^{a,b}, Jimmy A. McGuire ^c, Jamie R. Oaks ^a, L. Lee Grismer ^d, Jesse L. Grismer ^d, Anchalee Aowphol ^e, and Jack W. Sites Jr. ^{b,f}

^aDepartment of Biological Sciences and Museum of Natural History, Auburn University, 331 Funchess Hall, Auburn, AL, 36849

^bDepartment of Biology and M.I. Bean Life Science Museum, Brigham Young University, 150 East Bulldog Boulevard, Provo, UT, 84602

^cMuseum of Vertebrate Zoology, University of California, 3101 Valley Life Sciences Building, Berkeley, CA, 94720

^dHerpetology Laboratory, Department of Biology, La Sierra University, 4500 Riverwalk Parkway, Riverside, CA, 92515

^eFaculty of Science, Department of Zoology, Kasetsart University, Chatuchak, Bangkok, Thailand

^fDepartment of Biology, Austin Peay State University, Clarksville, TN, 37044

Abstract

Southeast Asia hosts a rich concentration of biodiversity within multiple biodiversity hotspots. Indochina, a region with remarkably high levels of *in situ* diversification, possesses five major rivers (Chiang Mai, Ayeyarwady, Mekong, Salween, and Red), several of which coincide with phylogenetic breaks of terrestrial taxa. *Draco maculatus* possesses a range that stretches across Indochina, which widespread geographic distribution along with potential discrete variation within subspecies alludes to the possibility of this taxon constituting multiple divergent lineages. Using sequence data from three mitochondrial (*12S*, *16S*, and *ND2*) and three nuclear (*BDNF*, *CMOS*, and *PNN*) genes, we provide the first estimated phylogeny of this hypothesized species complex and examine its phylogeographic architecture with maximum likelihood and Bayes factor delimitation (BFD) approaches. Our results support multiple divergent lineages with phylogenetic breaks coincident with rivers, indicating that river barriers may be contributing to the elevated levels of *in situ* diversification of Indochina.

2.1 Introduction

Speciation, biodiversification, and the natural processes driving these phenomena have been of great interest to biologists since the dawn of evolutionary biology (Wallace, 1852). Southeast Asia is a region with concentrated biodiversity and a complex geological history, and is therefore a prime setting for testing hypotheses of biodiversification and biogeography (Hall, 2009).

Within the continental core of Southeast Asia, Indochina (composed of far-eastern India, Myanmar, Thailand, Cambodia, Laos, Vietnam, southwestern China, and northern Peninsular Malaysia; Fig. 2.1) is a major evolutionary hotspot noted for its remarkable species richness and *in situ* diversification (de Bruyn et al., 2014; Myers et al., 2000). Indochina is home to the Indo-Burma biodiversity hotspot, a region noted among the ‘hottest hotspots’ for its high degree of endemism and threatened habitat (Myers et al., 2000; Tordoff et al., 2012). Species descriptions from this area are accumulating at a high rate, yet the forces driving speciation within this megadiverse landscape have not been well characterized. While its montane contour has remained stable over the past 20 million years (Bain and Hurley, 2011; Workman, 1975), multiple studies of Indochina’s biodiversity allude to a significant evolutionary impact from riverine barriers (Campbell et al., 2004; Takacs et al., 2005; Johansson et al., 2007; Reddy, 2008; Zhang et al., 2010a; Zhang et al., 2010b; Wang et al., 2012; Wood Jr. et al., 2012; Hartmann et al., 2013; Agarwal et al., 2014; Wang et al., 2015; Yuan et al., 2016). Here we assess the importance of major SE Asian rivers as potential drivers of lineage diversification in a widespread lizard species, *Draco maculatus*, that is distributed across the entirety of Indochina.

2.1.1 Geological History of Indochina

The current topography of Indochina was largely shaped by the collision of the Indian subcontinent with Eurasia ~50 million years ago (mya), resulting in the uplift of the Tibetan Plateau and concomitant creation of Indochina's major waterways, including the Chao Phraya, Ayeyarwady, Mekong, Red, and Salween rivers (Nie et al., 2018; Horton et al., 2002). All five of these watersheds originate on the Tibetan Plateau, and contiguously gather in close proximity within the Hengduan Mountains before radiating across Indochina (Fig. 2.1).

Although the headwaters of each of these rivers have remained geographically static since their onset, for several of these rivers downstream paths and volumes changed drastically beginning in the Miocene (Lacassin et al., 1998; Meijaard and Groves, 2006). Until recently (post-Pleistocene), the Mekong and Salween rivers flowed via the Chao Phraya River Basin southward beyond the current delta at the Gulf of Thailand in a vast watershed known as the Siam River (Carbannel, 1965; Hutchinson, 1989). While the Ayeyarwady River has maintained its course within the Myanmar Central Basin since the Late Eocene, the volume of this river decreased after its connection with the northern Yarlung Tsangpo River was severed in the Early Miocene (Robinson et al., 2014). Similarly, the course of the Red River within the Song Hong fault has been constant since its origin, but it was fed by the Middle Yangtze from the north until this connection was broken between 24 – 12 mya, resulting in a substantial reduction in flow (Clift et al., 2008; Robinson et al., 2014). Riverine vicariance, wherein waterways act as barriers to gene-flow (Wallace, 1852), has been suggested as a mechanism of allopatric speciation in Indochinese taxa (Meijaard and Groves, 2006; Johansson et al., 2007; Corlett, 2009; Woodruff, 2010; Wang et al., 2015), including multiple species of herpetofauna (Zhang et al., 2010a; Zhang et al., 2010b; Bain and Hurley, 2011; Wood et al., 2012; Hartmann et al., 2013; Grismer et al., 2014,

Yuan et al., 2016). As major geographical barriers within Indochina, these rivers may act as significant agents of allopatric isolation and drivers of biodiversification.

2.1.2 Species Complex of Indochinese Agamid Lizards

Found throughout Southeast Asia, the genus *Draco* (family Agamidae) stands unique among Earth's extant lifeforms due to their wing-like patagia attached to elongated ribs that can be extended voluntarily to generate lift as they glide between trees (Herre, 1958; Klingel, 1965; Colbert, 1967; McGuire *et al.*, 2011). *Draco maculatus* (Gray, 1945) has a large range covering all of Indochina (Taylor, 1963; Musters, 1983). This species is listed as a species of least concern by IUCN (although it is protected in Thailand) despite 95 percent of primary rainforest being lost due to human activity (Myers et al., 2000). Despite the loss of primary forest edge habitat, *D. maculatus* remain abundant due to their ability to persist in disturbed habitats with similar structure, such as rubber tree plantations.

Draco maculatus may constitute multiple evolutionary lineages given (A) its expansive geographic range across Indochina and (B) the occurrence of potentially discrete morphological variation that has been used in the diagnoses of subspecies (Taylor 1963; Musters, 1983). The geographic setting and widespread nature of this putative species complex makes it ideal for testing biogeographic hypotheses, specifically concerning the influence of riverine vicariance on the reported high levels of *in situ* diversification within Indochina. Using molecular phylogenetics and Bayes factor delimitation (BFD), we test hypotheses about the population structure across *Draco maculatus* based on (A) riverine barriers, (B) mitochondrial DNA (mtDNA) variation, and (C) current taxonomy, and recover statistical support for significant

divergence within this group with models of riverine barriers separating lineages as the best explanation for the genetic variation within this Indochinese species complex.

2.2 Materials and Methods

2.2.1 Sample Collection

We obtained a total of 115 *Draco maculatus* samples for this study, which were collected from the field between 1997 and 2013. Many of these were collected by members of our team, with the remainder obtained from museum collections or GenBank (see Appendix 2.B). In sampling individuals we did our best to cover the entire species range, and collected within three of the four subspecies ranges (the only exception being *D. maculatus divergens*, which is restricted to the Chiang Mai valley; Fig. 2.1). To assess the level of lineage structure within *D. maculatus* compared to other *Draco* species, we downloaded ND2 sequence data for most *Draco* species and several outgroup species (see Appendix 2.B).

2.2.2 DNA Extraction and Amplification

We extracted genomic DNA from liver samples (stored in 95% ethanol or RNA-Later) using the animal tissue protocol in the Qiagen DNeasyTM tissue kit (Qiagen). We first amplified a region of NADH dehydrogenase subunit 2 (*ND2*) for all samples to guide our subsampling of individuals from which to collect more loci (to minimize redundancy and lower sequencing cost). Our subsampling method is detailed in section 2.3. We then amplified genetic regions of six loci for all subsampled individuals, including three mitochondrial (mitochondrially encoded 12S ribosomal RNA [*12S*], mitochondrially encoded 16S ribosomal RNA *16S*, and *ND2*) and three nuclear (brain-derived neurotrophic factor [*BDNF*], oocyte maturation factor Mos [*CMOS*], and

pinin [*PNN*]) genes using double stranded polymerase chain reaction (PCR) under the following conditions: 1.0 μ L 5X buffer (1.5 μ M), 0.1 μ L Taq polymerase, 1.0 μ L Bovine Serum Albumin (BSA [0.05 mg/mL]), 1.0 μ L dinucleotide pairs (dNTP [1.5 μ L]), 1.0 μ L primer (Appendix 2.C), and 6.5 μ L diH₂O. Each PCR was carried out on an MJ Research PTC-2000 Peltier Thermal Cycler using the following protocol: Initial denaturation at 94° C for 2 min, followed by 35 cycles of a secondary denaturation at 94°C for 30 s, annealing for 30 s (temperatures varied by locus, see Appendix 2.C), elongation at 72°C for 1.5 min, with a final extension at 72°C for 10 min. We visualized the PCR product on agarose gels, and purified PCR products that successfully amplified the targeted region using PrepEase PCR Purification 96-well (Ultrafiltration) plates. We resuspended purified PCR products in diH₂O, and sequenced all samples using an ABI Big-Dye Terminator v3.1 Cycle Sequencing Kit in an ABI GeneAmp PCR 9700 thermal cycler. We purified sequencing reaction products using Sephadex G-50 Fine (GE Healthcare), and generated sequences on an ABI 3730xl DNA Analyzer. After assembling sample-specific contigs and edited sequences in GENEIOUSv8.1.5 (Kearse et al., 2012), we aligned the DNA fragments using the MUSCLEv3.831 algorithm (Edgar, 2004) and adjusted the alignments manually in Mesquite v3.02 (Maddison and Maddison, 2015). Multiple individuals had loci sequenced prior to the onset of this study, and these sequence data were also included in the sub-sample set (Appendix 2.B).

2.2.3 Maximum Likelihood Tree Inference and Sequence Analysis

We estimated all gene trees using maximum likelihood (ML) in IQ-TREE with 1,000 nonparametric bootstrap replicates; we used the “-m TEST” function to perform model selection for each locus followed by subsequent tree inference (Nguyen et al., 2015; Kalyaanamoorthy et

al, 2017). Because *ND2* is a rapidly evolving protein-coding gene, we used it to guide our selection of samples for which to collect nuclear DNA sequences. More specifically, we used the ML tree estimated from the *ND2* sequences to sub-sample at least one individual from each clade of identical or nearly identical sequences, effectively reducing our original sample set of 115 individuals to a sub-sample set of 62 individuals. All analyses described below used this sub-sample of 62 individuals.

We created a concatenated alignment of all the loci using the program Sequence Matrix (Vaidya et al., 2011). When estimating the tree from the concatenated alignment, we allowed each gene to evolve under a different model of nucleotide substitution along a shared genealogy. We then separated mitochondrial and nuclear genes into respective nexus files and estimated ML trees for each of these (Fig. 2.2). Finally, we estimated ML gene trees for each of the six genes (Figs. 2.S4–2.S9). All ML trees were visualized using the program FigTree (Rambaut, 2012).

After determining the number of variable sites (S) and nucleotide diversity (π) for each locus (treating the three mtDNA regions as a single locus) within the program MEGA7 (Kumar et al., 2016), we identified haplotypes using the package Haplotypes in R (Aktas, 2015). We obtained the proportion of variable sites within each gene from the IQ-TREE output log (Table 2.1).

To make intrageneric comparisons between *Draco maculatus* and other *Draco* species, we aligned our *ND2* sequence data with that of 32 other *Draco* species and several additional genera for outgroup and rooting purposes (Appendix 2.B; McGuire and Kiew, 2001; Grismer et al., 2016). After estimating the ML gene tree (Fig. 2.2A), we calculated the phylogenetic distance

between each pair of sister lineages by summing the branch lengths that separate the sister pairs using the R package *ape* (calculating the mean branch length for lineages composed of multiple tips; Paradis and Schliep, 2018). We plotted phylogenetic distances for lineage pairs using the R package *ggplot2* (Wickham, 2009).

2.2.4 Testing Hypotheses of Phylogeographic Structure

We used the nuclear DNA (nDNA) sequence data to estimate marginal likelihoods of 15 models, designing each model based on one of three categories: (1) riverine barriers, (2) mtDNA clustering, or (3) current taxonomy. The riverine barrier models utilize contemporary GIS data for mapping these rivers, with the inclusion/exclusion of specific rivers in differing models reflects historical paths that may have been involved in biodiversification. These models are described in detail within the supplementary methods, with graphical summaries of each model provided in Fig. 2.4.

Stepping-stone sampling (Xie et al., 2011) is a reliable approach for estimating the marginal likelihoods of phylogenetic models, and it outperforms other widely used approaches in phylogenetics (Baele et al., 2012, Oaks et al., 2019). The program BEAST2 (Bouckaert et al., 2014) facilitates stepping-stone sampling within a fully Bayesian framework during species-tree estimation (Leaché et al., 2014). We estimated the marginal likelihoods for each model, verifying our results with a total of seven uniquely seeded BEAST runs per model.

For each model, we designed an XML file using the BEAST graphic user-interface BEAUTi, using the *BEAST template. Nuclear sequence data was phased using a python script written by

Jamie Oaks (Oaks, 2012). Mitochondrial sequence data was not included in marginal likelihood estimation. Samples were grouped into taxon sets based on the criteria from each model. Because our NULL model hypothesized a single lineage and *BEAST requires at least two taxa for species tree estimation, we used sequence data from a single *D. fimbriatus* as the outgroup taxon in all hypotheses. Each nuclear locus was treated as an independent (unlinked) evolutionary unit, and a model of nucleotide substitution was selected for each using the BIC criterion within the program Partition Finder (Table 2.1; Lanfear *et al*, 2012). We used a strict molecular clock for each locus, and fixed the mutationRate parameter to be equal (mean = 1.0) among the nuclear loci. We used an exponential distribution with a mean of 0.001 as the hyperprior on the scale parameter of the gamma-distributed prior on the effective population sizes, with the population function set to “linear_with_constant_root.” A Yule process model was used for the species tree prior, which assumes a constant lineage birthrate for each branch of the tree; we used an exponential distribution with a mean of 300 for the prior on the birthrate (initialized at 300). For a prior on kappa values (transition-transversion parameter of the HKY model or rate AG/CT parameter of partition of the TN93 model, depending on the partition) and the shape parameter of the gamma-distributed rates among sites, we used a lognormal (log mean = 1, log standard deviation = 1.25) and exponential (mean = 1) distribution, respectively. To approximately sample from the posterior distribution, we ran a Markov Chain Monte Carlo (MCMC) simulation for 80,000,000 generations (with a preburnin of 20,000,000), and sampled every 5,000 generations. Following the tutorial by Leaché and Ogilvie (2016), we then manually edited the XML file for stepping-stone sampling (100 steps; 10,000,000 iterations per step; total chain length of 80,000,000). We ran 4 randomly seeded BEAST (v2.4.3) replicates for each model, and combined log (only step1– the final step) and tree files. We assessed convergence of all

parameters visually and using effective sample sizes > 200 using the program Tracer v1.6.0 (Rambaut *et al.*, 2014).

Bayes factors (BF) are measurements obtained via Bayesian model comparison to identify the model that best explains empirical data among a group of hypothesized models (Baele *et al.*, 2013; Grummer *et al.*, 2014; Oaks *et al.* 2019). We calculated Bayes factors on a $2\ln(\text{BF})$ scale by finding twice the pairwise difference between median log marginal likelihood values of the models being compared. We evaluated Bayes factors under the criteria provided by Kass and Raftery (1995): [A] $0 < 2\ln(\text{BF}) < 2$ is not more than a bare mention, [B] $2 < 2\ln(\text{BF}) < 6$ is positive support, [C] $6 < 2\ln(\text{BF}) < 10$ is strong support, and [D] $2\ln(\text{BF}) > 10$ is decisive.

2.2.5 Divergence Time Estimation

We estimated the timing of *Draco maculatus* diversification using *BEAST (in BEAST2) and the species-tree model that best explained the sequence data according to our BFD analyses. The mitochondrial rate of evolution for other agamid lizards has been estimated at $\sim 0.65\%$ per million years (Macey *et al.*, 1998). However, this value was obtained using parsimony distances that did not account for multiple substitutions per site, and therefore may have underestimated the true evolutionary rate. Using the *Paralaudakia* dataset from Macey *et al.* (1998), we aligned and truncated the mtDNA sequence data (*Paralaudakia*) to only include the *ND2* region used in our study. We estimated pairwise distances by calculating maximum composite likelihoods (sum of the log-likelihoods) between all *Paralaudakia* sequence pairs in MEGA7, and applied these values to the same vicariance-based calibration using statistical regression (Fig. 2.S1; Appendix

2.G). Our analysis estimates an evolutionary rate of 1% per million years for *Paralaudakia*, and we fixed this substitution rate for the *ND2* partition. Rates for the remaining mitochondrial and nuclear regions were estimated based on the sequence data relative to the *ND2* rate.

With both phased nuclear sequence data and mitochondrial sequence data, we used the *BEAST template in BEAUTi for XML file formatting. We linked the tree models for the mitochondrial loci (due to the lack of recombination in the mtDNA genome), but left the tree model unlinked for nuclear loci and site and clock models unlinked for all loci due to the time calibration rate being specific to *ND2* region and the differences in the number of variable sites and selected models for each locus. We selected site models using the BIC criterion within the program Partition Finder (Table 2.1; Lanfear *et al*, 2012). We assumed a strict molecular clock for each locus; we fixed the substitution rate of *ND2* to 0.01, and used an exponentially distributed prior (with means and starting values at 0.01) for the other loci. For the scale parameter of the gamma-distributed prior on effective population sizes, we used a 1/x prior, with the population function set to “linear_with_constant_root.” A Yule process model was used for the species tree prior, which assumes a constant lineage birthrate for each branch of the tree; we assumed a 1/x distributed prior on the birthrate. For priors on the kappa parameters (transition-transversion parameter of the HKY model or rate AG/CT parameter of the TN93 model, depending on the partition) and shape parameters of the gamma-distributed rates among sites, we used lognormal (log mean = 1, log standard deviation = 1.25) and exponential (mean = 1) distributions, respectively. To sample from the posterior distribution, we ran a Markov Chain Monte Carlo (MCMC) simulation for 50,000,000 generations, and sampled every 5,000 generations. We ran five replicate BEAST (v2.4.8) analyses with random number starting seeds. Within the program

Tracer v1.6.0 (Rambaut *et al*, 2014) we examined stationary and convergence for all parameters visually and using effective sample sizes (ESS) greater than 200 (after discarding the first 25% of each run as burn-in), and visualized species trees in Figtree (Rambaut, 2012). We consider Bayesian posterior probabilities (PP) ≥ 0.95 as statistically significant (Wilcox *et al*, 2002).

2.3 Results

2.3.1 Sequence Variation

Our concatenated alignment of six gene regions included 4311 base pairs, and aligned sequences revealed that while mtDNA data comprise 56 percent of the complete dataset, it contributes 85 percent of the total variable sites. Further, mtDNA nucleotide diversity is more than an order of magnitude greater than that of each nDNA locus. These results were expected due to the high mutation rate of the mitochondrial genome. Table 2.1 summarizes locus diversity by clade and best-fit models of nucleotide substitution.

2.3.2 ND2 Guide Tree Recovers Five Major Clades

The ND2 ML tree estimated using all individuals (we will refer to this tree as the “ND2 guide tree” in subsequent references within this paper) recovered five well-supported clades (Fig. 2.S2), which may represent instances of local endemism in five regions of Indochina: (1) Central [bootstrap proportion (BP) = 81], (2) Northeast [a.k.a. China-Vietnam Coastal Plain; BP = 100], (3) Hainan Island [BP = 100], (4) Southeast [BP = 70], and (5) West [BP = 100].

2.3.3 Discordant ML Trees from Sub-sample Set

Topological comparisons between trees indicate that the six-locus ML tree appears to be primarily influenced by the mitochondrial rather than the nuclear data (Fig. 2.S3; Fig. 2.2). Further, this mtDNA topology ML tree seems to be largely influenced by ND2 sequences, the 12S and 16S gene trees reflect similar relationships where nodes are well-supported, but between-group relationships are unsupported (Figs. 2.S4–2.S6). The concatenated nDNA ML tree (Fig. 2.2B) and each nDNA gene tree (Figs. 2.S4–2.S8) are discordant with the 6-locus (Fig. 2.S3) and mitochondrial (Fig. 2.2A) ML trees. However, the concatenated nDNA ML tree contains very few (3 of 58) well-supported nodes. The tip labeled “?” in both Figs. 2.2A and 2.2B (and supplemental figures) corresponds to an individual that was purchased at a market (ROM47738), and therefore the locality could not be specified.

2.3.4 Genus-wide Analysis Shows Species-level Divergence

In order to compare the divergence between mitochondrial lineages in a genus-wide context, we examined phylogenetic distance between sister pairs across *Draco*. Our resulting genus-wide ND2 ML tree (Fig. 2.3) largely reflects the relationships estimated by McGuire and Kiew (2001). Differences in topology between the trees include two regions where bootstrap support values are low in both analyses (relationships within the Philippine *volans* group, and the placement of *D. sumatranus* relative to *D. volans* and the *D. boschmai* group). Reasons for variation between these trees may include (A) shorter sequence region (our alignment did not include the tRNA regions that were obtained by McGuire and Kiew [2001]), (B) decreased tip diversity (we did not include undescribed species or geographic variants used by McGuire and Kiew [2001]), and (C) different methodology (PAUP* vs IQ-TREE, GTR vs TIM2). Because all points of disagreement

involve poorly supported relationships, we do not address other species placements within *Draco*.

Draco maculatus within the genus-wide ND2 ML tree subdivides into the same five clades we see in Fig. 2.2A, albeit with reduced support for the Central (BP = 64) and Southeast (BP = 56) clades (it is worth noting that the Southeast clade receives strong support [BP = 100] at the node excluding ROM32033 [locality G; Fig. 2.S10]). Our genus-wide comparisons show that the phylogenetic distance between the lineages described in this study ([West]-[Central]: 0.18; [Hainan Island]-[Northeast]:0.14; [[Hainan Island]-[Northeast]]-[Southeast]: 0.16) are greater than 27% of our included *Draco* sister pairs (Appendix 2.E). When species richness per sister group is accounted for (i.e., when the number of species in the sister clade is taken into account), the [West]-[Central] lineages are separated by a phylogenetic distance greater than 55 percent of other *Draco* species (Fig. 2.3B). The [[Hainan Island]-[Northeast]]-[Southeast] group is separated by a phylogenetic distance greater than 16 percent of other *Draco* species (Fig. 2.3B). These data imply that, according to this mitochondrial region (*ND2*), these *D. maculatus* lineages have experienced divergence commensurate with the majority of included comparable *Draco* species.

2.3.5 BFD Supports Eight Lineages

The MCMC computational running time for marginal likelihood estimation (MLE) is described in Appendix 2.F. It is worth re-emphasizing that although information from mtDNA sequence data was used in model design (models F, H, and M), we did not use these sequence data for testing any of the models (MLE analyses were performed using nDNA sequence data only). The

River-8 hypothesis (Model A) yielded the largest median MLE value (see Table 2.2). Bayes factor analyses of model A with each other model indicates the following: decisive support compared to models L–O, strong support compared to model D–K, and positive support compared to model B and C (Fig. 2.4). Based on these findings we infer that (1) *Draco maculatus* is not a single, panmictic population, and (2) current subspecies taxonomy does not best reflect evolutionary history. We found a strong positive correlation between the lineage number and the estimated marginal likelihood ($p < 1.0 \times 10^{-5}$), indicating that our BFD may favor splitting.

The River-8 hypothesis (Model A): Aquatic barriers best explain the nDNA variation. Riverine and marine bodies acting as agents of geographic isolation include the Chao Phraya, Ayeyarwady, Mekong, Red, and Salween rivers and the Qiongzhou Strait. Lineage descriptions: (1) Western Myanmar Hills [east of Ayeyarwady], (2) Central Myanmar Lowlands [west of Ayeyarwady, east of Salween], (3) Thai-Malay Peninsula, (4) Shan Hills [west of Salween, east of Mekong, north of Chao Phraya], (5) Cardamom Mountains [west of Mekong, east of Chao Phraya], (6) Annamite Mountains [east of Mekong, south of Red], (7) China-Vietnam Coastal Plain [north of Red], and (8) Hainan Island [separated by Qiongzhou Strait].

2.3.6 Species Trees and Divergence Time Estimations

Because BFD indicated strongest support for model A, we used this model for estimating a species tree using all of the data (mtDNA and nDNA; Fig. 2.5). Our time-calibrated species tree estimate (Fig. 2.5) indicates strong support for the monophyly of *Draco maculatus* with respect to *D. fimbriatus*, with the split between the two occurring 30 – 16 mya (placing the most recent

common ancestor [MRCA] within the Late Oligocene – Early Miocene). Estimates for the most recent common ancestor for all extant *D. maculatus* range between 10 – 7 mya, placing the basal node of the *D. maculatus* radiation during the Late Miocene. Timing of diversification events within *D. maculatus* ranges throughout the Late Miocene, Pliocene, and Pleistocene. We include biogeographic implications with respect to estimated node ages within the discussion.

2.4 Discussion

In order to provide a biogeographical analysis of Indochina and assess the lineage content of the *Draco maculatus* complex, we performed a comprehensive phylogenetic analysis of *D. maculatus* with samples from across Indochina. Our ML results indicate that that this widespread species is actually composed of at least five lineages whose mitochondria are as divergent from one another as most other currently recognized *Draco* species pairs, and our BFD results support eight divergent lineages. Contemporary river courses for the major rivers of Indochina, which have been suggested as agents of diversification in other Indochinese taxa, appear to correlate with phylogeographic breaks between divergent *D. maculatus* lineages.

2.4.1 Biogeography

Musters (1983) suggested that the variation observed in *Draco maculatus* may be a result of population isolation during the last glacial period. Our divergence time estimations based on molecular clock analysis suggest that most divergence within *D. maculatus* predates the onset of the Quaternary ice age (Fig. 2.5), therefore rendering Musters' (1983) hypothesis unlikely. While Pleistocene glacial cycles likely had little influence on diversification of the *D. maculatus* species complex, older forces within this region may have played important roles. Changes in

forest structure and microhabitat offer one potential explanation. However, throughout most of the Miocene and Pliocene predominantly warm and humid conditions persisted throughout much of Indochina (Meijaard and Groves, 2006) and montane dynamics remained mostly stable (Bain and Hurley, 2011; Workman, 1975). Water barrier dynamics offer another explanation, wherein shifting river courses and marine channels may have restricted gene flow and led to population divergence (see Wallace, 1852). To evaluate the potential role of riverine barriers in *D. maculatus* diversification, we examined alternative associations of six water channels (Chao Phraya, Ayeyarwady, Mekong, Salween, and Red rivers, and the Qiongzhou Strait) relative to lineage boundaries within the *D. maculatus* complex (see Table 2.3). While each of these rivers has an origin on the Tibetan Plateau dating back to the Miocene (Brookfield, 1998), their volumes and downstream courses have experienced substantial shifts over time.

The Mekong river is Southeast Asia's largest river and one of the ten largest rivers in the world today. The upper Mekong river valley became deeply incised within the Qiantang, Lhasa, and Himalaya terranes during the middle Miocene (Nie et al., 2018). This date for rapid river cutting coincides with a period of strong monsoon activity within Southeast Asia (Clift, 2006; Guo et al., 2002), in which intensified erosion likely contributed to the reported increase in incision rate of the Mekong River Valley. Geological and biological evidence indicates that until recently (post-Pleistocene), the Mekong and Salween rivers joined the Siam River system and flowed to the presently recognized Gulf of Thailand through the Chao Phraya River Basin, dividing Indochina into East and West regions (Sawamura and Laming, 1974; Carbonnel, 1965; Hutchinson, 1989; Attwood and Johnston, 2001; Glaubrecht and Köhler, 2004; Lukoschek et al., 2011). The basal *D. maculatus* bifurcation reflected in both the mtDNA ML tree (Fig. 2.2A) and time-calibrated

species tree (Fig. 2.5) correspond with the presence of this longitudinal riverine boundary (9.9 – 6.5 mya; Late Miocene), and may explain the geographic and phylogenetic division within the species tree into east/west groups.

The Red River fault zone forms the boundary between the South China and Indochina plates, where water flow has remained constant through the Red River since the India-Asia collision ~30 mya, (Searle, 2006; Leloup et al., 1995; Hall, 1998). This shear zone largely delimits the boundary between subtropical and tropical climates (Chen and Chen, 2013; Peel et al., 2007). The Red River is hypothesized to inhibit gene flow in the bird *Parus monticolus* (Wang et al., 2012), and in the frogs *Microhyla fissipes* (Yuan et al., 2016), *Leptobrachium ailaonicum* (Zhang et al., 2010b), and *Nanorana yunnanensis* (Zhang et al., 2010a), *Ichthyophis bannanicus* (Wang et al., 2015), and appears to coincide with a barrier for the two northeast lineages (China-Vietnam Coastal Plain + Hainan Island) from populations to the west (Cardamom Mountains + Annamite Mountains + Shan Hills; Fig. 2.5). Divergence time estimation in other organisms around the Red River are consistent our estimates for *Draco maculatus* (8.6 – 6.0 mya; Yuan et al., 2016; Zhang et al., 2010a; Zhang et al., 2010b), which is contemporaneous with a significant extrusion event along the Red River fault zone (Xiang et al., 2004; Leloup et al., 1995).

While the geological (volcanic vs. continental) and geographic (Leizhou Peninsula vs Beibu Gulf) history of Hainan Island is debated (Zhao et al., 2007; Zhu, 2016), it is widely accepted that the Qiongzhou Strait has separated the island from the Leizhou Peninsula for the past ~2.5 million years with periodic land bridges during glacial maxima. MtDNA indicates significant divergence between the Northeast clade (China-Vietnam Coastal Plain) and Hainan Island clade

(Figs. 2.2A, 2.6), and divergence time estimation suggests that these lineages split between 7.0–2.1 mya (Fig. 2.5). While a portion of this range predates the estimated origin of the Qiongzhou Strait, it reflects similar squamate colonization times across the same marine channel in the snakes *Trimeresurus albolabris* (Zhu et al., 2015), *T. stejnegeri* (Liang et al., 2018), and *Protothrops mucrosquamatus* (Guo et al., 2019), and the gecko *Goniurosaurus lichtenfelderi* (Guo et al., 2016).

The Salween and Ayeyarwady Rivers also appear to be associated with population structure in *Draco maculatus*, an observation consistent with the century-old “Mekong-Salween Divide” and “Salween-Ayeyarwady Divide” hypotheses (Li et al., 2011; Ward, 1921). The Central Myanmar Lowland and Western Myanmar Hills lineages, divided by the Salween, diverged between 3.5 – 1.5 mya, and these two lineages diverged from the Thai-Malay Peninsula group between 3.9 – 2.7 mya. These divergence times coincide with the hypothesized capture of a major drainage in the Southeastern region (~4 mya), which increased the volume of the Salween via capture of the Mekong (Clark et al., 2004). Other Indochinese vertebrates also appear to experience population divergence corresponding to the course of the Salween and Ayeyarwady river basins, including in the gecko genus *Cyrtodactylus* (Wood Jr. et al., 2012; Agarwal et al., 2014), bird genera *Phylloscopus* and *Seicercus* (Johansson et al., 2007), gibbon genera *Bunopithecus* and *Hylobates* (Takacs et al., 2005), and bat genus *Cynopterus* (Campbell et al., 2004).

The current course of the Mekong River has received much attention in studies of Indochinese biogeography (Geissler et al., 2015; Bain and Hurley, 2011; Tantrawatpan, 2011; Meijaard and Groves, 2006; Fooden, 1996; Long et al., 1994), with mixed results regarding its impact on gene

flow. Contrasting biogeographic findings within this region may be due in part to the dynamic history of the Mekong, which has experienced a great West to East geographical shift (over 700 km) since the Pleistocene (Carbonnel, 1965; Hutchinson, 1989). Because of the unique histories of the Upper (stable) and Lower (west-east shifting) Mekong, the geographically entangled relationships between Cardamom (4.7–7.8 mya), Annamite (2.1–6.1 mya), and Shan regions are not totally unsurprising (Fig. 2.5). These results echo similar findings in other draconinae lizards (e.g., genus *Calotes*), where the lower Mekong River appears to coincide with a barrier to gene flow (Hartmann et al., 2013).

Several instances of river boundaries separating color-morph races occur within species of *Heliconius*, a neotropical butterfly genus that has become a model for understanding processes of speciation (Lamas, 1982; Beltrán et al., 2002; Arias et al., 2008; Muñoz et al., 2011). Similarly, with the *Draco maculatus* species complex exhibiting widespread color pattern polymorphism, population structure, and phylogenetic boundaries coinciding with riverine barriers, we see potential for the *Draco maculatus* species complex being used as a terrestrial vertebrate system to understand tropical speciation and biodiversification. Understanding whether the rivers themselves versus other co-occurring ecological factors (i.e., the distinct Chinese floristic provinces found on the west and east sides of the upper Salween [Wu and Wu, 1998]) requires further examination involving the specific geographical history of these rivers.

2.4.2 Phylogenetic Discordance and the “Out of China” Hypothesis

The smaller effective population size of the mitochondrial genome (1/4 that of the nuclear genome) is expected to lead to more rapid lineage sorting of the mitochondrion (Funk and

Omland, 2003; Lynch, 2006). Furthermore, phylogenetic estimation of the ND2 genealogy is aided by the high mutation rate of this region and its length relative to other loci included in this study, two characteristics which resulted in a high number of variable sites (Appendix 2.D). Additionally, all three mitochondrial regions evolved along a single gene tree (due to the lack of recombination in the mitochondrial genome; Nichols, 2001). This would explain the well-supported, divergent clades in the mitochondrial tree (Fig. 2.2A), whereas a slower sorting rate of ancestral variation would explain the lack of monophyletic clustering in the nuclear tree (Fig. 2.2B). The larger effective population size and slower mutation rate of nDNA results in a greater time requirement for identifiable lineage emergence (Nichols, 2001). While the effect of incomplete lineage sorting makes it difficult to recover more strongly supported relationships, studies using genome-wide SNP data have had success in resolving previously obscure species boundaries (e.g. Rubin et al., 2012; Leaché et al., 2014; Hou et al., 2017). We suggest that further studies seeking to resolve the *D. maculatus* phylogeny at a finer scale implement a high-throughput SNP dataset.

If the observed mitonuclear phylogenetic discordance is due to incomplete lineage sorting, the region with the most widespread variation throughout the tree may hint at the geographic origin of *Draco maculatus*. Samples from the Northeast clade occur as a monophyletic group sister to all other *D. maculatus* samples, and the remaining Northeast samples are scattered throughout the nDNA ML tree (Fig. 2.2B). This extensive variation could be due to this region being the ancestral origin of the *D. maculatus* lineages, but is contrary to Musters' (1983) hypothesis of *D. maculatus* originating in Thailand. Furthermore, the species within the sister clade of *D. maculatus* (*D. cristatellus*, *D. fimbriatus*, and *D. hennigi*, *D. punctatus*) are all found south of the

Isthmus of Kra, suggesting the initial split may have occurred here. Future testing of biogeographic hypotheses would benefit from inclusion of samples from Thailand.

2.4.3 Regarding Taxonomy

In this study, we provide the first fine-scale molecular examination of the *Draco maculatus* species complex. While our findings from BFD support eight divergent lineages, the multispecies coalescent implemented in *BEAST assumes no population structure within species, and thus can be biased toward lineage division. We therefore opt against elevating subspecies to species status or creating new names for these divergent *D. maculatus* lineages.

Major examinations implementing morphological and molecular analyses have aided in understanding species richness and composition of the Phillipine, Sulawesi, and *Draco fimbriatus* groups (McGuire and Alcala, 2000; McGuire et al., 2007; McGuire et al., 2018). However, the *Draco maculatus* complex has received no such treatment and has consistently puzzled taxonomists (Honda et al., 1999; McGuire and Kiew, 2001). Although the four subspecies were originally described as full species (Gray, 1845; Taylor, 1934; Boettger, 1893; Boulenger, 1899; Appendix 2.A), lack of information regarding the presence or absence of intergradation at the subspecies boundaries has left taxonomists reluctant to elevate subspecies to species status (see, for example, McGuire and Kiew, 2001; Taylor, 1963, Musters, 1983). McGuire and Kiew (2001) suggest the need for a phylogeographic and/or fine-scale morphological study to better understand relationships among the *D. maculatus* subspecies. We echo this suggestion, and recommend that future studies sample individuals from the contact zones of the divergent lineages described in this study.

While superimposing our sampling localities over the subspecies ranges described by Taylor (1963) and Musters (1983) indicates general geographic clustering of some subspecies within unique clades of the ML tree (Fig. 2.1), our results reject the current recognition “four subspecies” of *Draco maculatus* restricted to West Indochina (*D. maculatus maculatus*), Southeast Indochina (*D. maculatus haasei*), Northeast Indochina (*D. maculatus whiteheadi*), and the Chiang Mai Valley (*D. maculatus divergens*) on the basis of (A) our concatenated ML trees recovering a divergent clade grouping samples of westernmost Myanmar with southeastern Laos [Central clade], and (B) our recovery of a divergent fifth clade [Hainan Island] sister to the mainland Northeast clade. The range of *D. maculatus divergens* is defined as being strictly within the Chiang Mai Valley, therefore the Central clade is either a unique group or the range of *D. maculatus divergens* is larger than previously assumed.

Studies of morphology in subspecies have been biased by geographical sampling limitations. For example, the key Musters (1983) developed for subspecies identification only examined *Draco maculatus whiteheadi* from Hainan Island (the type locality), but no material from the mainland. Our molecular results question the taxonomic proximity of island individuals with those of the adjacent mainland, indicating that *D. maculatus* of Hainan may be endemic to the island. Lastly, our genus-wide gene tree (Fig. 2.3) supports the hypothesis that species-level diversity is present within this group, waiting to be described.

2.5 Conclusion

Our results support the hypothesis that there are multiple divergent lineages within *Draco maculatus*. Comparing the fit of a suite of hypotheses for explaining the distribution of genetic variation across this species indicate that the major rivers of Indochina coincide with boundaries between divergent lineages. The rapid origin of the massive Siam River (sourced by the current Upper Mekong) along its Pre-Pleistocene course within the Chao Phraya river valley may have acted as a major geographic barrier, resulting in the phylogenetic split of west and east lineages. These groups may be phenotypically distinct based on dewlap color, with the West group possessing a sky-blue patch at the base of the dewlap and the east group lacking such a patch. Given similar patterns have been observed in other phylogeographic studies in Indochina, our results suggest riverine vicariance as a mechanism for diversification of *D. maculatus* within this megadiverse hotspot.

2.6 Acknowledgements

For tissues, we thank the La Sierra University Herpetological Collection, the Museum of Vertebrate Zoology, the National Museum of Natural History (Daniel C. Mulcahy), the Royal Ontario Museum (Bob Murphy), the North Carolina Museum of Natural Sciences (Bryan L. Stuart), and the California Academy of Sciences (Jens Vindum). We extend a special thanks to Tonia Schwartz, Rory Telemeco, members and alumni of the Phyletica lab at Auburn University who provided comments (Breanna Siple, Kerry Cobb, and Brian Folt), Bryan Stuart, and Paulina Klabacka for their constructive feedback on the manuscript. Computationally intensive phylogenetic methods were carried out on the Auburn University Hopper Cluster. Funding: This work was supported by the National Science Foundation [EF 1241885, to J.W. Sites, Jr], the

Brigham Young University Bean Life Science Museum [to J.W. Sites, Jr], and Office of Research and Creative Activities [#4594, to R.L. Klabacka].

2.7 References

- Agarwal, I., Bauer, A.M., Jackman, T.R., Karanth, K.P. (2014). Insight into Himalayan biogeography from geckos: A molecular phylogeny of *Cyrtodactylus* (Squamata: Gekkonidae). *Molecular Phylogenetics and Evolution*. 80:145-155. DOI: 10.1016/j.ympev.2014.07.018
- Aktas, C. (2015). R package haplotypes: Haplotype inference and statistical analysis of genetic variation. Stable URL: cran.r-project.org/web/packages/haplotypes
- Arias, C.F., Muñoz, A.G., Jiggins, C.D., Mavárez, J., Bermingham, E., Linares, M. (2008). A hybrid zone provides evidence for incipient ecological speciation in *Heliconius* butterflies. *Molecular Ecology*. 17(21):3799-4712. DOI: 10.1111/j.1365-294X.2008.03934.x
- Attwood, S.W., Johnston, D.A. (2001). Nucleotide sequence differences reveal genetic variation in *Neotricula aperta* (Gastropoda: Pomatiopsidae), the snail host of schistosomiasis in the lower Mekong Basin. *Biological Journal of the Linnean Society*. 73:23–24. DOI: 10.1006/bijl.2000.0520
- Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., et al. (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*. 29(9):2157–2167. DOI: 10.1093/molbev/mss084
- Baele, G., Lemey, P., Vansteelandt, S. (2013). Make the most of your samples: Bayes factor estimators for high-dimensional models of sequence evolution. *BMC Bioinformatics*. 14(85). DOI: 10.1186/1471-2105-14-85

- Bain, R.H. and Hurley, M.M. (2011). A biogeographic synthesis of the amphibians and reptiles of Indochina. *Bulletin of the American Museum of Natural History*. 46(2). DOI: 10.1206/360.1
- Beltrán, M., Jiggins, C.D., Bull, V., Linares, M., Mallet, J. et al. (2002). Phylogenetic discordance at the species boundary: Comparative gene genealogies among rapidly radiating *Heliconius* butterflies
- Boettger, O. (1893). Ein neuer *Drache* (*Draco*) aus Siam. *Zoologischer Anzeiger*. 16:429–430
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., et al. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLOS Computational Biology*, 10(4), 1–6. DOI: 10.1371/journal.pcbi.1003537
- Boulenger, G.A. (1900). On the reptiles, batrachians (and fishes) collected by the late Mr. John Whitehead in the interior of Hainan. *Proceedings of the Zoological Society of London*. 1899:956–959
- Brookfield, M.E. (1998). The evolution of the great river systems of southern Asia during the Cenozoic India-Asia collision: rivers draining southwards. *Geomorphology*. 22(3–4):285–312. DOI:10.1016/S0169-555X(97)00082-2
- de Bruyn, M., Stelbrink, B., Morley, R.J., Hall, R., Carvalho, G.R., et al. (2014). Borneo and Indochina are Major Evolutionary Hotspots for Southeast Asian Biodiversity. *Systematic Biology*. 63(6):879–901. DOI: 10.1093/sysbio/syu047

- Campbell, P., Schneider, C.J., Adnan, A.M., Zubaid, A., Kunz, T.H. (2004). Phylogeny and phylogeography of Old World fruit bats in the *Cynopterus brachyotis* complex. *Molecular Phylogenetics and Evolution*. 33:764-781. DOI: 10.1016/j.ympev.2004.06.019
- Carbonnel, J.P. (1965). Essai d'interpretation morphotectonique ' de la cuvette Cambodgienne. *Revue de Geographie Phisique et de Geologie Dynamique*. 7:277–281
- Chen, D.L., Chen, H.W. (2013). Using the Köppen classification to quantify climate variation and change: An example for 1901–2010. *Experimental Development*. 6:69–79. DOI: 10.16/j.envdev.2013.03.007
- Clark, M.K., Schoenbohm, L.M., Royden, L.H., Whipple, K.X., Burchfiel, B.C., et al. (2004). Surface uplift, tectonics, and erosion of eastern Tibet from large-scale drainage patterns. *Tectonics*. 23(TC1006):1-6. DOI: 10.1029/2002TC001402
- Clift, P.D. (2006). Controls on the erosion of Cenozoic Asia and the flux of clastic sediment to the ocean. *Earth and Planetary Science Letters*. 241:571–580. DOI: 10.1016/j.epsl.2005.11.028
- Corlett, R.T. (2009). *The ecology of tropical East Asia*. Oxford University Press, Oxford.
- Edgar, R.C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 32(5): 1792–97
- Fooden, J. (1996). Zoogeography of Vietnamese primates. *International Journal of Primatology*. 17(5):845-899. DOI: 10.1007/BF02735268

- Funk, D.J., Omland, K.E. (2003). Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*. 34:397–423. DOI: 10.1146/annurev.ecolsys.34.011802.132421
- Geissler, P., Hartmann, T., Ihlow, F., Rödder, D., Poyarkov Jr., N.A., et al. (2015). The lower Mekong: An insurmountable barrier to amphibians in southern Indochina? *Biological Journal of the Linnean Society*. 114(4):905-914. DOI: 10.1111/bij.12444
- Glaubrecht, M., Köhler, F. (2004). Radiating in a river: systematics, molecular genetics and morphological differentiation of viviparous freshwater gastropods endemic to the Kaek River, central Thailand (Cerithioidea, Pachychilidae). *Biological Journal of the Linnean Society*. 82(3):275–311. DOI: 10.1111/j.1095-8312.2004.00361.x
- Gray, J.E. (1845). Catalogue of the specimens of lizards in the collection of the British Museum. Trustees of die British Museum/Edward Newman, London: xxvii + 289 pp. DOI: 10.5962/bhl.title.5499
- Grismer, J.L., Bauer, A.M., Grismer, L.L., Thirakhupt, K., Aowphol, A., et al. (2014). Multiple origins of parthenogenesis, and a revised species phylogeny of the Southeast Asian butterfly lizards, *Leiolepis*. *Biological Journal of the Linnean Society*. 113(4):1080–1093. DOI: 10.1111/bij.12367
- Grismer, J.L., Schulte II, J.A., Alexander, A., Wagner, P., Travers, S.L., et al. (2016). The Eurasian invasion: phylogenomic data reveal multiple Southeast Asian origins for Indian Dragon Lizards. *BMC Evolutionary Biology*. 16(43). DOI: 10.1186/s12862-016-0611-6

- Grummer, J.A., Bryson, R.W. Jr., Reeder, T.W. (2014). Species delimitation using Bayes factors: simulations and application to the *Sceloporus scalaris* species group. *Systematic Biology*. 63(2):119–33. DOI: 10.1093/sysbio/syt069
- Guo, P., Liu, Q., Zhu, F., Zhong, G.H., Burbrink, F.T., et al. (2016). Complex longitudinal diversification across South China and Vietnam in Stejneger’s pit viper, *Viridovipera stejnegeri*. *Molecular Ecology*. 25:2920-2936. DOI: 10.1111/mec.13658
- Guo, P., Liu, Q., Zhu, F., Zhong, G.H., Malhotra, A., et al. (2019). Multilocus phylogeography of the brown-spotted pitviper *Protobothrops mucrosquamatus* sheds a new light on the diversification pattern in Asia. *Molecular Ecology*. 133:82-91. DOI: 10.1016/j.ympev.2018.12.028
- Guo, Z.T., Ruddiman, W.F., Hao, Q.Z., Wu, H.B., Qiao, Y.S., et al. (2002). Onset of Asian desertification by 22 Myr ago inferred from loess deposits in China. *Nature*. 416(6877):159–163. DOI: 10.1038/416159a
- Hall, R. (1998). The plate tectonics of Cenozoic Asia and the distribution of land and sea. In: Hall, R., Holloway, J.D. (editors). *Biogeography and Geological Evolution of Southeast Asia*. Leiden, Backhuys, pp. 99–131
- Hall, R. (2009). Southeast Asia’s changing palaeogeography. *Blumea: Journal of Plant Taxonomy and Plant Geography*, 54(1–3), 148–161. DOI: 10.3767/000651909X475941
- Hartmann, T., Geissler, P., Poyarkov Jr., N.A., Ihlow, F., Galoyan, E.A., et al. (2013). A new species of the genus *Calotes* (Cuvier, 1817; Squamata: Agamidae) from southern Vietnam. *Zootaxa*. 3599(3):246–260. DOI: 10.11646/zootaxa.3599.3.3

- Honda, M., Ota, H., Kobayashi, M., Nabhitabhata, J., Yong, H.-S., and Hikida, T. (1999b). Phylogenetic Relationships of the Flying Lizards, Genus *Draco* (Reptilia, Agamidae). *Zoological Science*, 16(3), 535–549. DOI: 10.2108/zsj.16.535
- Hou, Y., Nowak, M.D., Mirré, V., Bjorå, C.S., Brochmann, C., et al. (2015). Thousands of RAD-seq loci fully resolve the phylogeny of the highly disjunct arctic-alpine genus *Diapensia* (Diapensiaceae). *PLOS ONE*. 10(10):e0140175. DOI: 10.1371/journal.pone.0140175
- Horton, B.K., Yin, A., Spurlin, M.S., Zhou, J., Wang, J. (2002). Paleocene–Eocene syncontractational sedimentation in narrow, lacustrine-dominated basins of east-central Tibet. *Geological Society of America*. 114(7):771–786. DOI: 10.1130/0016-7606(2002)114<0771:PESSIN>2.0.CO;2
- Hutchinson, C.S. (1989). Geological evolution of South-east Asia. Oxford: Clarendon Press.
- Johansson, U.S., Alström, P., Olsson, U., Ericson, P.G., Sundberg, P., et al. (2007). Build-up of the Himalayan avifauna through immigration: A biogeographical analysis of the *Phylloscopus* and *Seicercus* warblers. *Evolution*. 61(2):324–333. DOI: 10.1111/j.1558-5646.2007.00024.x
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermini, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*. 14(6):587–589. DOI: 10.1038/nmeth.4285
- Kass, R.E., Raftery, A.E. (1995). Bayes factors. *Journal of the American Statistical Association*. 90(430):773–795

- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., et al (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), 1647-1649
- Kumar, S., Stecher, G., Tamura, K. (2016). MEGA7: Molecular Evolutionary Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*. 33(7):1870–1874. DOI: 10.1093/molbev/msw054
- Lacassin, R., Replumaz, A., Leloup, P.H. (1998). Hairpin river loops and slip-sense inversion on southeast Asian strike-slip faults. *Geology*. 26(8):703-706. DOI: 10.1130/0091-7613(1998)026<0703:HRLASS>2.3.CO;2
- Lamas, G. (1982). A Preliminary Zoogeographical Division of Peru, Based on Butterfly Distributions (Lepidoptera, Papilionoidea). *Biological Diversification in the Tropics*. Columbia University Press, New York, pp. 336-357
- Lanfear, R., Calcott, B., Ho, S.Y.M., Guindon, S. (2012). PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29(6): 1695–1701. DOI: 10.1093/molbev/mss020
- Leaché, A.D., Fujita, M.K., Minin, V.N., Bouckaert, R.R. (2014). Species delimitation using genome-wide SNP data. *Systematic Biology*. 63(4):534–542. DOI: 10.1093/sysbio/syu018
- Leaché, A.D., Ogilvie, H.A. (2016). Bayes factor delimitation of species: A tutorial and worked example. *2016 Workshop on Population and Speciation Genomics*

- Leloup, P.H., Lacassin, R., Tapponnier, P., Schärer, U., Dalai, Z., et al. (1995). The Ailao Shan-Red River shear zone (Yunnan, China), Tertiary transform boundary of Indochina. *Tectonophysics*. 251(1–4):3–10, 13–84. DOI: 10.1016/0040-1951(95)00070-4
- Li, Y., Zhai, S.N., Qiu, Y.X., Guo, Y.P., Ge, X.J., et al. (2011). Glacial survival east and west of the ‘Mekong-Salween Divide’ in the Himalaya-Hengduan Mountains region as revealed by AFLPs and cpDNA sequence variation in *Sinopodophyllum hexandrum* (Berberidaceae). *Molecular Phylogenetics and Evolution*. 59(2):412-424. DOI: 10.1016/j.ympev.2011.01.009
- Liang, B., Zhou, R., Liu, Y., Chen, B., Grismer, L., et al. (2018). Renewed classification within *Goniurosaurus* (Squamata: Eublepharidae) uncovers the dual roles of a continental island (Hainan) in species evolution. *Molecular Phylogenetics and Evolution*. 127:646-654. DOI: 10.1016/j.ympev.2018.06.011
- Long, Y., Kirkpatrick, C.R., Zhongtai, X. (1994). Report on the distribution, population, and ecology of the Yunnan snub-nosed monkey (*Rhinopithecus bieti*). *Primates*. 3(2):241-250. DOI: 10.1007/BF02382060
- Lukoschek, V., Osterhage, J.L., Karns, D.R., Murphy, J.C., Voris, H.K. (2011). Phylogeography of the Mekong mud snake (*Enhydris subtaeniata*): the biogeographic importance of dynamic river drainages and fluctuating sea levels for semiaquatic taxa in Indochina. *Ecology and Evolution*. 1(3):330–342. DOI: 10.1002/ece3.29
- Lynch, M., Koskella, B., Schaak, S. (2006). Mutation pressure and the evolution of organelle genomic architecture. *Science*, 311(5768), 1727–1730. DOI: 10.1126/science.1118884

Maddison, W.P., Maddison, D.R. (2017). Mesquite: a modular system for evolutionary analysis.

Version 3.02. <http://mesquiteproject.org>

Macey, R.J., Schulte II, J.A., Ananjeva, N.B., Larson, A., Rastegar-Pouyani, N., et al. (1998).

Phylogenetic relationships among Agamid Lizards of the *Laudakia caucasia* species group:

Testing hypotheses of biogeographic fragmentation and an area cladogram for the Iranian

Plateau. *Molecular Phylogenetics and Evolution*. 10(1):118–131. DOI:

10.1006/mpev.1997.0478

McGuire, J.A., Alcala, A.C. (2000). A taxonomic revision of the flying lizards (Iguania:

Agamidae: *Draco*) of the Philippine islands, with a description of a new species.

Herpetological Monographs. 14:81-138. DOI: 10.2307/1467046

McGuire, J.A., Brown, R.M., Mumpuni, Riyanto, A., Andayani, N. (2007). The flying lizards of

the *Draco linneatus* group (Squamata: Iguania: Agamidae): A taxonomic revision with

descriptions of two new species. *Herpetological monographs*. 21(1): 179–212. DOI:

10.1655/07-012.1

McGuire, J.A., Kiew, B.H. (2001). Phylogenetic systematics of Southeast Asian flying lizards

(Iguania: Agamidae: *Draco*) as inferred from mitochondrial DNA sequence data. *Biological*

Journal of the Linnean Society. 72(203–229). DOI: 10.1006/bijl.2000.0487

McGuire, J.A., Cotoras, D.D., O’Connell, B., Lawalata, S.Z.S., Wang-Claypool, C.Y., et al.

(2018). Squeezing water from a stone: high-throughput sequencing from a 145-year old

holotype resolves (barely) a cryptic species problem in flying lizards. *Peer J*. 6:24470. DOI:

10/7717/peerj.4470

- Meijaard, E., Groves, C.P. (2006). The geography of mammals and rivers in Southeast Asia. In Lehman, S.M. and Fleagle, J.G. (Eds.) *Primate Biogeography*. New York, USA, Springer, pp. 305–329
- Muñoz, A.G., Baxter, S.W., Linares, M., Jiggins, C.D. (2011). Deep mitochondrial divergence within a *Heliconius* butterfly species is not explained by cryptic speciation or endosymbiotic bacteria. *BMC Evolutionary Biology*. 11:358. DOI: 10.1186/1471-2148-11-358
- Musters, C.J.M. (1983). Taxonomy of the genus *Draco* L. (Agamidae, Lacertilia, Reptilia). *Zoologische Verhandelingen*. 199(1):1–120
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403(6772), 853–858. DOI: 10.1038/35002501
- Nguyen, L.T., Schmidt H.A., von Haeseler A. and Minh B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol Biol Evol*, 32:268–274. DOI: 10.1093/molbev/msu300
- Nichols, R. 2001. Gene trees and species trees are not the same. *Trends in Ecology and Evolution*. 16(7):358–364. DOI: 10.1016/S0169-5347(01)02203-0
- Nie, J., Ruetenik, G., Gallagher, K., Hoke, G., Garziona, C.N., Wang, W., Stockli, D., Hu, X., Wang, Z., Stevens, T., Danišík, M., Liu, S. (2018). Rapid incision of the Mekong River in the middle Miocene linked to monsoonal precipitation. *Nature geoscience*. DOI: 10.1038/s41561-018-0244-z

- Oaks, J.R. (2012). Phase.py. https://github.com/joaks1/phase_py.
- Oaks, J.R., Cobb, K.A., Minin, V.N., Leaché, A.D. (2019). Marginal Likelihoods in Phylogenetics: A Review of Methods and Applications. *Systematic Biology*. syz003. DOI: 10.1093/sysbio/syz003
- Paradis, E., Schliep, K. (2018). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. 35:526-528. DOI: 10.1093/bioinformatics/bty633
- Peel, M.C., Finlayson, B.L., McMahon, T.A., (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*. 11:1633–1644. DOI: 10.5194/hess-11-1633-2007
- Rambaut, A. (2012). FigTree v1.4.2: A graphical viewer of phylogenetic trees. Available from <http://tree.bio.ed.ac.uk/software/figtree/>
- Rambaut A., Suchard M.A., Xie D., & Drummond A.J. (2014) Tracer v1.6: A visual tool for visualization and diagnostics of MCMC output. Available from <http://beast.bio.ed.ac.uk/Tracer>
- Reddy, S. (2008). Systematics and biogeography of the shrike-babblers (*Pteruthius*): species limits, molecular phylogenetics, and diversification patterns across southern Asia. *Molecular Phylogenetics and Evolution*. 47(1): 54–72. DOI: 10.1016/j.ympev.2008.01.014
- Robinson, A.J., Brezina, C.A., Parrish, R.R., Horstwood, M.S.A., Oo, N.W., et al. (2014). Large rivers and orogens: The evolution of the Yarlung Tsangpo-Irrawaddy system and the eastern Himalayan syntaxis. *Gondwana Research*. 26:112–121. DOI: 10.1016/j.gr.2013.07.002

- Rubin, B.E.R., Ree, R.H., Moreau, C.S. (2012). Inferring phylogenies from RAD sequence data. *PLOS ONE*. 7(4):e33394. DOI: 10.1371/journal.pone.0033394
- Sawamura, K., Laming, D.J. (1974). Sea-floor valleys in the Gulf of Thailand and quaternary sea-level changes. *CCOP Newsletter*. 1:23–27
- Searle, M.P. (2006). Role of the Red River Shear zone, Yunnan and Vietnam, in the continental extrusion of SE Asia. *Journal of the Geological Society, London*. 163(6):1025–1036. DOI: 10.1144/0016-76492005
- Takacs, Z., Morales, J.C., Geissmann, T., Melnick, D.J. (2005). A complete species-level phylogeny of the Hylobatidae based on mitochondrial *ND3-ND4* gene sequences. *Molecular Phylogenetics and Evolution*. 36:456-467. DOI: 10.1016/j.ympev.2005.03.032
- Tantrawatpan, C., Saijuntha, W., Pilab, W., Sakdakham, K., Pasorn, P., et al. (2011). Genetic differentiation among populations of *Brachytrupes portentosus* (Lichtenstein 1796) (Orthoptera: Gryllidae) in Thailand and the Lao PDR: the Mekong River as a biogeographic barrier. *Bulletin of Entomological Research*. 101:687–696. DOI: 10.1017/S000748531100023X
- Taylor, E.H. (1934). Zoological results of the third De Schauensee Siamese Expedition, Part III. Amphibians and Reptiles. *Proceedings of the Academy of Natural Sciences of Philadelphia*. 86:281–310
- Taylor, E.H. (1963). The lizards of Thailand. *Kansas University Scientific Bulletin*. 44:687–1077

- Tordoff, A.W., Bezuijen, M.R., Duckworth, J.W., Fellowes, J.R., Koenig, K. et al. (2012). Ecosystem Profile: Indo-Burma Biodiversity Hotspot (2011 Update). *Critical Ecosystem Partnership Fund*
- Vaidya, G., Lohman, D. J. and Meier, R. (2011). SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, 27(2): 171–180. DOI:10.1111/j.1096-0031.2010.00329.x
- Wallace, A.R. (1852). On the monkeys of the Amazon. *Proceedings of the Zoological Society of London*. 20:107-110
- Wang, H., Luo, X., Meng, S., Bei, Y., Song, T., et al. (2015). The phylogeography and population demography of the Yunnan Caecilian (*Ichthyophis bannanicus*): Massive rivers as barriers to gene flow. *PLOS ONE*. 10(4): e0125770. DOI: 10.1371/journal.pone.0125770
- Wang, W., McKay, B.D., Dai, C., Zhao, N., Zhang, R., et al. (2012). Glacial expansion and diversification of an East Asian montane bird, the green-backed tit (*Parus monticolus*). *Journal of Biogeography*. 40:1156–1169. DOI: 10.1111/jbi.12055
- Ward, F.K. (1921). The Mekong-Salween divide as a geographical barrier. *The Geographical Journal*. 58(1):49-56. DOI: 10.2307/1780720
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York
- Wilcox, T.P., Zwickl, D.J., Heath, T.A., Hillis, D.M. (2002). Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support.

Molecular Phylogenetics and Evolution 25(2002): 361–371. DOI: 10.1016/S1055-7903(02)00244-0

Wood Jr., P.L., Heinicke, M.P., Jackman, T.R., Bauer, A.M. (2012). Phylogeny of bent-toed geckos (*Cyrtodactylus*) reveals a west to east pattern of diversification. *Molecular Phylogenetics and Evolution*. 65(3):992–1003. DOI: 10.1016/j.ympev.2012.08.025

Woodruff, D.S. (2010). Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodiversity Conservation*. 19(4):919–941. DOI: 10.1007/s10531-010-9783-3

Workman, D.R. (1975). Tectonic evolution of Indochina. *Journal of the Geological Society of Indochina*. 1:3–19

Wu, Z.Y., Wu, S.G. (1996). A proposal for a new floristic kingdom (realm) – the E. Asiatic kingdom, its delimitation and characteristics. *Floristic characteristics and diversity of east Asian plants*

Xiang, H.F., Han, Z.J., Guo, S.M., Cheng, L.C., Zhang, W.X. (2004). Processing about quantitative study of large-scale strike-slip movement on Red River fault zone. *Advance in Earth Science*. 19:56-59. DOI: 10.1007/s11430-007-2037-x

Xie, X., Lewis, P.O., Fan, Y., Kuo, L., Chen, M. (2011). Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology*. 60(2):150-160. DOI: 10.1093/sysbio/syq085

- Yuan, Z.Y., Suwannapoom, C., Yan, F., Poyarkov Jr., N.A., Nguyen, S.N. (2016). Red River barrier and Pleistocene climatic fluctuations shaped the genetic structure of *Microhyla fissipes* complex (Anura: Microhylidae) in southern China and Indochina. *Current Zoology*. 62(6):531–543. DOI: 10.1093/cz/zow042
- Zhao, H., Wang, L., Yuan, J. (2007). Origin and time of Qiongzhou Strait. *Marine Geology and Quaternary Geology*
- Zhang, D.R., Chen, M.Y., Murphy, R.W., Che, J., Pang, J.F., et al. (2010a). Genealogy and palaeodrainage basins in Yunnan Province: phylogeography of the Yunnan spiny frog, *Nanorana yunnanensis* (Dicroglossidae). *Molecular Ecology*. 19:3406–3420. DOI: 10.1111/j.1365-294X.2010.04747.x
- Zhang, M.W., Rao, D.Q., Yang, J.X., Yu, G.H., Wilkinson, J.A. (2010b). Molecular phylogeography and population structure of a mid-elevation montane frog *Leptobrachium ailaonicum* in a fragmented habitat of southwest China. *Molecular Phylogenetics and Evolution*. 54(1):47-58. DOI: 10.1016/j.ympev.2009
- Zhu, F., Liu, Q., Che, J., Zhang, L., Guo, P. (2015). Molecular phylogeography of white-lipped tree viper (*Trimeresurus*; Viperidae). *Zoologica Scripta*. 45(3):252-262. DOI: 10.1111/zsc.12156
- Zhu, H. (2016). Biogeographical evidences help revealing the origin of Hainan Island. *PLOS ONE*. 11(4):e0151941. DOI: 10.1371/journal.pone.0151941

Chapter 2: Figures

Figure 2.1 Indochina and Study Sampling

Figure 2.2 Lineage Divergence in *Draco maculatus*

Figure 2.3 Species Richness and Lineage Divergence in *Draco*

Figure 2.4 Bayes Factor Results

Figure 2.5 Divergence Time Estimates

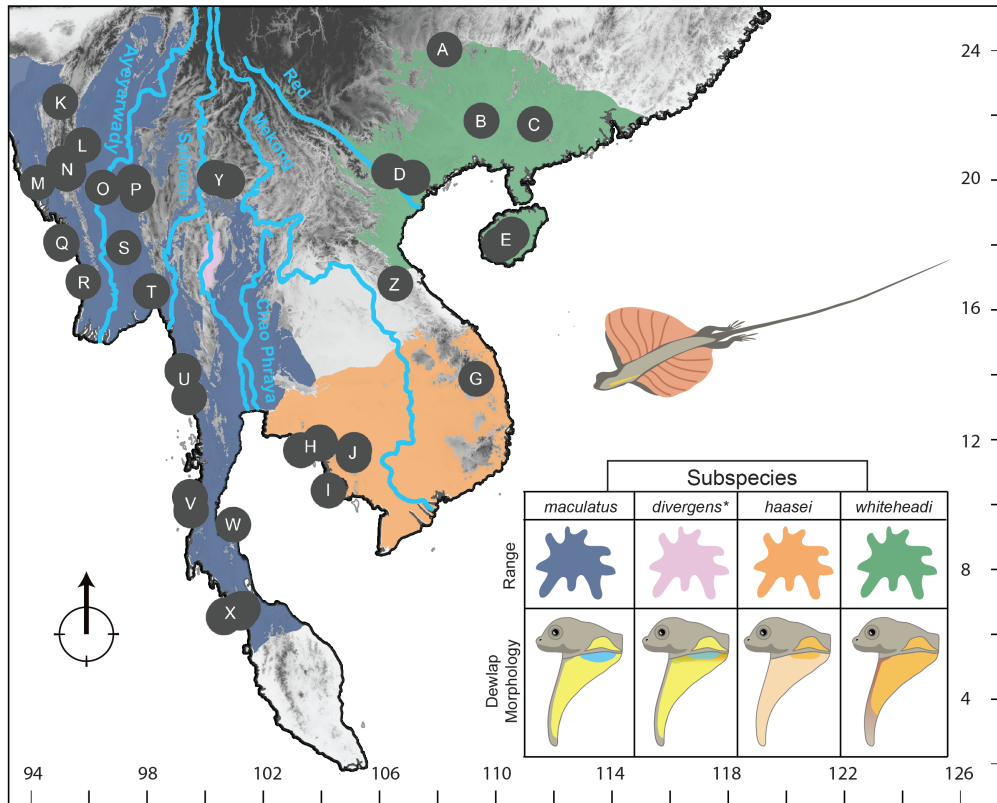


Figure 2.1: **Indochina and Study Sampling**

Map of Indochina with sample localities (grey circles, labeled alphabetically), geographic ranges of *Draco maculatus* subspecies (colored), and the major rivers within the region. Dewlap illustrations for each subspecies are based on diagnostic characters from live- animal and specimen examinations by Musters (1983) and Taylor (1963).

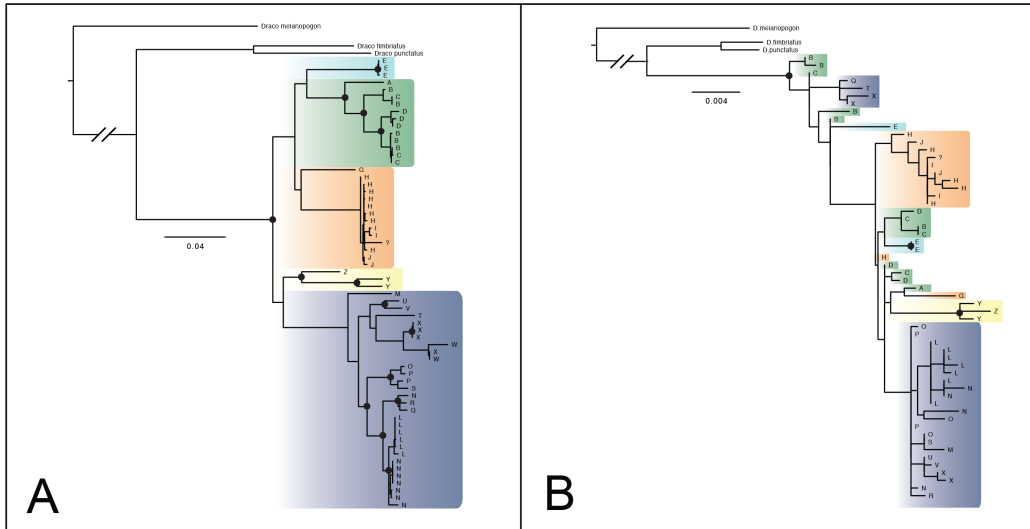


Figure 2.2: **Lineage Divergence in *Draco maculatus***

Maximum likelihood phylogenies for (A) mitochondrial sequence data (12S, 16S, ND2) and (B) nuclear sequence data (BDNF, CMOS, PNN). Filled circles indicate nodes with BV greater than or equal to 70. Tip letters correspond to sampling localities in Fig. 1. Colored lineages represent divergent clades of *Draco maculatus*: Navy blue = West, green = China-Vietnam Coastal Plain, light blue = Hainan Island, orange = Southeast, light yellow = Central.

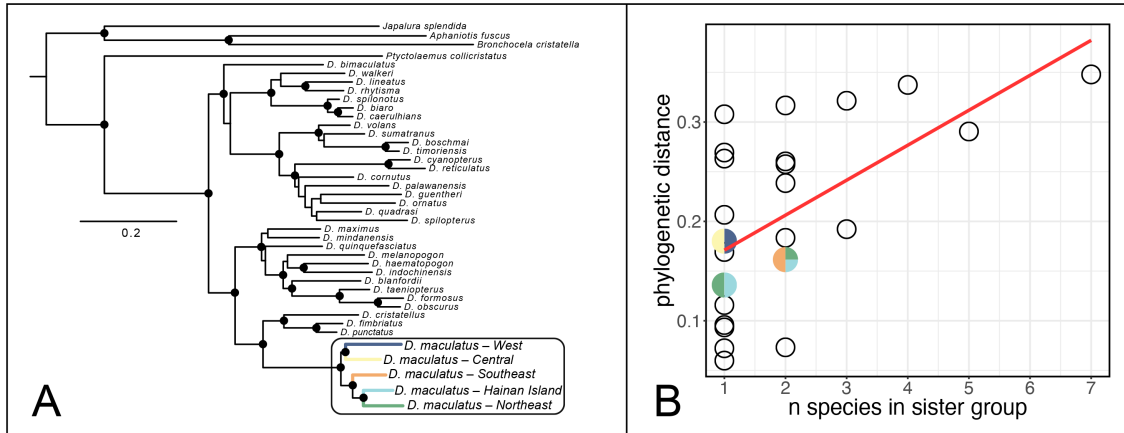


Figure 2.3: Species Richness and Lineage Divergence in *Draco*

Genus-wide analysis of phylogenetic distance using ND2 sequence data. (A) Gene tree. Filled circles at nodes indicate BV greater than or equal to 70. Colored lineages represent divergent clades of *Draco maculatus*: Navy blue = West, green = China-Vietnam Coastal Plain, light blue = Hainan Island, orange = Southeast, light yellow = Central. Each *D. maculatus* lineage is an average branch length for all samples pertaining to the respective clade (Sample sizes: West = 28, Central = 3, Southeast = 11, Hainan Island = 3, China-Vietnam Coastal Plain = 12). (B) Plot showing phylogenetic distance between *Draco* species and sister lineages. The x-axis groups species by the number of species in their sister lineage. We plotted the data in this manner to show that phylogenetic distance is positively associated with the number of species in the sister lineage, thus showing that the divergence between a species with a single sister species is not directly comparable to a species with three species in its sister lineage. Colored circles correspond to *D. maculatus* lineages in (A). The line of best fit is shown by the diagonal red line (Adjusted R² = 0.3021, p-value = 0.003)

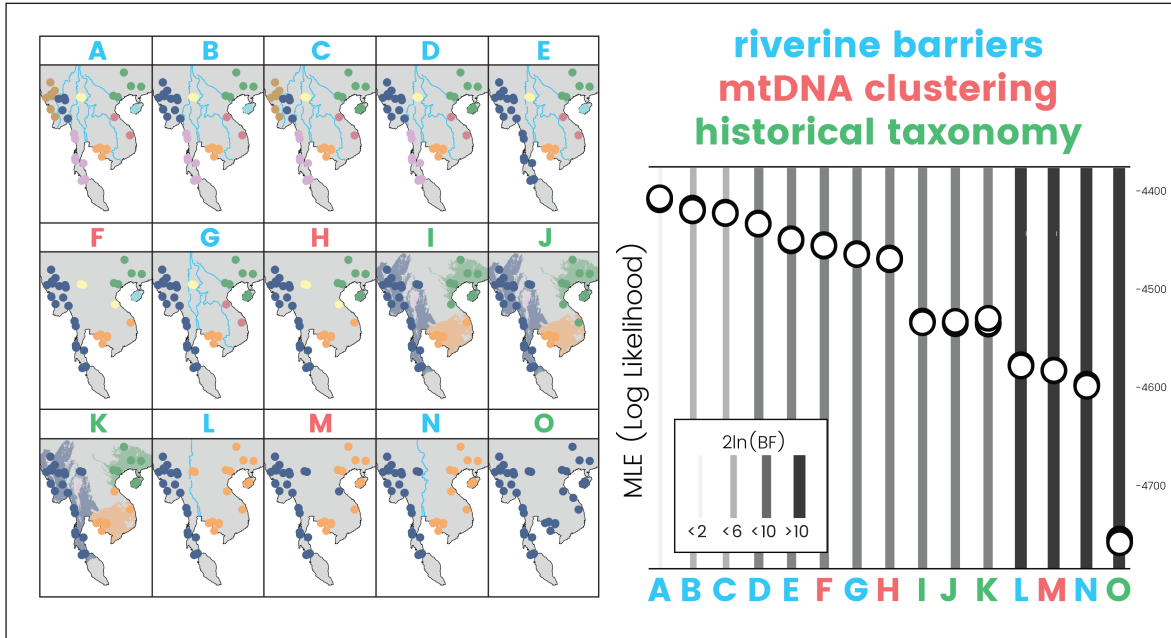


Figure 2.4: **Bayes Factor Results**

Experimental Design and Results from Bayes Factor Delimitation (BFD). Left: Models depicted in map form, with colored letters corresponding to model category (riverine barriers, mtDNA clustering, or historical taxonomy). Right: Plot for the estimated marginal log-likelihood scores of eight replicates. Model letters on x-axis correspond to those above each map. Vertical lines indicate $2\ln(\text{BF})$ value for model A compared to each other model.

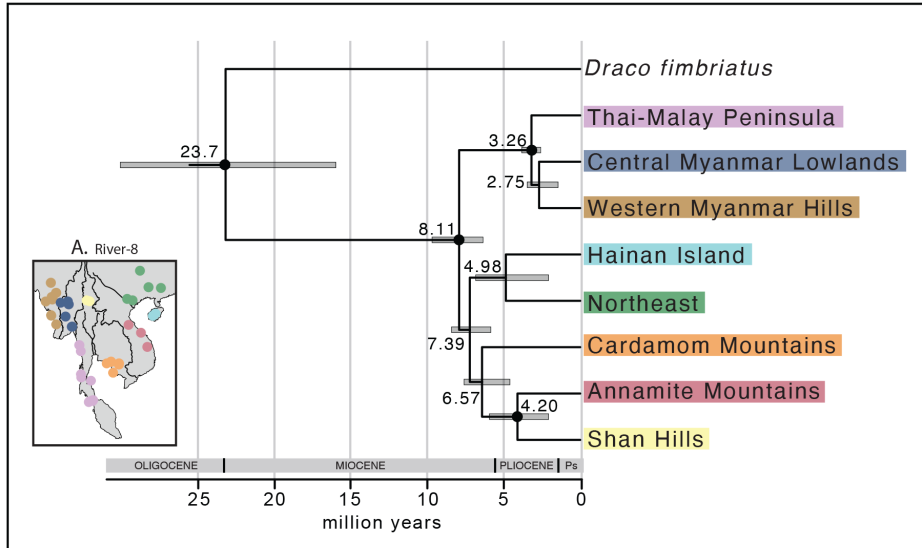


Figure 2.5: **Divergence Time Estimates**

Time-calibrated maximum clade credibility trees for the best-supported model (River-8 [Model A]), estimated from DNA sequence data of six loci. Filled circles indicate nodes greater than or equal to 0.95 BPP. Grey bars show the 95% HPD (Bayesian credible interval) for timing of divergence. Tree tip colors correspond to the localities shown on the map.

Chapter 2: Tables

Table 2.1: Summary statistics and selected models for each of the four loci sequenced: *12S*, *16S*, *ND2*, *BDNF*, *CMOS*, and *PNN*. (treating the mtDNA regions as a single locus). S = number of segregating sites, π = nucleotide diversity.

| Locus | Function | n | Nucleotide Sites | Variable Sites | Percent Variable | Site Model Selected | Haplotypes | π |
|--------------|----------------------|----|------------------|----------------|------------------|---------------------|------------|-------|
| <i>mtDNA</i> | Protein & RNA Coding | 53 | 2470 | 601 | 24% | TrN+G | 58 | 0.077 |
| <i>BDNF</i> | Protein Coding | 54 | 730 | 51 | 7% | K80+G | 42 | 0.007 |
| <i>CMOS</i> | Protein Coding | 54 | 422 | 29 | 7% | K80+G | 26 | 0.007 |
| <i>PNN</i> | Protein Coding | 54 | 689 | 44 | 6% | TrN+G | 30 | 0.007 |
| Total | – | 54 | 4311 | 704 | 16% | NA | NA | NA |

Table 2.2: Marginal-likelihood estimation (MLE) for each species hypothesis and pairwise Bayes factor delimitation (BFD) between each of the four tested hypotheses. BFD interpretation is as follows: $0 < 2\ln(\text{BF}) < 2$ is not more than a bare mention, $2 < 2\ln(\text{BF}) < 6$ is positive support, $6 < 2\ln(\text{BF}) < 10$ is strong support, and $10 < 2\ln(\text{BF})$ is decisive.

| Hypothesis | Type | N Lineages | Median MLE | 2ln BF (Hyp A) |
|------------|---------------------|------------|------------|----------------|
| A | riverine barriers | 8 | -4407.60 | - |
| B | riverine barriers | 7 | -4419.38 | 4.93 |
| C | riverine barriers | 7 | -4422.14 | 5.35 |
| D | riverine barriers | 6 | -4433.22 | 6.49 |
| E | riverine barriers | 6 | -4449.44 | 7.47 |
| F | mtDNA clustering | 5 | -4455.48 | 7.74 |
| G | riverine barriers | 5 | -4464.10 | 8.07 |
| H | mtDNA clustering | 4 | -4469.22 | 8.24 |
| I | historical taxonomy | 3 | -4532.99 | 9.66 |
| J | historical taxonomy | 3 | -4534.02 | 9.68 |
| K | historical taxonomy | 3 | -4534.75 | 9.69 |
| L | riverine barriers | 2 | -4577.66 | 10.27 |
| M | riverine barriers | 2 | -4582.74 | 10.33 |
| N | mtDNA clustering | 2 | -4597.62 | 10.49 |
| O | historical taxonomy | 1 | -4754.80 | 11.70 |

Table 2.3: Divergent lineages from the River-8 Hypothesis (Model A) along with the riverine barriers corresponding to phylogenetic breaks and relevant geographical history (references cited in section 4.1). Lineages include China-Vietnam Coastal Plain (CVCP), Hainan Island (HI), Central Myanmar Lowlands (CML), Western Myanmar Hills (WMH), and Thai-Malay Peninsula (TML) lineages, and the two major branches from the basal node of *Draco maculatus* (East & West). EDT = Estimated divergence time.

| Lineages | EDT | Barrier | Geographical relevance to est. divergence time |
|--------------------|-----------|--------------------------------------|--|
| East / West | 9.9 – 6.5 | Siam River (Chao Phraya River Basin) | Corresponds with Middle Miocene rapid river cutting for Siam River |
| (CVCP + HI) / West | 8.6 – 6.0 | Red River | Corresponds with extrusion event along the Song Hong fault zone |
| CVCP / HI | 7.0–2.1 | Qiongzhou Strait | Predates assumed strait formation, but reflects other studies |
| CML / WMH | 3.5 – 1.5 | Salween | Corresponds with river capture event (Mekong -> Salween) |

Chapter 2: Appendices

Appendix 2.A: Subspecies names (with description reference), distribution, and dewlap coloration based on previous studies of multiple individuals.

| Subspecies | Distribution | Dewlap Color |
|---|--|--|
| <i>Draco maculatus divergens</i> (Taylor, 1934) | Chiang Mai, Thailand | Dewlap dirty greenish with a dim blue spot (Taylor, 1963). Blue spot on base of gular pouch (Musters, 1983) |
| <i>Draco maculatus haasei</i> (Boettger, 1893) | Indochinese Peninsula (Eastern Thailand, Cambodia, Southern Vietnam) | Base of dewlap deep orange and lacking blue spot. (Taylor, 1963). Base of dewlap yellow/white and lacking blue spot (Musters, 1983). |
| <i>Draco maculatus maculatus</i> (Gray, 1845) | Western Indochina (Eastern India, Myanmar, Western & Southern Thailand, Peninsular Malaysia) | Bright blue spot on base of dewlap, with front and terminals pearl grey (Taylor, 1963). Blue spot at base of dewlap (Musters, 1983) |
| <i>Draco maculatus whiteheadi</i> (Boulenger, 1900) | Northeastern Indochina (Northern Vietnam, Hainan Island) | Dewlap blue at end, red behind base (Taylor, 1963). Base of gular pouch brownish (Musters, 1983). |

Appendix 2.B: List of samples with voucher ID (VID), locality ID from Fig. 1 (LID), country codes (C), latitude & longitude coordinates, and GenBank numbers for all sequences used in this study.

| VID | Taxon | LID | C | Lat | Long | ND2 | 12S | 16S | BDNF | CMOS | PNN |
|------------|------------------------------------|-----|-----|-------|--------|------------|-----|-----|------|------|-----|
| TNHC57874 | <i>Aphanotis fuscus</i> | - | MYS | 3.33 | 101.77 | AF288228.1 | - | - | - | - | - |
| TNHC56517 | <i>Bronchocela cristatella</i> | - | MYS | 1.72 | 110.33 | AF288229.1 | - | - | - | - | - |
| LSUMZ81212 | <i>Japalura splendida</i> | - | - | - | - | AF288230.1 | - | - | - | - | - |
| TNHC56803 | <i>Draco maximus</i> | - | MYS | 6.05 | 116.70 | AF288231.1 | - | - | - | - | - |
| TNHC56829 | <i>Draco quinquefasciatus</i> | - | MYS | 3.33 | 101.77 | AF288232.1 | - | - | - | - | - |
| TNHC58527 | <i>Draco spilopterus</i> | - | PHI | 9.22 | 123.57 | AF288240.1 | - | - | - | - | - |
| TNHC57786 | <i>Draco bimaculatus</i> | - | PHI | 11.80 | 125.27 | AF288241.1 | - | - | - | - | - |
| TNHC56531 | <i>Draco blanfordii</i> | - | MYS | 6.68 | 100.18 | AF288242.1 | - | - | - | - | - |
| ROM31987 | <i>Draco indochinensis</i> | - | VNM | 13.18 | 108.68 | AF288243.1 | - | - | - | - | - |
| TNHC56769 | <i>Draco cornutus</i> | - | MYS | 6.05 | 116.70 | AF288244.1 | - | - | - | - | - |
| TNHC56842 | <i>Draco cyanopterus</i> | - | PHI | 6.97 | 125.42 | AF288245.1 | - | - | - | - | - |
| TNHC56702 | <i>Draco reticulatus</i> | - | PHI | 9.83 | 124.14 | AF288247.1 | - | - | - | - | - |
| TNHC58848 | <i>Draco mindanensis</i> | - | PHI | 7.19 | 125.41 | AF288249.1 | - | - | - | - | - |
| TNHC56814 | <i>Draco obscurus</i> | - | MYS | 6.05 | 116.70 | AF288250.1 | - | - | - | - | - |
| TNHC56685 | <i>Draco taeniopterus</i> | - | MYS | 6.68 | 100.18 | AF288251.1 | - | - | - | - | - |
| TNHC55072 | <i>Draco ornatus</i> | - | PHI | 11.86 | 125.07 | AF288252.1 | - | - | - | - | - |
| TNHC56763 | <i>Draco cristatellus</i> | - | MYS | 1.73 | 110.33 | AF288255.1 | - | - | - | - | - |
| LSUMZ81446 | <i>Draco punctatus</i> | - | IDN | - | - | AF288257.1 | - | - | - | - | - |
| TNHC56847 | <i>Draco haematopogon</i> | - | MYS | 4.84 | 100.78 | AF288259.1 | - | - | - | - | - |
| TNHC58847 | <i>Draco guentheri</i> | - | PHI | 7.19 | 125.41 | AF288260.1 | - | - | - | - | - |
| TNHC55067 | <i>Draco quadrasi</i> | - | PHI | 13.37 | 121.06 | AF288261.1 | - | - | - | - | - |
| TNHC56719 | <i>Draco palawanensis</i> | - | PHI | 14.58 | 120.98 | AF288262.1 | - | - | - | - | - |
| TNHC56540 | <i>Draco formosus</i> | - | MYS | 3.33 | 101.77 | AF288263.1 | - | - | - | - | - |
| TNHC56728 | <i>Draco sumatranus</i> | - | MYS | 3.33 | 101.77 | AF288265.1 | - | - | - | - | - |
| LSUMZ81441 | <i>Draco volans</i> | - | IDN | -7.45 | 110.52 | AF288267.1 | - | - | - | - | - |
| WAM104530 | <i>Draco boschmai</i> | - | IDN | 8.67 | 120.81 | AF288269.1 | - | - | - | - | - |
| WAM107005 | <i>Draco timoriensis</i> | - | IDN | 10.17 | 123.60 | AF288275.1 | - | - | - | - | - |
| LSUMZ81223 | <i>Draco walkeri</i> | - | IDN | 2.64 | 120.18 | AF288276.1 | - | - | - | - | - |
| LSUMZ81270 | <i>Draco biaro</i> | - | IDN | 2.10 | 125.38 | AF288277.1 | - | - | - | - | - |
| LSUMZ81297 | <i>Draco bourouiensis</i> | - | IDN | 3.75 | 126.79 | AF288279.1 | - | - | - | - | - |
| LSUMZ81327 | <i>Draco rhytisma</i> | - | IDN | -1.59 | 123.22 | AF288280.1 | - | - | - | - | - |
| LSUMZ81307 | <i>Draco caerulhians</i> | - | IDN | 3.61 | 125.48 | AF288281.1 | - | - | - | - | - |
| LSUMZ81375 | <i>Draco spilonotus</i> | - | IDN | 1.43 | 124.98 | AF288282.1 | - | - | - | - | - |
| USNM559811 | <i>Ptyctolaemus collicristatus</i> | - | MMR | 21.37 | 93.98 | AY555837.1 | - | - | - | - | - |

| | | | | | | | | | | | |
|------------|------------------------|---|-----|-------|--------|----------|----------|----------|----------|----------|----------|
| CAS210160 | <i>Draco maculatus</i> | L | MMR | 22.32 | 94.48 | MT041888 | MT031857 | MT040978 | MK754266 | MK754325 | MT041831 |
| CAS210245 | <i>Draco maculatus</i> | L | MMR | 22.32 | 94.48 | MT041889 | MT031858 | MT040979 | MK754267 | MK754326 | MT041832 |
| CAS210502 | <i>Draco maculatus</i> | L | MMR | 22.32 | 94.47 | MT041890 | MT031859 | MT040980 | MK754268 | MK754327 | MT041833 |
| CAS214083 | <i>Draco maculatus</i> | O | MMR | 20.91 | 95.24 | MT041891 | MT031860 | MT040981 | MK754269 | MK754328 | MT041834 |
| CAS215259 | <i>Draco maculatus</i> | P | MMR | 20.70 | 96.51 | MT041892 | MT031861 | MT040982 | MK754270 | MK754329 | MT041835 |
| CAS215538 | <i>Draco maculatus</i> | L | MMR | 22.32 | 94.44 | MT041893 | MT031862 | MT040983 | MK754271 | MK754330 | MT041836 |
| CAS215634 | <i>Draco maculatus</i> | L | MMR | 22.32 | 94.49 | MT041894 | MT031863 | MT040984 | MK754272 | MK754331 | MT041837 |
| CAS215637 | <i>Draco maculatus</i> | L | MMR | 22.32 | 94.49 | MT041895 | MT031864 | MT040985 | MK754273 | MK754332 | MT041838 |
| CAS220002 | <i>Draco maculatus</i> | N | MMR | 21.38 | 93.97 | MT041896 | MT031865 | MT040986 | MK754274 | MK754333 | MT041839 |
| CAS220005 | <i>Draco maculatus</i> | N | MMR | 21.38 | 93.97 | MT041897 | MT031866 | MT040987 | MK754275 | MK754334 | MT041840 |
| CAS220006 | <i>Draco maculatus</i> | N | MMR | 21.38 | 93.97 | MT041898 | MT031867 | MT040988 | MK754276 | MK754335 | MT041841 |
| CAS220007 | <i>Draco maculatus</i> | N | MMR | 21.38 | 93.97 | MT041899 | MT031868 | MT040989 | MK754277 | MK754336 | MT041842 |
| CAS220018 | <i>Draco maculatus</i> | N | MMR | 21.37 | 93.98 | MT041900 | MT031869 | MT040990 | MK754278 | MK754337 | MT041843 |
| CAS220019 | <i>Draco maculatus</i> | N | MMR | 21.37 | 93.98 | MT041901 | MT031870 | MT040991 | MK754279 | MK754338 | MT041844 |
| CAS220050 | <i>Draco maculatus</i> | N | MMR | 21.33 | 93.92 | MT041902 | MT031871 | MT040992 | MK754280 | MK754339 | MT041845 |
| CAS220257 | <i>Draco maculatus</i> | Q | MMR | 17.70 | 94.65 | MT041903 | MT031872 | MT040993 | MK754281 | MK754340 | MT041846 |
| CAS221127 | <i>Draco maculatus</i> | M | MMR | 21.00 | 92.88 | MT041904 | MT031873 | MT040994 | MK754282 | MK754341 | MT041847 |
| CAS222144 | <i>Draco maculatus</i> | S | MMR | 18.90 | 96.06 | MT041905 | MT031874 | MT040995 | MK754283 | MK754342 | MT041848 |
| CAS228463 | <i>Draco maculatus</i> | P | MMR | 21.12 | 96.35 | MT041906 | MT031875 | MT040996 | MK754284 | MK754343 | MT041849 |
| CAS228472 | <i>Draco maculatus</i> | U | MMR | 13.84 | 98.45 | MT041907 | MT031876 | MT040997 | MK754285 | MK754344 | MT041850 |
| CAS228473 | <i>Draco maculatus</i> | Y | MMR | 21.20 | 99.76 | MT041908 | MT031877 | MT040998 | MK754286 | MK754345 | MT041851 |
| CAS228474 | <i>Draco maculatus</i> | Y | MMR | 21.32 | 99.30 | MT041909 | MT031878 | MT040999 | MK754287 | MK754346 | MT041852 |
| CAS228475 | <i>Draco maculatus</i> | V | MMR | 10.00 | 98.54 | MT041910 | MT031879 | MT041000 | MK754288 | MK754347 | MT041853 |
| CAS235050 | <i>Draco maculatus</i> | N | MMR | 21.69 | 93.80 | MT041911 | - | - | - | - | - |
| CAS235106 | <i>Draco maculatus</i> | N | MMR | 21.60 | 93.94 | MT041912 | MT031880 | MT041001 | MK754289 | MK754348 | MT041854 |
| CAS235961 | <i>Draco maculatus</i> | V | MMR | 10.46 | 98.50 | MT041913 | - | - | - | - | - |
| CAS239914 | <i>Draco maculatus</i> | Q | MMR | 19.03 | 93.81 | MT041914 | - | - | - | - | - |
| CAS239915 | <i>Draco maculatus</i> | Q | MMR | 19.03 | 93.81 | MT041915 | MT031881 | MT041002 | MK754290 | MK754349 | MT041855 |
| CAS239934 | <i>Draco maculatus</i> | Q | MMR | 18.95 | 93.83 | MT041916 | - | - | - | - | - |
| CAS243205 | <i>Draco maculatus</i> | K | MMR | 23.74 | 93.57 | MT041917 | - | - | - | - | - |
| CAS243219 | <i>Draco maculatus</i> | K | MMR | 23.77 | 93.57 | MT041918 | - | - | - | - | - |
| FMNH263343 | <i>Draco maculatus</i> | J | KHM | 11.33 | 104.07 | MT041919 | - | - | - | - | - |
| KU311487 | <i>Draco maculatus</i> | C | CHN | 22.78 | 111.05 | MT041920 | MT031892 | MT041013 | MK754298 | MK754359 | MT041861 |
| KU311482 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041921 | MT031888 | MT041009 | MK754294 | MK754353 | MT041858 |
| MVZ226483 | <i>Draco maculatus</i> | D | VNM | 21.45 | 105.64 | MT041922 | - | - | - | - | - |
| MVZ226484 | <i>Draco maculatus</i> | D | VNM | 21.45 | 105.64 | MT041923 | - | - | MK754318 | MK754376 | MT076080 |
| MVZ241447 | <i>Draco maculatus</i> | E | CHN | 19.13 | 109.96 | MT041924 | - | - | MK754319 | MK754377 | MT041881 |
| NCSM85190 | <i>Draco maculatus</i> | Z | LAO | 17.65 | 105.74 | MT041925 | MT031910 | MT041031 | MK754320 | MK754378 | MT041882 |

| | | | | | | | | | | | |
|------------|------------------------|---|-----|-------|--------|----------|----------|----------|----------|----------|----------|
| ROM35898 | <i>Draco maculatus</i> | D | VNM | 21.21 | 106.48 | MT041926 | MT031912 | MT041032 | MK754322 | MK754380 | MT041884 |
| ROM35899 | <i>Draco maculatus</i> | D | VNM | 21.21 | 106.48 | MT041927 | - | - | - | - | - |
| ROM47737 | <i>Draco maculatus</i> | F | VNM | ? | ? | MT041928 | - | - | - | - | - |
| ROM47738 | <i>Draco maculatus</i> | F | VNM | ? | ? | MT041929 | MT031913 | MT041033 | MK754323 | MK754381 | MT041885 |
| TNHC56576 | <i>Draco maculatus</i> | X | MYS | 6.63 | 100.21 | MT041930 | MT031886 | MT041007 | MK754292 | MK754351 | MT041886 |
| LSUMZ81826 | <i>Draco maculatus</i> | E | CHN | 18.91 | 109.69 | MT041931 | MT031914 | MT041034 | MK754324 | MK754382 | MT041887 |
| USNM587772 | <i>Draco maculatus</i> | T | MMR | 17.44 | 97.10 | MT041932 | - | - | - | - | - |
| ROM32033 | <i>Draco maculatus</i> | G | VNM | 14.34 | 108.48 | MT041933 | MT031911 | - | MK754321 | MK754379 | MT041883 |
| LSUHC5617 | <i>Draco punctatus</i> | - | MYS | 5.33 | 101.37 | MT041934 | MT031884 | MT041005 | MK754303 | - | MT041866 |
| ROM32032 | <i>Draco maculatus</i> | G | VNM | 14.34 | 108.48 | MT041935 | - | - | - | - | - |
| LSUHC4191 | <i>Draco maculatus</i> | E | CHN | 18.82 | 109.51 | MT041936 | - | - | - | - | - |
| LSUHC7852 | <i>Draco maculatus</i> | H | KHM | 12.26 | 103.00 | MT041937 | MT031906 | MT041027 | MK754311 | MK754369 | MT041875 |
| LSUHC7919 | <i>Draco maculatus</i> | H | KHM | 12.26 | 103.00 | MT041938 | MT031907 | MT041028 | MK754312 | MK754370 | MT041876 |
| LSUHC7851 | <i>Draco maculatus</i> | H | KHM | 12.26 | 103.00 | MT041939 | MT031905 | MT041026 | MK754310 | MK754368 | MT041874 |
| KU311488 | <i>Draco maculatus</i> | C | CHN | 22.78 | 111.05 | MT041940 | - | - | - | - | - |
| LSUHC7117 | <i>Draco maculatus</i> | X | MYS | 6.38 | 99.67 | MT041941 | - | - | - | - | - |
| MVZ236739 | <i>Draco maculatus</i> | E | CHN | 18.91 | 109.68 | MT041942 | - | - | - | - | - |
| KU311481 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041943 | - | - | - | - | - |
| KU311485 | <i>Draco maculatus</i> | C | CHN | 22.78 | 111.05 | MT041944 | MT031890 | MT041011 | MK754296 | MK754355 | MT041860 |
| KU351002 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041945 | - | - | - | - | - |
| KU351004 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041946 | - | - | - | - | - |
| KU311486 | <i>Draco maculatus</i> | C | CHN | 22.78 | 111.05 | MT041947 | MT031891 | MT041012 | MK754297 | MK754356 | - |
| KU311479 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041948 | MT031887 | MT041008 | MK754293 | MK754352 | MT041857 |
| KU311484 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041949 | - | - | - | - | - |
| LSUHC4190 | <i>Draco maculatus</i> | E | CHN | 18.82 | 109.51 | MT041950 | MT031898 | MT041019 | MK754302 | MK754361 | MT041865 |
| LSUHC8798 | <i>Draco maculatus</i> | X | MYS | 6.52 | 100.23 | MT041951 | - | - | - | - | - |
| LSUHC8800 | <i>Draco maculatus</i> | X | MYS | 6.63 | 100.18 | MT041952 | - | - | - | - | - |
| LSUHC10524 | <i>Draco maculatus</i> | J | KHM | 10.67 | 103.27 | MT041953 | - | - | - | - | - |
| USNM587770 | <i>Draco maculatus</i> | L | MMR | 22.32 | 94.49 | MT041954 | - | - | - | - | - |
| LSUHC9426 | <i>Draco maculatus</i> | X | MYS | 6.43 | 99.71 | MT041955 | - | - | - | - | - |
| CAS235962 | <i>Draco maculatus</i> | V | MMR | 10.44 | 98.50 | MT041956 | - | - | - | - | - |
| CAS239951 | <i>Draco maculatus</i> | Q | MMR | 18.96 | 93.84 | MT041957 | - | - | - | - | - |
| CAS240511 | <i>Draco maculatus</i> | T | MMR | 17.42 | 97.05 | MT041958 | - | - | - | - | - |
| CAS240615 | <i>Draco maculatus</i> | T | MMR | 17.43 | 97.10 | MT041959 | MT031882 | MT041003 | MK754291 | MK754350 | MT041856 |
| CAS243212 | <i>Draco maculatus</i> | K | MMR | 23.76 | 93.55 | MT041960 | - | - | - | - | - |
| CAS247990 | <i>Draco maculatus</i> | U | MMR | 14.74 | 98.20 | MT041961 | - | - | - | - | - |
| MVZ226481 | <i>Draco maculatus</i> | D | VNM | 21.45 | 105.64 | MT041962 | - | - | - | - | - |
| MVZ226482 | <i>Draco maculatus</i> | D | VNM | 21.45 | 105.64 | MT041963 | - | - | MK754317 | MK754375 | - |
| ROM35902 | <i>Draco maculatus</i> | D | VNM | 21.21 | 106.48 | MT041964 | - | - | - | - | - |
| KU351001 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041965 | - | - | - | - | - |

| | | | | | | | | | | | |
|------------|--------------------------|---|-----|-------|--------|----------|----------|----------|----------|----------|----------|
| KU351003 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041966 | MT031894 | MT041015 | MK754300 | MK754358 | MT041863 |
| KU351005 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041967 | MT031895 | MT041016 | MK754301 | MK754360 | MT041864 |
| LSUHC6791 | <i>Draco maculatus</i> | X | MYS | 6.37 | 99.67 | MT041968 | - | - | - | - | - |
| LSUHC6824 | <i>Draco maculatus</i> | X | MYS | 6.37 | 99.67 | MT041969 | - | - | - | - | - |
| LSUHC6825 | <i>Draco maculatus</i> | X | MYS | 6.37 | 99.67 | MT041970 | MT031900 | MT041021 | MK754305 | MK754362 | MT041868 |
| LSUHC6827 | <i>Draco maculatus</i> | X | MYS | 6.37 | 99.67 | MT041971 | - | - | - | - | - |
| LSUHC7103 | <i>Draco maculatus</i> | X | MYS | 6.38 | 99.67 | MT041972 | - | - | - | - | - |
| LSUHC7104 | <i>Draco maculatus</i> | X | MYS | 6.38 | 99.67 | MT041973 | - | - | - | - | - |
| LSUHC7105 | <i>Draco maculatus</i> | X | MYS | 6.38 | 99.67 | MT041974 | MT031901 | MT041022 | MK754306 | MK754363 | MT041869 |
| LSUHC7343 | <i>Draco maculatus</i> | I | KHM | 12.03 | 104.17 | MT041975 | MT031903 | MT041024 | MK754308 | MK754365 | MT041871 |
| LSUHC7411 | <i>Draco maculatus</i> | I | KHM | 12.03 | 104.17 | MT041976 | - | - | - | - | - |
| LSUHC7613 | <i>Draco fimbriatus</i> | - | MYS | 6.38 | 99.67 | MT041977 | MT031883 | MT041004 | MK754309 | MK754366 | MT041872 |
| LSUHC10305 | <i>Draco melanopogon</i> | - | MYS | 5.98 | 100.95 | MT041978 | MT031885 | MT041006 | MK754314 | MK754372 | MT041878 |
| USNM587773 | <i>Draco maculatus</i> | T | MMR | 17.44 | 97.10 | MT041979 | - | - | - | - | - |
| KU311480 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041980 | - | - | - | - | - |
| KU311483 | <i>Draco maculatus</i> | B | CHN | 23.01 | 109.10 | MT041981 | MT031889 | MT041010 | MK754295 | MK754354 | MT041859 |
| LSUHC7321 | <i>Draco maculatus</i> | I | KHM | 11.86 | 104.15 | MT041982 | MT031902 | MT041023 | MK754307 | MK754364 | MT041870 |
| LSUHC8981 | <i>Draco maculatus</i> | X | MYS | 6.66 | 100.32 | MT041983 | - | - | - | - | - |
| LSUHC9936 | <i>Draco melanopogon</i> | - | MYS | 2.04 | 103.56 | MT041984 | - | - | - | - | - |
| KU312113 | <i>Draco maculatus</i> | A | CHN | 25.48 | 107.88 | MT041985 | MT031893 | MT041014 | MK754299 | MK754357 | MT041862 |
| LSUHC6823 | <i>Draco maculatus</i> | X | MYS | 6.37 | 99.67 | MT041986 | MT031899 | MT041020 | MK754304 | - | MT041867 |
| LSUHC6826 | <i>Draco maculatus</i> | X | MYS | 6.37 | 99.67 | MT041987 | - | - | - | - | - |
| LSUHC6828 | <i>Draco maculatus</i> | X | MYS | 6.37 | 99.67 | MT041988 | - | - | - | - | - |
| LSUHC7322 | <i>Draco maculatus</i> | I | KHM | 11.86 | 104.15 | MT041989 | - | - | - | - | - |
| LSUHC7342 | <i>Draco maculatus</i> | I | KHM | 12.03 | 104.17 | MT041990 | - | - | - | - | - |
| LSUHC7344 | <i>Draco maculatus</i> | I | KHM | 12.03 | 104.17 | MT041991 | - | - | - | - | - |
| LSUHC7490 | <i>Draco maculatus</i> | X | MYS | 6.36 | 99.67 | MT041992 | - | - | - | - | - |
| LSUHC7849 | <i>Draco maculatus</i> | H | KHM | 12.26 | 103.00 | MT041993 | MT031904 | MT041025 | - | MK754367 | MT041873 |
| LSUHC8411 | <i>Draco maculatus</i> | H | KHM | 12.26 | 103.00 | MT041994 | MT031908 | MT041029 | MK754313 | MK754371 | - |
| LSUHC8980 | <i>Draco maculatus</i> | X | MYS | 6.66 | 100.32 | MT041995 | - | - | - | - | - |
| LSUHC10525 | <i>Draco maculatus</i> | J | KHM | 10.67 | 103.27 | MT041996 | MT031896 | MT041017 | MK754315 | MK754373 | MT041879 |
| LSUHC10526 | <i>Draco maculatus</i> | J | KHM | 10.67 | 103.27 | MT041997 | MT031897 | MT041018 | MK754316 | MK754374 | MT041880 |
| AB023727 | <i>Draco maculatus</i> | H | THA | 12.05 | 102.33 | NA | AB023727 | - | - | - | - |
| AB023739 | <i>Draco maculatus</i> | H | THA | 12.05 | 102.33 | NA | - | AB023739 | - | - | - |
| AB023758 | <i>Draco maculatus</i> | W | THA | 9.52 | 99.99 | NA | - | AB023758 | - | - | - |
| AB023759 | <i>Draco maculatus</i> | H | THA | 12.05 | 102.33 | NA | AB023759 | - | - | - | - |
| LSUHC9304 | <i>Draco maculatus</i> | H | KHM | 12.31 | 102.99 | NA | MT031909 | MT041030 | - | - | MT041877 |

Appendix 2.C: PCR conditions for each of the six loci used in this study

| Locus | Primers | Sequence | Annealing Temp | Reference |
|-------|-----------------|------------------------------------|----------------|-----------------------|
| 12S | L10091 | AAACTGGGATTAGATACCCCACTAT | 55 | Kocher et al., 1989 |
| | H1478 | GAGGGTGACGGGCGGTGTGT | | |
| 16S | L2606 | CTGACCGTGCAAAGGTAGCGTAATCACT | 50 | Hass et al., 1993 |
| | H3056 | CTCCGGTCTGAACTCAGATCACGTAGG | | |
| ND2 | L4437a (Metf.6) | AAGCTTTCGGGCCCATACC | 50 | Macey et al., 1997 |
| | ALAr.2m* | AAAGTGTCTGAGTTGCATTTCRG | | |
| BDNF | BDNF-F | GACCATCCTTTTCCTKACTATGGTTATTTCACTT | 56 | Townsend et al., 2008 |
| | BDNF-R | CTATCTTCCCCTTTTAATGGTCAGTGACAAAC | | |
| CMOS | G73 | GCGGTAAGCAGGTGAAGAAA | 54 | Saint et al., 1998 |
| | G74 | TGAGCATCCAAAGTCTCCAATC | | |
| PNN | PNN_F1 | TTTGCAGARCARATAAAAYAAAATGGA | 50.5 | Townsend et al., 2008 |
| | PNN_R1 | AACGCCTTTTGTCTTTCCTGTCTGATT | | |

* Primer modified from Macey et al. (1997); see McGuire and Kiew, 2001.

Hass, C.A., Hedges, S.B., Maxson, L.R. (1993). Molecular insights into the relationships and biogeography of West Indian Anoline lizards. *Biochemical Systematics and Ecology*. 21(1):97-114. DOI: 10.1016/0305-1978(93)90015-J

Kocher T.D., Thomas W.K., Meyer A., Edwards S.V., Pääbo S., et al. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *PNAS*. 86:6196–6200. DOI: 10.1073/pnas/86.16.6196

Macey, J.R., Schulte II, J.A., Larson, A., Fang, Z., Wang, Y., Tuniyev, B.S., Papenfuss, T.J. (1998). Phylogenetic Relationships of Toads in the *Bufo bufo* Species Group from the Eastern Escarpment of the Tibetan Plateau: A Case of Vicariance and Dispersal. *Molecular Phylogenetics and Evolution*. 9(1):80-87. DOI: 10.1006/mpev.1997.0440

McGuire, J.A., Kiew, B.H. (2001). Phylogenetic systematics of Southeast Asian flying lizards (Iguania: Agamidae: Draco) as inferred from mitochondrial DNA sequence data. *Biological Journal of the Linnean Society*. 72(2):203-229. DOI: 10.1006./bjl.2000.0487

Saint, K.M., Austin, C.C., Donnellan, S.C., Hutchinson, M.N. (1998). C-mos, a nuclear marker useful for squamate phylogenetic analysis. *Molecular Phylogenetics and Evolution*. 10(2):259-63. DOI: 10.1006/mpev/1998.0515

Townsend, T.M., Alegre, R.E., Kelley, S.T., Wiens, J.J., Reeder, T.W. (2008). Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: An example from squamate reptiles. *Molecular Phylogenetics and Evolution*. 47(1)129-42. DOI: 10.1016/j.ympv.2008.01.008

Appendix 2.D: Summary statistics and selected models for each of the three mtDNA regions sequenced: *12S*, *16S*, and *ND2*. The clade column indicates the monophyletic groups from the full concatenated phylogeny.

| Locus | Function | n | Nucleotide Sites | Variable Sites | Percent Variable | Site Model Selected | Haplotypes | π |
|------------|----------------|----|------------------|----------------|------------------|---------------------|------------|-------|
| <i>ND2</i> | Protein Coding | 53 | 1032 | 545 | 53% | TrN+I+G | 46 | 0.113 |
| <i>12S</i> | rRNA | 57 | 439 | 221 | 50% | HKY+G | 33 | 0.035 |
| <i>16S</i> | rRNA | 56 | 999 | 141 | 14% | HKY+G | 43 | 0.042 |

Appendix 2.E: Sister pairs and phylogenetic distance from *ND2* ML gene tree (Fig. 7A).

Accession numbers for each species can be found in Appendix B.

| Sister Pair | Phylogenetic Distance |
|---|------------------------------|
| D.biaro-D.caerulhians | 0.0599 |
| D.boschmai-D.timoriensis | 0.0723 |
| D.spilonotus-(D.biaro-D.caerulhians) | 0.0733 |
| D.fimbriatus-D.punctatus | 0.0933 |
| D.formosus-D.obscurus | 0.0957 |
| D.cyanopterus-D.reticulatus | 0.1158 |
| D.maculatus.ChinaVietnamCoastalPlain- D.maculatus.HainanIsland | 0.1362 |
| D.maculatus.Southeast- (D.maculatus.ChinaVietnamCoastalPlain- D.maculatus.HainanIsland) | 0.1618 |
| D.lineatus-D.rhytisma | 0.1692 |
| D.maculatus.West-D.maculatus.Central | 0.1795 |
| D.taeniopterus-(D.formosus-D.obscurus) | 0.1835 |
| D.volans-(D.sumatranus-(D.boschmai-D.timoriensis)) | 0.1922 |
| D.maximus-D.mindanensis | 0.2066 |
| D.sumatranus-(D.boschmai-D.timoriensis) | 0.2386 |
| D.cristatellus-(D.fimbriatus/D.punctatus) | 0.2576 |
| D.walkerii-(D.lineatus-D.rhytisma) | 0.2604 |
| D.haematopogon-D.indochinensis | 0.2632 |
| D.quadrasi-D.spilopterus | 0.2693 |
| D.cornutus-(D.palawanensis-((D.ornatus-D.guentheri)- (D.quadrasi-D.spilopterus))) | 0.2905 |
| D.ornatus-D.guentheri | 0.3078 |
| D.melanopogon-(D.haematopogon-D.indochinensis) | 0.3167 |
| D.blanfordii-(D.taeniopterus-(D.formosus-D.obscurus)) | 0.3214 |
| D.palawanensis-((D.ornatus-D.guentheri)-(D.quadrasi- D.spilopterus)) | 0.3374 |
| D.quinquefasciatus-((D.melanopogon-(D.haematopogon- D.indochinensis))-(D.blanfordii-(D.taeniopterus- (D.formosus-D.obscurus)))) | 0.3479 |
| D.bimaculatus-((D.volans-(D.sumatranus-(D.boschmai- D.timoriensis)))-(D.cyanopterus-D.reticulatus)- (D.cornutus-(D.palawanensis-((D.ornatus-D.guentheri)- (D.quadrasi-D.spilopterus))))-(D.spilonotus-(D.biaro- D.caerulhians))-(D.walkerii-(D.lineatus-D.rhytisma))) | 0.4996 |

Appendix 2.F: Replicate runs for each BFD hypothesis. Table indicates hypothesis abbreviation, rationale, species number (n), CPU Time (seconds), and log marginal likelihood estimation (MLE) values.

| Hypothesis | Rationale | n | Replicate | CPU Time | MLE |
|------------|------------------------------|---|-----------|----------|--------------|
| A | Historical Aquatic Geography | 8 | rep1 | 214160 | -4597.617158 |
| | | | rep2 | 208372 | -4597.720615 |
| | | | rep3 | 209697 | -4596.598447 |
| | | | rep4 | 374806 | -4597.029853 |
| | | | rep5 | 377002 | -4596.56537 |
| B | Historical Aquatic Geography | 7 | rep1 | 395097 | -4593.675208 |
| | | | rep2 | 374383 | -4597.530861 |
| | | | rep3 | 213651 | -4596.879477 |
| | | | rep4 | 372869 | -4596.808071 |
| | | | rep5 | 391229 | -4597.253378 |
| C | Historical Aquatic Geography | 7 | rep1 | 387899 | -4597.252168 |
| | | | rep2 | 201704 | -4597.545547 |
| | | | rep3 | 200302 | -4598.237985 |
| | | | rep4 | 211689 | -4598.834807 |
| | | | rep5 | 214806 | -4594.887357 |
| D | Historical Aquatic Geography | 6 | rep1 | 214027 | -4577.645098 |
| | | | rep2 | 375033 | -4577.089999 |
| | | | rep3 | 212527 | -4577.662676 |
| | | | rep4 | 214236 | -4579.212076 |
| | | | rep5 | 214156 | -4578.1297 |
| E | Historical Aquatic Geography | 6 | rep1 | 377698 | -4576.574852 |
| | | | rep2 | 378138 | -4577.888589 |
| | | | rep3 | 394398 | -4576.788208 |
| | | | rep4 | 214561 | -4577.698038 |
| | | | rep5 | 391686 | -4577.3389 |
| F | mtDNA Molecular Markers | 5 | rep1 | 390011 | -4578.122939 |
| | | | rep2 | 360315 | -4577.419931 |
| | | | rep3 | 203948 | -4578.141564 |
| | | | rep4 | 204505 | -4578.791566 |
| | | | rep5 | 200797 | -4576.25828 |
| G | Historical Aquatic Geography | 5 | rep1 | 371081 | -4453.935685 |
| | | | rep2 | 211816 | -4455.739569 |
| | | | rep3 | 374055 | -4456.014235 |
| | | | rep4 | 392188 | -4453.660578 |
| | | | rep5 | 386381 | -4454.551731 |

| | | | | | |
|---|------------------------------|---|------|--------|--------------|
| H | mtDNA Molecular Markers | 4 | rep1 | 211365 | -4455.272653 |
| | | | rep2 | 387163 | -4456.05453 |
| | | | rep3 | 390064 | -4453.251308 |
| | | | rep4 | 355560 | -4455.483302 |
| | | | rep5 | 201861 | -4456.411531 |
| I | Current taxonomy | 3 | rep1 | 201865 | -4455.606074 |
| | | | rep2 | 198743 | -4455.016161 |
| | | | rep3 | 372704 | -4457.68514 |
| | | | rep4 | 424806 | -4460.945711 |
| | | | rep5 | 212183 | -4469.06854 |
| J | Current taxonomy | 3 | rep1 | 373583 | -4469.498338 |
| | | | rep2 | 388249 | -4468.147914 |
| | | | rep3 | 388132 | -4469.94697 |
| | | | rep4 | 211908 | -4464.400202 |
| | | | rep5 | 386660 | -4469.556508 |
| K | Current taxonomy | 3 | rep1 | 215004 | -4467.895765 |
| | | | rep2 | 199973 | -4468.577431 |
| | | | rep3 | 202725 | -4469.812417 |
| | | | rep4 | 199160 | -4466.499783 |
| | | | rep5 | 211592 | -4469.604985 |
| L | Historical Aquatic Geography | 2 | rep1 | 210705 | -4469.217751 |
| | | | rep2 | 376643 | -4583.611463 |
| | | | rep3 | 374253 | -4582.358616 |
| | | | rep4 | 374253 | -4582.358616 |
| | | | rep5 | 214682 | -4581.90713 |
| M | mtDNA Molecular Markers | 2 | rep1 | 378195 | -4581.578355 |
| | | | rep2 | 377165 | -4581.859893 |
| | | | rep3 | 392489 | -4581.971004 |
| | | | rep4 | 394023 | -4582.744517 |
| | | | rep5 | 214062 | -4583.13061 |
| N | Historical Aquatic Geography | 2 | rep1 | 390845 | -4581.705425 |
| | | | rep2 | 200813 | -4583.677188 |
| | | | rep3 | 201189 | -4582.940865 |
| | | | rep4 | 372319 | -4582.897457 |
| | | | rep5 | 200628 | -4582.743971 |
| O | Null | 1 | rep1 | 213951 | -4582.286366 |
| | | | rep2 | 213089 | -4581.425775 |
| | | | rep3 | 391542 | -4757.010739 |
| | | | rep4 | 385114 | -4754.391687 |
| | | | rep5 | 231071 | -4756.783417 |

Appendix 2.G: Pairwise Maximum Composite Likelihood values for *ND2* region used in this study, with average likelihood values for each known geological event (10, 9, 3.5, 2.5, 1.5) listed beneath columns. Data and dates for geological events described in Macey et al (1998).

| <u>Species 1</u> | <u>Species 2</u> | <u>MAXCOMLIK</u> | <u>10</u> | <u>9</u> | <u>3.5</u> | <u>2.5</u> | <u>1.5</u> |
|------------------|------------------|------------------|-----------|----------|------------|------------|------------|
| P.himalayana | P.lehmanni | 0.159 | NA | NA | NA | NA | NA |
| P.himalayana | P.lehmanni | 0.165 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0.058 | NA | NA | NA | NA | NA |
| P.himalayana | P.lehmanni | 0.165 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0.058 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.himalayana | P.lehmanni | 0.165 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0.058 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.himalayana | P.lehmanni | 0.165 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0.058 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.himalayana | P.lehmanni | 0.165 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0.058 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.himalayana | P.lehmanni | 0.165 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0.058 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.lehmanni | P.lehmanni | 0 | NA | NA | NA | NA | NA |
| P.himalayana | P.microlepis | 0.161 | 0.161 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.192 | 0.192 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.himalayana | P.microlepis | 0.161 | 0.161 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.192 | 0.192 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0.189 | 0.189 | NA | NA | NA | NA |
| P.lehmanni | P.microlepis | 0 | NA | NA | NA | NA | NA |
| P.himalayana | P.erythrogaster | 0.164 | 0.164 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.193 | 0.193 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.196 | 0.196 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.196 | 0.196 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.196 | 0.196 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.196 | 0.196 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.196 | 0.196 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.196 | 0.196 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.196 | 0.196 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.196 | 0.196 | NA | NA | NA | NA |
| P.lehmanni | P.erythrogaster | 0.196 | 0.196 | NA | NA | NA | NA |
| P.microlepis | P.erythrogaster | 0.151 | NA | 0.151 | NA | NA | NA |
| P.microlepis | P.erythrogaster | 0.151 | NA | 0.151 | NA | NA | NA |

| | | | | | | | |
|-------------------|------------------|-------|-------|-------|-------|-------|----|
| P.lehmanni | SouthCaspian | 0.202 | 0.202 | NA | NA | NA | NA |
| P.lehmanni | SouthCaspian | 0.202 | 0.202 | NA | NA | NA | NA |
| P.lehmanni | SouthCaspian | 0.202 | 0.202 | NA | NA | NA | NA |
| P.microlepis | SouthCaspian | 0.163 | NA | 0.163 | NA | NA | NA |
| P.microlepis | SouthCaspian | 0.163 | NA | 0.163 | NA | NA | NA |
| P.erythrogaster | SouthCaspian | 0.059 | NA | NA | 0.059 | NA | NA |
| P.erythrogaster | SouthCaspian | 0.065 | NA | NA | 0.065 | NA | NA |
| CaucasusMountains | SouthCaspian | 0.034 | NA | NA | NA | 0.034 | NA |
| CaucasusMountains | SouthCaspian | 0.033 | NA | NA | NA | 0.033 | NA |
| CaucasusMountains | SouthCaspian | 0.034 | NA | NA | NA | 0.034 | NA |
| P.himalayana | Balkhan | 0.165 | 0.165 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.198 | 0.198 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.microlepis | Balkhan | 0.16 | NA | 0.16 | NA | NA | NA |
| P.microlepis | Balkhan | 0.16 | NA | 0.16 | NA | NA | NA |
| P.erythrogaster | Balkhan | 0.057 | NA | NA | 0.057 | NA | NA |
| P.erythrogaster | Balkhan | 0.06 | NA | NA | 0.06 | NA | NA |
| CaucasusMountains | Balkhan | 0.032 | NA | NA | NA | 0.032 | NA |
| CaucasusMountains | Balkhan | 0.029 | NA | NA | NA | 0.029 | NA |
| CaucasusMountains | Balkhan | 0.03 | NA | NA | NA | 0.03 | NA |
| SouthCaspian | Balkhan | 0.023 | NA | NA | NA | NA | NA |
| P.himalayana | Balkhan | 0.165 | 0.165 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.198 | 0.198 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | Balkhan | 0.2 | 0.2 | NA | NA | NA | NA |
| P.microlepis | Balkhan | 0.16 | NA | 0.16 | NA | NA | NA |
| P.microlepis | Balkhan | 0.16 | NA | 0.16 | NA | NA | NA |
| P.erythrogaster | Balkhan | 0.057 | NA | NA | 0.057 | NA | NA |
| P.erythrogaster | Balkhan | 0.06 | NA | NA | 0.06 | NA | NA |
| CaucasusMountains | Balkhan | 0.032 | NA | NA | NA | 0.032 | NA |
| CaucasusMountains | Balkhan | 0.029 | NA | NA | NA | 0.029 | NA |
| CaucasusMountains | Balkhan | 0.03 | NA | NA | NA | 0.03 | NA |
| SouthCaspian | Balkhan | 0.023 | NA | NA | NA | NA | NA |
| Balkhan | Balkhan | 0 | NA | NA | NA | NA | NA |
| P.himalayana | CaspianKopetDagh | 0.17 | 0.17 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.198 | 0.198 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.microlepis | CaspianKopetDagh | 0.162 | NA | 0.162 | NA | NA | NA |
| P.microlepis | CaspianKopetDagh | 0.162 | NA | 0.162 | NA | NA | NA |

| | | | | | | | |
|-------------------|------------------|-------------|---------|---------|---------|---------|---------|
| P.erythrogaster | CaspianKopetDagh | 0.068 | NA | NA | 0.068 | NA | NA |
| P.erythrogaster | CaspianKopetDagh | 0.074 | NA | NA | 0.074 | NA | NA |
| CaucasusMountains | CaspianKopetDagh | 0.048 | NA | NA | NA | 0.048 | NA |
| CaucasusMountains | CaspianKopetDagh | 0.044 | NA | NA | NA | 0.044 | NA |
| CaucasusMountains | CaspianKopetDagh | 0.045 | NA | NA | NA | 0.045 | NA |
| SouthCaspian | CaspianKopetDagh | 0.036 | NA | NA | NA | NA | NA |
| Balkhan | CaspianKopetDagh | 0.024 | NA | NA | NA | NA | 0.024 |
| Balkhan | CaspianKopetDagh | 0.024 | NA | NA | NA | NA | 0.024 |
| P.himalayana | CaspianKopetDagh | 0.169 | 0.169 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.2 | 0.2 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.207 | 0.207 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.207 | 0.207 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.207 | 0.207 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.207 | 0.207 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.207 | 0.207 | NA | NA | NA | NA |
| P.microlepis | CaspianKopetDagh | 0.16 | NA | 0.16 | NA | NA | NA |
| P.microlepis | CaspianKopetDagh | 0.16 | NA | 0.16 | NA | NA | NA |
| P.erythrogaster | CaspianKopetDagh | 0.066 | NA | NA | 0.066 | NA | NA |
| P.erythrogaster | CaspianKopetDagh | 0.072 | NA | NA | 0.072 | NA | NA |
| CaucasusMountains | CaspianKopetDagh | 0.045 | NA | NA | NA | 0.045 | NA |
| CaucasusMountains | CaspianKopetDagh | 0.042 | NA | NA | NA | 0.042 | NA |
| CaucasusMountains | CaspianKopetDagh | 0.043 | NA | NA | NA | 0.043 | NA |
| SouthCaspian | CaspianKopetDagh | 0.034 | NA | NA | NA | NA | NA |
| Balkhan | CaspianKopetDagh | 0.022 | NA | NA | NA | NA | 0.022 |
| Balkhan | CaspianKopetDagh | 0.022 | NA | NA | NA | NA | 0.022 |
| CaspianKopetDagh | CaspianKopetDagh | 0.002 | NA | NA | NA | NA | NA |
| P.himalayana | CaspianKopetDagh | 0.167 | 0.167 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.197 | 0.197 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.lehmanni | CaspianKopetDagh | 0.205 | 0.205 | NA | NA | NA | NA |
| P.microlepis | CaspianKopetDagh | 0.158 | NA | 0.158 | NA | NA | NA |
| P.microlepis | CaspianKopetDagh | 0.158 | NA | 0.158 | NA | NA | NA |
| P.erythrogaster | CaspianKopetDagh | 0.064 | NA | NA | 0.064 | NA | NA |
| P.erythrogaster | CaspianKopetDagh | 0.07 | NA | NA | 0.07 | NA | NA |
| CaucasusMountains | CaspianKopetDagh | 0.047 | NA | NA | NA | 0.047 | NA |
| CaucasusMountains | CaspianKopetDagh | 0.045 | NA | NA | NA | 0.045 | NA |
| CaucasusMountains | CaspianKopetDagh | 0.046 | NA | NA | NA | 0.046 | NA |
| SouthCaspian | CaspianKopetDagh | 0.035 | NA | NA | NA | NA | NA |
| Balkhan | CaspianKopetDagh | 0.022 | NA | NA | NA | NA | 0.022 |
| Balkhan | CaspianKopetDagh | 0.022 | NA | NA | NA | NA | 0.022 |
| CaspianKopetDagh | CaspianKopetDagh | 0.013 | NA | NA | NA | NA | NA |
| CaspianKopetDagh | CaspianKopetDagh | 0.011 | NA | NA | NA | NA | NA |
| | | 0.128457895 | 0.19287 | 0.15663 | 0.06194 | 0.03822 | 0.02266 |

Chapter 2: Supplementary Methods

2.S1 Model Design

(1) NULL– **Historical taxonomy:** *D. maculatus* may be a widespread taxon with divergent mitochondrial lineages, but it is nevertheless a single species. McGuire and Kiew (2001) proposed this as a potential scenario following examination of the subspecies [*Model code: O*].

(2) The data is best explained by two species. **Riverine barriers:** Lineages are divided by either (A) the Chao Phraya and Upper Salween [*Model code: L*], or (B) the Chao Phraya and Upper Mekong [*Model code: N*]. Until recently (post Pleistocene) the Mekong and Salween rivers joined and flowed to the Gulf of Thailand through the presently known Chao Phraya River Basin (Lehman and Fleagle, 2006; Woodruff, 2010). This large river may have played a role in maintaining post-Miocene geographic isolation between the West and Southeast clades. **MtDNA clustering:** The most basal node of the MtDNA clustering splits the two major lineages [*Model code: M*].

(3) The data is best explained by three species. **Historical taxonomy:** The currently recognized subspecies *D. maculatus haasei*, *D. maculatus maculatus*, and *D. maculatus whiteheadi*. While another subspecies is currently recognized (*D. maculatus divergens*), we have left it out of this model because it is located strictly in Chiang Mai, Thailand (a locality missing from our sample set). We sorted samples into the three subspecies by geographic location based on the subspecific distribution described by Musters (1983; Appendix A). The lineage grouping of samples from the Annamite region (localities Z, F, & G) led us to create three subspecies hypotheses: All Annamite samples included in *D. maculatus whiteheadi* [*Model code: J*], all Annamite samples

included in *D. maculatus haasei* [*Model code: K*], and Annamite samples included in *D. maculatus whiteheadi* [*Model code: I*].

(4) The data is best explained by four species. **MtDNA clustering:** The most basal split of the ML tree is followed by a subsequent split on each branch, resulting in four lineages (reciprocally monophyletic pairs). We name these monophyletic groups the (1) West Indochina clade, (2) Central Indochina clade, (3) Southeast Indochina clade, and (4) Northeast Indochina clade. While the topology of the species tree may not reflect these relationships, this model poses that the data is best explained by four lineages grouped according to monophyly from the MtDNA clustering [*Model code: H*].

(5) The data is best explained by five species. **MtDNA clustering:** This model is almost identical to *H* (4 species), with the difference being the splitting of monophyletic Hainan samples and the monophyletic China-Vietnam Coastal Plain samples from the original Northeast Indochina clade, creating five total lineages [*Model code: F*]. **Riverine barriers:** Lineages are divided by four major rivers: Chao Phraya, Mekong, Red, and Upper Salween, four of Indochina's most prominent rivers [*Model code: G*].

(6) The data is best explained by six species. **Riverine barriers:** Lineages are divided by either (A) four major rivers: Chao Phraya, Mekong, Red, and Salween, four of Indochina's most prominent rivers [*Model code: D*], or (B) four major rivers: Chao Phraya, Mekong, Red, and Upper Salween, along with the Quiongzhou Strait [*Model code: E*].

(7) The data is best explained by seven species. **Riverine barriers:** Lineages are divided by either (A) five major rivers: Chao Phraya, Irrawady, Mekong, Red, and Salween, five of

Indochina's most prominent rivers [*Model code: B*], or (B) four major rivers: Chao Phraya, Mekong, Red, and Salween, along with the Quiongzhou Strait [*Model code: C*].

(8) The data is best explained by eight species. **Riverine barriers:** Lineages are divided by five major rivers: Chao Phraya, Irrawady, Mekong, Red, and Salween, along with the Quiongzhou Strait [*Model code: A*]

Chapter 2: Supplementary Figures

Figure S1 Molecular Clock Estimation

Figure S2 Complete ND2 Tree

Figure S3 ML Tree from Six Concatenated Loci

Figure S4 ML gene tree from 12S sequence data

Figure S5 ML gene tree from 16S sequence data

Figure S5 ML gene tree from ND2 sequence data

Figure S5 ML gene tree from BDNF sequence data

Figure S5 ML gene tree from CMOS sequence data

Figure S5 ML gene tree from PNN sequence data

Figure S10 Genus ML gene tree from ND2 sequence data

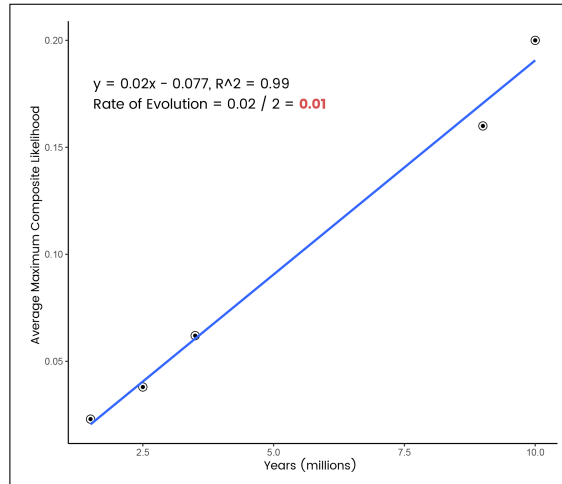


Figure S1: Molecular Clock Estimation

Agamid rate of evolution for region of ND2 used in this study estimated from vicariance-based calibration within *Paralaudakia* species of the Iranian Plateau. Plot shows relationship of average maximum composite likelihood scores estimated from sequences from Macey et al (1998) using truncated ND2 sequence data aligned to the locus used in this study (y axis) and dated geological events corresponding to cladogenesis in *Paralaudakia* (x axis). Figure is analogous to Fig. 7 in Macey et al (1998), which used parsimony distances.

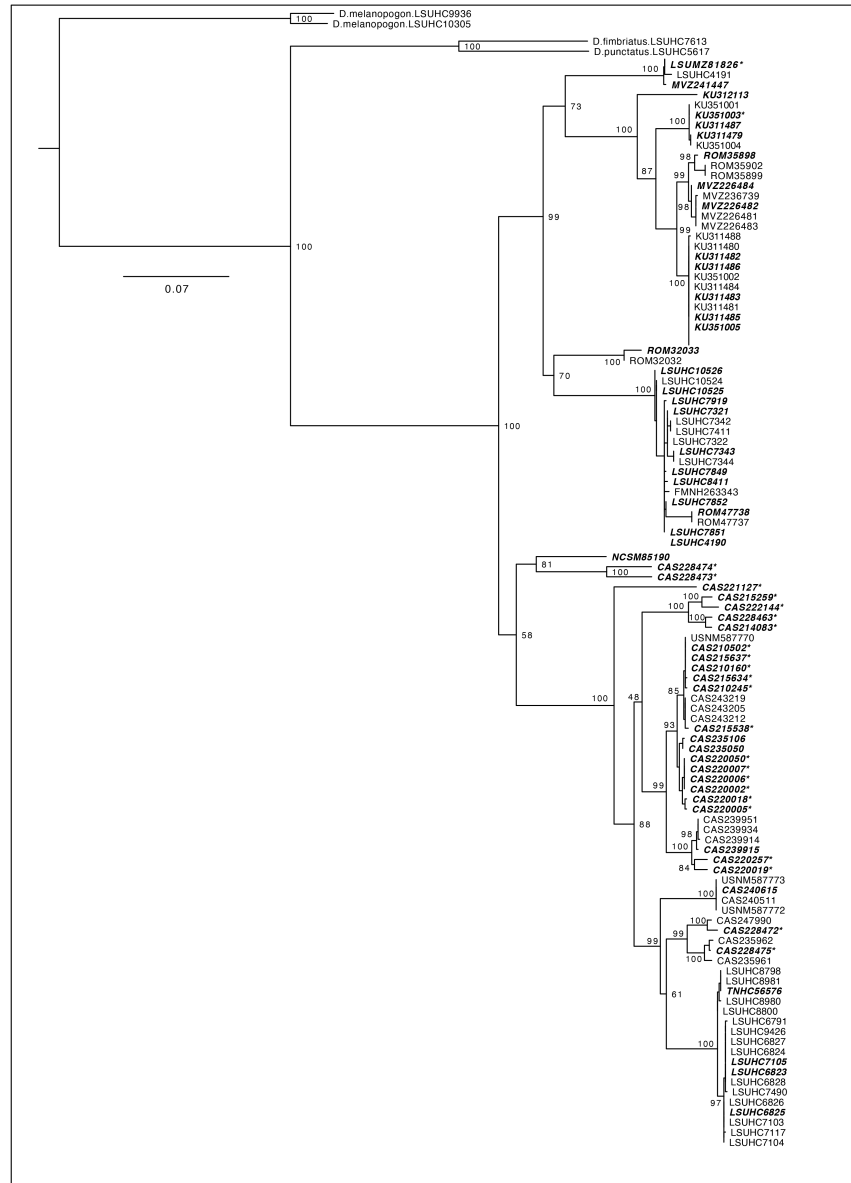


Figure S2: Complete ND2 Tree

Maximum-likelihood (ML) ND2 gene tree with all samples from this study. From this tree we subsampled individuals for future phylogenetic analysis. Bold, Italicized tip names indicate subsampled individuals. Tips are labeled by voucher name (Appendix B). Voucher names followed by * indicate samples with molecular data sequenced before this study.

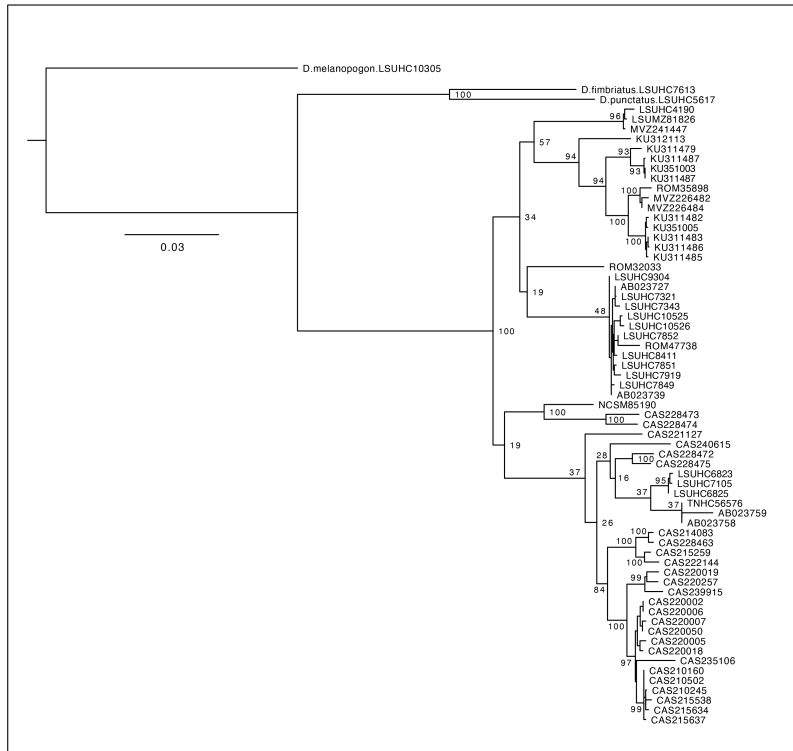


Figure S3: ML Tree from Six Concatenated Loci

ML tree estimated from six-locus (12S, 16S, ND2, BDNF, CMOS, and PNN) concatenated dataset.

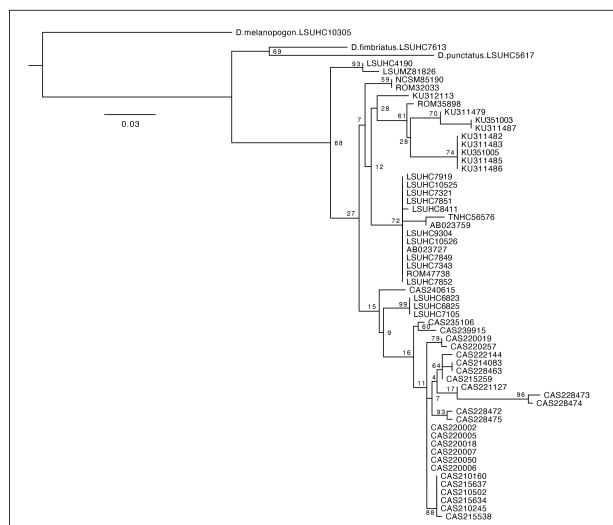


Figure S4: ML gene tree from 12S sequence data.

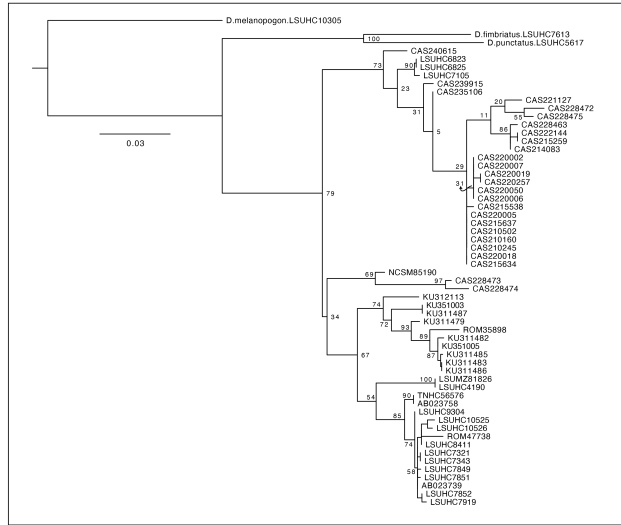


Figure S5: ML gene tree from 16S sequence data.

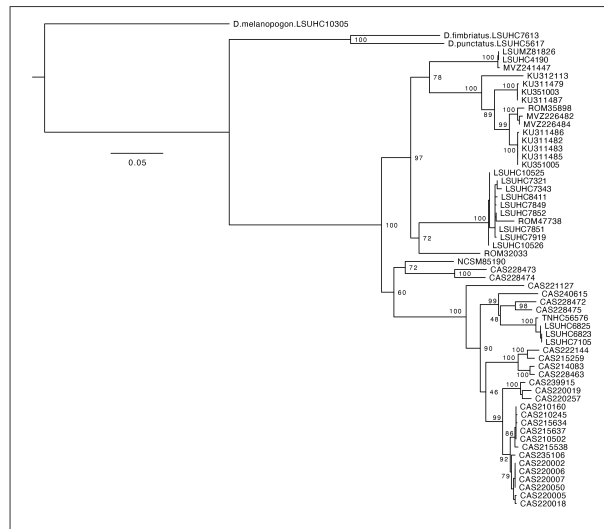


Figure S6: ML gene tree from ND2 sequence data.

Chapter 3: Reduced mitochondrial respiration in hybrid asexual lizards

Published in The American Naturalist (Klabacka et al., 2022)

Randy L. Klabacka^{a,*}, Hailey A. Parry^a, Kang Nian Yap^a, Ryan A. Cook^{a,b}, Victoria A. Herron^{a,c},
L. Miles Horne^{a,d}, Matthew E. Wolak^a, Jose A. Maldonado^e, Matthew K. Fujita^e, Andreas N.
Kavazis^a, Jamie R. Oaks^{a,†}, Tonia S. Schwartz^{a,†}

^aAuburn University and Auburn University Museum of Natural History, Auburn, AL 36849

^bVillanova University, Villanova, PA, 19085

^cUniversity of Missouri College of Veterinary Medicine, Columbia, MO, 65211

^dThe University of Texas at El Paso, El Paso, TX, 79968

^eThe University of Texas at Arlington, Amphibian and Reptile Diversity Research Center,
Arlington, TX, 76019

† These authors contributed equally

Abstract

The scarcity of asexual reproduction in vertebrates alludes to an inherent cost. Several groups of asexual vertebrates exhibit lower endurance capacity (a trait predominantly sourced by mitochondrial respiration) compared to congeneric sexual species. Here we measure endurance capacity in five species of *Aspidoscelis* lizards and examine mitochondrial respiration between sexual and asexual species using mitochondrial respirometry. Our results show reduced endurance capacity, mitochondrial respiration, and phenotypic variability in asexual species compared to parental sexual species along with a positive relationship between endurance capacity and mitochondrial respiration. Results of lower endurance capacity and lower mitochondrial respiration in asexual *Aspidoscelis* are consistent with hypotheses involving mitonuclear incompatibility.

3.1 Introduction

The fitness advantages of asexual reproduction predict an abundance of asexual species (Maynard Smith 1958, 1978). However, the prevalence of sexual reproduction in animals suggests that the evolutionary costs of asexual reproduction outweigh the benefits (Speijer et al. 2015). Although theoretical and empirical studies over the past century have proposed and tested hypotheses regarding these costs, much remains to be understood regarding the direct consequences of asexual reproduction in vertebrates (Fujita et al. 2020).

For the purposes of this article, we refer to “asexuality” as a reproductive strategy where all progeny are produced without male genetic contribution (as opposed to facultative asexuality). Asexual vertebrates, virtually all of which are of hybrid origin (Dawley and Bogart 1989; Avise 2008, 2015; Fujita et al., 2020; but see Sinclair et al. 2010), reproduce by premeiotically doubling their ploidy (Lutes et al. 2010). The subsequent pairing of conspecific homologous chromosomes in meiosis I results in the perpetual preservation of genome-wide heterozygosity (Vrijenhoek and Pfeiler 2008; Warren et al. 2018). With ploidy restored after the completion of meiosis, the cells are ready to develop without variation introduced via fertilization or recombination, thus maintaining the genome in a "frozen" hybrid state (Vrijenhoek and Pfeiler 2008; Warren et al. 2018; but see Hillis et al. (1991) and Warren et al. (2018) for evidence of some gene conversion).

The effect of this unique evolutionary strategy on intracellular bioenergetics is unclear, but examining the effect of heterozygosity on mitochondrial function and overall fitness can inform predictions. Higher rates of coupled mitochondrial respiration and increased fitness (interpreted

as heterosis) have been observed in F₁ hybrids from inbred *Drosophila melanogaster* lines (McDaniel and Grimwood 1971; Martinez and McDaniel 1979) and natural *Tigriopus californicus* populations (Ellison and Burton 2008) compared to their parental lineages, whereas lower values for these traits have been observed when backcrosses lead to mismatched mitochondrial and introgressed nuclear genomes in natural populations of *Tigriopus californicus* (Ellison and Burton 2008) and *Urosaurus* (Haenel and Moore 2018). High heterozygosity of asexual vertebrates led numerous researchers to predict an increase in performance compared to sexual parental species (White 1970; Schultz 1971; Cole 1975; Mitton and Grant 1984; Bullini 1994; Cullum 1997), yet the results from several studies have contradicted these predictions by showing reduced aerobic performance in asexual lineages (Cullum 1997; Mee et al. 2011; Denton et al. 2017).

Aerobic activities requiring endurance (continuous exertion) are powered by oxidative phosphorylation. This catalytic conversion process occurs in the mitochondrion, where a proton gradient powered by nutrient-donated electrons facilitates the phosphorylation of ADP. Production of ATP via this electron transport system (ETS) produces the vast majority of energy used for cellular functions. While the link between mitochondrial function and endurance may seem intuitive, studies examining the association of endurance capacity with mitochondrial respiration have been primarily in the context of biomedical and exercise physiology rather than evolution (e.g., Davies et al. 1981; Gollnick and Saltin 1982; Mercier et al. 1995; Bouchard et al. 1999; Eynon et al. 2011; Jacobs and Lundby 2013; Scott et al. 2018).

Reductions in asexual aerobic performance may be explained in part by mitonuclear incompatibility— the result of interactions between poorly co-adapted gene products from mitochondrial and nuclear genomes that can result in reduced mitochondrial function, organismal performance, and fitness (Ryan and Hoogenraad 2007; Meiklejohn et al. 2013; Hill et al. 2019; Healy and Burton 2020; Rand and Mossman 2020; Moran et al. 2021). We test the hypothesis that the reduced aerobic performance previously observed in several groups of hybrid asexual vertebrates is due to decreased mitochondrial function, as would be predicted with mitonuclear incompatibility. The evolutionary mechanisms leading to incompatibility in these F₁ hybrids could result from mismatched genomes with dominance effects as a result of Darwin’s corollary (Turelli and Moyle 2007) and/or the reduced efficiency of selection on nuclear mutations imposed by the lack of recombination in asexual species (Fisher 1930; Muller 1932). Because asexual vertebrates are of hybrid origin, rather than seeking to disentangle the effects of these traits (asexual reproduction and hybrid origin) we strictly examine hypothesized contributions of an intracellular process (mitochondrial respiration) to an organismal phenomenon (reduced endurance capacity). Squamata (snakes and lizards) is the only vertebrate clade with lineages that reproduce primarily through parthenogenesis, a mode of asexual reproduction with no male input. Using the whiptail lizard genus *Aspidoscelis* as a model system (in which roughly one third of species reproduce parthenogenetically), we quantify endurance capacity and mitochondrial respiration to contrast a sample of sexual and asexual species with two independent origins of parthenogenesis (Reeder et al. 2002; Densmore et al. 1989).

3.2 Methods

3.2.1 Animal Capture

We collected individuals of three sexual (n=6 *Aspidoscelis inornatus*, n=6 *A. marmoratus*, n=7 *A. septemvittatus*) and two asexual (n=4 *A. neomexicanus*, n=7 *A. tessellatus*) species along the Rio Grande basin between Las Cruces, NM and Big Bend National Park, TX (Table 3.S1). The estimated evolutionary relationships of these species (from Reeder et al., 2002) are depicted in Figure 1A (note: although asexual lineages are not “species” in the typical sense [originating via cladogenesis], we join others in referring to them as such given their independent evolutionary trajectory). We caught lizards via lasso or by hand and transported all individuals to Auburn University for temporary housing. All collection and animal care procedures were approved by the United States Department of the Interior, state departments, and the Auburn University IACUC (2018-3286). Additional sampling information is included in the Supplemental Methods.

3.2.2 Endurance Capacity and Mitochondrial Respirometry

We quantified endurance by measuring the time a lizard maintained forward progression at 1 km/hr (on a treadmill), following previously established protocols (Garland 1994; Cullum 1997; see Supplemental Methods for more details). One week later, we measured mitochondrial respiration following previously established protocols (Palmer et al. 1977; Hood et al. 2019; see Supplementary Methods for more details). To measure oxygen consumption through the electron transport chain, we added isolated mitochondria with electron-donating substrates to electrode chamber A (for starting electron transport from CI and continuing through CIII, CIV, and CV) and to electrode chamber B (for starting electron transport from CII and continuing through CIII, CIV, and CV). Measuring mitochondrial respiration via the electron transport chain using these two ports of entry provides independent avenues with different starting substrates to quantify respiration. Because both avenues (starting with CI or CII) comprise interacting mitochondrial

and nuclear gene products (CIII, CIV, and CV), it is inappropriate to use these measures to draw conclusions regarding mitonuclear compatibility.

To initiate coupled, ADP-stimulated respiration (State 3), we added ADP to each chamber. After the phosphorylation of ADP was complete and any oxygen being consumed was driven by protons moving across the inner membrane without facilitation from ATP synthase, we recorded basal respiration (State 4). We normalized respiration rates to mitochondrial protein concentration. To calculate the respiratory control ratio (RCR), we divided State 3 respiration by State 4 respiration.

1.2.2.1 Predictions

To test our hypothesis of reduced mitochondrial respiration in hybrid asexual species compared to their sexual congeneric progenitors, we used the six mitochondrial respiration response variables (State 3, State 4, and RCR initiated from either CI or CII). State 3 respiration measures the rate of oxygen consumption when ATP is being produced (i.e., oxygen consumption is coupled with proton movement through ATP synthase [CV]). If coupled electron transport and ATP synthesis is associated with endurance capacity, we predict that State 3 respiration would be lower in hybrid asexual species. State 4 respiration measures the rate of oxygen consumption when ATP is not being produced (i.e., oxygen consumption is coupled with proton leak across the inner membrane). In this context, we predict no differences in State 4 respiration. As RCR is an indicator of respiration efficiency (coupled respiration controlling for leak), we predicted to see lower RCR in hybrid asexual species if they have lower endurance capacity.

3.2.3 Phylogenetic Network Estimation

Accounting for evolutionary history is critical for accurate comparative methods when multiple lineages are present in a sample set. However, in study systems where lineage history is reticulate rather than bifurcate, models with a phylogenetic network (rather than a tree) more appropriately account for evolutionary history. To estimate the history of diversification and hybridization of the five species of *Aspidoscelis*, we sequenced mitochondrial genomes (following Roelke et al. 2018) and downloaded available mitochondrial sequence data from GenBank (Table 3.S2). We used several software packages to estimate the phylogenetic network (Than and Nakhleh, 2008; Nguyen et al. 2015; Solís-Lemus et al. 2017); we provide details in the Supplementary Methods.

3.2.4 Statistical Analyses

We analyze the data in three ways. First, we use phylogenetic network linear models that include reticulate evolutionary relationships within the model to estimate (1) the effect of hybrid asexuality on each response variable (endurance and mitochondrial respiration [State 3, State 4, and RCR initiated from either CI or CII]) and (2) the effect of mitochondrial respiration on endurance capacity. Second, we use linear mixed-effects models with species random effects to test for (1) an effect of hybrid asexuality on each response variable, (2) the effect of mitochondrial respiration on endurance capacity, and (3) differences in variability between hybrid asexual and sexual species for each response variable. Third, we use linear models for subgroups without needing to account for ancestry to test the effect of hybrid asexuality on each response variable (we made subgroup assignments based on mitochondrial history and parentage). More details for each of these approaches are provided in the Supplementary

Methods. Data and code are available on GitHub (Klabacka, 2021; <https://doi.org/10.5281/zenodo.5784646>) and Dryad (Klabacka et al., 2021; <https://doi.org/10.5061/dryad.zs7h44j8n>).

3.3 Results

3.3.1 Effect of Hybrid Asexuality on Endurance and Mitochondrial Respiration

We found reduced endurance capacity and mitochondrial respiration in hybrid asexual species when using either the phylogenetic network or mixed-effects linear models (Fig. 3.1; Table 3.1; summary statistics in Table 3.S3). We observed that hybrid asexual species had reduced endurance capacity and rates of oxygen consumption when starting from either CI or CII for State 3 and State 4 respiration. We see no support for differences in RCR (for either complex) between sexual and asexual species (Fig. 3.S2). This is not surprising given that both State 3 and State 4 changed in the same direction, resulting in no changes in the ratio between the two measures (RCR). The effect sizes for each response variable are similar between phylogenetic network and mixed-effects linear models (Table 3.1), providing evidence for little phylogenetic signal for response variables. Within-group comparisons show the same general pattern without a statistically significant effect for each response variable (potentially due to lower sample size; Table 3.S4). Details on the within-group comparisons are included in the supplemental materials.

3.3.2 Positive Relationship Between Endurance and Mitochondrial Respiration

We observed a positive relationship between endurance and rate of oxygen consumption when starting from either CI or CII for State 3 and State 4 respiration with either the phylogenetic network or mixed-effects linear models; each of these relationships are statistically significant

except for the phylogenetic network model for CII State 3 (Table 3.1; Fig. 3.2). We see no support for a relationship between endurance and RCR.

3.3.3 Greater Variation in Sexual Species

We found that models incorporating different residual variation parameters for sexual and hybrid asexual groups were preferred for endurance, CII State 3 respiration, and CI and CII State 4 respiration (Tables 1 and S5, Figure S3; also see pink rows vs gray rows in Fig. 3.1B). The approximate posterior probability that sexual species have a greater mean-corrected variance than asexual species was 75 percent, 96 percent, 83 percent, and 94 percent for endurance, CI State 4, CII State 3, and CII State 4, respectively (Fig. 3.S4).

3.4 Discussion

We present novel findings of reduced mitochondrial respiration in hybrid asexual species, along with results reproducing those of previous studies indicating reduced endurance capacity in these asexual species relative to parental sexual species (Cullum 1997; Mee et al. 2011; Denton et al. 2017). A positive relationship between mitochondrial respiration and endurance capacity is evident in our results, which matches our prediction given that aerobic activities require a large amount of ATP and reflects a similar correlation between endurance and mitochondrial genotype in *Drosophila* (Sujkowski et al. 2019). The lower variability in endurance, CII State 3, and CI/CII State 4 in the hybrid asexual species supports the hypothesis that asexual species have lower phenotypic variability due to decreased genetic variation (Ghiselin 1974; Williams 1975; Maynard Smith 1978). While lower phenotypic variability in locomotor performance has been previously documented in asexual *Aspidoscelis* species (Cullum 2000), our study is the first to

report decreased variability in mitochondrial respiration of hybrid asexual species relative to respective parental sexual species.

Despite the high heterozygosity of hybrid asexual species, which led numerous researchers to predict an increase in performance of hybrid asexual species compared to parental sexual species ("hybrid vigor"; see White 1970; Schultz 1971; Cole 1975; Mitton and Grant 1984; Bullini 1994; Cullum 1997), several studies have shown reduced aerobic performance in asexual vertebrate species (Cullum 1997; Mee et al. 2011; Denton et al. 2017). Historically this decrease in performance has been attributed to either (A) genomic incompatibility [consequence of hybridization via negative epistasis (Cullum 1997; Denton et al. 2017) and/or subsequent gene conversion], (B) mutational erosion [consequence of asexuality via Muller's Ratchet (Muller 1964; Leslie and Vrijenhoek 1978; Cullum 1997; Vorburger 2001)], or (C) the inability of the organism to "keep up" with the evolution of parasites due to lack of variation [consequence of asexuality via Red Queen (Valen 1973; Hamilton et al. 1990; Lively et al. 1990; Moritz et al. 1991; Mee and Rowe 2006; Mata-Silva et al. 2008)]. It is also possible that the inability of asexual lineages to combine beneficial alleles that arise in a population via sexual recombination (Maynard Smith 1978) results in the failure of the nuclear genome to efficiently compensate for deleterious mutations that arise in the mitochondrial genome. This hypothesis, an extension of the Hill-Robertson effect (Fisher 1930; Muller 1932; Felsenstein 1974; Hill and Robertson 2007) in the context of accelerated compensatory evolution in nuclear-encoded mitochondrial genes, was originally posed to explain the origin and prevalence of sexual reproduction among eukaryotes (Havird et al. 2015). Additional biological factors such as demography, ecology, and/or life history strategies specific to *Aspidoscelis* sexual mode may contribute to the

differences in endurance and mitochondrial respiration, although several studies have found little to no differences in factors such as response to habitat disturbance (Cosentino et al. 2019), thermal preference (Díaz de la Vega-Pérez et al. 2013), reproductive strategies (Schall 1993), and diet (Smith 1989; Mata-Silva et al. 2013). This study is unable to identify which of the previously described non-mutually exclusive hypotheses best explain our observation of reduced mitochondrial respiration in hybrid asexual species. We recommend that future work integrates genomic sequencing approaches with physiological and cellular measurements (e.g., RNA-seq and individual ETS complex activity) to disentangle potential contributions from these hypotheses.

While reduced endurance is observed in several groups of asexual vertebrates (including this study), examination of parthenogenetic geckos in the *Heteronotia binoei* complex has shown no difference (Roberts et al. 2012) and increased endurance (Kearney et al. 2005) in hybrid asexual species compared to sexual progenitors in different studies. Variation in results between asexual groups may be due to differences in (1) age of asexual lineages, (2) divergence between parental species at the time of hybridization, (3) ploidy, or (4) ecology. Compared to *Heteronotia*, *Aspidoscelis* possesses (1) younger asexual lineages (Reeder et al. 2002; Kearney et al. 2006), (2) greater divergence times between parental species (Strasburg and Kearney 2005; Zheng and Wiens 2016), (3) diploid asexuals [in this study; however, see Cullum (1997) for diploid and triploid asexual *Aspidoscelis* species] (Kearney et al. 2005; Roberts et al. 2012), and (4) a more active foraging strategy (Milstead 1957; Bauer 2007). We do not refer to these differences between taxa as factors that wholly explain our observations, rather we point out that

complexities within these biological systems may be responsible for the seemingly contrasting results.

Decreased RCR between sexual parent and asexual hybrid species is not evident given our data, which contradicts our predictions. We attribute this to the significant differences in State 4 respiration, which we did not predict. State 4, commonly called the "leak" or "basal" state, occurs when ADP has been exhausted. Oxygen consumption occurring via the ETS is being driven by protons "leaking" across the mitochondrial inner membrane rather than via ATP synthase. Low State 4 respiration tends to lead to high reactive oxygen species (ROS) production (Brand 2000), therefore a higher State 4 in sexual compared to hybrid asexual species may be an adaptive trait to mitigate oxidative damage. Investigation into potential differences in ROS production and oxidative damage between sexual and hybrid asexual species is needed to test this hypothesis. We also recommend that future studies include additional respiration states (e.g., State 4 induced by oligomycin, State 3u) for examination of respiratory ratios (Gnaiger 2020).

Lower State 3 respiration in hybrid asexual species suggests decreased mitochondrial respiratory capacity and, as a result, diminished ATP production. The positive relationship we observed between mitochondrial respiration and endurance capacity affirms our predicted relationship between these traits and supports the hypothesis that efficient oxidative phosphorylation increases endurance capacity. Reduced variability of endurance and mitochondrial respiration in hybrid asexual species, a potential result of decreased genetic variation, may be an evolutionary disadvantage. Here we show novel evidence for costs incurred by hybrid asexual species on mitochondrial respiration and reproduce findings of their reduced endurance capacity.

Determining the evolutionary underpinnings of these phenomena, and thus shedding light on which hypotheses are responsible, will require integrating of physiological and genomic sequencing approaches. While the benefits of asexual reproduction can explain the genesis of asexual lineages, incurred costs for this strategy may explain their short evolutionary existence. Reduced mitochondrial respiration and variability in hybrid asexual species may be evolutionary disadvantages when performance and variation are important factors in the realm of natural selection.

3.5 Acknowledgements

We thank the Texas Parks & Wildlife Department [SPR-0318-100] (including the State Parks Division [2019_R1R2_06]), the New Mexico Department of Game and Fish and Energy [3670], the New Mexico Department of Minerals and Natural Resources (State Parks Division), and The United States Department of the Interior (Big Bend National Park [BIBE-2019-SCI-0025]). We also thank the Indio Mountains Research Station for their help with field collecting, the Schwartz, Phyletica, Hill, and Hood labs at Auburn University for their comments on the manuscript, and Paulina Klabacka for her help with field collecting and animal care. This research was funded in-part by an Auburn University Cellular and Molecular Biosciences Peaks of Excellence Research Fellowship. This is contribution no. 952 of the Auburn University Museum of Natural History.

3.6 Statement of Authorship

RLK: Conceptualization, Funding acquisition, Experimental design, Permit acquisition, Sample collection, Animal care, Data collection, Data processing, Data analysis, Data visualization,

Provided resources [field equipment], Code scripting, Writing – original draft, Writing – review and editing

HAP: Data collection, Data processing, Writing – review and editing

KNY: Data collection, Data processing, Writing – review and editing

RAC: Sample collection, Animal care, Data collection

VAH: Sample collection, Animal care, Data collection

LMH: Sample collection, Writing – review and editing

MEW: Data analysis, Code scripting, Data visualization, Writing – review and editing

JAM: Sample collection, Data collection

MKF: Conceptualization, Provided resources [lab equipment, reagents], Writing – review and editing

ANK: Provided resources [lab equipment], Supervision, Writing – review and editing

JRO: Conceptualization, Experimental design, Provided resources [field equipment, reagents], Supervision, Writing – review and editing

TSS: Conceptualization, Experimental design, Provided resources [field equipment, reagents], Supervision, Writing – review and editing

3.7 Data and Code Accessibility

GitHub: <https://doi.org/10.5281/zenodo.5784646> (Klabacka, 2021)

Dryad: <https://doi.org/10.5061/dryad.zs7h44j8n> (Klabacka et al., 2021)

3.8 References

- Avise, J. C. 2008. *Clonality*. Oxford University Press, New York.
- Avise, J. C. 2015. Evolutionary perspectives on clonal reproduction in vertebrate animals. *Proceedings of the National Academy of Sciences*. 112(29):8867-8873.
- Bauer, A. M. 2007. The foraging biology of the Gekkota: life in the middle. Pages 371–404 *in* *Lizard Ecology*. Cambridge University Press, Cambridge.
- Bouchard, C., P. An, T. Rice, J. S. Skinner, J. H. Wilmore, J. Gagnon, L. Pérusse, A. S. Leon, and D. C. Rao. 1999. Familial aggregation of VO_{2max} response to exercise training: results from the HERITAGE Family Study. *Journal of Applied Physiology* 87:1003–1008.
- Brand, M. D. 2000. Uncoupling to survive? The role of mitochondrial inefficiency in ageing. *Experimental Gerontology* 35:811-820.
- Bullini, L. 1994. Origin and evolution of animal hybrid species. *Trends in Ecology and Evolution* 9:422-426.
- Cole, C. J. 1975. Evolution of Parthenogenetic Species of Reptiles. Pages 340–355 *in* R. Reinboth, ed. *Intersexuality in the Animal Kingdom*. Springer, Berlin Heidelberg.
- Cosentino, B. J., R. L. Schooley, B. T. Bestelmeyer, H. Campos, and L. M. Burkett. 2019. Does habitat disturbance promote geographical parthenogenesis in whiptail lizards? *Evolutionary Ecology* 33:839–853.
- Cullum, A. J. 1997. Comparisons of physiological performance in sexual and asexual whiptail lizards (genus *Cnemidophorus*): implications for the role of heterozygosity. *The American naturalist* 150:24–47.
- Cullum, A. J. 2000. Phenotypic variability of physiological traits in populations of sexual and asexual whiptail lizards (genus *Cnemidophorus*). *Evolutionary Ecology Research* 2:841-855.

- Davies, K. J. A., L. Packer, and G. A. Brooks. 1981. Biochemical adaptation of mitochondria, muscle, and whole-animal respiration to endurance training. *Archives of Biochemistry and Biophysics* 209:539–554.
- Dawley, R., and J. Bogart. 1989. *Evolution and Ecology of Unisexual Vertebrates*. New York State Museum, Albany.
- Díaz de la Vega-Pérez, A. H., V. H. Jiménez-Arcos, N. L. Manríquez-Morán, and F. R. Méndez-de la Cruz. 2013. Conservatism of thermal preferences between parthenogenetic *Aspidoscelis cozumela* complex (Squamata: Teiidae) and their parental species. *Herpetological Journal* 23:93–104.
- Densmore, L. D., J. W. Wright, and W. M. Brown. 1989. Mitochondrial-DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*) II. *C. neomexicanus* and the *C. tesselatus* complex. *Evolution* 43:943-957.
- Denton, R. D., K. R. Greenwald, and H. L. Gibbs. 2017. Locomotor endurance predicts differences in realized dispersal between sympatric sexual and unisexual salamanders. *Functional Ecology* 31:915–926.
- Ellison, C. K., and R. S. Burton. 2008. Interpopulation hybrid breakdown maps to the mitochondrial genome. *Evolution; international journal of organic evolution* 62:631–8.
- Eynon, N., M. Morán, R. Birk, and A. Lucia. 2011. The champions' mitochondria: Is it genetically determined? A review on mitochondrial DNA and elite athletic performance. *Physiological Genomics* 43:789-798.
- Felsenstein, J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737–756.
- Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- Fujita, M. K., S. Singhal, T. O. Brunes, and J. A. Maldonado. 2020. *Evolutionary Dynamics and*

- Consequences of Parthenogenesis in Vertebrates. *Annual Review of Ecology, Evolution, and Systematics* 51:191-214.
- Garland, T. 1994. Phylogenetic analyses of lizard endurance capacity in relation to body size and body temperature. Pages 237–259 *in* L. J. Vitt and E. R. Pianka, ed. *Lizard Ecology: Historical and Experimental Perspectives*. Princeton University Press, Princeton.
- Ghiselin, M. T. 1974. The Economy of Nature and the Evolution of Sex. *Journal of the History of Biology* 9:324.
- Gnaiger, E. 2020. Mitochondrial pathways and respiratory control. An introduction to OXPHOS analysis. *Bioenergetics Communications* 5.
- Gollnick, P. D., and B. Saltin. 1982. Significance of skeletal muscle oxidative enzyme enhancement with endurance training. *Clinical Physiology* 2:1–12.
- Haenel, G. J., and V. D. G. Moore. 2018. Functional divergence of mitochondria and coevolution of genomes: Cool mitochondria in hot lizards. *Physiological and Biochemical Zoology* 91:1068–1081.
- Hamilton, W. D., R. Axelrod, and R. Tanese. 1990. Sexual reproduction as an adaptation to resist parasites (A review). *Proceedings of the National Academy of Sciences of the United States of America* 87:3566–3573.
- Havird, J. C., M. D. Hall, and D. K. Dowling. 2015. The evolution of sex: A new hypothesis based on mitochondrial mutational erosion: Mitochondrial mutational erosion in ancestral eukaryotes would favor the evolution of sex, harnessing nuclear recombination to optimize compensatory nuclear coadaptation. *BioEssays* 37:951–958.
- Healy, T. M., and R. S. Burton. 2020. Strong selective effects of mitochondrial DNA on the nuclear genome. *Proceedings of the National Academy of Sciences of the United States of*

America 117:6616–6621.

Hill, G. E., J. C. Havird, D. B. Sloan, R. S. Burton, C. Greening, and D. K. Dowling. 2019.

Assessing the fitness consequences of mitonuclear interactions in natural populations.

Biological Reviews 94:1089–1104.

Hill, W. G., and A. Robertson. 2007. The effect of linkage on limits to artificial selection.

Genetical research 89:311–36.

Hillis, D. H., C. Moritz, C. A. Porter, and R. J. Baker. 1991. Evidence for biased gene conversion

in concerted evolution of ribosomal DNA. Science. 251(4991):308-310.

Hood, W. R., Y. Zhang, H. A. Taylor, N. R. Park, A. E. Beatty, R. J. Weaver, K. N. Yap, and A.

N. Kavazis. 2019. Prior reproduction alters how mitochondria respond to an oxidative event.

Journal of Experimental Biology 222.

Jacobs, R. A., and C. Lundby. 2013. Mitochondria express enhanced quality as well as quantity

in association with aerobic fitness across recreationally active individuals up to elite athletes.

Journal of Applied Physiology 114:344–350.

Kearney, M., M. J. Blacket, J. L. Strasburg, and C. Moritz. 2006. Waves of parthenogenesis in

the desert: evidence for the parallel loss of sex in a grasshopper and a gecko from Australia.

Molecular Ecology 15:1743–1748.

Kearney, M., R. Wahl, and K. Autumn. 2005. Increased Capacity for Sustained Locomotion at

Low Temperature in Parthenogenetic Geckos of Hybrid Origin. Physiological and

Biochemical Zoology 78:316-324.

Klabacka, R. L., H. A. Parry, K. N. Yap, R. A. Cook, V. A. Herron, L. M. Horne, M. E. Wolak,

J. A. Maldonado, M. K. Fujita, A. N. Kavazis, J. R. Oaks, and T. S. Schwartz. 2021.

American Naturalist, Dryad Digital Repository, doi: 10.5061/dryad.zs7h44j8n.

- Klabacka, R. L. 2021. AmNat-Aspidoscelis-2021, GitHub Repository. doi:
10.5281/zenodo.5784646.
- Laskowski, K. L., C. Doran, D. Bierbach, J. Krause, and M. Wolf. 2019. Naturally clonal vertebrates are an untapped resource in ecology and evolution research. *Nature Ecology & Evolution*. 3:161-169.
- Leslie, J. F., and R. C. Vrijenhoek. 1978. Genetic dissection of clonally inherited genomes of poeciliopsis. I. Linkage analysis and preliminary assessment of deleterious gene loads. *Genetics* 90:801–811.
- Lively, C. M., C. Craddock, and R. C. Vrijenhoek. 1990. Red Queen hypothesis supported by parasitism in sexual and clonal fish. *Nature* 344:864–866.
- Lutes, A. a, W. B. Neaves, D. P. Baumann, W. Wiegraebe, and P. Baumann. 2010. Sister chromosome pairing maintains heterozygosity in parthenogenetic lizards. *Nature* 464:283–286.
- Martinez, A. O., and R. G. McDaniel. 1979. Mitochondrial heterosis in aging *Drosophila* hybrids. *Experimental Gerontology* 14:231–238.
- Mata-Silva, V., C. R. Bursey, and J. D. Johnson. 2008. Gut parasites of two syntopic species of whiptail lizards, *Aspidoscelis marmorata* and *Aspidoscelis tessellata* from the Northern Chihuahuan Desert. *Sociedad Herpetológica Mexicana* 16:1–4.
- Mata-Silva, V., J. D. Johnson, and A. Ramirez-Bautista. 2013. Comparison of Diets of Two Syntopic Lizards, *Aspidoscelis marmorata* and *Aspidoscelis tessellata* (Teiidae), from the Northern Chihuahuan Desert of Texas. *The Southwestern Naturalist* 58:209–215.
- Maynard Smith, J. 1958. *The Theory of Evolution*. Cambridge University Press, Cambridge.
- Maynard Smith, J. 1978. *The Evolution of Sex*. Cambridge University Press, Cambridge.

- McDaniel, R. G., and B. G. Grimwood. 1971. Hybrid vigor in *Drosophila*: Respiration and mitochondrial energy conservation. *Comparative Biochemistry and Physiology* 38:309-314.
- Mee, J. A., C. J. Brauner, and E. B. Taylor. 2011. Repeat Swimming Performance and Its Implications for Inferring the Relative Fitness of Asexual Hybrid Dace (Pisces: *Phoxinus*) and Their Sexually Reproducing Parental Species. *Physiological and Biochemical Zoology* 84:306–315.
- Mee, J. A., and L. Rowe. 2006. A comparison of parasite loads on asexual and sexual *Phoxinus* (Pisces: Cyprinidae). *Canadian Journal of Zoology* 84:808–816.
- Meiklejohn, C. D., M. A. Holmbeck, M. A. Siddiq, D. N. Abt, D. M. Rand, and K. L. Montooth. 2013. An Incompatibility between a Mitochondrial tRNA and Its Nuclear-Encoded tRNA Synthetase Compromises Development and Fitness in *Drosophila*. *PLoS Genetics* 9:e1003238.
- Mercier, J. G., J. F. Hokanson, and G. A. Brooks. 1995. Effects of cyclosporine A on skeletal muscle mitochondrial respiration and endurance time in rats. *American Journal of Respiratory and Critical Care Medicine* 151:1532–1536.
- Milstead, W. W. 1957. Observations on the Natural History of Four Species of Whiptail Lizard, *Cnemidophorus* (Sauria, Teiidae) in Trans-Pecos Texas. *The Southwestern Naturalist* 2:105.
- Mitton, J. B., and M. C. Grant. 1984. Associations Among Protein Heterozygosity, Growth Rate, and Developmental Homeostasis. *Annual Review of Ecology and Systematics* 15:479–499.
- Moran, B. M., C. Y. Payne, D. L. Powell, E. N. K. Iverson, S. M. Banerjee, Q. K. Langdon, T. R. Gunn, F. Lui, R. Matney, K. Singhal, R. D. Leib, O. Hernandez-Perez, R. Corbett-Detig, M. Schartl, J. C. Havird, and M. Schumer. 2021. A Lethal Genetic Incompatibility between Naturally Hybridizing Species in Mitochondrial Complex I. *bioRxiv* 2021.07.13.452279.

- Moritz, C., H. Mccallum, S. Donnellan, and J. D. Roberts. 1991. Parasite loads in parthenogenetic and sexual lizards (*Heteronotia binoei*): support for the Red Queen hypothesis. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 244:145–149.
- Muller, H. J. 1932. Some Genetic Aspects of Sex. *The American Naturalist* 66:118–138.
- Muller, H. J. 1964. The relation of recombination to mutational advance. *Mutation Research* 1:2–9.
- Nguyen, L. T., H. A. Schmidt, A. Von Haeseler, and B. Q. Minh. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32:268–274.
- Palmer, J. W., B. Tandler, and C. L. Hoppel. 1977. Biochemical properties of subsarcolemmal and interfibrillar mitochondria isolated from rat cardiac muscle. *Journal of Biological Chemistry* 252:8731–8739.
- Rand, D. M., and J. A. Mossman. 2020. Mitonuclear conflict and cooperation govern the integration of genotypes, phenotypes and environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 375.
- Reeder, T. W., C. J. Cole, and H. C. Dessauer. 2002. Phylogenetic relationships of whiptail lizards of the genus *Cnemidophorus* (Squamata: Teiidae): a test of monophyly, reevaluation of karyotypic evolution, and review of hybrid origins. *American Museum Novitates* 2002:1–61.
- Roberts, J. A., H. D. Vo, M. K. Fujita, C. Moritz, and M. Kearney. 2012. Physiological implications of genomic state in parthenogenetic lizards of reciprocal hybrid origin. *Journal of Evolutionary Biology* 25:252–263.

- Roelke, C. E., Maldonado, J. A., Pope, B. W., Firreno Jr., T. J., Laduc, T. J., Hibbitts, T. J., Ryberg, W. A., Rains, N. R., and Fujita, M. K. 2018. Mitochondrial genetic variation within and between *Holbrookia lacerta lacerta* and *Holbrookia lacerta subcaudalis*, the spot-tailed earless lizards of Texas. *Journal of Natural History*. 52:1017-1027.
- Ryan, M. T., and N. J. Hoogenraad. 2007. Mitochondrial-nuclear communications. *Annual Review of Biochemistry* 77:701-722.
- Schall, J. J. 1993. Community ecology of *Cnemidophorus* lizards in southwestern Texas: a test of the weed hypothesis. Pages 319-343 in *Biology of Whiptail Lizards (Genus Cnemidophorus)* Oklahoma Museum of Natural History, Norman.
- Schultz, R. J. 1971. Special adaptive problems associated with unisexual fishes. *Integrative and Comparative Biology* 11:351–360.
- Scott, G. R., K. H. Guo, and N. J. Dawson. 2018. The Mitochondrial Basis for Adaptive Variation in Aerobic Performance in High-Altitude Deer Mice. *Integrative and Comparative Biology* 58:506–518.
- Sinclair, E. A., J. B. Pramuk, R. L. Bezy, K. A. Crandall, and J. W. Sites. 2010. Dna evidence for nonhybrid origins of parthenogenesis in natural populations of vertebrates. *Evolution* 64:1346–1357.
- Smith, D. D. 1989. A Comparison of Food Habits of Sympatric *Cnemidophorus exsanguis* and *Cnemidophorus gularis* (Lacertilia, Teiidae). Source: *The Southwestern Naturalist* 34:418–420.
- Solís-Lemus, C., P. Bastide, and C. Ané. 2017. PhyloNetworks: A Package for Phylogenetic Networks. *Molecular Biology and Evolution* 34:3292–3298.
- Speijer, D., J. Lukeš, and M. Eliáš. 2015. Sex is a ubiquitous, ancient, and inherent attribute of

- eukaryotic life. *Proceedings of the National Academy of Sciences of the United States of America* 112:8827–34.
- Strasburg, J. L., and M. Kearney. 2005. Phylogeography of sexual *Heteronotia binoei* (Gekkonidae) in the Australian arid zone: climatic cycling and repetitive hybridization. *Molecular Ecology* 14:2755–2772.
- Sujkowski, A., A. N. Spierer, T. Rajagopalan, B. Bazzell, M. Safdar, D. Imsirovic, R. Arking, D. M. Rand, and R. Wessells. 2019. Mito-nuclear interactions modify *Drosophila* exercise performance. *Mitochondrion* 47:188–205.
- Than, C., D. Ruths, and L. Nakhleh. 2008. PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics* 9:322.
- Turelli, M., and L. Moyle. 2007. Asymmetric postmating isolation: Darwin’s corollary to Haldane’s rule. *Genetics* 176:1059–1088.
- Valen, L. Van. 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- Vorburger, C. 2001. Fixation Of Deleterious Mutations In Clonal Lineages: Evidence From Hybridogenetic Frogs. *Evolution* 55:2319–2332.
- Vrijenhoek, R. C., and E. Pfeiler. 2008. Differential Survival of Sexual and Asexual *Poeciliopsis* During Environmental Stress. *Evolution* 51:1593-1600.
- Warren, W. C., R. García-Pérez, S. Xu, K. P. Lampert, D. Chalopin, M. Stöck, L. Loewe, Y. Lu, L. Kuderna, P. Minx, M. J. Mantague, C. Tomlinson, L. W. Hillier, D. N. Murphy, J. Wang, Z. Wang, C. M. Garcia, G. C. W. Thomas, J.N. Volff, F. Farias, B. Aken, R. B. Walter, K. D. Pruitt, T. Marques-Bonet, M. W. Hahn, S. Kneitz, M. Lynch, and M. Schartl. 2018. Clonal polymorphism and high heterozygosity in the celibate genome of the Amazon molly. *Nature Ecology and Evolution* 2:669–679.

White, M. J. D. 1970. Heterozygosity and Genetic Polymorphism in Parthenogenetic Animals.

Pages 237–262 *in* Essays in Evolution and Genetics in Honor of Theodosius Dobzhansky.

Springer, New York.

Williams, C. G. 1975. Sex and Evolution. Princeton University Press, Princeton.

Zheng, Y., and J. J. Wiens. 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.

Chapter 3: Figures

Figure 3.1 Reduced endurance and mitochondrial respiration in hybrid parthenogens

Figure 3.2 Positive relationship between mitochondrial respiration states and endurance

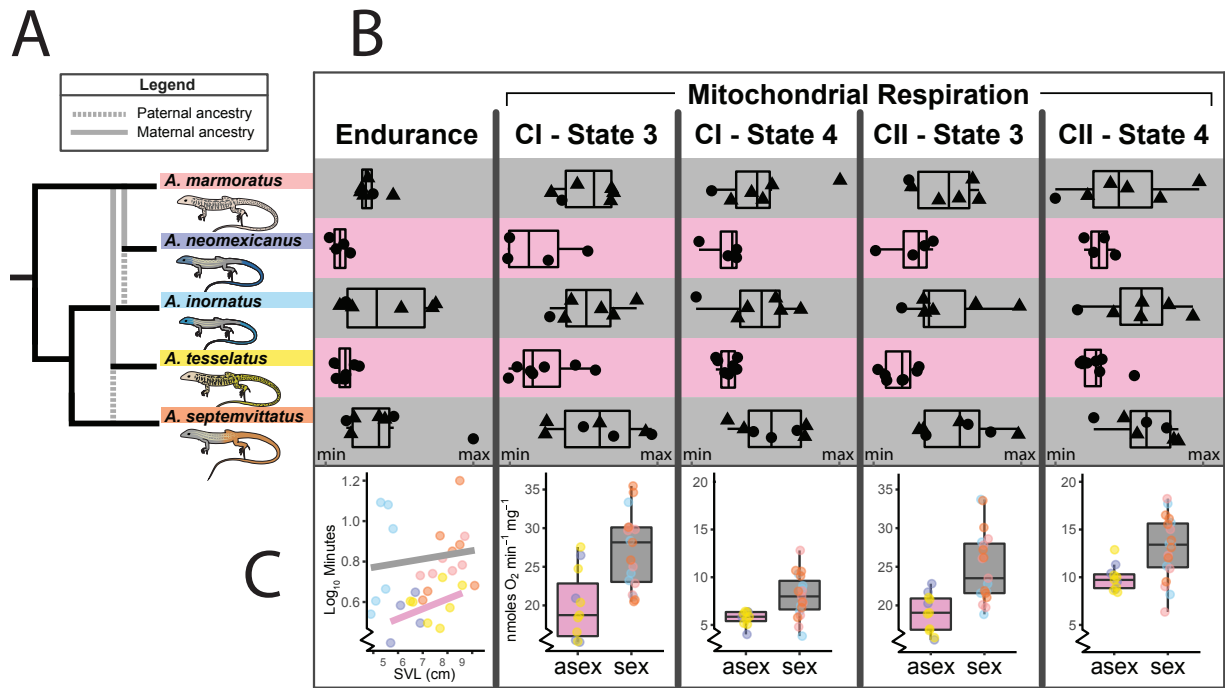


Figure 3.1: **Reduced endurance and mitochondrial respiration in hybrid parthenogens**

A- Evolutionary network showing relationships of all species in the study (from Reeder et al. [2002]; branch lengths and divergence times are not informative). Reticulations (representing hybridization events) leading to asexual hybrids are shown by gray bars, with solid bars showing maternal ancestry and dashed bars showing paternal ancestry. **B-** Table of box plots showing the variation and scaled response variable values for endurance, state 3 and 4 respiration starting at mt complex I, and state 3 and 4 respiration starting at mt complex II. Hybrid asexual species are indicated with pink-filled rows whereas sexual species are indicated with gray-filled rows. Females are indicated with circles and males with triangles. Except for CI State 3, all shown response variables show significant differences in variation between sexual and hybrid asexual species (see Table 1). **C-** Comparisons of all shown response variables were significant between reproductive modes (sexual vs asexual) for each of the response variables shown (with $p < 0.05$). Plots in C correspond to the above columns of B. The endurance data is shown with a scatterplot since we used body size [SVL] as a covariate in the model, and the lines of best fit represent hybrid asexual species (pink) and sexual species (gray). The y axes for each of the mitochondrial respiration response variables is the same. Values for effect sizes, variance, and p can be found in Table 1. Plots with full y axes (along with RCR) are shown in figure S2.

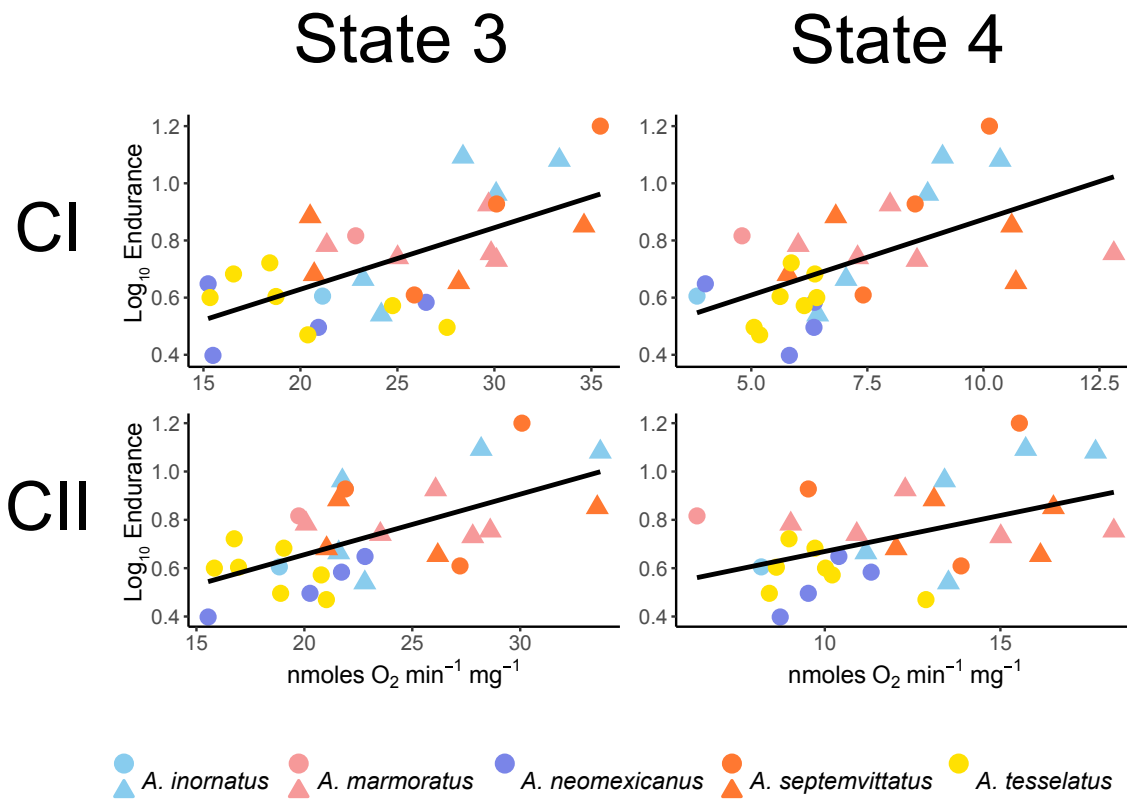


Figure 3.2: **Positive relationship between mitochondrial respiration states and endurance**

Scatterplot showing effect of mitochondrial respiration for the initiation complexes (CI and CII) and respiration states (State 3 and State 4) on \log_{10} endurance. Colors and shapes correspond to species and sex, respectively (circles=females, squares=males). Values for effect size (slope), p , and r^2 can be found in Table 1

Chapter 3: Tables

Table 3.1: Results from PhyloNetwork and Mixed-effects linear models

| | Hybrid Parthenogens | | | | Log Endurance | | | |
|---------------------|---------------------|-------|----------|----------------------|---------------|--------|---------|--------|
| | β | SE | p | σ_s, σ_a | β | SE | p | r^2 |
| PhyloNetwork Model | | | | | | | | |
| Log Endurance | -0.22 | 0.040 | 0.031 | – | – | – | – | – |
| CI State 3 | -6.88 | 0.44 | 6.03e-04 | – | 0.033 | 0.0036 | 0.0028 | 0.97 |
| CI State 4 | -2.33 | 0.23 | 0.0020 | – | 0.095 | 0.016 | 0.011 | 0.92 |
| CI RCR | 0.066 | 0.13 | 0.66 | – | -0.13 | 0.55 | 0.84 | 0.017 |
| CII State 3 | -6.12 | 1.13 | 0.013 | – | 0.031 | 0.11 | 0.060 | 0.74 |
| CII State 4 | -2.90 | 0.47 | 0.0087 | – | 0.072 | 0.017 | 0.023 | 0.86 |
| CII RCR | -0.13 | 0.09 | 0.26 | – | 0.52 | 0.59 | 0.44 | 0.20 |
| Mixed-effects Model | | | | | | | | |
| Log Endurance | -0.24 | 0.050 | 0.017 | 0.18, 0.090* | – | – | – | – |
| CI State 3 | -7.10 | 1.77 | 0.028 | 4.77, 4.51 | 0.021 | 0.0050 | 0.0002 | 0.40 |
| CI State 4 | -2.31 | 0.57 | 0.026 | 2.26, 0.75* | 0.049 | 0.014 | 0.0023 | 0.29 |
| CI RCR | 0.021 | 0.29 | 0.95 | 0.69, 0.88 | -0.042 | 0.042 | 0.33 | 0.23 |
| CII State 3 | -5.91 | 1.28 | 0.019 | 4.53, 2.49* | 0.021 | 0.0060 | 4.0e-04 | 0.37 |
| CII State 4 | -3.17 | 0.85 | 0.034 | 3.29, 1.32* | 0.023 | 0.011 | 0.042 | 0.14 |
| CII RCR | 0.036 | 0.13 | 0.81 | 0.36, 0.21 | 0.039 | 0.11 | 0.71 | 0.0035 |

NOTE. – Within columns we show the deviation of hybrid asexual species from sexual species for all response variables (left) and the effect of mitochondrial respiration states on endurance (right). The table is broken into two horizontal sections showing results from the PhyloNetwork linear model (top) and the mixed-effects linear model (bottom). Effect sizes (β), standard error (SE), p-values (p), standard deviations for reproductive modes (σ_s, σ_a ; standard deviation for sexual and hybrid asexual species, respectively. Asterisks [*] indicate models where two residual variances were selected. Confidence intervals, results from the likelihood ratio test, and coefficients of variation are in Table S5.), and coefficient of determination (r^2) are shown for the models.

Chapter 3: Supplementary Methods

Animal Capture

We collected individuals along the Rio Grande basin between Las Cruces, New Mexico and Big Bend National Park, Texas. We caught thirty lizards (Table S1) either via lasso or by hand, and transported all individuals to Auburn University for temporary housing. Lizards were housed in laboratory conditions reflecting their typical desert habitat (25° C burrowing conditions, 40° C sunning conditions), and were fed ad libitum (crickets and mealworms) and watered daily (Townsend 1979). Individuals were accessioned as alcohol-preserved specimens in the Auburn University Museum of Natural History.

Endurance Capacity

After one month of acclimation to lab conditions, we randomly ordered individuals for treadmill running. After fasting the lizards 24 hours, we measured endurance capacity by timing the number of seconds an individual maintained forward progression on a treadmill moving at 1 km/hr (measured with a stopwatch). We terminated each trial when an individual could no longer keep pace with the treadmill following five repeated prompts (light pinching at base of tail). We log (Log_{10}) transformed the endurance measurement for all statistical models including endurance following visual inspection of the distribution of residuals (log transformation showed a more normal distribution of residuals). For all models comparing the effect of reproductive mode on endurance, we included snout-vent length (SVL) as a covariate for body size (Garland, 1994).

Mitochondrial Isolation and Respirometry

One week after endurance trials, we randomly ordered individuals and assigned them to days for live mitochondrial respirometry. After fasting lizards one day, we euthanized animals via decapitation and immediately excised and transferred skeletal muscle from the front and hind limbs to 10 w/v of isolation buffer (100 mM l^{-1} KCl, 40 mM l^{-1} Tris-HCl, 10 mM l^{-1} Tris base, 1 mM l^{-1} MgCl₂, 1 mM l^{-1} EGTA, 0.2 mM l^{-1} ATP and 0.15 percent [w/v] free fatty acid bovine serum albumin [BSA], pH 7.50). After mincing muscle tissue, we homogenized the sample with VITRIS-5 homogenizer at medium power for five seconds and added fresh protease (Trypsin-T1426: 5 mg/g wet muscle). Homogenate was then mixed every 30 seconds for seven minutes before digestion was terminated with an equal volume of isolation buffer. We then centrifuged the sample at 500 xG for 10 minutes at 4°C, and passed the supernatant through gauze. We performed two steps of centrifugation at 4969 xG for 15 minutes at 4° C, with resuspension of the mitochondrial pellet following each of these two centrifugations (using isolation buffer with a volume equal to that of the original isolation buffer used, but the second resuspension used isolation buffer lacking

BSA). Following centrifugation, we suspended the final mitochondrial pellet in 100 μ l mannitol/sucrose solution (220 mmol l⁻¹ mannitol, 70 mmol l⁻¹ sucrose, 10 mmol l⁻¹ Tris-HCl, and 1 mmol l⁻¹ EGTA; pH 7.40).

We added isolated mitochondria to respiration buffer (100 mM KCl, 10 mM KH₂PO₄, 1 mM EGTA, 50 mM MOPS, 10 mM MgCl₂, 20 mM glucose, and 0.2 percent BSA) in two water-jacketed respiratory chambers (Hansatech Oxytherm; hereafter referred to as chambers A and B) with continuous stirring at 40°C. We measured respiration by quantifying oxygen consumption with 2 mM pyruvate, 2 mM malate and 10 mM glutamate substrates in chamber A for respiration initiated through complex I (CI), and 5 mM succinate as a substrate with 5 μ M of rotenone to inhibit CI in chamber B for respiration initiated through complex II (CII). We began ADP-stimulated respiration (State 3) by adding 0.25 mM ADP in each chamber. Basal respiration (State 4), which occurs when oxygen consumption is driven by protons "leaking" across the inner membrane, was recorded after the phosphorylation of ADP was complete. Respiration rates were normalized to mitochondrial protein content determined via Bradford assay. Respiratory control ratio (RCR) was calculated by dividing state 3 respiration by state 4 respiration. Mitochondria isolation and respirometry were both executed in a manner where the researchers and data recorders were blind to the species and mode of reproduction.

Phylogenetic Network Estimation

We sequenced mitochondrial genomes following methods described by (Roelke et al., 2015) and downloaded available mitochondrial sequence data from GenBank for phylogenetic estimation (Table S2). We aligned sequences using the MAFFT v7.388 algorithm (Kato et al., 2002) within the sequence editing program Geneious Prime v2019.0.4. We bootstrapped the sequence alignment (1000 nonparametric replicates), estimated the maximum likelihood tree for each bootstrap replicate, and used the maximum likelihood trees (with the -m TEST command to perform standard model selection) to construct a consensus tree using the program IQ-Tree (Nguyen et al., 2015; see Fig S1). We pruned the tree to include only the focal taxa for this study using the R package ape. Because this gene tree only represented the maternal mitochondrial ancestry, to approximate the paternal history we pruned copies of the asexual hybrid lineages from their maternal sister lineage and grafted these in with their paternal sister lineage while preserving the branch lengths. The reticulation at the base of an asexual lineage occurs at the hybridization event between two sexual species. Shortly after this event, the asexual species' mitochondrial gene tree (which generally sorts [coalesces] faster than the nuclear ancestry due to the smaller effective population size) diverged from the maternal sexual ancestor, providing an upper bound on the time of hybridization. While using this paternal tree would not be an appropriate reference for understanding the evolution of the nuclear genome in the hybrid lineages, it is adequate to reconstruct the reticulate phylogeny that is currently supported from previous studies with these species (Densmore et al., 1989). With the maternal and paternal trees as input, we estimated the phylogenetic network using the InferNetwork_ML (maximum

likelihood) command in the program PhyloNet (Than and Nakhleh, 2008; see Fig S1).

Linear Modeling and Analyses

We analyze the data in three ways:

1) Phylogenetic Linear Models

First, in order to account for reticulate evolutionary history, we constructed phylogenetic network linear models to test (1) the effect of hybrid asexuality on each response variable (Log_{10} Endurance, CI State 3, CI State 4, CI RCR, CII State 3, CII State 4, and CII RCR) and (2) the effect of mitochondrial respiration on endurance capacity using the function `phylolm` within the Julia package `PhyloNetworks` (Solis-Lemus et al., 2017). These models included the phylogenetic network to compute the variance matrix for the linear regression. Within-species variation was captured using the `y_mean_std` flag to incorporate standard deviation into the model.

2) Linear Mixed-Effects Models

Second, we designed linear mixed-effects models with species random effects to again test for (1) an effect of hybrid asexuality on each response variable, (2) differences in variation between hybrid asexual and sexual species for each response variable, and (3) the effect of mitochondrial respiration on endurance capacity. In this approach we did not include phylogenetic relatedness, therefore all species are independent. We created linear models using the R packages `nlme` and `MCMCglmm`. Here we briefly describe the model architecture; model specifics can be seen in the annotated code file `StatisticalAnalysis.R`. Results from the analyses are shown in Table S5.

For each response variable, we created two linear models using the function `lme`: (1) a model with the sexual mode as a fixed-effect independent variable and species as a random variable and (2) a model with the same components as model 1, with an additional residual variation parameter [allowing for different variability for sexual and hybrid asexual species]. We fit the models to the data and performed a likelihood ratio test to compare model fit (to test whether a model with two parameters of variation fit better than a model with a single parameter). For the models including two residual variation parameters, we obtained estimates of uncertainty by performing nonparametric bootstrapping (1000 replicates) and re-estimating the standard deviation.

Because there can be a linear relationship between the mean and variance on the scale of measurement, estimates of variance using mean-corrected approaches can provide a conservative approach to assessing heteroscedasticity. To account for the possibility that differences in variance between sexual and asexual species is due to differences in scale

of the response variable, we also used a Bayesian approach to compare mean-corrected standard deviations. Within models constructed using MCMCglmm, we used a non-informative prior, a burn-in of 3000, a thinning interval of 50, and a chain length of 53,000 (giving us 1,000 samples of the posterior). Each model contained unique residual variance parameters for each of the reproductive modes to allow for differences in variance between sexual and hybrid asexual species (Fig S3). For each MCMC sample, we calculated the difference between hybrid asexual and sexual species in the coefficient of variation (i.e., the standard deviation divided by the mean). We then compared the coefficients of variation between reproductive modes by subtracting the posterior distribution of the hybrid asexual species from that of the sexual species (Fig S4).

3) Subgroup Linear Models

Third, in order to examine the same effects we described in the linear mixed-effects model within individual subgroups (i.e., without needing to account for ancestry), we constructed linear models for (A) the tessellatus and neomexicanus groups independently (comparing each hybrid asexual to its parental species) and (B) all species with the same mitochondrial ancestry (*Aspidoscelis marmoratus*, *A. neomexicanus*, and *A. tessellatus*). While ideally the latter would include information regarding which of the two asexual groups first arose, such information is yet to be estimated (thus we assume they arose at the same time).

Chapter 2: Supplementary Results

Within-group ANOVAs

We observed the same general pattern in the *Tesselatus* and *Neomexicanus* groups (reduced endurance and respiration [State 3 and 4] in hybrid asexual species compared to parental sexual species), however, many of the models were not statistically significant (Table S4). RCR for both CI and CII showed no differences between sexual parents and hybrid asexuals. The models including the three species with similar mitochondrial haplotypes (*Aspidoscelis marmoratus*, *A. neomexicanus*, *A. tessellatus*) showed reduced endurance and maximal and basal respiration (State 3 and 4 for CI and CII) of both hybrid asexuals compared to *A. marmoratus*, although models for CI State 4 and CII State 4 were not statistically significant (Table S4). RCR for both CI and CII showed no differences between *A. marmoratus* and hybrid asexuals.

Study Limitations

Ideally our sampling would have focused entirely on female individuals and been larger than 30 individuals per species. Although we aimed to collect only female individuals, challenges of capturing sufficient females within our timeframe led to our inclusion of males in the dataset (the lizards within this study system are particularly difficult to capture). We found this justifiable based on previous work showing marginal sexual dimorphism within the same species examined in this study (Cullum, 1998), and on this basis we determined that bias due to sex would be minor. Although removing males severely reduces our sample size and dramatically decreases the power for our analyses, when we perform analyses in a female-only dataset we see that the direction of the patterns remain consistent with our conclusions.

Because post-hoc tests can impose an unnecessarily strict procedure to an underpowered dataset (Nakagawa, 2004), we do not perform post-hoc corrective tests. We recognize the possible contribution of false positives within our results and encourage replication.

References

- Cullum, A. J. , 1998. Sexual dimorphism in physiological performance of whiptail lizards (genus *Cnemidophorus*). *Physiological Zoology* 71:541–552.
- Katoh, K., K. Misaawa, K. Kuma, and T. Miyata. , 2002. Mafft: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Research* 30:3059–3066.
- Nakagawa, S. , 2004. A farewell to bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology* 15:1044–1045.
- Solis-Lemus, C., P. Bastide, and C. Ane. , 2017. Phylonetworks: A package for phylogenetic networks. *Molecular Biology and Evolution* 34.

Supplementary Tables and Figures

Table 3.S1 Sample information.

Table 3.S2 List of individuals used for phylogenetic inference.

Table 3.S3 Summary statistics

Table 3.S4 Within-group linear models

Table 3.S4 Differences in variance between sexual and hybrid asexual species

Figure 3.S1 Evolutionary relationships

Figure 3.S2 Endurance and mitochondrial respiration of sexual and hybrid asexual species

Figure 3.S3 Standard deviations for sexual and hybrid asexual species

Figure 3.S4 Posterior distributions of coefficient of variation

Table 3.S1: Sample Information

| Field ID | Species | Capture Date | Reprod | SVL | Sex | State | County | Lat | Long | Endurance | CI State 3 | CI State 4 | CI RCR | CII State 3 | CII State 4 | CII RCR |
|----------|--------------------------|--------------|--------|-----|-----|-------|----------|-----------|-------------|-----------|------------|------------|--------|-------------|-------------|---------|
| RLK178 | <i>A. inornatus</i> | 5/16/2019 | sex | 4.9 | M | TX | Brewster | 29.3024 | -103.1735 | 3.47 | 24.17 | 6.43 | 3.76 | 22.80 | 13.52 | 1.69 |
| RLK198 | <i>A. inornatus</i> | 5/24/2019 | sex | 5.3 | M | TX | Hudspeth | 30.776757 | -105.016495 | 12.37 | 28.38 | 9.12 | 3.11 | 28.19 | 15.72 | 1.79 |
| RLK200 | <i>A. inornatus</i> | 5/25/2019 | sex | 5.1 | F | TX | Hudspeth | 30.777342 | -105.016564 | 4.03 | 21.14 | 3.83 | 5.52 | 18.84 | 8.19 | 2.30 |
| RLK204 | <i>A. inornatus</i> | 5/26/2019 | sex | 5.8 | M | TX | Hudspeth | 30.776901 | -105.016291 | 9.17 | 30.09 | 8.8 | 3.42 | 21.76 | 13.41 | 1.62 |
| RLK210 | <i>A. inornatus</i> | 5/29/2019 | sex | 5.5 | M | NM | Doña Ana | 32.249198 | -106.821715 | 4.62 | 23.23 | 7.04 | 3.30 | 21.59 | 11.17 | 1.93 |
| RLK211 | <i>A. inornatus</i> | 5/29/2019 | sex | 5.6 | M | NM | Doña Ana | 32.249134 | -106.822207 | 12.05 | 33.35 | 10.36 | 3.22 | 33.71 | 17.71 | 1.90 |
| RLK186 | <i>A. marmoratus</i> | 5/18/2019 | sex | 8.2 | M | TX | Brewster | 29.1775 | -102.9989 | 5.68 | 29.83 | 12.81 | 2.33 | 28.61 | 18.23 | 1.57 |
| RLK187 | <i>A. marmoratus</i> | 5/18/2019 | sex | 6.9 | M | TX | Brewster | 29.1781 | -102.9977 | 5.38 | 30.11 | 8.57 | 3.51 | 27.80 | 15.01 | 1.85 |
| RLK191 | <i>A. marmoratus</i> | 5/19/2019 | sex | 7.8 | F | TX | Brewster | 29.1781 | -102.9974 | 6.55 | 22.85 | 4.8 | 4.76 | 19.74 | 6.36 | 3.11 |
| RLK193 | <i>A. marmoratus</i> | 5/19/2019 | sex | 7.4 | M | TX | Brewster | 29.1781 | -102.9974 | 5.50 | 25.01 | 7.29 | 3.43 | 23.52 | 10.91 | 2.16 |
| RLK205 | <i>A. marmoratus</i> | 5/25/2019 | sex | 8.6 | M | TX | Hudspeth | 30.775838 | -105.015899 | 6.07 | 21.36 | 6.01 | 3.55 | 20.08 | 9.03 | 2.22 |
| RLK206 | <i>A. marmoratus</i> | 5/28/2019 | sex | 8.7 | M | NM | Doña Ana | 32.25602 | -107.63207 | 8.42 | 29.70 | 7.99 | 3.72 | 26.08 | 12.29 | 2.12 |
| RLK208 | <i>A. neomexicanus</i> | 5/28/2019 | asex | 5.7 | F | NM | Doña Ana | 32.24994 | -106.82133 | 2.50 | 15.49 | 5.82 | 2.66 | 15.54 | 8.73 | 1.78 |
| RLK212 | <i>A. neomexicanus</i> | 5/29/2019 | asex | 6.9 | F | NM | Doña Ana | 32.24704 | -106.82327 | 3.13 | 20.93 | 6.35 | 3.30 | 20.26 | 9.53 | 2.13 |
| RLK213 | <i>A. neomexicanus</i> | 5/29/2019 | asex | 6.1 | F | NM | Doña Ana | 32.24631 | -106.82299 | 3.83 | 26.48 | 6.36 | 4.16 | 21.72 | 11.32 | 1.92 |
| RLK214 | <i>A. neomexicanus</i> | 5/30/2019 | asex | 6.7 | F | NM | Doña Ana | 32.25612 | -106.84009 | 4.45 | 15.24 | 4.01 | 3.80 | 22.81 | 10.40 | 2.19 |
| RLK165 | <i>A. septemvittatus</i> | 5/15/2019 | sex | 9.1 | M | TX | Brewster | 29.2589 | -103.2987 | 4.80 | 20.71 | 5.79 | 3.58 | 21.03 | 12.03 | 1.75 |
| RLK170 | <i>A. septemvittatus</i> | 5/15/2019 | sex | 8.5 | M | TX | Brewster | 29.254819 | -103.300742 | 7.65 | 20.50 | 6.82 | 3.01 | 21.58 | 13.11 | 1.65 |
| RLK171 | <i>A. septemvittatus</i> | 5/15/2019 | sex | 8.2 | M | TX | Brewster | 29.260019 | -103.297041 | 7.10 | 34.62 | 10.61 | 3.26 | 33.57 | 16.51 | 2.03 |
| RLK172 | <i>A. septemvittatus</i> | 5/15/2019 | sex | 8.5 | F | TX | Brewster | 29.259492 | -103.297782 | 15.85 | 35.46 | 10.13 | 3.50 | 30.08 | 15.54 | 1.94 |
| RLK174 | <i>A. septemvittatus</i> | 5/16/2019 | sex | 7.0 | F | TX | Brewster | 29.2959 | -103.1771 | 4.07 | 25.87 | 7.41 | 3.49 | 27.21 | 13.88 | 1.96 |
| RLK176 | <i>A. septemvittatus</i> | 5/16/2019 | sex | 7.2 | M | TX | Brewster | 29.29646 | -103.17565 | 4.50 | 28.16 | 10.7 | 2.63 | 26.18 | 16.14 | 1.62 |
| RLK177 | <i>A. septemvittatus</i> | 5/16/2019 | sex | 7.7 | F | TX | Brewster | 29.3024 | -103.1735 | 8.47 | 30.11 | 8.53 | 3.53 | 21.90 | 9.53 | 2.30 |

Table 3.S1: Sample Information (cont.)

| Field ID | Species | Capture Date | Reprod | SVL | Sex | State | County | Lat | Long | Endurance | CI State 3 | CI State 4 | CI RCR | CII State 3 | CII State 4 | CII RCR |
|----------|-----------------------|--------------|--------|-----|-----|-------|----------|-----------|-------------|-----------|------------|------------|--------|-------------|-------------|---------|
| RLK162 | <i>A. tessellatus</i> | 5/15/2019 | asex | 8.1 | F | TX | Brewster | 29.324813 | -104.035785 | 3.73 | 24.75 | 6.14 | 4.03 | 20.78 | 10.21 | 2.04 |
| RLK181 | <i>A. tessellatus</i> | 5/17/2019 | asex | 7.7 | F | TX | Brewster | 29.1622 | -103.6143 | 2.95 | 20.38 | 5.18 | 3.93 | 21.02 | 12.88 | 1.63 |
| RLK182 | <i>A. tessellatus</i> | 5/17/2019 | asex | 6.6 | F | TX | Brewster | 29.3055 | -103.1781 | 3.98 | 15.34 | 6.41 | 2.39 | 15.83 | 10.02 | 1.58 |
| RLK183 | <i>A. tessellatus</i> | 5/17/2019 | asex | 7.2 | F | TX | Brewster | 29.3055 | -103.1743 | 3.13 | 27.56 | 5.06 | 5.45 | 18.90 | 8.42 | 2.25 |
| RLK185 | <i>A. tessellatus</i> | 5/18/2019 | asex | 7.8 | F | TX | Brewster | 29.1775 | -102.9989 | 5.27 | 18.42 | 5.86 | 3.14 | 16.75 | 8.98 | 1.87 |
| RLK189 | <i>A. tessellatus</i> | 5/18/2019 | asex | 6.5 | F | TX | Brewster | 29.1779 | -102.9981 | 4.02 | 18.74 | 5.62 | 3.34 | 16.95 | 8.62 | 1.97 |
| RLK194 | <i>A. tessellatus</i> | 5/19/2019 | asex | 8.6 | F | TX | Brewster | 29.177572 | -102.997929 | 4.82 | 16.55 | 6.37 | 2.60 | 19.05 | 9.72 | 1.96 |

Table 3.S2: Individuals used for phylogenetic inference

| GenBank Accession | Species | Loci | Field ID |
|-------------------|------------------------------------|--------------|----------|
| OK104662 | <i>Aspidoscelis gularis</i> | MtDNA Genome | ASH22 |
| OK104663 | <i>Aspidoscelis gularis</i> | MtDNA Genome | ASH25 |
| OK104664 | <i>Aspidoscelis gularis</i> | MtDNA Genome | ASH63 |
| AY620808.1 | <i>Aspidoscelis inornatus</i> | ND4 | NA |
| AY620811.1 | <i>Aspidoscelis inornatus</i> | ND4 | NA |
| AY620813.1 | <i>Aspidoscelis inornatus</i> | ND4 | NA |
| AY620812.1 | <i>Aspidoscelis inornatus</i> | ND4 | NA |
| OK104676 | <i>Aspidoscelis inornatus</i> | MtDNA Genome | ASH120 |
| OK104677 | <i>Aspidoscelis inornatus</i> | MtDNA Genome | ASH121 |
| MZ673806 | <i>Aspidoscelis inornatus</i> | MtDNA Genome | ASH124 |
| OK104678 | <i>Aspidoscelis inornatus</i> | MtDNA Genome | ASH123 |
| OK104715 | <i>Aspidoscelis sexlineatus</i> | MtDNA Genome | KLC154 |
| OK104716 | <i>Aspidoscelis sexlineatus</i> | MtDNA Genome | KLC156 |
| OK104717 | <i>Aspidoscelis sexlineatus</i> | MtDNA Genome | KLC157 |
| OK104681 | <i>Aspidoscelis marmoratus</i> | MtDNA Genome | ASH131 |
| OK104685 | <i>Aspidoscelis marmoratus</i> | MtDNA Genome | ASH145 |
| OK104686 | <i>Aspidoscelis marmoratus</i> | MtDNA Genome | ASH146 |
| OK104730 | <i>Aspidoscelis neomexicanus</i> | MtDNA Genome | RLK89 |
| OK104731 | <i>Aspidoscelis neomexicanus</i> | MtDNA Genome | RLK90 |
| OK104668 | <i>Aspidoscelis tessellatus</i> | MtDNA Genome | ASH80 |
| OK104669 | <i>Aspidoscelis tessellatus</i> | MtDNA Genome | ASH97 |
| OK104665 | <i>Aspidoscelis tessellatus</i> | MtDNA Genome | ASH70 |
| OK104670 | <i>Aspidoscelis tessellatus</i> | MtDNA Genome | ASH98 |
| OK104687 | <i>Aspidoscelis marmoratus</i> | MtDNA Genome | ASH148 |
| AF026179.1 | <i>Aspidoscelis septemvittatus</i> | ND4 | NA |
| AF026181.1 | <i>Aspidoscelis septemvittatus</i> | ND4 | NA |
| AF026182.1 | <i>Aspidoscelis septemvittatus</i> | ND4 | NA |
| AF026170 | <i>Teius teyou</i> | ND4 | NA |
| AF151207 | <i>Kentropyx viridistriga</i> | ND4 | NA |

Table 3.S3: Summary statistics for species and reproductive modes

| | Species | | | | | Repro Mode | |
|--------------------|---------------------|----------------------|------------------------|--------------------------|-----------------------|--------------------|--------------------|
| | <i>A. inornatus</i> | <i>A. marmoratus</i> | <i>A. neomexicanus</i> | <i>A. septemvittatus</i> | <i>A. tessellatus</i> | Hybrid Asexual | Sexual |
| Endurance | | | | | | | |
| \bar{x} | 7.62 | 6.27 | 3.48 | 7.49 | 3.99 | 3.80 | 7.14 |
| σ | 4.09 | 1.14 | 0.85 | 4.06 | 0.84 | 0.84 | 3.30 |
| 95% CI | $\bar{x} \pm 1.67$ | $\bar{x} \pm 0.46$ | $\bar{x} \pm 0.42$ | $\bar{x} \pm 1.53$ | $\bar{x} \pm 0.32$ | $\bar{x} \pm 0.25$ | $\bar{x} \pm 0.76$ |
| CI State 3 | | | | | | | |
| \bar{x} | 26.73 | 26.48 | 19.54 | 27.92 | 20.25 | 20.00 | 27.09 |
| σ | 4.65 | 3.91 | 5.32 | 6.03 | 4.42 | 4.51 | 4.77 |
| 95% CI | $\bar{x} \pm 1.90$ | $\bar{x} \pm 1.59$ | $\bar{x} \pm 2.66$ | $\bar{x} \pm 1.28$ | $\bar{x} \pm 1.67$ | $\bar{x} \pm 1.36$ | $\bar{x} \pm 1.09$ |
| CI State 4 | | | | | | | |
| \bar{x} | 7.60 | 7.91 | 5.64 | 8.57 | 5.81 | 5.74 | 8.05 |
| σ | 2.34 | 2.76 | 1.11 | 1.97 | 0.55 | 0.75 | 2.26 |
| 95% CI | $\bar{x} \pm 0.95$ | $\bar{x} \pm 1.13$ | $\bar{x} \pm 0.556$ | $\bar{x} \pm 0.744$ | $\bar{x} \pm 0.21$ | $\bar{x} \pm 0.23$ | $\bar{x} \pm 0.52$ |
| CI RCR | | | | | | | |
| \bar{x} | 3.72 | 3.55 | 3.48 | 3.29 | 3.55 | 3.53 | 3.51 |
| σ | 0.91 | 0.77 | 0.65 | 0.35 | 1.04 | 0.88 | 0.69 |
| 95% CI | $\bar{x} \pm 0.37$ | $\bar{x} \pm 0.32$ | $\bar{x} \pm 0.33$ | $\bar{x} \pm 0.13$ | $\bar{x} \pm 0.39$ | $\bar{x} \pm 0.27$ | $\bar{x} \pm 0.16$ |
| CII State 3 | | | | | | | |
| \bar{x} | 24.48 | 24.31 | 20.08 | 25.94 | 18.47 | 19.06 | 24.96 |
| σ | 5.47 | 3.83 | 3.20 | 4.77 | 2.02 | 2.49 | 4.53 |
| 95% CI | $\bar{x} \pm 2.23$ | $\bar{x} \pm 1.56$ | $\bar{x} \pm 1.60$ | $\bar{x} \pm 1.80$ | $\bar{x} \pm 0.77$ | $\bar{x} \pm 0.75$ | $\bar{x} \pm 1.04$ |
| CII State 4 | | | | | | | |
| \bar{x} | 13.29 | 11.97 | 10.00 | 13.82 | 9.84 | 9.89 | 13.07 |
| σ | 3.35 | 4.24 | 1.12 | 2.51 | 1.51 | 1.32 | 3.29 |
| 95% CI | $\bar{x} \pm 1.37$ | $\bar{x} \pm 1.73$ | $\bar{x} \pm 0.56$ | $\bar{x} \pm 0.95$ | $\bar{x} \pm 0.57$ | $\bar{x} \pm 0.40$ | $\bar{x} \pm 0.75$ |
| CII RCR | | | | | | | |
| \bar{x} | 1.87 | 2.17 | 2.00 | 1.89 | 1.90 | 1.94 | 1.97 |
| σ | 0.24 | 0.52 | 0.19 | 0.24 | 0.23 | 0.21 | 0.36 |
| 95% CI | $\bar{x} \pm 0.10$ | $\bar{x} \pm 0.21$ | $\bar{x} \pm 0.10$ | $\bar{x} \pm 0.09$ | $\bar{x} \pm 0.09$ | $\bar{x} \pm 0.06$ | $\bar{x} \pm 0.08$ |

\bar{x} = mean, σ = standard deviation. Endurance is time in minutes. Respiration states are nmoles O₂ consumed per minute per mg of protein.

Table 3.S4: Within-group linear models

| | | tess model | tess-marm | tess-sept | neom model | neom-marm | neom-inor | marm model | neom-marm | tess-marm |
|-------------------------------|---------|------------|-----------|-----------|------------|-----------|-----------|------------|-----------|-----------|
| Log₁₀ Endur | | | | | | | | | | |
| | β | NA | -0.17 | -0.20 | NA | -0.043 | -0.43 | NA | -0.17 | -0.18 |
| | SE | NA | 0.076 | 0.074 | NA | 0.16 | 0.13 | NA | 0.074 | 0.11 |
| | p | 0.011 | 0.041 | 0.016 | 0.032 | 0.80 | 0.0060 | 8.60E-04 | 0.034 | 0.0031 |
| CI State 3 | | | | | | | | | | |
| | β | NA | -6.23 | -7.67 | NA | -6.94 | -7.19 | NA | -6.94 | -6.23 |
| | SE | NA | 2.74 | 2.63 | NA | 2.94 | 2.94 | NA | 2.88 | 2.48 |
| | p | 0.023 | 0.036 | 0.0096 | 0.056 | 0.034 | 0.024 | 0.039 | 0.030 | 0.025 |
| CI State 4 | | | | | | | | | | |
| | β | NA | -2.11 | -2.76 | NA | -2.28 | -1.96 | NA | -2.28 | -2.11 |
| | SE | NA | 1.07 | 1.03 | NA | 1.49 | 1.49 | NA | 1.14 | 0.98 |
| | p | 0.041 | 0.066 | 0.016 | 0.31 | 0.15 | 0.21 | 0.086 | 0.066 | 0.050 |
| CI RCR | | | | | | | | | | |
| | β | NA | -0.0044 | 0.27 | NA | -0.070 | -0.24 | NA | -0.070 | -0.0044 |
| | SE | NA | 0.43 | 0.41 | NA | 0.52 | 0.52 | NA | 0.56 | 0.49 |
| | p | 0.77 | 0.99 | 0.52 | 0.88 | 0.90 | 0.65 | 0.99 | 0.90 | 0.99 |
| CII State 3 | | | | | | | | | | |
| | β | NA | -5.84 | -7.47 | NA | -4.22 | -4.40 | NA | -4.22 | -5.84 |
| | SE | NA | 2.07 | 1.98 | NA | 2.85 | 2.85 | NA | 1.96 | 1.69 |
| | p | 0.0040 | 0.012 | 0.0016 | 0.28 | 0.16 | 0.15 | 0.012 | 0.049 | 0.0038 |
| CII State 4 | | | | | | | | | | |
| | β | NA | -2.14 | -3.98 | NA | -1.98 | -3.29 | NA | -1.98 | -2.14 |
| | SE | NA | 1.60 | 1.54 | NA | 2.19 | 2.19 | NA | 1.79 | 1.54 |
| | p | 0.059 | 0.20 | 0.019 | 0.35 | 0.38 | 0.16 | 0.36 | 0.29 | 0.19 |
| CII RCR | | | | | | | | | | |
| | β | NA | -0.28 | 0.0061 | NA | 0.13 | -0.17 | NA | -0.17 | -0.27 |
| | SE | NA | 0.19 | 0.18 | NA | 0.24 | 0.24 | NA | 0.23 | 0.20 |
| | p | 0.28 | 0.17 | 0.97 | 0.39 | 0.59 | 0.49 | 0.41 | 0.48 | 0.19 |

Effect sizes (β), standard error (SE), and p-values (p) for each response variable. The tess model (left) includes effects between maternal-hybrid (tess-marm) and paternal-hybrid (tess-sept). The neom model (center) includes effects between maternal-hybrid (neom-marm) and paternal-hybrid (neom-sept). The marm model (right) includes effects between hybrids and the maternal offspring (neom-marm and tess-marm). We considered effects with $p < 0.05$ as statistically significant.

Table 3.S5: Differences in variance between sexual and hybrid asexual species

| | lme | | | | MCMCglmm | | | |
|-------------|----------------------|--------------|---------|-------|----------------------|----------------------------|-------------|------|
| | σ_s, σ_a | SEs, SEa | L Ratio | p | σ_s, σ_a | (95% HPDs), (95% HPD a) | ΔCV | P |
| Endurance | 0.18, 0.090 | 0.030, 0.016 | 5.19 | 0.023 | 0.20, 0.11 | (0.13-0.26), (0.067-0.18) | 0.027 | 0.75 |
| CI State 3 | 4.77, 4.51 | 0.56, 0.84 | 0.041 | 0.84 | 5.27, 5.29 | (3.38-7.20), (3.09-8.34) | -0.076 | 0.21 |
| CI State 4 | 2.26, 0.75 | 0.36, 0.19 | 11.44 | 7e-04 | 2.47, 0.89 | (1.63-3.43), (0.512-1.37) | 0.14 | 0.96 |
| CI RCR | 0.69, 0.88 | 0.17, 0.21 | 0.81 | 0.37 | 0.75, 0.103 | (0.49-1.02), (0.59-1.63) | -0.08 | 0.20 |
| CII State 3 | 4.53, 2.49 | 0.62, 0.48 | 3.95 | 0.047 | 4.95, 2.86 | (3.469-6.89), (1.63-4.39) | 0.069 | 0.83 |
| CII State 4 | 3.29, 1.32 | 0.51, 0.36 | 8.28 | 0.004 | 3.55, 1.58 | (2.55-4.99), (0.90-2.48) | 0.11 | 0.94 |
| CII RCR | 0.36, 0.21 | 0.084, 0.039 | 8.55 | 0.088 | 0.38, 0.25 | (0.26-0.52), (0.14-0.38) | 0.063 | 0.89 |

LME SECTION: Standard deviations for the lme model fitting two residual variation parameters to the data (for the reproductive modes) are shown under σ_s, σ_a (standard deviation for sexual species, hybrid asexual species). The standard errors for each reproductive mode (SEs, SEa) were estimated using a nonparametric bootstrap approach. For each response variable, a likelihood ratio test was performed between a model fitting a single residual variation parameters and a model fitting two residual variation parameters (for the reproductive modes). L Ratio is the likelihood ratio score. p is the p-value for the likelihood ratio test, with $p < 0.05$ indicating the model with two residual variation parameters is a better fit for the data. MCMCglmm SECTION: Standard deviations determined by calculating the mean of the posterior distribution for the MCMCglmm model fitting two residual variance parameters to the data (for the reproductive modes) are shown under σ_s, σ_a (standard deviation for sexual species, hybrid asexual species). (95% HPDs), (95% HPD a) is the 95% credible interval for the standard deviations estimated from the posterior distribution (see Fig S3 for a posterior distribution of the standard deviations estimated from the MCMCglmm approach). ΔCV is the mean difference for the coefficient of variation (sexual - asexual hybrid). P is the posterior probability that the sexual coefficient of variation is greater than the hybrid asexual coefficient of variation Fig S4.

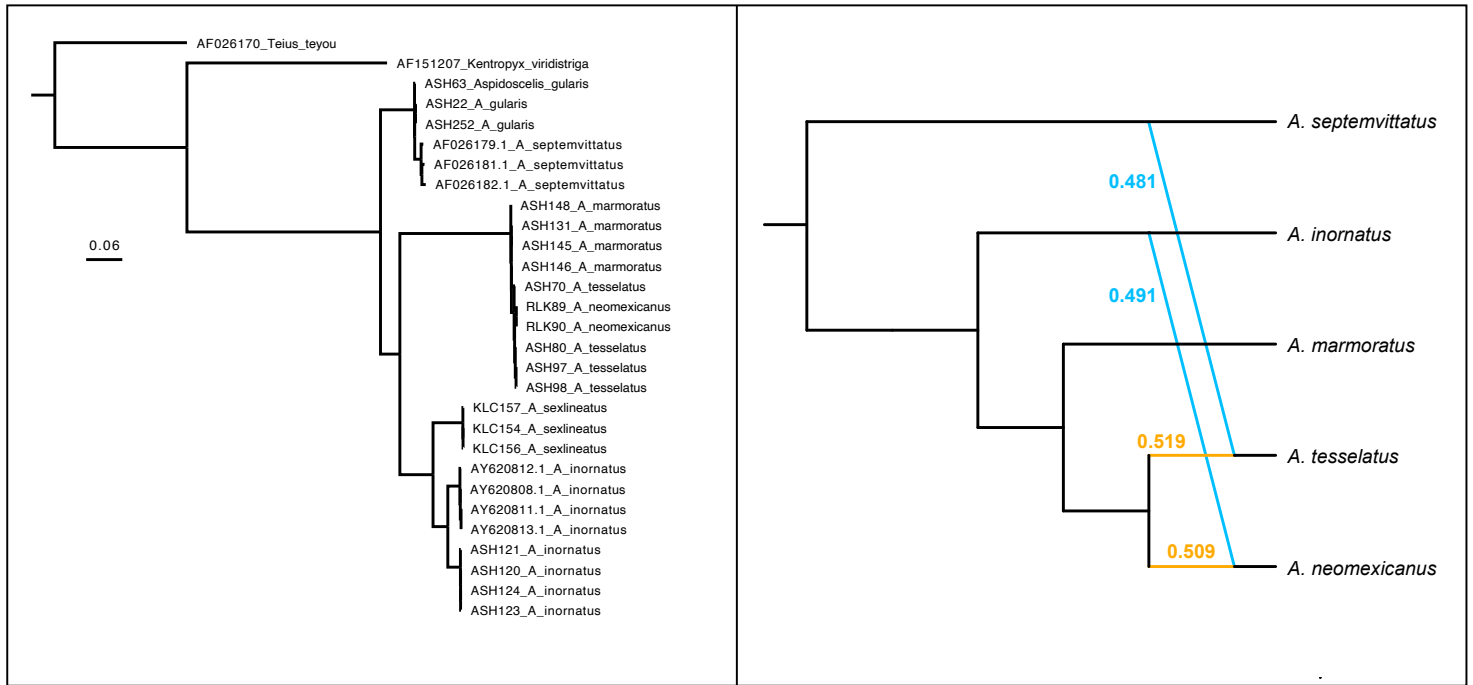


Figure 3.S1: Left: mitochondrial consensus tree constructed from 1000 bootstrap replicates of the mitochondrial alignment with subsequent maximum likelihood tree estimation using IQ-Tree with individuals described in Table S3. Right: Phylogenetic network estimated using PhyloNet using mitochondrial relationships from the mitochondrial consensus tree and an estimated paternal ancestry. Blue lines represent estimated contribution from paternal ancestor, and yellow lines represent estimated contribution from maternal ancestor.

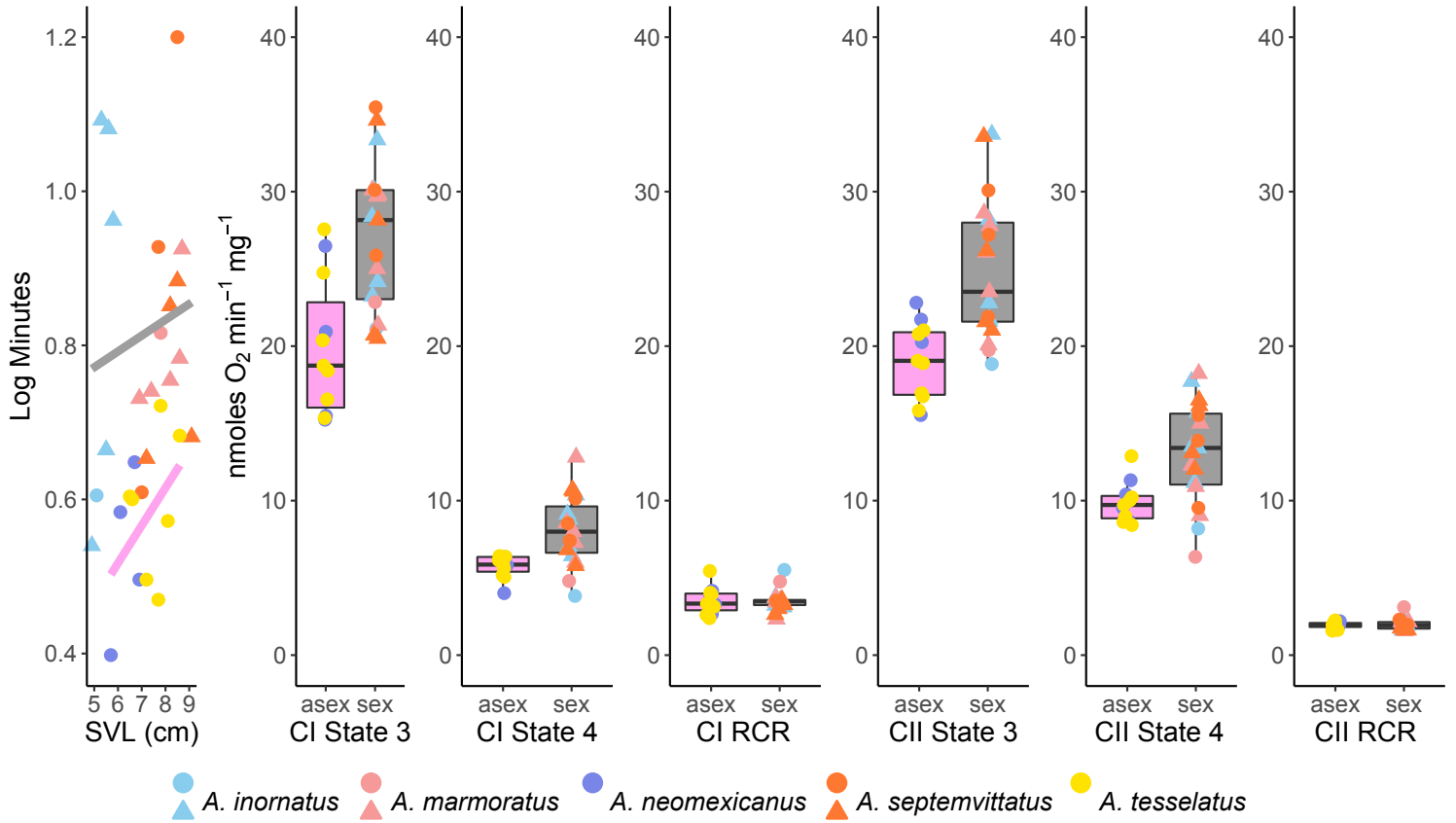


Figure 3.S2: Plots showing effect of hybrid asexuality on mitochondrial respiration. Each plot corresponds to the individual data from Table S1. All models showed a significant difference between sexual and hybrid asexuals for the response variables for both lme and phylonetworklme approaches ($p < 0.05$) except for the two models for RCR. Point symbols represent males (triangles) and females (circles)

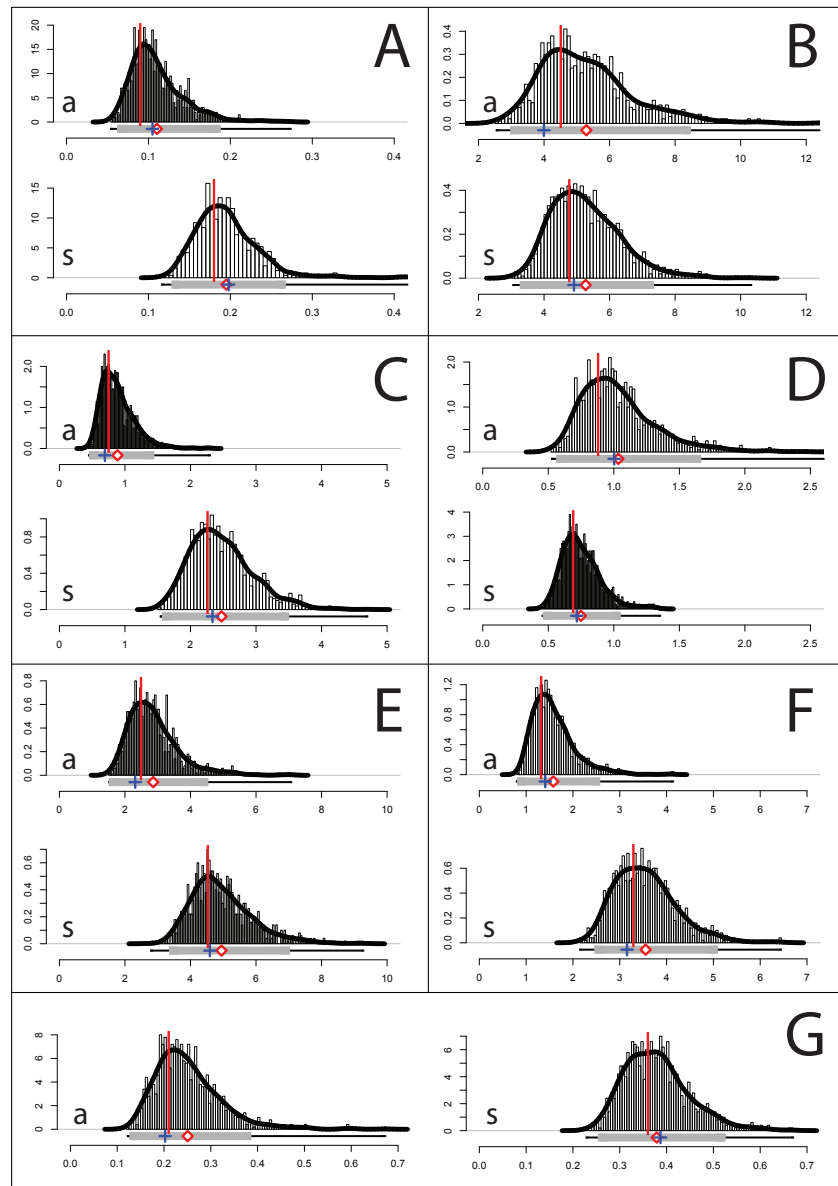


Figure 3.S3: Plots showing the approximate posterior distribution for the standard deviation for sexual (s) and hybrid asexual (a) species for each of the response variables: (A) Endurance, (B) CI State 3 Respiration, (C) CI State4 Respiration, (D) CI RCR, (E) CII State 3 Respiration, (F) CII State 4 Respiration, and (G) CII RCR. Posteriors distributions were estimated using the MCMCglmm package in R. The gray bar on the x-axis is the 95% credible interval, with the red diamond and blue cross marking the mean and mode, respectively. The red line marks the standard deviation estimated from the nlme::lme approach.

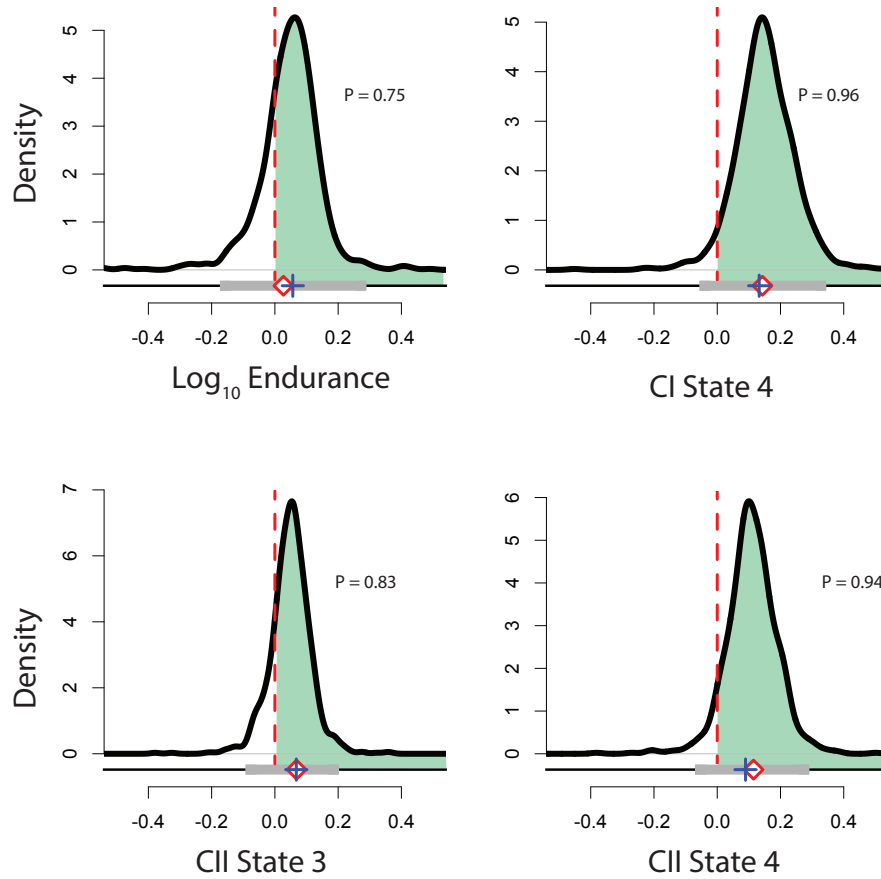


Figure 3.S4: Plots showing the approximate posterior distribution for the sexual-aseual difference between coefficient of variance posterior distributions. The gray bar on the x-axis is the 95% credible interval, with the red diamond and blue cross marking the mean and mode, respectively. P = posterior probability that the sexual coefficient of variation is greater than the asexual coefficient of variation. The dashed red line marks the line where the sexual and asexual coefficient of variations are equal, with the green area under the curve showing the area of the posterior distribution with greater coefficient of variation in sexual species compared to hybrid asexual species.

Chapter 4: Divergence in genetic variation and expression in molecular networks that underlie aging between fast- and slow-aging garter snakes.

Formatted for Aging Cell

Randy L. Klabacka¹, Anne M. Bronikowski^{2,3,4}, Suzanne E. McGaugh⁵, Dawn Reding^{2,6}, Andrew Lithio⁷, Daniel Nettleton⁷, Laurie S. Stevison¹, Jessica Judson⁴, Tonia S. Schwartz^{1,2*}

¹Department of Biological Sciences, Auburn University, Auburn, AL USA

²Department of Ecology, Evolution, & Organismal Biology, Iowa State University, Ames, IA USA

³Department of Integrative Biology, Michigan State University, Hickory Corners, MI USA

⁴Kellogg Biological Station, Michigan State University, Hickory Corners, MI USA

⁵Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, Minnesota, USA

⁶Department of Biology, Luther College, Decorah, Iowa, USA

⁷Department of Statistics, Iowa State University, Ames, Iowa, USA

Abstract

Understanding the genetic underpinnings of natural variation in rates of aging can identify mechanisms for how such variation evolves. The process of organismal aging has been explained by many gene networks in laboratory model organisms. However, a thorough understanding of the mechanisms of aging requires examination of patterns in wild populations. Phenotypic divergence between fast- and slow-aging ecotypes of western terrestrial garter snakes (*Thamnophis elegans*) is associated with differences in many aging-related cell signaling pathways. Here we test whether there are genetic underpinnings for these divergent ecotypes by examining gene networks associated with aging: Metabolic Function (including Insulin and Insulin-like Signaling [IIS] and Oxidative Phosphorylation), Macromolecule Damage and Repair, and Stress Adaptation. We assess (1) divergent transcriptomic responses between the ecotypes in response to acute heat stress, and (2) divergence in sequence variation of putative aging loci through elevated genetic distance (F_{ST} and D_{XY}) and amino acid polymorphisms between fast- and slow-aging snake ecotypes. We find significant divergence between the two aging ecotypes in regulation of the IIS, Oxidative Phosphorylation, and DNA repair networks, particularly in response to stress, as well as genetic divergence in key nodes in these networks. We incorporate findings of previous work conducted in the Eagle Lake garter snake system and compare our findings to what is known from model organisms. This research provides insight into the evolution of life-history traits, including aging, and highlights the importance of studying aging within and across species.

4.1 Introduction

Demographic aging is measured at the population level and defined by accelerating age-specific adult mortality. Such aging is widespread across animals (Reinke et al., 2022) and is underpinned by conserved cellular phenotypes (e.g., accumulation of damaged macromolecules, genomic instability, deregulation of cellular signaling) (López-Otín et al., 2013). These cellular aging phenotypes lead to deterioration of an individual's functional traits, with resultant increased frailty and organismal aging. Variation in cellular aging phenotypes can originate from sequence variation and altered gene expression in evolutionarily conserved molecular networks. Integrated signaling across these molecular networks gives rise to shared “pillars/hallmarks” of aging (*sensu* (López-Otín et al., 2013; Kennedy et al., 2014)), which can be defined at cellular and organismal levels (e.g., genomic instability and sarcopenia, respectively). Of particular interest are the hallmark-related cellular phenotypes that are robust to animal diversity, many of which were discovered in model laboratory species. Interesting questions arise as to (i) whether homologous molecular networks govern aging in natural populations (i.e., where aging evolved); and (ii) whether the genetic underpinnings of aging are shared across animals. Answers to these questions are critical for understanding which nodes within molecular networks of aging are flexible – and thereby recommend themselves as targets for biomedical intervention – versus those that are constrained or species-specific, and unlikely to result in meaningful increases to healthspan.

Comparisons between long- and short-lived species have further corroborated that conserved cellular aging phenotypes underlie organismal aging (Kennedy et al., 2014) – and, ultimately, demographic aging and lifespan – due to variation in the genetic mechanisms controlling these

phenotypes. For example, Tian et al. (2017) reviewed genes with roles in DNA repair, tumor suppression, or insulin and insulin-like cellular signaling (IIS) that distinguish short- and long-lived mammals. Additionally, McGaugh et al. (2015) and Passow et al. (2019) have quantified variation in selection across genes in these same molecular networks both within mammals and reptiles, and across amniotes. In a larger comparative framework, Hoekstra et al. (2020) reviewed studies on cellular aging mechanisms in reptiles and further defined their use in compliment to mammals. As the sister clade to mammals, similarities and differences in aging mechanisms between reptiles and mammals suggest evolvability in these molecular networks and additional networks that may account for variation in cellular aging hallmarks and aging demography within and between amniotes.

Intraspecific genetic variation that correlates with fast versus slow demographic aging can lead to the discovery of novel genes and gene networks involved in aging in addition to corroborating such variation observed across species. While much of this research has focused on differences between the sexes in aging and lifespan (Bronikowski et al., 2022), a complimentary approach is to study the evolution and function of molecular networks among populations within a species that have polymorphism in aging rates. Our own work in natural populations of western terrestrial garter snake (*Thamnophis elegans*) has utilized this latter approach to better understand the evolution of aging at both demographic and cellular scales. These populations have diverged in many aging traits including growth, maturation, reproductive rate, and mortality (Bronikowski and Arnold, 1999; Miller et al., 2014; Schwartz et al., 2015). These suites of co-evolved life-history traits establish that populations in close proximity (ca. 25km) can evolve independently in aging and lifespan, and their divergence has resulted in Fast- and Slow- aging

ecotypes (FA and SA, hereafter). Snakes of the FA ecotype have high mortality and lower median and maximum lifespan than SA snakes (age-in-years at 50th and 95th percentile of survivorship: FA = 5, 9; SA = 9, 13) (Miller et al., 2014; Reinke et al., 2022). Previous work in this system has used classic quantitative genetic approaches (i.e., common gardens, laboratory reciprocal transplants) to understand the contributions of genetic (Bronikowski, 2000; Gangloff et al., 2015) and environmental (Gangloff et al., 2016; Addis et al., 2017) variation to the life-history divergence between FA and SA snakes. Remarkably, genetic differentiation between FA and SA populations is relatively low at neutral nuclear markers (microsatellites) and at the whole-genome level (ca. $F_{ST} \sim 0.05$), yet high at the mitochondrial genome (ca. $F_{ST} \sim 0.45$) (Manier and Arnold, 2005; Schwartz et al., 2015; Gangloff et al., 2020). The many putative cellular mechanisms underlying the divergence between FA and SA snakes have been reviewed in Schwartz and Bronikowski (2011) and Hoekstra et al. (2020). FA and SA snakes are known to differ in measures representing three main cellular hallmarks of aging (*sensu* Kennedy et al., 2014): (1) Metabolism, (2) Macromolecule Damage, and (3) Stress Adaptation (summarized in Table 4.1). Regarding metabolism, we found differences that often interact with temperature in whole-animal metabolism, cellular respiration, and aspects of nutrient sensing pathways between these two ecotypes (Schwartz and Bronikowski, 2013). Moreover, the two ecotypes have diverged in circulating hormone levels and gene expression of their receptors within the Insulin and Insulin-like cellular signaling network (IIS) (Reding et al., 2016). Regarding macromolecular damage, we have found ecotype differences in the production of free radicals and in the DNA-damage-repair response (Schwartz and Bronikowski, 2013). Finally, regarding stress adaptations, we have found consistent differences between FA and SA snakes in both innate and acquired immunity (Sparkman and Palacios, 2009; Palacios et al., 2013).

Here we interrogate these natural populations for divergence in gene expression and segregating allelic variation to test the hypothesis that the molecular genetic networks regulating three aging hallmarks (metabolism, macromolecule damage, and stress adaptation) are diverging between the FA and SA ecotypes. First, using transcriptome data we test for divergence in the regulation of molecular pathways related to aging between the ecotypes under control laboratory conditions and in response to heat stress. Second, we use genetic sequence data to evaluate allelic variation within these candidate networks and test for molecular network divergence between the FA and SA snakes in the wild.

4.2 Materials and Methods

4.2.1 Description of study species and habitats

Here we interrogate these natural populations for divergence in gene expression and segregating allelic variation to test the hypothesis that the molecular genetic networks regulating three aging hallmarks (metabolism, macromolecule damage, and stress adaptation) are diverging between the FA and SA ecotypes. First, using transcriptome data we test for divergence in the regulation of molecular pathways related to aging between the ecotypes under control laboratory conditions and in response to heat stress. Second, we use genetic sequence data to evaluate allelic variation within these candidate networks and test for molecular network divergence between the FA and SA snakes in the wild.

4.2.2 Genetic Data

We use two data types to address our goals: RNAseq data for liver gene expression (detailed methods in Schwartz and Bronikowski [2013] and Schwartz et al. [2015]); and sequence variation data from three sources – the RNAseq data, DNA allele sequence capture (detailed methods in Gangloff et al. [2020]) and whole genome sequences.

4.2.2.1 Gene Expression Data, Differential Expression, and Enrichment Analyses

Data Collection. Our first experiment tests for differential gene expression between FA and SA snakes and their responses to an acute metabolic stress (including an interaction between ecotype and stress). We conducted two heat stress experiments in 2008 and 2012 (hereafter referred to as HS2008 and HS2012). The physiological responses to heat stress in these two ecotypes from these experiments are described in-detail in Schwartz and Bronikowski (2013). Each heat stress experiment was conducted on lab-born juveniles at age 1.2 or 1.3 years with a 2 x 2 factorial design (temperature treatment x ecotype). We assigned animals (siblings split across treatments) to either a control temperature treatment (27 °C – i.e., a temperature within their normal rearing range) or a heat stress treatment (37 °C) for two hours, after which organs were dissected and flash-frozen. For HS2008, we assayed three unrelated females for each temperature-by-ecotype combination (total N = 12). For HS2012, we assayed five unrelated individuals (three males and two females) for each temperature-by-ecotype combination (total N = 20). Liver RNA was extracted with an RNAeasy kit (Qiagen) for 80-cycle single read sequencing on the Illumina GXII platform (one library per lane for HS2008), or 100 cycle paired-end sequencing on the Illumina HiSeq platform (HS2012). Population and ecotype sampling for this dataset are shown in Figure 4.1.

Differential Expression. The bioinformatic analysis pipeline for the RNAseq data is depicted in Figure 4.S1, and the scripts can be found on GitHub (https://github.com/rklabacka/ThamnophisElegans_FunctionalGenomics2021). For the two RNAseq datasets we removed reads shorter than 36bp, and cleaned raw sequencing reads using Trimmomatic (version 0.36) (Bolger et al., 2014) – removing leading and trailing bases with quality value lower than 25 (for HS2012 data) or 20 (for HS2008 data), and a sliding window of six with a quality value lower than 30. We mapped the clean reads to the *Thamnophis elegans* reference genome (rThaEle1.pri; GenBank GCA_009769535.1, Bronikowski et al., 2019) using HiSat2 (version 2.1.0; Pertea et al., 2016) and counted reads mapped to genes defined in the genome annotation file (.gff) with Stringtie (Pertea et al., 2015, 2016) using the PrepDE.py script to generate read counts. Genes with low or no expression were removed for the analysis of differential gene expression in EdgeR (Robinson et al., 2010). To identify individual genes that were differentially expressed between FA and SA snakes in response to heat stress treatment, the two datasets were first analyzed independently in the R packages EdgeR and Limma-Voom (Law et al., 2014, 2018; Ritchie et al., 2015). Log-CPM data were normalized using the TMM method in EdgeR, and variance was modeled in Limma-Voom. We assessed five linear models using Limma-Voom with the following fixed effects: (1) Ecotype (FA, SA), (2) Temperature (Control, Heat stress), (3) the Interaction between Ecotype and Temperature; and two analyses on the following subsets of data, (4) Temperature within FA, (5) Temperature within SA. We used an Empirical Bayes smoothing of standard errors and a false discovery rate of $FDR < 0.05$.

To identify genes that were consistently differentially expressed across both RNAseq datasets, we used the results from the independent analyses of the datasets described above to conduct a

meta-analysis using the R package metaSeq (Tsuyuzaki and Nikaido, 2021). For each linear model described above, the results for HS2008 and HS2012 datasets were combined by converting p-values from each experiment into z-values via the inverse normal cumulative distribution function, also matching the sign with the direction of the estimated change. This procedure provides additional evidence for differential expression when the estimated direction of a test is the same in both experiments, while weakening evidence for differential expression when the estimated directions of a test are different for each experiment.

Functional Pathway Enrichment. To test for divergence in gene expression at the level of molecular pathways, we used Gene Set Enrichment Analysis (GSEA) (Mootha et al., 2003; Subramanian et al., 2005). For this analysis we only used HS2012 due to higher sample size, and deeper and more consistent read coverage across samples. Using the results from the linear models described above, for each model we calculated a rank value for each gene using the sign of the fold change multiplied by the negative log of the p-value (Reimand et al., 2019). As a result, highly upregulated genes with small p-values are ranked at the top of the list and highly downregulated genes with small p-values are ranked at the bottom of the list. For this analysis we used 10,602 genes that passed the no/low expression filter and had annotation that mapped to a human gene ID in the GSEA database. In GSEA, we tested 146 KEGG pathways (Gene Set: c2.cp.kegg.v7.5.1.symbols.gmt [Curated]) for enrichment of genes that are at the leading edge of each ranked list (i.e., Ecotype, Temperature, Ecotype-by-Temperature Interaction, and Temperature within each Ecotype). To visualize the relationships among the KEGG pathways that were enriched in the leading edge of the ranked genes for each model, the GSEA results were imported into Cytoscape (Shannon et al., 2003) using the Enrichment Map App (Merico et

al., 2010) keeping all pathways that were significantly enriched at FDR 0.1 with edges set for overlap similarity coefficient between pathways > 0.3 . For the resulting network of enriched KEGG pathways for each model, we highlight pathways that were associated with our cellular hallmarks of aging (i.e., metabolism, macromolecule damage, and stress adaptation).

4.2.2.2 Sequence Variation

Sequence Capture Data Collection. We designed MyBaits probes to capture DNA sequences that targeted exons for 455 nuclear genes in molecular pathways of relevance to various life-history and behavioral differences between the two ecotypes. The full methodology for this sequence capture is published in Gangloff et al. (2020) where we focused on the divergence of the mitochondrial genome between the ecotypes. We sequenced 100bp paired-end reads on Illumina HiSeq 2000. Probes were blasted (blastx) against the *Thamnophis elegans* reference genome (rThaEle1.pri; GenBank Assembly Accession GCA_009769535.1; Bronikowski et al., 2019) to confirm the identification of the putative annotation. We used these blast results to create new annotation files for three categories of data: (1) *gene*: genic regions (SeqCap_CapturedGenes.gff) including exons and parts of introns that had coverage, (2) *exon*: transcribed regions (the exons of genic regions; SeqCap_CapturedExons.gff), and (3) *CDS*: translated regions (the coding sequence (CDS) of genic regions; SeqCap_CapturedCDS.gff). These reference annotation files were used downstream to bin variants into the gene, exon, and CDS categories. In a few instances, a probe matched to multiple genes in the genome and was removed from the analyses. We probed the genomes of 94 individuals from our FA and SA populations (Figure 4.1). The probes successfully matched genic region in 397 genes in the *T. elegans* reference genome Table 4.S1. We refer to this dataset as “Seq-Cap” hereafter.

Whole Genome Data Collection. To expand our sample size for our genes of interest, we used a subset of data from a larger whole genome sequencing project (Judson, 2021). Specifically, we included the variable sites (see “SNP identification and filtering” section below) identified in our Seq-Cap focal genes from whole-genome resequencing of 122 additional individuals (Fig. 1). In brief, we extracted DNA from blood cells using either a Qiagen DNeasy Blood and Tissue Kit or a phenol-chloroform DNA extraction protocol (Sambrook et al. 1989). To confirm high DNA quality before sequencing, we quantified DNA with a NanoDrop™ 2000 Spectrophotometer and assessed DNA purity and quality using a 1% agarose gel. For library preparation of DNA for whole genome resequencing, we prepared libraries of 250bp insert size and sequenced libraries on a DNBSEQ-G400 with 100bp paired-end reads (BGI Genomics). We refer to this dataset as “WGS” hereafter.

SNP identification and filtering. We identified SNPs from the two RNA-Seq datasets and the Seq-Cap dataset as described in Figure 4.S2. For the two RNAseq datasets and the Seq-Cap dataset, we trimmed adapters and cleaned raw sequencing reads using Trimmomatic (version 0.36). For mapping, we used the *Thamnophis elegans* reference genome (rThaEle1.pri, GenBank Assembly Accession: GCA_009769535.1, (Bronikowski et al., 2019)). We mapped the cleaned RNAseq reads to the reference using the splice-aware aligner HiSat2 (version 2.1.0; Pertea et al., 2016), and mapped the sequence capture data to the reference using the BWA -mem algorithm (version 0.7.15; Li and Durbin, 2009). The resulting SAM files for both the RNA-Seq and Seq-Cap data were processed for calling SNPs by first adding read groups to reads and marking PCR duplicates using Picard (version 2.1.0; Broad Institute, 2016). We called SNPs using GATK

(version 4.1.7.0; Van der Auwera and O’Conner, 2020) following the GATK best practices guidelines for non-model organisms (DePristo et al., 2011; Poplin et al., 2018; see **Figure 4.S2**). We then used GATK to filter SNPs following the best practices guidelines, and we performed additional filtering to remove: genotypes with low quality ($GQ < 20$; VCFtools version 0.1.17 [Danecek et al., 2011]); genotypes with low depth ($DP < 10$; VCFtools v0.1.17); multiallelic SNPs (more than 2 alleles; BCFtools version 1.2.3 [Danecek et al., 2021]); low frequency alleles (singletons; VCFtools 0.1.17); and sites with low sample representation (>70 percent samples missing; VCFtools v0.1.17). Following filtering, we annotated the SNPs using the reference genome annotation and BCFtools (v1.3.2), and we reduced the SNPs to include only sites within the genes of our final probe set (Table 4.S1). We then categorized variants as within a gene, within an exon, and/or within a CDS using the annotate tool from BCFtools (v1.2.3).

Obtaining Focal SNP calls from Whole Genome Sequencing Data. After calling SNPs from our RNAseq and Seq-Cap datasets, we called genotypes of these same variant sites from the WGS dataset to increase sample size and broaden the number of populations for population genetic analyses. The raw reads from the whole genome sequencing dataset were trimmed for adapter sequences, contained more than 10% unknown bases, or had more than 50% low quality bases ($Q \leq 12$) with cutadapt v2.5 (Martin 2011). We downloaded the NCBI *Thamnophis elegans* reference genome (rThaEle1.pri, GenBank Assembly Accession: GCA_009769535.1, Bronikowski et al., 2019) and used the -mem algorithm (Li 2013) in bwa v0.7.17 (Li and Durbin 2009; Li and Durbin 2010) to align all reads to the reference. Before variant calling, we used SAMtools v1.9 (Li et al. 2009) to remove unmapped reads and reads not in primary alignment. We used the Sentieon DNaseq workflow (v 201808.01; Kendig et al., 2019) to call variants across individuals. For each individual, we marked and removed duplicate reads with the –LocusCollector and –Dedup algorithms, then realigned reads around insertions and deletions (indels) with the –Realigner algorithm before using the –Haplotype algorithm to generate genomic variant call format files (GVCFs). We joint-called variants across all GVCFs with a minimum base quality > 20 using the –GVCFTyper algorithm, which results in a VCF file of variants called across all individuals. We limited this VCF to only SNPs using GATK v.4.0.4.0 (McKenna et al. 2010; DePristo et al. 2011) and filtered sites from the whole-genome VCF based on the SNPs recovered from the sequence capture VCF with VCFtools v0.1.14 (Danecek et al. 2011).

Merging Datasets. We used the variant sites from the RNA-Seq + Seq-Cap VCF as a guide for SNP isolation from the whole-genome sequencing data. We then merged the RNA-Seq + Seq-

Cap VCF file and the reduced WGS VCF file using the merge tool in BCFtools (v1.2.3; Danecek et al., 2021). Predicting Peptide Sequences and SNP functions. We created nucleotide sequences of our targeted regions for each sample by inserting sample-specific SNPs into the reference genome for each individual from the RNA-Seq and Seq-Cap dataset using the GATK tool FastaAlternateReferenceMaker. To eliminate bias from unsequenced regions, we masked each alternate reference for individual genotypes with read depth less than two using BEDtools (v2.30.0 (Quinlan and Hall, 2010) sub-commands genomecov and maskfasta. We then extracted coding sequences using the program gffreader (v0.12.7; Pertea and Pertea, 2020) and translated them using the custom python script parse_and_translate.py within our GitHub repository. We created multiple sequence alignments for each gene of interest (including all RNA-Seq and Seq-Cap samples).

Functionally Annotating Variants. To identify nonsynonymous and synonymous SNPs, we functionally annotated the merged CDS VCF file using snpeff (v27; Cingolani et al., 2012) and created VCF files for synonymous and nonsynonymous SNPs. We obtained additional functional information for nonsynonymous SNPs of interest using the software package SIFT4G (commit 8fd9f40; Vaser et al., 2016). The lack of a *Thamnophis* genome annotation stored in the Ensembl genomics database required us to create a SIFT database from our local genome file and annotation (Ng, 2020). In brief, we downloaded the uniref90 protein database and used the perl script make-SIFT-db-all.pl with our configuration file. We then used the SIFT4G_Annotator.jar script (Ng, 2019) to annotate our VCF with SIFT scores and categories. In analyses of functional implications for variants of interest, we filtered out any SNPs with low confidence SIFT scores (as reported in the SIFT4G output).

Population Genetic Analyses. To estimate relative (F_{ST}) and absolute (D_{XY}) genetic divergence, we calculated pairwise F_{ST} and D_{XY} between populations for each gene (i.e., using SNPs from exon regions in target genes) and for each site with a nonsynonymous SNP. For these calculations, we used the `popgenWindows.py` script in the `genomics_general` software package (Martin, 2019). This script conducts pairwise calculations of F_{ST} (more specifically K_{st} [Hudson et al., 1992]) and D_{XY} in sliding windows across sites and per each site. In addition, this software calculates pairwise averages of nucleotide diversity (π) for each specified group. Pairwise population calculations were categorized as either “within-ecotype” or “between-ecotype” comparisons (e.g., comparing two FA populations versus comparing an FA and an SA population). We then used R to perform two-tailed T-tests of the distribution of F_{ST} or D_{XY} of within and between ecotype comparisons for both the per gene and per site estimates (R Core Team, 2020). We identified loci with a significant (< 0.05) p-value for sites of interest from the t-test. For per SNP estimates, we constructed a linear mixed effects model to (1) test for differences in divergence estimates between sub-categories of the Hallmarks of Aging, (2) test for an overall difference in the pairwise population divergence estimates from within-ecotype comparisons vs between-ecotype population comparisons, and (3) test for an interaction between ecotype comparison and gene network on F_{ST}/D_{XY} values. considered p-values < 0.05 as statistically significant. We also calculated Tajima’s D statistic for each locus using `VCFTools` (v0.1.14; Danecek et al., 2011).

4.3 Results

4.3.1 Differential Gene Expression

The HS2008 and HS2012 RNA-Seq datasets varied in quality and depth of sequencing. The HS2008 data set had 5 to 31 million reads per library with average quality scores ranging from 21 to 36, and the HS2012 dataset had 42 to 90 million reads per library with quality scores ranging from 35 to 36, due to advances in high-throughput sequencing (Table 4.S1). Of the 21,554 annotated genes in the *T. elegans* genome, 13,067 passed the low/no expression filter in HS2008, and 14,431 in the HS2012 dataset, with 12,897 genes expressed in both datasets (Figure 4.S3).

In general, the meta-analysis increased the number of differentially expressed genes (DEGs) identified and had a large overlap with the HS2012 dataset (Figure 4.2, Figure 4.S3), providing confidence in the two datasets and demonstrating the increased power of the HS2012 dataset relative to HS2008. As expected, acute heat stress had a strong transcriptomic response (the meta-analysis identified 962 DEG), with approximately twice as many genes up-regulated compared to down-regulated (Figure 4.2). The list of differentially expressed genes contained the expected heat shock proteins (including our candidate gene HSP70; (Schwartz and Bronikowski, 2013), validating the 37 °C temperature treatment as a heat stress on these garter snakes.

Between ecotypes overall, we found only 11 DEG. As expected with a small sample size and a large number of genes reducing the statistical power, we found no DEGs at FDR 0.05 in the Interaction model. However, when we separated ecotypes to test for the effect of heat stress, we found that the SA ecotype had twice as many DEG in response to heat stress (340 DEG) compared to the FA ecotype (136 DEG). This separated analysis of heat treatment showed only 50 shared DEG between ecotypes (Figure 4.2).

4.3.2 Gene Set Enrichment Analysis

By using the information from all the genes in a pathway, gene set enrichment analysis provides more power to detect biological relevant molecular pathways that are divergent between groups relative to conducting individual gene tests. The GSEA identified 14 KEGG Pathways that were significantly enriched in high-ranking genes in response to heat treatment, 12 of which were upregulated (Figure 4.3). Of our focal pathways, the Insulin Signaling Network (of the Metabolic Function hallmark) was upregulated in a highly connected network including EERB signaling, neurotrophin signaling, and T-cell signaling pathways. Interestingly, two pathways related to the Macromolecule Damage and Repair hallmark (Nucleotide Excision Repair and DNA Replication) were downregulated in response to heat (Figure 4.3, Figure 4.S4).

The between-ecotype comparison revealed six significantly enriched pathways (Figure. 3), including networks related to the Macromolecule Damage and Repair hallmark (Base Excision Repair) and the Metabolic Function hallmark (Oxidative Phosphorylation). Despite not having the power to detect individual DEGs with a significant interaction between ecotype and response to heat stress, the GSEA interaction model found many significantly enriched pathways including two related to the Metabolic Function hallmark (Insulin Signaling [again in a tight network of overlapping pathways] and Oxidative Phosphorylation) and one related to the Macromolecule Damage and Repair hallmark (DNA Replication) that responded differently to heat between the ecotypes (Figures 4.3 and 4.S4). We separated datasets by ecotype to better understand the expression patterns of these interactions. By contrasting networks of enriched pathways between the SA ecotype and the FA ecotype responses to heat stress, we see a highly connected network of 20 pathways (including Insulin Signaling) upregulated in response to heat

in the SA ecotype, compared to a few minimally connected pathways in the FA ecotype (Figures 4.3 and 4.S4); this suggests a more highly coordinated cellular signaling response in the SA ecotype relative to the FA ecotype. Examining the significant interaction with Oxidative Phosphorylation, in response to heat stress SA is downregulating Oxidative Phosphorylation whereas FA is upregulating this pathway (when viewing pathways at FDR of 0.17 instead of 0.1) (Figures 4.3 and 4.S4). Interestingly, both ecotypes are downregulating pathways for DNA repair in response to heat, although SA shows a stronger response with more DNA repair-related pathways reaching significance.

4.3.3 SNP Analysis

Of the 397 captured genes, 351 genes contain at least one SNP, 333 contain at least one SNP within an exon, 304 contain at least one SNP within the coding region, and 211 contain at least one nonsynonymous SNP. Our final SNP dataset contained 12,290 SNPs, of which 3,868 were within exons, 2,664 were within CDS, and 790 were nonsynonymous SNPs. Information on all genes is contained within Table 4.S1. For the remainder of this manuscript, we will be focusing on 297 focal genes from the aging hallmark categories: Macromolecule Damage Repair (83), Stress Adaptation (44), and Metabolic Function (170).

4.3.4 Population Genetics for Focal Genes

Average pairwise F_{ST} and D_{XY} between FA and SA ecotype populations (ignoring population assignment) for exons of focal genes was 0.0169 and 0.00011, respectively (Figure 4.S9, center) demonstrating how closely related the populations are across this landscape. The frequency distributions for pairwise F_{ST} and D_{XY} are right-skewed, with some genes (albeit few) having F_{ST}

values of 0.05 to 0.24 and D_{XY} values of 0.00025 to 0.000876, suggesting resisted gene flow between the ecotypes at these genes (or genomic locations). The shapes of these distributions were also reflected in average pairwise differences for each of the respective populations (Figure 4.S9, bottom). Average Tajima's D for focal genes is 0.85, indicating no obvious evidence of selection across the dataset as a whole (Figure 4.S9, top right). However, the right-skew in the wide-ranging distribution (-1.22 to 6.16) may be due to some genes experiencing selection, with those genes >2 potentially diverging between the ecotypes.

Effect sizes for “within” and “between” ecotype interpopulation comparisons are shown in Figure 4.4 for both per-gene and per-ecotype analyses. The number, categorical distribution, and gene names of statistically significant findings from analyses of the within-ecotype vs between-ecotype F_{ST} and D_{XY} comparisons for both exon regions and nonsynonymous sites are shown in Figure 4.5B. The distribution of F_{ST} estimates from between-ecotype comparisons were significantly different than within-ecotype comparisons for exon regions in 33 of the 397 genes. The distribution of D_{XY} estimates from between-ecotype were significantly different than within-ecotype interpopulation comparisons for exon regions of 13 genes. Of the 33 and 13 genes significant for F_{ST} and D_{XY} (respectively), 12 overlapped. For nonsynonymous SNPs, per-site F_{ST} estimates for 35 SNPs from 26 genes were significantly higher in the between-ecotype than within-ecotype comparisons. Per-site D_{XY} estimates for 12 SNPs from 11 genes were significantly different for between-ecotype as compared to within-ecotype interpopulation comparisons. Of the 26 and 11 genes containing sites significant for F_{ST} and D_{XY} , 15 sites overlapped from eight genes. It is worth noting that for every case where F_{ST} or D_{XY} was shown to be significantly different, the higher value was always for the between-ecotype comparison

(Figure 4.4). Four genes contained significant F_{ST} and D_{XY} values for both exon regions and nonsynonymous sites (HIF1A, NR4A1, PYROXD2, and XRCC3). Details of the genes/sites are contained within Tables 4.S1, 4.S2, and 4.S3.

Across most of the SNPs within the focal genes, the alternate allele frequency was similar for both the FA and SA ecotypes (Figure 4.5C; Table 4.S2). However, the frequency distribution of the difference between allele frequencies (ΔAF) is right-skewed, showing that some alleles are more prevalent in populations of one ecotype compared to the other. Most of the nonsynonymous SNPs with significant differences in F_{ST} or D_{XY} values are contained within the right tail of the distribution (Figure 4.5C).

Within the nonsynonymous site data binned into finer-scale groups of each hallmark category (Macromolecule Damage and Repair [DNA Repair, Oxidative Stress, p53], Metabolic Function [Oxidative Phosphorylation, Nutrient Sensing with Insulin Signaling, Metabolism], and Stress Adaptation [Heat, Hypoxia, General Stress]), we found no difference in average F_{ST} between these categories, between population comparison type (within-ecotype populations vs between-ecotype populations), or an interaction between these. However, we did find a significant difference between the population comparison type for D_{XY} of the nonsynonymous sites. Within-ecotype comparisons had D_{XY} values 9.04×10^{-3} ($\pm 7.89 \times 10^{-3}$) less than those between-ecotypes ($p = 0.021$). However, when examined visually it is apparent that this difference is driven primarily by the Heat and General Stress groups, with only slight differences in D_{XY} values (Figure 4.S10). The model examining D_{XY} showed no difference between hallmark categories or in the interaction between population comparison type and the categories.

4.3.5 Predicting Functional Effects of SNPs

Annotation from SIFT4G for the nonsynonymous SNPs significant for F_{ST} or D_{XY} are included in the Table 4.S3. Only one SNP (gene *NDUFV1* pos. 69298519) was reported to have a low confidence SIFT score due to insufficient coverage in the multiple sequence alignment constructed by SIFT4G. Of the remaining 37 nonsynonymous SNPs, three were labeled as “deleterious” with SIFT scores below 0.05 (*PIK3C2A* pos. 83291589, *NFKBIA* pos. 128350420, and *CYP4F22* pos. 876503), suggesting these SNPs likely have a strong impact on the function of the protein. SIFT scores for these SNPs are plotted in Figure 4.S8.

We point out that the “deleterious” designation is a prediction of functional impact based on evolutionary constraint that does not account for evolutionary relatedness (i.e., prevalence of the allele in a multiple sequence alignment across taxa, all taxa are treated as equally related), thus permitting biased assignments depending on taxonomic representation within the alignment (which are heavily mammal-biased). It is possible that variants marked as “deleterious” have functional implications that aren’t necessarily harmful, but likely have an impact on how that protein functions. Therefore, we recommend that scores deviating from baseline be examined for functional impact using additional approaches, and in the discussion we comment on potential effects that some of these mutations may have on phenotypic divergence between ecotypes.

4.5 Discussion

In this study we set out to determine whether the hallmarks of aging identified in laboratory model species can translate to natural populations where aging phenotypes are evolving. We provide evidence that supports our hypothesis that life history differences between ecotypes are

underpinned by differences in transcriptional regulation and sequence variation (including nonsynonymous sequence variation) in gene networks underlying the hallmarks of aging (Figure 4.6).

We used a whole-transcriptome approach to understand the hepatic response to heat stress and test for divergence between the ecotypes in this response. Strikingly, we found that $< 1/6$ of the transcriptomic response to heat stress was shared by the ecotypes, which are closely related populations of the same species (Fig 1, 3, S5). Further, we found the SA ecotype, relative to the FA ecotype, had a more robust response in terms of the number of genes differentially regulated, and a more coordinated response with a larger and more tightly connected network of molecular pathways responding to heat stress. Taken together these results suggests that the ecotypes are responding to acute heat stress using largely different molecular mechanisms. The more robust and coordinated transcriptomics response to an environmental stressor in the slow-aging ecotype may be a purview into the coevolutionary process for stress-resistance and senescence-resistance.

Within the network of pathways enriched for ecotype-specific responses to heat stress we found pathways that fit within the hallmarks of aging that we targeted for assessing genetic variation (Figure 4.3). Among our targeted set of genes evaluated for sequence variation, 47 showed divergence between the ecotypes. The differences in gene regulation, particularly those in response to stress, appear more extensive than those of genetic sequence (see “nutrient sensing” and “oxidative phosphorylation” boxes of Figure 4.6). This is expected given that regulatory differences often drive phenotypic divergence over genetic sequence dissimilarity (King and Wilson, 1975; Whitehead and Crawford, 2006; Abolins-Abols et al., 2018), and prominent

differences in gene expression may be due to sequence variation in cis-regulatory regions (see Jin et al., 2001; Brem et al., 2002; Cheung et al., 2003; Stamatoyannopoulos, 2004; Gibson and Weir, 2005). In the integration of the gene expression and sequence variation results in the context of the hallmarks of aging pathways we see many examples of proteins that are divergent between the aging ecotypes in their gene regulation that are interacting with proteins divergent in their sequence variation (Figure 4.6), illustrating the cooperative nature of selection acting on expression regulation and protein structure/function to manipulate molecular networks.

To our knowledge, this is the first study to examine the potential genomic (at both sequence and expression levels) underpinnings of aging in a natural, conspecific vertebrate system using targeted gene networks and a controlled experimental design. Below we discuss the relevance of the patterns we observed in genes and gene networks with significant expression and/or sequence divergence within their respective hallmark categories. We illuminate the contributions of genetic variation in several molecular networks to these ecotypes with divergent aging phenotypes by integrating results from gene expression and nucleotide sequence variation with previously published life-history traits from the Eagle Lake terrestrial garter snake study system. Given that many physiological and molecular pathways associated with aging are shared across vertebrate lineages (Hoekstra et al., 2020), we discuss the extent to which ecotypes within a species differ in expression and allelic variation for genes of these pathways and highlight findings that are consistent with patterns of aging seen across vertebrates.

4.5.1 Macromolecule Damage and Repair

Unrepaired DNA damage is linked to rate of aging and overall longevity. It has been hypothesized that molecular pathways for DNA repair are responsible for aging due to deleterious somatic mutation accumulation in the blueprint for essential cellular machinery (Gladyshev, 2013). Previous findings in the SA and FA ecotypes show divergence in DNA repair phenotypes, with the FA ecotype showing less-efficient DNA repair (Bronikowski, 2008) and a decreased capacity to repair DNA (Robert and Bronikowski, 2010; Schwartz and Bronikowski, 2013) compared to the SA ecotype. Our pathway enrichment analyses indicated that the Base Excision Repair pathway is upregulated in SA compared to FA, a pattern which is consistent with observations of longer-living organisms exhibiting upregulation in DNA-repair genes (MacRae et al., 2015). The interaction between ecotype and heat stress response indicates that the DNA replication pathway is more downregulated in SA in response to acute heat stress relative to the FA response. Focusing on the response of each ecotype to acute heat stress separately shows that four pathways involved with DNA repair are downregulated by SA compared to only one by FA (Figure 4.3). Overall this interaction pattern of gene expression at the pathway level is consistent with Schwartz and Bronikowski (2013, Fig. 5) that showed SA has lower background DNA damage but more DNA damage in response to acute heat stress, and then faster, more efficient repair.

In addition to regulatory modifications influencing aging, there is evidence for an association between sequence variation in DNA repair genes and longevity across species (Tollis et al., 2019). Eleven genes in this study that are involved with macromolecule damage and repair have significant divergence in exon regions and /or nonsynonymous SNPs between ecotypes (XRCC3, RAD54B, CCS, GADD45B, GDAP1, GPX3, MYC, PRDX6, PXDN, and PYROXD2, MPO).

Because the decreased ability of somatic cells to properly respond to damaging oxidative agents such as ROS is an integral part of aging, and alterations to antioxidant expression can affect aging (Landis and Tower, 2005), the significance of these genes (many of which are involved in oxidative stress response) matched our predictions.

Our pathway enrichment analysis showed that two of these genes, XRCC3 and RAD54B, were on the leading-edges of the Homologous Recombination DNA repair pathways for the heat treatment in the SA ecotype. The gene XRCC3, a RAD51 homolog, is required to recruit RAD51 to breakage sites for double-strand DNA repair. In addition to XRCC3 having significant differences between ecotypes in absolute and relative divergence for SNPs in the exon regions (16 total SNPs), we also found that one of the two non-synonymous SNPs (Q129R, based on human amino acid number) showed absolute and relative divergence significantly different between ecotypes. Additionally, we found a high Tajima's D value for XRCC3 (2.05), indicating the possibility of balancing selection occurring within this gene. Similarly, the gene RAD54B, a subunit of the enzyme RAD54, is involved in double-strand DNA repair. In our study, this gene is on the expression leading edge of the homologous recombination pathway for the response to heat, and it contains four significant nonsynonymous SNPs; one is within the superfamily II DNA/RNA helicase and the other three are within the DEXH-box helicase.

Variation in genes associated with double-strand break (DSB) repair is associated with longevity in humans (Debrabant et al., 2014) and the long-living American lobster (Polinski et al., 2021). Further, differences in lifespan between queen and workers in eusocial insects (specifically, ants and termites) are associated with higher expression of DNA repair genes (including XRCC3 and RAD54B; Lucas et al., 2016; Tasaki et al., 2018). Overexpression of RAD54 also resulted in

longer median lifespans in *Drosophila* lab strains (Shaposhnikov et al., 2015). While sequence variation in XRCC3 has not been directly associated with aging, paralogs and factors within the homologous recombination repair regime cause progeroid phenotypes in mammals (Lombard et al., 2005). It has also been proposed that the well-established relationship between Sir2 function and senescence involves the DSB repair function of Sir2 (Hasty, 2001).

Silencing of the copper chaperone CCS, which has been proposed as a candidate gene for longevity in pigs (Metodiev et al., 2006), is shown to cause significant decreases in viability and performance (Theotoki et al., 2019). An association between the expression and activity of superoxide dismutase (the copper recipient of CCS) and aging has been observed in *Drosophila* (Dudas and Arking, 1995; Spencer et al., 2003; Tower, 2015), yeast (Fabrizio et al., 2003), nematodes (Larsen, 1993; Yen et al., 2009), lab mice (Levin et al., 2005), peas (del Río et al., 2003), and bean weevils (Šešlija et al., 1999). Increased expression of PRDX6, which in our study contains a significant nonsynonymous SNP with potential functional impact (SIFT – 0.19), is associated with increased aging in mice (Pacifici et al., 2020; Soriano-Arroquia et al., 2021), rats (Lubec et al., 2019) and human cells (Chhunchha et al., 2017, 2020, 2022; Salovska et al., 2022). Reduced expression of GPX3 has also been associated with increased aging in mice (Qi et al., 2018) and humans (Pastori et al., 2016). Involvement of MYC, a transcription factor that is estimated to regulate up to 15 percent of all genes in the genome (Dang et al., 2006), with maintaining cellular redox balance has been proposed (Benassi et al., 2006). MYC expression is associated with that of PRDX6 in human cells (Li et al., 2021). Expression of the antioxidant MPO, which in our study contains a significant nonsynonymous SNP that may affect function (SIFT = 0.08), is associated with aging in lab rats (Gen Son et al., 2005; Liu et al., 2015), lab

mice (Shen et al., 2018; Marquez-Exposito et al., 2022), humans (Vianello et al., 2016; Lee et al., 2021), and human cells (Liu et al., 2015; Lee et al., 2022).

Decreased expression of GADD45B, a transcriptional target of NF κ B involved with cellular stress response to regulate proliferation and apoptosis (Papa et al., 2004, 2007), is associated with premature aging in mouse cells (Magimaidas et al., 2016), and deletion of GADD45B is associated with decreased long-term memory (Leach et al., 2012; Sultan et al., 2012) and reduced neurogenesis (Ma et al., 2009) in mice. In our study, we observed a significant nonsynonymous SNP that likely influences function (SIFT = 0.05). Sequence variation in the NF κ B pathway and related pathways are associated with long lifespans in mammals (Kowalczyk et al., 2020).

4.5.2 Metabolic Function

Nutrient Sensing. Pathways involved with nutrient sensing and energy conversion have been implicated with the life-history trade-off hypothesis, which suggests that investment in growth/reproduction is inversely linked with investment in maintenance/survival. The nutrient sensing pathway IIS/Tor is involved in variation in longevity across animals (Kenyon et al., 1993; Fontana et al., 2010; Kenyon, 2010; Sparkman et al., 2012; Allison et al., 2014; McGaugh et al., 2015; Sanders et al., 2018; reviewed in Hoekstra et al., 2020), and although the pathways are highly conserved, variation indicative of positive selection has been found within squamate reptiles (McGaugh et al., 2015). In this study we found the SA and FA ecotypes differentially regulate this pathway (Figure 4.3), and some of the top regulators of this pathway evidence genetic divergence between ecotypes of the pathway, including IGF2, INSR, and IRS. This bias in sequence divergence in top regulator genes reflects predictions from the hypothesis that top

regulators of the IIS network and critical intracellular genes experience positive selection while downstream genes experience purifying selection (McGaugh et al., 2015).

We found evidence for significant divergence in both relative nucleotide divergence and gene expression for IGF1R, a main IGF transport protein in the bloodstream that is associated with human longevity and aging (He et al., 2014; Teumer et al., 2016), and suppression of insulin-like growth factors is associated with increased longevity (Kenyon, 2005; Tullet et al., 2008). A nonsynonymous SNP in the IGF2 E-peptide shows significant relative divergence between ecotypes, although it is unclear how this affects IIS signaling since (in mice) the E-peptide is cleaved prior to IGF2 secretion for endocrine function (Buchanan et al., 2001). INSR, a gene with a nonsynonymous SNP with significant relative divergence between ecotypes, has been identified as a longevity-associated gene in long-living mammals (Yu et al., 2021), and sequence variation in INSR is associated with longevity in human centenarians (Kojima et al., 2004). Reduced activity of PI3K and its upstream regulator IRS have been associated with increases in lifespan in *Drosophila*, nematodes, and mice (Clancy et al., 2001; Ayyadevara et al., 2008; Foukas et al., 2013). We found lower SA expression in both of these genes in response to heat and evidence for significant relative divergence between ecotypes in IRS and three nonsynonymous SNPs of PIK3C2A (an enzyme within the PI3K family), demonstrating consistency in the role of these genes in longevity across lab model organisms and our natural populations of snakes.

The PI3K pathway, which plays an important role in regulating the cell cycle, is activated by the same upstream components as the IIS network. Genetic variation in PI3K is associated with

longevity in nematodes and *Drosophila* (Ayyadevara et al., 2008; Moskalev and Shaposhnikov, 2008). PI3K directly converts PIP₂ to PIP₃, which is an upstream activator of both AKT and mTOR. INPP5A catalyzes the reversal of PIP₃ to PIP₂, thus inhibiting activation of AKT and mTOR. We found a nonsynonymous SNP within INPP5A that is significantly differentiated between the ecotypes and possesses a predicted functional impact (SIFT score of 0.05).

The forkhead box (FOX) proteins are downstream transcription factors in the IIS gene network that regulate genes associated with life history and are associated with variation in longevity (Passtoors et al., 2013; Stefanetti et al., 2018; Li et al., 2019). Variation in FOXO3, a gene with significant divergence between ecotypes, is linked to longevity in several metazoan lineages. FOXO3 is one of the few genes with multiple studies showing genetic variation associated with increased longevity in humans (Willcox et al., 2008; Anselmi et al., 2009; Flachsbarth et al., 2009; Pawlikowska et al., 2009; Soerensen et al., 2010; Bao et al., 2014; Sun et al., 2015; Teumer et al., 2016). Upstream of FOXO3 is FOXA3, a gene with significant relative divergence between ecotypes that also contains a nonsynonymous SNP with significant relative divergence between ecotypes. Located within the forkhead N-terminal region, this SNP has a low SIFT score (0.06), which we interpret as high probability of impacting protein function.

While the effects of variation in the PI3K-mediated IIS pathway on aging are better understood, evidence for a link between the Ras/MAPK-mediated IIS pathway and variation in aging also exist (Slack et al., 2015). We found evidence for SA downregulation in response to heat (SHC, GRB2, SOS, Ras, Raf, and ELK) and genetic divergence (Ras, MAPK3, RIT1) within the Ras/MAPK-mediated IIS pathway. MAP3K1, a gene downstream of Ras with significant relative

divergence between ecotypes, contains a nonsynonymous SNP with significant relative and absolute divergence between ecotypes. However, the high SIFT score (1.0) suggests no indication of functional impact from this SNP. MAP3K1 activates NR4A1 expression downstream, a gene strongly associated with aging in nematodes (Akhoon et al., 2019) that in our study shows significant relative and absolute divergence between ecotypes and contains a nonsynonymous SNP with significant relative and absolute divergence that may affect protein function (SIFT 0.09).

We also observed a nonsynonymous SNP with significant absolute divergence between the ecotypes within the gene TSC1, a gene that inhibits cell growth and regulates the mTOR signaling pathway. This SNP has a low SIFT score (0.15), suggesting the possibility of a functional impact. Knocking out TSC1 is associated with accelerated aging in lab mice (Deng et al., 2021; Rao et al., 2021) and differences in expression are associated with longevity in *Drosophila* (Li et al., 2019), lab mice (Zhang et al., 2017), and humans (Passtoors et al., 2013).

Mitochondrial Function. Oxidative phosphorylation, the primary metabolic pathway for energy production in animals, is widely recognized as a pathway whose dysfunction is directly associated with senescence (López-Otín et al., 2013; Kennedy et al., 2014). Our pathway enrichment analyses show differences in OXPHOS as a network between ecotypes (upregulated in SA) and in the heat x ecotype interaction (Figure 4.2), and we found five nonsynonymous sites were significantly divergent between ecotypes (Figure 4.5C).

Evidence for divergence in OXPHOS expression between ecotypes is harmonious with previously reported differences in mitochondrial function between the two ecotypes showing physiological divergence in mitochondrial efficiency, ROS production, antioxidant expression, organismal metabolic rate, and cellular oxygen consumption rate (Bronikowski and Vleck, 2010; Robert and Bronikowski, 2010; Schwartz and Bronikowski, 2013; Gangloff et al., 2015, 2020), and this also ties in with the ecotypic segregation of a nonsynonymous mitochondrial-encoded SNP in Cytochrome-C Reductase (Complex III, ; Schwartz et al., 2015). Interestingly, at the expression level the directionality is reversed for nuclear-encoded mitochondrial-targeted genes, which show higher expression in SA compared to FA, while mitochondrial-encoded genes show higher expression in FA compared to SA (Schwartz et al., 2015). However, we do not know whether the difference in mitochondrial transcript abundance is due to more mitochondrial genomes or increased transcription.

Genes with significant relative and/or absolute divergence between ecotypes in either exon regions or nonsynonymous SNPs include members of the electron transport chain; the location of the affected subunits and amino acids within the complexes are shown in Figure 4.S11 We see no evidence for significant genetic divergence between ecotypes for any of the nuclear-encoded genes in CIII corresponding to the ecotype-segregating SNP. In addition to the genetic divergence we observed between ecotypes, it is possible that genetic variation in upstream regulators or differences in mitochondrial behavior/morphology may be responsible for the significant differences in enrichment that we observed between ecotypes and their gene expression response to heat.

4.5.3 Stress Adaptation

The effectiveness of organismal and cellular responses to external stressors, which occur via pathways that are largely conserved across animals, generally declines with age. Examples of stressors and the intracellular groups involved in stress responses include (1) hypoxia and hypoxia-inducible factors (HIFs), and (2) heat stress and heat shock proteins (HSPs).

Hypoxia. The association of longevity genes and those involved with the hypoxic response is documented in humans (Passtoors et al., 2013) and nematode worms (Mabon et al., 2009). Two primary responders to hypoxia are VHL and its target protein HIF1-A, whose inactivation results in an increase in longevity (Mehta et al., 2009; Müller et al., 2009). HIF1A and VHL were significantly divergent between the ecotypes across the exon regions, and HIF1A has a significant nonsynonymous SNP that may affect function (SIFT = 0.35). RORC, whose expression is regulated by HIF1A, also has a divergent nonsynonymous SNP that may affect function (SIFT = 0.23). Activity of the NF- κ B transcription factor, which responds to a variety of cellular stressors including hypoxia, is modulated by the IKK complex and nucleoporin 88 (Takahashi et al., 2008). Sequence variation in NFKBIA, which encodes a subunit in the IKK complex, is associated with longevity in humans. In NFKBIA, we observed significant absolute and relative divergence between ecotypes in exon regions as a whole and a nonsynonymous SNP within the ANK 1 region that is predicted to affect protein function (SIFT = 0.04) with relative divergence significantly different between ecotypes. We also observed a nonsynonymous SNP in nucleoporin 88 (NUP88) with relative divergence significantly different between ecotypes that may affect function (SIFT = 0.22). NFKB1A and VHL, both of which show significant absolute and relative divergence between ecotypes, are associated with aging and longevity in nematodes

(HIF-1 [Leiser and Kaeberlein, 2010]; VHL [Müller et al., 2009]), and humans (Ryu et al., 2021). A nonsynonymous SNP in NFKB1A with significant relative divergence is predicted to affect protein function (SIFT = 0.04).

Heat Stress. The efficiency of a cell to respond to misfolded proteins is linked to cellular health, and this efficiency decreases with aging (Tower, 2009). This understanding, based on the differential gene expression of heat shock proteins (HSPs), is observed in *Drosophila* (King and Tower, 1999; Yang and Tower, 2009), humans (Fonager et al., 2002), lab mice (Jurivich et al., 2020), nematodes (Manière et al., 2014), and Rhesus macaques (Schultz et al., 2001). Heat treatment showed significant differential gene expression compared to control in six out of twelve targeted heat shock proteins. We found significant sequence divergence between ecotypes for several HSP genes, which is not a widely reported observation in studies on aging. This distinction may be due to environmental conditions differing between wild populations and lab strains. FA and SA ecotypes inhabit unique environments with unique selection pressures, including differences in climate. For instance, the ephemeral ponds of the meadows (where SA snakes reside) are more prone to shifts in temperature than the stable lakeshore (where FA snakes reside).

4.5.4 Implications for Life History Evolution:

While multiple studies have examined the genetic underpinnings of life-history variation between species (e.g., Fushan et al., 2015; McGaugh et al., 2015; Passow et al., 2019) and within lab model organisms (e.g., Flatt, 2020; Evans et al., 2021; Bou Sleiman et al., 2022), less is known about the genetic basis for life-history variation in conspecific natural populations of non-

model species. In this study we identify nucleotide sequence and genetic regulatory differences between phenotypically divergent, conspecific ecotypes in gene networks that have been associated with aging variation on a deeper evolutionary scale and within experimental lab settings. The converse environmental selective pressures experienced by the two ecotypes (including predation, food availability, temperature, water permanence) have been implicated as sources of life history phenotypic divergence in the Eagle Lake garter snake system, and they may also be the drivers of genetic divergence and regulatory differences between the ecotypes. Numerous lines of evidence support the hypothesis that caloric restriction increases longevity via decreased activity of the IIS pathway, and the fluctuating food availability in SA habitats (Bronikowski and Arnold, 1999) may be responsible for the genetic divergence in IGF2 and FOXO3 along with the difference in SA IIS regulatory response to heat. Enhanced risk of predation (or other sources of extrinsic mortality) is associated with decreased longevity and senescence (Magnhagen, 1990; Austad, 1993; Dudycha, 2001; Bryant and Reznick, 2004; Bronikowski and Promislow, 2005; Carlson et al., 2007; Costanzo et al., 2011; Chandrasegaran et al., 2018) due to inefficient selection on old-individual phenotypes and/or deleterious mutation accumulation, and genetic ecotypic divergence and expression differences in DNA repair genes XRCC3 and RAD54B may be due in-part to divergence in extrinsic mortality in the FA and SA habitats (Sparkman and Bronikowski, 2013). Although low levels of overall genetic differentiation in this and other studies suggest that gene flow occurs between populations of these ecotypes, strong selection could result in restriction to gene flow in genes and genomic regions that maintain ecotype life-history divergence.

4.5.5 Value of Studying Aging Within and Across Species

Understanding general processes of aging and their underlying genetic mechanisms requires a broad understanding of aging variation. Knowledge on commonalities across systems can provide insights into life history theory, evolution of aging hypotheses, and inform decisions in medical research. While research in model organisms has pioneered aging research, restricting efforts to a handful of taxa limits knowledge capacity to a few tips of the tree of life. Expanding the scope to include more diverse taxa has provided previously hidden perspectives into phenotypic traits associated with longevity (Reinke et al., 2022), sequence diversity of aging genes (Opazo et al., 2022), and unique expression patterns of aging genes (Beatty et al., 2022). These two approaches (model species vs diverse taxa) provide finite and expanded perspectives, yet they both lack a critical vantage point found within conspecific populations with naturally occurring variation in aging where evolution in real-time, rather than the end result, can be examined. By working with a populations such as *Thamnophis elegans* around Eagle Lake to quantify aging-related traits, document (or control) environmental conditions, and capture the genomic diversity, connections can be made between these three measures, thus shedding light on the source of aging phenotypes in a natural population.

4.6 Conclusion

Using natural population of garter snakes that are divergent in their physiology, life history, and lifespan, we have documented divergence in regulation and nucleotide sequence of candidate gene networks underpinning three hallmarks of aging: metabolic processes, macromolecule damage and repair, and stress adaptation. In doing so, we have demonstrated these candidate molecular networks regulating aging translate from lab models to natural populations. In addition, the interaction between ecotypes in response to heat indicate many more pathways are

diverging in these garter snakes than the ones highlighted in this study, suggesting natural populations and non-model organisms likely have mechanisms for diversification in aging processes beyond those currently identified in lab models.

4.7 Reproducibility

Raw sequencing reads are available on GenBank (SRA052923, SRA062606) and upon request.

Additional sequence files (BAM, FASTA, and VCF) are available upon request. Code is available on github:

https://github.com/rklabacka/ThamnophisElegans_FunctionalGenomics2021).

4.8 Funding

Grants: This research was supported by the National Science foundation [DEB-DDIG 1011350 to TSS & AMB, IOS-1558071 to AMB], GRFP (SEM); the Iowa Academy of Sciences [ISF 09-21], and the Iowa State University (ISU) Center for Integrated Animal Genomics. TSS was supported by Graduate Fellowships from the NSF during data collection for this project (IGERT in Computational Biology 0504304, and GK-12 DGE-0947929) and acknowledges additional support from the James S. McDonnell Foundation, and Auburn University. We are grateful for resources from the University of Minnesota Supercomputing Institute and Auburn University's High Performance Computing. This work was made possible in part by a grant of high performance computing resources and technical support from the Alabama Supercomputer Authority.

4.9 Acknowledgments

We would like to acknowledge and thank K. Roberts and M. Manes for the collection and care of animals; M. Manes and A. Sparkman for assistance on experiment days; M. Baker and the ISU DNA Sequencing Facility and O. Fedrigo and the Duke Genome Sequencing and Analysis Core; A. Severin for assistance with initial sequence data analyses. Thank you to Samuel Weaver for manuscript comments and other members of the Schwartz lab for general feedback on the figures.

4.10 References

- Abolins-Abols, M., Kornobis, E., Ribeca, P., Wakamatsu, K., Peterson, M. P., Ketterson, E. D., et al. (2018). Differential gene regulation underlies variation in melanic plumage coloration in the dark-eyed junco (*Junco hyemalis*). *Molecular Ecology* 27, 4501–4515. doi: 10.1111/mec.14878.
- Addis, E. A., Gangloff, E. J., Palacios, M. G., Carr, K. E., and Bronikowski, A. M. (2017). Merging the “Morphology–Performance–Fitness” Paradigm and Life-History Theory in the Eagle Lake Garter Snake Research Project. *Integrative and Comparative Biology* 57, 423–435. doi: 10.1093/icb/ix079.
- Akhood, B. A., Gupta, S. K., Tiwari, S., Rathor, L., Pant, A., Singh, N., et al. (2019). *C. elegans* protein interaction network analysis probes RNAi validated pro-longevity effect of *nhr-6*, a human homolog of tumor suppressor *Nr4a1*. *Sci Rep* 9, 15711. doi: 10.1038/s41598-019-51649-0.
- Allison, D. B., Antoine, L. H., Ballinger, S. W., Bamman, M. M., Biga, P., Darley-Usmar, V. M., et al. (2014). Aging and energetics’ ‘Top 40’ future research opportunities 2010–2013. *F1000Res* 3, 219. doi: 10.12688/f1000research.5212.1.
- Anselmi, C. V., Malovini, A., Roncarati, R., Novelli, V., Villa, F., Condorelli, G., et al. (2009). Association of the FOXO3A Locus with Extreme Longevity in a Southern Italian Centenarian Study. *Rejuvenation Research* 12, 95–104. doi: 10.1089/rej.2008.0827.
- Austad, S. N. (1993). Retarded senescence in an insular population of Virginia opossums (*Didelphis virginiana*). *Journal of Zoology* 229, 695–708. doi: 10.1111/j.1469-7998.1993.tb02665.x.

- Ayyadevara, S., Alla, R., Thaden, J. J., and Reis, R. J. S. (2008). Remarkable longevity and stress resistance of nematode PI3K-null mutants. *Aging Cell* 7, 13–22. doi: 10.1111/j.1474-9726.2007.00348.x.
- Bao, J.-M., Song, X.-L., Hong, Y.-Q., Zhu, H.-L., Li, C., Zhang, T., et al. (2014). Association between FOXO3A gene polymorphisms and human longevity: a meta-analysis. *Asian J Androl* 16, 446–452. doi: 10.4103/1008-682X.123673.
- Beatty, A., Rubin, A. M., Wada, H., Heidinger, B., Hood, W. R., and Schwartz, T. S. (2022). Postnatal expression of IGF2 is the norm in amniote vertebrates. *Proceedings of the Royal Society B: Biological Sciences* 289, 20212278. doi: 10.1098/rspb.2021.2278.
- Benassi, B., Fanciulli, M., Fiorentino, F., Porrello, A., Chiorino, G., Loda, M., et al. (2006). c-Myc Phosphorylation Is Required for Cellular Response to Oxidative Stress. *Molecular Cell* 21, 509–519. doi: 10.1016/j.molcel.2006.01.009.
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170.
- Bou Sleiman, M., Roy, S., Gao, A. W., Sadler, M. C., von Alvensleben, G. V. G., Li, H., et al. (2022). Sex- and age-dependent genetics of longevity in a heterogeneous mouse population. *Science* 377, eabo3191. doi: 10.1126/science.abo3191.
- Brem, R. B., Yvert, G., Clinton, R., and Kruglyak, L. (2002). Genetic Dissection of Transcriptional Regulation in Budding Yeast. *Science* 296, 752–755. doi: 10.1126/science.1069516.
- Broad Institute (2016). Picard: A set of command line tools (in Java) for manipulating high-throughput sequencing (HTS) data and formats such as SAM/BAM/CRAM and VCF.

- Bronikowski, A. M., Fedrigo, O., Fungtammasan, C., Rhie, A., Mountcastle, J., Haase, B., et al. *Thamnophis elegans* (Western terrestrial garter snake) genome, rThaEle1, primary haplotype. Accession No. GCA_009769535.1. Available at: https://www.ncbi.nlm.nih.gov/assembly/GCA_009769535.1 [Accessed January 7, 2021].
- Bronikowski, A. M. (2000). Experimental Evidence for the Adaptive Evolution of Growth Rate in the Garter Snake *Thamnophis Elegans*. *Evolution* 54, 1760–1767. doi: 10.1111/j.0014-3820.2000.tb00719.x.
- Bronikowski, A. M. (2008). The evolution of aging phenotypes in snakes: a review and synthesis with new data. *AGE* 30, 169. doi: 10.1007/s11357-008-9060-5.
- Bronikowski, A. M., and Arnold, S. J. (1999). The Evolutionary Ecology of Life History Variation in the Garter Snake *Thamnophis Elegans*. *Ecology* 80, 2314–2325. doi: 10.1890/0012-9658(1999)080[2314:TEEOLH]2.0.CO;2.
- Bronikowski, A. M., Meisel, R. P., Biga, P. R., Walters, J. R., Mank, J. E., Larschan, E., et al. (2022). Sex-specific aging in animals: Perspective and future directions. *Aging Cell* 21, e13542. doi: 10.1111/accel.13542.
- Bronikowski, A. M., and Promislow, D. E. L. (2005). Testing evolutionary theories of aging in wild populations. *Trends in Ecology & Evolution* 20, 271–273. doi: 10.1016/j.tree.2005.03.011.
- Bronikowski, A., and Vleck, D. (2010). Metabolism, Body Size and Life Span: A Case Study in Evolutionarily Divergent Populations of the Garter Snake (*Thamnophis elegans*). *Integrative and Comparative Biology* 50, 880–887. doi: 10.1093/icb/icq132.
- Bryant, M. J., and Reznick, D. (2004). Comparative Studies of Senescence in Natural Populations of Guppies. *The American Naturalist* 163, 55–68. doi: 10.1086/380650.

- Buchanan, C. M., Phillips, A. R. J., and Cooper, G. J. S. (2001). Preptin derived from proinsulin-like growth factor II (proIGF-II) is secreted from pancreatic islet β -cells and enhances insulin secretion. *Biochemical Journal* 360, 431–439. doi: 10.1042/bj3600431.
- Carlson, S. M., Hilborn, R., Hendry, A. P., and Quinn, T. P. (2007). Predation by Bears Drives Senescence in Natural Populations of Salmon. *PLOS ONE* 2, e1286. doi: 10.1371/journal.pone.0001286.
- Chandrasegaran, K., Kandregula, S. R., Quader, S., and Juliano, S. A. (2018). Context-dependent interactive effects of non-lethal predation on larvae impact adult longevity and body composition. *PLOS ONE* 13, e0192104. doi: 10.1371/journal.pone.0192104.
- Cheung, V. G., Conlin, L. K., Weber, T. M., Arcaro, M., Jen, K.-Y., Morley, M., et al. (2003). Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat Genet* 33, 422–425. doi: 10.1038/ng1094.
- Chhunchha, B., Kubo, E., and Singh, D. P. (2020). Clock Protein Bmal1 and Nrf2 Cooperatively Control Aging or Oxidative Response and Redox Homeostasis by Regulating Rhythmic Expression of Prdx6. *Cells* 9, 1861. doi: 10.3390/cells9081861.
- Chhunchha, B., Kubo, E., and Singh, D. P. (2022). Switching of Redox Signaling by Prdx6 Expression Decides Cellular Fate by Hormetic Phenomena Involving Nrf2 and Reactive Oxygen Species. *Cells* 11, 1266. doi: 10.3390/cells11081266.
- Chhunchha, B., Singh, P., Stamer, W. D., and Singh, D. P. (2017). Prdx6 retards senescence and restores trabecular meshwork cell health by regulating reactive oxygen species. *Cell Death Discov.* 3, 1–12. doi: 10.1038/cddiscovery.2017.60.

- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* 6, 80–92. doi: 10.4161/fly.19695.
- Clancy, D. J., Gems, D., Harshman, L. G., Oldham, S., Stocker, H., Hafen, E., et al. (2001). Extension of Life-Span by Loss of CHICO, a Drosophila Insulin Receptor Substrate Protein. *Science* 292, 104–106. doi: 10.1126/science.1057991.
- Costanzo, K. S., Muturi, E. J., and Alto, B. W. (2011). Trait-mediated effects of predation across life-history stages in container mosquitoes. *Ecological Entomology* 36, 605–615. doi: 10.1111/j.1365-2311.2011.01302.x.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–8. doi: 10.1093/bioinformatics/btr330.
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, giab008. doi: 10.1093/gigascience/giab008.
- Dang, C. V., O'Donnell, K. A., Zeller, K. I., Nguyen, T., Osthus, R. C., and Li, F. (2006). The c-Myc target gene network. *Seminars in Cancer Biology* 16, 253–264. doi: 10.1016/j.semcancer.2006.07.014.
- Debrabant, B., Soerensen, M., Flachsbar, F., Dato, S., Mengel-From, J., Stevnsner, T., et al. (2014). Human longevity and variation in DNA damage response and repair: study of the contribution of sub-processes using competitive gene-set analysis. *Eur J Hum Genet* 22, 1131–1136. doi: 10.1038/ejhg.2013.299.

- del Río, L. A., Sandalio, L. M., Altomare, D. A., and Zilinskas, B. A. (2003). Mitochondrial and peroxisomal manganese superoxide dismutase: differential expression during leaf senescence. *Journal of Experimental Botany* 54, 923–933. doi: 10.1093/jxb/erg091.
- Deng, Y., Yang, Q., Yang, Y., Li, Y., Peng, H., Wu, S., et al. (2021). Conditional knockout of Tsc1 in ROR γ t-expressing cells induces brain damage and early death in mice. *Journal of Neuroinflammation* 18, 107. doi: 10.1186/s12974-021-02153-8.
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43, 491–498. doi: 10.1038/ng.806.
- Dudas, S. P., and Arking, R. (1995). A Coordinate Upregulation of Antioxidant Gene Activities Is Associated With the Delayed Onset of Senescence in a Long-Lived Strain of *Drosophila*. *The Journals of Gerontology: Series A* 50A, B117–B127. doi: 10.1093/gerona/50A.3.B117.
- Dudycha (2001). The senescence of *Daphnia* from risky and safe habitats. *Ecology Letters* 4, 102–105. doi: 10.1046/j.1461-0248.2001.00211.x.
- Evans, K. S., van Wijk, M. H., McGrath, P. T., Andersen, E. C., and Sterken, M. G. (2021). From QTL to gene: *C. elegans* facilitates discoveries of the genetic mechanisms underlying natural variation. *Trends in Genetics* 37, 933–947. doi: 10.1016/j.tig.2021.06.005.
- Fabrizio, P., Liou, L.-L., Moy, V. N., Diaspro, A., Valentine, J. S., Gralla, E. B., et al. (2003). SOD2 Functions Downstream of Sch9 to Extend Longevity in Yeast. *Genetics* 163, 35–46. doi: 10.1093/genetics/163.1.35.

- Flachsbart, F., Caliebe, A., Kleindorp, R., Blanché, H., von Eller-Eberstein, H., Nikolaus, S., et al. (2009). Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proceedings of the National Academy of Sciences* 106, 2700–2705. doi: 10.1073/pnas.0809594106.
- Flatt, T. (2020). Life-History Evolution and the Genetics of Fitness Components in *Drosophila melanogaster*. *Genetics* 214, 3–48. doi: 10.1534/genetics.119.300160.
- Fonager, J., Beedholm, R., Clark, B. F. C., and Rattan, S. I. S. (2002). Mild stress-induced stimulation of heat-shock protein synthesis and improved functional ability of human fibroblasts undergoing aging in vitro. *Experimental Gerontology* 37, 1223–1228. doi: 10.1016/S0531-5565(02)00128-6.
- Fontana, L., Partridge, L., and Longo, V. D. (2010). Extending Healthy Life Span—From Yeast to Humans. *Science* 328, 321–326. doi: 10.1126/science.1172539.
- Foukas, L. C., Bilanges, B., Betti, L., Pearce, W., Ali, K., Sancho, S., et al. (2013). Long-term p110 α PI3K inactivation exerts a beneficial effect on metabolism. *EMBO Molecular Medicine* 5, 563–571. doi: 10.1002/emmm.201201953.
- Fushan, A. A., Turanov, A. A., Lee, S.-G., Kim, E. B., Lobanov, A. V., Yim, S. H., et al. (2015). Gene expression defines natural changes in mammalian lifespan. *Aging Cell* 14, 352–365. doi: 10.1111/accel.12283.
- Gangloff, E. J., Holden, K. G., Telemeco, R. S., Baumgard, L. H., and Bronikowski, A. M. (2016). Hormonal and metabolic responses to upper temperature extremes in divergent life-history ecotypes of a garter snake. *J Exp Biol* 219, 2944–2954. doi: 10.1242/jeb.143107.

- Gangloff, E. J., Schwartz, T. S., Klabacka, R., Huebschman, N., Liu, A.-Y., and Bronikowski, A. M. (2020). Mitochondria as central characters in a complex narrative: Linking genomics, energetics, pace-of-life, and aging in natural populations of garter snakes. *Experimental Gerontology* 137, 110967. doi: 10.1016/j.exger.2020.110967.
- Gangloff, E. J., Vleck, D., and Bronikowski, A. M. (2015). Developmental and Immediate Thermal Environments Shape Energetic Trade-Offs, Growth Efficiency, and Metabolic Rate in Divergent Life-History Ecotypes of the Garter Snake *Thamnophis elegans*. *Physiological and Biochemical Zoology* 88, 550–563. doi: 10.1086/682239.
- Gen Son, T., Zou, Y., Pal Yu, B., Lee, J., and Young Chung, H. (2005). Aging effect on myeloperoxidase in rat kidney and its modulation by calorie restriction. *Free Radical Research* 39, 283–289. doi: 10.1080/10715760500053461.
- Gibson, G., and Weir, B. (2005). The quantitative genetics of transcription. *Trends in Genetics* 21, 616–623. doi: 10.1016/j.tig.2005.08.010.
- Gladyshev, V. N. (2013). The origin of aging: imperfectness-driven non-random damage defines the aging process and control of lifespan. *Trends in Genetics* 29, 506–512. doi: 10.1016/j.tig.2013.05.004.
- Hasty, P. (2001). The impact energy metabolism and genome maintenance have on longevity and senescence: lessons from yeast to mammals. *Mechanisms of Ageing and Development* 122, 1651–1662. doi: 10.1016/S0047-6374(01)00294-9.
- He, Y.-H., Lu, X., Yang, L.-Q., Xu, L.-Y., and Kong, Q.-P. (2014). Association of the insulin-like growth factor binding protein 3 (IGFBP-3) polymorphism with longevity in Chinese nonagenarians and centenarians. *Aging (Albany NY)* 6, 944–951.

- Hoekstra, L. A., Schwartz, T. S., Sparkman, A. M., Miller, D. A. W., and Bronikowski, A. M. (2020). The untapped potential of reptile biodiversity for understanding how and why animals age. *Functional Ecology* 34, 38–54. doi: 10.1111/1365-2435.13450.
- Hudson, R. R., Boos, D. D., and Kaplan, N. L. (1992). A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution* 9, 138–151. doi: 10.1093/oxfordjournals.molbev.a040703.
- Jin, W., Riley, R. M., Wolfinger, R. D., White, K. P., Passador-Gurgel, G., and Gibson, G. (2001). The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nat Genet* 29, 389–395. doi: 10.1038/ng766.
- Judson, J. (2021). Selection and population demography shape evolution of two long-lived ectotherms. *Iowa State University Theses and Dissertations*.
- Jurivich, D. A., Manocha, G. D., Trivedi, R., Lizakowski, M., Rakoczy, S., and Brown-Borg, H. (2020). Multifactorial Attenuation of the Murine Heat Shock Response With Age. *The Journals of Gerontology: Series A* 75, 1846–1852. doi: 10.1093/gerona/glz204.
- Kendig, K. I., Baheti, S., Bockol, M. A., Drucker, T. M., Hart, S. N., Heldenbrand, J. R., et al. (2019). Sentieon DNaseq Variant Calling Workflow Demonstrates Strong Computational Performance and Accuracy. *Frontiers in Genetics* 10. Available at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00736> [Accessed October 3, 2022].
- Kennedy, B. K., Berger, S. L., Brunet, A., Campisi, J., Cuervo, A. M., Epel, E. S., et al. (2014). Geroscience: Linking Aging to Chronic Disease. *Cell* 159, 709–713. doi: 10.1016/j.cell.2014.10.039.

- Kenyon, C. (2005). The Plasticity of Aging: Insights from Long-Lived Mutants. *Cell* 120, 449–460. doi: 10.1016/j.cell.2005.02.002.
- Kenyon, C., Chang, J., Gensch, E., Rudner, A., and Tabtiang, R. (1993). A *C. elegans* mutant that lives twice as long as wild type. *Nature* 366, 461–464. doi: 10.1038/366461a0.
- Kenyon, C. J. (2010). The genetics of ageing. *Nature* 464, 504–512. doi: 10.1038/nature08980.
- King, M.-C., and Wilson, A. C. (1975). Evolution at Two Levels in Humans and Chimpanzees. *Science* 188, 107–116. doi: 10.1126/science.1090005.
- King, V., and Tower, J. (1999). Aging-Specific Expression of *Drosophila hsp22*. *Developmental Biology* 207, 107–118. doi: 10.1006/dbio.1998.9147.
- Kojima, T., Kamei, H., Aizu, T., Arai, Y., Takayama, M., Nakazawa, S., et al. (2004). Association analysis between longevity in the Japanese population and polymorphic variants of genes involved in insulin and insulin-like growth factor 1 signaling pathways. *Experimental Gerontology* 39, 1595–1598. doi: 10.1016/j.exger.2004.05.007.
- Kowalczyk, A., Partha, R., Clark, N. L., and Chikina, M. (2020). Pan-mammalian analysis of molecular constraints underlying extended lifespan. *eLife* 9, e51089. doi: 10.7554/eLife.51089.
- Landis, G. N., and Tower, J. (2005). Superoxide dismutase evolution and life span regulation. *Mechanisms of Ageing and Development* 126, 365–379. doi: 10.1016/j.mad.2004.08.012.
- Larsen, P. L. (1993). Aging and resistance to oxidative damage in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences* 90, 8905–8909. doi: 10.1073/pnas.90.19.8905.

- Law, C. W., Alhamdoosh, M., Su, S., Dong, X., Tian, L., Smyth, G. K., et al. (2018). RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res* 5, 1408. doi: 10.12688/f1000research.9005.3.
- Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014). voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 15, R29. doi: 10.1186/gb-2014-15-2-r29.
- Leach, P. T., Poplawski, S. G., Kenney, J. W., Hoffman, B., Liebermann, D. A., Abel, T., et al. (2012). Gadd45b knockout mice exhibit selective deficits in hippocampus-dependent long-term memory. *Learn. Mem.* 19, 319–324. doi: 10.1101/lm.024984.111.
- Lee, S. F., Hirpara, J. L., Qu, J., Yadav, S. K., Sachaphibulkij, K., and Pervaiz, S. (2022). Identification of a novel catalytic inhibitor of topoisomerase II alpha that engages distinct mechanisms in p53wt or p53-/- cells to trigger G2/M arrest and senescence. *Cancer Lett* 526, 284–303. doi: 10.1016/j.canlet.2021.11.025.
- Lee, T. X. Y., Wu, J., Jean, W.-H., Condello, G., Alkhatib, A., Hsieh, C.-C., et al. (2021). Reduced stem cell aging in exercised human skeletal muscle is enhanced by ginsenoside Rg1. *Aging (Albany NY)* 13, 16567–16576. doi: 10.18632/aging.203176.
- Leiser, S. F., and Kaeberlein, M. (2010). The hypoxia-inducible factor HIF-1 functions as both a positive and negative modulator of aging. 391, 1131–1137. doi: 10.1515/bc.2010.123.
- Levin, E. D., Christopher, N. C., and Crapo, J. D. (2005). Memory decline of aging reduced by extracellular superoxide dismutase overexpression. *Behav Genet* 35, 447–453. doi: 10.1007/s10519-004-1510-y.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760.

- Li, H., Zhang, D., Li, B., Zhen, H., Chen, W., and Men, Q. (2021). PRDX6 Overexpression Promotes Proliferation, Invasion, and Migration of A549 Cells in vitro and in vivo. *Cancer Manag Res* 13, 1245–1255. doi: 10.2147/CMAR.S284195.
- Li, J., Duan, D., Zhang, J., Zhou, Y., Qin, X., Du, G., et al. (2019). Bioinformatic prediction of critical genes and pathways involved in longevity in *Drosophila melanogaster*. *Mol Genet Genomics* 294, 1463–1475. doi: 10.1007/s00438-019-01589-1.
- Liu, W.-Q., Zhang, Y.-Z., Wu, Y., Zhang, J.-J., Li, T.-B., Jiang, T., et al. (2015). Myeloperoxidase-derived hypochlorous acid promotes ox-LDL-induced senescence of endothelial cells through a mechanism involving β -catenin signaling in hyperlipidemia. *Biochemical and Biophysical Research Communications* 467, 859–865. doi: 10.1016/j.bbrc.2015.10.053.
- Lombard, D. B., Chua, K. F., Mostoslavsky, R., Franco, S., Gostissa, M., and Alt, F. W. (2005). DNA Repair, Genome Stability, and Aging. *Cell* 120, 497–512. doi: 10.1016/j.cell.2005.01.028.
- López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., and Kroemer, G. (2013). The Hallmarks of Aging. *Cell* 153, 1194–1217. doi: 10.1016/j.cell.2013.05.039.
- Lubec, J., Smidak, R., Malikovic, J., Feyissa, D. D., Korz, V., Höger, H., et al. (2019). Dentate Gyrus Peroxiredoxin 6 Levels Discriminate Aged Unimpaired From Impaired Rats in a Spatial Memory Task. *Frontiers in Aging Neuroscience* 11. Available at: <https://www.frontiersin.org/article/10.3389/fnagi.2019.00198> [Accessed June 10, 2022].
- Lucas, E. R., Privman, E., and Keller, L. (2016). Higher expression of somatic repair genes in long-lived ant queens than workers. *Aging (Albany NY)* 8, 1940–1949. doi: 10.18632/aging.101027.

- Ma, D. K., Jang, M.-H., Guo, J. U., Kitabatake, Y., Chang, M., Pow-anpongkul, N., et al. (2009). Neuronal Activity–Induced Gadd45b Promotes Epigenetic DNA Demethylation and Adult Neurogenesis. *Science* 323, 1074–1077. doi: 10.1126/science.1166859.
- Mabon, M. E., Mao, X., Jiao, Y., Scott, B. A., and Crowder, C. M. (2009). Systematic Identification of Gene Activities Promoting Hypoxic Death. *Genetics* 181, 483–496. doi: 10.1534/genetics.108.097188.
- MacRae, S. L., Zhang, Q., Lemetre, C., Seim, I., Calder, R. B., Hoeijmakers, J., et al. (2015). Comparative analysis of genome maintenance genes in naked mole rat, mouse, and human. *Aging Cell* 14, 288–291. doi: 10.1111/accel.12314.
- Magimaidas, A., Madireddi, P., Maifrede, S., Mukherjee, K., Hoffman, B., and Liebermann, D. A. (2016). Gadd45b deficiency promotes premature senescence and skin aging. *Oncotarget* 7, 26935–26948. doi: 10.18632/oncotarget.8854.
- Magnhagen, C. (1990). Reproduction under predation risk in the sand goby, *Pomatoschistus minutus*, and the black goby, *Gobius niger*: the effect of age and longevity. *Behav Ecol Sociobiol* 26, 331–335. doi: 10.1007/BF00171098.
- Manier, M. K., and Arnold, S. J. (2005). Population genetic analysis identifies source–sink dynamics for two sympatric garter snake species (*Thamnophis elegans* and *Thamnophis sirtalis*). *Molecular Ecology* 14, 3965–3976. doi: 10.1111/j.1365-294X.2005.02734.x.
- Manière, X., Krisko, A., Pellay, F. X., Di Meglio, J.-M., Hersen, P., and Matic, I. (2014). High transcript levels of heat-shock genes are associated with shorter lifespan of *Caenorhabditis elegans*. *Experimental Gerontology* 60, 12–17. doi: 10.1016/j.exger.2014.09.005.

- Marquez-Exposito, L., Tejedor-Santamaria, L., Valentijn, F. A., Tejera-Muñoz, A., Rayego-Mateos, S., Marchant, V., et al. (2022). Oxidative Stress and Cellular Senescence Are Involved in the Aging Kidney. *Antioxidants* 11, 301. doi: 10.3390/antiox11020301.
- Martin, S. (2019). General tools for genomic analyses. Contribute to [simonhmartin/genomics_general](https://github.com/simonhmartin/genomics_general) development by creating an account on GitHub. Available at: https://github.com/simonhmartin/genomics_general [Accessed June 15, 2019].
- McGaugh, S. E., Bronikowski, A. M., Kuo, C.-H., Reding, D. M., Addis, E. A., Flagel, L. E., et al. (2015). Rapid molecular evolution across amniotes of the IIS/TOR network. *Proc Natl Acad Sci USA* 112, 7055–7060. doi: 10.1073/pnas.1419659112.
- Mehta, R., Steinkraus, K. A., Sutphin, G. L., Ramos, F. J., Shamieh, L. S., Huh, A., et al. (2009). Proteasomal regulation of the hypoxic response modulates aging in *C. elegans*. *Science* 324, 1196–1198. doi: 10.1126/science.1173507.
- Merico, D., Isserlin, R., Stueker, O., Emili, A., and Bader, G. D. (2010). Enrichment Map: A Network-Based Method for Gene-Set Enrichment Visualization and Interpretation. *PLOS ONE* 5, e13984. doi: 10.1371/journal.pone.0013984.
- Metodiev, S., Mote, B., and Rothschild, M. F. (2006). Mapping of the Porcine CCS and CCR7 Genes. *Bulgarian Journal of Agricultural Science* 12, 153–157.
- Miller, D. A., Janzen, F. J., Fellers, G. M., Kleeman, P. M., and Bronikowski, A. M. (2014). Biodemography of ectothermic tetrapods provides insights into the evolution and plasticity of mortality patterns. *Sociality, Hierarchy, Health: Comparative Biodemography: A Collection of Papers*, 295–313.

- Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34, 267–273. doi: 10.1038/ng1180.
- Moskalev, A. A., and Shaposhnikov, M. V. (2008). [Longevity extension by specific inhibition of pI3K in *Drosophila melanogaster*]. *Adv Gerontol* 21, 602–606.
- Müller, R.-U., Fabretti, F., Zank, S., Burst, V., Benzing, T., and Schermer, B. (2009). The von Hippel Lindau Tumor Suppressor Limits Longevity. *JASN* 20, 2513–2517. doi: 10.1681/ASN.2009050497.
- Ng, P. (2019). SIFT4G Annotator. Available at: https://github.com/pauline-ng/SIFT4G_Annotator.
- Ng, P. (2020). SIFT4G Create Genomic DB. Available at: https://github.com/pauline-ng/SIFT4G_Create_Genomic_DB/.
- Opazo, J. C., Vandewege, M. W., Hoffmann, F. G., Zavala, K., Meléndez, C., Luchsinger, C., et al. (2022). How many sirtuin genes are out there? evolution of sirtuin genes in vertebrates with a description of a new family member. 2020.07.17.209510. doi: 10.1101/2020.07.17.209510.
- Pacifici, F., Della-Morte, D., Piermarini, F., Arriga, R., Scioli, M. G., Capuani, B., et al. (2020). Prdx6 Plays a Main Role in the Crosstalk between Aging and Metabolic Sarcopenia. *Antioxidants* 9, 329. doi: 10.3390/antiox9040329.
- Palacios, M. G., Cunnick, J. E., and Bronikowski, A. M. (2013). Complex Interplay of Body Condition, Life History, and Prevailing Environment Shapes Immune Defenses of Garter Snakes in the Wild. *Physiological and Biochemical Zoology* 86, 547–558. doi: 10.1086/672371.

- Papa, S., Monti, S. M., Vitale, R. M., Bubici, C., Jayawardena, S., Alvarez, K., et al. (2007). Insights into the Structural Basis of the GADD45 β -mediated Inactivation of the JNK Kinase, MKK7/JNKK2*. *Journal of Biological Chemistry* 282, 19029–19041. doi: 10.1074/jbc.M703112200.
- Papa, S., Zazzeroni, F., Bubici, C., Jayawardena, S., Alvarez, K., Matsuda, S., et al. (2004). Gadd45 β mediates the NF- κ B suppression of JNK signalling by targeting MKK7/JNKK2. *Nat Cell Biol* 6, 146–153. doi: 10.1038/ncb1093.
- Passow, C. N., Bronikowski, A. M., Blackmon, H., Parsai, S., Schwartz, T. S., and McGaugh, S. E. (2019). Contrasting Patterns of Rapid Molecular Evolution within the p53 Network across Mammal and Sauropsid Lineages. *Genome Biology and Evolution* 11, 629–643. doi: 10.1093/gbe/evy273.
- Passtoors, W. M., Beekman, M., Deelen, J., van der Breggen, R., Maier, A. B., Guigas, B., et al. (2013). Gene expression analysis of mTOR pathway: association with human longevity. *Aging Cell* 12, 24–31. doi: 10.1111/accel.12015.
- Pastori, D., Pignatelli, P., Farcomeni, A., Menichelli, D., Nocella, C., Carnevale, R., et al. (2016). Aging-Related Decline of Glutathione Peroxidase 3 and Risk of Cardiovascular Events in Patients With Atrial Fibrillation. *Journal of the American Heart Association* 5, e003682. doi: 10.1161/JAHA.116.003682.
- Pawlikowska, L., Hu, D., Huntsman, S., Sung, A., Chu, C., Chen, J., et al. (2009). Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. *Aging Cell* 8, 460–472. doi: 10.1111/j.1474-9726.2009.00493.x.
- Pertea, G., and Pertea, M. (2020). GFF Utilities: GffRead and GffCompare. doi: 10.12688/f1000research.23297.2.

- Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols* 11, 1650–1667. doi: 10.1038/nprot.2016.095.
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 33, 290–295. doi: 10.1038/nbt.3122.
- Polinski, J. M., Zimin, A. V., Clark, K. F., Kohn, A. B., Sadowski, N., Timp, W., et al. (2021). The American lobster genome reveals insights on longevity, neural, and immune adaptations. *Science Advances* 7, eabe8290. doi: 10.1126/sciadv.abe8290.
- Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Auwera, G. A. V., et al. (2018). Scaling accurate genetic variant discovery to tens of thousands of samples. 201178. doi: 10.1101/201178.
- Qi, X., Ng, K. T.-P., Lian, Q., Li, C. X., Geng, W., Ling, C. C., et al. (2018). Glutathione Peroxidase 3 Delivered by hiPSC-MSCs Ameliorated Hepatic IR Injury via Inhibition of Hepatic Senescence. *Theranostics* 8, 212–222. doi: 10.7150/thno.21656.
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033.
- R Core Team (2020). R: A language and environment for statistical computing. Available at: <https://www.R-project.org/>.
- Rao, Y.-Q., Zhou, Y.-T., Zhou, W., Li, J.-K., Li, B., and Li, J. (2021). mTORC1 Activation in Chx10-Specific Tsc1 Knockout Mice Accelerates Retina Aging and Degeneration. *Oxidative Medicine and Cellular Longevity* 2021, e6715758. doi: 10.1155/2021/6715758.

- Reding, D. M., Addis, E. A., Palacios, M. G., Schwartz, T. S., and Bronikowski, A. M. (2016). Insulin-like signaling (IIS) responses to temperature, genetic background, and growth variation in garter snakes with divergent life histories. *General and Comparative Endocrinology* 233, 88–99. doi: 10.1016/j.ygcen.2016.05.018.
- Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., et al. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols* 14, 482–517. doi: 10.1038/s41596-018-0103-9.
- Reinke, B. A., Cayuela, H., Janzen, F. J., Lemaître, J.-F., Gaillard, J.-M., Lawing, A. M., et al. (2022). Diverse aging rates in ectothermic tetrapods provide insights for the evolution of aging and longevity. *Science* 376, 1459–1466. doi: 10.1126/science.abm0151.
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43, e47. doi: 10.1093/nar/gkv007.
- Robert, K. A., and Bronikowski, A. M. (2010). Evolution of senescence in nature: Physiological evolution in populations of garter snake with divergent life histories. *The American Naturalist* 175, 147–159. doi: 10.1086/649595.
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616.
- Ryu, S., Han, J., Norden-Krichmar, T. M., Zhang, Q., Lee, S., Zhang, Z., et al. (2021). Genetic signature of human longevity in PKC and NF- κ B signaling. *Aging Cell* 20, e13362. doi: 10.1111/accel.13362.

- Salovska, B., Kondelova, A., Pimkova, K., Liblova, Z., Pribyl, M., Fabrik, I., et al. (2022). Peroxiredoxin 6 protects irradiated cells from oxidative stress and shapes their senescence-associated cytokine landscape. *Redox Biology* 49, 102212. doi: 10.1016/j.redox.2021.102212.
- Sanders, J. L., Guo, W., O'Meara, E. S., Kaplan, R. C., Pollak, M. N., Bartz, T. M., et al. (2018). Trajectories of IGF-I Predict Mortality in Older Adults: The Cardiovascular Health Study. *The Journals of Gerontology: Series A* 73, 953–959. doi: 10.1093/gerona/glx143.
- Schultz, C., Dick, E. J., Cox, A. B., Hubbard, G. B., Braak, E., and Braak, H. (2001). Expression of stress proteins alpha B-crystallin, ubiquitin, and hsp27 in pallido-nigral spheroids of aged rhesus monkeys. *Neurobiology of Aging* 22, 677–682. doi: 10.1016/S0197-4580(01)00229-9.
- Schwartz, T. S., Arendsee, Z. W., and Bronikowski, A. M. (2015). Mitochondrial divergence between slow- and fast-aging garter snakes. *Experimental Gerontology* 71, 135–146. doi: 10.1016/j.exger.2015.09.004.
- Schwartz, T. S., and Bronikowski, A. M. (2011). “Molecular stress pathways and the evolution of life histories in reptiles,” in *Mechanisms of Life History Evolution* (Oxford: Oxford University Press). doi: 10.1093/acprof:oso/9780199568765.003.0015.
- Schwartz, T. S., and Bronikowski, A. M. (2013). Dissecting molecular stress networks: identifying nodes of divergence between life-history phenotypes. *Mol Ecol* 22, 739–756. doi: 10.1111/j.1365-294X.2012.05750.x.
- Šešlija, D., Blagojević, D., Spasić, M., and Tucić, N. (1999). Activity of superoxide dismutase and catalase in the bean weevil (*acanthoscelides obtectus*) selected for postponed

- senescence. *Experimental Gerontology* 34, 185–195. doi: 10.1016/S0531-5565(98)00078-3.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303.
- Shaposhnikov, M., Proshkina, E., Shilova, L., Zhavoronkov, A., and Moskalev, A. (2015). Lifespan and Stress Resistance in *Drosophila* with Overexpressed DNA Repair Genes. *Sci Rep* 5, 15299. doi: 10.1038/srep15299.
- Shen, D., Li, H., Zhou, R., Liu, M., Yu, H., and Wu, D.-F. (2018). Pioglitazone attenuates aging-related disorders in aged apolipoprotein E deficient mice. *Experimental Gerontology* 102, 101–108. doi: 10.1016/j.exger.2017.12.002.
- Slack, C., Alic, N., Foley, A., Cabecinha, M., Hoddinott, M. P., and Partridge, L. (2015). The Ras-Erk-ETS-Signaling Pathway Is a Drug Target for Longevity. *Cell* 162, 72–83. doi: 10.1016/j.cell.2015.06.023.
- Soerensen, M., Dato, S., Christensen, K., McGue, M., Stevnsner, T., Bohr, V. A., et al. (2010). Replication of an association of variation in the FOXO3A gene with human longevity using both case–control and longitudinal data. *Aging Cell* 9, 1010–1017. doi: 10.1111/j.1474-9726.2010.00627.x.
- Soriano-Arroquia, A., Gostage, J., Xia, Q., Bardell, D., McCormick, R., McCloskey, E., et al. (2021). miR-24 and its target gene Prdx6 regulate viability and senescence of myogenic progenitors during aging. *Aging Cell* 20, e13475. doi: 10.1111/acel.13475.

- Sparkman, A. M., and Bronikowski, A. M. (2013). Avian Predation and the Evolution of Life Histories in the Garter Snake *Thamnophis elegans*. *amid* 170, 66–85. doi: 10.1674/0003-0031-170.1.66.
- Sparkman, A. M., and Palacios, M. G. (2009). A test of life-history theories of immune defence in two ecotypes of the garter snake, *Thamnophis elegans*. *Journal of Animal Ecology* 78, 1242–1248. doi: 10.1111/j.1365-2656.2009.01587.x.
- Sparkman, A. M., Schwartz, T. S., Madden, J. A., Boyken, S. E., Ford, N. B., Serb, J. M., et al. (2012). Rates of molecular evolution vary in vertebrates for insulin-like growth factor-1 (IGF-1), a pleiotropic locus that regulates life history traits. *General and Comparative Endocrinology* 178, 164–173. doi: 10.1016/j.ygcen.2012.04.022.
- Spencer, C. C., Howell, C. E., Wright, A. R., and Promislow, D. E. L. (2003). Testing an ‘aging gene’ in long-lived *Drosophila* strains: increased longevity depends on sex and genetic background. *Aging Cell* 2, 123–130. doi: 10.1046/j.1474-9728.2003.00044.x.
- Stamatoyannopoulos, J. A. (2004). The genomics of gene expression. *Genomics* 84, 449–457. doi: 10.1016/j.ygeno.2004.05.002.
- Stefanetti, R. J., Voisin, S., Russell, A., and Lamon, S. (2018). Recent advances in understanding the role of FOXO3. *F1000Res* 7, F1000 Faculty Rev-1372. doi: 10.12688/f1000research.15258.1.
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *PNAS* 102, 15545–15550.
- Sultan, F. A., Wang, J., Tront, J., Liebermann, D. A., and Sweatt, J. D. (2012). Genetic Deletion of *gadd45b*, a Regulator of Active DNA Demethylation, Enhances Long-Term Memory

- and Synaptic Plasticity. *J. Neurosci.* 32, 17059–17066. doi: 10.1523/JNEUROSCI.1747-12.2012.
- Sun, L., Hu, C., Zheng, C., Qian, Y., Liang, Q., Lv, Z., et al. (2015). FOXO3 variants are beneficial for longevity in Southern Chinese living in the Red River Basin: A case-control study and meta-analysis. *Sci Rep* 5, 9852. doi: 10.1038/srep09852.
- Takahashi, N., van Kilsdonk, J. W. J., Ostendorf, B., Smeets, R., Bruggeman, S. W. M., Alonso, A., et al. (2008). Tumor marker nucleoporin 88kDa regulates nucleocytoplasmic transport of NF- κ B. *Biochemical and Biophysical Research Communications* 374, 424–430. doi: 10.1016/j.bbrc.2008.06.128.
- Tasaki, E., Mitaka, Y., Nozaki, T., Kobayashi, K., Matsuura, K., and Iuchi, Y. (2018). High expression of the breast cancer susceptibility gene BRCA1 in long-lived termite kings. *Aging (Albany NY)* 10, 2668–2683. doi: 10.18632/aging.101578.
- Teumer, A., Qi, Q., Nethander, M., Aschard, H., Bandinelli, S., Beekman, M., et al. (2016). Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits. *Aging Cell* 15, 811–824. doi: 10.1111/accel.12490.
- Theotoki, E. I., Velentzas, A. D., Katarachia, S. A., Papandreou, N. C., Kalavros, N. I., Pasadaki, S. N., et al. (2019). Targeting of copper-trafficking chaperones causes gene-specific systemic pathology in *Drosophila melanogaster*: prospective expansion of mutational landscapes that regulate tumor resistance to cisplatin. *Biology Open* 8, bio046961. doi: 10.1242/bio.046961.
- Tian, X., Seluanov, A., and Gorbunova, V. (2017). Molecular Mechanisms Determining Lifespan in Short- and Long-Lived Species. *Trends in Endocrinology & Metabolism* 28, 722–734. doi: 10.1016/j.tem.2017.07.004.

- Tollis, M., Robbins, J., Webb, A. E., Kuderna, L. F. K., Caulin, A. F., Garcia, J. D., et al. (2019). Return to the Sea, Get Huge, Beat Cancer: An Analysis of Cetacean Genomes Including an Assembly for the Humpback Whale (*Megaptera novaeangliae*). *Molecular Biology and Evolution* 36, 1746–1763. doi: 10.1093/molbev/msz099.
- Tower, J. (2009). Hsps and aging. *Trends in Endocrinology & Metabolism* 20, 216–222. doi: 10.1016/j.tem.2008.12.005.
- Tower, J. (2015). “Superoxide Dismutase (SOD) Genes and Aging in *Drosophila*,” in *Life Extension: Lessons from Drosophila* Healthy Ageing and Longevity., eds. A. M. Vaiserman, A. A. Moskalev, and E. G. Pasyukova (Cham: Springer International Publishing), 67–81. doi: 10.1007/978-3-319-18326-8_3.
- Tsuyuzaki, K., and Nikaido, I. (2021). metaSeq: Meta-analysis of RNA-Seq count data in multiple studies. doi: 10.18129/B9.bioc.metaSeq.
- Tullet, J. M. A., Hertweck, M., An, J. H., Baker, J., Hwang, J. Y., Liu, S., et al. (2008). Direct Inhibition of the Longevity-Promoting Factor SKN-1 by Insulin-like Signaling in *C. elegans*. *Cell* 132, 1025–1038. doi: 10.1016/j.cell.2008.01.030.
- Van der Auwera, G. A., and O’Conner, B. D. (2020). *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. 1st ed.
- Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M., and Ng, P. C. (2016). SIFT missense predictions for genomes. *Nat Protoc* 11, 1–9. doi: 10.1038/nprot.2015.123.
- Vianello, E., Dozio, E., Rigolini, R., Marrocco-Trischitta, M. M., Tacchini, L., Trimarchi, S., et al. (2016). Acute phase of aortic dissection: a pilot study on CD40L, MPO, and MMP-1, -2, 9 and TIMP-1 circulating levels in elderly patients. *Immunity & Ageing* 13, 9. doi: 10.1186/s12979-016-0063-2.

- Whitehead, A., and Crawford, D. L. (2006). Variation within and among species in gene expression: raw material for evolution. *Molecular Ecology* 15, 1197–1211. doi: 10.1111/j.1365-294X.2006.02868.x.
- Willcox, B. J., Donlon, T. A., He, Q., Chen, R., Grove, J. S., Yano, K., et al. (2008). FOXO3A genotype is strongly associated with human longevity. *Proceedings of the National Academy of Sciences* 105, 13987–13992. doi: 10.1073/pnas.0801030105.
- Yang, J., and Tower, J. (2009). Expression of hsp22 and hsp70 Transgenes Is Partially Predictive of Drosophila Survival Under Normal and Stress Conditions. *The Journals of Gerontology: Series A* 64A, 828–838. doi: 10.1093/gerona/glp054.
- Yen, K., Patel, H. B., Lublin, A. L., and Mobbs, C. V. (2009). SOD isoforms play no role in lifespan in ad lib or dietary restricted conditions, but mutational inactivation of SOD-1 reduces life extension by cold. *Mechanisms of Ageing and Development* 130, 173–178. doi: 10.1016/j.mad.2008.11.003.
- Yu, Z., Seim, I., Yin, M., Tian, R., Sun, D., Ren, W., et al. (2021). Comparative analyses of aging-related genes in long-lived mammals provide insights into natural longevity. *The Innovation* 2, 100108. doi: 10.1016/j.xinn.2021.100108.
- Zhang, H.-M., Diaz, V., Walsh, M. E., and Zhang, Y. (2017). Moderate lifelong overexpression of tuberous sclerosis complex 1 (TSC1) improves health and survival in mice. *Sci Rep* 7, 834. doi: 10.1038/s41598-017-00970-7.

Figures

Figure 4.1 Map with sample geographic and dataset distribution

Figure 4.2 Differentially Expressed Genes from Meta Analysis

Figure 4.3 Networks of Enriched Pathways

Figure 4.4 Sequence divergence for "within" vs "between" ecotype population comparisons

Figure 4.5 Targeted Genes and Divergent Loci

Figure 4.6 Summary of differential gene expression and sequence variation results

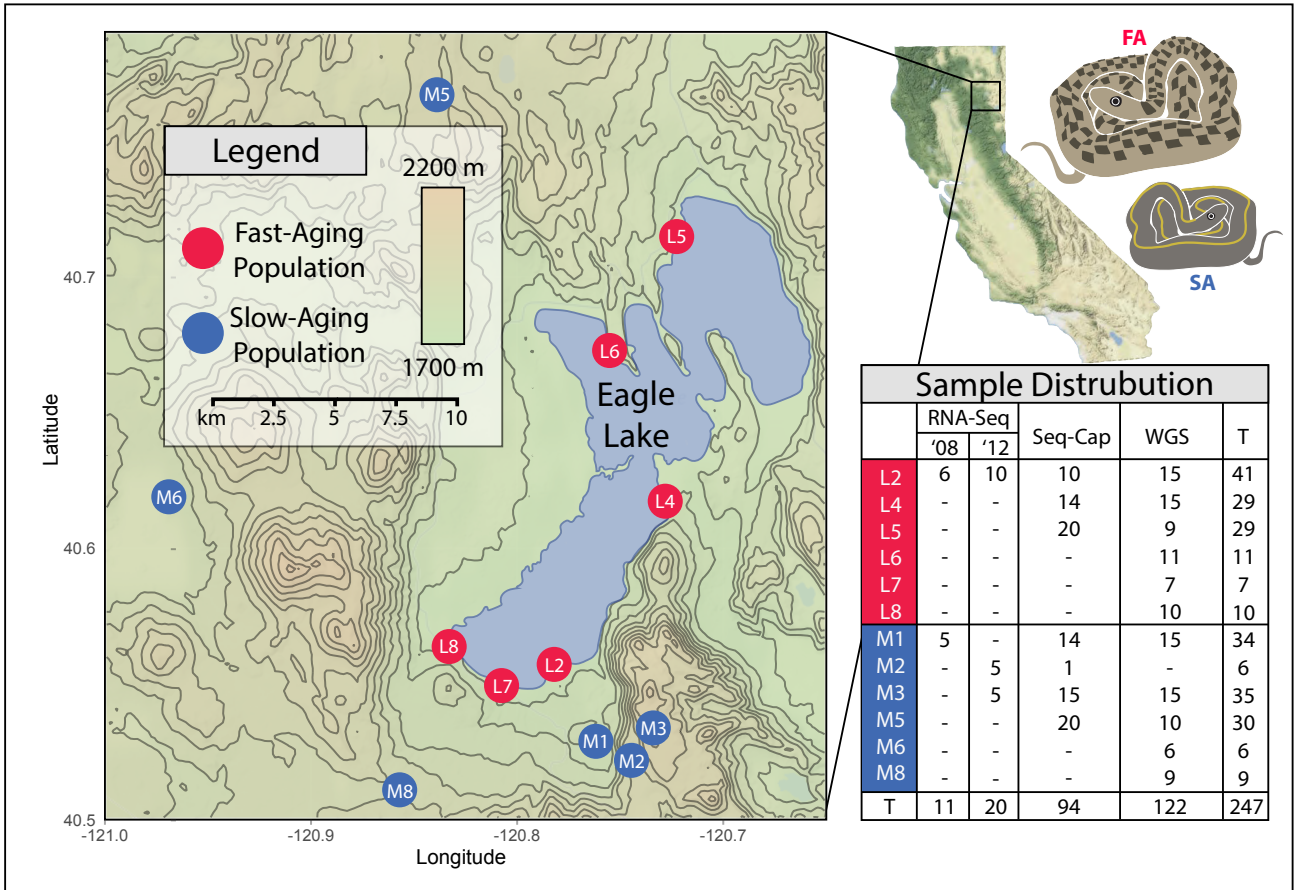


Figure 4.1: Map with sample geographic and dataset distribution

Map of Eagle Lake and surrounding topography (CA, USA) with garter snake population localities. Ecotype populations are shown in red (FA) and blue (SA). Sample distribution within datasets of populations are shown within the table.

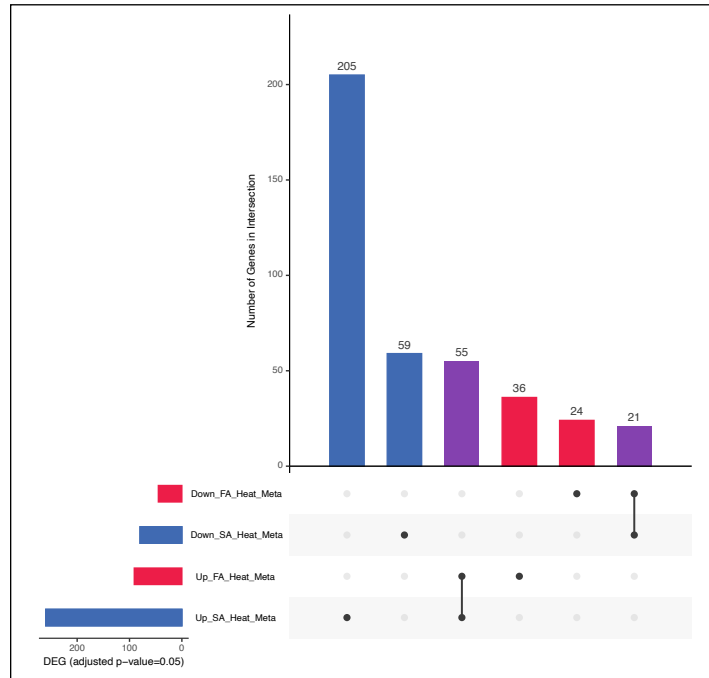


Figure 4.2: **Differentially Expressed Genes from Meta Analysis**

Upset plot showing the number of differentially expressed genes from the meta analysis of RNA-seq data. The genes found to be differentially expressed in the meta analysis were divided amongst categories for up/down regulation in each of the ecotypes (FA/SA). The horizontal bar chart (labeled "DEG") shows the number of differentially expressed genes contained within each of the sets. The vertical bar chart (labeled "number of genes in intersection") shows the number of genes from the DEG sets that are unique to each of the intersections between sets.

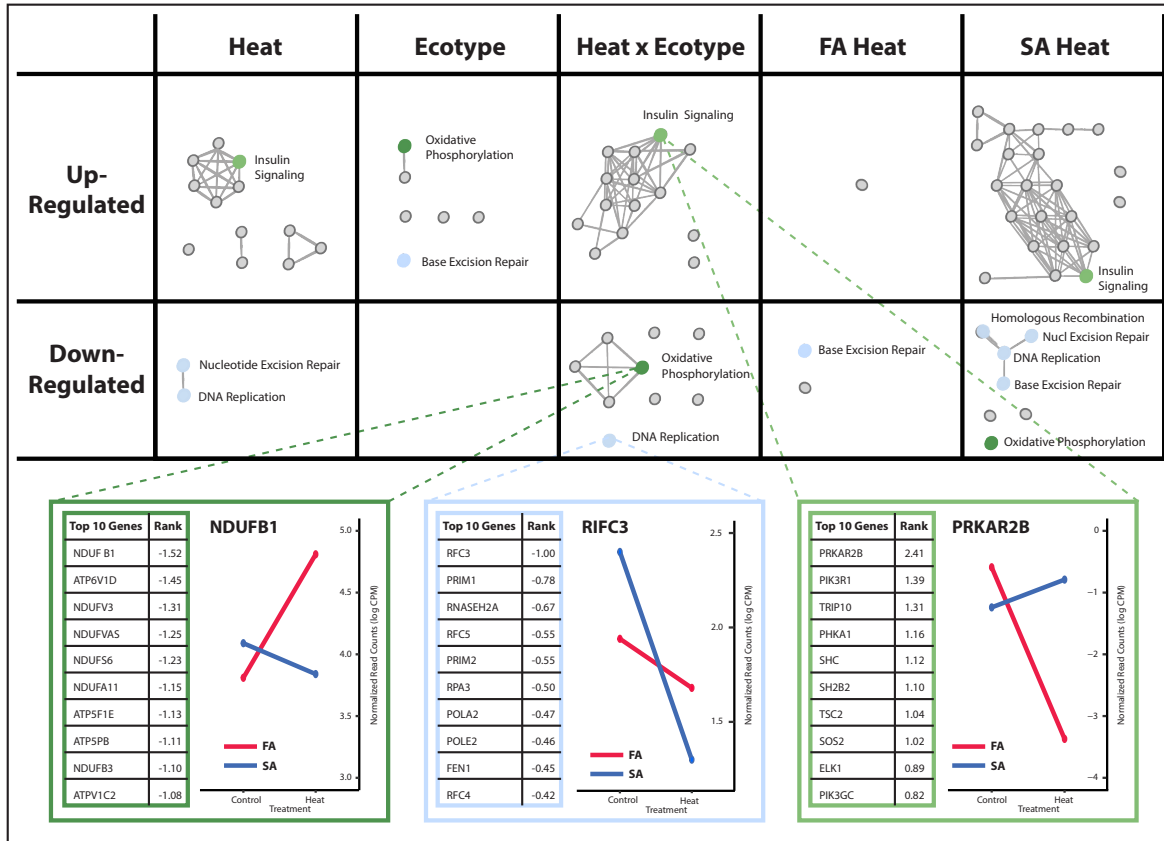


Figure 4.3: Networks of Enriched Pathways

Results from Gene Set Enrichment Analysis where enriched pathways from each model were imported into Cytoscape for visualization as a network of pathways. Each node is a KEGG pathway that was enriched (FDR ≤ 0.1) for differentially expressed genes in the following statistical models: Heat relative to Control, Ecotype (SA relative to FA), Heat x Ecotype interaction (SA relative to FA), Heat for FA ecotype, Heat for SA ecotype. Here, the Cytoscape plot was simplified- node sizes, edge lengths, and edge widths do not convey information. For the full Cytoscape plot, along with each pathway name, see Figure S5. Colors correspond to focal gene networks/pathways that we found to be enriched (non-focal gene networks/pathways are grey).

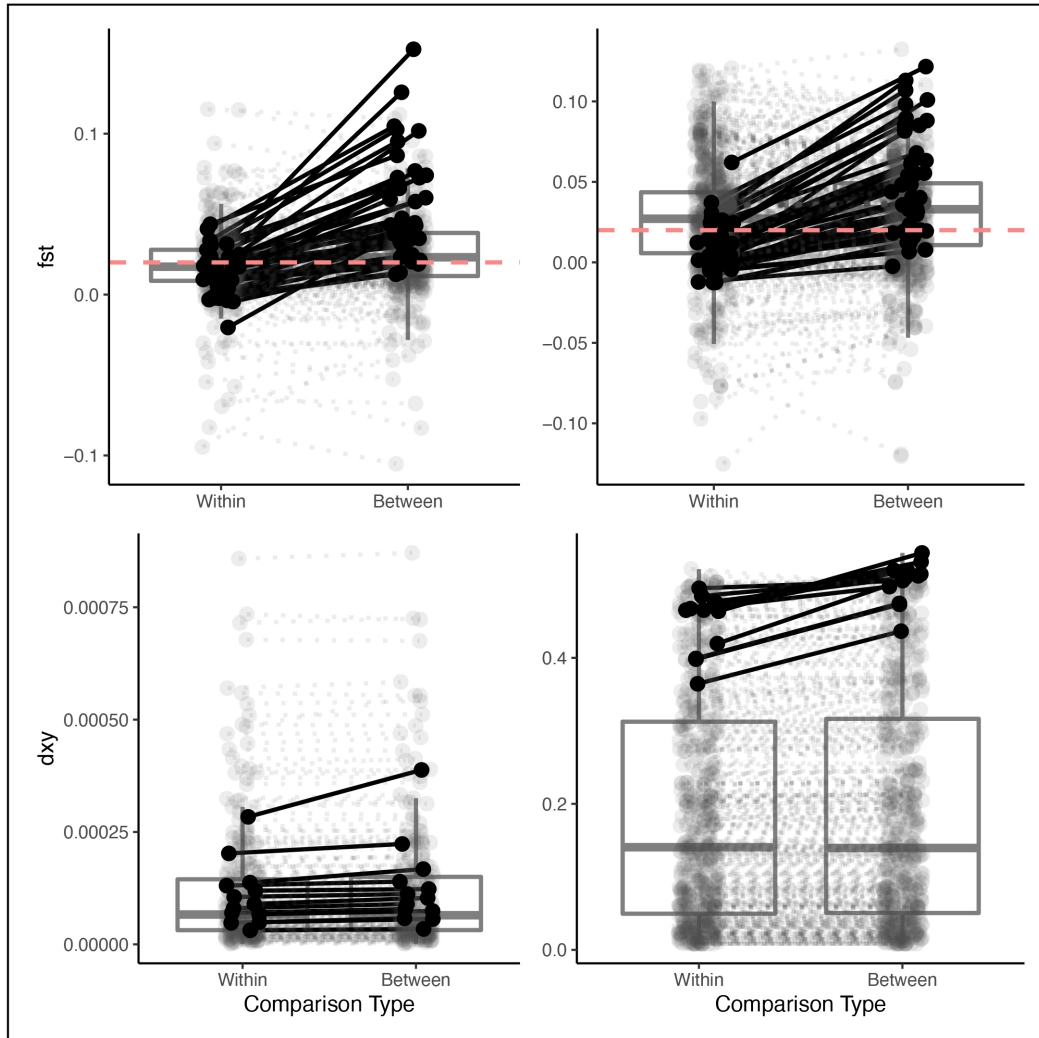


Figure 4.4: Average Fst and dxy for "within" vs "between" ecotype population comparisons

The number and categorical distribution of statistically significant findings from analyses of the within-ecotype vs between-ecotype Fst/Dxy comparisons for both exon regions and missense sites are shown. Left column: Each point corresponds to the Fst/dxy value calculated from the exonic regions of a focal gene. The red dotted line marks the genome wide Fst value between ecotypes (0.02). Right column: Each point corresponds to the Fst/dxy value calculated from a nonsynonymous SNP from a focal gene. For each plot, the "Within" category refers to pairwise comparisons for populations of the same ecotype (i.e., within-ecotype comparisons) whereas the "Between" category refers to pairwise comparisons for populations of differing ecotype (i.e., between-ecotype comparisons). The lines between points connect the same genes for between and within comparisons (i.e., the effect size). Black points/lines are the genes/sites with statistically significant differences in Fst or dxy between comparison type ("within" or "between"). Transparent points with dotted lines are genes/sites from targeted genes with no statistically significant differences.

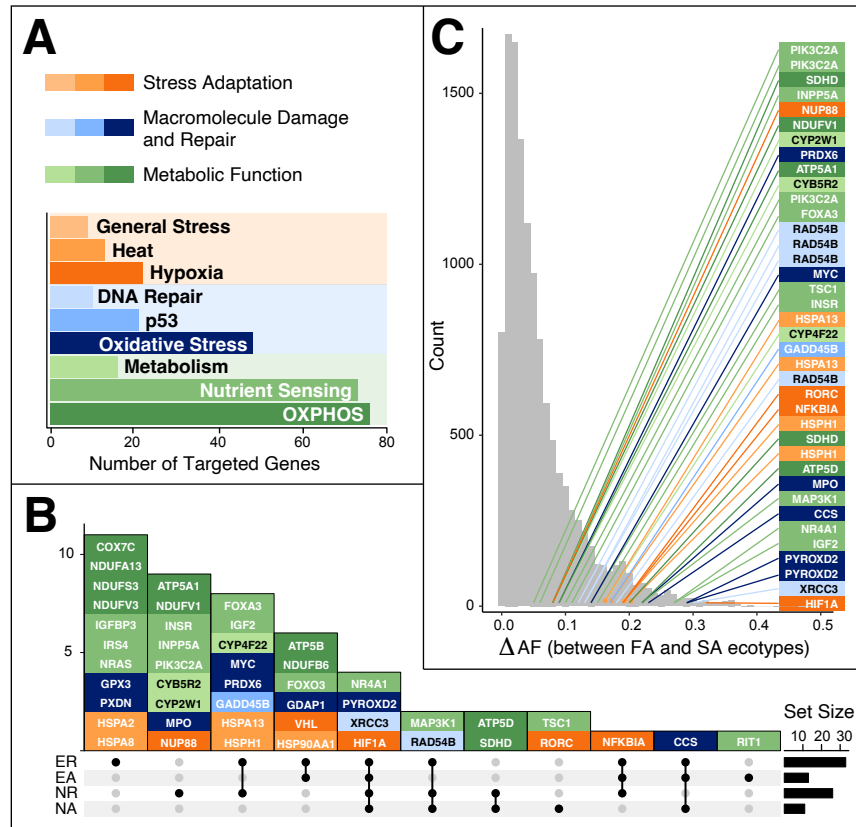


Figure 4.5: Targeted Genes and Significant Loci

A: Number of genes we targeted, shown within their categories. The color categories (orange, blue, green) correspond to three pillars of aging (stress adaptation, macromolecule damage and repair, and metabolic function). **B:** Upset plot for genes of interest based on statistical significance in *Fst* and/or *dxy* from exon regions or nonsynonymous sites. Sets: ER/EA sets contain the focal genes whose exon regions have significantly different relative (*fst*; ER) or absolute (*dxy*; EA) divergence values for between- vs within-ecotype population comparisons. NR/NA sets contain the nonsynonymous SNPs from the focal genes with significantly different relative (*fst*; NR) or absolute (*dxy*; NA) divergence values for between- vs within-ecotype population comparisons. The different intersections of genes within these categories are shown within the sets by dots and lines, with the number of genes pertaining to each intersection displayed in the top barchart. The side barchart shows the number of genes within each set. **C:** Differences in allele frequency (ΔAF) of nonsynonymous SNPs between ecotypes. Across the great majority of significant nonsynonymous SNPs within the focal genes, the alternate allele frequency was similar for both the FA and SA ecotypes. Distribution of ΔAF between FA and SA ecotypes for all nonsynonymous SNPs from focal genes is shown in grey. Significant nonsynonymous SNPs with *Fst* and/or *dxy* values are listed with a line directed to their respective ΔAF ; most of these are contained within the right tail of the distribution. The colors of the genes correspond to the hallmarks of aging to which each gene corresponds.

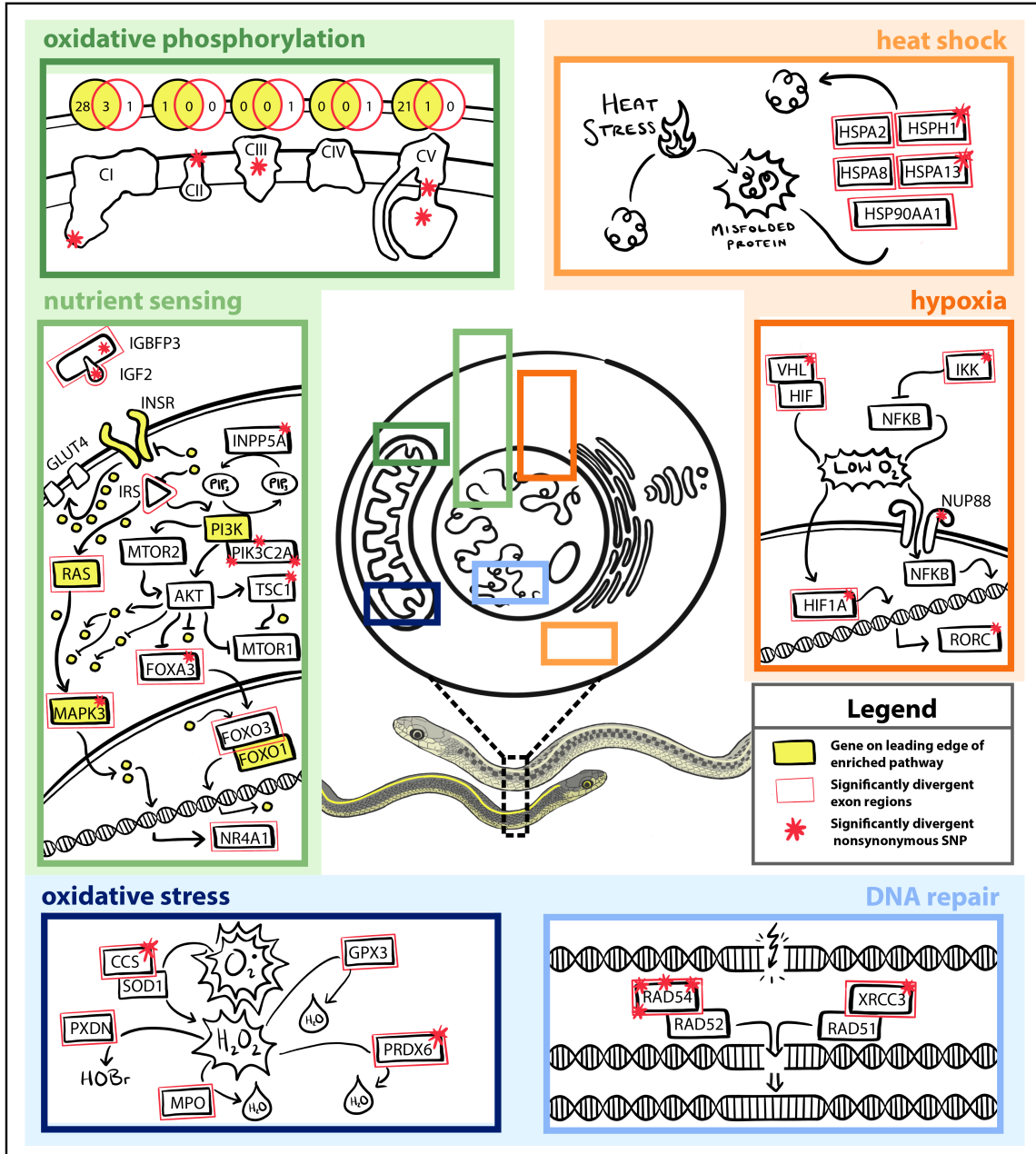


Figure 4.6: Summary of differential gene expression and sequence variation results

Summary of ecotype divergence at the gene network level in the context of primary cellular localities involved. Focal pathways of interest are colored similar to Figure 4.5 within their respective hallmarks are shown. Genes on leading edge of enriched pathways (either up- or down-regulated in Ecotype treatment or Heat x Ecotype interaction- see Fig. 2) are highlighted in yellow. The unmarked yellow circles are genes on the leading edge that we didn't focus on, but show the extent of the enrichment within the pathway. Genes with sequence divergence significantly different in exon regions between ecotypes are outlined in red. Nonsynonymous SNPs with divergence significantly different between ecotypes are marked with red asterisks.

Ch 4: Tables

Table 4.1: Three Cellular Hallmarks of Aging and Differences Between FA and SA Snakes

| Cellular Hallmark | Characteristic | FA | SA | References |
|----------------------------------|--|---|---|---|
| Metabolism | Mass-independent VO ₂ | Higher | Lower | Bronikowski & Vleck, 2010 Gangloff et al. 2015, 2016 |
| | Basal & Maximal Cellular VO ₂ | Increases with age | Decreases with age | Gangloff et al. 2020 |
| | Mitochondrial transcription | Higher | Lower | Schwartz et al. 2015 |
| | Mitochondrial efficiency (P:O) | Lower | Higher | Robert & Bronikowski 2010 |
| | Liver expression of IGF hormones | Higher IGF1R & IGF2R | Lower IGF1R & IGF2R | Reding et al. 2016 |
| | Blood circulating IGF1 | Higher IGF1 at cool temp Constant over adult age | Lower IGF1 at cool temp Decreases with adult age | Reding et al. 2016 Addis et al. 2017 Sparkman et al. 2009 |
| | Blood circulating IGF2 | Higher IGF2 at warm temp | Lower IGF2 at warm temp | Reding et al. 2016 |
| Macromolecular Damage | Mitochondrial superoxide | Higher under heat stress | Lower under heat stress | Schwartz et al. 2015 Schwartz & Bronikowski 2013 |
| | Hydrogen peroxide production | Higher | Lower | Robert et al. 2010 |
| | DNA damage and repair | Lower repair capacity | Higher repair capacity | Robert & Bronikowski 2010 Bronikowski 2008 |
| Adaptation to Stress | Natural antibodies | Higher | Lower | Sparkman & Palacios 2009 |
| | Complement-mediated lysis | Higher | Lower | Sparkman & Palacios 2009 |
| | Bactericidal competence | Higher or same | Lower or same | Palacios et al. 2013 |
| | T-lymphocyte proliferation | Higher | Lower | Palacios et al. 2013 |
| | B-lymphocyte proliferation | Higher | Lower | Palacios et al. 2013 |

Supplementary Figures

Figure 4.S1 Bioinformatics Pipeline - Gene Expression

Figure 4.S2 Bioinformatics Pipeline - Sequence Variation

Figure 4.S3 Cytoscape networks of enriched pathways

Figure 4.S4 Insulin Signaling Gene Expression enplot

Figure 4.S5 Oxidative Phosphorylation FA Heat Treatment Gene Expression enplot

Figure 4.S6 Oxidative Phosphorylation SA Heat Treatment Gene Expression enplot

Figure 4.S7 Nonsynonymous SNP Functional Impact Scores

Figure 4.S7 SNP frequency distributions

Figure 4.S8 Comparison of Dxy between ecotypes for each category

Figure 4.S9 Electron Transport Chain Significant Subunits and Nonsynonymous SNPs

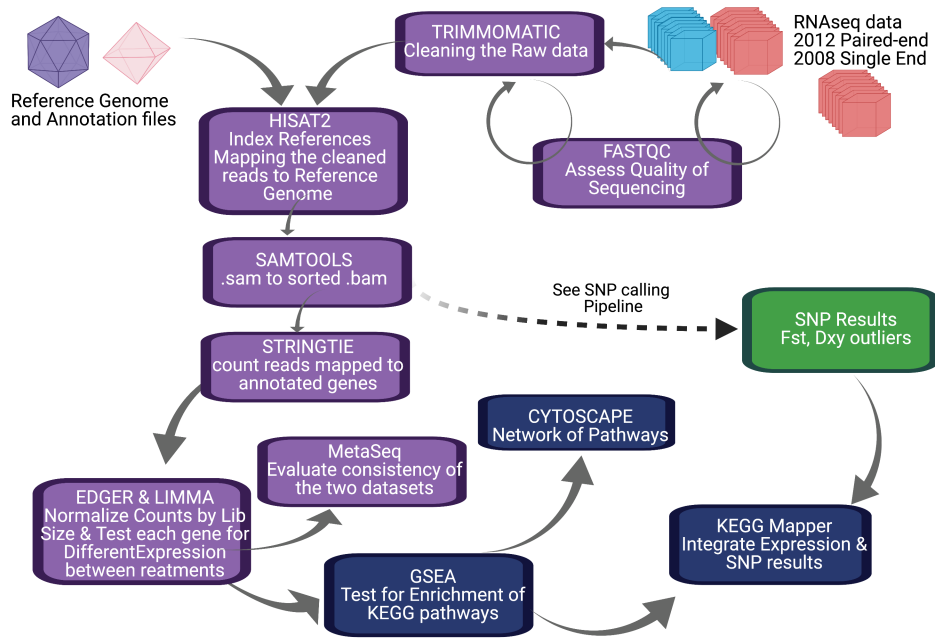


Figure 4.S1: Bioinformatics Pipeline - Gene Expression

The bioinformatics pipeline used to process genomic data to examine gene expression for this study.

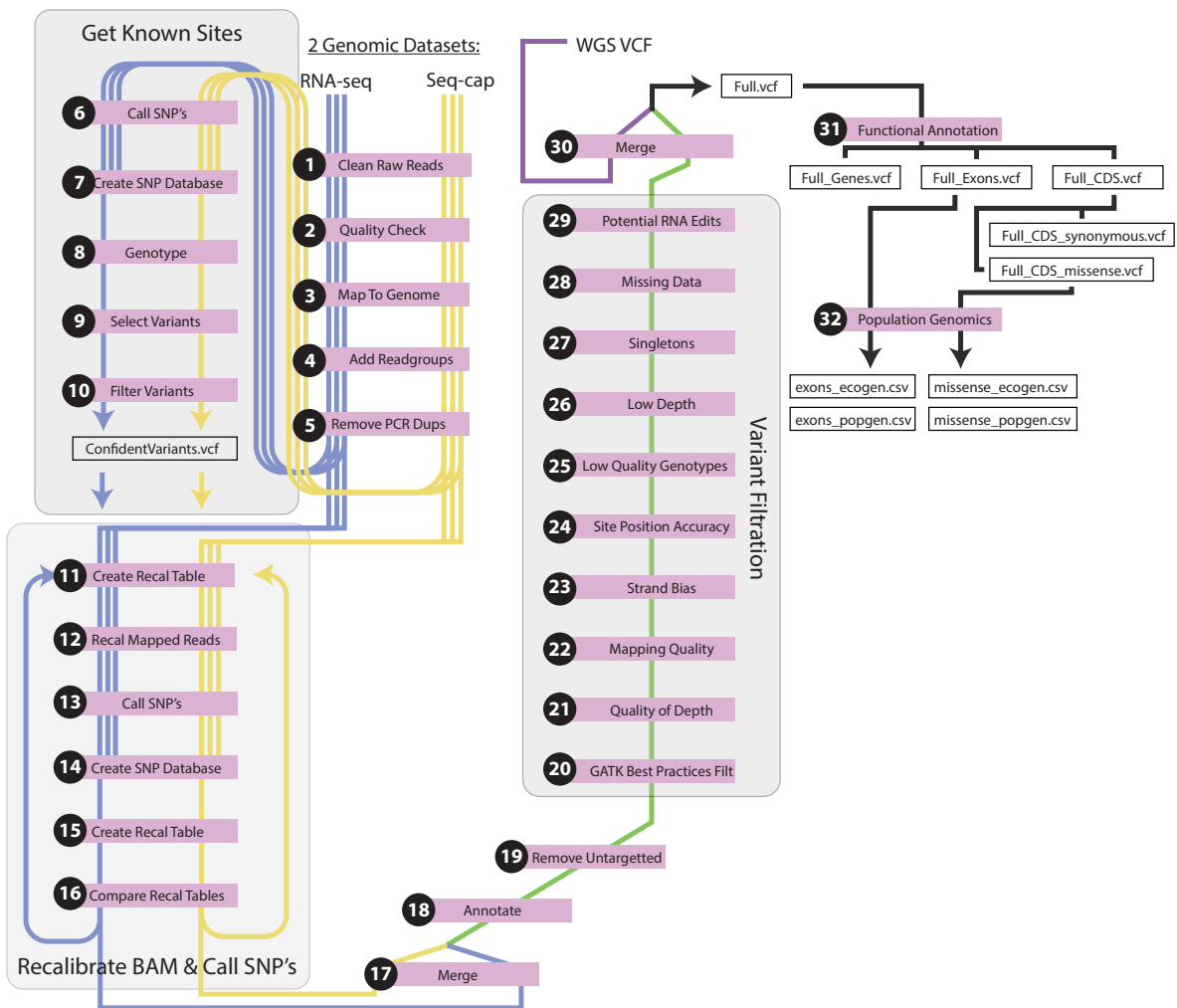


Figure 4.S2: **Bioinformatics Pipeline - Sequence Variation**

The bioinformatics pipeline used to process genomic data to examine sequence variation for this study.

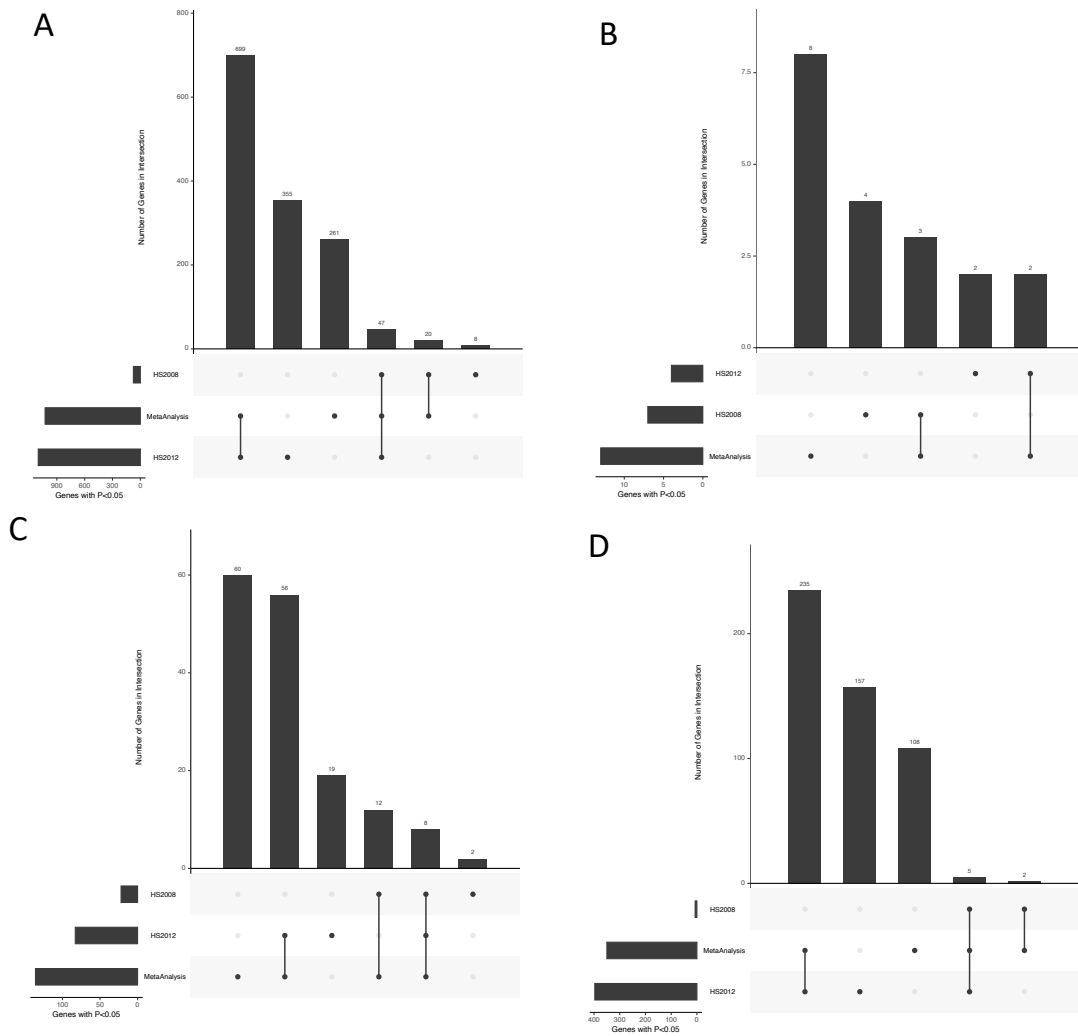


Figure 4.S3: Gene Expression Upset Plots

Upset plots summarizing the results from the differential gene expression analysis of four statistical models using the 2008 Heat Stress Dataset (HS2008), the 2012 Heat Stress Dataset (HS2012), or the meta-analysis across both datasets. Each panel displays the number of differentially expressed genes for each dataset (horizontal bars) and their overlap among datasets (vertical bars). A. Testing for effect of treatment (Heat vs Control Treatments). B. Testing for effect of Ecotype (Slow-living vs Fast-living Ecotypes). C. Testing for the effect of Treatment within the Fast-living Ecotype only. D. Testing for the effect of Treatment within the Slow-living Ecotype only. Note the Y-axis for the vertical bar graph, and the X-axis for the horizontal bar plot are not standardized across the panels.

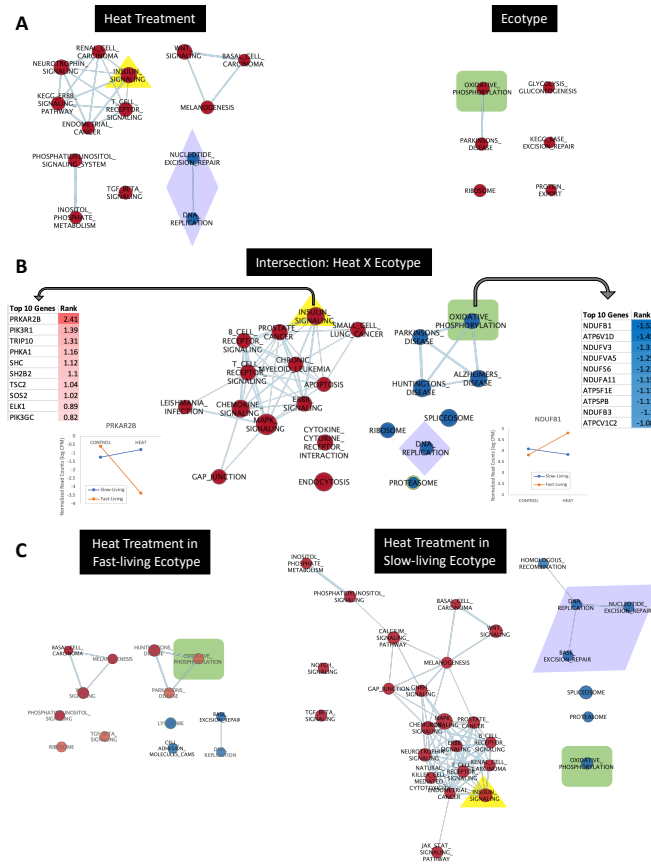


Figure 4.S4: Cytoscape networks of enriched pathways

Results from Gene Set Enrichment Analysis where enriched pathways from each statistical model were imported into Cytoscape for visualization as a network of pathways for each model. Within each network, each node is a KEGG pathway that was enriched ($FDR < 0.1$) for high ranking (upregulated pathways, Red) or low ranking genes (downregulated pathways, blue). The size of node is relative to the number of genes in that pathway. The width of the edges between KEGG pathways represented the number of genes shared between the pathways with overlap set to 0.3, and the length of the edges represent the connectedness of the node to the network. The background shapes highlight the same focal pathways affected by the different models: yellow triangle=Insulin Signaling; green square = Oxidative Phosphorylation; lavender diamond= DNA damage/repair. A. Right, the network of KEGG pathways enriched in response to Heat Treatment relative to Control Treatment; Left, pathways enriched between ecotypes, Slow-living relative relative to Fast-living. B. Network of pathways enriched under the interaction between Heat Treatment and Ecotype. Tables on sides provide examples of the top ten leading edge genes for two enriched pathways: Insulin signaling or Oxidative Phosphorylation. The line graphs provide examples of the interaction for a single gene, the top ranking gene, in each of those pathways. C. Ecotypes are separated to further illustrate their unique responses to heat stress. Right, the network of KEGG pathways enriched in response to Heat Treatment relative to Control Treatment in the Fast-living Ecotype. Here the $FDR = 0.17$ to illustrate the pattern for Oxidative Phosphorylation. All pathway names with $FDR 0.10$ to 0.17 are in gray. Left, the network of KEGG pathways enriched in response to Heat Treatment relative to Control Treatment in the Fast-living Ecotype.

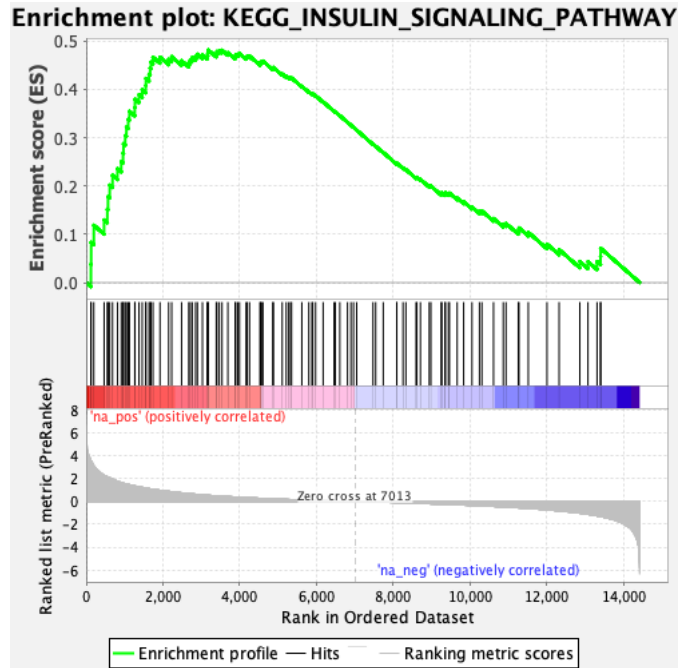


Figure 4.S5: **Insulin Signaling Gene Expression enplot**

Plot of enrichment score (ES) for IIS genes in SA ecotype in response to heat. ES indicates degree to which a gene is overrepresented at the top or bottom of a ranked list. The top plot shows the ES calculated along the gene set, with the leading edge subset of genes occurring prior to the maximum ES (the ES for the gene set). The gene position for each of the genes in the set among all ranked genes from the study is shown in the middle plot. The bottom plot shows the ranking metric, which measures the signal-to-noise ratio for each gene.

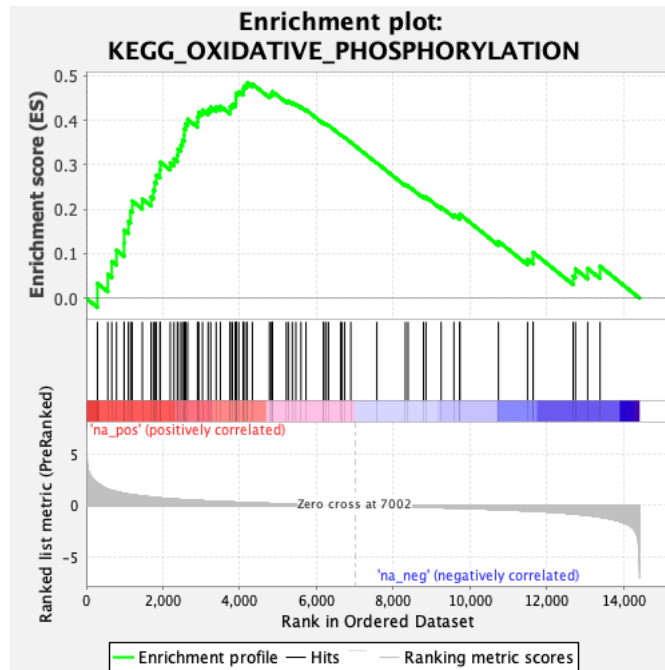


Figure 4.S6: **Oxidative Phosphorylation FA Heat Treatment Gene Expression enplot**

Plot of enrichment score (ES) for OXPPOS genes in FA ecotype in response to heat. ES indicates degree to which a gene is overrepresented at the top or bottom of a ranked list. The top plot shows the ES calculated along the gene set, with the leading edge subset of genes occurring prior to the maximum ES (the ES for the gene set). The gene position for each of the genes in the set among all ranked genes from the study is shown in the middle plot. The bottom plot shows the ranking metric, which measures the signal-to-noise ratio for each gene.

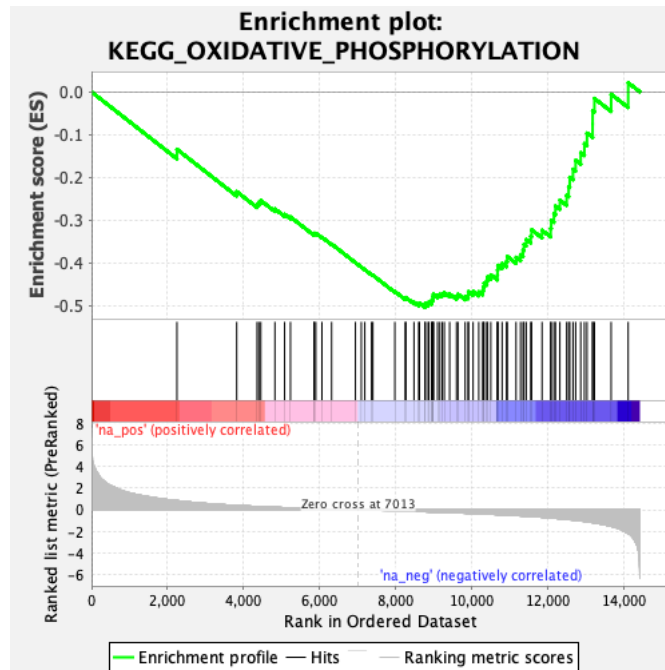


Figure 4.S7: **Oxidative Phosphorylation SA Heat Treatment Gene Expression enplot**

Plot of enrichment score (ES) for OXPPOS genes in SA ecotype in response to heat. ES indicates degree to which a gene is overrepresented at the top or bottom of a ranked list. The top plot shows the ES calculated along the gene set, with the leading edge subset of genes occurring prior to the maximum ES (the ES for the gene set). The gene position for each of the genes in the set among all ranked genes from the study is shown in the middle plot. The bottom plot shows the ranking metric, which measures the signal-to-noise ratio for each gene.

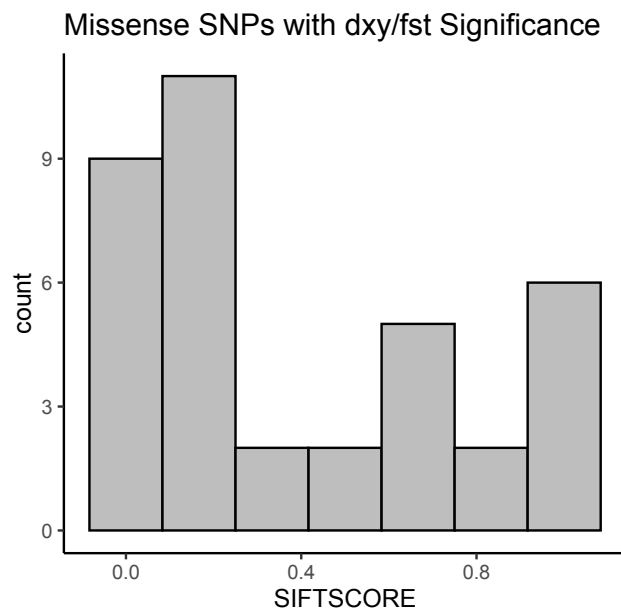


Figure 4.S8: Nonsynonymous SNP Functional Impact Scores

SIFT4g scores for each significant nonsynonymous SNP. SIFT scores ≤ 0.05 are predicted to be “deleterious”.

Scores for each site are available in the supplemental file “topfst_dxy_missense_complete.csv”.

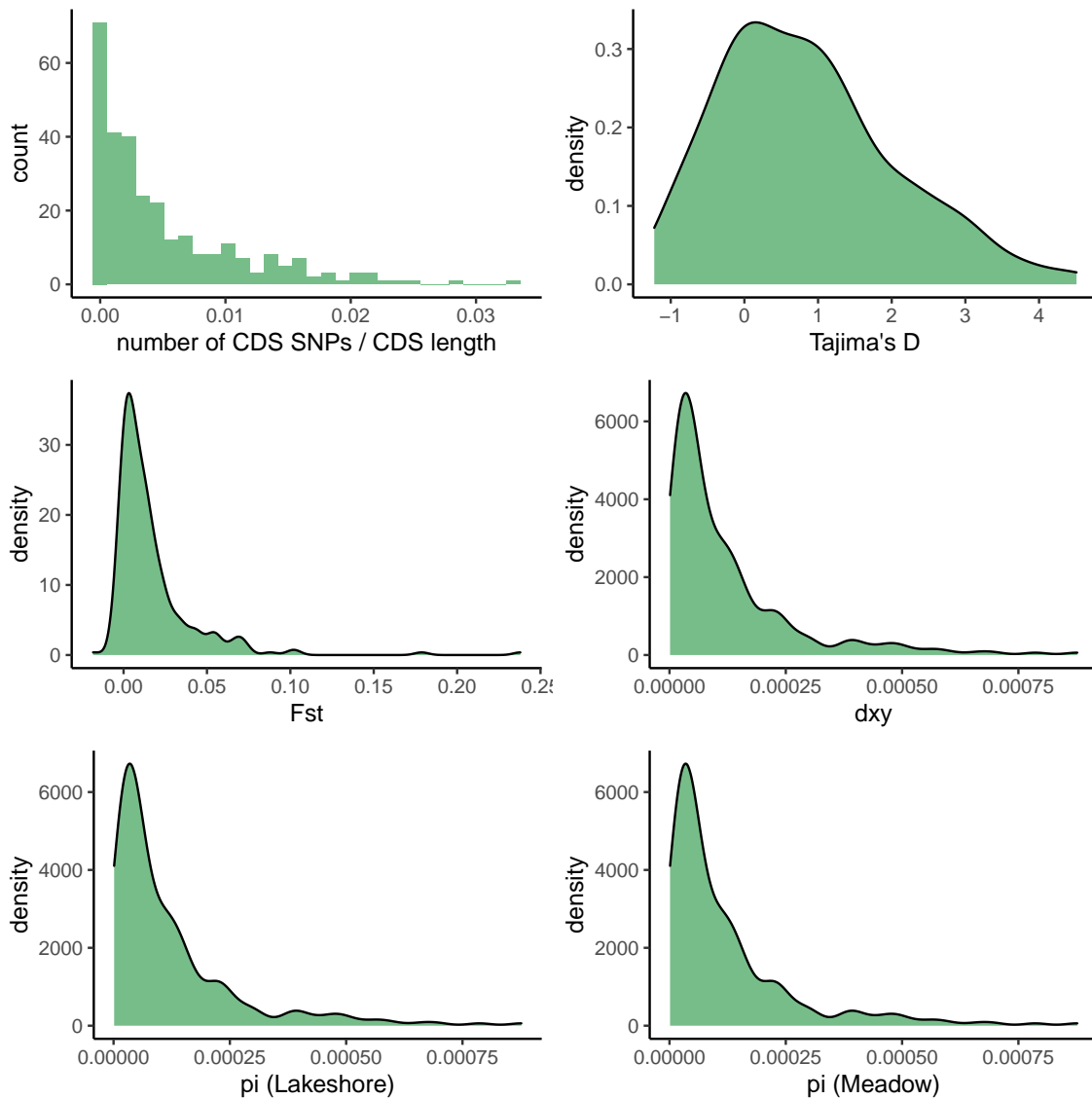


Figure 4.S9: SNP frequency distributions

Frequency distribution for a few perspectives of the genomic dataset. Top left: Number of SNPs within the coding sequence (CDS) normalized by the length of the CDS for each focal gene. Top right: Tajima's D for exon regions of each focal gene. Center: Calculated Fst (left) and dxy (right) from exon regions of each focal gene. Bottom: Calculated pi for each ecotype (FA, SA) from exon regions of each focal gene.

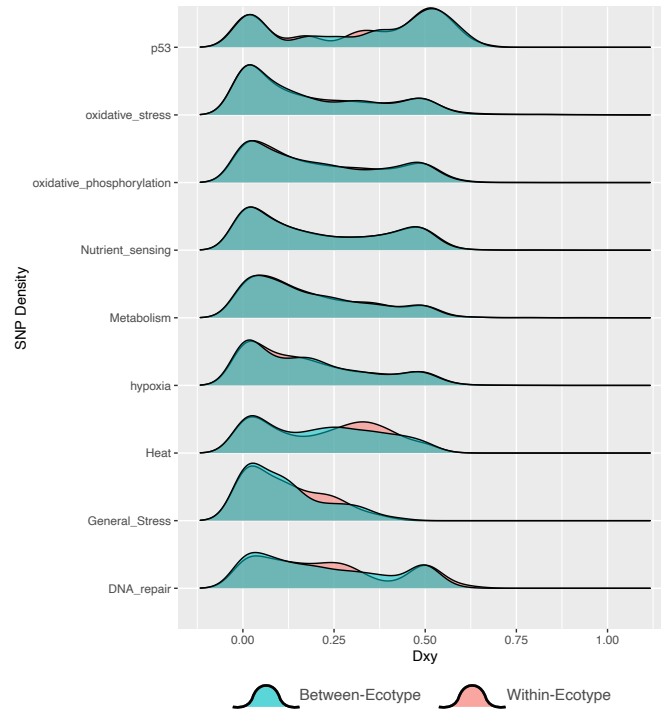


Figure 4.S10: Comparison of Dxy between ecotypes for each category

Statistically significant differences between ecotypes for Dxy shown here; we interpret these differences as not biologically significant. The only evident non-overlapping segment of distribution is in the slightly noticeable difference in the Heat and General Stress categories.

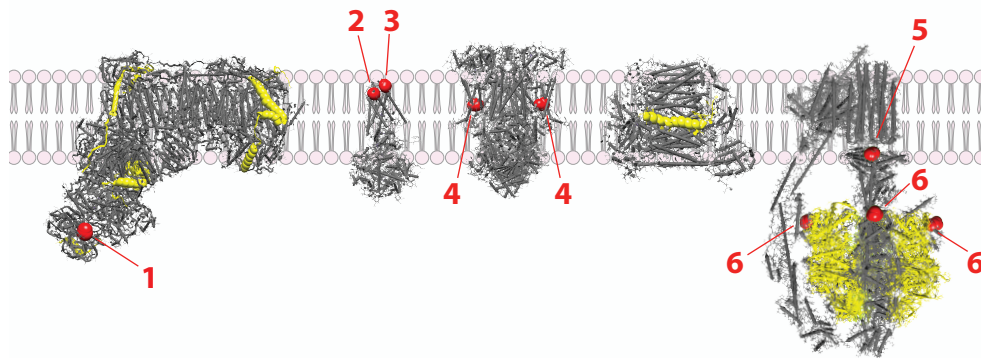


Figure 4.S11: Electron Transport Chain Significant Subunits and Nonsynonymous SNPs

Complexes of the electron transport chain with subunits/residues with significant fst/dxy values highlighted (subunits: yellow; residues: red). Complex I (subunits with significant divergence: NDUFA13, NDUFB6, NDUFS3, NDUFV3, and NDUFV1), Complex II (subunits with significant divergence: SDHD), Complex IV (subunits with significant divergence: COX7C), and Complex V (subunits with significant divergence: ATP5A1, ATP5B, and ATP5D)

Chapter 4: Supplementary Tables

Table 4.S1 Targeted Genes

Table 4.S2 Significant Gene Regions

Table 4.S3 Significant Nonsynonymous SNPs

Table 4.S1: Statistics for Targeted Genes

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|--------------|--------------|----------|--------------|-------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| DNAA1 | DNAA1 | MDR | DR | NC_045543.1 | 451014 | 45109315 | 7900 | 19 | 5379 | 4 | 1186 | 2 | 2 | 0 | 0.002 | 0.002 | 1.1E-05 | 1.1E-05 | 0.00 | 1.1E-05 | -0.94 |
| DNAA3 | DNAA3 | MDR | DR | NC_045544.1 | 263769 | 270867 | 6597 | 43 | 3377 | 7 | 1420 | 4 | 2 | 2 | 0.003 | 0.001 | 2.1E-05 | 2.1E-05 | 0.00 | 2.1E-05 | -0.96 |
| DNAA22 | DNAA22 | MDR | DR | NC_045542.1 | 3552308 | 35562127 | 9818 | 7 | 1712 | 4 | 997 | 4 | 2 | 2 | 0.004 | 0.002 | 1.2E-04 | 1.2E-04 | 0.00 | 1.2E-04 | 0.08 |
| RAD51 | RAD51 | MDR | DR | NC_045541.1 | 118430697 | 118441466 | 10768 | 24 | 3351 | 2 | 1011 | 2 | 0 | 2 | 0.002 | 0.000 | 8.7E-05 | 8.7E-05 | 0.02 | 8.7E-05 | 1.11 |
| RAD51B | RAD51B | MDR | DR | NC_045541.1 | 111552087 | 111966535 | 414447 | 13 | 4274 | 5 | 1169 | 2 | 0 | 2 | 0.002 | 0.000 | 1.1E-04 | 1.1E-04 | 0.05 | 1.1E-04 | 3.65 |
| RAD51C | RAD51C | MDR | DR | NC_045544.1 | 137968221 | 138019047 | 50825 | 83 | 2566 | 19 | 1101 | 17 | 5 | 12 | 0.015 | 0.005 | 1.0E-04 | 1.0E-04 | 0.00 | 1.0E-04 | 0.74 |
| RAD54B | RAD54B | MDR | DR | NC_045548.1 | 53653379 | 53696736 | 43356 | 83 | 13390 | 25 | 2656 | 22 | 14 | 8 | 0.008 | 0.005 | 1.8E-05 | 1.8E-05 | 0.02 | 1.8E-05 | 1.39 |
| TERT | TERT | MDR | DR | NC_045546.1 | 77517130 | 77561430 | 44299 | 0 | 3956 | 0 | 3854 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| XRCC2 | XRCC2 | MDR | DR | NC_045538.1 | 9543105 | 9551339 | 8233 | 9 | 1282 | 6 | 831 | 6 | 2 | 4 | 0.007 | 0.002 | 1.6E-04 | 1.6E-04 | 0.00 | 1.6E-04 | 0.48 |
| XRCC3 | XRCC3 | MDR | DR | NC_045541.1 | 14897494 | 149011978 | 14483 | 29 | 5581 | 10 | 1106 | 7 | 2 | 5 | 0.006 | 0.002 | 5.6E-05 | 5.6E-05 | 0.07 | 5.6E-05 | 2.05 |
| AIFM1 | AIFM1 | MDR | OS | NC_045552.1 | 47589958 | 47617405 | 27446 | 69 | 2096 | 24 | 1804 | 17 | 4 | 13 | 0.009 | 0.002 | 7.6E-05 | 7.6E-05 | 0.00 | 7.6E-05 | 0.12 |
| AIFM2 | AIFM2 | MDR | OS | NC_045555.1 | 13454546 | 13466706 | 12159 | 41 | 4662 | 25 | 1114 | 9 | 5 | 4 | 0.008 | 0.004 | 3.6E-05 | 3.6E-05 | 0.00 | 3.6E-05 | 0.34 |
| ALPK1 | ALPK1 | MDR | OS | NC_045549.1 | 26291217 | 26349280 | 58062 | 23 | 7789 | 10 | 3662 | 6 | 5 | 1 | 0.002 | 0.001 | 1.2E-05 | 1.2E-05 | 0.01 | 1.2E-05 | -0.90 |
| AIF2 | AIF2 | MDR | OS | NC_045541.1 | 23631386 | 23679141 | 4754 | 14 | 4819 | 10 | 1453 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| CCS | CCS | MDR | OS | NC_045557.1 | 9237017 | 9246390 | 9372 | 21 | 1231 | 7 | 823 | 5 | 3 | 2 | 0.006 | 0.004 | 2.3E-04 | 2.3E-04 | 0.04 | 2.3E-04 | 1.26 |
| GDAF1 | GDAF1 | MDR | OS | NC_045548.1 | 46590053 | 46597877 | 7823 | 9 | 3305 | 5 | 1059 | 4 | 0 | 4 | 0.004 | 0.000 | 8.0E-05 | 8.0E-05 | 0.03 | 8.0E-05 | 1.01 |
| GPX1 | GPX1 | MDR | OS | NC_045542.1 | 100194069 | 100197177 | 3107 | 31 | 1344 | 22 | 580 | 9 | 4 | 5 | 0.016 | 0.007 | 3.1E-04 | 3.1E-04 | 0.01 | 3.1E-04 | 4.41 |
| GPX3 | GPX3 | MDR | OS | NC_045542.1 | 91669017 | 91679582 | 10564 | 22 | 2120 | 8 | 750 | 4 | 2 | 2 | 0.005 | 0.003 | 1.1E-04 | 1.1E-04 | 0.06 | 1.1E-04 | 0.53 |
| GPX4 | GPX4 | MDR | OS | NC_045541.1 | 17570782 | 175727532 | 19949 | 68 | 861 | 17 | 578 | 12 | 2 | 10 | 0.021 | 0.003 | 4.2E-04 | 4.2E-04 | 0.00 | 4.2E-04 | 3.17 |
| GPX7 | GPX7 | MDR | OS | NC_045545.1 | 33759070 | 33759381 | 8510 | 5 | 1266 | 3 | 750 | 3 | 3 | 0 | 0.004 | 0.004 | 4.1E-05 | 4.1E-05 | 0.00 | 4.1E-05 | -1.03 |
| GPX8 | GPX8 | MDR | OS | NC_045543.1 | 122551246 | 122559083 | 7836 | 7 | 1263 | 2 | 636 | 2 | 0 | 2 | 0.003 | 0.000 | 2.3E-04 | 2.3E-04 | 0.01 | 2.3E-04 | 1.26 |
| GSTA1 | GSTA1 | MDR | OS | NC_045544.1 | 4469255 | 4479935 | 10679 | 64 | 1022 | 18 | 660 | 9 | 2 | 7 | 0.014 | 0.003 | 2.5E-04 | 2.5E-04 | 0.01 | 2.5E-04 | 1.44 |
| GSTP1 | GSTP1 | MDR | OS | NC_045541.1 | 69278307 | 69290696 | 12388 | 36 | 1221 | 12 | 629 | 7 | 1 | 6 | 0.011 | 0.002 | 2.2E-04 | 2.2E-04 | 0.00 | 2.2E-04 | 1.26 |
| HAGHL | HAGHL | MDR | OS | NC_045545.1 | 17331572 | 17349013 | 17440 | 13 | 4359 | 7 | 838 | 5 | 2 | 3 | 0.006 | 0.002 | 4.9E-05 | 4.9E-05 | 0.00 | 4.9E-05 | 0.28 |
| JUN | JUN | MDR | OS | NC_045545.1 | 29151374 | 29154108 | 2733 | 1 | 2733 | 1 | 944 | 0 | 0 | 0 | 0.000 | 0.000 | 9.1E-05 | 9.1E-05 | 0.03 | 9.1E-05 | 0.55 |
| LOC116505938 | PLA2G4E | MDR | OS | NC_045541.1 | 116273315 | 116318685 | 45369 | 15 | 2810 | 3 | 2419 | 1 | 1 | 0 | 0.000 | 0.000 | 7.7E-05 | 7.7E-05 | 0.02 | 7.7E-05 | 0.70 |
| LOC116507099 | GSTA4 | MDR | OS | NC_045544.1 | 4439408 | 4450843 | 11434 | 58 | 980 | 19 | 663 | 13 | 5 | 8 | 0.020 | 0.008 | 2.7E-04 | 2.7E-04 | 0.01 | 2.7E-04 | 1.59 |
| LOC116507918 | GSTA4 | MDR | OS | NC_045544.1 | 4567425 | 4583606 | 16180 | 64 | 1576 | 21 | 687 | 8 | 2 | 6 | 0.012 | 0.003 | 1.4E-04 | 1.4E-04 | 0.01 | 1.4E-04 | 0.87 |
| LOC116508979 | GSTM1 | MDR | OS | NC_045545.1 | 45374295 | 45383509 | 9213 | 6 | 1226 | 3 | 655 | 2 | 2 | 0 | 0.003 | 0.003 | 2.9E-04 | 2.9E-04 | 0.00 | 2.9E-04 | 0.91 |
| LOC116514261 | CAT | MDR | OS | NC_045541.1 | 100771070 | 100816651 | 45880 | 31 | 5417 | 10 | 1568 | 5 | 2 | 3 | 0.003 | 0.001 | 4.5E-05 | 4.5E-05 | 0.01 | 4.5E-05 | 1.15 |
| LOC116514334 | GSTO1 | MDR | OS | NC_045550.1 | 11311766 | 11325877 | 14110 | 5 | 3166 | 3 | 768 | 2 | 0 | 2 | 0.003 | 0.000 | 4.7E-05 | 4.7E-05 | 0.02 | 4.7E-05 | -0.08 |
| LOC116516566 | GSTT1 | MDR | OS | NC_045553.1 | 18402410 | 18468697 | 6286 | 3 | 892 | 1 | 774 | 1 | 1 | 0 | 0.001 | 0.001 | 4.1E-04 | 4.1E-04 | 0.04 | 4.1E-04 | 1.27 |
| LOC116517285 | HAGH | MDR | OS | NC_045554.1 | 17328110 | 17381010 | 2828 | 8 | 2294 | 2 | 918 | 0 | 0 | 0 | 0.000 | 0.000 | 5.2E-05 | 5.2E-05 | 0.01 | 5.2E-05 | -0.27 |
| LOC116520753 | PON2 | MDR | OS | NC_045538.1 | 31868699 | 31882815 | 14115 | 8 | 1775 | 2 | 1062 | 1 | 0 | 1 | 0.001 | 0.000 | 1.0E-04 | 1.0E-04 | 0.02 | 1.0E-04 | 0.23 |
| LOC116522295 | MAP3K3 | MDR | OS | NC_045558.1 | 5847710 | 5941343 | 9632 | 60 | 4932 | 18 | 1844 | 14 | 3 | 11 | 0.008 | 0.002 | 4.1E-05 | 4.1E-05 | 0.01 | 4.1E-05 | 0.82 |
| MAPK11 | MAPK11 | MDR | OS | NC_045547.1 | 64415882 | 64460425 | 44542 | 23 | 3802 | 7 | 1074 | 3 | 1 | 2 | 0.003 | 0.001 | 4.7E-05 | 4.7E-05 | 0.00 | 4.7E-05 | 0.27 |
| MAPK13 | MAPK13 | MDR | OS | NC_045545.1 | 47114596 | 47145436 | 30839 | 8 | 1565 | 2 | 1086 | 0 | 0 | 0 | 0.000 | 0.000 | 2.4E-04 | 2.4E-04 | 0.02 | 2.4E-04 | 0.89 |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|--------------|--------------|----------|--------------|----------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| MAPK8 | MAPK8 | MDR | OS | NC_045551.1 | 71218361 | 71240583 | 2221 | 67 | 7177 | 14 | 1273 | 12 | 1 | 11 | 0.009 | 0.001 | 1.6E-05 | 1.6E-05 | 0.01 | 1.6E-05 | -0.45 |
| MAPK9 | MAPK9 | MDR | OS | NC_045542.1 | 8904944 | 89089094 | 3949 | 13 | 8623 | 0 | 1260 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| MAPKAPK5 | MAPKAPK5 | MDR | OS | NC_045553.1 | 1083663 | 1097507 | 13823 | 59 | 7118 | 21 | 1408 | 19 | 2 | 17 | 0.013 | 0.001 | 2.8E-05 | 2.8E-05 | 0.01 | 2.8E-05 | 1.30 |
| MAX | MAX | MDR | OS | NC_045541.1 | 66356537 | 66380027 | 23489 | 2 | 5303 | 1 | 523 | 0 | 0 | 0 | 0.000 | 0.000 | 2.2E-05 | 2.2E-05 | 0.00 | 2.2E-05 | -0.19 |
| MKNK1 | MKNK1 | MDR | OS | NC_045545.1 | 38273402 | 38309476 | 36073 | 2 | 16764 | 1 | 1388 | 0 | 0 | 0 | 0.000 | 0.000 | 1.1E-05 | 1.1E-05 | 0.00 | 1.1E-05 | 0.20 |
| MKNK2 | MKNK2 | MDR | OS | NC_045541.1 | 17063889 | 170690469 | 51629 | 36 | 19026 | 8 | 1385 | 4 | 0 | 4 | 0.003 | 0.000 | 1.1E-05 | 1.1E-05 | 0.00 | 1.1E-05 | 0.62 |
| MPO | MPO | MDR | OS | NC_045544.1 | 118634537 | 118669946 | 35408 | 85 | 3332 | 23 | 2138 | 21 | 8 | 13 | 0.010 | 0.004 | 4.0E-05 | 4.0E-05 | 0.04 | 4.0E-05 | -0.38 |
| MYC | MYC | MDR | OS | NC_045548.1 | 69065584 | 69071587 | 6002 | 29 | 4990 | 21 | 1288 | 15 | 2 | 13 | 0.012 | 0.002 | 5.5E-05 | 5.5E-05 | 0.01 | 5.5E-05 | 1.99 |
| PLA2G4A | PLA2G4A | MDR | OS | NC_045551.1 | 3606659 | 3684985 | 78325 | 174 | 2797 | 37 | 2205 | 28 | 6 | 22 | 0.013 | 0.003 | 1.1E-04 | 1.1E-04 | 0.01 | 1.1E-04 | 2.71 |
| PRDX1 | PRDX1 | MDR | OS | NC_045545.1 | 41310774 | 41318938 | 8163 | 2 | 4212 | 0 | 595 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| PRDX2 | PRDX2 | MDR | OS | NC_045542.1 | 67330740 | 67344583 | 13842 | 9 | 1753 | 2 | 592 | 1 | 1 | 0 | 0.002 | 0.002 | 2.5E-05 | 2.5E-05 | 0.01 | 2.5E-05 | -0.85 |
| PRDX3 | PRDX3 | MDR | OS | NC_045550.1 | 10982922 | 10986385 | 13462 | 61 | 1260 | 16 | 863 | 5 | 1 | 4 | 0.006 | 0.001 | 1.7E-04 | 1.7E-04 | 0.01 | 1.7E-04 | 0.78 |
| PRDX4 | PRDX4 | MDR | OS | NC_045551.1 | 17708462 | 17716753 | 8290 | 3 | 990 | 0 | 857 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| PRDX5 | PRDX5 | MDR | OS | NW_022473598.1 | 5363 | 17244 | 11880 | 45 | 1235 | 14 | 633 | 8 | 0 | 8 | 0.013 | 0.000 | 3.1E-04 | 3.1E-04 | 0.02 | 3.1E-04 | 3.65 |
| PRDX6 | PRDX6 | MDR | OS | NC_045551.1 | 11997739 | 12015594 | 17854 | 43 | 1217 | 14 | 661 | 7 | 1 | 6 | 0.011 | 0.002 | 1.4E-04 | 1.4E-04 | 0.01 | 1.4E-04 | -1.15 |
| PXDN | PXDN | MDR | OS | NC_045543.1 | 718410 | 798190 | 79779 | 230 | 17845 | 68 | 4420 | 64 | 7 | 57 | 0.014 | 0.002 | 1.5E-05 | 1.5E-05 | 0.02 | 1.5E-05 | 2.16 |
| PXMP2 | PXMP2 | MDR | OS | NC_045553.1 | 14731491 | 14742612 | 11120 | 5 | 1073 | 1 | 565 | 0 | 0 | 0 | 0.000 | 0.000 | 3.2E-05 | 3.2E-05 | -0.02 | 3.2E-05 | -0.69 |
| PYROXD1 | PYROXD1 | MDR | OS | NC_045547.1 | 1038902 | 10397732 | 16829 | 129 | 1998 | 44 | 1481 | 35 | 8 | 27 | 0.024 | 0.005 | 1.5E-04 | 1.5E-04 | 0.00 | 1.5E-04 | 2.54 |
| PYROXD2 | PYROXD2 | MDR | OS | NC_045550.1 | 25919165 | 25931512 | 12346 | 29 | 2188 | 7 | 1733 | 7 | 4 | 3 | 0.004 | 0.002 | 1.1E-04 | 1.1E-04 | 0.07 | 1.1E-04 | 0.89 |
| SOD1 | SOD1 | MDR | OS | NC_045546.1 | 50805425 | 50813527 | 8101 | 4 | 697 | 0 | 487 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| SOD2 | SOD2 | MDR | OS | NC_045544.1 | 5864218 | 58644474 | 10255 | 15 | 2780 | 4 | 673 | 0 | 0 | 0 | 0.000 | 0.000 | 1.4E-04 | 1.4E-04 | 0.04 | 1.4E-04 | 2.20 |
| CASP3 | CASP3 | MDR | P53 | NC_045549.1 | 48814636 | 48829477 | 14840 | 0 | 10913 | 0 | 896 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| CASP9 | CASP9 | MDR | P53 | NC_045555.1 | 188223 | 191546 | 3322 | 0 | 1628 | 0 | 1200 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| CDK1 | CDK1 | MDR | P53 | NC_045555.1 | 28553926 | 28564304 | 10377 | 0 | 2392 | 0 | 905 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| CDKN1A | CDKN1A | MDR | P53 | NC_045545.1 | 47896244 | 47905926 | 9681 | 0 | 10419 | 0 | 573 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| CHEK1 | CHEK1 | MDR | P53 | NC_045553.1 | 37419384 | 37433319 | 13934 | 76 | 1834 | 21 | 1415 | 20 | 3 | 17 | 0.014 | 0.002 | 1.6E-04 | 1.6E-04 | 0.01 | 1.6E-04 | 2.53 |
| CHEK2 | CHEK2 | MDR | P53 | NC_045553.1 | 31101 | 38024 | 6922 | 0 | 4885 | 0 | 1540 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| DDR2 | DDR2 | MDR | P53 | NC_045541.1 | 109679738 | 109721802 | 42063 | 0 | 18044 | 0 | 1502 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| E2F6 | E2F6 | MDR | P53 | NC_045543.1 | 6310249 | 6319756 | 9506 | 0 | 4274 | 0 | 611 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| GABARAPL1 | GABARAPL1 | MDR | P53 | NC_045542.1 | 24922 | 294605 | 46682 | 0 | 2282 | 0 | 350 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| GADD45A | GADD45A | MDR | P53 | NC_045545.1 | 2511096 | 25115039 | 3942 | 0 | 1388 | 0 | 476 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| HDAC3 | HDAC3 | MDR | P53 | NC_045546.1 | 6797873 | 68016832 | 38278 | 0 | 2357 | 0 | 1272 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116507365 | PERP | MDR | P53 | NC_045544.1 | 46461409 | 46471643 | 10233 | 0 | 2445 | 0 | 567 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116523000 | BAX | MDR | P53 | NC_045558.1 | 116899066 | 116918708 | 19641 | 0 | 1745 | 0 | 612 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116523426 | TFEB | MDR | P53 | NA | NA | NA | 14320 | 0 | 4352 | 0 | 1263 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| MDM2 | MDM2 | MDR | P53 | NC_045547.1 | 48027451 | 48042974 | 15322 | 0 | 7582 | 0 | 1376 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| MTA2 | MTA2 | MDR | P53 | NC_045557.1 | 6330996 | 6355765 | 24768 | 0 | 2956 | 0 | 1995 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|--------------|--------------|----------|--------------|----------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| PPM1D | PPM1D | MDR | p53 | NC_045544.1 | 13721190 | 13724852 | 37361 | 0 | 3093 | 0 | 1713 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| RRM2B | RRM2B | MDR | p53 | NC_045548.1 | 56812986 | 56828645 | 15658 | 0 | 4019 | 0 | 1062 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| TPM1 | TPM1 | MDR | p53 | NC_045556.1 | 671834 | 692620 | 20785 | 0 | 36314 | 0 | 1080 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| WIP1 | WIP1 | MDR | p53 | NW_022472647.1 | 15785 | 42135 | 26349 | 0 | 3787 | 0 | 1302 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| WRN | WRN | MDR | p53 | NC_045543.1 | 44805637 | 44878660 | 73022 | 0 | 18135 | 0 | 3886 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| GADD45B | GADD45B | MDR | | NC_045541.1 | 186465076 | 186466768 | 1691 | 11 | 1316 | 9 | 476 | 1 | 1 | 0 | 0.002 | 0.002 | 1.3E-04 | 1.3E-04 | 0.04 | 1.3E-04 | -0.17 |
| GADD45G | GADD45G | MDR | | NC_045543.1 | 88419595 | 88421971 | 2375 | 10 | 1370 | 4 | 512 | 2 | 0 | 2 | 0.004 | 0.000 | 3.2E-05 | 3.2E-05 | 0.01 | 3.2E-05 | -1.08 |
| MSRA | MSRA | MDR | | NC_045544.1 | 111416456 | 111583387 | 168930 | 5 | 14382 | 3 | 942 | 1 | 1 | 0 | 0.001 | 0.001 | 1.2E-05 | 1.2E-05 | 0.01 | 1.2E-05 | 1.02 |
| MSRB3 | MSRB3 | MDR | | NC_045547.1 | 49594196 | 49661938 | 67741 | 9 | 7007 | 4 | 549 | 1 | 0 | 1 | 0.002 | 0.000 | 3.0E-05 | 3.0E-05 | 0.01 | 3.0E-05 | 0.61 |
| AACS | AACS | MF | M | NC_045553.1 | 14206605 | 14256726 | 5020 | 11 | 4548 | 1 | 1986 | 1 | 0 | 1 | 0.001 | 0.000 | 9.6E-05 | 9.6E-05 | 0.04 | 9.6E-05 | 1.65 |
| DECR1 | DECR1 | MF | M | NC_045548.1 | 51973102 | 51989892 | 16789 | 42 | 1124 | 2 | 1001 | 2 | 0 | 2 | 0.002 | 0.000 | 4.4E-04 | 4.4E-04 | 0.00 | 4.4E-04 | 2.81 |
| LEPR | LEPR | MF | M | NC_045545.1 | 25982819 | 26029188 | 46368 | 0 | 3531 | 0 | 3531 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116502538 | CYP4F22 | MF | M | NC_045542.1 | 861851 | 894981 | 33129 | 144 | 1987 | 26 | 1572 | 24 | 15 | 9 | 0.015 | 0.010 | 1.1E-04 | 1.1E-04 | 0.02 | 1.1E-04 | 0.90 |
| LOC116505343 | FADS1 | MF | M | NC_045541.1 | 71961730 | 71975426 | 13695 | 0 | 908 | 0 | 908 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116507994 | CYB5R4 | MF | M | NC_045544.1 | 18380605 | 18420517 | 40011 | 31 | 6380 | 9 | 1556 | 3 | 0 | 3 | 0.000 | 0.000 | 3.1E-05 | 3.1E-05 | 0.00 | 3.1E-05 | 0.45 |
| LOC116509327 | CYP4B1 | MF | M | NC_045545.1 | 38102387 | 38127812 | 25242 | 7 | 7404 | 3 | 1560 | 0 | 0 | 0 | 0.000 | 0.000 | 3.5E-05 | 3.5E-05 | 0.00 | 3.5E-05 | 1.01 |
| LOC116513013 | CYB5R2 | MF | M | NC_045541.1 | 89520080 | 89533769 | 13688 | 73 | 3030 | 18 | 909 | 2 | 1 | 1 | 0.002 | 0.001 | 6.1E-05 | 6.1E-05 | 0.01 | 6.1E-05 | 0.58 |
| LOC116515748 | CYP2G1 | MF | M | NC_045552.1 | 419901 | 425513 | 5611 | 70 | 6558 | 23 | 1476 | 23 | 6 | 17 | 0.016 | 0.004 | 2.5E-05 | 2.5E-05 | 0.02 | 2.5E-05 | 0.06 |
| LOC116515749 | CYP2A13 | MF | M | NC_045552.1 | 402549 | 418388 | 15838 | 178 | 2105 | 39 | 1469 | 30 | 16 | 14 | 0.020 | 0.011 | 7.9E-05 | 7.9E-05 | 0.01 | 7.9E-05 | -0.26 |
| LOC116517450 | CYP2W1 | MF | M | NC_045544.1 | 1803700 | 1815654 | 11953 | 103 | 4557 | 42 | 1491 | 23 | 9 | 14 | 0.015 | 0.006 | 7.0E-05 | 7.0E-05 | 0.02 | 7.0E-05 | 3.15 |
| LOC116519097 | CYP1A1 | MF | M | NC_045556.1 | 5272984 | 5283825 | 10840 | 53 | 2211 | 24 | 1545 | 15 | 6 | 9 | 0.010 | 0.004 | 1.1E-04 | 1.1E-04 | 0.02 | 1.1E-04 | 1.16 |
| GAPDH | GAPDH | MF | M | NC_045543.1 | 42145657 | 42151782 | 6124 | 21 | 1232 | 13 | 991 | 9 | 3 | 6 | 0.009 | 0.003 | 1.0E-04 | 1.0E-04 | 0.01 | 1.0E-04 | -0.30 |
| LOC116509241 | LBP | MF | M | NC_045545.1 | 99166491 | 99185664 | 19172 | 68 | 852 | 17 | 662 | 15 | 6 | 9 | 0.023 | 0.009 | 2.2E-04 | 2.2E-04 | 0.00 | 2.2E-04 | 0.54 |
| NUAK1 | NUAK1 | MF | M | NC_045547.1 | 25247867 | 25309079 | 61211 | 28 | 4735 | 6 | 2069 | 6 | 1 | 5 | 0.003 | 0.000 | 4.0E-05 | 4.0E-05 | 0.02 | 4.0E-05 | 0.50 |
| NUAK2 | NUAK2 | MF | M | NC_045545.1 | 54090049 | 54111232 | 21182 | 25 | 4165 | 5 | 1931 | 3 | 3 | 0 | 0.002 | 0.002 | 3.5E-05 | 3.5E-05 | 0.00 | 3.5E-05 | 0.00 |
| CCN2 | CCN2 | MF | NS | NC_045544.1 | 42953499 | 42959286 | 5786 | 6 | 3546 | 4 | 1033 | 2 | 0 | 2 | 0.002 | 0.000 | 8.2E-05 | 8.2E-05 | 0.00 | 8.2E-05 | 0.87 |
| FOS | FOS | MF | NS | NC_045541.1 | 13147755 | 131480876 | 3340 | 21 | 2238 | 10 | 1100 | 8 | 3 | 5 | 0.007 | 0.003 | 1.4E-04 | 1.4E-04 | 0.06 | 1.4E-04 | 2.11 |
| FOXA1 | FOXA1 | MF | NS | NC_045541.1 | 129714464 | 129720820 | 6355 | 2 | 6697 | 2 | 1330 | 2 | 1 | 1 | 0.002 | 0.001 | 3.8E-05 | 3.8E-05 | 0.01 | 3.8E-05 | 0.54 |
| FOXA3 | FOXA3 | MF | NS | NC_045552.1 | 18830356 | 1906575 | 23518 | 9 | 2495 | 7 | 1258 | 6 | 2 | 4 | 0.005 | 0.002 | 3.4E-05 | 3.4E-05 | 0.01 | 3.4E-05 | -0.67 |
| FOXF1 | FOXF1 | MF | NS | NW_022472632.1 | 50839 | 53368 | 2528 | 26 | 1666 | 27 | 1300 | 27 | 4 | 23 | 0.021 | 0.003 | 2.1E-04 | 2.1E-04 | 0.00 | 2.1E-04 | 3.14 |
| FOXP1 | FOXP1 | MF | NS | NC_045542.1 | 55147783 | 55158438 | 16654 | 5 | 3996 | 3 | 1339 | 3 | 1 | 2 | 0.002 | 0.001 | 3.4E-05 | 3.4E-05 | 0.04 | 3.4E-05 | -0.21 |
| FOXP2 | FOXP2 | MF | NS | NC_045549.1 | 77428351 | 77445081 | 16729 | 29 | 15535 | 14 | 1672 | 11 | 2 | 9 | 0.007 | 0.001 | 1.3E-05 | 1.3E-05 | 0.00 | 1.3E-05 | 0.88 |
| FOXP2 | FOXP2 | MF | NS | NC_045544.1 | 10176038 | 101730370 | 44331 | 5 | 34492 | 1 | 1318 | 1 | 1 | 0 | 0.001 | 0.001 | 3.4E-06 | 3.4E-06 | 0.00 | 3.4E-06 | -0.20 |
| FOXP3 | FOXP3 | MF | NS | NC_045541.1 | 139123513 | 13901006 | 177492 | 13 | 26595 | 4 | 1482 | 3 | 1 | 2 | 0.002 | 0.001 | 1.2E-05 | 1.2E-05 | 0.03 | 1.2E-05 | 1.71 |
| FOXO1 | FOXO1 | MF | NS | NC_045543.1 | 16454312 | 16501192 | 46879 | 21 | 5133 | 18 | 1909 | 8 | 2 | 6 | 0.004 | 0.001 | 4.9E-05 | 4.9E-05 | 0.00 | 4.9E-05 | 1.46 |
| FOXO3 | FOXO3 | MF | NS | NC_045544.1 | 30931338 | 31059005 | 127666 | 7 | 6216 | 5 | 1963 | 0 | 0 | 0 | 0.000 | 0.000 | 5.7E-05 | 5.7E-05 | 0.10 | 5.7E-05 | 2.45 |
| FOXO4 | FOXO4 | MF | NS | NC_045552.1 | 55418945 | 55432196 | 13250 | 12 | 9547 | 6 | 1648 | 1 | 0 | 1 | 0.001 | 0.000 | 2.2E-05 | 2.2E-05 | 0.02 | 2.2E-05 | 0.32 |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|---------|--------------|----------|--------------|----------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| FOXP4 | FOXP4 | MF | NS | NC_045545.1 | 50767990 | 50938364 | 176673 | 24 | 7861 | 10 | 2000 | 0 | 0 | 0 | 0.000 | 0.000 | 1.3E-06 | 1.3E-06 | 0.00 | 1.3E-06 | -0.90 |
| CSK3A | CSK3A | MF | NS | NC_045552.1 | 5330886 | 5357461 | 26574 | 52 | 5763 | 3 | 1375 | 3 | 0 | 3 | 0.002 | 0.000 | 2.3E-05 | 2.3E-05 | 0.02 | 2.3E-05 | -0.13 |
| CSK3B | CSK3B | MF | NS | NC_045543.1 | 38025983 | 38157869 | 131885 | 17 | 47273 | 8 | 1414 | 0 | 0 | 0 | 0.000 | 0.000 | 6.7E-06 | 6.7E-06 | 0.01 | 6.7E-06 | 2.24 |
| IRS4 | IRS4 | MF | NS | NC_045552.1 | 40406058 | 40423168 | 17109 | 8 | 7081 | 8 | 3574 | 2 | 2 | 2 | 0.001 | 0.001 | 2.6E-05 | 2.6E-05 | 0.07 | 2.6E-05 | -0.37 |
| KL | KL | MF | NS | NC_045546.1 | 25030481 | 25050118 | 19636 | 19 | 7630 | 14 | 3043 | 10 | 6 | 4 | 0.003 | 0.002 | 3.1E-05 | 3.1E-05 | 0.00 | 3.1E-05 | 1.14 |
| MAP2K1 | MAP2K1 | MF | NS | NC_045556.1 | 22743235 | 22795477 | 52341 | 33 | 2626 | 18 | 1177 | 5 | 0 | 5 | 0.004 | 0.000 | 5.7E-05 | 5.7E-05 | 0.02 | 5.7E-05 | -0.18 |
| MAP2K3 | MAP2K3 | MF | NS | NC_045554.1 | 4717938 | 4778731 | 60792 | 69 | 7826 | 33 | 1034 | 17 | 0 | 17 | 0.016 | 0.000 | 3.4E-05 | 3.4E-05 | 0.00 | 3.4E-05 | 2.04 |
| MAP3K1 | MAP3K1 | MF | NS | NC_045543.1 | 121693300 | 121770410 | 77109 | 35 | 7006 | 12 | 4597 | 7 | 4 | 3 | 0.002 | 0.001 | 1.8E-05 | 1.8E-05 | 0.03 | 1.8E-05 | -0.40 |
| MDM4 | MDM4 | MF | NS | NC_045545.1 | 52978167 | 53010554 | 32386 | 10 | 9787 | 4 | 1430 | 4 | 1 | 3 | 0.003 | 0.001 | 1.4E-05 | 1.4E-05 | 0.00 | 1.4E-05 | 0.08 |
| PDK2 | PDK2 | MF | NS | NC_045558.1 | 104306189 | 104345243 | 39053 | 11 | 16248 | 3 | 1245 | 0 | 0 | 0 | 0.000 | 0.000 | 7.7E-06 | 7.7E-06 | 0.00 | 7.7E-06 | 0.39 |
| PDK3 | PDK3 | MF | NS | NC_045551.1 | 17250731 | 17348708 | 97976 | 19 | 5548 | 8 | 1210 | 1 | 0 | 1 | 0.001 | 0.000 | 5.2E-05 | 5.2E-05 | 0.02 | 5.2E-05 | 1.58 |
| PDK4 | PDK4 | MF | NS | NC_045558.1 | 31955467 | 31968343 | 12875 | 3 | 2675 | 1 | 1273 | 1 | 1 | 0 | 0.001 | 0.001 | 5.6E-05 | 5.6E-05 | 0.04 | 5.6E-05 | -0.03 |
| PPP1K1 | PPP1K1 | MF | NS | NC_045554.1 | 21275095 | 21313242 | 38146 | 11 | 4384 | 4 | 1807 | 3 | 2 | 1 | 0.002 | 0.001 | 3.4E-05 | 3.4E-05 | 0.01 | 3.4E-05 | 0.05 |
| PIK3R5 | PIK3R5 | MF | NS | NC_045542.1 | 62912035 | 62969537 | 57301 | 31 | 8227 | 6 | 2628 | 5 | 4 | 1 | 0.002 | 0.002 | 3.1E-05 | 3.1E-05 | 0.00 | 3.1E-05 | 1.33 |
| PIK3R6 | PIK3R6 | MF | NS | NC_045542.1 | 62975883 | 63009270 | 33386 | 5 | 2723 | 3 | 222 | 3 | 2 | 1 | 0.001 | 0.001 | 1.3E-04 | 1.3E-04 | 0.00 | 1.3E-04 | 1.13 |
| PTEN | PTEN | MF | NS | NC_045555.1 | 42970557 | 43035925 | 65367 | 35 | 4892 | 8 | 1206 | 6 | 0 | 6 | 0.005 | 0.000 | 2.9E-05 | 2.9E-05 | 0.00 | 2.9E-05 | 0.35 |
| RAF1 | RAF1 | MF | NS | NC_045542.1 | 109442707 | 109500604 | 57896 | 30 | 14044 | 4 | 1981 | 3 | 1 | 2 | 0.002 | 0.001 | 1.1E-05 | 1.1E-05 | 0.03 | 1.1E-05 | 0.03 |
| RASA1 | RASA1 | MF | NS | NC_045543.1 | 109203790 | 10925262 | 48771 | 43 | 3780 | 5 | 3020 | 4 | 2 | 2 | 0.001 | 0.001 | 3.9E-05 | 3.9E-05 | 0.02 | 3.9E-05 | -0.67 |
| RPS6KA1 | RPS6KA1 | MF | NS | NC_045552.1 | 18513405 | 18634783 | 121377 | 48 | 16630 | 7 | 2245 | 7 | 2 | 5 | 0.003 | 0.001 | 1.8E-05 | 1.8E-05 | 0.03 | 1.8E-05 | 1.06 |
| RPS6KA3 | RPS6KA3 | MF | NS | NC_045551.1 | 19153938 | 19233066 | 79127 | 21 | 13758 | 2 | 2210 | 2 | 1 | 1 | 0.001 | 0.000 | 1.7E-05 | 1.7E-05 | 0.01 | 1.7E-05 | 0.52 |
| RPS6KA4 | RPS6KA4 | MF | NS | NW_022472598.1 | 63562 | 92977 | 29414 | 89 | 2988 | 22 | 2327 | 22 | 3 | 19 | 0.009 | 0.001 | 7.0E-05 | 7.0E-05 | 0.00 | 7.0E-05 | 1.34 |
| RPS6KA5 | RPS6KA5 | MF | NS | NC_045541.1 | 139960942 | 139990614 | 29671 | 12 | 3866 | 6 | 2350 | 4 | 2 | 2 | 0.002 | 0.001 | 3.2E-05 | 3.2E-05 | 0.00 | 3.2E-05 | -0.46 |
| RPS6KA6 | RPS6KA6 | MF | NS | NC_045552.1 | 44489613 | 4456784 | 67170 | 59 | 6243 | 10 | 2339 | 8 | 1 | 7 | 0.003 | 0.000 | 3.4E-05 | 3.4E-05 | 0.00 | 3.4E-05 | 0.97 |
| RPS6KB1 | RPS6KB1 | MF | NS | NC_045544.1 | 135617126 | 135650931 | 33804 | 44 | 4951 | 12 | 1494 | 9 | 1 | 8 | 0.006 | 0.001 | 3.8E-05 | 3.8E-05 | 0.01 | 3.8E-05 | 0.62 |
| RPS6KB2 | RPS6KB2 | MF | NS | NC_045541.1 | 6669203 | 66690977 | 11773 | 14 | 2179 | 3 | 1413 | 1 | 0 | 1 | 0.001 | 0.000 | 6.2E-05 | 6.2E-05 | 0.01 | 6.2E-05 | 0.05 |
| TSC1 | TSC1 | MF | NS | NC_045556.1 | 33964652 | 34006230 | 41577 | 96 | 34084 | 36 | 3453 | 34 | 13 | 21 | 0.010 | 0.004 | 9.6E-06 | 9.6E-06 | 0.00 | 9.6E-06 | 3.06 |
| TSC2 | TSC2 | MF | NS | NC_045554.1 | 21430962 | 21497065 | 66102 | 48 | 6107 | 15 | 5317 | 14 | 8 | 6 | 0.003 | 0.002 | 2.2E-05 | 2.2E-05 | 0.01 | 2.2E-05 | -0.60 |
| MTOR | MTOR | MF | NS | NC_045555.1 | 7204873 | 7287667 | 82793 | 352 | 16274 | 64 | 7550 | 51 | 2 | 49 | 0.007 | 0.000 | 1.9E-05 | 1.9E-05 | 0.00 | 1.9E-05 | 3.17 |
| NRFI | NRFI | MF | NS | NA | NA | NA | 35658 | 28 | 1155 | 8 | 954 | 8 | 0 | 8 | 0.008 | 0.000 | NA | NA | NA | NA | 0.10 |
| RHEB | RHEB | MF | NS | NC_045558.1 | 8980554 | 9011531 | 30976 | 5 | 2290 | 1 | 550 | 1 | 0 | 1 | 0.002 | 0.000 | 3.1E-05 | 3.1E-05 | 0.00 | 3.1E-05 | -0.46 |
| RPTOR | RPTOR | MF | NS | NC_045542.1 | 46106034 | 46543388 | 438353 | 63 | 13123 | 13 | 4021 | 8 | 0 | 8 | 0.002 | 0.000 | 2.3E-05 | 2.3E-05 | 0.00 | 2.3E-05 | 1.86 |
| SIRT1 | SIRT1 | MF | NS | NC_045555.1 | 21493444 | 21519272 | 19227 | 16 | 14874 | 5 | 2439 | 5 | 1 | 4 | 0.002 | 0.000 | 1.1E-05 | 1.1E-05 | 0.02 | 1.1E-05 | 0.12 |
| SIRT2 | SIRT2 | MF | NS | NC_045552.1 | 2108487 | 2117076 | 12288 | 86 | 1633 | 29 | 1160 | 17 | 3 | 14 | 0.015 | 0.003 | 1.7E-04 | 1.7E-04 | 0.01 | 1.7E-04 | 2.32 |
| SIRT3 | SIRT3 | MF | NS | NC_045541.1 | 72495530 | 72509176 | 13645 | 11 | 18257 | 4 | 1226 | 2 | 1 | 1 | 0.002 | 0.001 | 6.1E-06 | 6.1E-06 | 0.05 | 6.1E-06 | -0.42 |
| SIRT4 | SIRT4 | MF | NS | NC_045553.1 | 16207863 | 16219377 | 11513 | 6 | 3173 | 1 | 951 | 1 | 1 | 0 | 0.001 | 0.001 | 3.3E-05 | 3.3E-05 | 0.00 | 3.3E-05 | -0.27 |
| SIRT5 | SIRT5 | MF | NS | NC_045548.1 | 10929095 | 10945684 | 14588 | 53 | 4685 | 24 | 919 | 9 | 2 | 7 | 0.010 | 0.002 | 2.2E-05 | 2.2E-05 | 0.00 | 2.2E-05 | -0.76 |
| SIRT6 | SIRT6 | MF | NS | NC_045541.1 | 180393527 | 180418543 | 25015 | 81 | 1529 | 32 | 1042 | 22 | 5 | 17 | 0.021 | 0.005 | 1.1E-04 | 1.1E-04 | 0.01 | 1.1E-04 | 0.38 |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|-------------|--------------|----------|--------------|----------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| SIRT7 | SIRT7 | MF | NS | NC_045542.1 | 4914042 | 49166303 | 25660 | 11 | 2623 | 5 | 1199 | 3 | 3 | 0 | 0.003 | 0.003 | 3.7E-05 | 3.7E-05 | 0.01 | 3.7E-05 | -0.55 |
| CCND1 | CCND1 | MF | NS | NC_045541.1 | 99316894 | 99339972 | 23077 | 12 | 4230 | 7 | 874 | 5 | 2 | 3 | 0.006 | 0.002 | 5.1E-05 | 5.1E-05 | 0.02 | 5.1E-05 | 0.45 |
| INPP5A | INPP5A | MF | NS | NC_045555.1 | 48753482 | 48986459 | 232976 | 93 | 3890 | 26 | 1216 | 18 | 7 | 11 | 0.015 | 0.006 | 6.2E-05 | 6.2E-05 | 0.00 | 6.2E-05 | 1.40 |
| JAK1 | JAK1 | MF | NS | NC_045545.1 | 26273465 | 26338926 | 65460 | 95 | 4864 | 16 | 3429 | 12 | 5 | 7 | 0.003 | 0.001 | 6.9E-05 | 6.9E-05 | 0.00 | 6.9E-05 | 2.42 |
| JAK2 | JAK2 | MF | NS | NC_045543.1 | 75033530 | 75109628 | 76097 | 36 | 14718 | 17 | 3370 | 15 | 12 | 3 | 0.004 | 0.004 | 1.1E-05 | 1.1E-05 | 0.04 | 1.1E-05 | 0.17 |
| KRAS | KRAS | MF | NS | NC_045547.1 | 8018869 | 8053229 | 34359 | 71 | 3183 | 22 | 678 | 9 | 1 | 8 | 0.013 | 0.001 | 9.9E-05 | 9.9E-05 | 0.02 | 9.9E-05 | 2.54 |
| MRAS | MRAS | MF | NS | NC_045541.1 | 56654969 | 56741032 | 86062 | 5 | 2133 | 3 | 619 | 0 | 0 | 0 | 0.000 | 0.000 | 1.0E-04 | 1.0E-04 | 0.02 | 1.0E-04 | 0.68 |
| NRAS | NRAS | MF | NS | NC_045545.1 | 61950623 | 61960076 | 9452 | 21 | 7767 | 4 | 643 | 2 | 1 | 1 | 0.003 | 0.002 | 4.5E-05 | 4.5E-05 | 0.07 | 4.5E-05 | 2.23 |
| PIK3C2A | PIK3C2A | MF | NS | NC_045541.1 | 83259130 | 83316287 | 57156 | 239 | 8935 | 55 | 5108 | 50 | 14 | 36 | 0.010 | 0.003 | 2.0E-05 | 2.0E-05 | 0.01 | 2.0E-05 | 0.65 |
| RIT1 | RIT1 | MF | NS | NC_045557.1 | 5408090 | 5416034 | 7943 | 15 | 3455 | 5 | 661 | 5 | 0 | 5 | 0.008 | 0.000 | 1.2E-04 | 1.2E-04 | 0.03 | 1.2E-04 | 3.12 |
| YWHA8 | YWHA8 | MF | NS | NC_045545.1 | 75996115 | 75112826 | 16710 | 10 | 3013 | 3 | 730 | 2 | 0 | 2 | 0.003 | 0.000 | 5.9E-05 | 5.9E-05 | 0.01 | 5.9E-05 | 0.31 |
| YWHAE | YWHAE | MF | NS | NC_045544.1 | 142687821 | 142702733 | 14911 | 11 | 2323 | 8 | 762 | 1 | 0 | 1 | 0.001 | 0.000 | 1.5E-04 | 1.5E-04 | 0.02 | 1.5E-04 | 2.81 |
| YWHA9 | YWHA9 | MF | NS | NC_045544.1 | 121763984 | 121783914 | 19929 | 9 | 3861 | 5 | 742 | 1 | 0 | 1 | 0.001 | 0.000 | 6.4E-05 | 6.4E-05 | 0.00 | 6.4E-05 | 0.21 |
| YWHAQ | YWHAQ | MF | NS | NC_045543.1 | 4962625 | 4975020 | 12394 | 2 | 1441 | 0 | 730 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| GHR | GHR | MF | NS | NC_045543.1 | 126304208 | 126466410 | 162201 | 9 | 47290 | 3 | 1873 | 3 | 1 | 2 | 0.002 | 0.001 | 6.6E-06 | 6.6E-06 | 0.01 | 6.6E-06 | 1.62 |
| IGF1 | IGF1 | MF | NS | NC_045547.1 | 23829988 | 23903413 | 73424 | 8 | 9704 | 2 | 505 | 2 | 1 | 1 | 0.004 | 0.002 | 3.5E-06 | 3.5E-06 | 0.00 | 3.5E-06 | -0.93 |
| IGF1R | IGF1R | MF | NS | NC_045556.1 | 21438666 | 2166937 | 225270 | 49 | 5251 | 14 | 4084 | 12 | 4 | 8 | 0.003 | 0.001 | 2.9E-05 | 2.9E-05 | 0.00 | 2.9E-05 | 0.05 |
| IGF2 | IGF2 | MF | NS | NC_045541.1 | 80154142 | 80174323 | 20180 | 5 | 3870 | 5 | 647 | 4 | 1 | 3 | 0.006 | 0.002 | 7.1E-05 | 7.1E-05 | 0.06 | 7.1E-05 | 1.35 |
| IGF2R | IGF2R | MF | NS | NC_045544.1 | 58952645 | 59064872 | 82226 | 62 | 8934 | 12 | 7400 | 12 | 6 | 6 | 0.002 | 0.001 | 2.7E-05 | 2.7E-05 | 0.03 | 2.7E-05 | 0.81 |
| IGFBP1 | IGFBP1 | MF | NS | NC_045558.1 | 79525953 | 79536695 | 10741 | 8 | 1986 | 6 | 779 | 2 | 1 | 1 | 0.003 | 0.001 | 1.8E-04 | 1.8E-04 | 0.00 | 1.8E-04 | 1.85 |
| IGFBP2 | IGFBP2 | MF | NS | NC_045541.1 | 60907109 | 60939328 | 86818 | 7 | 3112 | 4 | 878 | 1 | 0 | 1 | 0.001 | 0.000 | 8.9E-05 | 8.9E-05 | 0.03 | 8.9E-05 | 0.69 |
| IGFBP3 | IGFBP3 | MF | NS | NC_045558.1 | 79580336 | 795884615 | 34278 | 1 | 2422 | 1 | 872 | 1 | 0 | 1 | 0.001 | 0.000 | 1.4E-04 | 1.4E-04 | 0.18 | 1.4E-04 | 0.94 |
| IGFBP4 | IGFBP4 | MF | NS | NC_045558.1 | 97177152 | 97207892 | 30739 | 3 | 4238 | 0 | 773 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| IGFBP7 | IGFBP7 | MF | NS | NC_045543.1 | 43720138 | 43746573 | 26434 | 8 | 1070 | 3 | 823 | 3 | 2 | 1 | 0.004 | 0.002 | 1.2E-04 | 1.2E-04 | 0.02 | 1.2E-04 | -0.56 |
| INSR | INSR | MF | NS | NC_045541.1 | 166100054 | 166184388 | 84333 | 71 | 7960 | 18 | 4053 | 17 | 4 | 13 | 0.004 | 0.001 | 3.7E-05 | 3.7E-05 | 0.02 | 3.7E-05 | 2.30 |
| LOC16599867 | NDUFC2 | MF | OP | NC_045546.1 | 3769293 | 3772841 | 3547 | 32 | 541 | 14 | 339 | 3 | 0 | 3 | 0.009 | 0.000 | 7.0E-04 | 7.0E-04 | 0.02 | 7.0E-04 | 3.74 |
| LOC16511712 | NDUFA9 | MF | OP | NC_045547.1 | 77416510 | 77446458 | 29947 | 0 | 1496 | 0 | 1063 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| NDUFA10 | NDUFA10 | MF | OP | NC_045553.1 | 19813620 | 19822847 | 9226 | 8 | 1178 | 2 | 829 | 2 | 0 | 2 | 0.002 | 0.000 | 4.0E-05 | 4.0E-05 | 0.00 | 4.0E-05 | -0.82 |
| NDUFA11 | NDUFA11 | MF | OP | NC_045541.1 | 166300839 | 166307546 | 6706 | 8 | 750 | 3 | 407 | 1 | 1 | 0 | 0.002 | 0.002 | 2.5E-04 | 2.5E-04 | 0.00 | 2.5E-04 | 0.18 |
| NDUFA13 | NDUFA13 | MF | OP | NC_045541.1 | 17952171 | 179164618 | 12446 | 40 | 598 | 15 | 436 | 8 | 0 | 8 | 0.018 | 0.000 | 2.9E-04 | 2.9E-04 | 0.02 | 2.9E-04 | 0.33 |
| NDUFA2 | NDUFA2 | MF | OP | NC_045546.1 | 6978235 | 69800461 | 2207 | 6 | 439 | 2 | 303 | 2 | 1 | 1 | 0.007 | 0.003 | 6.6E-04 | 6.6E-04 | 0.00 | 6.6E-04 | 0.97 |
| NDUFA3 | NDUFA3 | MF | OP | NC_045558.1 | 120008654 | 120015146 | 6601 | 14 | 442 | 6 | 254 | 3 | 0 | 3 | 0.012 | 0.000 | 5.7E-04 | 5.7E-04 | 0.07 | 5.7E-04 | 0.97 |
| NDUFA4 | NDUFA4 | MF | OP | NC_045558.1 | 2278157 | 2283756 | 5598 | 13 | 576 | 5 | 317 | 3 | 0 | 3 | 0.009 | 0.000 | 3.8E-04 | 3.8E-04 | 0.01 | 3.8E-04 | 0.68 |
| NDUFA5 | NDUFA5 | MF | OP | NC_045547.1 | 61934189 | 61940369 | 6179 | 17 | 540 | 3 | 346 | 1 | 0 | 1 | 0.003 | 0.000 | 7.9E-04 | 7.9E-04 | 0.00 | 7.9E-04 | 2.69 |
| NDUFA7 | NDUFA7 | MF | OP | NW_022473609.1 | 29445 | 38040 | 8594 | 31 | 630 | 16 | 332 | 6 | 3 | 3 | 0.018 | 0.009 | 3.8E-04 | 3.8E-04 | 0.01 | 3.8E-04 | 1.26 |
| NDUFAF1 | NDUFAF1 | MF | OP | NC_045541.1 | 117627659 | 117689387 | 11727 | 7 | 1881 | 3 | 896 | 1 | 0 | 1 | 0.001 | 0.000 | 5.5E-05 | 5.5E-05 | 0.00 | 5.5E-05 | 0.17 |
| NDUFB1 | NDUFB1 | MF | OP | NC_045541.1 | 140828363 | 140832332 | 3968 | 2 | 654 | 2 | 172 | 0 | 0 | 0 | 0.000 | 0.000 | 4.7E-04 | 4.7E-04 | 0.01 | 4.7E-04 | 1.57 |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|--------------|--------------|----------|--------------|-------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| NDUFB10 | NDUFB10 | MF | OP | NC_045554.1 | 21852477 | 21855605 | 3127 | 5 | 746 | 2 | 521 | 0 | 0 | 0 | 0.000 | 0.000 | 1.1E-04 | 1.1E-04 | 0.01 | 1.1E-04 | -0.55 |
| NDUFB11 | NDUFB11 | MF | OP | NC_045542.1 | 12714830 | 12766025 | 5194 | 6 | 534 | 3 | 450 | 3 | 0 | 3 | 0.007 | 0.000 | 4.9E-04 | 4.9E-04 | 0.01 | 4.9E-04 | 1.15 |
| NDUFB2 | NDUFB2 | MF | OP | NC_045547.1 | 21874376 | 21878460 | 4083 | 6 | 658 | 2 | 312 | 0 | 0 | 0 | 0.000 | 0.000 | 5.4E-05 | 5.4E-05 | 0.00 | 5.4E-05 | -0.92 |
| NDUFB3 | NDUFB3 | MF | OP | NC_045541.1 | 8245864 | 8254340 | 8475 | 30 | 476 | 7 | 280 | 5 | 2 | 3 | 0.018 | 0.007 | 2.5E-04 | 2.5E-04 | 0.00 | 2.5E-04 | -0.46 |
| NDUFB4 | NDUFB4 | MF | OP | NC_045543.1 | 38276126 | 38282371 | 6244 | 18 | 2663 | 5 | 408 | 0 | 0 | 0 | 0.000 | 0.000 | 5.2E-05 | 5.2E-05 | 0.02 | 5.2E-05 | 0.01 |
| NDUFB5 | NDUFB5 | MF | OP | NC_045550.1 | 6081485 | 60489488 | 8002 | 26 | 827 | 7 | 564 | 4 | 1 | 3 | 0.007 | 0.002 | 2.3E-04 | 2.3E-04 | 0.00 | 2.3E-04 | 0.39 |
| NDUFB6 | NDUFB6 | MF | OP | NC_045543.1 | 60318894 | 60330100 | 11205 | 4 | 795 | 3 | 437 | 2 | 1 | 1 | 0.005 | 0.002 | 4.6E-04 | 4.6E-04 | 0.24 | 4.6E-04 | 1.40 |
| NDUFB7 | NDUFB7 | MF | OP | NC_045542.1 | 65285676 | 65291857 | 6180 | 5 | 639 | 4 | 378 | 0 | 0 | 0 | 0.000 | 0.000 | 2.3E-04 | 2.3E-04 | 0.00 | 2.3E-04 | 0.82 |
| NDUFB9 | NDUFB9 | MF | OP | NC_045548.1 | 67199710 | 67208013 | 8302 | 22 | 1480 | 11 | 521 | 6 | 3 | 3 | 0.012 | 0.006 | 9.3E-05 | 9.3E-05 | 0.01 | 9.3E-05 | -0.25 |
| NDUFS1 | NDUFS1 | MF | OP | NC_045541.1 | 4677652 | 4720243 | 42390 | 131 | 4968 | 35 | 2169 | 30 | 4 | 26 | 0.014 | 0.002 | 4.2E-05 | 4.2E-05 | 0.00 | 4.2E-05 | 1.03 |
| NDUFS2 | NDUFS2 | MF | OP | NC_045557.1 | 5463386 | 5480255 | 16868 | 154 | 1617 | 36 | 1398 | 30 | 3 | 27 | 0.021 | 0.002 | 1.3E-04 | 1.3E-04 | 0.00 | 1.3E-04 | 1.22 |
| NDUFS3 | NDUFS3 | MF | OP | NC_045541.1 | 110221834 | 110236057 | 14222 | 15 | 954 | 3 | 770 | 1 | 0 | 1 | 0.001 | 0.000 | 1.9E-04 | 1.9E-04 | 0.06 | 1.9E-04 | 0.38 |
| NDUFS4 | NDUFS4 | MF | OP | NC_045543.1 | 123242728 | 123282804 | 40075 | 12 | 707 | 2 | 538 | 1 | 1 | 0 | 0.002 | 0.002 | 2.0E-04 | 2.0E-04 | 0.00 | 2.0E-04 | -0.04 |
| NDUFS5 | NDUFS5 | MF | OP | NC_045552.1 | 30616987 | 30620349 | 3361 | 2 | 1168 | 0 | 334 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| NDUFS6 | NDUFS6 | MF | OP | NC_045546.1 | 77305359 | 77311071 | 5711 | 33 | 516 | 10 | 398 | 10 | 5 | 5 | 0.025 | 0.013 | 5.2E-04 | 5.2E-04 | 0.01 | 5.2E-04 | 1.57 |
| NDUFS8 | NDUFS8 | MF | OP | NC_045541.1 | 66450440 | 66465211 | 14770 | 9 | 2716 | 3 | 608 | 1 | 0 | 1 | 0.002 | 0.000 | 1.2E-04 | 1.2E-04 | 0.00 | 1.2E-04 | 0.97 |
| NDUFV1 | NDUFV1 | MF | OP | NC_045541.1 | 69294488 | 69299998 | 5509 | 75 | 1638 | 22 | 1388 | 14 | 5 | 9 | 0.010 | 0.004 | 2.2E-04 | 2.2E-04 | 0.00 | 2.2E-04 | 3.84 |
| NDUFV2 | NDUFV2 | MF | OP | NC_045548.1 | 25037255 | 25053578 | 16322 | 31 | 924 | 9 | 730 | 5 | 2 | 3 | 0.007 | 0.003 | 3.8E-04 | 3.8E-04 | 0.00 | 3.8E-04 | 2.91 |
| NDUFV3 | NDUFV3 | MF | OP | NC_045546.1 | 56321323 | 56333543 | 12219 | 2 | 2144 | 1 | 357 | 1 | 0 | 1 | 0.003 | 0.000 | 2.2E-04 | 2.2E-04 | 0.09 | 2.2E-04 | 1.71 |
| LOC116523446 | SDHC | MF | OP | NA | NA | NA | 8837 | 11 | 1539 | 3 | 504 | 3 | 0 | 3 | 0.006 | 0.000 | NA | NA | NA | NA | 1.00 |
| SDHA | SDHA | MF | OP | NC_045558.1 | 83290806 | 83338650 | 47843 | 4 | 2466 | 0 | 2001 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| SDHB | SDHB | MF | OP | NC_045555.1 | 7599488 | 7608503 | 9014 | 13 | 1086 | 4 | 853 | 4 | 0 | 4 | 0.005 | 0.000 | 2.5E-04 | 2.5E-04 | 0.00 | 2.5E-04 | 1.14 |
| SDHD | SDHD | MF | OP | NC_045553.1 | 48073716 | 48083134 | 9417 | 47 | 952 | 32 | 546 | 18 | 6 | 12 | 0.033 | 0.011 | 3.9E-04 | 3.9E-04 | 0.02 | 3.9E-04 | 3.83 |
| LOC116503820 | UQCRC1 | MF | OP | NC_045542.1 | 115920792 | 115943491 | 22698 | 62 | 1634 | 14 | 1436 | 11 | 5 | 6 | 0.008 | 0.003 | 1.5E-04 | 1.5E-04 | 0.01 | 1.5E-04 | 1.48 |
| LOC116509003 | UQCRC1 | MF | OP | NC_045545.1 | 38933362 | 38937238 | 3875 | 0 | 496 | 0 | 272 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116512557 | UQCRC1 | MF | OP | NC_045548.1 | 54446995 | 54454152 | 7156 | 20 | 1184 | 5 | 332 | 0 | 0 | 0 | 0.000 | 0.000 | 2.3E-04 | 2.3E-04 | 0.01 | 2.3E-04 | 1.33 |
| LOC116512786 | CYC1 | MF | OP | NC_045548.1 | 80866313 | 80869214 | 2900 | 36 | 1196 | 13 | 935 | 9 | 3 | 6 | 0.010 | 0.003 | 1.6E-04 | 1.6E-04 | 0.01 | 1.6E-04 | 0.40 |
| LOC116516714 | UQCRC1 | MF | OP | NC_045553.1 | 4097459 | 4101064 | 3604 | 0 | 632 | 0 | 423 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116517709 | UQCRC2 | MF | OP | NC_045554.1 | 6799444 | 6814582 | 15137 | 38 | 1685 | 7 | 1351 | 6 | 1 | 5 | 0.004 | 0.001 | 1.0E-04 | 1.0E-04 | 0.01 | 1.0E-04 | -0.07 |
| LOC116518020 | UQCRC1 | MF | OP | NC_045554.1 | 40121544 | 40128624 | 5079 | 10 | 1020 | 9 | 823 | 8 | 1 | 7 | 0.010 | 0.001 | 1.7E-04 | 1.7E-04 | 0.01 | 1.7E-04 | -0.03 |
| LOC116520028 | UQCRC1 | MF | OP | NC_045557.1 | 5480389 | 5484627 | 4237 | 0 | 491 | 0 | 240 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116520018 | COX11 | MF | OP | NC_045542.1 | 32845043 | 32850545 | 5491 | 16 | 1340 | 3 | 737 | 1 | 0 | 1 | 0.001 | 0.000 | 2.2E-04 | 2.2E-04 | 0.04 | 2.2E-04 | 1.21 |
| LOC116506575 | COX7C | MF | OP | NC_045543.1 | 109406721 | 10941773 | 11051 | 7 | 871 | 2 | 190 | 2 | 1 | 1 | 0.011 | 0.005 | 1.8E-04 | 1.8E-04 | 0.07 | 1.8E-04 | 0.03 |
| LOC116506766 | COX17 | MF | OP | NC_045543.1 | 37988454 | 37994728 | 6273 | 12 | 553 | 3 | 208 | 1 | 0 | 1 | 0.005 | 0.000 | 8.8E-04 | 8.8E-04 | 0.02 | 8.8E-04 | 3.13 |
| LOC116506863 | COX7A2 | MF | OP | NC_045544.1 | 14911935 | 14915577 | 3641 | 1 | 593 | 1 | 224 | 0 | 0 | 0 | 0.000 | 0.000 | 2.3E-04 | 2.3E-04 | 0.02 | 2.3E-04 | -0.07 |
| LOC116507992 | COX7A2L | MF | OP | NC_045544.1 | 94082703 | 94090734 | 8030 | 16 | 1029 | 6 | 345 | 2 | 2 | 0 | 0.006 | 0.006 | 1.9E-04 | 1.9E-04 | 0.00 | 1.9E-04 | 0.58 |
| LOC116512326 | COX6C | MF | OP | NC_045548.1 | 55935588 | 55939528 | 3939 | 3 | 496 | 0 | 229 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|--------------|--------------|----------|--------------|-------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| LOC116514227 | COX15 | MF | OP | NC_045550.1 | 26298153 | 26294530 | 6476 | 16 | 1770 | 5 | 1215 | 5 | 5 | 0 | 0.004 | 0.004 | 4.4E-05 | 4.4E-05 | 0.02 | 4.4E-05 | -0.99 |
| LOC116516211 | COX6B1 | MF | OP | NC_045552.1 | 9988938 | 9993328 | 4369 | 24 | 553 | 6 | 258 | 1 | 0 | 1 | 0.004 | 0.000 | 5.6E-04 | 5.6E-04 | 0.01 | 5.6E-04 | 1.59 |
| LOC116516217 | COX7B | MF | OP | NC_045552.1 | 37548307 | 37550835 | 2527 | 7 | 790 | 4 | 237 | 3 | 2 | 1 | 0.013 | 0.008 | 1.2E-04 | 1.2E-04 | 0.00 | 1.2E-04 | -0.75 |
| LOC116517048 | COX6A1 | MF | OP | NC_045553.1 | 16412457 | 16415643 | 3185 | 1 | 1037 | 0 | 312 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116517973 | COX4I1 | MF | OP | NC_045554.1 | 47273043 | 47276904 | 3860 | 33 | 797 | 7 | 533 | 5 | 1 | 4 | 0.009 | 0.002 | 4.8E-04 | 4.8E-04 | 0.00 | 4.8E-04 | 2.80 |
| LOC116519415 | COX5A | MF | OP | NC_045556.1 | 5393184 | 5401460 | 8275 | 57 | 810 | 9 | 428 | 3 | 1 | 2 | 0.007 | 0.002 | 2.1E-04 | 2.1E-04 | 0.02 | 2.1E-04 | 0.34 |
| LOC116520181 | COX18 | MF | OP | NC_045558.1 | 89705647 | 89725159 | 19511 | 5 | 4944 | 2 | 1023 | 2 | 1 | 1 | 0.002 | 0.001 | 3.5E-05 | 3.5E-05 | 0.05 | 3.5E-05 | 0.13 |
| ATP5F1A | ATP5A1 | MF | OP | NC_045543.1 | 142013791 | 142028579 | 14787 | 79 | 2011 | 23 | 1650 | 17 | 3 | 14 | 0.010 | 0.002 | 1.6E-04 | 1.6E-04 | 0.01 | 1.6E-04 | 2.71 |
| ATP5F1B | ATP5B | MF | OP | NC_045542.1 | 28179118 | 28186474 | 7355 | 20 | 2370 | 6 | 1583 | 5 | 1 | 4 | 0.003 | 0.001 | 1.4E-04 | 1.4E-04 | 0.03 | 1.4E-04 | 2.26 |
| ATP5F1C | ATP5C1 | MF | OP | NC_045547.1 | 5606677 | 5620049 | 16371 | 67 | 2161 | 17 | 815 | 11 | 0 | 11 | 0.013 | 0.000 | 6.1E-05 | 6.1E-05 | 0.01 | 6.1E-05 | -0.23 |
| ATP5F1D | ATP5D | MF | OP | NC_045541.1 | 175081193 | 175091426 | 10232 | 63 | 749 | 18 | 497 | 8 | 4 | 4 | 0.016 | 0.008 | 3.3E-04 | 3.3E-04 | 0.01 | 3.3E-04 | 1.20 |
| ATP5F1E | ATP5E | MF | OP | NC_045545.1 | 86268313 | 86273824 | 5510 | 9 | 2048 | 5 | 151 | 1 | 1 | 0 | 0.007 | 0.007 | 6.5E-05 | 6.5E-05 | 0.01 | 6.5E-05 | -0.16 |
| ATP5M2 | ATP5G2 | MF | OP | NC_045542.1 | 22459222 | 22471798 | 12573 | 18 | 711 | 6 | 446 | 2 | 1 | 1 | 0.004 | 0.002 | 5.0E-04 | 5.0E-04 | 0.00 | 5.0E-04 | 1.61 |
| ATP5M5 | ATP5I | MF | OP | NC_045543.1 | 49250743 | 49255584 | 4840 | 4 | 442 | 1 | 209 | 0 | 0 | 0 | 0.000 | 0.000 | 6.2E-05 | 6.2E-05 | 0.00 | 6.2E-05 | -0.72 |
| ATP5M6 | ATP5J2 | MF | OP | NC_045554.1 | 4627104 | 4630500 | 3395 | 42 | 480 | 6 | 266 | 2 | 0 | 2 | 0.008 | 0.000 | 6.0E-04 | 6.0E-04 | 0.01 | 6.0E-04 | 1.75 |
| ATP5P5 | ATP5F1 | MF | OP | NC_045545.1 | 58431365 | 58444921 | 13555 | 8 | 1548 | 2 | 761 | 0 | 0 | 0 | 0.000 | 0.000 | 1.5E-04 | 1.5E-04 | 0.01 | 1.5E-04 | 0.72 |
| ATP6A1 | ATP5H | MF | OP | NC_045542.1 | 52924691 | 52931199 | 6507 | 4 | 1259 | 1 | 481 | 1 | 1 | 0 | 0.002 | 0.002 | 1.0E-04 | 1.0E-04 | 0.00 | 1.0E-04 | -0.11 |
| ATP6A1 | ATP6A1 | MF | OP | NC_045542.1 | 9510780 | 9528946 | 18165 | 37 | 2030 | 21 | 1340 | 9 | 0 | 9 | 0.007 | 0.000 | 1.4E-04 | 1.4E-04 | 0.01 | 1.4E-04 | 2.13 |
| ATP6A1 | ATP6A1 | MF | OP | NC_045543.1 | 38605536 | 38637546 | 32189 | 57 | 3616 | 14 | 1840 | 9 | 0 | 9 | 0.005 | 0.000 | 1.2E-04 | 1.2E-04 | 0.00 | 1.2E-04 | 4.46 |
| ATP6A1 | ATP6A1 | MF | OP | NC_045553.1 | 34564841 | 34585588 | 18746 | 67 | 1857 | 6 | 1513 | 5 | 0 | 5 | 0.003 | 0.000 | 4.7E-05 | 4.7E-05 | 0.01 | 4.7E-05 | -0.79 |
| ATP6A1 | ATP6A1 | MF | OP | NC_045548.1 | 56988198 | 57025562 | 37363 | 26 | 5318 | 4 | 1274 | 2 | 0 | 2 | 0.002 | 0.000 | 4.8E-05 | 4.8E-05 | 0.02 | 4.8E-05 | 1.01 |
| ATP6A1 | ATP6A1 | MF | OP | NC_045547.1 | 14662149 | 14674540 | 12390 | 26 | 1843 | 6 | 672 | 2 | 1 | 1 | 0.003 | 0.001 | 1.3E-04 | 1.3E-04 | 0.02 | 1.3E-04 | 0.64 |
| ATP6A1 | ATP6A1 | MF | OP | NC_045548.1 | 38891196 | 38910014 | 18817 | 21 | 1704 | 3 | 1418 | 1 | 0 | 1 | 0.001 | 0.000 | 1.3E-04 | 1.3E-04 | 0.01 | 1.3E-04 | 0.09 |
| ISCA1 | ISCA1 | MF | OP | NC_045543.1 | 84051615 | 84058362 | 6746 | 7 | 1752 | 4 | 386 | 2 | 1 | 1 | 0.005 | 0.003 | 3.6E-05 | 3.6E-05 | 0.01 | 3.6E-05 | -0.83 |
| ISCA2 | ISCA2 | MF | OP | NC_045541.1 | 130904989 | 130912489 | 7499 | 18 | 818 | 3 | 521 | 1 | 0 | 1 | 0.002 | 0.000 | 2.9E-04 | 2.9E-04 | 0.01 | 2.9E-04 | 0.69 |
| SIC25A28 | SIC25A28 | MF | OP | NC_045550.1 | 26259342 | 26263919 | 4576 | 5 | 1572 | 3 | 1091 | 2 | 1 | 1 | 0.002 | 0.001 | 9.1E-05 | 9.1E-05 | 0.00 | 9.1E-05 | -0.08 |
| SIC25A37 | SIC25A37 | MF | OP | NC_045553.1 | 27767622 | 27819226 | 51603 | 7 | 2298 | 2 | 1097 | 2 | 0 | 2 | 0.002 | 0.000 | 1.3E-05 | 1.3E-05 | 0.00 | 1.3E-05 | -0.95 |
| NR4A1 | NR4A1 | MF | OP | NC_045542.1 | 28294691 | 28339866 | 45174 | 24 | 7958 | 14 | 1967 | 9 | 3 | 6 | 0.005 | 0.002 | 3.4E-05 | 3.4E-05 | 0.05 | 3.4E-05 | 1.70 |
| NR4A2 | NR4A2 | MF | OP | NC_045541.1 | 34952046 | 34963239 | 11192 | 16 | 7591 | 5 | 1923 | 3 | 2 | 1 | 0.002 | 0.001 | 1.1E-05 | 1.1E-05 | 0.02 | 1.1E-05 | -0.90 |
| NR4A3 | NR4A3 | MF | OP | NC_045558.1 | 82102224 | 82125409 | 23184 | 5 | 8777 | 2 | 1845 | 1 | 1 | 0 | 0.001 | 0.001 | 3.3E-05 | 3.3E-05 | 0.01 | 3.3E-05 | 1.04 |
| SMAD6 | SMAD6 | MF | OP | NC_045556.1 | 22977881 | 22979481 | 57099 | 12 | 2550 | 9 | 1394 | 9 | 2 | 7 | 0.006 | 0.001 | 2.7E-05 | 2.7E-05 | 0.01 | 2.7E-05 | -1.22 |
| SMAD7 | SMAD7 | MF | OP | NC_045543.1 | 139726632 | 139794226 | 57393 | 14 | 2779 | 5 | 1379 | 2 | 0 | 2 | 0.001 | 0.000 | 7.8E-05 | 7.8E-05 | 0.03 | 7.8E-05 | 0.89 |
| ALCAM | ALCAM | O | BR | NC_045546.1 | 34350297 | 3478971 | 128673 | 13 | 6685 | 5 | 1764 | 3 | 0 | 3 | 0.002 | 0.000 | 3.0E-05 | 3.0E-05 | 0.08 | 3.0E-05 | 0.40 |
| AMDHD1 | AMDHD1 | O | BR | NC_045547.1 | 37666474 | 37678064 | 11389 | 13 | 3110 | 10 | 1293 | 3 | 2 | 1 | 0.002 | 0.002 | 8.5E-05 | 8.5E-05 | 0.04 | 8.5E-05 | 1.46 |
| BGCAT3 | BGCAT3 | O | BR | NC_045557.1 | 6271786 | 6280201 | 8414 | 22 | 3755 | 6 | 973 | 6 | 1 | 5 | 0.006 | 0.001 | 1.4E-05 | 1.4E-05 | 0.00 | 1.4E-05 | -1.17 |
| CA5A | CA5A | O | BR | NC_045554.1 | 47994229 | 48011053 | 16823 | 93 | 4346 | 35 | 977 | 29 | 6 | 23 | 0.030 | 0.006 | 5.8E-05 | 5.8E-05 | 0.01 | 5.8E-05 | 2.59 |
| CC2D1A | CC2D1A | O | BR | NC_045542.1 | 4181033 | 4257445 | 76391 | 188 | 3715 | 80 | 2995 | 78 | 43 | 35 | 0.026 | 0.014 | 6.7E-05 | 6.7E-05 | 0.01 | 6.7E-05 | 1.78 |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|--------------|--------------|----------|--------------|-------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| CTSL | CTSL | O | BR | NC_045543.1 | 84641238 | 84652874 | 11635 | 18 | 3114 | 5 | 1016 | 1 | 0 | 1 | 0.001 | 0.000 | 6.5E-05 | 6.5E-05 | 0.02 | 6.5E-05 | 0.42 |
| D2HGDH | D2HGDH | O | BR | NC_045550.1 | 32707353 | 32720172 | 12818 | 4 | 6919 | 2 | 1715 | 1 | 1 | 0 | 0.001 | 0.001 | 4.3E-05 | 4.3E-05 | 0.13 | 4.3E-05 | 0.65 |
| DHX29 | DHX29 | O | BR | NC_045543.1 | 122457999 | 122490672 | 32672 | 42 | 4963 | 7 | 4071 | 7 | 3 | 4 | 0.002 | 0.001 | 1.4E-05 | 1.4E-05 | 0.02 | 1.4E-05 | -1.04 |
| E1F4H | E1F4H | O | BR | NC_045544.1 | 124744258 | 124766850 | 22391 | 17 | 864 | 3 | 749 | 2 | 0 | 2 | 0.003 | 0.000 | 6.2E-05 | 6.2E-05 | 0.00 | 6.2E-05 | -0.91 |
| ELK3 | ELK3 | O | BR | NC_045547.1 | 37532494 | 37584089 | 51594 | 15 | 5399 | 7 | 1214 | 3 | 0 | 3 | 0.002 | 0.000 | 4.9E-05 | 4.9E-05 | 0.03 | 4.9E-05 | 1.43 |
| FOXP1 | FOXP1 | O | BR | NC_045554.1 | 3926145 | 3984172 | 58026 | 63 | 7296 | 24 | 2073 | 24 | 3 | 21 | 0.012 | 0.001 | 2.6E-05 | 2.6E-05 | 0.01 | 2.6E-05 | 0.72 |
| G1GYF2 | G1GYF2 | O | BR | NC_045551.1 | 28492447 | 28591018 | 98570 | 68 | 5457 | 15 | 3825 | 15 | 3 | 12 | 0.004 | 0.001 | 6.6E-05 | 6.6E-05 | 0.12 | 6.6E-05 | 2.30 |
| GPATCH1 | GPATCH1 | O | BR | NC_045554.1 | 37546856 | 37572480 | 25623 | 36 | 2889 | 15 | 2729 | 15 | 3 | 12 | 0.005 | 0.001 | 1.3E-04 | 1.3E-04 | 0.00 | 1.3E-04 | 3.44 |
| GRB2 | GRB2 | O | BR | NC_045542.1 | 53268855 | 53351195 | 82339 | 2 | 2900 | 1 | 808 | 1 | 0 | 1 | 0.001 | 0.000 | 1.7E-05 | 1.7E-05 | 0.00 | 1.7E-05 | -0.59 |
| HDAC10 | HDAC10 | O | BR | NC_045547.1 | 64580905 | 64610740 | 29834 | 28 | 19003 | 6 | 2116 | 4 | 2 | 2 | 0.002 | 0.001 | 1.6E-05 | 1.6E-05 | 0.00 | 1.6E-05 | 1.59 |
| KLHL24 | KLHL24 | O | BR | NC_045550.1 | 26653625 | 26669351 | 15725 | 20 | 9240 | 10 | 1794 | 9 | 1 | 8 | 0.005 | 0.001 | 1.1E-05 | 1.1E-05 | 0.00 | 1.1E-05 | -0.52 |
| KMT2C | KMT2C | O | BR | NC_045558.1 | 9320906 | 9492161 | 171254 | 13 | 160957 | 5 | 14871 | 2 | 0 | 2 | 0.000 | 0.000 | 1.3E-06 | 1.3E-06 | 0.04 | 1.3E-06 | 0.40 |
| LIN7C | LIN7C | O | BR | NC_045541.1 | 94415494 | 94431850 | 16355 | 39 | 3996 | 34 | 589 | 0 | 0 | 0 | 0.000 | 0.000 | 4.1E-05 | 4.1E-05 | 0.00 | 4.1E-05 | 0.10 |
| LOC116503212 | TUBA1B | O | BR | NC_045542.1 | 35893622 | 35903516 | 9893 | 4 | 1688 | 0 | 1352 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116503236 | H2AC1 | O | BR | NC_045542.1 | 116187553 | 116192067 | 4713 | 22 | 9045 | 23 | 380 | 2 | 0 | 2 | 0.005 | 0.000 | 3.4E-05 | 3.4E-05 | 0.00 | 3.4E-05 | 2.65 |
| LOC116504544 | Zip160 | O | BR | NC_045542.1 | 160893567 | 160915567 | 21999 | 19 | 7690 | 0 | 1567 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116505849 | CYP2K21 | O | BR | NC_045543.1 | 30512609 | 30548372 | 35762 | 16 | 5610 | 0 | 1497 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116505884 | T1SZ | O | BR | NC_045545.1 | 8620874 | 8622547 | 14472 | 16 | 1321 | 1 | 900 | 1 | 0 | 1 | 0.001 | 0.000 | 2.7E-05 | 2.7E-05 | 0.00 | 2.7E-05 | -0.67 |
| LOC116514694 | ARGLU1 | O | BR | NC_045551.1 | 41274693 | 41282695 | 8001 | 22 | 1677 | 8 | 875 | 2 | 0 | 2 | 0.002 | 0.000 | 8.4E-05 | 8.4E-05 | 0.02 | 8.4E-05 | -0.58 |
| LOC116515224 | AGC4 | O | BR | NC_045552.1 | 21954615 | 21982503 | 27887 | 51 | 19447 | 23 | 2608 | 7 | 0 | 7 | 0.003 | 0.000 | 9.3E-06 | 9.3E-06 | 0.01 | 9.3E-06 | -0.07 |
| LOC116521616 | ZNF629 | O | BR | NC_045558.1 | 123023699 | 123047376 | 23676 | 176 | 6024 | 151 | 4189 | 115 | 65 | 50 | 0.027 | 0.016 | 7.0E-05 | 7.0E-05 | 0.00 | 7.0E-05 | 6.16 |
| LITBP1 | LITBP1 | O | BR | NC_045544.1 | 75721901 | 75937105 | 216203 | 13 | 17227 | 1 | 4833 | 1 | 0 | 1 | 0.000 | 0.000 | 1.2E-05 | 1.2E-05 | 0.00 | 1.2E-05 | 0.32 |
| MAK | MAK | O | BR | NC_045548.1 | 7870292 | 7902523 | 34960 | 113 | 4080 | 28 | 1868 | 17 | 5 | 12 | 0.009 | 0.003 | 7.0E-05 | 7.0E-05 | 0.00 | 7.0E-05 | 2.35 |
| MAP2K6 | MAP2K6 | O | BR | NC_045542.1 | 40369878 | 40490821 | 120942 | 15 | 5006 | 2 | 1020 | 0 | 0 | 0 | 0.000 | 0.000 | 4.4E-05 | 4.4E-05 | 0.00 | 4.4E-05 | 0.59 |
| MAP3K11 | MAP3K11 | O | BR | NC_045557.1 | 9357203 | 9391243 | 34039 | 113 | 3140 | 44 | 2609 | 44 | 6 | 38 | 0.017 | 0.002 | 9.7E-05 | 9.7E-05 | 0.01 | 9.7E-05 | 2.89 |
| MAPK1 | MAPK1 | O | BR | NC_045553.1 | 9071472 | 9089594 | 18121 | 16 | 4101 | 4 | 1099 | 2 | 0 | 2 | 0.002 | 0.000 | 4.8E-05 | 4.8E-05 | 0.01 | 4.8E-05 | 0.88 |
| MPST | MPST | O | BR | NC_045547.1 | 29055806 | 29061398 | 5591 | 8 | 1457 | 8 | 892 | 7 | 2 | 5 | 0.008 | 0.002 | 1.2E-04 | 1.2E-04 | 0.00 | 1.2E-04 | -0.21 |
| NR0B2 | NR0B2 | O | BR | NC_045552.1 | 18271410 | 18274315 | 2904 | 7 | 2019 | 4 | 763 | 0 | 0 | 0 | 0.000 | 0.000 | 6.1E-05 | 6.1E-05 | 0.02 | 6.1E-05 | -0.07 |
| PODN | PODN | O | BR | NC_045545.1 | 33441998 | 33464822 | 22823 | 17 | 2891 | 7 | 1755 | 6 | 2 | 4 | 0.003 | 0.001 | 3.2E-05 | 3.2E-05 | 0.01 | 3.2E-05 | -0.81 |
| PRKDC | PRKDC | O | BR | NC_045548.1 | 36295762 | 36421690 | 125927 | 60 | 34576 | 17 | 12304 | 17 | 13 | 4 | 0.001 | 0.001 | 3.4E-06 | 3.4E-06 | 0.03 | 3.4E-06 | -0.55 |
| PRKCG | PRKCG | O | BR | NC_045549.1 | 10018132 | 10012390 | 104259 | 148 | 3188 | 35 | 2325 | 23 | 5 | 18 | 0.010 | 0.002 | 5.7E-05 | 5.7E-05 | 0.01 | 5.7E-05 | 0.56 |
| PSAT1 | PSAT1 | O | BR | NC_045543.1 | 80652527 | 80652390 | 27132 | 27 | 3248 | 19 | 1085 | 1 | 0 | 1 | 0.001 | 0.000 | 8.3E-05 | 8.3E-05 | 0.09 | 8.3E-05 | 1.53 |
| RAB3IP | RAB3IP | O | BR | NC_045547.1 | 47753981 | 47783958 | 29976 | 13 | 8521 | 2 | 1376 | 2 | 1 | 1 | 0.001 | 0.001 | 4.9E-05 | 4.9E-05 | 0.01 | 4.9E-05 | 2.18 |
| RFX5 | RFX5 | O | BR | NC_045557.1 | 1894835 | 1902664 | 7828 | 101 | 3091 | 30 | 2091 | 22 | 7 | 15 | 0.011 | 0.003 | 9.0E-05 | 9.0E-05 | 0.00 | 9.0E-05 | 2.19 |
| RNHI | RNHI | O | BR | NC_045541.1 | 75584582 | 75605590 | 21007 | 12 | 6098 | 2 | 1568 | 1 | 0 | 1 | 0.001 | 0.000 | 1.4E-05 | 1.4E-05 | 0.02 | 1.4E-05 | -0.40 |
| RUVBL1 | RUVBL1 | O | BR | NC_045542.1 | 116951956 | 116976901 | 24964 | 7 | 2577 | 3 | 1360 | 0 | 0 | 0 | 0.000 | 0.000 | 1.4E-04 | 1.4E-04 | 0.00 | 1.4E-04 | 1.68 |
| SLC20A1 | SLC20A1 | O | BR | NC_045553.1 | 20284716 | 20293724 | 9007 | 10 | 3122 | 6 | 2018 | 2 | 1 | 1 | 0.001 | 0.000 | 8.5E-05 | 8.5E-05 | 0.00 | 8.5E-05 | 1.02 |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|---------------|--------------|----------|--------------|-------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| SMC6 | SMC6 | O | BR | NC_045544.1 | 142303558 | 14269289 | 65730 | 75 | 10625 | 44 | 4199 | 37 | 6 | 31 | 0.009 | 0.001 | 2.5E-05 | 2.5E-05 | 0.01 | 2.5E-05 | 2.19 |
| SNRPA1 | SNRPA1 | O | BR | NC_045556.1 | 16269376 | 16277989 | 8612 | 27 | 1090 | 3 | 759 | 2 | 0 | 2 | 0.003 | 0.000 | 4.1E-04 | 4.1E-04 | 0.01 | 4.1E-04 | 1.67 |
| SPATA20 | SPATA20 | O | BR | NC_045542.1 | 51187917 | 51241667 | 53749 | 23 | 2330 | 3 | 2330 | 3 | 2 | 1 | 0.001 | 0.001 | 1.3E-04 | 1.3E-04 | 0.03 | 1.3E-04 | 0.82 |
| TBX18 | TBX18 | O | BR | NC_045544.1 | 18733030 | 18759682 | 26651 | 0 | 1861 | 0 | 1834 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| TMEM184A | TMEM184A | O | BR | NC_045554.1 | 2358630 | 2376466 | 17815 | 104 | 39332 | 26 | 1395 | 25 | 4 | 21 | 0.018 | 0.003 | 7.1E-06 | 7.1E-06 | 0.01 | 7.1E-06 | 2.09 |
| TRPC6 | TRPC6 | O | BR | NC_045546.1 | 14776937 | 14866688 | 89750 | 0 | 3153 | 0 | 2679 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| WDR12 | WDR12 | O | BR | NC_045541.1 | 7262402 | 7294038 | 31635 | 116 | 1985 | 35 | 1265 | 18 | 1 | 17 | 0.014 | 0.001 | 1.3E-04 | 1.3E-04 | 0.01 | 1.3E-04 | 2.02 |
| ADRB2 | ADRB2 | O | B | NC_045542.1 | 76394605 | 76463823 | 69217 | 20 | 1948 | 9 | 1225 | 8 | 1 | 7 | 0.007 | 0.001 | 1.0E-04 | 1.0E-04 | 0.00 | 1.0E-04 | 0.31 |
| AVPR2 | AVPR2 | O | B | NC_045542.1 | 7914403 | 7919720 | 5316 | 5 | 1071 | 5 | 1071 | 5 | 1 | 4 | 0.005 | 0.001 | 1.2E-04 | 1.2E-04 | 0.00 | 1.2E-04 | -0.36 |
| CKKAR | CKKAR | O | B | NC_045549.1 | 6899777 | 68427029 | 27251 | 94 | 1723 | 32 | 1335 | 21 | 9 | 12 | 0.016 | 0.007 | 1.7E-04 | 1.7E-04 | 0.00 | 1.7E-04 | 2.28 |
| DRD4 | DRD4 | O | B | NC_045541.1 | 76374179 | 76405856 | 31676 | 7 | 1127 | 4 | 1127 | 4 | 3 | 1 | 0.004 | 0.003 | 3.3E-04 | 3.3E-04 | 0.01 | 3.3E-04 | 2.49 |
| EDA | EDA | O | B | NC_045552.1 | 35702579 | 35795587 | 93007 | 1 | 2746 | 0 | 1111 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116510561 | MAOA | O | B | NC_045546.1 | 59089885 | 59145594 | 55708 | 4 | 10071 | 4 | 1581 | 2 | 1 | 1 | 0.001 | 0.001 | 5.3E-06 | 5.3E-06 | 0.01 | 5.3E-06 | -1.06 |
| LOC116510618 | MAOB | O | B | NC_045546.1 | 58981266 | 59083947 | 102680 | 14 | 7930 | 1 | 1554 | 1 | 1 | 0 | 0.001 | 0.001 | 2.8E-05 | 2.8E-05 | 0.08 | 2.8E-05 | 0.30 |
| SLC6A12 | SLC6A12 | O | B | NC_045547.1 | 17117517 | 17197177 | 79659 | 51 | 5888 | 15 | 1935 | 7 | 5 | 2 | 0.004 | 0.003 | 2.1E-05 | 2.1E-05 | 0.01 | 2.1E-05 | -0.46 |
| KITLG | KITLG | O | C | NC_045547.1 | 39920623 | 39973368 | 52744 | 0 | 1222 | 0 | 834 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116516483 | BCDO2 | O | C | NC_045553.1 | 48103300 | 48127202 | 23901 | 0 | 4075 | 0 | 1728 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| MCIIR | MCIIR | O | C | NC_045554.1 | 49215175 | 49216120 | 944 | 0 | 944 | 0 | 944 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| MITF | MITF | O | C | NC_045542.1 | 132520584 | 132703024 | 182439 | 0 | 8530 | 0 | 1717 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| OCA2 | OCA2 | O | C | NC_045551.1 | 31253402 | 31450074 | 196671 | 0 | 2573 | 0 | 2573 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| PMEL | PMEL | O | C | NC_045542.1 | 36558826 | 36575519 | 16692 | 0 | 3389 | 0 | 2481 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| POMC | POMC | O | C | NC_045544.1 | 114460473 | 114464283 | 3809 | 0 | 790 | 0 | 790 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| SOX10 | SOX10 | O | C | NC_045547.1 | 30459632 | 30483595 | 23962 | 0 | 3745 | 0 | 1383 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| TYRP1 | TYRP1 | O | C | NC_045543.1 | 71213337 | 71232140 | 18802 | 8 | 1577 | 1 | 1577 | 1 | 1 | 0 | 0.001 | 0.001 | 2.0E-05 | 2.0E-05 | 0.00 | 2.0E-05 | -0.70 |
| CFH | CFH | O | FE | NC_045545.1 | 259128 | 351600 | 92471 | 0 | 11068 | 0 | 3476 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116512145 | NCR3L1 | O | FE | NC_045541.1 | 83025668 | 83109855 | 87286 | 0 | 67185 | 0 | 4397 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116516798 | JGLL5 | O | FE | NC_045553.1 | 8134218 | 8146494 | 12275 | 0 | 597 | 0 | 375 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| SV2A | SV2A | O | FE | NC_045557.1 | 1045941 | 1054510 | 8568 | 0 | 2772 | 0 | 2178 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116519949 | MYO1E | O | G | NC_045557.1 | 1462667 | 1481256 | 18388 | 62 | 3771 | 18 | 3278 | 18 | 2 | 16 | 0.005 | 0.001 | 6.2E-05 | 6.2E-05 | 0.03 | 6.2E-05 | 1.31 |
| TSHB | TSHB | O | G | NC_045545.1 | 62155614 | 62159733 | 4138 | 6 | 480 | 0 | 480 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| TSHR | TSHR | O | G | NC_045541.1 | 135259579 | 135299383 | 39803 | 17 | 6640 | 3 | 2385 | 2 | 1 | 1 | 0.001 | 0.000 | 1.9E-05 | 1.9E-05 | 0.00 | 1.9E-05 | -0.23 |
| VIPR1 | VIPR1 | O | G | NC_045558.1 | 4786194 | 4818310 | 32115 | 51 | 1633 | 16 | 1206 | 13 | 5 | 8 | 0.011 | 0.004 | 1.4E-04 | 1.4E-04 | 0.03 | 1.4E-04 | 1.26 |
| CD74 | CD74 | O | I | NC_045542.1 | 91168637 | 91196013 | 27375 | 34 | 6269 | 20 | 831 | 11 | 8 | 3 | 0.013 | 0.010 | 3.0E-05 | 3.0E-05 | 0.00 | 3.0E-05 | 0.62 |
| LOC1165103377 | MRI | O | I | NC_045542.1 | 170768524 | 170886972 | 118447 | 0 | 6266 | 0 | 1362 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC1165103386 | MRI | O | I | NC_045542.1 | 17100170 | 171505786 | 504075 | 0 | 10422 | 0 | 1113 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC1165103390 | MRI | O | I | NC_045542.1 | 169611652 | 169618702 | 7049 | 0 | 1327 | 0 | 1195 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC1165103394 | MRI | O | I | NC_045542.1 | 170899655 | 170915629 | 15973 | 0 | 2397 | 0 | 1084 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|--------------|--------------|----------|--------------|-------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| LOC116508405 | MRI | O | I | NC_045542.1 | 17284745 | 172901387 | 16641 | 57 | 4228 | 15 | 1057 | 14 | 6 | 8 | 0.013 | 0.006 | 3.0E-05 | 3.0E-05 | 0.00 | 3.0E-05 | -0.50 |
| LOC116507754 | FSHR | O | R | NC_045544.1 | 101979094 | 102013729 | 34634 | 3 | 3806 | 3 | 1646 | 1 | 0 | 1 | 0.001 | 0.000 | 9.4E-05 | 9.4E-05 | 0.02 | 9.4E-05 | 1.88 |
| | DMRT1 | O | R | NC_045543.1 | 76602171 | 76682763 | 80591 | 9 | 1102 | 1 | 1102 | 1 | 0 | 1 | 0.001 | 0.000 | 3.9E-05 | 3.9E-05 | 0.00 | 3.9E-05 | -0.64 |
| | RBMX | O | R | NC_045552.1 | 57104939 | 57114934 | 9994 | 36 | 3968 | 9 | 1144 | 4 | 0 | 4 | 0.003 | 0.000 | 1.7E-05 | 1.7E-05 | -0.01 | 1.7E-05 | -1.19 |
| | SOX4 | O | R | NC_045548.1 | 12502775 | 12507109 | 4333 | 3 | 4333 | 3 | 1286 | 2 | 1 | 1 | 0.002 | 0.001 | 3.3E-05 | 3.3E-05 | 0.00 | 3.3E-05 | 0.03 |
| | SOX9 | O | R | NC_045542.1 | 42012178 | 42048058 | 26779 | 6 | 7585 | 4 | 1500 | 3 | 1 | 2 | 0.002 | 0.001 | 2.2E-05 | 2.2E-05 | 0.03 | 2.2E-05 | -0.02 |
| | USP24 | O | R | NC_045545.1 | 31450839 | 31541678 | 90838 | 46 | 31045 | 11 | 7742 | 7 | 3 | 4 | 0.001 | 0.000 | 6.7E-06 | 6.7E-06 | 0.00 | 6.7E-06 | 0.45 |
| | BTBD16 | O | R | NC_045550.1 | 867733 | 8734744 | 56990 | 36 | 1976 | 1 | 1837 | 1 | 0 | 1 | 0.001 | 0.000 | 3.2E-05 | 3.2E-05 | 0.00 | 3.2E-05 | -0.51 |
| | CATSPER1 | O | R | NC_045557.1 | 8957182 | 8964303 | 7120 | 0 | 886 | 0 | 444 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| | CATSPER2 | O | R | NC_045548.1 | 75263269 | 75264739 | 1469 | 12 | 1469 | 12 | 1469 | 12 | 6 | 6 | 0.008 | 0.004 | 2.3E-04 | 2.3E-04 | 0.00 | 2.3E-04 | 2.93 |
| | CATSPER3 | O | R | NC_045542.1 | 85985786 | 86002071 | 16284 | 23 | 2000 | 9 | 1221 | 6 | 1 | 5 | 0.005 | 0.001 | 1.6E-04 | 1.6E-04 | 0.03 | 1.6E-04 | 1.98 |
| | CATSPER4 | O | R | NC_045552.1 | 16116755 | 16164788 | 48712 | 8 | 2450 | 3 | 2267 | 3 | 0 | 3 | 0.001 | 0.000 | 5.6E-05 | 5.6E-05 | 0.00 | 5.6E-05 | 0.06 |
| | CATSPERG | O | R | NC_045552.1 | 2866949 | 2920727 | 53777 | 74 | 3635 | 25 | 3307 | 25 | 14 | 11 | 0.008 | 0.004 | 6.9E-05 | 6.9E-05 | 0.01 | 6.9E-05 | 1.80 |
| | SPATA18 | O | R | NC_045549.1 | 6759202 | 67626213 | 36010 | 28 | 5954 | 9 | 1509 | 9 | 2 | 7 | 0.006 | 0.001 | 6.2E-05 | 6.2E-05 | 0.02 | 6.2E-05 | 3.05 |
| | SPATA22 | O | R | NC_045544.1 | 139211603 | 139223592 | 11988 | 71 | 1290 | 16 | 1078 | 16 | 5 | 11 | 0.015 | 0.005 | 3.0E-04 | 3.0E-04 | 0.00 | 3.0E-04 | 3.65 |
| | SYCF2 | O | R | NC_045545.1 | 86910991 | 86922277 | 11285 | 14 | 1017 | 1 | 959 | 1 | 0 | 1 | 0.001 | 0.000 | 2.9E-05 | 2.9E-05 | 0.02 | 2.9E-05 | -0.72 |
| | TEX261 | O | R | NC_045549.1 | 1126145 | 1136503 | 10357 | 5 | 1230 | 5 | 585 | 0 | 0 | 0 | 0.000 | 0.000 | 1.8E-04 | 1.8E-04 | 0.00 | 1.8E-04 | 1.01 |
| | DDX4 | O | R | NC_045543.1 | 122241130 | 122257920 | 16789 | 15 | 2321 | 1 | 1963 | 1 | 1 | 0 | 0.001 | 0.001 | 8.6E-06 | 8.6E-06 | 0.09 | 8.6E-06 | -0.76 |
| | EIF4A3 | O | R | NC_045544.1 | 100836144 | 100847412 | 11267 | 18 | 1649 | 4 | 1221 | 1 | 0 | 1 | 0.001 | 0.000 | 1.6E-04 | 1.6E-04 | 0.01 | 1.6E-04 | 1.18 |
| | PGK1 | O | R | NC_045552.1 | 37394430 | 37415101 | 20670 | 10 | 1760 | 6 | 1243 | 2 | 1 | 1 | 0.002 | 0.001 | 8.7E-05 | 8.7E-05 | 0.01 | 8.7E-05 | -0.24 |
| | RBM15B | O | R | NC_045542.1 | 103731498 | 103735778 | 4279 | 12 | 4279 | 12 | 2774 | 11 | 2 | 9 | 0.004 | 0.001 | 4.8E-05 | 4.8E-05 | 0.00 | 4.8E-05 | 0.67 |
| | MAPKAPK2 | SA | CS | NC_045545.1 | 55461929 | 55538600 | 76670 | 10 | 3267 | 3 | 1187 | 0 | 0 | 0 | 0.000 | 0.000 | 2.6E-05 | 2.6E-05 | 0.00 | 2.6E-05 | -0.20 |
| | MOK | SA | CS | NC_045541.1 | 147405811 | 147425113 | 19301 | 10 | 10303 | 1 | 1251 | 0 | 0 | 0 | 0.000 | 0.000 | 1.3E-05 | 1.3E-05 | 0.00 | 1.3E-05 | -0.11 |
| | NFKB1 | SA | CS | NC_045549.1 | 21785778 | 21866873 | 81094 | 13 | 3736 | 3 | 2896 | 1 | 0 | 1 | 0.000 | 0.000 | 3.4E-05 | 3.4E-05 | 0.00 | 3.4E-05 | -0.36 |
| | NFKBIA | SA | CS | NC_045541.1 | 128347484 | 128352147 | 4662 | 17 | 1208 | 3 | 966 | 3 | 1 | 2 | 0.003 | 0.001 | 1.6E-04 | 1.6E-04 | 0.06 | 1.6E-04 | 0.37 |
| | NFKBIB | SA | CS | NC_045552.1 | 2095672 | 2104406 | 8733 | 55 | 3335 | 24 | 1083 | 15 | 3 | 12 | 0.014 | 0.003 | 7.2E-05 | 7.2E-05 | 0.00 | 7.2E-05 | 1.50 |
| | NUP88 | SA | CS | NC_045544.1 | 122767228 | 122787174 | 19945 | 43 | 3227 | 18 | 2173 | 8 | 4 | 4 | 0.004 | 0.002 | 6.6E-05 | 6.6E-05 | 0.00 | 6.6E-05 | 1.01 |
| | PPIA | SA | CS | NC_045547.1 | 73961252 | 74008911 | 47658 | 28 | 2473 | 9 | 1299 | 7 | 1 | 6 | 0.005 | 0.001 | 1.5E-04 | 1.5E-04 | 0.02 | 1.5E-04 | 3.03 |
| | PPARD | SA | CS | NC_045545.1 | 50366967 | 50418617 | 51649 | 16 | 39949 | 4 | 2369 | 1 | 0 | 1 | 0.000 | 0.000 | 9.0E-06 | 9.0E-06 | 0.00 | 9.0E-06 | 2.28 |
| | TRAP1 | SA | CS | NC_045554.1 | 41069 | 48930 | 7860 | 55 | 4642 | 12 | 2149 | 10 | 4 | 6 | 0.005 | 0.002 | 2.7E-05 | 2.7E-05 | 0.01 | 2.7E-05 | -0.26 |
| | HSF1 | SA | H | NC_045548.1 | 80150444 | 80182729 | 28684 | 62 | 4479 | 16 | 1606 | 16 | 7 | 9 | 0.010 | 0.004 | 5.6E-05 | 5.6E-05 | 0.01 | 5.6E-05 | 1.57 |
| | HSF2 | SA | H | NC_045544.1 | 38052949 | 38079562 | 26612 | 4 | 9226 | 2 | 1676 | 2 | 0 | 2 | 0.001 | 0.000 | 1.6E-05 | 1.6E-05 | 0.05 | 1.6E-05 | -0.07 |
| | HSP90AA1 | SA | H | NC_045541.1 | 147315906 | 147322852 | 6945 | 43 | 2769 | 15 | 2177 | 9 | 0 | 9 | 0.004 | 0.000 | 1.1E-04 | 1.1E-04 | 0.07 | 1.1E-04 | 2.19 |
| | HSP90B1 | SA | H | NC_045547.1 | 24421892 | 24434095 | 12202 | 25 | 3453 | 9 | 2570 | 3 | 2 | 1 | 0.001 | 0.001 | 4.6E-05 | 4.6E-05 | 0.01 | 4.6E-05 | 0.01 |
| | HSPA13 | SA | H | NC_045546.1 | 45595703 | 45595126 | 7222 | 14 | 3071 | 5 | 1405 | 3 | 3 | 0 | 0.002 | 0.002 | 6.2E-05 | 6.2E-05 | 0.03 | 6.2E-05 | 0.30 |
| | HSPA2 | SA | H | NC_045541.1 | 15250382 | 152506232 | 2349 | 6 | 2349 | 6 | 1907 | 4 | 0 | 4 | 0.002 | 0.000 | 5.8E-05 | 5.8E-05 | 0.03 | 5.8E-05 | -0.18 |
| | HSPA4 | SA | H | NC_045542.1 | 82303217 | 82330594 | 27376 | 64 | 3909 | 7 | 2510 | 5 | 2 | 3 | 0.002 | 0.001 | 2.8E-05 | 2.8E-05 | 0.01 | 2.8E-05 | -0.53 |

Table 4.S1: Statistics for Targeted Genes (cont.)

| Gene ID | Human Symbol | Category | Sub-category | Scaffold | Start | End | Gene Len | Genes SNPs | Exons Len | Exons SNPs | CDS Len | CDS SNPs | Mis SNPs | Syn SNPs | SPCL | MPCL | PIM | PiL | F _{ST} LM | D _{XY} LM | TajD |
|--------------|--------------|----------|--------------|-------------|-----------|-----------|----------|------------|-----------|------------|---------|----------|----------|----------|-------|-------|---------|---------|--------------------|--------------------|-------|
| HSPA5 | HSPA5 | SA | H | NC_045556.1 | 35104000 | 35106660 | 6459 | 78 | 2139 | 25 | 1951 | 23 | 0 | 23 | 0.012 | 0.000 | 1.3E-04 | 1.3E-04 | 0.00 | 1.3E-04 | 2.10 |
| HSPA8 | HSPA8 | SA | H | NC_045553.1 | 44843810 | 44854285 | 10474 | 102 | 2236 | 31 | 1933 | 28 | 0 | 28 | 0.014 | 0.000 | 7.2E-05 | 7.2E-05 | 0.01 | 7.2E-05 | -0.01 |
| HSPA9 | HSPA9 | SA | H | NC_045543.1 | 17768052 | 17791685 | 23632 | 92 | 2934 | 19 | 2020 | 10 | 0 | 10 | 0.005 | 0.000 | 8.4E-05 | 8.4E-05 | 0.02 | 8.4E-05 | 1.08 |
| HSPB1 | HSPB1 | SA | H | NC_045544.1 | 121716293 | 121726986 | 10692 | 17 | 956 | 11 | 630 | 11 | 3 | 8 | 0.017 | 0.005 | 3.8E-04 | 3.8E-04 | 0.00 | 3.8E-04 | 3.18 |
| HSPH1 | HSPH1 | SA | H | NC_045546.1 | 24397277 | 24420818 | 23540 | 57 | 1853 | 8 | 2544 | 8 | 5 | 3 | 0.003 | 0.002 | 1.8E-05 | 1.8E-05 | 0.06 | 1.8E-05 | 2.11 |
| LOC116515182 | HSPA8-like | SA | H | NC_045552.1 | 6958345 | 6977329 | 18983 | 146 | 2443 | 35 | 1939 | 30 | 2 | 28 | 0.015 | 0.001 | 1.3E-04 | 1.3E-04 | 0.01 | 1.3E-04 | 2.85 |
| AKT1 | AKT1 | SA | HY | NC_045541.1 | 151083181 | 151180361 | 97179 | 26 | 7479 | 5 | 1430 | 2 | 0 | 2 | 0.001 | 0.000 | 1.8E-05 | 1.8E-05 | 0.01 | 1.8E-05 | -0.23 |
| AKT2 | AKT2 | SA | HY | NC_045552.1 | 242902 | 252080 | 9177 | 42 | 13163 | 19 | 1475 | 16 | 5 | 11 | 0.011 | 0.003 | 1.8E-05 | 1.8E-05 | 0.03 | 1.8E-05 | 1.20 |
| AKT3 | AKT3 | SA | HY | NC_045544.1 | 71039357 | 71205040 | 169982 | 13 | 17954 | 3 | 1427 | 2 | 0 | 2 | 0.001 | 0.000 | 2.1E-05 | 2.1E-05 | 0.00 | 2.1E-05 | 2.17 |
| ARNT | ARNT | SA | HY | NC_045557.1 | 1721974 | 1737863 | 15888 | 6 | 2448 | 4 | 2314 | 4 | 0 | 4 | 0.002 | 0.000 | 6.3E-05 | 6.3E-05 | 0.00 | 6.3E-05 | 0.17 |
| CREBBP | CREBBP | SA | HY | NC_045554.1 | 28 | 31460 | 31431 | 57 | 6642 | 12 | 5682 | 12 | 4 | 8 | 0.002 | 0.001 | 2.3E-05 | 2.3E-05 | 0.01 | 2.3E-05 | -0.06 |
| EP300 | EP300 | SA | HY | NC_045547.1 | 32680338 | 32755615 | 75276 | 10 | 8504 | 2 | 7201 | 2 | 1 | 1 | 0.000 | 0.000 | 2.4E-05 | 2.4E-05 | 0.00 | 2.4E-05 | 0.57 |
| HIF1A | HIF1A | SA | HY | NC_045541.1 | 154059133 | 154105065 | 45931 | 41 | 2889 | 10 | 2451 | 8 | 4 | 4 | 0.003 | 0.002 | 9.7E-05 | 9.7E-05 | 0.10 | 9.7E-05 | 1.65 |
| HIF1AN | HIF1AN | SA | HY | NC_045555.1 | 18919275 | 18937425 | 18149 | 24 | 1957 | 2 | 1063 | 2 | 1 | 1 | 0.002 | 0.001 | 1.3E-04 | 1.3E-04 | 0.00 | 1.3E-04 | 0.91 |
| LOC116506478 | TXN | SA | HY | NC_045543.1 | 59501904 | 59524365 | 22460 | 2 | 740 | 0 | 565 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| LOC116506537 | PIK3C3 | SA | HY | NC_045543.1 | 144287033 | 14437346 | 90312 | 173 | 6073 | 44 | 1577 | 26 | 3 | 23 | 0.016 | 0.002 | 5.0E-05 | 5.0E-05 | 0.00 | 5.0E-05 | 2.74 |
| LOC116510063 | HBE1 | SA | HY | NC_045546.1 | 399498 | 400622 | 1123 | 10 | 693 | 9 | 441 | 6 | 0 | 6 | 0.014 | 0.000 | 3.0E-04 | 3.0E-04 | 0.02 | 3.0E-04 | 0.69 |
| LOC116512681 | HBA2 | SA | HY | NC_045548.1 | 6323926 | 6325300 | 1373 | 24 | 890 | 19 | 426 | 12 | 6 | 6 | 0.028 | 0.014 | 1.6E-04 | 1.6E-04 | 0.01 | 1.6E-04 | -0.12 |
| PIK3C2B | PIK3C2B | SA | HY | NC_045545.1 | 52853706 | 52964169 | 110462 | 49 | 41491 | 16 | 4891 | 13 | 6 | 7 | 0.003 | 0.001 | 4.3E-06 | 4.3E-06 | 0.01 | 4.3E-06 | 0.23 |
| PIK3CA | PIK3CA | SA | HY | NC_045550.1 | 60206996 | 60250578 | 43381 | 82 | 9877 | 15 | 3184 | 15 | 0 | 15 | 0.005 | 0.000 | 4.4E-05 | 4.4E-05 | 0.00 | 4.4E-05 | 4.51 |
| PIK3CB | PIK3CB | SA | HY | NC_045550.1 | 34487265 | 34569141 | 81875 | 9 | 6174 | 0 | 3165 | 0 | 0 | 0 | 0.000 | 0.000 | NA | NA | NA | NA | NA |
| PIK3CD | PIK3CD | SA | HY | NC_045555.1 | 607188 | 6055387 | 48398 | 96 | 12390 | 18 | 3110 | 18 | 5 | 13 | 0.006 | 0.002 | 1.6E-05 | 1.6E-05 | 0.00 | 1.6E-05 | 0.87 |
| PIK3R1 | PIK3R1 | SA | HY | NC_045543.1 | 116808097 | 116871157 | 63059 | 19 | 33008 | 7 | 2157 | 3 | 0 | 3 | 0.001 | 0.000 | 9.0E-06 | 9.0E-06 | 0.05 | 9.0E-06 | 1.53 |
| PIK3R2 | PIK3R2 | SA | HY | NC_045541.1 | 165596196 | 165643817 | 49620 | 75 | 3265 | 14 | 2169 | 11 | 2 | 9 | 0.005 | 0.001 | 3.9E-05 | 3.9E-05 | 0.00 | 3.9E-05 | -0.28 |
| RORC | RORC | SA | HY | NC_045557.1 | 3749241 | 3756948 | 7706 | 111 | 2021 | 51 | 1357 | 29 | 13 | 16 | 0.021 | 0.010 | 1.2E-04 | 1.2E-04 | 0.03 | 1.2E-04 | 1.74 |
| TXN2 | TXN2 | SA | HY | NC_045547.1 | 28611605 | 28628204 | 16398 | 25 | 3457 | 7 | 516 | 1 | 0 | 1 | 0.002 | 0.000 | 4.6E-05 | 4.6E-05 | 0.02 | 4.6E-05 | -0.01 |
| VHL | VHL | SA | HY | NC_045542.1 | 122495983 | 122503014 | 7030 | 24 | 4767 | 11 | 537 | 4 | 1 | 3 | 0.007 | 0.002 | 7.4E-05 | 7.4E-05 | 0.05 | 7.4E-05 | 2.18 |

Table 4.S2: Genes with Significant Interpopulation-Ecotype Divergence

| Human Symbol | Category | Sub-Category | N SNPs | Total Length | $\overline{F_{ST}}$ Within | $\overline{F_{ST}}$ Between | F_{ST} p-val | $\overline{D_{XY}}$ Within | $\overline{D_{XY}}$ Between | D_{XY} p-val |
|--------------|----------|--------------|--------|--------------|----------------------------|-----------------------------|----------------|----------------------------|-----------------------------|----------------|
| HSPA13 | SA | H | 5 | 3071 | -0.002 | 0.034 | 1.1E-06 | 5.4E-05 | 5.9E-05 | 1.2E-01 |
| IGFBP3 | MF | NS | 1 | 2422 | 0.017 | 0.126 | 2.0E-05 | 8.8E-05 | 1.1E-04 | 3.4E-01 |
| NDUFS3 | MF | OP | 3 | 954 | -0.004 | 0.043 | 2.4E-05 | 1.8E-04 | 2.0E-04 | 1.8E-01 |
| NDUFB6 | MF | OP | 3 | 795 | 0.031 | 0.152 | 2.4E-05 | 2.8E-04 | 3.6E-04 | 8.7E-02 |
| NFKBIA | SA | GS | 3 | 1208 | 0.004 | 0.077 | 5.3E-05 | 1.4E-04 | 1.7E-04 | 3.0E-03 |
| PYROXD2 | MDR | OS | 7 | 2188 | 0.016 | 0.073 | 1.2E-04 | 9.0E-05 | 1.0E-04 | 7.3E-04 |
| COX7C | MF | OP | 2 | 871 | 0.000 | 0.036 | 4.1E-04 | 1.5E-04 | 1.5E-04 | 8.1E-01 |
| VHL | SA | HY | 11 | 4767 | 0.017 | 0.060 | 5.9E-04 | 6.7E-05 | 7.4E-05 | 8.0E-05 |
| GPX3 | MDR | OS | 8 | 2120 | 0.031 | 0.069 | 6.2E-04 | 9.3E-05 | 1.1E-04 | 4.1E-02 |
| FOXO3 | MF | NS | 5 | 6216 | 0.034 | 0.103 | 7.6E-04 | 4.8E-05 | 5.4E-05 | 7.5E-02 |
| HIF1A | SA | HY | 10 | 2889 | 0.018 | 0.066 | 1.5E-03 | 8.0E-05 | 8.9E-05 | 4.4E-02 |
| MAP3K1 | MF | NS | 12 | 7006 | 0.011 | 0.034 | 1.8E-03 | 1.7E-05 | 1.8E-05 | 2.0E-01 |
| IGF2 | MF | NS | 5 | 3870 | 0.010 | 0.035 | 1.8E-03 | 5.7E-05 | 5.8E-05 | 8.9E-01 |
| GDAP1 | MDR | OS | 5 | 3305 | 0.024 | 0.073 | 1.9E-03 | 7.0E-05 | 7.8E-05 | 1.3E-02 |
| XRCC3 | MDR | DR | 10 | 5581 | 0.025 | 0.074 | 2.4E-03 | 4.9E-05 | 5.7E-05 | 1.3E-05 |
| RAD54B | MDR | DR | 25 | 13390 | -0.004 | 0.021 | 3.8E-03 | 1.7E-05 | 1.8E-05 | 2.7E-01 |
| IRS4 | MF | NS | 8 | 7081 | 0.019 | 0.047 | 4.1E-03 | 2.4E-05 | 2.5E-05 | 2.6E-01 |
| PXDN | MDR | OS | 68 | 17845 | 0.017 | 0.035 | 5.7E-03 | 1.5E-05 | 1.6E-05 | 1.5E-01 |
| ATP5B | MF | OP | 6 | 2370 | 0.018 | 0.043 | 5.9E-03 | 1.3E-04 | 1.4E-04 | 1.6E-01 |
| HSP90AA1 | SA | H | 15 | 2769 | 0.014 | 0.043 | 7.8E-03 | 1.1E-04 | 1.1E-04 | 1.5E-02 |
| NR4A1 | MF | | 14 | 7958 | 0.023 | 0.058 | 8.0E-03 | 3.1E-05 | 3.4E-05 | 6.0E-02 |
| NRAS | MF | NS | 4 | 7767 | 0.007 | 0.043 | 1.4E-02 | 4.0E-05 | 4.3E-05 | 1.0E-01 |
| HSPA2 | SA | H | 6 | 2349 | -0.001 | 0.014 | 1.8E-02 | 5.9E-05 | 5.9E-05 | 8.9E-01 |
| CYP4F22 | MF | M | 26 | 1987 | 0.018 | 0.030 | 1.9E-02 | 1.1E-04 | 1.1E-04 | 8.7E-02 |
| NDUFB3 | MF | OP | 1 | 2144 | -0.020 | 0.045 | 2.1E-02 | 1.8E-04 | 2.1E-04 | 1.7E-02 |
| FOXA3 | MF | NS | 7 | 2495 | -0.003 | 0.013 | 2.5E-02 | 3.3E-05 | 3.6E-05 | 3.4E-01 |
| HSPH1 | SA | H | 8 | 18513 | -0.002 | 0.023 | 2.5E-02 | 1.6E-05 | 1.7E-05 | 3.6E-01 |
| PRDX6 | MDR | OS | 14 | 1217 | 0.009 | 0.019 | 2.6E-02 | 1.4E-04 | 1.4E-04 | 8.4E-01 |
| CCS | MDR | OS | 7 | 1231 | 0.028 | 0.068 | 2.9E-02 | 2.0E-04 | 2.2E-04 | 2.5E-02 |
| MYC | MDR | OS | 21 | 4990 | 0.011 | 0.020 | 3.0E-02 | 5.4E-05 | 5.4E-05 | 7.3E-01 |
| GADD45B | MDR | | 9 | 1316 | 0.044 | 0.087 | 3.5E-02 | 1.4E-04 | 1.6E-04 | 4.1E-01 |
| NDUFA13 | MF | OP | 15 | 598 | 0.009 | 0.025 | 4.4E-02 | 2.6E-04 | 2.6E-04 | 9.1E-01 |
| HSPA8 | SA | H | 31 | 2236 | 0.019 | 0.032 | 4.4E-02 | 7.4E-05 | 7.6E-05 | 6.0E-01 |
| RIT1 | MF | NS | 5 | 3455 | -0.003 | 0.009 | 5.6E-02 | 1.2E-04 | 1.2E-04 | 4.1E-03 |
| PIK3R1 | SA | HY | 7 | 33008 | 0.038 | 0.069 | 6.3E-02 | 8.4E-06 | 9.3E-06 | 1.8E-02 |

Table 4.S3: Nonsynonymous SNPs with Significant Interpopulation-Ecotype Divergence

| Human Symb | Scaffold | Pos | Category | Sub-Category | Ref | Alt | AA | AA Pos | AC-L | AC-M | AF-L | AF-M | Δ AF | $\overline{F_{ST}}$ Within | $\overline{F_{ST}}$ Between | F_{ST} p-val | $\overline{D_{XY}}$ Within | $\overline{D_{XY}}$ Between | D_{XY} p-val | SIFT Score | SIFT Med | SIFT n |
|------------|-------------|-----------|----------|--------------|-----|-----|-----|--------|------|------|------|------|-------------|----------------------------|-----------------------------|----------------|----------------------------|-----------------------------|----------------|------------|----------|--------|
| CYP4F22 | NC.045542.1 | 876503 | MF | M | G | A | R/C | 266 | 77 | 35 | 0.31 | 0.15 | 0.16 | 0.012 | 0.044 | 1.9E-02 | 0.339 | 0.368 | 2.7E-01 | 0 | 3.16 | 398 |
| CYP2W1 | NC.045554.1 | 1810567 | MF | M | G | C | L/V | 155 | 82 | 61 | 0.37 | 0.29 | 0.08 | -0.012 | -0.002 | 5.1E-03 | 0.432 | 0.441 | 3.5E-01 | 0.63 | 2.51 | 179 |
| FOXA3 | NC.045552.1 | 1884906 | MF | NS | C | T | V/I | 71 | 36 | 8 | 0.14 | 0.03 | 0.11 | 0.001 | 0.018 | 9.5E-03 | 0.146 | 0.154 | 7.1E-01 | 0.06 | 2.76 | 206 |
| RORC | NC.045557.1 | 3749701 | SA | HY | A | T | S/T | 442 | 155 | 109 | 0.69 | 0.5 | 0.19 | -0.047 | -0.036 | 7.4E-01 | 0.465 | 0.498 | 5.9E-04 | 0.23 | 2.7 | 373 |
| CCS | NC.045557.1 | 9246067 | MDR | OS | A | G | K/R | 262 | 136 | 79 | 0.57 | 0.34 | 0.23 | 0.032 | 0.074 | 7.6E-02 | 0.467 | 0.520 | 1.5E-03 | 0.05 | 2.74 | 211 |
| PRDX6 | NC.045551.1 | 12002316 | MDR | OS | G | T | L/I | 121 | 6 | 27 | 0.02 | 0.11 | 0.09 | 0.014 | 0.048 | 7.8E-03 | 0.142 | 0.153 | 6.9E-01 | 0.15 | 3.1 | 398 |
| HSPH1 | NC.045546.1 | 24404921 | SA | H | A | G | I/T | 506 | 26 | 68 | 0.11 | 0.3 | 0.19 | -0.004 | 0.019 | 3.9E-02 | 0.320 | 0.346 | 2.4E-01 | 0.64 | 3.45 | 394 |
| HSPH1 | NC.045546.1 | 24412139 | SA | H | T | A | I/F | 208 | 28 | 71 | 0.12 | 0.32 | 0.2 | -0.005 | 0.019 | 3.0E-02 | 0.338 | 0.372 | 1.8E-01 | 0.74 | 3.45 | 394 |
| PYROXD2 | NC.045550.1 | 2921051 | MDR | OS | T | C | H/R | 475 | 116 | 176 | 0.48 | 0.77 | 0.29 | 0.024 | 0.085 | 1.1E-03 | 0.399 | 0.475 | 1.5E-02 | 0.16 | 2.82 | 383 |
| PYROXD2 | NC.045550.1 | 25925545 | MDR | OS | T | G | M/L | 352 | 111 | 167 | 0.48 | 0.77 | 0.29 | 0.000 | 0.082 | 1.5E-03 | 0.398 | 0.473 | 1.5E-02 | 1 | 2.82 | 397 |
| NR4A1 | NC.045542.1 | 28310789 | MF | OP | C | T | V/I | 87 | 102 | 34 | 0.43 | 0.16 | 0.27 | 0.037 | 0.098 | 1.0E-02 | 0.364 | 0.437 | 1.7E-02 | 0.09 | 2.65 | 339 |
| TSC1 | NC.045556.1 | 33998360 | MF | NS | T | A | M/L | 40 | 121 | 88 | 0.55 | 0.41 | 0.14 | -0.048 | -0.038 | 7.2E-01 | 0.495 | 0.506 | 3.1E-03 | 0.15 | 3.01 | 337 |
| HSPA13 | NC.045546.1 | 45589746 | SA | H | T | C | E/G | 340 | 37 | 70 | 0.15 | 0.31 | 0.16 | 0.000 | 0.054 | 8.7E-05 | 0.337 | 0.383 | 1.1E-01 | 0.23 | 2.75 | 101 |
| HSPA13 | NC.045546.1 | 45591271 | SA | H | T | C | K/R | 154 | 53 | 12 | 0.22 | 0.05 | 0.17 | -0.008 | 0.033 | 1.1E-04 | 0.198 | 0.217 | 5.1E-01 | 0.14 | 2.47 | 277 |
| SDHD | NC.045553.1 | 48077603 | MF | OP | C | T | T/I | 57 | 98 | 132 | 0.43 | 0.63 | 0.2 | 0.002 | 0.023 | 4.7E-02 | 0.485 | 0.510 | 8.3E-03 | 0.19 | 2.92 | 172 |
| SDHD | NC.045553.1 | 48080392 | MF | OP | A | G | T/A | 114 | 162 | 169 | 0.74 | 0.82 | 0.08 | -0.012 | 0.012 | 7.0E-03 | 0.329 | 0.342 | 5.4E-01 | 0.18 | 2.8 | 305 |
| INPP5A | NC.045555.1 | 48843477 | MF | NS | C | T | R/K | 214 | 21 | 4 | 0.09 | 0.01 | 0.08 | 0.005 | 0.054 | 1.4E-02 | 0.093 | 0.099 | 7.6E-01 | 0.05 | 2.77 | 399 |
| RAD54B | NC.045548.1 | 53660338 | MDR | DR | C | G | E/D | 536 | 105 | 141 | 0.46 | 0.64 | 0.18 | 0.005 | 0.048 | 8.3E-05 | 0.465 | 0.520 | 4.4E-05 | 0.24 | 2.25 | 343 |
| RAD54B | NC.045548.1 | 53664004 | MDR | DR | T | C | M/V | 308 | 41 | 12 | 0.18 | 0.05 | 0.13 | 0.015 | 0.061 | 1.9E-02 | 0.178 | 0.196 | 5.1E-01 | 0.5 | 2.27 | 343 |
| RAD54B | NC.045548.1 | 53668476 | MDR | DR | T | C | N/S | 242 | 41 | 12 | 0.18 | 0.05 | 0.13 | 0.014 | 0.062 | 1.6E-02 | 0.178 | 0.196 | 5.1E-01 | 0.88 | 2.32 | 306 |
| RAD54B | NC.045548.1 | 53672176 | MDR | DR | C | T | A/T | 104 | 41 | 13 | 0.18 | 0.06 | 0.12 | 0.011 | 0.062 | 1.4E-02 | 0.170 | 0.188 | 5.1E-01 | 0.58 | 2.29 | 241 |
| MYC | NC.045548.1 | 69067706 | MDR | OS | C | T | A/V | 147 | 45 | 20 | 0.25 | 0.11 | 0.14 | 0.000 | 0.016 | 3.1E-02 | 0.240 | 0.246 | 7.9E-01 | 0.06 | 2.56 | 172 |
| NDUFB1 | NC.045541.1 | 69298519 | MF | OP | C | T | R/C | 270 | 22 | 2 | 0.08 | 0 | 0.08 | 0.011 | 0.031 | 1.2E-02 | 0.077 | 0.081 | 8.4E-01 | 0 | 3.84 | 399 |
| IGF2 | NC.045541.1 | 80157313 | MF | NS | G | C | T/S | 158 | 92 | 26 | 0.38 | 0.11 | 0.27 | 0.016 | 0.055 | 4.0E-03 | 0.316 | 0.352 | 2.7E-01 | 1 | 2.54 | 213 |
| PIK3C2A | NC.045541.1 | 83288816 | SA | HY | T | A | Y/N | 584 | 59 | 35 | 0.27 | 0.17 | 0.1 | 0.026 | 0.068 | 3.7E-02 | 0.293 | 0.315 | 4.8E-01 | 1 | 3.05 | 339 |
| PIK3C2A | NC.045541.1 | 83291589 | SA | HY | A | T | Y/F | 682 | 25 | 11 | 0.1 | 0.05 | 0.05 | -0.001 | 0.030 | 1.5E-02 | 0.122 | 0.127 | 7.4E-01 | 0.03 | 3.06 | 374 |
| PIK3C2A | NC.045541.1 | 83302247 | SA | HY | A | C | I/L | 1049 | 25 | 11 | 0.11 | 0.05 | 0.06 | 0.002 | 0.033 | 4.4E-02 | 0.126 | 0.132 | 7.4E-01 | 0.69 | 3.05 | 393 |
| CYB5R2 | NC.045541.1 | 89321298 | MF | M | G | A | V/I | 8 | 13 | 33 | 0.05 | 0.14 | 0.09 | 0.001 | 0.031 | 1.5E-02 | 0.189 | 0.200 | 6.5E-01 | 0.79 | 3.23 | 208 |
| MPO | NC.045544.1 | 118637004 | MDR | OS | G | A | G/S | 12 | 64 | 16 | 0.29 | 0.07 | 0.22 | 0.023 | 0.085 | 4.5E-03 | 0.250 | 0.294 | 2.1E-01 | 0.08 | 2.76 | 196 |
| MAP3K1 | NC.045543.1 | 121713512 | MF | NS | C | T | G/S | 1002 | 135 | 79 | 0.57 | 0.35 | 0.22 | 0.012 | 0.040 | 2.7E-02 | 0.477 | 0.512 | 7.9E-03 | 1 | 3.75 | 184 |
| NUP88 | NC.045544.1 | 122777754 | SA | GS | C | T | T/M | 502 | 46 | 25 | 0.19 | 0.11 | 0.08 | 0.000 | 0.020 | 1.2E-03 | 0.225 | 0.238 | 5.2E-01 | 0.22 | 3.06 | 141 |
| NFKBIA | NC.045541.1 | 128350420 | SA | GS | T | C | N/S | 114 | 61 | 14 | 0.25 | 0.06 | 0.19 | 0.005 | 0.055 | 1.9E-05 | 0.214 | 0.245 | 3.1E-01 | 0.04 | 2.46 | 257 |
| ATP5A1 | NC.045543.1 | 142028312 | MF | OP | G | A | V/I | 543 | 56 | 33 | 0.22 | 0.13 | 0.09 | -0.004 | 0.008 | 4.2E-02 | 0.296 | 0.305 | 6.3E-01 | 1 | 3.74 | 378 |
| XRCC3 | NC.045541.1 | 149009348 | MDR | DR | C | T | R/Q | 142 | 120 | 53 | 0.55 | 0.25 | 0.3 | 0.062 | 0.122 | 4.4E-02 | 0.419 | 0.515 | 7.9E-04 | 0.71 | 2.39 | 352 |
| HIF1A | NC.045541.1 | 154070464 | SA | HY | C | A | A/S | 492 | 160 | 76 | 0.65 | 0.33 | 0.32 | 0.012 | 0.063 | 8.5E-04 | 0.473 | 0.532 | 1.4E-05 | 0.35 | 2.67 | 281 |
| INSR | NC.045541.1 | 166109368 | MF | NS | C | G | G/A | 1025 | 62 | 93 | 0.25 | 0.4 | 0.15 | 0.001 | 0.020 | 2.5E-02 | 0.441 | 0.461 | 2.6E-01 | 0.35 | 3.53 | 400 |
| ATP5D | NC.045541.1 | 175086965 | MF | OP | A | G | N/D | 144 | 92 | 136 | 0.36 | 0.58 | 0.22 | 0.024 | 0.088 | 4.3E-03 | 0.464 | 0.544 | 3.5E-06 | 1 | 2.67 | 377 |
| GADD45B | NC.045541.1 | 186466105 | MDR | C | G | C | P/A | 158 | 19 | 56 | 0.08 | 0.25 | 0.17 | 0.024 | 0.101 | 2.1E-03 | 0.312 | 0.360 | 1.9E-01 | 0.05 | 3.05 | 313 |

Chapter 5: Summary and Concluding Remarks

While the reader of this dissertation may find the goals and methods within each data chapter diverse from one another, it should be recognized that the mitochondrion is a consistent character within each study. I examined hypotheses that implicated mitochondria as players driving patterns in lineage diversification (Chapter 2), aerobic performance (Chapter 3), and ecotype divergence (Chapter 4). In this final chapter I review the historical connections between mitochondriology (the study of mitochondria) and evolutionary biology, I summarize how research findings from the data chapters impact society as a whole, and I discuss contributions of this dissertation to our current understanding regarding the role of mitochondria in driving patterns in evolution.

5.1 History and Current Understanding of Mitochondriology

Perceived importance of the mitochondrion to life and the evolution thereof has fluctuated over the past two centuries. Their initial discovery was so underappreciated that it is difficult to assign an exact time (sometime in the 1850's) or person (Lehninger 1964), and the lack of a widely-broadcasted description resulted in scientists referring to the same structure with variety of names over the next half century (sarcosomes, fila, blepharoblasts, chondriokonts, chondriomites, chondrioplasts, chondriosomes, chondriospheres, fuchsinophilic granules, interstitial bodies, Körner, Fädenkörner, mitogel, parabasal bodies, plasmasomes, plastochondria, plastosomes, spheroblasts, vermicules, bioblasts, mitochondria, and others). However, occasionally a scientist during the late 19th / early 20th century would hypothesize regarding their crucial function, causing an ephemeral buzz in the scientific community (Fig. 5.1); Richard

Altmann referred to them as the ultimate “elementary living particles” (Altmann 1894), Friedrich Meves contended they were bearers of heredity characteristics (Meves 1908), Benjamin Kingsbury suggested that they possessed respiratory capabilities (Kingsbury 1912), Paul Portier hypothesized that they were bacterial ‘symbiotes’ living within every animal cell (Portier 1918), and Ivan Wallin claimed they play a fundamental role in speciation (Wallin 1927). Yet the lack of widespread acceptance and examination of these organelles, potentially caused by insufficient interest within the general scientific community, resulted in a relatively limited focus on mitochondria for several decades (Figure 5.1).

Recognition of mitochondria as critical players in physiology began in the late 1940’s with interdisciplinary discussion between cytologists, biochemists, and enzymologists (Lehninger 1964) who discovered the localization of respiratory enzymes exclusively within the mitochondrion (Kennedy and Lehninger 1948). Subsequent work identified the mitochondrion as the site of ATP production via oxidative phosphorylation, a process driven by a hydrogen ion gradient generated from an electron transport system (Mitchell 1961; Criddle et al. 1962; Hatefi et al. 1962). Although these findings were originally met with much skepticism, with time each was completely accepted (Lehninger 1964; Tzagoloff 1982). Frequency of mitochondrial-involved research began to increase around this time (Fig 5.1), with the majority of these studies conducted within the fields of physiology and biochemistry.

In an entirely separate field whose early work rarely interdigitated with that of physiology, patterns consistent with non-mendelian inheritance (Baur 1908; Correns 1909) led geneticists to accept that “certain forms of inheritance are the outcome of self-perpetuating bodies in the

cytoplasm” (Morgan 1919). Despite this spot-on claim written by renowned geneticist Thomas Hunt Morgan, he later stated that, based on the “rare cases” of cytoplasmic inheritance, “the cytoplasm may be ignored genetically” (Morgan 1926). Yet as the accumulation of evidence for cytoplasmically inherited traits increased, several scientists over the next quarter century refused to accept this statement (East 1934; Wright 1941; Ephrussi 1949). During the middle of the twentieth century, a larger group of biologists began to suggest the importance of the mitochondrion as a functional and heritable intracellular entity (Sonneborn 1950; Newcomer 1951; Lederberg 1952). In 1963 Margit and Sylvan Nass discovered the presence of DNA within the mitochondrion (Nass and Nass 1963), opening the door for collaboration between physiologists, population geneticists, and phylogeneticists. Just four years later, Lynn Margulis published a landmark study describing the endosymbiotic nature of eukaryotes (Sagan 1967).

Following the discovery of mitochondrial DNA and a description of their prokaryotic ancestry, geneticists and molecular biologists dissected and characterized the mitochondrial genome in yeast for the next couple of decades (Tzagoloff 1982). These studies, along with the mitochondrial physiological studies that took off in the 50s, resulted in a dramatic increase in publications on the mitochondrion (Fig 5.1).

Once the power of nucleotide sequence data became available to evolutionary biologists, genes in the mitochondria were found to be useful genetic markers (Avice 2012). Those who utilized coalescent theory recognized the usefulness of a non-recombining, haploid, maternally inherited unit with a relatively elevated mutation rate (in the case of most metazoans) for estimating evolutionary history. The abundance of mitochondrial DNA is much greater than that of the

nucleus, making it much easier to extract and amplify with PCR. Additionally, the presence of conserved regions of the mitochondrial genome adjacent to variable sites allowed for efficient primer design at informative areas (Ladoukakis and Zouros 2017). This led to widespread use during the 1990's in studies of population genetics and systematics (Vigilant et al. 1991; Rubinoff and Holland 2005; Figure 5.1), but interestingly few publications focused on additional questions regarding the interactive nature of the mitochondrial and nuclear genomes at the functional and coevolutionary level (in fact, it seemed that many evolutionary biologists simply looked at genes within the mitochondrial genome as useful genetic markers without considering function at all). Further, physiologists continued to provide a more detailed picture of the molecular underpinnings of cellular respiration, but rarely were broad-scale connections made regarding evolution. Somehow it seemed the perspectives of the early and late 20th century were completely flipped— the broad scope of early scientists focused on the forest, whereas more recent scientists focused on the trees.

In the 21st century, recognition of the mitochondrion as more than just an ATP-producing organelle with a neutrally-evolving genome has widened our understanding of mitochondrial biology (Towarnicki and Ballard 2020). This shift was catalyzed by an association of mitochondria with various traits and functions, including innate immune response (Wang et al. 2011), intracellular signaling (Rizzuto et al. 2012), reactive oxygen species production (Hirst et al. 2008), and programmed cell death (Bossy-Wetzel et al. 1998), in addition to a more thorough understanding of mitochondrial respiration via oxidative phosphorylation. Medical researchers also discovered associations between multiple diseases and mitochondrial mutations (Singh et al. 1989; Agostino et al. 2003; McFarland et al. 2007; Koopman et al. 2012; Lightowers et al.

2015). These findings, along with a more vivid picture of the nuclear and mitochondrial genomes in several organisms, allowed integrative biologists to hypothesize a central role of the mitochondrion in driving biological patterns (Lane and Martin 2010; Lane 2014), including the three central themes of this dissertation: (1) speciation (Hill 2016, 2017; Visinoni and Delneri 2022), (2) sexual reproduction (Hörandl and Hadacek 2013; Havird et al. 2015; Radzvilavicius and Blackstone 2015; Speijer 2015; Garg and Martin 2016), and (3) aging (Short et al. 2005; Schulz et al. 2007; López-Otín et al. 2013; Kang et al. 2016; Sun et al. 2016). For the remainder of this concluding chapter I will focus on implications and future directions of my chapters that addressed these themes in the context of the findings described in their respective data chapters (chapters two, three, and four), and at the end (in section 5.5) I will summarize how this dissertation contributes to our understanding of the mitochondrion's role in contributing to the evolution of these biological patterns.

5.2 Riverine Barriers as Potential Drivers of Lineage Diversification in Indochina

5.2.1 Relevance to Biodiversification Hypotheses and Conservation Efforts

Biodiversity benefits humanity, as humanity has learned step by step through breakthroughs in biomedicine and engineering. Whether or not one recognizes the inherent value of earth's organisms (Curry 2011), quantifying existing biodiversity is critical to understand the severity of our current mass extinction (Ceballos et al. 2020) and assess the future impact on vulnerable taxa (Costello et al. 2013). The study presented in Chapter 2 provides novel empirical support for the riverine barrier hypothesis and suggests a potential mechanism for the remarkable species richness generated in-situ within this biodiversity hotspot. We discovered multiple, independent lineages within the flying lizard *Draco maculatus* species complex, and the discovery of this

hidden diversity may be relevant to conservation efforts. These lineages are potential species that have yet to be described, and we cannot protect what we don't know exists. Identifying these lineages is a first step in the process of assessing whether any require conservation efforts.

5.2.2 Paleo-Rivers and/or Paleo-Niches Driving Divergence: Suggestions for Future Work

I make recommendations in Chapter 2 regarding future directions for this work, including sampling from contact zones of the divergent lineages in the study, collecting genome-wide sequencing data (e.g., RAD-seq) for more accurate phylogenetic estimation, and measuring fine-scale morphological data for species descriptions. Here I make an additional recommendation regarding the riverine barrier hypothesis as an agent of isolation within Indochina involving paleo-river and paleo-niche reconstructions.

My results from Chapter 2 are published in the journal *Molecular Phylogenetics and Evolution* (Klabacka et al. 2020). During the peer-review process, one reviewer recommended changing the manuscript title to “Ancient rivers drive lineage diversification of flying lizards in tropical Indochina.” However, I settled with the final title of “Rivers of Indochina as *potential* drivers of lineage diversification in spotted flying lizards.” We retained the word “potential” since this study doesn't provide direct evidence that the rivers themselves are driving diversification (rather that they coincide with phylogenetic breaks). We omitted the word “Ancient” since we don't test the historical paths of these rivers, only the contemporary paths. Because of this, I recommend direct testing whether paleo-river dynamics correspond with evolutionary history and making comparisons with niche evolution to determine whether rivers are true barriers or if they simply

co-vary with changes in habitat distribution. This can be achieved by estimating evolutionary history, reconstructing ancestral niches, and testing paleo-river hypotheses.

The riverine barrier hypothesis suggests speciation resulting from vicariance via river formation (allopatry) or dispersion with subsequent isolation (peripatry). Two hypotheses of geographic river history in Indochina suggest either of these mechanisms of biodiversification (Figure 5.2)—peripatry in the Ancient Paleo-river Hypothesis, which suggests the rivers of Indochina have been in the same location since at least the middle Miocene (Workman 1975; Attwood and Johnston 2001; Clark et al. 2004; Jamaluddin et al. 2019), and allopatry in the Dynamic Paleo-river Hypothesis, which suggests most of the rivers in Indochina formed in or just before the Pliocene (Hallet and Molnar 2001; Bolotov et al. 2017; Nie et al. 2018; Wang et al. 2020).

An alternative explanation for lineage boundaries corresponding with riverine barriers focuses on the connectivity of suitable habitat. Systems where divergent populations with historically continuous habitats are bisected by rivers (Continuous Paleo-niche Hypothesis) lend support to biodiversification occurring directly as a result of riverine vicariance (Fig. 5.2). However, paleo-niche fragmentation where gaps roughly correlate geographically with rivers can cause population divergence that may appear directly caused by riverine barriers (Fragmented Paleo-niche Hypothesis), especially when current suitable habitat is unfragmented.

Using (1) phylogenetic divergence time estimation, (2) paleo-niche building, (3) ancestral state reconstruction, and (4) Bayes factor comparison of models built from hypotheses, the contributions of previously listed hypotheses to patterns of biodiversity could be elucidated.

However, because information gained from approaches 1–3 would be used to construct models for approach 4, independently analyzed datasets should be used for approaches 1–3 (Dataset 1) and approach 4 (Dataset 2) to avoid multiple use of data.

The species assignments for each model for approach 4 could be based on (1) riverine barriers from the Ancient Paleo-river Hypothesis, (2) riverine barriers from the Dynamic Paleo-river Hypothesis, or (3) suitable habitat that was historically fragmented (Fig. 5.2). In addition to having species assignments, models could also include information on geographic connectivity over time from approaches 1-3 with Dataset 1 (such as separation due to river formation or distribution fragmentation). Thus models would include both extant lineage groupings and historical constraints based on past events that promote/suppress the potential for divergence. To determine the model that best explains the current distribution of genomic diversity across the range of *Draco maculatus*, the marginal likelihood can be estimated for each model and support can be quantified for the best model using Bayes factors.

If approaches 1–3 indicate (A) divergence time corresponds to proposed recent river capture and (B) no evidence for niche fragmentation (or if niche fragmentation is evident, it does not correspond with divergence times), this would support the Dynamic Paleo-river Hypothesis. Alternatively, if results from approaches 1–3 show (A) divergence times do not correspond with proposed recent river capture and (B) no evidence for niche fragmentation (or if niche fragmentation is evident, it does not correspond with divergence times), this would support the Ancient Paleo-river Hypothesis. Results from approaches 1–3 with evidence that niche fragmentation corresponds to divergence times would support the Fragmented Paleo-niche

Hypothesis. Marginal likelihood estimates could then be obtained using Dataset 2 for pairwise model comparison.

As a last note on recommendations for future investigation, replication across taxa with distributions spanning Indochina would provide additional support for the riverine barrier hypothesis as a major agent of biodiversification within Indochina. For example, an examination of multiple taxa across these putative barriers, potentially testing for shared divergence times (such as with the software EcoEvolity (Oaks 2019)) would provide a useful perspective on how river capture affects diversification across taxa.

5.2.3 Relevance of Lineage Diversification Research to Society

Looming over my time as a PhD student, the global COVID-19 pandemic shows the relevance of this work to science and the public on a broad scale. The general population has had a challenging wake-up call regarding lineage diversification; since early 2020, we have experienced the rapid dispersion and radiation of a novel human-vectored coronavirus, Sars-coV2. This resulted in a global pandemic of a dangerous disease, COVID-19. Humankind has witnessed the power of evolution as a force, and those without previous knowledge in molecular evolution have become familiar with the terms “mutation” and “variant”. Understanding the broad processes driving evolution (e.g., vicariance in this dissertation chapter) is critical to defending humankind from biological pathogens and misinformation dissimilated during pandemic.

5.3 Reduced Mitochondrial Respiration in Hybrid Asexual Lineages

The study presented in Chapter 3 provides empirical evidence for reduced physiological performance in hybrid asexual lineages compared to their parental sexual species. On a broader scale, it also contains implications for understanding why the vast majority of vertebrate organisms reproduce sexually. If a shift from sexual reproduction to asexual reproduction results in a decrease in performance that negatively impacts fitness, this may explain the scarcity of parthenogenesis as a primary mode of reproduction across vertebrates.

5.3.1 Implications for Mitonuclear Ecology

In Chapter 2, I suggest that reduced mitochondrial respiration in hybrid asexual lizards relative to their sexual parental species may be a result of genomic incompatibility. Given the hybrid ancestry of these organisms, this incompatibility may be inter-specific (between the divergent genomes of the parental species). These interspecific interactions may occur between nuclear genomes or between the paternal nuclear genome and the maternal mitochondrial genome. Alternatively, given the lack of recombination in an asexual lineage, the incompatibility may also be intra-specific (within the genome of a parental species). The mechanism of this latter hypothesis is based on the mutational erosion principle, wherein due to the inability to purge deleterious mutations via genetic recombination, asexual species retain these mutations within the genomes of their descendants. The continuation of this process results in the compiling acquisition of deleterious mutations, a.k.a. Muller's Ratchet. I deem this intraspecific incompatibility given that the incompatible interactions can occur within the genome of a single parental species. More specifically, these interactions could be between (A) genes of the nuclear genome or (B) those of the nuclear genome and those of the mitochondrial genome, although it is worth noting that this process could also affect interactions between the divergent parental

genomes. A process that may explain intraspecific incompatibility between the nuclear and mitochondrial genomes invokes both Muller's Ratchet (described above) and the Red Queen Hypothesis (the requirement of organisms to use sexual recombination to "keep up" with fast-evolving parasites). I've deemed this process the "Red Ratchet" Hypothesis.

Before describing the Red Ratchet hypothesis, it is important to note some fundamentals of mitonuclear ecology. First, in bilaterian animals, the mitochondrion has a higher substitution rate than the nuclear genome. Second, the mitochondrial genome acquires deleterious mutations. Third, sexually-reproducing organisms can utilize variation within their population's allele pool to recombine nuclear mutations that mitigate the deleterious effects of the mitochondrial mutation (a process known as "compensatory coevolution").

In the Red Ratchet Hypothesis, intergenomic incompatibility arises gradually as the coadaptation between intraspecific nuclear and mitochondrial genomes deteriorates via accumulation of deleterious mutations in the mitochondrial genome (Muller's Ratchet) and the inability of the nuclear genome to respond to said deleterious mutations via sexual recombination (Red Queen). If this is the primary driver of reduced mitochondrial respiration in hybrid asexual lineages, then younger lineages would be predicted to have a less-severe effect size compared to older lineages. It is worth noting that none of these hypotheses are mutually exclusive, but they can be somewhat disentangled by integrating mitochondrial physiology and targeted genomics. I describe some of these approaches below:

5.3.1.1 *Time Machine*

Ideally, identifying the contributions of interspecific and intraspecific (including Red Ratchet) incompatibilities to reduced mitochondrial respiration would involve an experimental design where a researcher could compare the parental sexual species with the hybrid asexuals from the timepoint of hybridization and compare them with today. Figure 5.3 provides a visual schematic of this. Samples from each lineage could be taken from the time of the hybridization event (T1) and today (T2), and a reaction norm could be created using a measure of mitochondrial function as a response variable.

Predicted responses in the context of each hypothesis are shown in Figure 5.3. Interspecific incompatibility would show reduced mitochondrial function at the time of hybridization, and this shouldn't change over time based on the frozen genome hypothesis (although it is possible that gene conversion could affect this [Hillis et al., 1991; Warren et al., 2018]). Intraspecific incompatibility would show no reduction in mitochondrial function at the time of hybridization (T1), but over time the accumulation of deleterious mutations with lack of compensation would result in reduced mitochondrial function at T2. A combination of these two hypotheses would result in both reduced mitochondrial function at the time of hybridization (T1) and a future reduction by T2. Hybrid vigor is essentially the opposite of interspecific incompatibility; essentially the high heterozygosity acquired at hybridization (T1) results in favorable allele combinations that boost mitochondrial function. Lastly, vigor + intraspecific incompatibility would result in an increased mitochondrial function at the time of hybridization (T1), but this would gradually decrease with the acquisition of deleterious mutations (T2). To repeat a previous point, results consistent with predictions of intraspecific incompatibility could be due either to

nuclear-nuclear interactions or nuclear-mitochondrial interactions. Disentanglement of these requires further approaches, some of which are described below.

5.3.1.2 *Complex Activity Assays*

One approach is to use complex activity assays to examine the efficiency of electron transport for each protein complex of the electron transport chain. Because all of the protein complexes except for succinate dehydrogenase (CII) are made up of both mitochondrial and nuclear gene products, comparing the effect sizes between protein complexes can shed light on the contributions of different hypotheses (Fig. 5.4). A scenario with reductions in enzyme efficiency of consistent effect size at all five complexes could be due to nuclear-nuclear incompatibilities, since every complex contains nuclear-encoded protein subunits. Alternatively, if all complexes have reduced enzyme efficiency except for CII, this would provide indirect evidence for mitonuclear incompatibility in genes coding for transcription or translation (i.e., ribosomal proteins [encoded by the nucleus] and rRNA [encoded by mitochondrion]). There is another scenario where both of these occur, wherein all complexes show reduced enzyme efficiency, but the reduction is less in CII compared to the other five complexes. Scenarios with alternative combinations of enzyme efficiency (e.g., reductions in activity of one or two complexes) would indicate incompatibility within the subunits themselves, which could be due to nuclear-nuclear, mito-mito, or mitonuclear interactions. All of the above scenarios assume the presence of genetic variation at or upstream of the ETS subunits. The precise genetic source of the reduced enzyme efficiency, if detected, can then be elucidated by examining genetic variation at the gene products involved with mitochondrial function (i.e., genes encoding RNA polymerase, tRNA synthetase, tRNA, mitochondrial ribosomal proteins, mitochondrial rRNA, and ETS subunits).

5.3.1.3 *Selection Ratios*

The Red Queen hypothesis poses that genetic recombination utilizes variants available in the gene pool to mix and match genotypes to keep pace with the high mutation rates of pathogens, thus sexual populations can draw from the pool of variants available within the population rather than only those that arise within their genome. However, it is worth noting that the necessity of recombination to compensate for pathogen evolution is only critical in the genes that interact (directly or indirectly) with the pathogen. Because of this, genes of the immune system in sexual species have higher rates of substitution compared to those in other regions of the genome. Similarly, gene products of the nuclear genome that interact with the gene products of the fast-evolving mitochondrial genome are shown to have evidence of higher positive selection compared to similar genes that only interact with nuclear-encoded gene products (Barreto et al. 2018). These genes include ribosomal proteins, Amino-acyl tRNA synthetases, and the protein complexes of the electron transport chain. To test if intragenomic mitonuclear incompatibilities (a result of the red ratchet hypothesis) may be occurring, the ratio of D_n/D_s for nuclear gene products that interact with other nuclear gene products (n) can be compared to that of nuclear gene products that interact with mitochondrial gene products (m). We'll refer to D_n/D_s as ω , with ω_n referring to the value calculated from n and ω_m referring to the value calculated from m. We can then compare the ω_n / ω_m ratio (hereafter Ω) of these values between sexually reproducing lineages (Ω_S) and asexually reproducing lineages (Ω_A), with the prediction that the ratio of the sexual lineages will be greater than that of the asexual lineages ($\Omega_S > \Omega_A$).

5.3.1.4 *Functional Genomics*

In addition to population genetics examination such as that described in 1.3.1.3, functional examination of the interacting gene sequences of nuclear and mitochondrial genomes can shed light on the genetic mechanisms underlying the observed reduction in mitochondrial function. This can be achieved by sequencing the mitochondrial genomes and the nuclear-encoded genes that interact with mitochondrial gene products for both sexual and asexual species and then examining sequence variation for these genes. If any variants modify the peptide sequence (i.e., nonsynonymous SNPs), these can be examined for functional significance using statistical software packages (e.g., using SNP annotation software such as SIFT (Ng and Henikoff 2003), Polyphen-2 (Adzhubei et al. 2010), or PHACT (Kuru et al. n.d.)), by comparing protein stability between peptide variants (e.g., using the modeling software SWISS-MODEL (Waterhouse et al. 2018)), by investigating whether the peptide variant is near areas of interest (e.g., using structural model visualization software such as Chimera (Pettersen et al. 2004)), and by performing in-silico docking assays (e.g., using VMD software (Humphrey et al. 1996)). The tests examining genomic interactions can be conducted between (1) gene products of the same genome (paternal or maternal) to assess intraspecific compatibility (including between nuclear and mitochondrial gene products), and (2) gene products of the divergent genomes to assess interspecific incompatibility.

5.3.2 Mitochondrial Biology – Recognizing the Complete Picture

Readers making inferences from the results I presented in Chapter 3 (which show differences in mitochondrial respiration between parental sexual species and hybrid asexual lineages) should keep in mind that this examined only one facet of mitochondrial physiology (mitochondrial respiration). Further, they should also recognize that only a portion of the available approaches

for measuring mitochondrial respiration were implemented. To be more specific, we examined State 3 and State 4 Respiration through Complex I and through Complex II, providing four total measurements of mitochondrial respiration (CIS3, CIIS3, CIS4, and CIIS4). We used these values to calculate the respiratory control ratio (RCR), providing two additional metrics (CIRCR and CIIRCR). However, other approaches are available and would provide further detail into the differences between these groups. These approaches include oligomycin-modulated state 4 (State 4o; achieved by supplying a CV inhibitor rather than depending on natural ADP depletion), and uncoupled respiration (State 3u; achieved by supplying an uncoupler that permeabilizes the inner membrane). State 4o provides the true leak state; oxygen consumption is mediated solely by the leaking of protons across the inner membrane. State 3u provides the full capacity of electron transport; the rate of transport is not limited by ATP production. Additionally, quantification of the membrane potential and free radical production, either of which can be measured in conjunction with mitochondrial respiration on the Oroboros O2k-FluoRespirometer, provide yet another angle of mitochondrial physiology.

In addition to mitochondrial physiology, mitochondrial behavior and morphology also play a role in mitochondrial performance (Heine and Hood 2020). The quantity, size, structure, positioning, and connectivity of mitochondria affect organelle performance (Mannella 2006; Zick et al. 2009; Rafelski 2013). Rate of ATP production, for example, can be affected by cristae structure within the mitochondrion (Mannella et al. 2013; Nielsen et al. 2017). It is possible that the pattern we observed in Chapter 2 are due to differences in mitochondrial behavior and morphology; ruling this out through examination using electron microscopy examination would be useful in disentangling the underpinning forces behind our observed pattern.

5.3.3 Implications for Evolutionary Physiology

Although the mitochondrion is widely recognized as the agent responsible for producing the great majority of energy used for active cellular processes, the connection between mitochondrial respiration and aerobic performance within the context of evolutionary ecology is understudied. While variation in endurance performance among organisms has been observed, lacking is empirical evidence that directly links this variation to mitochondrial function. A Web of Science query used to find studies that examine the relationship between mitochondria and endurance in the context of either evolution or ecology (syntax: ALL="mitochondria*" AND "respiration" AND "endurance" AND ("evolution" OR "ecology") NOT ("exercise" OR "train*" OR "disease*" OR "athlet*")) yielded only our paper published in *The American Naturalist* from Chapter 2. Several related studies found using variations of this search criteria (e.g., excluding “respiration”, “endurance”, or “evolution”) include (1) an association between whole-organism metabolic rate and running in birds (Bundle et al. 1999), (2) an association between hydrophobicity of mitochondrial membrane proteins and aerobic capacity in tetrapods (Kitazoe et al. 2011), and (3) an association between endurance capacity and mitochondrial protein abundance (Wisløff et al. 2005) and respiratory capacity (Aon et al. 2021) in rats selected for high- and low-capacity running. Our study provides valuable insight into the connection between aerobic performance and cellular energy production, yet there is a need for future studies to examine this relationship in a broader phylogenetic context.

5.4 Divergence in Molecular Networks that Underly Aging

5.4.1 Relevance to Evolutionary Biology

Understanding genomic underpinnings of aging in natural populations provides insight to the evolution of life history strategies. All fields of biology are unified by evolution. The diverse forms of life on earth are explained through evolution by natural selection. And an understanding of natural selection is achieved through the lens of life history theory (Stearns 1992). To explain this in more detail, it is helpful to think of natural selection as a force that requires two connected characteristics in its subjects, which are (1) variation in heritable matter (genes) that cause (2) variation in reproductive success (fitness). Fitness is influenced by the underlying genes inherited from ancestors, but it is important to recognize that fitness is relative to the environment. In other words, the efficacy of an organism to reproduce varies depending on its surroundings. For example, a population with high prevalence of parasites may have a delayed age of sexual maturity due to an increased allocation of energy to combat parasites compared to a population with less parasites which reaches sexual maturity sooner. If the age of maturity is genetically determined, placing an individual from either population in the alternate environment could result in decreased fitness. The individuals and their genes haven't changed, but their environment has.

Consider a population residing in an environment with an abundance of food sources; we'll call this Env I. A few individuals from this population disperse into a neighboring region which has less food availability; we'll call this Env II. While individuals in Env II experience reduced access to food, they don't have to compete with the greater number of individuals in the source population. Now imagine that a mutation occurs in an individual of Env II that results in a decreased investment in growth and a reduced need for food uptake. This individual is not as dependent on a high concentration of food for survival, and therefore is more likely to survive to

reproductive maturity than the individuals with greater investment in growth. Selection acts on this differential fitness, resulting in an increase in frequency of the mutation among the Env II sub-population. However, this increase does not occur in the Env I sub-population, because the difference in food availability results in a change in fitness for individuals with this mutation. Therefore, a balance in allele frequencies (for the original and mutated versions of the gene) is reached between the two subpopulations, which will flutter due to migration and drift.

While this simple example has outlined how variation in a life history strategy might arise and become prevalent, it does not address a fundamental question relevant to both life history and evolutionary theory: What are the specific gene networks that underly this phenotypic variation, and is variation in the same gene networks found in independent populations with similar phenotypic divergence? In his seminal work, Stephen Stearns wrote that students of life history seek to explain the variation in reproductive traits (Stearns 1992). Within earlier frameworks of life history theory, genetic details were treated as a black box (Flatt and Heyland 2011). Yet a thorough explanation of variation in life histories requires investigation of intracellular processes, including the blueprint of said processes (DNA). Therefore, more specific questions for those interested in molecular life history evolution should include: Where did the mutation occur (what genomic region)? Did it occur within a region that encodes a product (e.g., protein, RNA)? If within a protein-coding gene, is it within an exon or an intron? If within an exon, does the mutation alter the amino acid? If it alters the amino acid, does the change result in a difference in polarity? Does the change alter the structure of the protein? Is it located at a region that interacts with other proteins? If it is not within the coding region of a gene, is it a gene expression regulatory region? Is the locus part of a gene network? How many gene products

interact with this gene and/or gene product (i.e., is it a top-regulator or at a key node)? Answers to these questions can help us understand the path from genotype to phenotype, which, as I mentioned at the beginning of this section, is central to our understanding of evolution (Lewontin 1974). On a broader scale, we can also examine the genomic sources of convergent evolution and determine whether the same gene networks, genes, and even locations within a gene underpin similarities in life history strategies across the tree of life. Addressing these questions using genomics is now feasible in non-model organisms due to the affordable cost of high-throughput sequencing, deposition of genomic resources in public databases, computational capacity of high-performance computers, and accessibility to software programs shared globally on the internet.

In Chapter 4, I presented differences in nucleotide sequences and gene expression patterns between two garter snake ecotypes with divergent life history strategies. The genes and gene networks with significant results reflect those associated with differences in life history (specifically aging) among other taxa. We see our findings as both (A) relevant for current understanding of life history evolution, providing a critical perspective from a natural population, and (B) instrumental in the collective effort of understanding whether general processes cause similar phenotypes across the tree of life.

5.4.2 Implications for Mitonuclear Ecology

The exact role of the mitochondrion in aging is not well understood- but the relationship between mitochondrial function and aging is undeniable (Son and Lee 2021). Several current hypotheses that seek to explain the evolution of aging include (1) antagonistic pleiotropy (alleles that benefit

the fitness of an organism early in life have a negative effect on survival later in life), (2) the disposable soma (allocation of energy to reproduction limits the amount available for repair), and (3) mutation accumulation (inefficient selection on individual old-body phenotypes due to extrinsic mortality). The bioenergetic efficiency of the mitochondrion throughout an organism's life is relevant to each of these hypotheses (Batalha et al. 2022). For example, reactive oxygen species, which originate at the mitochondrion, early in an organism's life are important molecular messengers and later in life become damaging intracellular molecules. Given that ATP is required for cellular maintenance, impaired ATP production at the mitochondrion would result in a reduced energy pool for reproduction and cell repair. And the mitochondrion acquires deleterious mutations whose damaging effects increase mortality.

Similarly, while the nature of the relationship between life history divergence and the mitochondrion in the terrestrial garter snakes is not well-defined, evidence for the existence of a relationship is well established (Schwartz et al. 2015; Gangloff et al. 2020). This published evidence to-date includes ecotypic divergence in mitochondrial respiration efficiency, ROS production, antioxidant expression, whole-organism metabolic rate, cellular oxygen consumption, mitochondrial gene expression, and mitochondrial sequence divergence (Bronikowski and Vleck 2010; Robert and Bronikowski 2010; Schwartz and Bronikowski 2013; Gangloff et al. 2015, 2020; Schwartz et al. 2015). The differences in mitochondrial gene expression between ecotypes along with the SNP in the mitochondrial genome that coincides with ecotype divergence are mirrored by differences in pathway-level expression and molecular divergence of the nuclear-encoded OXPHOS genes (see results in Chapter 4). I recommend future comparison of the evolutionary history of these nuclear-encoded OXPHOS genes with (A)

the mitochondrial genome and (B) the rest of the nuclear genome to identify whether patterns in life history are reflected by the phylogeny for genes involved with mitochondrial respiration. I predict that the evolutionary history of the nuclear-encoded OXPHOS genes of significance from Chapter 4 is more similar to that of the mitochondrial genome compared to the rest of the nuclear genome.

5.4.3 Relevance to Medicine

Whether aging is (Bulterijs et al. 2015) or isn't (Rattan 2014) a disease, its connection to increased mortality is undeniable. While human lifespan (average length of life) has increased significantly over the past century (Oeppen and Vaupel 2002), healthspan (average length of healthy life) has remained unchanged (Crimmins 2015; Olshansky 2018). A desire to maintain an active, healthy lifestyle late into the "third age" has created a growing public interest in understanding the processes responsible for aging and how to slow them. This has led to many pills and creams promising to reduce the effects of aging (usually with no scientific evidence) and multiple best-selling books focused on the causes of aging and how they can be mitigated (some written by prominent scientific researchers, many written by amateurs, and others written by charlatans). Knowledge of processes and development of treatments for aging and age-related disease requires the acquisition of basic knowledge regarding the genes and gene networks associated with aging.

Understanding the genomic underpinnings of aging can help illuminate areas worth targeting for medical research. Scientific research over the past century has revealed that aging is plastic, and that variation in environments (e.g., McCay et al. 1935; Lamb 1968) and genotypes (e.g.,

Friedman and Johnson 1988; Kenyon et al. 1993) contribute to aging and longevity. Genetic research of aging prompted the pharmaceutical targeting of several candidate genes, which have shown some promise in preventing or reducing age-related disease (Nadon et al. 2017; Garay 2021). However, these developments are based almost entirely on research in model organisms; much information with medical relevance remains to be discovered in non-model organisms. Given similar patterns in a hallmark of aging across vertebrates (Remot et al. 2021), presence of sequence diversity in aging genes of vertebrates (Opazo et al. 2022), widespread variation of aging rates in ectothermic tetrapods (Reinke et al. 2022), and differential expression of genes associated with aging in tetrapods (Beatty et al. 2022), seeking out genomic underpinnings of aging on a taxonomically broad scale can inform our understanding of the general processes relevant to individual lineages (including humans). This includes comparative approaches both within and between species and comparing these results with those of model organisms in controlled labs with knock-out and artificially-selected strains.

As shown in Chapter 4, patterns of divergence in nucleotide sequence and gene expression between garter snake ecotypes are similar to those found in other organisms with variation in aging (including humans). Although these similarities may not lead to direct application within human medicine, identification of patterns among diverse taxa will help illuminate the molecular pathways that are critical for aging across the tree of life. Reptiles may not be considered by the general public to have relevance to human medicine, but compounds derived from lizard and snake venom are currently used to treat diabetes, stroke, hypertension, cardiovascular diseases, and acute peripheral arterial occlusion (El-Aziz et al. 2019; Bordon et al. 2020). When it comes to medicine, the distance between basic and applied research may not be as distant as some may

perceive. Solutions to many human health concerns may be hidden within the biology of our scaly relatives.

5.5 Dissertation Contributions and Implications

Although the essential role mitochondria play in sustaining life is undeniable and has been solidified over the past two centuries by theorists and empiricists alike, this dissertation shows that the sources of biological patterns across the tree of life are complex and may not be best explained by a singular source. I hypothesized that patterns of lineage diversification, consequences of parthenogenesis, and ecotype divergence in three independent reptile systems were all results of mitonuclear ecology, but my results contained mixed support for a central role of the mitochondrion in driving these patterns. In the *Draco maculatus* species complex (Chapter 2), we tested the hypothesis that mitochondrial phylogeny best explained patterns in nuclear lineage structure. However, the hypothesis receiving the strongest support implicated riverine barriers as the agents driving isolation rather than mitochondrial lineages. In *Aspidoscelis* (Chapter 3), we did find support for reduced mitochondrial function as a basis for lower endurance capacity in asexual hybrid species. Determining the nature of this relationship will require further work integrating physiology and genomics. And lastly, in *Thamnophis elegans* (Chapter 4), we hypothesized that ecotype divergence in physiology (metabolic rate, ROS production), life history (rates of aging and longevity) and mitochondrial genetics (nonsynonymous SNP) would be reflected by patterns in the nuclear-encoded OXPHOS genes. While OXPHOS genes showed evidence of expression and sequence variation that matched our predictions, these did not stand out from patterns we observed in other candidate networks (e.g., insulin signalling and DNA repair) as drivers of divergence.

These findings do not preclude or resolve any of the forementioned hypotheses involving the mitochondria, but they do supply some perspective on the role of mitochondria as ubiquitous players in eukaryotic evolution. The “endless forms most beautiful” referenced in the Introduction (Chapter 1) of this dissertation are the result of interactions within a “tangled bank” (Darwin 1859). Similarly, the patterns of variation observed in lineage diversification, sexual/asexual reproduction, and life history strategies are the result of an intricate and complex network of intracellular products working together within an external environment. Depending on the system, some of these networks (including the gene products involved in mitochondrial function) may play a more influential role, and variation in underlying contributors can be largely determined by environmental influences.

5.6 Concluding Remarks

“Why does receiving answers to curiosity-based research questions matter?” I imagine that scientists Mary Anning, Gregor Mendel, Charles Darwin, Alfred Russell Wallace, among other greats, faced a question similar to this. After all, to many they were simply collecting curios, picking peas, playing with barnacles, and catching butterflies. Today, a century later, we recognize these individuals for their ground-breaking contributions to evolution and genetics. The rippling effect of their findings have impacted diverse areas of basic and applied biological research. In this concluding chapter, I addressed the relevance and implications of the basic science within this dissertation to different aspects of biological theory and society. As mentioned in the Introduction to this dissertation, reptiles as models are valuable resources to finding answers for key questions in evolutionary biology. Findings from projects designed

around these questions are relevant to diverse fields and topics in biology, such as conservation, speciation, mitonuclear ecology, sex evolution, evolutionary physiology, mitochondrial biology, life history evolution, and human medicine. While much remains to be understood regarding the forces driving lineage diversification, the scarcity of parthenogenetic organisms, and the genetic underpinnings of variation in life history strategies, these findings provide some pieces to these puzzles. In the words of Steven J. Gould (Gould 1985), “We have made some sense and order of nature’s confusion.”

4.6 References

- Adzhubei, I. A., S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov, et al. 2010. A method and server for predicting damaging missense mutations. *Nature Methods* 7:4 7:248–249.
- Agostino, A., F. Invernizzi, C. Tiveron, G. Fagiolari, A. Prella, E. Lamantea, A. Giavazzi, et al. 2003. Constitutive knockout of Surf1 is associated with high embryonic lethality, mitochondrial disease and cytochrome c oxidase deficiency in mice. *Human Molecular Genetics* 12:399–413.
- Altmann, R. 1894. *Die Elementarorganismen und ihre Beziehungen zu den Zellen*.
- Aon, M. A., S. Cortassa, M. Juhaszova, J. A. González-Reyes, M. Calvo-Rubio, J. M. Villalba, A. D. Lachance, et al. 2021. Mitochondrial health is enhanced in rats with higher vs. lower intrinsic exercise capacity and extended lifespan. *npj Aging and Mechanisms of Disease* 2021 7:1 7:1–16.
- Attwood, S. W., and D. A. Johnston. 2001. Nucleotide sequence differences reveal genetic variation in *Neotricula aperta* (Gastropoda: Pomatiopsidae), the snail host of schistosomiasis in the lower Mekong Basin. *Biological Journal of the Linnean Society* 73:23–41.
- Avise, J. 2012. *Molecular markers, natural history and evolution*.
- Barreto, F. S., E. T. Watson, T. G. Lima, C. S. Willett, S. Edmands, W. Li, and R. S. Burton. 2018. Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nature Ecology and Evolution* 2:1250–1257.
- Batalha, C. M. P. F., A. E. Vercesi, and N. C. Souza-Pinto. 2022. The Many Roles Mitochondria Play in Mammalian Aging. *Antioxidants and Redox Signaling* 36:824–843.

- Baur, E. 1908. Das Wesen und die Erblchkeitsverhältnisse der “Varietates albomarginatae hort.”
Zeitschrift für induktive Abstammungs-und Vererbungslehre 330–351.
- Beatty, A., A. M. Rubin, H. Wada, B. Heidinger, W. R. Hood, and T. S. Schwartz. 2022.
Postnatal expression of IGF2 is the norm in amniote vertebrates. *Proceedings of the Royal Society B* 289.
- Bolotov, I. N., A. V. Kondakov, I. V. Vikhrev, O. V. Aksenova, Y. V. Bespalaya, M. Y. Gofarov, Y. S. Kolosova, et al. 2017. Ancient River Inference Explains Exceptional Oriental Freshwater Mussel Radiations. *Scientific Reports* 7:1–14.
- Bordon, K. de C. F., C. T. Cologna, E. C. Fornari-Baldo, E. L. Pinheiro-Júnior, F. A. Cerni, F. G. Amorim, F. A. P. Anjolette, et al. 2020. From Animal Poisons and Venoms to Medicines: Achievements, Challenges and Perspectives in Drug Discovery. *Frontiers in Pharmacology* 11:1132.
- Bossy-Wetzel, E., D. D. Newmeyer, and D. R. Green. 1998. Mitochondrial cytochrome c release in apoptosis occurs upstream of DEVD-specific caspase activation and independently of mitochondrial transmembrane depolarization. *EMBO Journal* 17:37–49.
- Bronikowski, A., and D. Vleck. 2010. Metabolism, Body Size and Life Span: A Case Study in Evolutionarily Divergent Populations of the Garter Snake (*Thamnophis elegans*). *Integrative and Comparative Biology* 50:880–887.
- Bulterijs, S., R. S. Hull, V. C. E. Björk, and A. G. Roy. 2015. It is time to classify biological aging as a disease. *Frontiers in Genetics* 6:205.
- Bundle, M. W., H. Hoppeler, R. Vock, J. M. Tester, and P. G. Weyand. 1999. High metabolic rates in running birds. *Nature* 1999 397:6714 397:31–32.
- Ceballos, G., P. R. Ehrlich, and P. H. Raven. 2020. Vertebrates on the brink as indicators of

- biological annihilation and the sixth mass extinction. *Proceedings of the National Academy of Sciences of the United States of America* 117:13596–13602.
- Clark, M. K., L. M. Schoenbohm, L. H. Royden, K. X. Whipple, B. C. Burchfiel, X. Zhang, W. Tang, et al. 2004. Surface uplift, tectonics, and erosion of eastern Tibet from large-scale drainage patterns. *Tectonics* 23.
- Correns, C. 1909. Zur Kenntnis der Rolle von Kern und Plasma bei der Vererbung. *Zeitschrift für Induktive Abstammungs- und Vererbungslehre* 2:331–340.
- Costello, M. J., R. M. May, and N. E. Stork. 2013. Can we name earth's species before they go extinct? *Science* 339:413–416.
- Criddle, R. S., R. M. Bock, D. E. Green, and H. Tisdale. 1962. Physical Characteristics of Proteins of the Electron Transfer System and Interpretation of the Structure of the Mitochondrion. *Biochemistry* 1:827–842.
- Crimmins, E. M. 2015. Lifespan and Healthspan: Past, Present, and Promise. *The Gerontologist* 55:901–911.
- Curry, P. 2011. *Ecological ethics: An introduction*.
- Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection, Or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.
- East, E. M. 1934. The nucleus-plasma problem. *The American Naturalist* 68:289–303.
- El-Aziz, T. M. A., A. G. Soares, and J. D. Stockand. 2019. Snake Venoms in Drug Discovery: Valuable Therapeutic Tools for Life Saving. *Toxins* 11.
- Ephrussi, B. 1949. Action de l'acriflavine sur les levures. I. La mutation "petite colonie". *Ann Inst Pasteur* 76:351–357.
- Flatt, T., and A. Heyland. 2011. Mechanisms of life history evolution: the genetics and

- physiology of life history traits and trade-offs.
- Friedman, D. B., and T. E. Johnson. 1988. A mutation in the age-1 gene in *Caenorhabditis elegans* lengthens life and reduces hermaphrodite fertility. *Genetics* 118:75–86.
- Gangloff, E. J., T. S. Schwartz, R. Klabacka, N. Huebschman, A. Y. Liu, and A. M. Bronikowski. 2020. Mitochondria as central characters in a complex narrative: Linking genomics, energetics, pace-of-life, and aging in natural populations of garter snakes. *Experimental Gerontology* 137.
- Gangloff, E. J., D. Vleck, and A. M. Bronikowski. 2015. Developmental and Immediate Thermal Environments Shape Energetic Trade-Offs, Growth Efficiency, and Metabolic Rate in Divergent Life-History Ecotypes of the Garter Snake *Thamnophis elegans*. *Physiological and Biochemical Zoology* 88:550–563.
- Garay, R. P. 2021. Investigational drugs and nutrients for human longevity. Recent clinical trials registered in ClinicalTrials.gov and clinicaltrialsregister.eu. <https://doi.org/10.1080/13543784.2021.1939306> 30:749–758.
- Garg, S. G., and W. F. Martin. 2016. Mitochondria, the Cell Cycle, and the Origin of Sex via a Syncytial Eukaryote Common Ancestor. *Genome Biology and Evolution* 8:1950–1970.
- Gould, S. J. 1985. Natural History. Pages 167–184 in *The Flamingo's Smile*. WW Norton & Company.
- Hallet, B., and P. Molnar. 2001. Distorted drainage basins as markers of crustal strain east of the Himalaya. *Journal of Geophysical Research: Solid Earth* 106:13697–13709.
- Hatefi, Y., A. G. Haavik, L. R. Fowler, and D. E. Griffiths. 1962. Studies on the Electron Transfer System: XLII. RECONSTITUTION OF THE ELECTRON TRANSFER SYSTEM. *Journal of Biological Chemistry* 237:2661–2669.

- Havird, J. C., M. D. Hall, and D. K. Dowling. 2015. The evolution of sex: A new hypothesis based on mitochondrial mutational erosion: Mitochondrial mutational erosion in ancestral eukaryotes would favor the evolution of sex, harnessing nuclear recombination to optimize compensatory nuclear coadaptation J. . *BioEssays* 37:951–958.
- Heine, K. B., and W. R. Hood. 2020. Mitochondrial behaviour, morphology, and animal performance. *Biological Reviews* 95:730–737.
- Hill, G. E. 2016. Mitonuclear coevolution as the genesis of speciation and the mitochondrial DNA barcode gap. *Ecology and Evolution* 6:5831–5842.
- . 2017. The mitonuclear compatibility species concept. *The Auk* 134:393–409.
- Hirst, J., M. S. King, and K. R. Pryde. 2008. The production of reactive oxygen species by complex I. *Biochemical Society Transactions* 36:976–980.
- Hörandl, E., and F. Hadacek. 2013. The oxidative damage initiation hypothesis for meiosis. *Plant Reproduction* 26:351–367.
- Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* 14:33–38.
- Jamaluddin, J. A. F., N. So, B. M. Tam, A. Ahmad, C. Grudpan, L. M. Page, M. Z. Khaironizam, et al. 2019. Genetic variation, demographic history and phylogeography of tire track eel, *Mastacembelus favus* (Synbranchiformes: Mastacembelidae) in Southeast Asia. *Hydrobiologia* 838:163–182.
- Kang, E., X. Wang, R. Tippner-Hedges, H. Ma, C. D. L. Folmes, N. M. Gutierrez, Y. Lee, et al. 2016. Age-Related Accumulation of Somatic Mitochondrial DNA Mutations in Adult-Derived Human iPSCs. *Cell Stem Cell* 18:625–636.
- Kennedy, E., and A. Lehninger. 1948. Intracellular structures and the fatty acid oxidase system

- of rat liver. *Journal of Biological Chemistry* 172:847.
- Kenyon, C., J. Chang, E. Gensch, A. Rudner, and R. Tabtiang. 1993. A *C. elegans* mutant that lives twice as long as wild type. *Nature* 1993 366:6454 366:461–464.
- Kingsbury, B. F. 1912. Cytoplasmic fixation. *The Anatomical Record* 6:39–52.
- Kitazoe, Y., H. Kishino, M. Hasegawa, A. Matsui, N. Lane, and M. Tanaka. 2011. Stability of Mitochondrial Membrane Proteins in Terrestrial Vertebrates Predicts Aerobic Capacity and Longevity. *Genome Biology and Evolution* 3:1233–1244.
- Klabacka, R. L., P. L. Wood, J. A. McGuire, J. R. Oaks, L. L. Grismer, J. L. Grismer, A. Aowphol, et al. 2020. Rivers of Indochina as potential drivers of lineage diversification in the spotted flying lizard (*Draco maculatus*) species complex. *Molecular Phylogenetics and Evolution* 150.
- Koopman, W. J. H., P. H. G. M. Willems, and J. A. M. Smeitink. 2012. Monogenic Mitochondrial Disorders. *New England Journal of Medicine* 366:1132–1141.
- Kuru, N., O. Dereli, E. Akkoyun, A. Bircan, O. Tastan, and O. Adebali. n.d. PHACT: Phylogeny-Aware Computing of Tolerance for Missense Mutations.
- Ladoukakis, E. D., and E. Zouros. 2017. Evolution and inheritance of animal mitochondrial DNA: Rules and exceptions. *Journal of Biological Research (Greece)* 24:1–7.
- Lamb, M. J. 1968. Temperature and Lifespan in *Drosophila*. *Nature* 1968 220:5169 220:808–809.
- Lane, N. 2014. Bioenergetic Constraints on the Evolution of Complex Life. *Cold Spring Harbor Perspectives in Biology* 6:a015982.
- Lane, N., and W. Martin. 2010. The energetics of genome complexity. *Nature* 2010 467:7318 467:929–934.

- Lederberg, J. 1952. Cell Genetics and Hereditary Symbiosis. *Physiological Reviews* 32:403–430.
- Lehninger, A. L. 1964. The mitochondrion. *Molecular basis of structure and function. The mitochondrion. Molecular basis of structure and function.*
- Lewontin, R. 1974. The genetic basis of evolutionary change.
- Lightowers, R. N., R. W. Taylor, and D. M. Turnbull. 2015. Mutations causing mitochondrial disease: What is new and what challenges remain? *Science* 349:1494–1499.
- López-Otín, C., M. A. Blasco, L. Partridge, M. Serrano, and G. Kroemer. 2013. The Hallmarks of Aging. *Cell* 153:1194–1217.
- Mannella, C. A. 2006. Structure and dynamics of the mitochondrial inner membrane cristae. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1763:542–548.
- Mannella, C. A., W. J. Lederer, and M. S. Jafri. 2013. The connection between inner membrane topology and mitochondrial function. *Journal of Molecular and Cellular Cardiology* 62:51–57.
- McCay, C. M., M. F. Crowell, and L. A. Maynard. 1935. The Effect of Retarded Growth Upon the Length of Life Span and Upon the Ultimate Body Size One Figure. *The Journal of Nutrition* 10:63–79.
- McFarland, R., P. Chinnery, E. Blakely, A. M. Schaefer, A. A. M. Morris, S. M. Foster, H. A. L. Tuppen, et al. 2007. Homoplasmy, heteroplasmy, and mitochondrial dystonia. *Neurology* 69:911–916.
- Meves, F. 1908. Die Chondriosomen als Träger erblicher Anlagen. *Cytologische Studien am Hühnerembryo. Archiv für mikroskopische Anatomie* 72:816–867.
- Mitchell, P. 1961. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* 191:144–148.

- Morgan, T. H. 1919. Cytoplasmic Inheritance. Page 219 *in* *The Physical Basis of Heredity*. JB Lippincott.
- . 1926. Genetics and the Physiology of Development. *The American Naturalist* 60:489–515.
- Nadon, N. L., R. Strong, R. A. Miller, and D. E. Harrison. 2017. NIA Interventions Testing Program: Investigating Putative Aging Intervention Agents in a Genetically Heterogeneous Mouse Model. *EBioMedicine* 21:3–4.
- Nass, M. M. K., and S. Nass. 1963. Intramitochondrial fibers with DNA characteristics: I. Fixation and electron staining reactions. *Journal of Cell Biology* 19:593–611.
- Newcomer, E. H. 1951. Mitochondria in plants. II. *The Botanical Review* 1951 17:2 17:53–89.
- Ng, P. C., and S. Henikoff. 2003. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research* 31:3812.
- Nie, J., G. Ruetenik, K. Gallagher, G. Hoke, C. N. Garziona, W. Wang, D. Stockli, et al. 2018. Rapid incision of the Mekong River in the middle Miocene linked to monsoonal precipitation. *Nature Geoscience* 11:944–948.
- Nielsen, J., K. D. Gejl, M. Hey-Mogensen, H. C. Holmberg, C. Suetta, P. Krstrup, C. P. H. Elemans, et al. 2017. Plasticity in mitochondrial cristae density allows metabolic capacity modulation in human skeletal muscle. *The Journal of Physiology* 595:2839–2847.
- Oaks, J. R. 2019. Full Bayesian Comparative Phylogeography from Genomic Data. *Systematic Biology* 68:371–395.
- Oeppen, J., and J. W. Vaupel. 2002. Demography: Broken limits to life expectancy. *Science* 296:1029–1031.
- Olshansky, S. J. 2018. From Lifespan to Healthspan. *JAMA* 320:1323–1324.

- Opazo, J. C., M. W. Vandewege, F. G. Hoffmann, K. Zavala, C. Meléndez, C. Luchsinger, V. A. Cavieres, et al. 2022. How many sirtuin genes are out there? evolution of sirtuin genes in vertebrates with a description of a new family member. *bioRxiv* 2020.07.17.209510.
- Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* 25:1605–1612.
- Portier, P. 1918. *Les Symbiotes*. Paris: Masson.
- Radzvilavicius, A. L., and N. W. Blackstone. 2015. Conflict and cooperation in eukaryogenesis: implications for the timing of endosymbiosis and the evolution of sex. *Journal of The Royal Society Interface* 12.
- Rafelski, S. M. 2013. Mitochondrial network morphology: Building an integrative, geometrical view. *BMC Biology* 11:1–9.
- Rattan, S. I. S. 2014. Aging Is Not a Disease: Implications for Intervention. *Aging and Disease* 5:196.
- Reinke, B. A., H. Cayuela, F. J. Janzen, J. F. Lemaître, J. M. Gaillard, A. M. Lawing, J. B. Iverson, et al. 2022. Diverse aging rates in ectothermic tetrapods provide insights for the evolution of aging and longevity. *Science (New York, N.Y.)* 376:1459–1466.
- Remot, F., V. Ronget, H. Froy, B. Rey, J. M. Gaillard, D. H. Nussey, and J. F. Lemaitre. 2021. Decline in telomere length with increasing age across nonhuman vertebrates: A meta-analysis. *Molecular Ecology*.
- Rizzuto, R., D. De Stefani, ... A. R.-... reviews M. cell, and undefined 2012. 2012. Mitochondria as sensors and regulators of calcium signalling. *nature.com*.
- Robert, K., and A. Bronikowski. 2010. Evolution of senescence in nature: physiological

- evolution in populations of garter snake with divergent life histories. *American Naturalist* 175:E47-159.
- Rubinoff, D., and B. S. Holland. 2005. Between Two Extremes: Mitochondrial DNA is neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference. *Systematic Biology* 54:952–961.
- Sagan, L. 1967. On the origin of mitosing cells. *Journal of Theoretical Biology* 14:225-IN6.
- Schulz, T. J., K. Zarse, A. Voigt, N. Urban, M. Birringer, and M. Ristow. 2007. Glucose Restriction Extends *Caenorhabditis elegans* Life Span by Inducing Mitochondrial Respiration and Increasing Oxidative Stress. *Cell Metabolism* 6:280–293.
- Schwartz, T. S., Z. W. Arendsee, and A. M. Bronikowski. 2015. Mitochondrial divergence between slow- and fast-aging garter snakes. *Experimental Gerontology* 71:135–146.
- Schwartz, T. S., and A. M. Bronikowski. 2013. Dissecting molecular stress networks: Identifying nodes of divergence between life-history phenotypes. *Molecular Ecology* 22:739–756.
- Short, K. R., M. L. Bigelow, J. Kahl, R. Singh, J. Coenen-Schimke, S. Raghavakaimal, and K. S. Nair. 2005. Decline in skeletal muscle mitochondrial function with aging in humans. *Proceedings of the National Academy of Sciences of the United States of America* 102:5618–5623.
- Singh, G., M. T. Lott, and D. C. Wallace. 1989. A Mitochondrial DNA Mutation as a Cause of Leber's Hereditary Optic Neuropathy. *New England Journal of Medicine* 320:1300–1305.
- Son, J. M., and C. Lee. 2021. Aging: All roads lead to mitochondria. *Seminars in Cell & Developmental Biology* 116:160–168.
- Sonneborn, T. M. 1950. The cytoplasm in heredity. *Heredity* 4:11–36.

- Speijer, D. 2015. Birth of the eukaryotes by a set of reactive innovations: New insights force us to relinquish gradual models. *BioEssays* 37:1268–1276.
- Stearns, S. C. 1992. *The evolution of life histories*. Oxford University Press, New York.
- Sun, N., R. J. Youle, and T. Finkel. 2016. The Mitochondrial Basis of Aging. *Molecular Cell* 61:654–666.
- Towarnicki, S. G., and J. W. O. Ballard. 2020. Towards understanding the evolutionary dynamics of mtDNA. <https://doi.org/10.1080/24701394.2020.1830076> 31:355–364.
- Tzagoloff, A. 1982. *Mitochondria*. Springer Science & Business Media.
- Vigilant, L., M. Stoneking, H. Harpending, K. Hawkes, and A. C. Wilson. 1991. African Populations and the Evolution of Human Mitochondrial DNA. *Science* 253:1503–1507.
- Visinoni, F., and D. Delneri. 2022. Mitonuclear interplay in yeast: from speciation to phenotypic adaptation. *Current Opinion in Genetics & Development* 76:101957.
- Wallin, I. E. 1927. *Symbiogenesis and the Origin of Species*. Рипол Классик.
- Wang, C., X. Liu, and B. Wei. 2011. Mitochondrion: An emerging platform critical for host antiviral signaling. *Expert Opinion on Therapeutic Targets* 15:647–665.
- Wang, L., L. Shen, C. Liu, and L. Ding. 2020. Evolution of the paleo-Mekong River in the Early Cretaceous: Insights from the provenance of sandstones in the Vientiane Basin, central Laos. *Palaeogeography, Palaeoclimatology, Palaeoecology* 545:109651.
- Waterhouse, A., M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, et al. 2018. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research* 46:W296–W303.
- Wisløff, U., S. M. Najjar, Ø. Ellingsen, P. M. Haram, S. Swoap, Q. Al-Share, M. Fernström, et al. 2005. Cardiovascular risk factors emerge after artificial selection for low aerobic

capacity. *Science* 307:418–420.

Workman, D. R. 1975. Tectonic evolution of Indochina. *Journal of the Geological Society, Thailand* 1:3–19.

Wright, S. 1941. The physiology of the gene. *Physiological Reviews* 21:487–527.

Zick, M., R. Rabl, and A. S. Reichert. 2009. Cristae formation-linking ultrastructure and function of mitochondria. *Biochimica et biophysica acta* 1793:5–19.

Ch 5: Figures

Figure 5.1 Publications involving mitochondria through time

Figure 5.2 Paleo-river and Paleo-niche hypotheses for *Draco maculatus* lineage diversification

Figure 5.3 Ideal experimental design for inter-genomic vs intra-genomic testing

Figure 5.4 Predictions for complex activity under different hypotheses

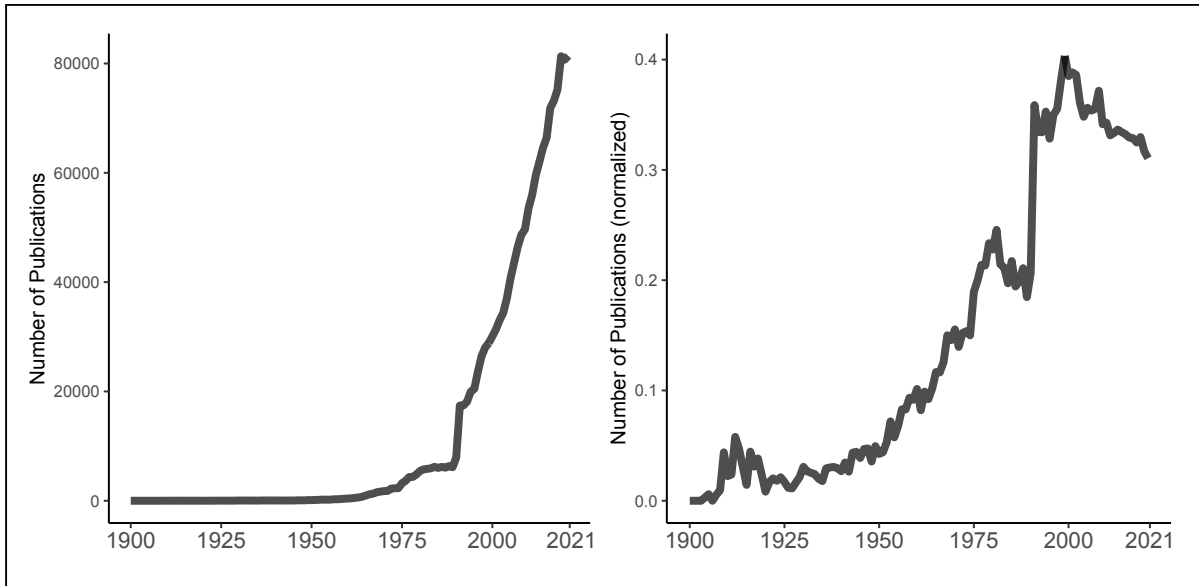


Figure 5.1: **Publications involving mitochondria through time**

Results from Web of Science literature search using the following search criteria: "sarcosomes" OR "film" OR "blepharoplasty" OR "chondriochonts" OR "chondriomites" OR "chondroblasts" OR "chondriosomes" OR "chondrosphere" OR "fuchsinophilic granules" OR "interstitial bodies" OR "Körsdr" OR "Fädenkörsdr" OR "mitogen" OR "parabasal bodies" OR "plastosomes" OR "plastochondria" OR "plastosomes" OR "spheroplasts" OR "vermiculites" OR "bioblasts" OR "mitochondria" OR "mitochondrion" OR "mitochondrial". **A-** Raw search results for each year. **B-** Search results for each year divided by the total number of publications that matched the search criteria "evolution" OR "ecology" OR "physiology" OR "biology".

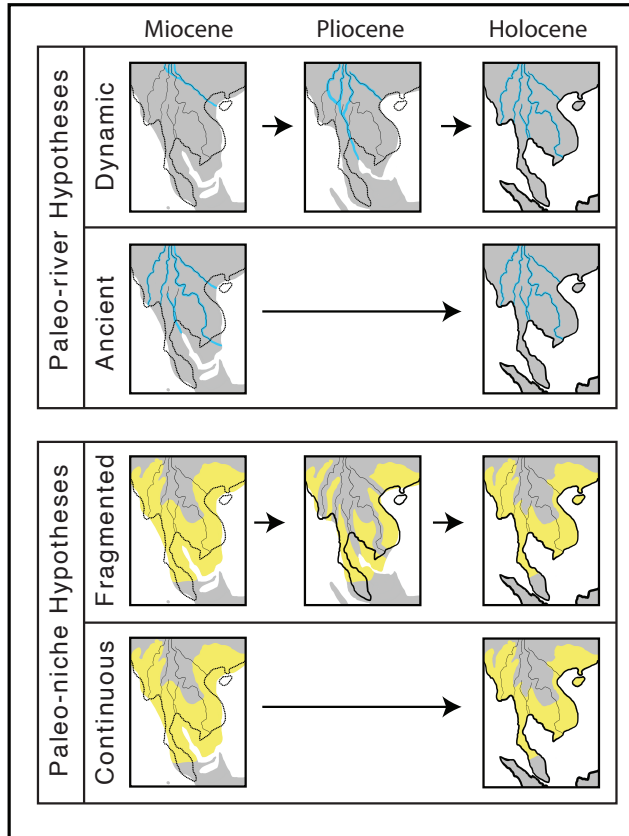


Figure 5.2: **Paleo-river and Paleo-niche hypotheses for *Draco maculatus* lineage diversification**

Cartoon depictions of Paleo-river (top) and Paleo-niche (bottom) hypotheses driving lineage diversification in *Draco maculatus*. Descriptions of these hypotheses can be found in section 5.2.2.

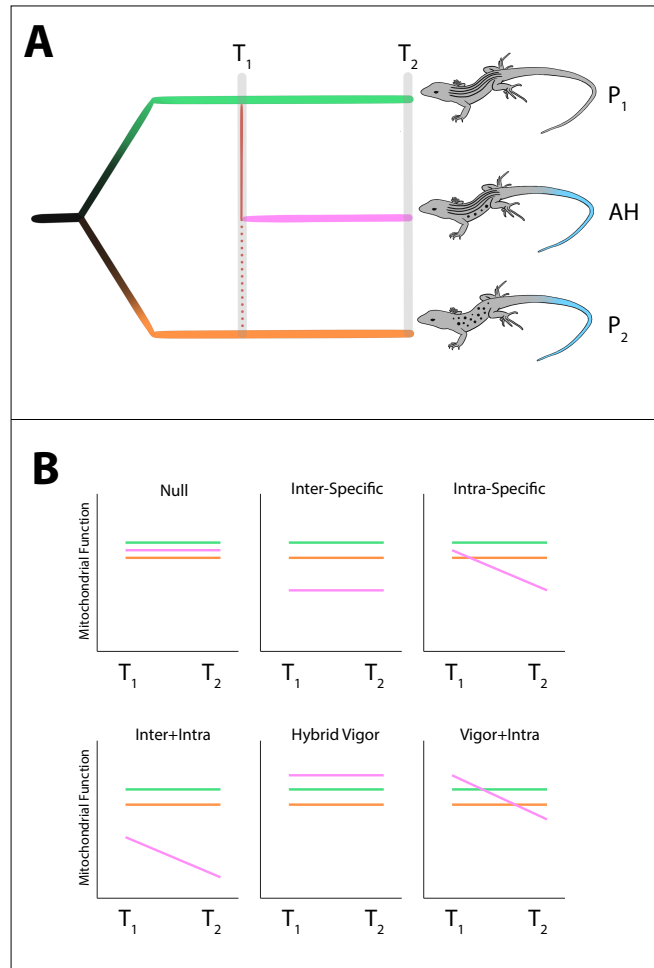


Figure 5.3: Ideal experimental design for inter-genomic vs intra-genomic testing

In order to understand the contributions of inter- vs intra-genomic contributions to patterns in mitochondrial function, it would be helpful to collect samples from the timepoint when the hybridization occurred (T_1) and also from today (T_2). **A:** Evolutionary tree depicting relationships of two sexual species (P_1 and P_2) and one asexual hybrid (AH) lineage. The parental ancestral lineages of the hybrid asexual lineage are depicted by the red lines; dotted line = paternal reticulation, solid line = maternal reticulation. **B:** Patterns predicted at different time points for the different lineages under six hypotheses. Null = the hypothesis where none of the time points are different from one another; Inter-Specific = reduced mitochondrial function began at the origin of the hybrid asexual lineage and is due to incompatibilities between the parental genomes; Intra-Specific = reduced mitochondrial function did not begin at the origin of the hybrid asexual lineage, rather a gradual reduction in mitochondrial function was incurred due to the effect of mutational erosion (via Muller's ratchet); Inter+Intra = a combination of hybrid incompatibility (beginning at initial hybridization event) and mutational erosion; Hybrid Vigor = the hybrid asexual lineage experiences increased mitochondrial function due to the beneficial combination of alleles from the divergent parental genomes; Vigor+Intra = a combination of hybrid vigor (increased mitochondrial performance at time of hybridization) and mutational erosion.

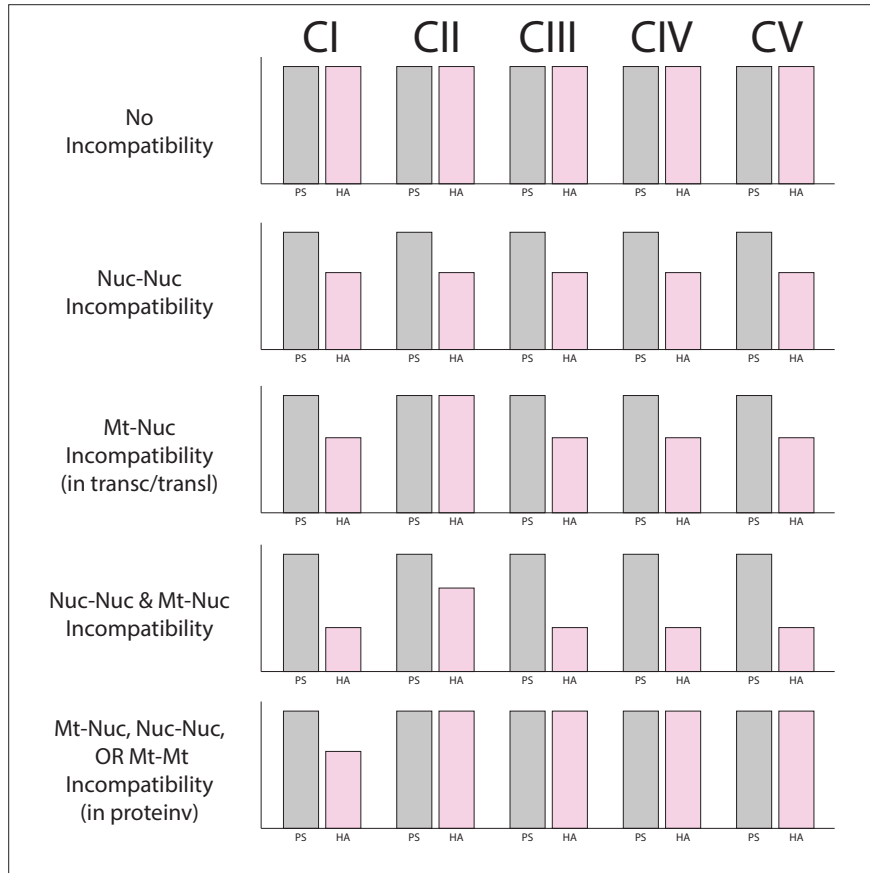


Figure 5.4: **Predictions for complex activity under different hypotheses**

Predictions of complex activity at each electron transport chain complex of the mitochondrion under four hypotheses.

CI-CV are the five primary protein complexes involved in oxidative phosphorylation. PS = parental sexual species,

HA = hybrid asexual lineage. The Y axis for each graph is enzyme activity (the rate at which the protein complex

catalyzes its respective action).