The legacy of macroevolution: understanding and predicting evolutionary responses to novel environments

Bailey K. Howell

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Biological Sciences

Josef C. Uyeda, Chair
Anuj Karpatne
Joel W. McGlothlin
Meryl C. Mims
Kristin M. Winchell

May 9, 2025
Blacksburg, Virginia

Keywords: macroevolution, morphological evolution, global change, constraint, evolutionary contingency

The legacy of macroevolution: understanding and predicting evolutionary responses to novel environments

Bailey K. Howell

(ABSTRACT)

Understanding how species will respond to global change is a major challenge for biologists. Two factors that play an important role are evolutionary constraints and contingencies. Evolutionary constraints are limitations on how species can adapt and change imposed by past evolutionary history, which could be driven by genetic, developmental, or functional variation. Evolutionary contingencies are chance events or evolutionary pathways that affect how species will respond to challenges and can lead to differences in species responses even when exposed to the same environmental pressures. Studying macroevolution of morphological traits and changes in microhabitat usage allows us to investigate the roles of these processes in generating variation in response to different habitats, and accounting for scale allows us to gain a more holistic picture of evolution. Understanding the role of evolutionary history is necessary to predict species responses to global change. In this work I focus on evolution within lizards as these species use a variety of different habitats and the relationship between morphology and microhabitat use has been studied extensively. We investigated the evolution of hindlimb allometry across lizards and identified that the majority of variation in lizards can be explained under a Brownian motion model of evolution, although we did identify contingent evolution in skinks. We then investigated the evolution of hindlimb length and adhesive toepad size in two "model clades" for studying urban tolerance and invasion biology. We identified evidence for evolutionary contingencies in the two clades with opposite patterns of trait evolution for each clade. We then used phylogenetic information, species traits, climate, and geographic data to predict species invasion probabilities across 486 lizard species. We found that incorporating phylogenetic information allowed us to account for complex or difficult to measure traits and improved our models performance. We then developed a feature selection procedure to compare species traits with randomly simulated phylogenetic traits in order to identify if these traits contained additional predictive power beyond phylogenetic information. Taken together these results highlight the importance of evolutionary constraint and contingency in the predictability of evolution and the utility of phylogenetic information for predicting future responses to environmental change.

The legacy of macroevolution: understanding and predicting evolutionary responses to novel environments

Bailey K. Howell

(GENERAL AUDIENCE ABSTRACT)

Understanding how species will react to a changing world is an important question for scientists. Two things that could influence these reactions are evolutionary constraints and chance events. Evolutionary constraints are the limitations or advantages that a species' past evolution has given it. Chance events are unpredictable things that happen along the way. Scientists study how traits have changed over long periods (macroevolution) to understand how these two forces, evolutionary constraints and chance, have shaped how different species have adapted to different environments. Looking at the big picture of evolution helps us understand these processes better. Knowing about a species' evolutionary history is crucial if we want to predict how it might cope with global changes like climate change or habitat loss. In this study, we focused on lizards. Lizards use many different environments, and scientists have already studied how their body shape and size changes in response to living in different environments. We looked at how the size of their back legs relative to their body size have evolved across different lizard species. We found that for most lizards, changes in relative leg length seem to follow a fairly predictable pattern of gradual evolution. However, we also saw some cases in skinks where evolution seemed to take a different, less predictable path depending on specific circumstances. Next, we zoomed in on two specific groups of lizards that are often used to study how animals adapt to city life and how they invade new areas. In these two groups, we looked at the evolution of leg length and the size of the adhesive pads on their toes. Interestingly, we found evidence of those "chance events" playing a role. The two groups showed opposite patterns of how these traits evolved, suggesting that different circumstances led to different evolutionary outcomes. Finally, we used information about lizard family trees (phylogeny), their physical traits, climate data, and where they live to predict how likely 486 different lizard species are to invade new areas. We discovered that including information about their evolutionary relationships improved our predictions. This suggests that evolutionary history carries important information about complex traits that might be hard to measure directly. We also developed a way to figure out if certain traits gave us extra predictive power beyond just knowing their evolutionary history, by comparing the real traits to randomly generated "fake" traits based on the family tree. Overall, our findings emphasize that both the limitations and opportunities passed down through evolution, as well as unpredictable events, are important for understanding and predicting how animals will evolve. Furthermore, knowing the evolutionary history of a species is a valuable tool for forecasting how they might respond to the environmental changes we are seeing today.

Dedication

This dissertation is dedicated to my late father, Bryan Howell. He was an unwavering source of support in my life. Although I am saddened he cannot be here with me today, I know he would be immensely proud. I love you, Dad.

Acknowledgments

I would like to first thank my advisor Dr. Josef Uyeda for his encouragement, guidance, and support throughout this process. Thank you, Josef. You have been a wonderful advisor and confidant. I would also like to thank my committee members Dr. Anuj Karpatne, Dr. Joel McGlothlin, Dr. Meryl Mims, and Dr. Kristin Winchell for their feedback and guidance throughout this process. In addition, I have many co-authors or planned co-authors for the various works included in this dissertation. I would like to thank them all for their feedback and guidance. I would like to especially thank my undergraduate advisor Dr. Travis Hagey and former postdoctoral scholar in my lab Dr. Fabio Machado for their encouragement and support throughout my scientific career. I would also like to thank my funding sources throughout my dissertation including the American Association of University Women and the Global Change Center here at Virginia Tech.

I would like to thank all of the present and past members of the Uyeda Lab who have made my experience at Virginia Tech so special including (but not limited to) Eliza Tarimo, Caleb Charpentier, Dr. Diego Porto, Dr. Katherine Corn, Dr. Mason Linscott, Dr. Orlando Schwery, and Dr. Brian Waldron. You all have given me such valuable feedback and advice over the years and I am very grateful to have your guidance. To the undergraduate students who allowed me to serve as their mentors, Emma White, Samantha Pham, and Lia Fontanella, thank you so much for your willingness to learn and grow with me.

My friends at Virginia Tech have also served as my unofficial academic committee at times and I have learned so much from them over the years and had the best time doing so. I would like to especially thank Camilo Alfonso Cuta, Kerry Gendreau, Sean McHugh, and Kait Malewicz who have given me the opportunity to be my entire self and who have loved me wholeheartedly in return. I love you all so much, and I could never have done this without you. I would also like to thank my friends from other walks of life who have listened to me drone on endlessly about my research over the past 6 years. Thank you Marylyn Crepps, Mariah Washington, Brian Washington, and Hallie Cole for seeing me through some of the most difficult years of my life. I love you all and cannot thank you enough for stepping up when I experienced great loss.

Thank you to my partner Nic Bone. You have improved my life immeasurably and I am so grateful for your unwavering support throughout this process. Despite finishing your dissertation at the same time, you never failed to offer me encouragement and to remind me of what is truly important. I love you more than you will ever know, and I look forward to our future together. Lastly, thank you to my family for their continued support over the years. To my mother, Tammy Howell, you are so loving and selfless. Thank you for

being an incredible role model and always laughing at my jokes. To my sister, Laken Howell, thank you for always trying to protect me and support me when life gets tough. I love you immensely. To my grandparents, Peggy and Roy Mullenix, thank you for your constant love, support, and advice. I love you more than words can say. And finally, to my late father, Bryan Howell, for his constant love and encouragement. Not a day goes by that I do not think of him and miss him dearly. I am so proud to be his daughter.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Global change is occurring rapidly and leading to biodiversity loss at an unprecedented rate [1, 2]. Human-driven impacts on the environment are exposing species to new challenges daily, whether through habitat fragmentation and urbanization, transport of invasive species, or spread of new diseases [3, 4]. Understanding a species' ability to adapt to these challenges, and the factors influencing these responses, is of great interest to evolutionary biologists and ecologists. It is becoming increasingly apparent that macroevolution and phylogeny can provide additional information about species adaptability and constraints when it comes to environmental challenges. Because environmental conditions, species traits, and constraints have shaped how species have evolved over time, studying species responses to novel environments on a macroevolutionary scale can help predict how species may adapt to changes in the future.

The role of stochastic and deterministic processes in evolutionary biology, particularly macroevolution, has historically been debated in the field [5, 6, 7, 8]. Stephen Jay Gould famously proposed that we imagine replaying the tape of life and pondered how different the outcomes might be [9]. The idea that current phenotypes depend on specific environments or events is termed evolutionary contingency [9, 10, 11]. Such contingency can lead to idiosyncrasies in species evolution that are important to understand when studying species responses to novel environments [12, 13]. For example, species that are exposed to similar environmental challenges may employ completely different strategies for coping with these challenges due to evolutionary contingencies.

However, this is not to say that evolution does not often act in predictable ways [14]. In addition to evolutionary contingency, which influences species' current phenotypes through past evolutionary events, species traits can be constrained by fundamental laws. In morphological traits, these can manifest as biomechanical, functional, developmental, or genetic constraints [15, 16, 17, 18, 19]. These constraints can often lead trait evolution to be surprisingly predictable [20].

In this dissertation, I sought to understand the relative roles of evolutionary contingency, constraint, and adaptation to novel environments in lizards, and whether macroevolutionary knowledge has utility for predicting species' responses to novel global change, particularly invasion capacity and adaptation to anthropogenic environments. Why might we expect macroevolution to be useful for such predictions? Human-mediated impacts to the environment are a recent historic phenomena and phylogenies primarily enable (imperfect) understanding of patterns in the macroevolutionary past, not future [21]. However this understanding of patterns in the past is important for making future predictions. While phylogeny does not reveal information about process, it supplements other data we already incorporate to make predictions about future responses to novelty. Particularly, phylogeny can provide information about traits that are difficult to define or measure that may have phylogenetic signal or provide additional information about the properties of traits such as lability or adaptability that may not be apparent from trait measurements alone.

Within lizards, many studies have investigated the role of adaptation in morphological traits to different habitat types, with highly variable results in different families or subfamilies, ranging from classic examples of evolutionary convergence to divergent responses highlighting evolutionary contingency [22, 23, 24, 25]. Within the context of species changes in response to urbanization or habitat fragmentation, recent studies have highlighted that adaptive responses often differ across species [26, 27, 28, 29, 30]. The combination of these many studies and the idiosyncrasies across species would suggest evolutionary contingency plays a large role in trait evolution over time. However, some results within species have suggested that species are constrained and often exhibit convergent or parallel evolution in response to urbanization [31, 32].

In addition to these results, the field of macroevolution has been subject to a debate over the role of scale or scope in comparative studies. Some have said that we should focus our studies on “model clades” [33] while others have suggested studying traits more broadly can alleviate ascertainment bias [34, 35]. While both sides have merit, it is necessary to perform studies at both scales to achieve a more holistic view of evolution over time [21, 36]. When focusing on “model clades” one can uncover important drivers of evolution, however, one may also overlook some of the processes that can generate these patterns at a larger scale. In this work, I sought to investigate and understand patterns in the tempo and mode of morphological evolution at various scales, studying both widespread adaptation to habitat types and macroevolution within “model clades” for studying species responses to urbanization and invasion. Understanding the influence of both evolutionary contingency and constraint in species responses to novel environments is crucial for recognizing the utility of macroevolution in predicting species responses to global change.

Predicting responses to novel environments is a major but pressing challenge, and maximizing predictive accuracy likely requires information on the environments species are adapted to, the traits they possess, and current distributions [37, 38, 39]. In addition to this data, this dissertation investigates the role of phylogeny in improving predictions and how to incorporate this information to achieve the desired result, whether that is improving predictions or increasing interpretability. Invasion biology provides a good framework for predicting species-level responses to novelty, as human-mediated transport introduces species to novel environments at a rapid pace [3]. Predicting species invasion capacities is crucial as it could help to limit the transport and spread of invasive species that are likely to invade and establish populations if introduced.

This collective work helps to identify the role of constraint and contingency in macroevolution and the importance of macroevolution in predicting responses to global change. In Chapter 2 we explore the evolutionary allometry of hindlimbs across lizards and the role of constraint and contingency in generating variation in response to habitat types across lizards. In Chapter 3, we explored evolution in traits related to urban adaptation in two "model clades" containing urban-tolerant and invasive species. This chapter investigates the role of evolutionary history in trait evolution of clades that are ecologically similar and possess similar traits. In Chapter 4 we use phylogeny along with trait and climate data to predict species invasion probabilities. We investigate the utility of including phylogenetic information to account for traits that are complex or difficult to measure. Traits that are adaptive in novel environments will likely hold additional information beyond phylogenetic information. To test this, in Chapter 5 we propose a novel algorithm for feature selection which allows the user to compare biological traits to randomly simulated phylogenetic traits and statistically compare feature importance, providing a useful metric for identifying when predictors may be driving changes in the trait of interest.

Bibliography

[1] Johan Rockström, Will Steffen, Kevin Noone, Åsa Persson, F. Stuart III Chapin, Eric Lambin, Timothy Lenton, Marten Scheffer, Carl Folke, Hans Joachim Schellnhuber, Björn Nykvist, Cynthia de Wit, Terry Hughes, Sander van der Leeuw, Henning Rodhe, Sverker Sörlin, Peter Snyder, Robert Costanza, Uno Svedin, Malin Falkenmark, Louise Karlberg, Robert Corell, Victoria Fabry, James Hansen, Brian Walker, Diana Liverman, Katherine Richardson, Paul Crutzen, and Jonathan Foley. Planetary boundaries: Exploring the safe operating space for humanity. *Ecology and Society*, 14(2), November 2009.

[2] Rowan F. Sage. Global change biology: A primer. *Global Change Biology*, 26(1):3–30, 2020.

[3] Emily V. Moran and Jake M. Alexander. Evolutionary responses to global change: lessons from invasive species. *Ecology Letters*, 17(5):637–649, 2014.

[4] Stuart H. M. Butchart, Matt Walpole, Ben Collen, Arco van Strien, Jörn P. W. Scharlemann, Rosamunde E. A. Almond, Jonathan E. M. Baillie, Bastian Bomhard, Claire Brown, John Bruno, Kent E. Carpenter, Geneviève M. Carr, Janice Chanson, Anna M. Chenery, Jorge Csirke, Nick C. Davidson, Frank Dentener, Matt Foster, Alessandro Galli, James N. Galloway, Piero Genovesi, Richard D. Gregory, Marc Hockings, Valerie Kapos, Jean-Francois Lamarque, Fiona Leverington, Jonathan Loh, Melodie A. McGeoch, Louise McRae, Anahit Minasyan, Monica Hernández Morcillo, Thomasina E. E. Oldfield, Daniel Pauly, Suhel Quader, Carmen Revenga, John R. Sauer, Benjamin Skolnik, Dian Spear, Damon Stanwell-Smith, Simon N. Stuart, Andy Symes,

Megan Tierney, Tristan D. Tyrrell, Jean-Christophe Vié, and Reg Watson. Global biodiversity: Indicators of recent declines. *Science*, 328(5982):1164–1168, May 2010.

[5] Sewall Wright. On the roles of directed and random changes in gene frequency in the genetics of populations. *Evolution*, 2(4):279–294, 1948.

[6] Motoo Kimura et al. Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626, 1968.

[7] David M Raup. Stochastic models in evolutionary palaeontology. In *Developments in palaeontology and stratigraphy*, volume 5, pages 59–78. Elsevier, 1977.

[8] Luis-Miguel Chevin. Species selection and random drift in macroevolution. *Evolution*, 70(3):513–525, March 2016.

[9] Stephen Jay Gould. *Wonderful Life: The Burgess Shale and the Nature of History.* W. W. Norton Company, September 1990. Google-Books-ID: q8BUAgAAQBAJ.

[10] Douglas H. Erwin. Evolutionary contingency. *Current Biology*, 16(19):R825–R826, October 2006.

[11] Zachary D. Blount, Richard E. Lenski, and Jonathan B. Losos. Contingency and determinism in evolution: Replaying life's tape. *Science*, 362(6415):eaam5979, November 2018.

[12] Kathleen L Foster, Jr Garland, Theodore, Lars Schmitz, and Timothy E Higham. Skink ecomorphology: forelimb and hind limb lengths, but not static stability, correlate with habitat use and demonstrate multiple solutions. *Biological Journal of the Linnean Society*, 125(4):673–692, November 2018.

[13] Anna Zimin, Sean V. Zimin, L. Lee Grismer, Aaron M. Bauer, David G. Chapple, Jacob Dembitzer, Uri Roll, and Shai Meiri. Microhabitat and adhesive toepads shape gecko limb morphology. *Integrative Zoology*, n/a(n/a).

[14] Alan de Queiroz. Contingent predictability in evolution: Key traits and diversification. *Systematic Biology*, 51(6):917–929, December 2002.

[15] Stephen Jay Gould. *Ontogeny and phylogeny.* Harvard University Press, 1985.

[16] Stephen Jay Gould. A developmental constraint in cerion, with comments on the definition and interpretation of constraint in evolution. *Evolution*, 43(3):516–539, 1989.

[17] James M Cheverud. Quantitative genetics and developmental constraints on evolution by selection. *Journal of theoretical biology*, 110(2):155–171, 1984.

[18] Dolph Schluter. Adaptive radiation along genetic lines of least resistance. *Evolution*, 50(5):1766–1774, 1996.

[19] Diogo Melo, Arthur Porto, James M. Cheverud, and Gabriel Marroig. Modularity: genes, development and evolution. *Annual Review of Ecology, Evolution, and Systematics*, 47:463–486, 2016.

[20] Miriam Leah Zelditch and Donald L. Swiderski. The predictable complexity of evolutionary allometry. *Evolutionary Biology*, 50(1):56–77, March 2023.

[21] Jonathan B Losos. Seeing the forest for the trees: The limitations of phylogenies in comparative biology: (american society of naturalists address). *The American Naturalist*, 177(6):709–727, 2011.

[22] Anthony Herrel, Jay J. Meyers, and Bieke Vanhooydonck. Relations between microhabitat use and limb shape in phrynosomatid lizards. *Biological Jounral of the Linnean Society*, 2002.

[23] Travis J. Hagey, Scott Harte, Mathew Vickers, Luke J. Harmon, and Lin Schwarzkopf. There's more than one way to climb a tree: Limb length and microhabitat use in lizards with toe pads. *PLoS ONE*, 12(9):1–17, September 2017.

[24] Jane Melville, Luke J Harmon, and Jonathan B Losos. Intercontinental community convergence of ecology and morphology in desert lizards. *Proceedings of the Royal Society B: Biological Sciences*, 273(1586):557–563, December 2005.

[25] Liam J. Revell, Michele A. Johnson, James A. II Schulte, Jason J. Kolbe, and Jonathan B. Losos. A phylogenetic test for adaptive convergence in rock-dwelling lizards. *Evolution*, 61(12):2898–2912, December 2007.

[26] Susannah S. French, Alison C. Webb, Spencer B. Hudson, and Emily E. Virgin. Town and country reptiles: A review of reptilian responses to urbanization. *Integrative and Comparative Biology*, 58(5):948–966, 2018.

[27] Kristin M. Winchell, R. Graham Reynolds, Sofia R. Prado-Irwin, Alberto R. Puente-Rolón, and Liam J. Revell. Phenotypic shifts in urban areas in the tropical lizard anolis cristatellus. *Evolution*, 70(5):1009–1022, May 2016.

[28] Breanna J. Putman, Maria Gasca, Daniel T. Blumstein, and Gregory B. Pauly. Downsizing for downtown: limb lengths, toe lengths, and scale counts decrease with urbanization in western fence lizards (sceloporus occidentalis). *Urban Ecosystems*, 22(6):1071–1081, December 2019.

[29] Breanna J. Putman and Zachary A. Tippie. Big city living: A global meta-analysis reveals positive impact of urbanization on body size in lizards. *Frontiers in Ecology and Evolution*, 8, November 2020.

[30] Kristopher W. Row, Oscar Villasana-Espinosa, Jaele Perez, Grace Urbina, and Luke O. Frishkoff. Morphological limitations imposed on lizards facing urbanization. page 2023.05.09.540039, May 2023.

[31] Amanda Sparkman, Stephen Howe, Stephanie Hynes, Brooke Hobbs, and Karina Handal. Parallel behavioral and morphological divergence in fence lizards on two college campuses. *PLOS ONE*, 13(2):e0191800, February 2018.

[32] Kristin M. Winchell, Shane C. Campbell-Staton, Jonathan B. Losos, Liam J. Revell, Brian C. Verrelli, and Anthony J. Geneva. Genome-wide parallelism underlies contemporary adaptation in urban lizards. *Proceedings of the National Academy of Sciences*, 120(3):e2216789120, January 2023.

[33] Michael J. Donoghue and Erika J. Edwards. Model clades are vital for comparative biology, and ascertainment bias is not a problem in practice: a response to beaulieu and o'meara (2018). *American Journal of Botany*, 106(3):327–330, 2019.

[34] Jeremy M. Beaulieu and Brian C. O'Meara. Can we build it? yes we can, but should we use it? assessing the quality and value of a very large phylogeny of campanulid angiosperms. *American Journal of Botany*, 105(3):417–432, 2018.

[35] Jeremy M. Beaulieu and Brian C. O'Meara. Diversity and skepticism are vital for comparative biology: a response to donoghue and edwards (2019). *American Journal of Botany*, 106(5):613–617, 2019.

[36] Marjorie G Weber and Anurag A Agrawal. Phylogeny, ecology, and the coupling of comparative and experimental approaches. *Trends in ecology & evolution*, 27(7):394–403, 2012.

[37] Nicola J. Van Wilgen and David M. Richardson. The roles of climate, phylogenetic relatedness, introduction effort, and reproductive traits in the establishment of non-native reptiles and amphibians. *Conservation Biology*, 26(2):267–277, 2012.

[38] Peter J. Mahoney, Karen H. Beard, Andrew M. Durso, Aimee G. Tallian, A. Lexine Long, Ryan J. Kindermann, Nicole E. Nolan, Daniel Kinka, and Harrison E. Mohn. Introduction effort, climate matching and species traits as predictors of global establishment success in non-native reptiles. *Diversity and Distributions*, 21(1):64–74, 2015.

[39] Alice Fournier, Caterina Penone, Maria Grazia Pennino, and Franck Courchamp. Predicting future invaders and future invasions. *Proceedings of the National Academy of Sciences*, 116(16):7905–7910, April 2019.

Chapter 2

Constraint in limb allometry across lizards

2.1 Attributions

2.1.1 Co-Authors

Fabio A. Machado, Professor, Department of Integrative Biology, Oklahoma State University, Stillwater, OK, USA

Philip J. Bergmann, Professor Department of Biology, Clark University, Worcester, MA, USA

Anthony Herrel, Professor, Mécanismes Adaptatifs et Evolution, UMR 7179, Département Adaptation du Vivant, Muséum national d'Histoire naturelle CNRS, Paris, France, Department of Biology, Evolutionary Morphology of Vertebrates, Ghent University, Ghent, Belgium, Department of Biology, University of Antwerp, Wilrijk, Belgium, Naturhistorisches Museum Bern, Bern, Switzerland

Jerry F. Husak, Professor, Department of Biology, University of St. Thomas, St. Paul, MN, USA

Antigoni Kaliontzopoulou, Professor, Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals (BEECA), Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Spain

Rishab R. Pillai, Research Scientist, College of Science and Engineering, James Cook University, Townsville, QLD, Australia, School of Biological Sciences University of Adelaide, Adelaide, SA, Australia

Jendrian Riedel, Research Scientist, College of Science and Engineering, James Cook University, Townsville, QLD, Australia, Core Facility for Multidisciplinary Structural Analysis, Hochschule Bremen – City University of Applied Sciences, Bremen, Germany

Ken S. Toyama, Research Scientist, Facultad de Ciencias Biológicas, Universidad Nacional Mayor de San Marcos, Lima, Peru

Travis J. Hagey, Professor, Department of Science and Mathematics, Mississippi University for Women, Columbus, MS, USA

Josef C. Uyeda, Professor, Department of Biological Sciences, Virginia Tech, Blacksburg, VA, USA

2.2 Abstract

Evolutionary allometries are shaped by adaptation and constraint over macroevolutionary time. Here, we examine macroevolutionary patterns of lizard hindlimb allometry to assess the relative impacts of adaptation and constraint. We estimated allometric relationships between hindlimb length and body size for over one thousand species of lizards using Mixed Gaussian Phylogenetic Models (MGPM), allowing shifts in allometric relationships among clades. We also modeled the evolution of hindlimb length and snout-vent length contingent upon habitat usage. Our model incorporating habitat was outperformed by our best MGPM. Under our best model, most species are contained within a single allometric regime. This suggests that hindlimb evolution is highly constrained, with strong phylogenetic inertia and relatively consistent dynamics across species. Variation across these species is most consistent with Brownian motion and appears to be driven by changes in overall body size, occurring along a conserved evolutionary trajectory. The remaining species were split among six different evolutionary regimes, which were often clade-specific. Within skinks, we found unique changes in evolutionary dynamics that preceded the decoupling of limb length and body size, which may have primed the clade to explore morphospace and access adaptive phenotypes associated with limb reduction and fossoriality.

2.3 Introduction

Evolutionary allometry refers to how traits scale across species or populations and has been suggested to arise from developmental constraints over evolutionary timescales [1, 2]. Mechanistically, these constraints are usually explained as a consequence of limited genetic variation in ontogenetic or static allometries [3, 4, 5]. Ontogenetic allometry dictates how traits scale throughout an organism's development, and static allometry governs how traits scale at the same life stage across individuals [6, 7, 8]. This nested relationship of evolutionary, static, and ontogenetic allometry can lead to evolutionary allometries becoming canalized by persistent patterns of selection on functionally or developmentally integrated traits, driving even stronger genetic correlations over time [9]. This scenario, in which selection and genetic correlations act in parallel and reinforce each other, can be referred to as "positive constraints", channeling variation along functional, adaptive axes [3, 10, 11]. This idea contradicts the traditional negative connotations of constraints restricting variation and evolution from achieving particular outcomes ("negative constraints"). Positive constraints can canalize traits along the most advantageous trajectory to evolve more rapidly, especially when functional and developmental rules align [11, 12, 13]. While we cannot explicitly discern between positive and negative constraints in macroevolutionary studies, analyses of evolutionary allometries enable deeper insight into the role of these two types of constraint and their effects on morphological evolution. Specifically, macroevolutionary analyses can reveal when and where allometric relationships shift [14, 15, 16], offering clues about constraints at lower biological levels (i.e., populational and developmental; [17]). A key question to which macroevolutionary data can contribute is the relative timing of shifts in allometry, its association with diversification within novel selective environments, and the emergence of new, evolutionarily favorable body plans. One potential use of macroevolutionary data is assessing how rapidly species respond to selection and whether constraints must be broken prior to adaptation. If constraints do not limit evolution, then transitions to a novel environment could drive specialization and variation across groups quickly. However, if constraints are limiting trait evolution, they must first be broken before species can respond to new selective regimes [18, 19]. Using comparative methods to model selective regimes through time [15, 20], we can ask if colonization of a novel environment is enough to prompt morphological evolution, or if the loosening of constraints is first required for a thorough exploration of adaptive solutions. In that context, the breaking of constraints might be seen as the evolution of key innovations that prompt new ecological opportunities during the invasion of new adaptive zones [18, 21, 22, 23]. To investigate this, we need a trait with a clear functional role—where the relative dimensions of its components are critical to its function—and a clade that has experienced multiple ecological and functional transitions.

Lizards are a highly diverse tetrapod group that occupies a wide range of habitats, making them an ideal system to study the interplay between biomechanical constraints and evolutionary adaptation on a macroevolutionary scale. Hindlimb length, a functional trait tightly linked to lizard locomotor performance, is particularly important as it modulates fitness across different substrates and microhabitats [24, 25, 26]. Performance, and hence fitness, often varies in different groups based on the substrates they use [27, 28, 29], with many groups having specialized hindlimbs for a given microhabitat [30, 31, 32]. Within certain groups, there is evidence for convergent evolution, where distantly related species that occupy similar structural niches evolve similar morphologies [33, 34, 35]. Within Caribbean anoles, different species that use the same microhabitats, termed ecomorphs, have evolved convergent morphologies particularly in their limb proportions [36, 37]. Hindlimb length is also tied to body size through developmental and genetic correlations [38, 39, 40], making it a key trait for studying evolutionary allometry and constraints.

We sought to understand how evolutionary allometry evolves across lizard clades and its relationship to adaptation in different habitats. To do so, we estimated allometric relationships between average species' hindlimb length and body size across lizards. We address three major questions: (1) How variable is hindlimb scaling with respect to body size across clades? (2) How (if at all) is allometric variation related to habitat use? (3) If affirmative, do changes in allometry appear to follow or precede major shifts in habitat use? We hypothesize that variation in allometric scaling could be driven by habitat specialization, which would be illustrated by a strong correlation of shifts in allometry with shifts in habitat. Alternatively, major changes in allometry may appear independently of habitat changes if allometric shifts predate major shifts in habitats. If niche adaptation and body plan diversification occur only after changes in evolutionary allometry, it may suggest that new body plans and adaptive strategies are constrained by the forces that promote allometric stability. To test whether specialization was present in our dataset, we modeled the evolution of limb length and body length contingent upon substrate usage or habitat type. This allows us to determine the relative timing of shifts in allometry and shifts in habitat use, which can inform whether species' allometries are constrained. To determine whether habitat specialization shapes evolutionary allometry across species, we compared models excluding habitat to models incorporating habitat use to determine which was a better fit for our data.

2.4 Methods

While many studies have looked for specialization in specific clades of lizards, we set out to study allometry broadly across families. We compiled data for hindlimb length and body size for over one thousand species of squamate lizards. To answer our first question we modeled evolutionary allometry using Mixed Gaussian Phylogenetic Models (as in [15]) allowing us to automatically detect shifts in limb:body allometry. We chose to use this method as it does not require the restriction of a single model of evolution to apply to the entire tree and it does not require specifying evolutionary shifts a priori. This allowed us to maximize our sample size and still achieve biologically realistic results without applying unnecessary restrictions or unconscious bias through our choice of models.

To answer our second and third questions, we modeled the evolution of limb length and body size contingent upon substrate usage and compared these with our best Mixed Gaussian Phylogenetic Model (MGPM). This allowed us to determine if changes in microhabitat led to changes in limb:body allometry. Investigating patterns produced by both models allowed us to determine how shifts in allometry relate to shifts in substrate usage–giving insight into the importance of evolutionary allometric constraints on enabling or limiting adaptation to novel habitats.

2.4.1 Data Collection

We compiled both published [41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51] and unpublished morphological data from various co-authors. For this study, we used measurements from hindlimbs as this trait's relationship to locomotion has been widely studied and quantified previously [26, 52], and it allowed us to maximize our sample size. We considered hindlimb length, from either side assuming bilateral symmetry, and snout-vent length (SVL) measured externally using digital calipers. We accepted data for limb lengths that were measured in segments (either as hip to knee and knee to longest toe or hip to knee, knee to ankle and ankle to longest toe) and we did not differentiate between museum specimens or live specimens. Although we are aware preservation can shrink specimens by 2-3 percent [53, 54], as well as the possible impact of clade specific inter-observer error, we felt this variation was minor compared to the expected between-species variation. We included squamate species with reduced hindlimbs in our dataset but excluded species that lack hind limbs entirely (e.g., snakes, limbless lizards). We used published data for our habitat type categorizations from [55] categorizing species as terrestrial, arboreal, fossorial, semi-aquatic, and saxicolous or some combination of these. For all analyses, we used a published phylogeny of lizards from [56]. Overall, our final dataset contained 1053 species. Limb length and body length were natural-log transformed prior to performing analyses. We also calculated within-species variance-covariance matrices for all species with measurements from 3 or more individuals to estimate measurement error for evolutionary modeling. For species with insufficient samples, we used the weighted average of other species as the error estimate.

2.4.2 How does allometry vary between clades?

To determine if limb length allometry differs across the lizard tree of life, we used the R package *PCMFit* [15] to fit MGPMs using the function PCMFitMixed. These MGPMs encompass multivariate versions of Brownian motion (BM) and Ornstein-Uhlenbeck (OU) models of trait evolution. This function uses recursive clade partitioning [15], allowing us to fit different models to different tree partitions (regimes). We ran PCMFitMixed using our two traits, limb length and SVL, setting a minimum number of species per partition to five for computational efficiency. Because the PCMFitMixed algorithm is a greedy algorithm that finds locally optimal steps rather than a global optimization [57], we expect some variation in the optimal configuration of shifts across runs. To account for this, we performed ten independent runs and compared them using Akaike Information Criterion (AIC) [58], of which the best-fitting model will hereafter be referred to as the clade model (because shifts may be nested within clades that experienced previous shifts, the clade model allows regimes to include paraphyletic groups). To determine how robust our results were to different tree topologies, we fit MGPMs to 10 different tree topologies randomly sampled from the pseudo-posterior distribution of Title et al. [56] (see Supplementary 2.7 figure 1). We used our clade model parameters to translate model fits into the allometric parameters detailed in the theory below.

Allometry refers to how traits scale with body size, in the form of a power law function as detailed by Huxley (Eq. 1), where X is mean body size, Z is the mean trait value, and a and b are empirical constants [59, 60].

$$Z = aX^b \tag{2.1}$$

When transformed to a log-scale, this relationship assumes its well-known linear form as

$$log(Z) = log(X)b + log(a) \tag{2.2}$$

Given Eq. 2, the slope and intercept can then be solved for as Eq. 3 and 4, respectively.

$$b = \frac{log(Z) - log(a)}{log(X)} \tag{2.3}$$

$$log(a) = log(Z) - b \times log(X) \tag{2.4}$$

To translate our model fits into macroevolutionary allometric relationships estimated using log-linear models, we calculated the slope (b) for each regime as the covariance of the two traits (limb length and SVL) divided by the variance in SVL (Eq. 5) using our estimated rate matrices from different model fits, the Σ matrix or the stationary variance for BM and OU models respectively.

$$b = \frac{Cov(X, Z)}{Var(X)} \tag{2.5}$$

We also calculated the intercept for each regime by solving for log(a) in Eq. 4, where log(X) is the value of SVL and log(Z) is the value of limb length estimated at the root of the phylogeny or the estimated optimal values, for BM and OU models respectively. To test for differences in allometric slopes across clades we used the R package *phylolm* [61] to perform phylogenetic least squares regression (PGLS), fitting a model of limb length and SVL evolution with an interaction across regimes. Although this model is not exactly equivalent to our MGPM, the recovered slopes are similar for most regimes and allow us to test the significance of shifts in slope.

2.4.3 Are shifts in allometry correlated with changes in habitat type?

To determine if habitat changes influence allometry, we simulated 100 stochastic character maps of habitat transitions based on data from [55]. Since many species are classified as occupying multiple habitat types, we used the R package *castor* [62] to establish probabilities for a single tip state (habitat type) by treating the observed habitat as ambiguous using prior probabilities. We then fit transition matrices between these states and recovered posterior probabilities for the tip states. We used the R package *phytools* [63] to simulate stochastic character maps using the best-fitting transition matrix from *castor* [62]. We then used the function PCMFit from the package *PCMFit* [15] to fit both BM and OU models of evolution using the same five habitat types as our designated regimes. In this framework, designated species with shared habitat types evolve under the same model and parameters.

Because many fossorial species were often also classified as both fossorial and terrestrial (see Supplementary Materials 2.7 for more discussion) we ran additional sensitivity analyses to determine how classification of these habitat types impacted our results by manually assigning these species to be only fossorial, only terrestrial, or using prior probabilities. While performance varied with different methods of classification, the primary results presented have all species that were classified as both terrestrial and fossorial manually assigned to the fossorial regime, since this model performed the best. We tested BM and OU models for all of our 100 stochastic character maps and compared them using AIC, of which the best-fitting model will hereafter be referred to as the habitat model. We compared our habitat model to our clade model using AIC to determine which was a better fit for our data.

2.5 Results

2.5.1 How does allometry vary across clades?

We compared our ten independent runs using AIC (see Supplementary, table 2.4). Our best clade model supported seven distinct evolutionary regimes, all evolving under Brownian motion but with varied parameters indicating variation primarily in evolutionary rates (see table 2.3 for all parameters). We tested the sensitivity of these results to varied phylogenies (see Supplementary, figure 2.4) and found similar results for most trees. Moving forward, all analyses were performed using a single phylogeny from Title et al. [56]. However, our discussion will primarily focus on only the most well-supported shifts. The phylogeny and regimes are shown in figure 2.1. We calculated the slope and intercept of each regime using Eq. 4 and Eq. 5, respectively (see table 2.1), and the allometric relationships are plotted along with our data in figure 2.2.

Most lizards were grouped into a single regime with a slope of 1.032 and intercept of -0.830 and three other regimes (2- Scincinae, 5- Eugongylinae & Mabuyinae, and 6- Amphibolurinae) also have similar slopes, but vary in evolutionary rates (figure 2.2, table 2.3). The prevalence of Brownian motion in recovered models suggests that any remaining visual discrepancies within regimes is explained by phylogenetic signal under the model. Many distinct evolutionary regimes were identified within Scincomorpha, the first of these exhibits a slope of 0.998 and intercept of -0.695 and contains the subfamilies of Scincinae and a subclade of Sphenomorphinae (figure 2.2, table 2.1, regime 2). Regime 4, which contains the Egerniinae subfamily and a subclade of Eugongylinae, exhibits decreased covariance of the two traits, while regime 5, containing the remainder of Eugongylinae and the Mabuyinae subfamily, has the opposite pattern with increased covariance between traits leading to an elevated slope. Notably, the split between these regimes (4 and 5) and regime 2 was not well-supported

---

[1] For all Σ matrices the off-diagonals represent co-variance between the two traits while diagonals represent variances in SVL and hindlimb length.

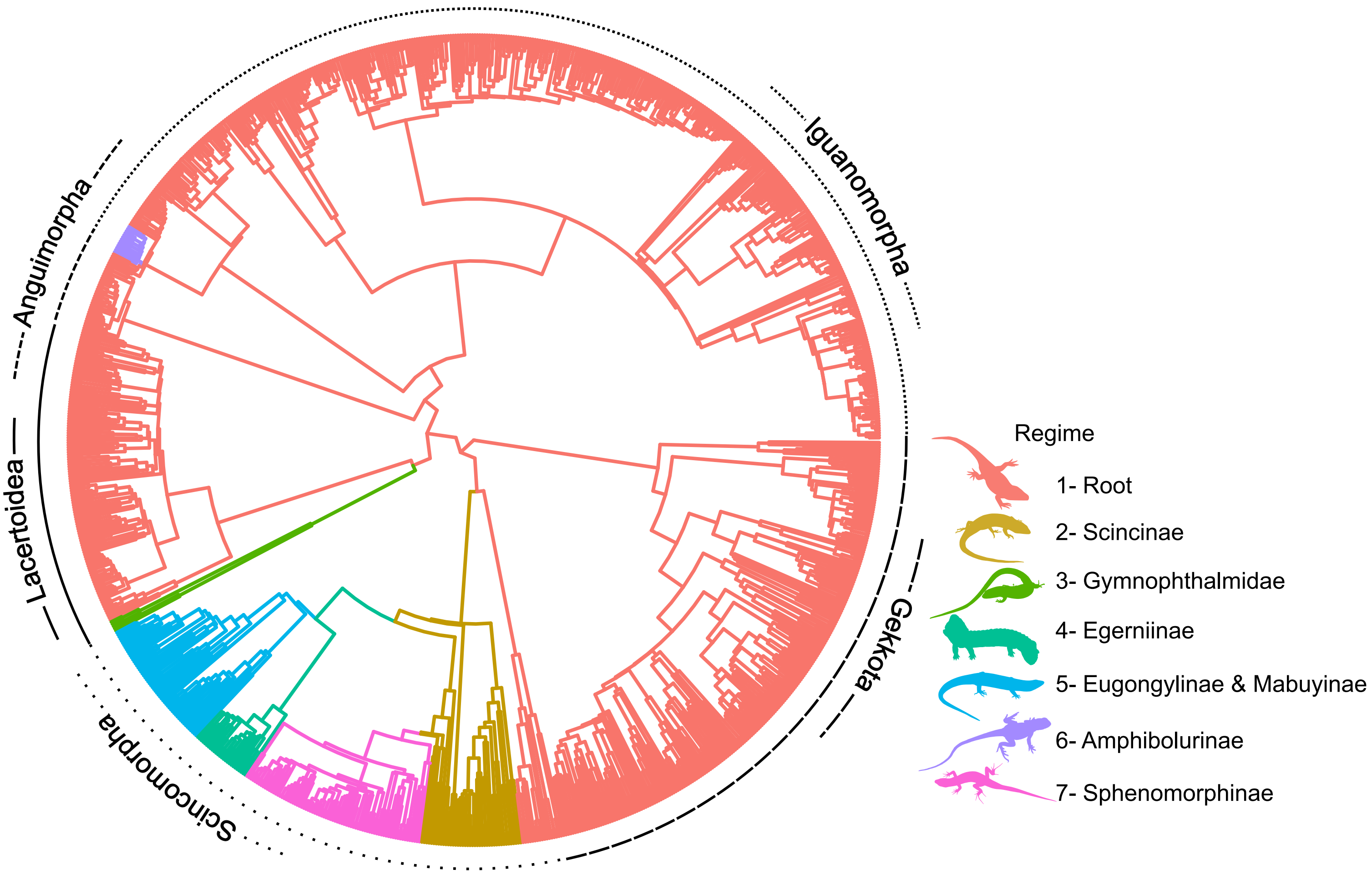


Figure 2.1: The best-fitting clade model contained 7 distinct evolutionary regimes, primarily occurring within skinks. Here we show all seven shifts with regimes denoted by color. The shifts towards different evolutionary regimes primarily occur within Scincomorpha (regime 2 - Scincinae, regime 4 - Egerniinae, regime 5 - Eugongylinae and Mabuyinae, and regime 7 - Sphenomorphinae). The remaining regimes occur in Gymnophthalmidae (regime 3) and Amphibolurinae (regime 6). Silhouettes from PhyloPic (https://www.phylopic.org/).

Table 2.1: The calculated slopes and intercepts of each regime from estimated model parameters (MGPM).

Regime	MGPM Slope	MGPM Intercept
1 – Root	1.03	-0.83
2 – Scincinae	1.00	-0.70
3 – Gymnophthalmidae	0.66	0.66
4 – Egerniinae	0.67	0.62
5 – Eugongylinae & Mabuyinae	1.13	-1.22
6 – Amphibolurinae	1.10	-1.09
7 – Sphenomorphinae	0.00	3.29

Table 2.2: The calculated slopes and intercepts with 95% confidence intervals and p-values for each regime from PGLS.

Regime	Slope	95CI	p-value	Intercept	95CI	p-value
1 – Root	1.16	1.11 – 1.21	<0.005*	-1.32	-1.91 – -0.72	<0.005*
2 – Scincinae	0.97	0.66 – 1.27	0.21	-1.17	-2.84 – 0.50	0.86
3 – Gymnophthalmidae	13.36	11.48 – 15.25	<0.005*	-55.29	-63.49 – -47.08	<0.005*
4 – Egerniinae	0.67	0.52 – 0.83	<0.005*	0.52	-1.03 – 2.06	0.02*
5 – Eugongylinae & Mabuyinae	1.10	0.83 – 1.36	0.64	-1.30	-2.90 – 0.30	0.98
6 – Amphibolurinae	0.97	0.45 – 1.51	0.50	-0.34	-2.64 – 1.97	0.40
7 – Sphenomorphinae	0.14	-0.04 – 0.32	<0.005*	2.04	0.64 – 3.44	<0.005*

Table 2.3: Model parameters for SVL and hindlimb length at the root ($X0$) and rate matrices (Σ) for each regime in the best fitting clade model.[1]

Regime	Type	$X0$	Σ
Root	NA	$\begin{bmatrix} 3.9940238 \\ 3.2896022 \end{bmatrix}$	
1 - Root	BM		$\begin{bmatrix} 0.0030744 & 0.0031713 \\ 0.0031713 & 0.0036955 \end{bmatrix}$
2 - Scincinae	BM		$\begin{bmatrix} 0.0017414 & 0.0017373 \\ 0.0017373 & 0.0089109 \end{bmatrix}$
3 - Gymophthalmidae	BM		$\begin{bmatrix} 0.1495117 & 0.0983790 \\ 0.0983790 & 0.1344341 \end{bmatrix}$
4 - Egerniinae	BM		$\begin{bmatrix} 0.0121640 & 0.0081263 \\ 0.0081263 & 0.0072055 \end{bmatrix}$
5 - Eugongylinae & Mabuyinae	BM		$\begin{bmatrix} 0.0019632 & 0.0022142 \\ 0.0022142 & 0.0033425 \end{bmatrix}$
6 - Amphibolurinae	BM		$\begin{bmatrix} 0.0030317 & 0.0033212 \\ 0.0033212 & 0.0056434 \end{bmatrix}$
7 - Sphenomorphinae	BM		$\begin{bmatrix} 0.0034947 & 0.0000000 \\ 0.0000000 & 0.0160004 \end{bmatrix}$

across different phylogenies (see Supplementary figure 2.4). However, one regime which was well-supported demonstrates a complete break between the covariance of limb length and body size leading to a slope of 0 (figure 2.2, table 2.1, regime 7). The remaining regimes are within Gymnophthalmidae and Amphibolurinae and contain fewer species, specifically Gymnophthalmidae which includes only five species. Of these regimes, the Amphibolurinae shift was not well-supported across different tree topologies.

We performed similar PGLS analyses, testing if slopes varied across regimes to assess the significance of our results. We only tested Brownian models since all regimes supported Brownian motion in our previous analyses. We found that there were significant differences in allometry across clades (see table 2.2 for p-values and confidence intervals and Supplementary table 2.8 for full results). The regimes that are primarily driving these differences appear to be regimes 3- Gymnophthalmidae, 4- Egerniinae, and 7- Sphenomorphinae (table 2.2). This suggests that within lizards, allometric relationships are mostly maintained, except within Scincomorpha and Gymnophthalmidae, where allometry varies (table 2.1, regimes 2, 3, 4, 5, and 7), even including a regime that breaks the allometric scaling of limbs altogether (figure 2.2, regime 7). Our results indicate that while many groups of lizards share a common allometric regime or exhibit similar scaling relationships, Scincomorpha

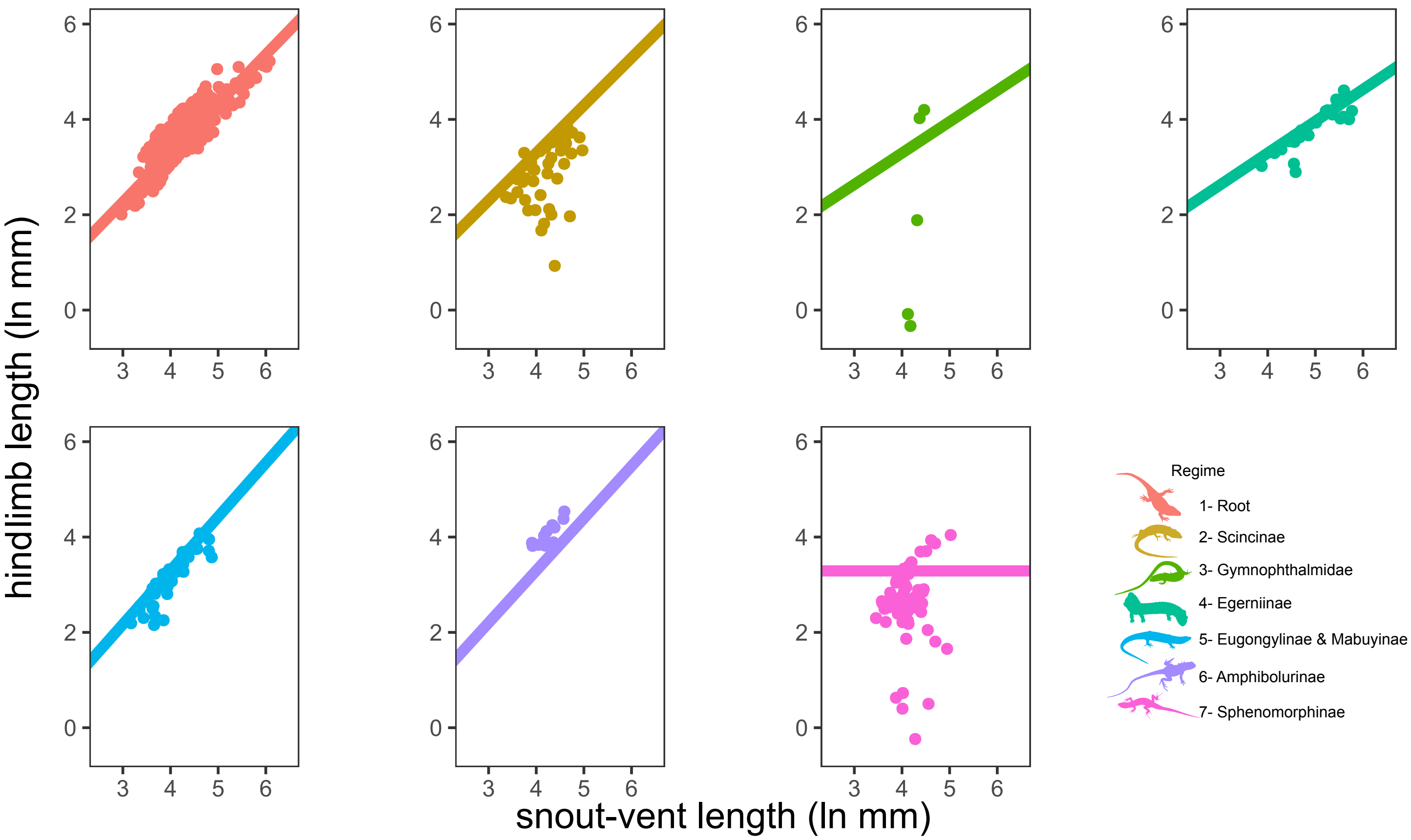


Figure 2.2: The allometric relationships between hind limb length and SVL as estimated by our model parameters are plotted above. Many regimes exhibit similar slopes (regimes 1, 2, 5, and 6). The remaining regimes have slightly shallower slopes (regimes 3 and 4) with the exception of regime 7 which has broken the scaling relationship between limb length and SVL. Silhouettes from PhyloPic (https://www.phylopic.org/).

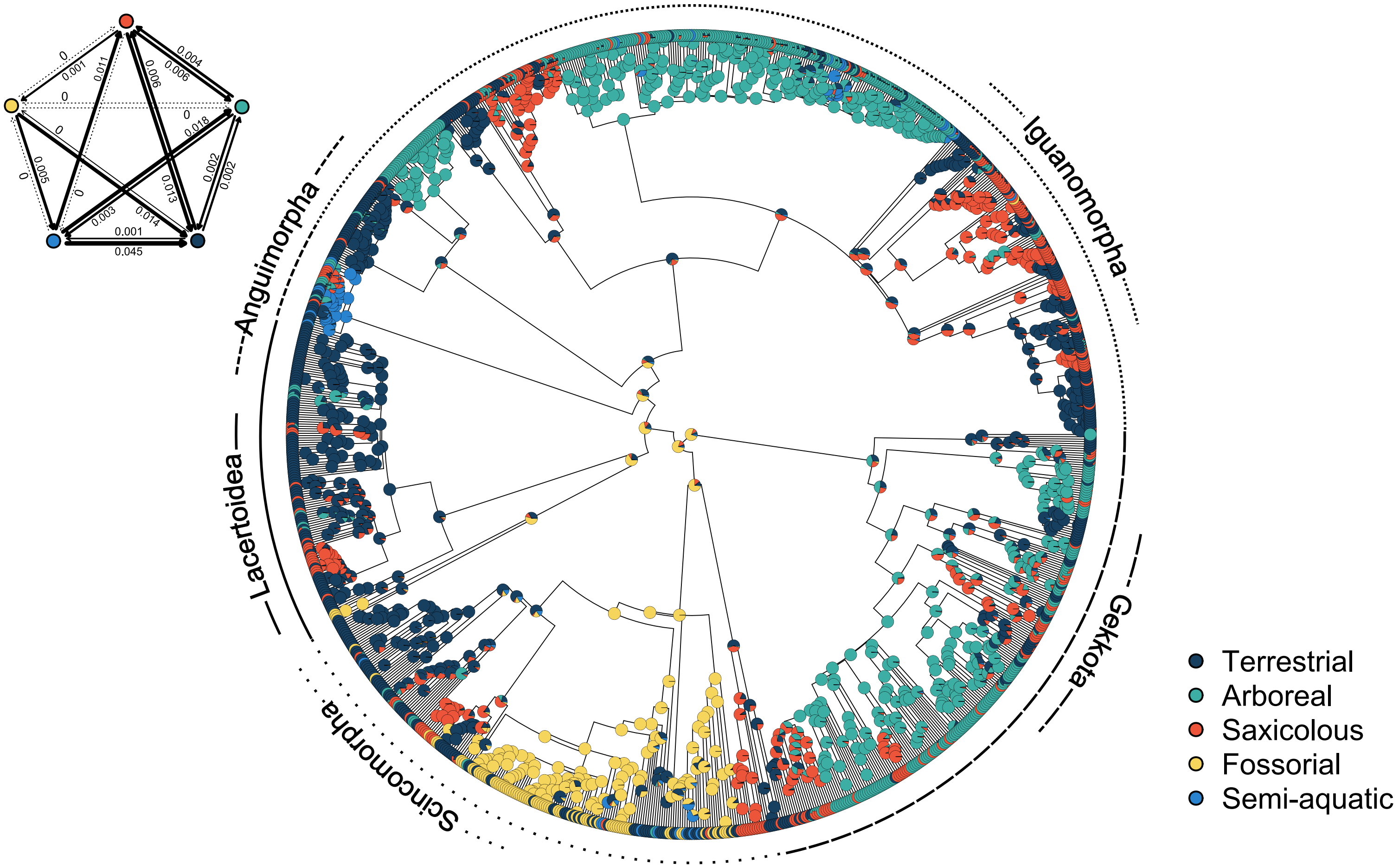


Figure 2.3: Summaries of 100 stochastic character maps which show habitat transitions across the phylogeny using the R package *phytools*. The relative proportions of each habitat type (color) at each node is demonstrated by pie charts. The best-fitting transition matrix from *castor* which was used for mapping is plotted in the upper left.

and Gymnophthalmidae vary with regard to limb scaling.

2.5.2 Are shifts in allometry correlated with changes in habitat type?

Because habitat type is a labile trait in lizards, there are many transitions in our 100 stochastic character maps (figure 2.3). While constrained, hypothesis-driven tests will generally underperform in model-selection against less constrained data-driven models under Maximum Likelihood, the MGPM limits convergence among regimes and penalizes a high number of shifts–giving the habitat model high potential to outperform the conservative transitions reconstructed by the clade model. The best-fitting model for reconstructing evolutionary transitions allowed all rates to differ (see Supplementary table 2.6 for full matrix, see upper left inset of figure 2.3 for visual summary). We compared different models, testing whether all regimes fit BM or OU models of evolution across our 100 maps of habitat type transitions using AIC. In the best-fitting habitat model, all regimes were evolving under Brownian motion with varied parameters (see Supplementary table 2.9 for all parameters). The slopes and intercepts for each regime (corresponding to the 5 habitat types) and plots of these allometries are reported in the Supplementary Materials (table 2.7, figure 2.5). We compared the AIC of our best habitat model (AIC=-845.44) with our best clade model (AIC=-878.58), and the best performing clade model was a better fit than the habitat model (ΔAIC=33.14). However, there was some overlap between our various clade and habitat model fits (see Supplementary Materials **??** for full details). We expect to see high variance in AIC for our habitat models because our methods for habitat categorizations (using prior probabilities) can lead to uncertainty in reconstructions across stochastic character maps.

Because most habitat types exhibited similar slopes with the exception of fossoriality, we compared a reduced habitat model (including only fossorial and non-fossorial classifications) to our best-fitting clade model as well as our previous model which included all habitat categorizations. We found that the clade model (AIC=-878.58) still outperformed this reduced complexity model (AIC=-863.01), although there was overlap in the distributions of AIC across iterations and stochastic maps (see Supplementary table 2.8).

2.6 Discussion

Many previous studies of limb length evolution in lizards have tested for adaptive specialization to different habitat types or convergent evolution with highly variable results [30, 31, 32, 33, 35, 49]. While we might expect to find shared regimes for limb length across species using similar microhabitats, studies have often identified idiosyncracies in morphological diversity due to evolutionary contingency [50, 64, 65, 66]. One reason for this could be due to negative constraints embedded in across-species allometry. Conserved allometric relationships are thought to be caused by canalization due to genetic or developmental linkage, which can be further strengthened by functional demands [9, 67, 68, 69]. While canalization can help maintain high fitness and performance by the generation of highly functional trait combinations, it can also limit exploration of morphospace, disallowing transitions to different adaptive peaks, acting as a negative constraint on evolution and resulting in strong patterns of contingency and phylogenetic signal [1, 5]. Our study reconciles these patterns found across studies of hindlimb allometry in lizards by answering the following questions:

2.6.1 How does allometry vary between clades?

Our results suggested that evolutionary allometry is highly conserved across lizard clades, as evident through most species being grouped into a single regime with common allometric scaling (figure 2.1). This suggests constraint, either due to functional (selective), developmental, or genetic correlations [8, 70]. While the presence of developmental and genetic correlations cannot be fully evaluated in macroevolutionary analyses [4, 71], it is likely that functional constraints might be playing a role in this system. Specifically, within species that locomote similarly (i.e., with an upright sprawling gait; figures 2.1,2.2) [26] there is a common allometric trend. Within several regimes (especially the Root regime which encompasses the majority of species), changes in average body size among species seem to be the main form of evolution. This means that most of the variation seen across the vast majority of lizard species is driven primarily by changes in body size. Changes in size can themselves lead to differences in performance without necessitating the need to change limb scaling [24]. Between regimes that exhibit similar slopes (regimes 1- Root, 2- Scincinae, 5- Eugongylinae & Mabuyinae, and 6- Amphibolurinae), we see changes in the evolutionary rates of limb length and body size that do not dramatically change the scaling relationship between the two traits but could lead to differentiation over time (table 2.3).

The two groups with dramatic changes in scaling contain species with reduced limbs (figure 2.2). Gymnophthalmidae and subclades of Scincomorpha each show significant evolutionary shifts (figure 2.1), and in Sphenomorphinae there is a complete decoupling of limb and body evolution (figure 2.2). Interestingly, there is evidence within Scincomorpha for shifts in allometry prior to the complete decoupling of limb and body size evolution (figures 2.1,2.2). Within Scincinae, the first observed regime shift, the slope of the allometry is slightly decreased from that of the Root regime (figure 2.2) but this group also exhibits higher rates of limb length evolution independently from body size (table 2.3). Thus, regimes that still exhibit strong correlated evolution between the two traits but differ in rates of evolution relative to their ancestral regimes suggest a potential precursor to limb loss, potentially priming this lineage for further exploration of morphospace. What drove this initial decoupling in Scincomorpha is unclear. However, independent of the original drivers, this initial decoupling is seemingly predated by novel shifts in evolutionary rates of limb length related to fossoriality and limb reduction. These subsequent shifts are suggestive that initial shifts in allometry in skinks, which are remarkably rare otherwise, may facilitate the evolution of limb reduction and body plans associated with fossoriality. Conversely, it suggests that other lizard groups may be largely barred from these regions of morphospace due to constraints imposed by allometric scaling.

One plausible reason for the initial and subsequent shifts within the skinks relates to their specialized mode of locomotion. There is evidence that early reptiles utilized a gait that was not unlike that of some extant skinks, which rely extensively on dragging the belly [72, 73]. There is also evidence of this strategy of locomotion being more energetically costly than the upright sprawling gait that most other species in our study exhibit [74]. This could be why extensive belly dragging only persists within some clades and that many species within these clades have transitioned to fossoriality. If early reptiles did rely on this belly-dragging gait, this would indicate why many of our shifts in allometry occur relatively deep in our phylogeny. This morphology could be ancestral and yet be largely selected against due to its energetic inefficiency compared to an upright sprawling gait. Studies of intermediate body plans from quadrupedal to snake-like and how performance varies in different habitats provide valuable insights into the evolution of fossoriality [75, 76, 77].Future studies could provide additional insights into the evolutionary allometry of limb length by incorporating fossil data.

2.6.2 Are shifts in allometry correlated with changes in habitat type?

We found little evidence for specialization or convergence among lizards that were terrestrial, arboreal, saxicolous, or semi-aquatic. Indeed, despite variation in habitat types, we observe very similar parameter estimates for the dynamics of evolution across habitat types which suggests that these lizards are using their limbs similarly (see Supplementary table 2.9). This could indicate that the same general morphologies are effective across multiple habitat types [78, 79]. Across habitats, species might exhibit specialization in other traits which compensate for their biomechanical needs in different environments or using different forms of locomotion, i.e. toe pads in arboreal environments [22, 66], tail specialization in semi-aquatic habitats [80], or musculature specialization for different locomotor modes [81]. Still, species might be specialized in particular limb segments or have variation in integration between limb segments [31, 82, 83]. Species might also be specialized by some other axis of habitat type such as openness or incline [84, 85]. While visually there does appear to be variation in limb:body allometry, particularly in intercept (figure 2.2, regime 1), the high support of Brownian motion models suggests that at broad phylogenetic scales, clade-specific effects can be explained by phylogenetically co-varying residuals, likely driven by shared factors related to ancestry, that nevertheless do not substantially alter the evolutionary dynamics over time. Smaller scale studies, i.e. within families or subfamilies, might reveal more evidence for specialization, as some of this may have been overshadowed by more dramatic shifts in allometry, like in the case of fossorial specialization. Our groupings of similar habitat types may override smaller-scale signs of directional evolutionary change within clades when we consider these traits together at broad phylogenetic scales as was demonstrated by [34].

While we did not find strong evidence of an association between habitat and evolutionary allometry, the associations that did exist appear to be driven by fossorial species that exhibit specialization in limb length [86]. Our data suggests that the most relevant habitat categorization for changes in limb length allometry is fossoriality and its precursors, and that additional shifts are relatively rare or modest in their effect when examined at broad phylogenetic scales.

2.6.3 Conclusions

Taken together our results primarily support the hypothesis that allometry is highly constrained across species, with strong phylogenetic inertia by relatively consistent dynamics across clades. Whether this is due to functional, genetic, or developmental constraints is less clear. Interestingly, while we do see evidence of specialization in fossorial lizards, it is not obvious what allowed this specialization to occur. Because limb length is tightly linked to performance and survival, it is possible that deviations too far from isometry lead to detrimental decreases in function. If this function is required for movement (excluding fossorial environments) it might not be possible to explore other alternative scaling relationships. However, we observe evidence within Scincomorpha that unique changes in allometry preceded complete decoupling associated with limb loss, priming them access adaptive phenotypes associated with limb reduction and subsequent specialization. Further exploration will require more studies incorporating both ontogenetic and static allometry within Scincomorpha to determine how this has influenced limb scaling within this group.

Bibliography

[1] Kjetil L Voje, Thomas F Hansen, Camilla K Egset, Geir H Bolstad, and Christophe Pélabon. Allometric constraints and the evolution of allometry. *Evolution*, 68(3):866–885, 2014.

[2] David Houle, Luke T Jones, Ryan Fortune, and Jacqueline L Sztepanacz. Why does allometry evolve so slowly? *Integrative and comparative biology*, 59(5):1429–1440, 2019.

[3] Stephen Jay Gould. *Ontogeny and phylogeny*. Harvard University Press, 1985.

[4] Dolph Schluter. Adaptive radiation along genetic lines of least resistance. *Evolution*, 50(5):1766–1774, 1996.

[5] Christophe Pélabon, Cyril Firmat, Geir H. Bolstad, Kjetil L. Voje, David Houle, Jason Cassara, Arnaud Le Rouzic, and Thomas F. Hansen. Evolution of morphological al-

lometry: The evolvability of allometry. *Annals of the New York Academy of Sciences*, 1320(1):58–75, July 2014.

[6] Stephen Jay Gould. Allometry and size in ontogeny and phylogeny. *Biological Reviews*, 1966.

[7] A. G. Cock. Genetical aspects of metrical growth and form in animals. *The Quarterly Review of Biology*, 41(2):131–190, June 1966.

[8] James M. Cheverud. Relationships among ontogenetic, static, and evolutionary allometry. *American Journal of Physical Anthropology*, 59(2):139–149, 1982.

[9] Diogo Melo, Arthur Porto, James M. Cheverud, and Gabriel Marroig. Modularity: genes, development and evolution. *Annual Review of Ecology, Evolution, and Systematics*, 47:463–486, 2016.

[10] Stephen Jay Gould. A developmental constraint in cerion, with comments on the definition and interpretation of constraint in evolution. *Evolution*, 43(3):516–539, 1989.

[11] Stephen Jay Gould. *The Structure of Evolutionary Theory*. Harvard University Press, 2002.

[12] Philip J Bergmann and Isabel Tonelli-Sippel. Many-to-many mapping: A simulation study of how the number of traits and tasks affect the evolution of form and function. *Journal of theoretical biology*, 581:111744, 2024.

[13] Fabio Andrade Machado. Selection and constraints in the ecomorphological adaptive evolution of the skull of living canidae (carnivora, mammalia). *The American Naturalist*, 196(2):197–215, 2020.

[14] Josef C Uyeda, Matthew W Pennell, Eliot T Miller, Rafael Maia, and Craig R McClain. The evolution of energetic scaling across the vertebrate tree of life. *The American Naturalist*, 190(2):185–199, 2017.

[15] Venelin Mitov, Krzysztof Bartoszek, and Tanja Stadler. Automatic generation of evolutionary hypotheses using mixed Gaussian phylogenetic models. *Proceedings of the National Academy of Sciences*, 116(34):16921–16926, August 2019. Publisher: National Academy of Sciences Section: Biological Sciences.

[16] Josef C Uyeda, Rosana Zenil-Ferguson, and Matthew W Pennell. Rethinking phylogenetic comparative methods. *Systematic Biology*, 67(6):1091–1109, 2018.

[17] Marjorie G Weber and Anurag A Agrawal. Phylogeny, ecology, and the coupling of comparative and experimental approaches. *Trends in ecology & evolution*, 27(7):394–403, 2012.

[18] Nathan M Young, Günter P Wagner, and Benedikt Hallgrímsson. Development and the evolvability of human limbs. *Proceedings of the National Academy of Sciences*, 107(8):3400–3405, 2010.

[19] George Gaylord Simpson. *Tempo and Mode in Evolution.* Columbia University Press, 1944.

[20] Josef C. Uyeda and Luke J. Harmon. A Novel Bayesian Method for Inferring and Interpreting the Dynamics of Adaptive Landscapes from Phylogenetic Comparative Data. *Systematic Biology*, 63(6):902–918, November 2014.

[21] Edward D. Burress and Martha M. Muñoz. Ecological opportunity from innovation, not islands, drove the anole lizard adaptive radiation. *Systematic Biology*, 2022.

[22] Aryeh H Miller and James T Stroud. Novel tests of the key innovation hypothesis: Adhesive toepads in arboreal lizards. *Systematic Biology*, 71(1):139–152, January 2022.

[23] Aryeh H. Miller, James T. Stroud, and Jonathan B. Losos. The ecology and evolution of key innovations. *Trends in Ecology Evolution*, 38(2):122–131, February 2023.

[24] Theodore Garland and Jonathan Losos. Ecological morphology of locomotor performance in squamate reptiles. In Peter C. Wainwright and Stephen M. Reilly, editors, *Ecological morphology: integrative organismal biology*, pages 240–302. University of Chicago Press, Chicago, IL, USA, 01 1994.

[25] Duncan J. Irschick and Theodore Garland Jr. Integrating function and ecology in studies of adaptation: Investigations of locomotor capacity as a model system. *Annual Review of Ecology, Evolution, and Systematics*, 32(Volume 32, 2001):367–396, November 2001.

[26] AP Russell and V Bels. Biomechanics and kinematics of limb-based locomotion in lizards: review, synthesis and prospectus. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 131(1):89–112, 2001.

[27] Kristin M. Winchell, Inbar Maayan, Jason R. Fredette, and Liam J. Revell. Linking locomotor performance to morphological shifts in urban lizards. *Proceedings of the Royal Society B: Biological Sciences*, 285(1880):20180229, June 2018.

[28] Barry Sinervo and Jonathan B. Losos. Walking the Tight Rope: Arboreal Sprint Performance Among Sceloporus Occidentalis Lizard Populations. *Ecology*, 72(4):1225–1233, 1991. _eprint: https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.2307/1941096.

[29] Jonathan B Losos and Barry Sinervo. The effecrs of morphology and perch diameter on sprint performance of *Anolis* lizards. page 8.

[30] Tiana Kohlsdorf, Theodore Garland Jr., and Carlos A. Navas. Limb and tail lengths in relation to substrate usage inTropidurus lizards. *Journal of Morphology*, 248(2):151–164, May 2001.

[31] Anthony Herrel, Jay J. Meyers, and Bieke Vanhooydonck. Relations between microhabitat use and limb shape in phrynosomatid lizards. *Biological Jounral of the Linnean Society*, 2002.

[32] Travis J. Hagey, Scott Harte, Mathew Vickers, Luke J. Harmon, and Lin Schwarzkopf. There's more than one way to climb a tree: Limb length and microhabitat use in lizards with toe pads. *PLoS ONE*, 12(9):1–17, September 2017.

[33] Jane Melville, Luke J Harmon, and Jonathan B Losos. Intercontinental community convergence of ecology and morphology in desert lizards. *Proceedings of the Royal Society B: Biological Sciences*, 273(1586):557–563, December 2005.

[34] Liam J. Revell, Michele A. Johnson, James A. II Schulte, Jason J. Kolbe, and Jonathan B. Losos. A phylogenetic test for adaptive convergence in rock-dwelling lizards. *Evolution*, 61(12):2898–2912, December 2007.

[35] Jendrian Riedel, L Lee Grismer, Timothy Higham, Joseph Wu, Quyen Hanh Do, Truong Quang Nguyen, Camila G Meneses, Rafe M Brown, Patrick D Campbell, Thomas Ziegler, et al. Ecomorphology of the locomotor apparatus in the genus cyrtodactylus (gekkota, squamata). *Evolutionary Biology*, 51(1):106–123, 2024.

[36] Jonathan B. Losos. The evolution of form and function: Morphology and locomotor performance in west indian anolis lizards. *Evolution; International Journal of Organic Evolution*, 44(5):1189–1203, August 1990.

[37] Jonathan B. Losos. *Lizards in an Evolutionary Tree Ecology and Adaptive Radiation of Anoles*. University of California Press, August 2009.

[38] Thomas J. Sanger, Liam J. Revell, Jeremy J. Gibson-Brown, and Jonathan B. Losos. Repeated modification of early limb morphogenesis programmes underlies the convergence of relative limb length in anolis lizards. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729):739–748, August 2011.

[39] Robin M. Andrews and Sable A. Skewes. Developmental origin of limb size variation in lizards. *Evolution Development*, 19(3):136–146, 2017.

[40] Joel W. McGlothlin, Megan E. Kobiela, Helen V. Wright, D. Luke Mahler, Jason J. Kolbe, Jonathan B. Losos, and Edmund D. III Brodie. Adaptive radiation along a deeply conserved genetic line of least resistance in anolis lizards. *Evolution Letters*, 2(4):310–322, August 2018.

[41] Graham G. Thompson, Christofer J. Clemente, Philip C. Withers, Bryan G. Fry, and Janette A. Norman. Is body shape of varanid lizards linked with retreat choice? *Australian Journal of Zoology*, 56(5):351–362, March 2009.

[42] C. J. Clemente, G. G. Thompson, and P. C. Withers. Evolutionary relationships of sprint speed in australian varanid lizards. *Journal of Zoology*, 278(4):270–280, 2009.

[43] Jerry F. Husak and Jimmy A. McGuire. Does 'gliding while gravid' explain Rensch's rule in flying lizards?: Size dimorphism in flying lizards. *Biological Journal of the Linnean Society*, 113(1):270–282, September 2014.

[44] N. C. Wu, L. A. Alton, C. J. Clemente, M. R. Kearney, and C. R. White. Morphology and burrowing energetics of semi-fossorial skinks (liopholis spp.). *Journal of Experimental Biology*, 218(15):2416–2426, August 2015.

[45] A. Kaliontzopoulou, M. A. Carretero, and D. C. Adams. Ecomorphological variation in male and female wall lizards and the macroevolution of sexual dimorphism in relation to habitat use. *Journal of Evolutionary Biology*, 28(1):80–94, 2015.

[46] Colin Donihue. Microgeographic variation in locomotor traits among lizards in a human-built environment. *PeerJ*, 4:e1776, March 2016.

[47] Colin M. Donihue, Anthony Herrel, Anne-Claire Fabre, Ambika Kamath, Anthony J. Geneva, Thomas W. Schoener, Jason J. Kolbe, and Jonathan B. Losos. Hurricane-induced selection on the morphology of an island lizard. *Nature*, 560(7716):88–91, August 2018.

[48] M. L. Haines, J. Melville, J. Sumner, N. Clemann, D. G. Chapple, and D. Stuart-Fox. Geographic variation in hybridization and ecological differentiation between three syntopic, morphologically similar species of montane lizards. *Molecular Ecology*, 25(12):2887–2903, 2016.

[49] Ken S. Toyama. Interaction between morphology and habitat use: A large-scale approach in tropidurinae lizards. *Breviora*, 554(1):1–20, May 2017.

[50] Philip J. Bergmann and Gen Morinaga. The convergent evolution of snake-like forms by divergent evolutionary pathways in squamate reptiles*. *Evolution*, 73(3):481–496, March 2019.

[51] Bethan L. Littleford-Colquhoun, Christofer Clemente, Graham Thompson, Romane H. Cristescu, Nicola Peterson, Kasha Strickland, Devi Stuart-Fox, and Celine H. Frere. How sexual and natural selection shape sexual size dimorphism: Evidence from multiple evolutionary scales. *Functional Ecology*, 33(8):1446–1458, 2019.

[52] Duncan J. Irschick and Bruce C. Jayne. Comparative three-dimensional kinematics of the hindlimb for high-speed bipedal and quadrupedal locomotion of lizards. *Journal of Experimental Biology*, 202(9):1047–1065, May 1999.

[53] Bart Vervust, Stefan Van Dongen, and Raoul Van Damme. The effect of preservation on lizard morphometrics – an experimental study. January 2009.

[54] Inbar Maayan, R Graham Reynolds, Rachel M Goodman, Paul M Hime, Ryan Bickel, E Allen Luck, and Jonathan B Losos. Fixation and preservation contribute to distortion in vertebrate museum specimens: a 10-year study with the lizard anolis sagrei. *Biological Journal of the Linnean Society*, 136(3):443–454, July 2022.

[55] Shai Meiri. Traits of lizards of the world: Variation around a successful evolutionary design. *Global ecology and biogeography*, 27(10):1168–1172, 2018.

[56] Pascal O. Title, Sonal Singhal, Michael C. Grundler, Gabriel C. Costa, R. Alexander Pyron, Timothy J. Colston, Maggie R. Grundler, Ivan Prates, Natasha Stepanova, Marc E. H. Jones, Lucas B. Q. Cavalcanti, Guarino R. Colli, Nicolas Di-Poï, Stephen C. Donnellan, Craig Moritz, Daniel O. Mesquita, Eric R. Pianka, Stephen A. Smith, Laurie J. Vitt, and Daniel L. Rabosky. The macroevolutionary singularity of snakes. *Science (New York, N.Y.)*, 383(6685):918–923, February 2024.

[57] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction To Algorithms*, chapter 16. Greedy Algorithms. MIT Press, 2001. Google-Books-ID: NLngYyWFl$_YC$.

[58] Hirotogu Akaike. Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike*, pages 199–213. Springer, 1998.

[59] Julian S. Huxley. Constant Differential Growth-Ratios and their Significance. *Nature*, 114(2877):895–896, December 1924. Publisher: Nature Publishing Group.

[60] J. S. Huxley and G. Teissier. Terminology of relative growth. *Nature*, 137(3471):780–781, May 1936.

[61] Lam si Tung Ho and Cécile Ané. A linear-time algorithm for gaussian and non-gaussian trait evolution models. *Systematic Biology*, 63(3):397–408, May 2014.

[62] Stilianos Louca and Michael Doebeli. Efficient comparative phylogenetics on large trees. *Bioinformatics*, 34(6):1053–1055, March 2018.

[63] Liam J. Revell. Phytools 2.0: An updated r ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, 12:e16505, January 2024.

[64] Gen Morinaga and Philip J. Bergmann. Convergent body shapes have evolved via deterministic and historically contingent pathways in lerista lizards. *Biological Journal of the Linnean Society*, 121(4):858–875, August 2017.

[65] Kathleen L Foster, Jr Garland, Theodore, Lars Schmitz, and Timothy E Higham. Skink ecomorphology: forelimb and hind limb lengths, but not static stability, correlate with habitat use and demonstrate multiple solutions. *Biological Journal of the Linnean Society*, 125(4):673–692, November 2018.

[66] Anna Zimin, Sean V. Zimin, L. Lee Grismer, Aaron M. Bauer, David G. Chapple, Jacob Dembitzer, Uri Roll, and Shai Meiri. Microhabitat and adhesive toepads shape gecko limb morphology. *Integrative Zoology*, n/a(n/a).

[67] Rupert Riedl. Order in living organisms: a systems analysis of evolution. *(No Title)*, 1978.

[68] Jeremy MV Rayner. Linear relations in biomechanics: the statistics of scaling functions. *Journal of Zoology*, 206(3):415–439, 1985.

[69] Miriam Leah Zelditch and Donald L. Swiderski. The predictable complexity of evolutionary allometry. *Evolutionary Biology*, 50(1):56–77, March 2023.

[70] James M Cheverud. Quantitative genetics and developmental constraints on evolution by selection. *Journal of theoretical biology*, 110(2):155–171, 1984.

[71] Joseph Felsenstein. Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics*, pages 445–471, 1988.

[72] Massimo Bernardi and Marco Avanzini. Locomotor behavior in early reptiles: Insights from an unusual erpetopus trackway. *Journal of Paleontology*, 85(5):925–929, September 2011.

[73] Stefan Curth, Martin S. Fischer, and John A. Nyakatura. Ichnology of an extant belly-dragging lizard—analogies to early reptile locomotion? *Ichnos*, 21(1):32–43, January 2014.

[74] Edwin Dickinson, Christopher S. Hanna, Hannah M. Fischer, Elizabeth C. Davoli, Allen A. Currier III, and Michael C. Granatosky. Locomotor energetics in the indonesian blue-tongued skink (tiliqua gigas) with implications for the cost of belly-dragging in early tetrapods. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, 337(4):329–336, 2022.

[75] Philip J. Bergmann, Gen Morinaga, Elyse S. Freitas, Duncan J. Irschick, Günter P. Wagner, and Cameron D. Siler. Locomotion and palaeoclimate explain the re-evolution of quadrupedal body form in brachymeles lizards. *Proceedings of the Royal Society B: Biological Sciences*, 287(1938):20201994, November 2020.

[76] Gen Morinaga and Philip J. Bergmann. Evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards. *Proceedings of the Royal Society B: Biological Sciences*, 287(1923):20200192, March 2020.

[77] Marco Camaiti, Alistair R. Evans, Christy A. Hipsley, Mark N. Hutchinson, Shai Meiri, Rodolfo de Oliveira Anderson, Alex Slavenko, and David G. Chapple. Macroecological and biogeographical patterns of limb reduction in the world's skinks. *Journal of Biogeography*, 50(2):428–440, 2023.

[78] María J. Tulli, Félix B. Cruz, Tiana Kohlsdorf, and Virginia Abdala. When a general morphology allows many habitat uses. *Integrative Zoology*, 11(6):483–499, 2016.

[79] Yuliya Kulyomina, Daniel S. Moen, and Duncan J. Irschick. The relationship between habitat use and body shape in geckos. *Journal of Morphology*, 280(5):722–730, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jmor.20979.

[80] Frank E Fish, Natalia Rybczynski, George V Lauder, and Christina M Duff. The role of the tail or lack thereof in the evolution of tetrapod aquatic propulsion. *Integrative and Comparative Biology*, 61(2):398–413, August 2021.

[81] Anthony Herrel, Bieke Vanhooydonck, Joanne Porck, and Duncan J. Irschick. Anatomical basis of differences in locomotor behavior in anolis lizards: A comparison between two ecomorphs. *Bulletin of the Museum of Comparative Zoology*, 159(4):213–238, December 2008.

[82] M. J. Tulli, V. Abdala, and F. B. Cruz. Relationships among morphology, clinging performance and habitat use in liolaemini lizards. *Journal of Evolutionary Biology*, 24(4):843–855, April 2011.

[83] Héctor Tejero-Cicuéndez, Iris Menéndez, Adrián Talavera, Gabriel Mochales-Riaño, Bernat Burriel-Carranza, Marc Simó-Riudalbas, Salvador Carranza, and Dean C Adams. Evolution along allometric lines of least resistance: morphological differentiation in pristurus geckos. *Evolution*, 77(12):2547–2560, December 2023.

[84] B. A. Goodman. Nowhere to run: the role of habitat openness and refuge use in defining patterns of morphological and performance evolution in tropical lizards. *Journal of Evolutionary Biology*, 22(7):1535–1544, July 2009.

[85] Timothy E. Higham and Bruce C. Jayne. Locomotion of lizards on inclines and perches: hindlimb kinematics of an arboreal specialist and a terrestrial generalist. *Journal of Experimental Biology*, 207(2):233–248, January 2004.

[86] Marco Camaiti, Alistair R. Evans, Christy A. Hipsley, and David G. Chapple. A farewell to arms and legs: a review of limb reduction in squamates. *Biological Reviews*, 96(3):1035–1050, 2021.

2.7 Supplementary Material

2.7.1 Methods

As mentioned in the main text, many species were classified as occupying multiple habitat types, so we used the R package *castor* [62] to establish probabilities for a single tip state (habitat type) by treating the observed habitat as ambiguous using prior probabilities. We then fit transition matrices between these states and recovered posterior probabilities for the tip states. We used the R package *phytools* [63] to simulate stochastic character maps using the best-fitting transition matrix from *castor* [62]. Under this framework, the ancestral state was often reconstructed as fossorial, which we suspected may be incorrect. We considered that this may have been driven by species that shared equal probabilities of being terrestrial and fossorial due to their ambiguous coding. To account for this we performed sensitivity analyses, by classifying all species coded as Fossorial&Terrestrial as only terrestrial and only fossorial and seeing how these models performed against ambiguous coding. We created 100 stochastic character maps under each model using *phytools* [63] and then fit BM and OU models of evolution using *PCMFit* [15]. We compared all of these models using AIC. We found that classifying all species as fossorial provided the best fit. Although this did not change the reconstructed ancestral state often being fossorial out of our 100 stochastic character maps, the best habitat model which we compared to the clade model did have a reconstructed ancestral state of terrestrial (see 2.6).

2.7.2 Figures

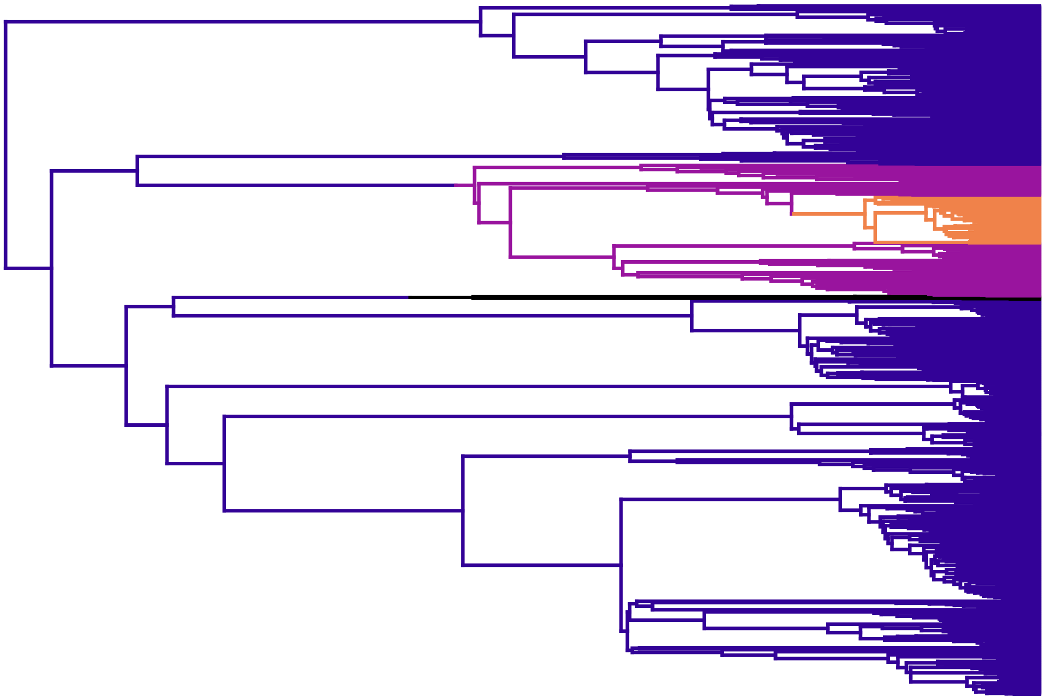


Figure 2.4: Summary of MGPMs fit across ten different tree topologies to determine sensitivity of detected shifts to topological uncertainty. We determined the most often observed regimes in the tree and plotted these on the best supported topology above. The colors of the regimes represent the proportion of times a branch was assigned to the same regime across all 10 topologies with darker colors being more well supported.

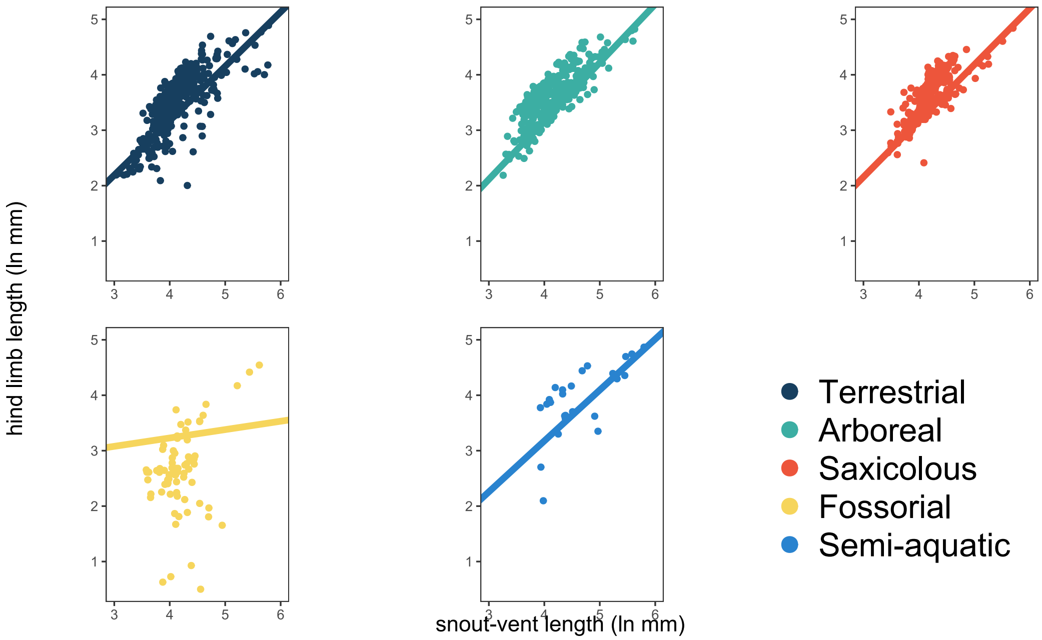


Figure 2.5: Estimated parameters from the habitat model were translated into allometric slopes and intercepts and plotted for each regime (corresponding to habitat types).

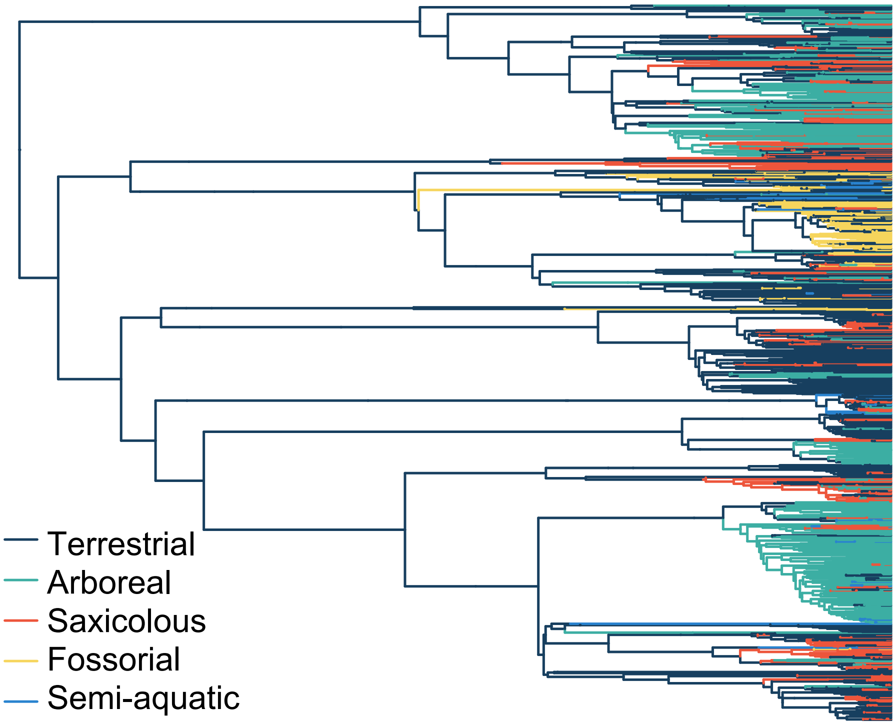


Figure 2.6: The best fitting model habitat model was fit using this reconstruction from 100 stochastic character maps classifying all species that were Fossorial&Terrestrial as Fossorial.

2.7.3 Tables

Table 2.4: Model fits of ten independent runs of PCMFitMixed.

Model	Number of Parameters	Log-Likelihood	AIC
1	13	414.0778	-788.1556
2	14	431.6580	-821.3160
3	17	455.0987	-858.1974
4	13	413.8373	-787.6747
5	16	441.0799	-832.1598
6	16	425.4476	-800.8951
7	16	438.9689	-827.9378
8	**22**	**474.2906**	**-878.5813**
9	19	451.5755	-843.1510
10	13	417.7084	-795.4169

Table 2.5: Model fit from PGLS with estimates and p-values.

Regime	Estimate	Standard Error	t-value	p-value
Root	-1.32	0.30	-4.36	<0.005*
Scincinae	-1.17	0.85	0.18	0.86
Gymnophthalmidae	-55.29	4.19	-12.90	<0.005*
Egerniinae	0.52	0.79	2.33	0.02*
Eugongylinae & Mabuyinae	-1.30	0.82	0.02	0.98
Amphibolurinae	-0.34	1.18	0.83	0.40
Sphenomorphinae	2.04	0.72	4.69	<0.005*
SVL:Root	1.16	0.03	45.94	<0.005*
SVL:Scincinae	0.97	0.16	-1.24	0.21
SVL:Gymnophthalmidae	13.36	0.96	12.69	<0.005*
SVL:Egerniinae	0.67	0.08	-6.01	<0.005*
SVL:Eugongylinae & Mabuyinae	1.10	0.13	-0.46	0.64
SVL:Amphibolurinae	0.98	0.27	-0.67	0.50
SVL:Sphenomorphinae	0.14	0.09	-10.87	<0.005*

Table 2.6: The estimated best-fitting transition matrix using the R package *castor* supported all rates differ (ARD).

State	Terrestrial	Arboreal	Saxicolous	Fossorial	Semi-aquatic
Terrestrial	-0.0088	0.0017	0.0060	0.0003	0.0007
Arboreal	0.0025	-0.0090	0.0039	0.0000	0.0025
Saxicolous	0.0130	0.0055	-0.0199	0.0014	0.0000
Fossorial	0.0139	0.0000	0.0000	-0.0188	0.0049
Semi-aquatic	0.0446	0.0178	0.0106	0.0000	-0.0730

Table 2.7: The calculated slopes and intercepts of each regime (corresponding to habitat types) using estimated model parameters.

Regime	Slope	Intercept
Terrestrial	0.981	-0.753
Arboreal	1.053	-1.045
Saxicolous	1.014	-0.886
Fossorial	0.151	2.626
Semi-aquatic	0.919	-0.498

Table 2.8: Minimum, mean, and maximum AIC values across different models. Models are described in the text.

Model	Min	Mean	Max
Fossorial&Terrestrial – BM	-633.15	-342.76	536.39
Fossorial&Terrestrial – OU	-656.60	-442.51	-280.38
Fossorial – BM	-845.44	-706.10	17.91
Fossorial – OU	-808.69	-775.57	-711.12
Terrestrial – BM	-301.90	-105.01	508.73
Terrestrial – OU	-335.19	-248.15	-203.59
Reduced Habitat – BM	-850.58	-722.75	-450.37
Reduced Habitat – OU	-863.01	-828.40	-779.65
Clade – Best Model	-878.58	-823.35	-787.67

Table 2.9: Model parameters for SVL and hindlimb length at the root ($X0$) and rate matrices (Σ) for the five regimes corresponding to habitat types in the best fitting habitat model. For all Σ matrices the off-diagonals represent co-variance between the two traits while diagonals represent variances in SVL and hindlimb length.

Regime	Type	$X0$	Σ
Root	NA	$\begin{bmatrix} 4.0697204 \\ 3.2398335 \end{bmatrix}$	
Terrestrial	BM		$\begin{bmatrix} 0.0032270 & 0.0031661 \\ 0.0031661 & 0.0037073 \end{bmatrix}$
Arboreal	BM		$\begin{bmatrix} 0.0031346 & 0.0033002 \\ 0.0033002 & 0.0039773 \end{bmatrix}$
Saxicolous	BM		$\begin{bmatrix} 0.0022993 & 0.0023310 \\ 0.0023310 & 0.0026600 \end{bmatrix}$
Fossorial	BM		$\begin{bmatrix} 0.0037315 & 0.0005632 \\ 0.0005632 & 0.0264533 \end{bmatrix}$
Semi-aquatic	BM		$\begin{bmatrix} 0.0038986 & 0.0035811 \\ 0.0035811 & 0.0043082 \end{bmatrix}$

Chapter 3

The role of evolutionary contingency in morphological evolution of two urban-tolerant, invasive lizard clades

3.1 Attributions

3.1.1 Co-Authors

Travis J. Hagey, Professor, Department of Science and Mathematics Mississippi University for Women, Columbus, MS

Samantha Pham, Undergraduate Researcher, Department of Biological Sciences, Virginia Tech, Blacksburg, VA

Josef C. Uyeda, Professor, Department of Biological Sciences, Virginia Tech, Blacksburg, VA

3.1.2 Acknowledgments

3.2 Abstract

Understanding how evolutionary contingency can lead to divergent evolutionary trajectories, even within ecologically similar species is important for predicting species responses to environmental changes. We sought to understand how two morphological traits, adhesive toepad size and hindlimb length, have evolved over time in two genera occupying similar breadths of ecologies. We tested multiple models of trait evolution, including early burst, Ornstein-Uhlenbeck, and Brownian motion. We found that *Hemidactylus* fit a pattern of Brownian, unconstrained evolution for toepad area, achieving greater divergence and diversification over relevant evolutionary timescales, while *Anolis* were more constrained in this trait. Conversely, we found that relative hindlimb length showed opposite patterns for the two lizards, indicating that they may differ in the traits that make this exploitation possible due to evolutionary contingency. These results suggest that these species may be using

different strategies to exploit and invade urban and diverse natural environments, however more studies need to be done to elucidate how (if at all) selection is acting differently upon these two traits in both genera.

3.3 Introduction

The range of morphology across species is often limited by genetic, developmental, or biomechanical constraints. These limitations in trait diversity and trait evolution can differ across clades due to evolutionary contingency [1, 2, 3]. Often we see idiosyncrasies in how species traits evolve and adapt despite species using similar environments [4, 5, 6]. Various combinations of trait values can often lead to similar performances—what has been termed many-to-one mapping of form to function [7, 8, 9].

To investigate the role of evolutionary contingency for morphological traits it is necessary to study traits related to performance in clades that are distantly related but use similar habitats or microhabitats. *Hemidactylus* geckos and *Anolis* lizards provide an excellent study system, where adhesive toepads used to exploit novel environments have evolved independently [10, 11]. Caribbean *Anolis* are already a model system for studying ecomorphology [12, 13], but in recent years have become an excellent example of a clade containing species that exploit urban environments and are excellent invaders [14]. Similarly *Hemidactylus* geckos heavily exploit urban environments and have introduced and established populations in tropical and subtropical environments across the globe [15]. Both clades contain species which have benefited from human-mediated introductions [16].

Previous studies have shown that species within these clades often occur in similar habitats [17]. Two traits which are thought to to be important for successful colonization and further adaptation of urban areas are adhesive toepads and relative limb proportions. Adhesive toepad size has been correlated with cling-force [18, 19, 20]. Toepads work synergistically with claw allowing species to cling to a variety of substrates, where toepads are thought to contribute more to clinging on smoother substrates [21, 22]. Relatively larger adhesive toepads could provide increased performance in urban environments where smooth substrates are more common [23, 24]. Increased relative limb lengths, particularly hindlimbs, are correlated with greater sprint speeds across some species [25, 26], which is important as greater habitat openness could select for the ability to more easily sprint between perches [23, 24, 27]. Previous work has shown that geckos and anoles differ in clade averages in clinging performance as well as in their tempo and mode of evolution [10], despite similar levels of performance being possible via different means in the two clades [28]. Specifically, *Anolis* appeared more constrained in their performance evolution than Geckos, which showed a signature of unconstrained evolution [10].

Given these differences in the patterns of evolution underlying adhesive performance among clades, we sought to understand how these two morphological traits, adhesive toepad size and hindlimb length, have evolved over time in these two clades in genera occupying similar breadths of ecologies. Specifically we wanted to investigate how trait diversity has changed through time for each trait (relative to body size as measured by snout-vent length). To do so, we fit continuous-trait evolutionary models to both traits with body size as a co-variate. We then sought to determine if patterns of trait evolution fit expectations of ongoing divergence and diversification, bounded or constrained evolution, or decelerating/early-burst evolution [29]. We hypothesized that despite ecological similarity, these species would differ in their trait evolution due to evolutionary contingencies and many-to-one mapping of morphology to adaptive performance to diverse ecological settings [10]. Although unidentifiable in the current study, we suspected these differences might be due to evolutionary constraints imposed by selection acting differently on these traits or genetic or developmental constraints. Understanding how these species have evolved to invade novel environments can help to untangle the relative roles of constraint and contingency in adaptation to global change.

3.4 Methods

3.4.1 Data collection

We used hindlimb length and snout-vent length (SVL) measurements taken externally using digital calipers that were compiled by Howell from both published and unpublished sources. We imaged toepads of preserved *Hemidactylus* specimens and estimated toepad area using the software ImageJ. We used published measurements of *Anolis* toepad area that were collected similarly [30].

3.4.2 Node height tests

We took the square-root of toepad area so that it represented a comparable linear measurement as body size and relative limb length. We then log-transformed all traits. To account for size differences we fit models of each trait with SVL as a predictor and used the residuals as our standardized data. We calculated phylogenetic independent constrasts for the residuals [31], which we then used to perform node height tests [32], plotting absolute value of the independent contrasts against the height of the node in the phylogeny (measured from the root). Node height tests are commonly used to detect changes in evolutionary rates over time and are complementary approaches to statistical model-fitting (see below). They are particularly useful for visualizing if patterns in trait evolution may be driven by outliers or measurement error at the tips. Early bursts of evolution show decreases in rates of evolution over time while increases in rates of evolution over time could indicate a late burst or Ornstein-Uhlenbeck model of trait evolution.

3.4.3 Model fitting

To validate our node height tests and provide further information, we fit models of trait evolution to residuals of the two traits in each clade separately. We tested multiple models of univariate trait evolution using the R package geiger [33], including early burst (EB), Ornstein-Uhlenbeck (OU), and Brownian motion (BM). We compared each of these models using AIC [34]. We also fit models of bivariate trait evolution using the R package mvMORPH [35]. We paired each trait with body size, to investigate patterns of trait evolution simultaneously, testing OU and BM models and comparing them with AIC [34]. We used these model fits to plot the stationary distributions of these traits for each group by calculating the stationary variance for OU models and using the rate matrix multiplied by the height of the tree for BM models. This was done to compare visually the predicted distributions under their respective evolutionary processes that account for uneven phylogenetic sampling and differences in clade ages.

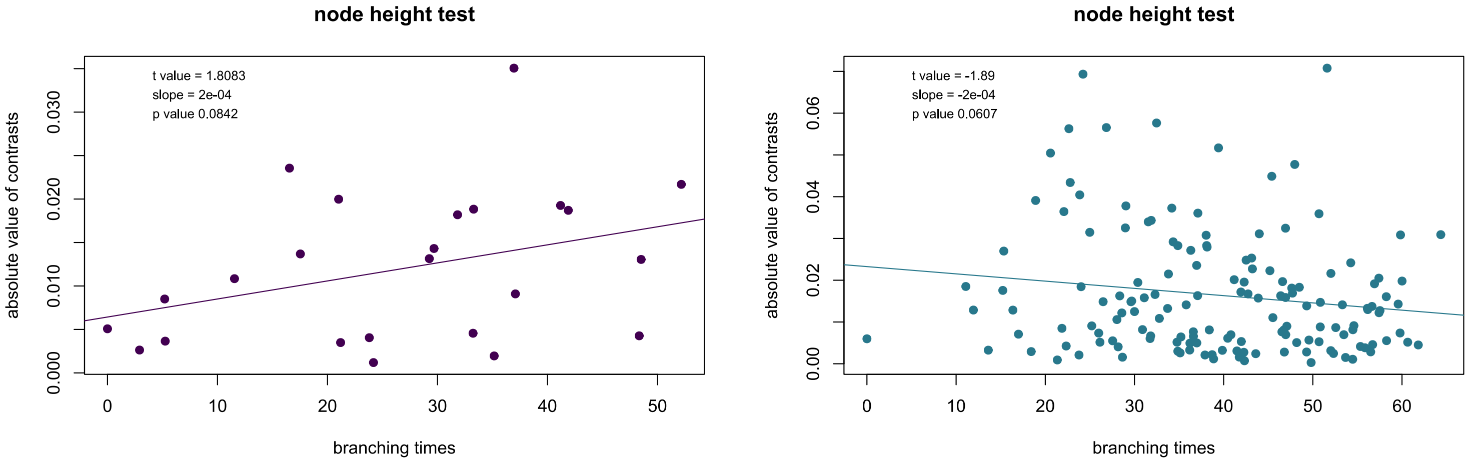
3.5 Results

In our comparisons of *Hemidactylus* and *Anolis* for both traits, relative toepad area and relative limb length, we found that the two clades exhibit opposite patterns regardless of the methods used.

Figure 3.1: Phylogeny of lizards showing species with data for limb length in *Hemidactylus* (purple) and *Anolis* (teal)

3.5.1 Node height tests

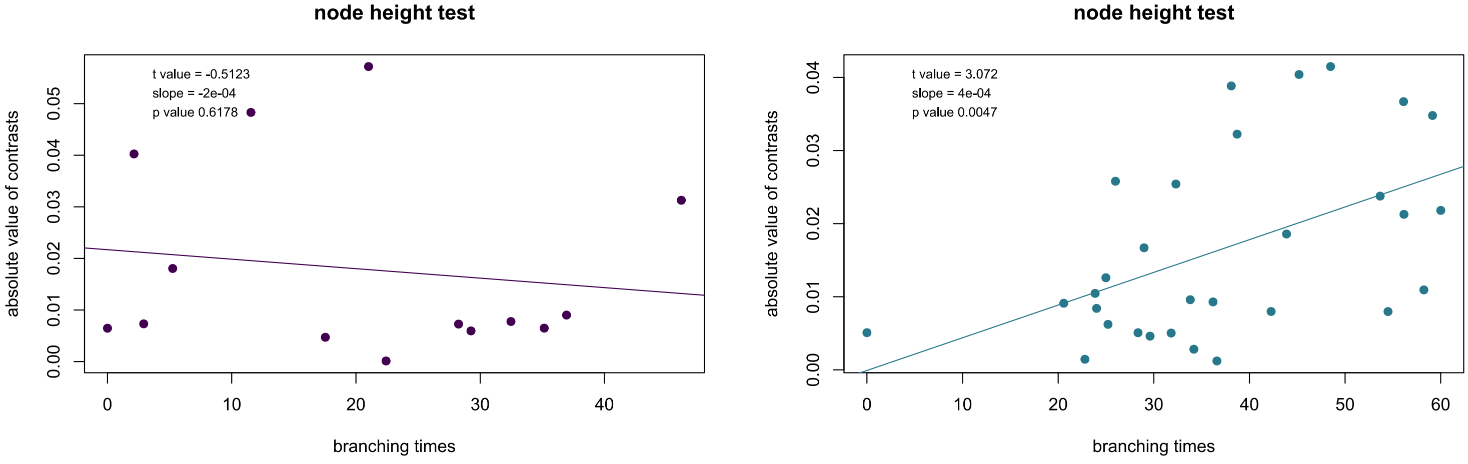
Node height tests of relative limb lengths showed opposite patterns for *Hemidactylus* and *Anolis*, with *Hemidactylus* showing an increase in trait contrasts for younger nodes (p = 0.0842) while *Anolis* showed decrease in trait contrasts for younger nodes (p = 0.0607). This would indicate that rates of evolution have decreased through time in *Anolis* and increased through time for *Hemidactylus*. This apparent increase in rates through time could also be a signature of a constrained model of evolution. Although the p-values for these node height tests are not significant (meaning we cannot definitively say evolutionary rates differ across different node ages), our later model-fitting results support these trends and provide a deeper understanding of the evolutionary dynamics in these clades.



(a) *Hemidactylus* (b) *Anolis*

Figure 3.2: Node height tests for the contrasts of limb length in *Hemidactylus* (purple) and *Anolis* (teal) from the root of the phylogeny (0 branching time) to the present.

Node height tests of independent contrasts of relative toepad area against divergence time again showed opposite patterns for *Hemidactylus* and *Anolis*. However the trends were reversed for each group. *Hemidactylus* showed similar changes in trait contrasts or a decrease in contrasts with decreasing node age ($p = 0.6178$), while *Anolis* showed an increasing contrasts with decreasing node age ($p = 0.0047$). While the regression is not significantly different from 0 in *Hemidactylus*, this could be because of our limited sample size. However, our model fitting analyses do still support these findings in both *Hemidactylus* and *Anolis*.



(a) *Hemidactylus* (b) *Anolis*

Figure 3.3: Node height tests for the contrasts of toepad area in *Hemidactylus* (purple) and *Anolis* (teal) from the root of the phylogeny (0 branching time) to the present.

3.5.2 Model fitting: limb length

We tested OU, BM, and EB models of evolution, fitting each model with our relative limb length data separately for the two clades. We compared models using AIC. For *Hemidactylus* the best-fitting model was OU, while for *Anolis* the best-fitting model was EB. In our bivariate models, *Anolis* fit BM (ΔAIC=2.4443, ΔAICc=2.7331) while *Hemidactylus* fit OU (ΔAIC=4.22011, ΔAICc=2.07155). We calculated stationary trait distributions from these model fits and identified that *Anolis* had greater variance in limb length achieved under Brownian evolution, while *Hemidactylus* was much more constrained in this trait under an OU model (see Figure 3.4).

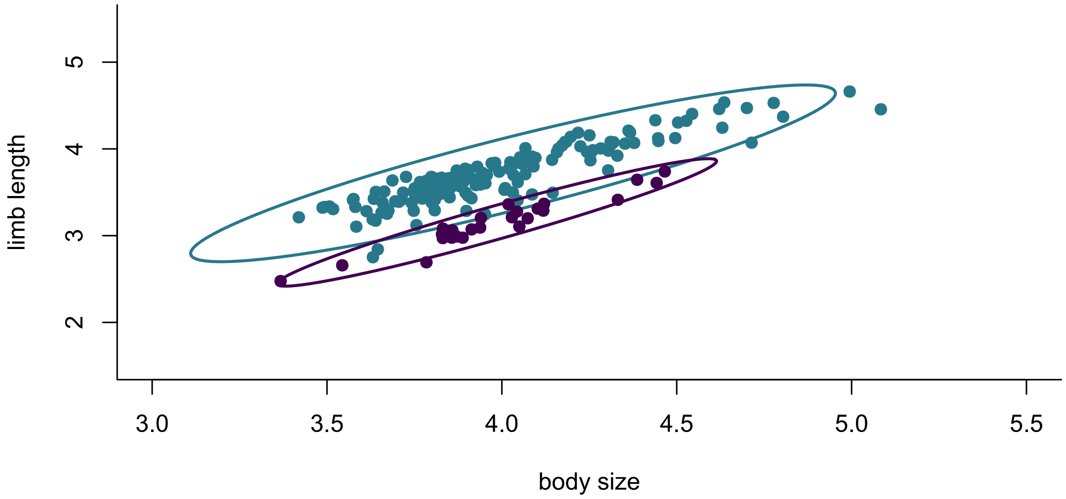


Figure 3.4: Ellipses represent that stationary trait distributions for hind limb length and body size (SVL) in *Hemidactylus* (purple) and *Anolis* (teal) from our model fits (either OU or BM).

Table 3.1: Model comparison for relative hindlimb length for *Hemidactylus* and *Anolis*.

Clade	Model	AIC	AICc
Hemidactylus	EB	-44.27	-43.12
Hemidactylus	BM	-46.27	-45.73
Hemidactylus	**OU**	**-48.83**	**-47.68**
Anolis	**EB**	**-212.57**	**-212.51**
Anolis	BM	-210.55	-210.46
Anolis	OU	-208.55	-208.39

Table 3.2: OU model parameters for relative limb length in *Hemidactylus*.

Parameter	Estimate
α	0.0556
σ^2	0.0007
Z_0	0.0015
Half-life	12.47 Myr

Table 3.3: EB model parameters for relative limb length in *Anolis*.

Parameter	Estimate
b	-0.0310
σ^2	0.0020
Z_0	-0.3165

3.5.3 Model fitting: toepad area

We saw essentially opposite patterns when fitting models of trait evolution to relative toepad area. We again tested OU, BM, and EB models of evolution, fitting each model separately for the two clades and compared models using AIC. For *Hemidactylus* the best-fitting model was BM, while for *Anolis* it was OU.

Table 3.4: Model comparison for relative toepad area for *Hemidactylus* and *Anolis*.

Clade	Model	AIC	AICc
Hemidactylus	EB	-5.96	-3.78
Hemidactylus	**BM**	**-7.65**	**-6.65**
Hemidactylus	OU	-5.65	-3.47
Anolis	EB	-39.31	-38.42
Anolis	BM	-41.31	-40.88
Anolis	**OU**	**-51.41**	**-50.52**

Table 3.5: BM model parameters for relative toepad area in *Hemidactylus*.

Parameter	Estimate
σ^2	0.0006
Z_0	0.00

Table 3.6: OU model parameters for relative toepad area in *Anolis*.

Parameter	Estimate
α	0.1247
σ^2	0.0023
Z_0	0.0042
Half-life	5.56 Myr

In our bivariate models using toepad area and body size we saw the opposite pattern from hind limb length. With toepad area, *Anolis* fit OU (ΔAIC=5.0227, ΔAICc=1.78041) while *Hemidactylus* fit BM (ΔAIC=7.144469, ΔAICc=16.22). We calculated stationary trait distributions from these model fits and identified that *Hemidactylus* had greater variance in toepad area achieved under Brownian evolution, while *Anolis* was more constrained in this trait under an OU model (see Figure 3.5). Additionally, our stationary trait distributions for *Anolis* suggest that the group has not yet reach its trait optima for body size and toepad area.

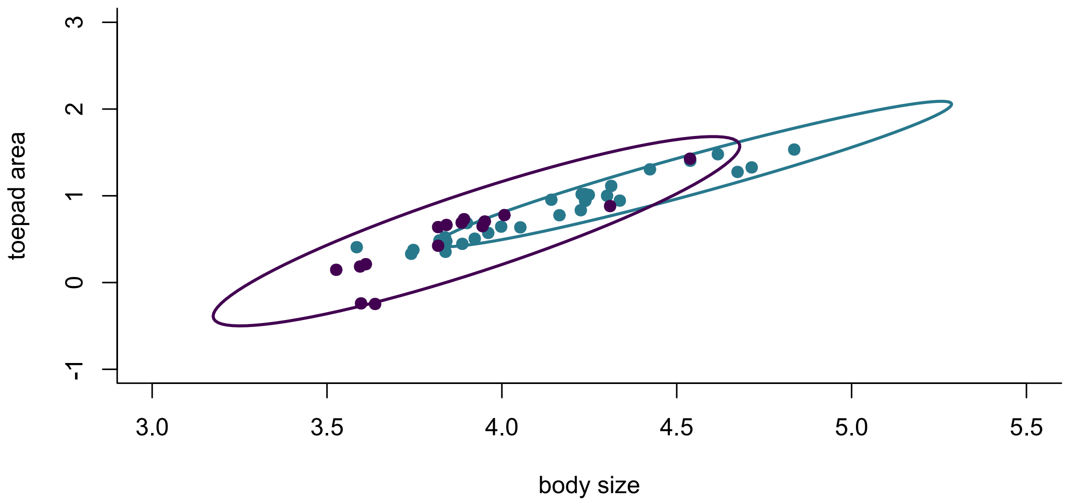


Figure 3.5: Ellipses represent that stationary trait distributions for toepad area and body size (SVL) in *Hemidactylus* (purple) and *Anolis* (teal) from our model fits (either OU or BM)

3.6 Discussion

Anolis lizards are rapidly becoming a model clade for studying urban tolerance and invasion biology [14]. *Hemidactylus* geckos are similarly successful urban exploiters and invaders [15]. These species have independently evolved adhesive toepads—a key innovation for exploiting arboreal environments or scansorial locomotion—and diverged and diversified across a similar range of natural and urban ecologies in tropical and subtropical environments [10, 11]. Toepads are also thought to be an exaptation to urban dwelling on smoother man-made surfaces [36]. Their proximity to urban environments means they are often inadvertently transported to novel environments and their ability to exploit urban habitats makes them excellent invaders once transported [16]. While these clades may have many ecological differences, they contain species which are similarly successful when it comes to urban exploitation and invasion [15, 23, 37, 38]. We sought to understand how these clades have evolved in two traits that are suspected to impact species urban tolerance and invasion, toepad size and limb length. Both traits affect performance, particularly in urban environments where habitat openness and broader substrates usage favors greater sprint speeds [23, 24, 25, 39], and the use of man-made surfaces drives a need for increased clinging performance [19, 20], especially via changes in toepads as claws may be less effective on these smoother substrates [21, 22, 40]. Previous work has shown that these clades differ in their clinging performance [10] and have differences in relative limb lengths as well [4]. We investigated whether these clades varied in rates of trait evolution in traits linked to performance to understand how this may influence their adaptation to novel environments in unique ways.

Across all models and tests, we find that *Hemidactylus* fit a pattern where toepad area evolved under Brownian, unconstrained evolution—achieving greater diversity than the apparently constrained evolution of *Anolis* toepad areas (see figures 3.3, 3.5 and table 3.4). By contrast, the exact opposite is true for relative hindlimb length. Our model-fitting showed that *Hemidactylus* fit a more constrained model of evolution for relative limb length, while *Anolis* fits a Brownian, unconstrained pattern of evolution that achieves greater divergence and diversification over relevant evolutionary timescales (see figures 3.2, 3.4 and table 3.1). These macroevolutionary results suggest opposing processes between traits and clades—*Hemidactylus* are constrained in relative limb length and unconstrained in toepad area, whereas *Anolis* are constrained in toepad area and unconstrained in relative limb length.

Our findings with respect to toepad area align well with previously reported differences in clinging performance. Previous studies of gecko adhesion have suggested that they are "overbuilt" with regards to their clinging performance [18, 41] particularly on smooth substrates [42], and that their clinging performance follows unconstrained Brownian evolution [10]. However, *Anolis* seem to be equally successful at adhesion on smooth substrates, albeit with a different setal morphology [28], and follows a constrained pattern of evolution in clinging performance at phylogenetic scales [4]. Differences in setal and toepad morphology, development, or evolvability could have allowed *Hemidactylus* to diversify more in toepad area than similar traits in *Anolis*.

However, it is also possible that toepad size in *Anolis* is simply not as labile which could be due to genetic [43] or developmental constraints [44] and/or selection. There is evidence that *Anolis* toepads develop in a similar way to those of geckos, in distal to proximal formation [45], although *Hemidactylus* toepads modify this developmental program further through elaboration leading to bifurcating lamellae [45, 46]. This derived morphology and additional developmental modification in *Hemidactylus* may allow them to exhibit additional variation in toepad area that is not accessible in *Anolis* lizards without further developmental modification, although more research is needed in this area.

While they are more labile in the evolution of toepad area, *Hemidactylus* geckos seem to lack lability in hindlimb length, unlike *Anolis* lizards. In fact, in Caribbean *Anolis* lizards, one of the major traits which separates species into ecomorphs based on microhabitat use is limb length [12, 13]. In some *Anolis* lizards it has been shown that trait divergence is linked to genetic variation [47] although this may be driven by selection [48]. Sanger et al. [49] identified that variation in relative limb lengths of ecomorphs is present early on in development. Perhaps *Anolis* lizards have unlocked important developmental or genetic mechanisms that allow greater variation in limb lengths than *Hemidactylus*, although this is not clear from our present study. Differentiation in relative limb lengths early on in development have also been identified in one species of *Hemidactylus* [50]. Even still, this could be driven by selection on hindlimbs rather than toepad size, which would be equally interesting. More work on developmental and genetic architecture across other groups could help provide additional context on the "uniqueness" of *Anolis*.

It is an open question what types of constraints may limit how species diversify over macroevolutionary scales, or whether many-to-one mapping simply provides multiple, largely equivalent avenues for adaptation to selective challenges. Within-population estimation of genetic variances for these traits compared across clades could suggest whether differences in constraints exist at the genetic level, while field studies of selection could indicate whether selective constraints impact the observed patterns [51]. Our results suggest that while the species may be using and exploiting similar habitats, they may differ in the traits that make this exploitation possible due to evolutionary contingency as some studies have suggested [52]. Alternatively, while the two clades are broadly similar in ecology, remaining selective differences could favor one trait or the other in adaptation. *Hemidactylus* are nocturnal, while *Anolis* are diurnal. Although species in these clades do overlap in the Americas and Caribbean islands, specifically due to human-mediated introductions [16, 17, 53, 54], *Hemidactylus* geckos may take greater advantage of human-modified habitats relying on lights which attract insects [53]. More research is needed to determine how these clades differ in ecology and behavior which may mediate some of these changes in trait evolution indirectly. Differences in trait adaptation would make sense if *Anolis* are under greater pressure regarding sprint speed while *Hemidactylus* might be under selection for changes in toepad size. A logical next step would be to determine if species within these clades are more variable in the same respective traits across populations.

3.6.1 Conclusions

Our results suggest that *Hemidactylus* and *Anolis* exhibit opposite trends of trait evolution, with *Hemidactylus* being less constrained in toepad size and *Anolis* being less constrained in limb length. This might suggest that these species are using different strategies to exploit and invade urban and diverse natural environments; however, more studies need to be done to elucidate how (if at all) selection is acting differently upon these two traits in both clades. Alternatively, these traits could theoretically be constrained differently by genetic, developmental, or functional mechanisms or any combination of these. Although these clades contain excellent examples of both urban-tolerant and invasive lizards, they may be using natural and novel urban environments quite differently, and with our analyses it is impossible to rule out whether these differences are due to selection, constraint, or evolutionary contingency. Nevertheless, our study identifies that the two clades exhibit opposite patterns in the lability of trait evolution for toepad size and limb length which is important for understanding how species within these clades might respond due to ongoing urbanization.

Bibliography

[1] Stephen Jay Gould. *Wonderful Life: The Burgess Shale and the Nature of History.* W. W. Norton Company, September 1990. Google-Books-ID: q8BUAgAAQBAJ.

[2] Alan de Queiroz. Contingent predictability in evolution: Key traits and diversification. *Systematic Biology*, 51(6):917–929, December 2002.

[3] Douglas H. Erwin. Evolutionary contingency. *Current Biology*, 16(19):R825–R826, October 2006.

[4] Travis J. Hagey, Scott Harte, Mathew Vickers, Luke J. Harmon, and Lin Schwarzkopf. There's more than one way to climb a tree: Limb length and microhabitat use in lizards with toe pads. *PLoS ONE*, 12(9):1–17, September 2017.

[5] Kathleen L Foster, Jr Garland, Theodore, Lars Schmitz, and Timothy E Higham. Skink ecomorphology: forelimb and hind limb lengths, but not static stability, correlate with habitat use and demonstrate multiple solutions. *Biological Journal of the Linnean Society*, 125(4):673–692, November 2018.

[6] Philip J. Bergmann and Gen Morinaga. The convergent evolution of snake-like forms by divergent evolutionary pathways in squamate reptiles*. *Evolution*, 73(3):481–496, March 2019.

[7] Peter C. Wainwright, Michael E. Alfaro, Daniel I. Bolnick, and C. Darrin Hulsey. Many-to-one mapping of form to function: A general principle in organismal design? *Integrative and Comparative Biology*, 45(2):256–262, April 2005.

[8] Cole J. Thompson, Newaz I. Ahmed, Thor Veen, Catherine L. Peichel, Andrew P. Hendry, Daniel I. Bolnick, and Yoel E. Stuart. Many-to-one form-to-function mapping weakens parallel morphological evolution. *Evolution*, 71(11):2738–2749, November 2017.

[9] Daniel S. Moen. What determines the distinct morphology of species with a particular ecology? the roles of many-to-one mapping and trade-offs in the evolution of frog ecomorphology and performance. *The American Naturalist*, 194(4):E81–E95, October 2019.

[10] Travis J. Hagey, Josef C. Uyeda, Kristen E. Crandell, Jorn A. Cheney, Kellar Autumn, and Luke J. Harmon. Tempo and mode of performance evolution across multiple independent origins of adhesive toe pads in lizards. *Evolution*, 71(10):2344–2358, October 2017.

[11] Aryeh H Miller and James T Stroud. Novel tests of the key innovation hypothesis: Adhesive toepads in arboreal lizards. *Systematic Biology*, 71(1):139–152, January 2022.

[12] Jonathan B. Losos. The evolution of form and function: Morphology and locomotor performance in west indian anolis lizards. *Evolution; International Journal of Organic Evolution*, 44(5):1189–1203, August 1990.

[13] Jonathan B. Losos. *Lizards in an Evolutionary Tree Ecology and Adaptive Radiation of Anoles*. University of California Press, August 2009.

[14] Martha M. Muñoz, Luke O. Frishkoff, Jenna Pruett, and D. Luke Mahler. Evolution of a model system: New insights from the study of anolis lizards. *Annual Review of Ecology, Evolution, and Systematics*, 54(Volume 54, 2023):475–503, November 2023.

[15] Robbie Weterings and Kai C Vetter. Invasive house geckos (hemidactylus spp.): their current, potential and future distribution. *Current Zoology*, 64(5):559–573, October 2018.

[16] Jesse B Borden and S Luke Flory. Urban evolution of invasive species. *Frontiers in Ecology and the Environment*, 19(3):184–191, 2021.

[17] Kevin M. Enge, Mark S. Robson, and Kenneth L. Krysko. Reptile surveys of pine rockland habitat in six miami-dade county parks. *Florida Scientist*, 67(3):194–204, 2004.

[18] Duncan J. Irschick, Christopher C. Austin, Ken Petren, Robert N. Fisher, Jonathan B. Losos, and Olaf Ellers. A comparative analysis of clinging ability among pad-bearing lizards. *Biological Journal of the Linnean Society*, 59(1):21–35, 1996.

[19] Zani. The comparative evolution of lizard claw and toe morphology and clinging performance. *Journal of Evolutionary Biology*, 13(2):316–325, March 2000.

[20] Duncan J. Irschick, Anthony Herrel, and Bieke Vanhooydonck. Whole-organism studies of adhesion in pad-bearing lizards: creative evolutionary solutions to functional problems. *Journal of Comparative Physiology A*, 192(11):1169–1177, November 2006.

[21] Emily R Naylor and Timothy E Higham. Attachment beyond the adhesive system: The contribution of claws to gecko clinging and locomotion. *Integrative and Comparative Biology*, 59(1):168–181, July 2019.

[22] Cleo H Falvey, Kevin J Aviles-Rodriguez, Travis J Hagey, and Kristin M Winchell. The finer points of urban adaptation: intraspecific variation in lizard claw morphology. *Biological Journal of the Linnean Society*, 131(2):304–318, September 2020.

[23] Kristin M. Winchell, R. Graham Reynolds, Sofia R. Prado-Irwin, Alberto R. Puente-Rolón, and Liam J. Revell. Phenotypic shifts in urban areas in the tropical lizard anolis cristatellus. *Evolution*, 70(5):1009–1022, May 2016.

[24] Kristin M. Winchell, Inbar Maayan, Jason R. Fredette, and Liam J. Revell. Linking locomotor performance to morphological shifts in urban lizards. *Proceedings of the Royal Society B: Biological Sciences*, 285(1880):20180229, June 2018.

[25] Jonathan B Losos and Barry Sinervo. The effecrs of morphology and perch diameter on sprint performance of *Anolis* lizards. page 8.

[26] Dirk Bauwens, Theodore Garland Jr., Aurora M. Castilla, and Raoul Van Damme. Evolution of sprint speed in lacertid lizards: Morphological, physiological, and behavioral covariation. *Evolution*, 49(5):848–863, 1995.

[27] B. A. Goodman. Nowhere to run: the role of habitat openness and refuge use in defining patterns of morphological and performance evolution in tropical lizards. *Journal of Evolutionary Biology*, 22(7):1535–1544, July 2009.

[28] Austin M. Garner, Michael C. Wilson, Caitlin Wright, Anthony P. Russell, Peter H. Niewiarowski, and Ali Dhinojwala. The same but different: setal arrays of anoles and geckos indicate alternative approaches to achieving similar adhesive effectiveness. *Journal of Anatomy*, 238(5):1143–1155, May 2021.

[29] Liam J Revell, Luke J Harmon, and David C Collar. Phylogenetic signal, evolutionary process, and rate. *Systematic biology*, 57(4):591–601, 2008.

[30] Camilo Cortés-Parra, Martha L. Calderón-Espinosa, and Adriana Jerez. Claws and toepads in mainland and island anolis (squamata: Dactyloidae): Different adaptive radiations with intersectional morphospatial zones. *Journal of Anatomy*, 240(5):850–866, 2022.

[31] Joseph Felsenstein. Phylogenies and the comparative method. *The American Naturalist*, 125(1):1–15, 1985.

[32] Robert P Freckleton and Paul H Harvey. Detecting non-brownian trait evolution in adaptive radiations. *PLoS biology*, 4(11):e373, 2006.

[33] M. W. Pennell, J. M. Eastman, G. J. Slater, J. W. Brown, J. C. Uyeda, R. G. Fitzjohn, M. E. Alfaro, and L. J. Harmon. geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, 30:2216–2218, 2014.

[34] Hirotogu Akaike. Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike*, pages 199–213. Springer, 1998.

[35] Julien Clavel, Gilles Escarguel, and Gildas Merceron. mvmorph: an r package for fitting multivariate evolutionary models to morphometric data. *Methods in Ecology and Evolution*, 6:1311–1319, 2015.

[36] Ruth A Hufbauer, Benoît Facon, Virginie Ravigné, Julie Turgeon, Julien Foucaud, Carol E Lee, Olivier Rey, and Arnaud Estoup. Anthropogenically induced adaptation to invade (aiai): contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications*, 5(1):89–101, January 2012.

[37] Walter E. Meshaka. Colonization dynamics of two exotic geckos (hemidactylus garnotii and h. mabouia) in everglades national park. *Journal of Herpetology*, 34(1):163–168, 2000.

[38] James T. Stroud, Marie Colom, Pedro Ferrer, Nicholas Palermo, Veronica Vargas, Martina Cavallini, Jesus Lopez, and Ian Jones. Behavioral shifts with urbanization may facilitate biological invasion of a widespread lizard. *Urban Ecosystems*, 22(3):425–434, June 2019.

[39] Jason J. Kolbe, Andrew C. Battles, and Kevin J. Avilés-Rodríguez. City slickers: poor performance does not deter anolis lizards from using artificial substrates in human-modified habitats. *Functional Ecology*, 30(8):1418–1429, 2016.

[40] Emily R. Naylor. *Function, Morphology, and Evolution of Gecko Locomotion: The Importance of Ecology and Substrate.* Ph.d., University of California, Riverside, United States – California, 2020.

[41] Kellar Autumn, Yiching A. Liang, S. Tonia Hsieh, Wolfgang Zesch, Wai Pang Chan, Thomas W. Kenny, Ronald Fearing, and Robert J. Full. Adhesive force of a single gecko foot-hair. *Nature*, 405(6787):681–685, June 2000.

[42] Bieke Vanhooydonck, Anemone Andronescu, Anthony Herrel, and Duncan J. Irschick. Effects of substrate structure on speed and acceleration capacity in climbing geckos. *Biological Journal of the Linnean Society*, 85(3):385–393, 2005.

[43] Yan Liu, Qian Zhou, Yongjun Wang, Longhai Luo, Jian Yang, Linfeng Yang, Mei Liu, Yingrui Li, Tianmei Qian, Yuan Zheng, Meiyuan Li, Jiang Li, Yun Gu, Zujing Han, Man Xu, Yingjie Wang, Changlai Zhu, Bin Yu, Yumin Yang, Fei Ding, Jianping Jiang, Huanming Yang, and Xiaosong Gu. Gekko japonicus genome reveals evolution of adhesive toe pads and tail regeneration. *Nature Communications*, 6(1):10033, November 2015.

[44] Priscila S. Rothier, Monique N. Simon, Gabriel Marroig, Anthony Herrel, and Tiana Kohlsdorf. Development and function explain the modular evolution of phalanges in gecko lizards. *Proceedings of the Royal Society B: Biological Sciences*, 289(1966):20212300, January 2022.

[45] Aaron H Griffing, Tony Gamble, Martin J Cohn, and Thomas J Sanger. Convergent developmental patterns underlie the repeated evolution of adhesive toe pads among lizards. *Biological Journal of the Linnean Society*, 135(3):518–532, March 2022.

[46] Aaron H Griffing, Tony Gamble, Ashmika Behere, Timothy E Higham, Greta M Keller, John Resener, and Thomas J Sanger. Developmental patterns underlying variation in form and function exhibited by house gecko toe pads. *Integrative and Comparative Biology*, 64(5):1494–1504, November 2024.

[47] Joel W. McGlothlin, Megan E. Kobiela, Helen V. Wright, D. Luke Mahler, Jason J. Kolbe, Jonathan B. Losos, and Edmund D. III Brodie. Adaptive radiation along a deeply conserved genetic line of least resistance in anolis lizards. *Evolution Letters*, 2(4):310–322, August 2018.

[48] Joel W. McGlothlin, Megan E. Kobiela, Helen V. Wright, Jason J. Kolbe, Jonathan B. Losos, and Edmund D. Brodie. Conservation and convergence of genetic architecture in the adaptive radiation of anolis lizards. *The American Naturalist*, 200(5):E207–E220, November 2022.

[49] Thomas J. Sanger, Liam J. Revell, Jeremy J. Gibson-Brown, and Jonathan B. Losos. Repeated modification of early limb morphogenesis programmes underlies the convergence of relative limb length in anolis lizards. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729):739–748, August 2011.

[50] Juliana Luzete, Lilian G Giugliano, and Julia Klaczko. Evaluating the drivers and engines of morphological diversification in the invasive gecko hemidactylus mabouia (moreau de jonnès, 1818) (squamata: Gekkonidae). *Biological Journal of the Linnean Society*, 137(3):475–495, November 2022.

[51] Stevan J Arnold. Constraints on phenotypic evolution. *The American Naturalist*, 140:S85–S107, 1992.

[52] Anna Zimin, Sean V. Zimin, L. Lee Grismer, Aaron M. Bauer, David G. Chapple, Jacob Dembitzer, Uri Roll, and Shai Meiri. Microhabitat and adhesive toepads shape gecko limb morphology. *Integrative Zoology*, n/a(n/a).

[53] Walter E. Meshaka, Betty Ferster, Maxwell L. Meshaka, and Molly L. Meshaka. Exotic species dominance in the herpetofaunal community of a new residential development in southern florida: Implications for faunal turnover rates. *Florida Scientist*, 71(1):65–75, 2008.

[54] Tomás Michel Rodríguez Cabrera, Rosalina Montes Espín, and Javier Torres. Predation attempt by a tropical house gecko, hemidactylus mabouia (gekkonidae), on a cuban blue anole, anolis allisoni (dactyloidae) in south-central cuba. *Reptiles & Amphibians*, 29(1):159–161, 2022.

Chapter 4

Predicting lizard invasion probabilities and incorporating phylogeny in random forests

4.1 Attributions

4.1.1 Co-Authors

Josef C. Uyeda, Professor, Department of Biological Sciences, Virginia Tech, Blacksburg, VA

4.1.2 Acknowledgments

4.2 Abstract

Accurate prediction of which species would be successful invasive species upon introduction would greatly facilitate management and mitigation, and help reveal what traits affect invasive potential. With the availability of large trait databases, such predictions are becoming tractable. However, such databases invariably miss some important traits that drive variation in invasiveness across species. One way of accounting for such traits when they are unavailable is to leverage the predictive power of the phylogeny—provided the underlying traits have phylogenetic signal. Here, we sought to predict species invasion probabilities across lizards using phylogenetic information, species traits, climate and geographic data. We used random forests, a dataset from the Global Register of Introduced and Invasive Species (GRIIS), and published trait data for lizards. We then used this model to predict the invasion capacity of species across 486 species and to identify traits that are important for predicting species invasion probabilities. Our results suggest that phylogenetic information can improve predictive accuracy of invasion potential by accounting for unobserved or difficult to measure traits in random forests. We incorporated species traits and climate data and of these features maximum body size, largest clutch size, mean monthly temperature range, and minimum annual precipitation were identified as some of the best predictors for species classified as introduced. We mapped invasive probabilities of native species across

the globe and identified that tropical and subtropical countries are acting as sources of many invasive species.

4.3 Introduction

Global change is increasing the spread and impacts of invasive species [1] and these invasive species are known to have significant detrimental impacts on native biodiversity [2, 3]. Furthermore, there is increasing evidence that human-driven impacts are homogenizing environments, with invasive species occupying a larger proportion of global species diversity—particularly in urban habitats [4, 5]. Improving our ability to anticipate potential invasive species before their introduction would significantly enhance efforts to assess risk, prevent introductions, and mitigate or prepare for their consequences.

Previous work has identified utility in predicting invasion success using a variety of methods, finding that species traits and climate are good predictors of species-level invasion [6, 7, 8]. Predicting future invasive species could allow management agencies to strengthen restrictions on the transport of species which have high probabilities of becoming invasive if introduced [9]. It is well established that predicting invasive species requires knowledge of current invaders, the traits they possess, and the environments they naturally occur in. Including this information helps us understand the environments these species may be adapted to and whether they possess traits which make it easier to invade and establish populations in novel environments [6]. While relying on current invasive species may not provide a comprehensive list given the many contingent events that must occur to lead to a successful invasion, it can help us to identify traits which make it more likely for a species to become invasive when introduced.

Various types of traits may make it more likely for a species to invade. Some of these traits can be considered intrinsic phenotypes of individuals within a species, like body size, reproductive traits, and thermal or physiological tolerances. Often traits related to dispersal and rapid reproductive capacity will increase species' invasion potential as it makes them more likely to spread and establish populations [10]. Traits related to generalism can make it easier for species to survive in novel environments [11, 12].

Other traits may represent species-level features that represent a combination of the traits of organisms and extrinsically-mediated conditions environmental [1]. For example, geographic range size can either affect invasion probability by simply increasing opportunities for the introduction of species to new areas globally or by reflecting broad, generalist tolerances to wide geographic areas. Species may be adapted to certain climatic conditions that make them pre-adapted to successfully invade new areas by serendipitous matching of the climatic conditions in their native range [6]. Many species could be physiologically adapted to the conditions in the areas they occur natively, which combines the constraints of intrinsic and extrinsic limitations on species invasion potential. However, species can often also invade areas that are climatically distinct from their native ranges [13]. Such invasion may suggest biological or geographic constraints rather than intrinsic physiological constraints [14], or may reflect high adaptive capacity for rapid evolutionary change in new environments.

Many traits which may be important for a species invasive capacity may be difficult to observe or measure directly. For instance, thermal physiology in lizards has been identified as an important predictor of success in urban habitats where many invasive species are present [15, 16]. While field body temperatures may be relatively easy to measure, key traits like critical thermal maximum or minimum temperatures are more difficult to measure [17, 18]. Even more difficult, traits with the same measured values in trait databases may nevertheless have different constraints on their development or variation that may impact their ability to respond to novel selective factors (Chapter 3), representing "hidden states" with different evolutionary properties [19].

Similarly, there are many traits that may influence which species are transported through the pet trade, which is a common way lizards may be introduced [20, 21]. These include docility, rarity, or desirability based on features of their appearance that may be difficult to define and measure, and often may encompass many multifactorial traits [22]. In these cases it may be hard to include these traits for making species-level predictions about invasion capacity. Nevertheless, these more complex or difficult to measure traits may still have phylogenetic signal [23, 24]. This means that closely related species will be more similar in these traits and might be clustered on the phylogeny. This makes it possible to use the phylogeny itself as a proxy for the myriad unmeasured traits that may improve prediction in the modeling of invasion probabilities.

Pili et al. [8] recently demonstrated the utility of species traits and macroecology for predicting species invasions across reptiles and amphibians. In this paper, we sought to add to this work by predicting invasion probabilities across lizards using species traits, climate, and geographic data (similar to as in Pili et al.)—but also by explicitly incorporating phylogenetic information (rather than using phylogenetic imputation). This distinction is important as incorporating the phylogeny itself can provide additional information, acting as a proxy for unmeasured or complex traits which would be otherwise difficult to capture. We tested various different methods of incorporating phylogenetic relationships between species to determine the method with highest predictive accuracy. Using our best model we identify traits that are important for predicting species invasion probabilities, and we verify that our incorporation of phylogenetic relationships improves predictions. We chose to limit our model to species with complete data rather than using phylogenetic imputation as we believe this could artificially influence feature importance for predicting invasion probabilities. We use this model to predict species invasion probabilities across 486 species. We use these probabilities to set a threshold for invasion capacity which could be useful for setting restrictions on the transport of species in the pet trade. We also map species invasion probabilities globally to determine areas that are acting as sinks and sources of invasive species in order to mitigate the spread of invasive species to and from these areas.

4.4 Methods

4.4.1 Data collection

We used data for species invasions from the Global Register of Introduced and Invasive Species (GRIIS) [25]. We compiled species occurrences from the Global Biodiversity Information Facility (GBIF) from 2013-2023 [26] and cleaned our data using the R package *CoordinateCleaner* [27]. We then used the country codes from our GBIF data and filtered these to exclude observations for species in countries where they were introduced in the GRIIS. We used our observations of species in their native ranges to acquire bioclimatic variables for each occurrence from WorldClim [28]. We then summarized these bioclimatic variables by taking the minimum, maximum, and mean values for each species. We excluded introduced occurrences to avoid over-inflating of climate variable importance in predicting invasion capacity. We combined these data, published trait data from Meiri 2018 [29], and a published phylogenetic tree [30]. The trait data includes measurements of body size, maximum and minumum body temperatures, and clutch sizes. It also includes categorizations of species microhabitat usage, diet, diel activity, foraging mode, and reproductive mode.

For our invasion dataset, we filtered the GRIIS data to label any species that was classified as introduced or invasive in any country as “introduced” to obtain species-level data on introductions. While this may not be a direct indicator of the invasive impact of a species, we felt it was the best way to categorize our data to allow us to predict invasiveness at the species level. This classification for each species was combined with our phylogenetic, climate, and trait data to produce the final dataset, where any species that was not included in our GRIIS dataset was classified as “native” in our dataset. We divided this data using 70 percent for training and 30 percent for testing performance. The use of the GRIIS and filtering it to produce species-level classifications means that our predictive model is splitting species into “native”—i.e. only present in their native range—and “introduced”—i.e. introduced in at least one location, and included in the GRIIS.

4.4.2 Incorporating phylogenetic information

We tested three methods for incorporating phylogenetic structure into random forests. For the first, we converted the tree into bipartitions using the R package *phangorn* [31] and removed all tip bipartitions (i.e. the minimum clade size on either side of a bipartition was 2). This method retains only the hierarchical structure of the phylogeny (i.e. the topology) and does not use branch lengths. For the second, we converted the phylogeny to a distance matrix using the R package *ape* [32]. This method distills the phylogenetic branch lengths between species into a pairwise distance for all species combinations. For the third method, we used phylogenetic eigenvector maps using the R package *MPSEM* [33]. This method incorporates branch lengths as in the case of the distance matrix, but distills the phylogeny into eigenfunctions that map tip species across progressive eigenvectors of the phylogenetic distance matrix. We compared the performance of each of these methods by using each method to predict invasiveness on our test dataset and measured performance using the area under receiver operating curves (ROC AUC).

4.4.3 Invasion probabilities and feature importance

We used the R package *Ranger* [34] to implement a probability forest algorithm [35]. To account for class imbalance, we up-sampled the minority class (introduced) and down-sampled the majority (native). We used permutation importance which compares the decrease in performance of our classifier when individual features are permuted to assess variable importance [36].

We used the best-performing method for incorporating phylogenetic data, along with trait and climate data on our training dataset and this model contains all features. We also trained a model which included only trait and climate data (without phylogenetic information) to test its performance. We trained another model with only selected features. To do this, we used the R package *Boruta* [37] to perform feature selection due to the large number of features our dataset contained. Boruta performs feature selection by creating "shadow features" from true features by permuting their values and statistically comparing the feature importance of true features against importance of shadow features to determine which features should be included. We then compared the performance of each of these models on our test dataset using ROC AUC.

In addition to using Boruta for feature selection, we developed a method which allows for comparison of true features to randomly simulated phylogenetic traits to determine if features are more informative than, random, phylogenetically-structured traits. Given the large number of features that can potentially be selected via random forests, there is an increased probability that traits that have phylogenetic signal could be selected as effective predictors merely due to their phylogenetic signal [38, 39, 40]. Thus, we sought to use a method that allows us to determine which traits have more predictive power related to our target classification relative to traits that are informative only due to their phylogenetic signal. Note that the overall goal of prediction may still be aided by characters that are informative only due to their phylogenetic signal. Therefore, we do not use the selected feature set in our model comparison. Rather, we use it to identify the key traits that have strong evidence for predictive power. More information and comparisons between Boruta and our method are provided in Chapter 5.

We used the R package *lime* [41] to fit reduced complexity models to help us understand how features were informing the model predictions. LIME, which fits a local model around specific predictions making the model more interpretable, allowed us to determine how trait values were used to classify native and introduced species and which traits were the most informative for a given prediction.

4.4.4 Mapping invasion probabilities

We used the best model to predict invasion probabilities for all species. We then divided the global map into grid cells and used range maps from Roll et al. [42] to determine which species were present in each grid cell, splitting species into their native and invasive ranges by filtering this data to exclude ranges in countries where species were introduced using the GRIIS. We took the mean invasion probability for the native species present in each grid cell and plotted this information globally to identify areas that are acting as sources of invasive species (i.e. containing many native species that have high invasion probabilities). We also plotted the same invasion probabilities but only for areas where species are invasive. This represents a summary of the invasiveness among the invasive species community in areas that are presently invaded. Notably, our summaries are sensitive to the number of invasive species present, and we therefore compared them to the species richness globally for the 486 lizards present in our dataset.

4.5 Results

4.5.1 Incorporating phylogenetic information

When comparing ROC AUC for the three different methods of incorporating phylogeny, we saw better performance when using all bipartitions of the phylogeny rather than the phylogeny as a distance matrix or phylogenetic eigenvector maps. For this reason we used only bipartitions of the phylogeny for the remainder of the analyses. Bipartitions also give the benefit of being easy to visualize and interpret, but carry no information about branch lengths or phylogenetic distance besides the hierarchical structure of the phylogeny.

Out of our 500+ features, Boruta selected a subset of 43 features, significantly reducing model complexity. A subset of trait, climate, and phylogenetic data was retained. See the supplementary material for a full list of original and selected features.

4.5.2 Invasion probabilities and feature importance

We compared the ROC AUC for all models—i.e. a model containing all features, a model with only trait and climate data, and a model with only selected features from Boruta—and our best performing model included all features (see figure 4.1 for model comparison). Thus, we retained this model for predicting invasion probabilities. Examples of species with relatively high invasion probabilities despite not being included in the GRIIS were *Thecadactylus rapicauda*, *varanus bengalensis*, and *Lampropholis guichenoti*.

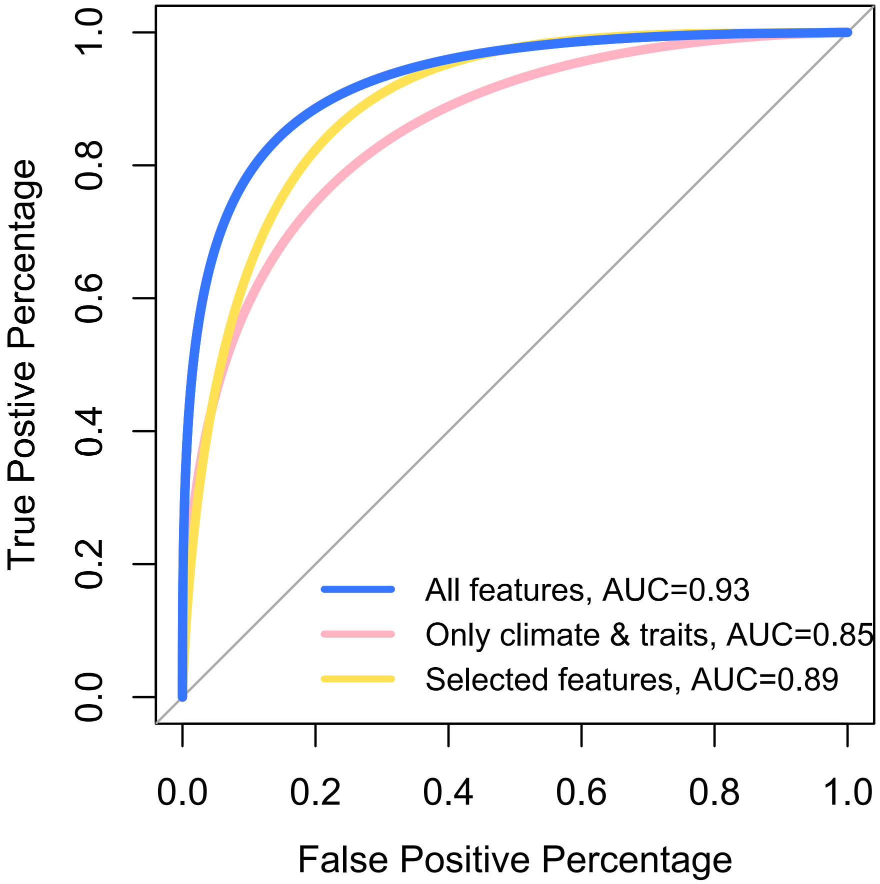
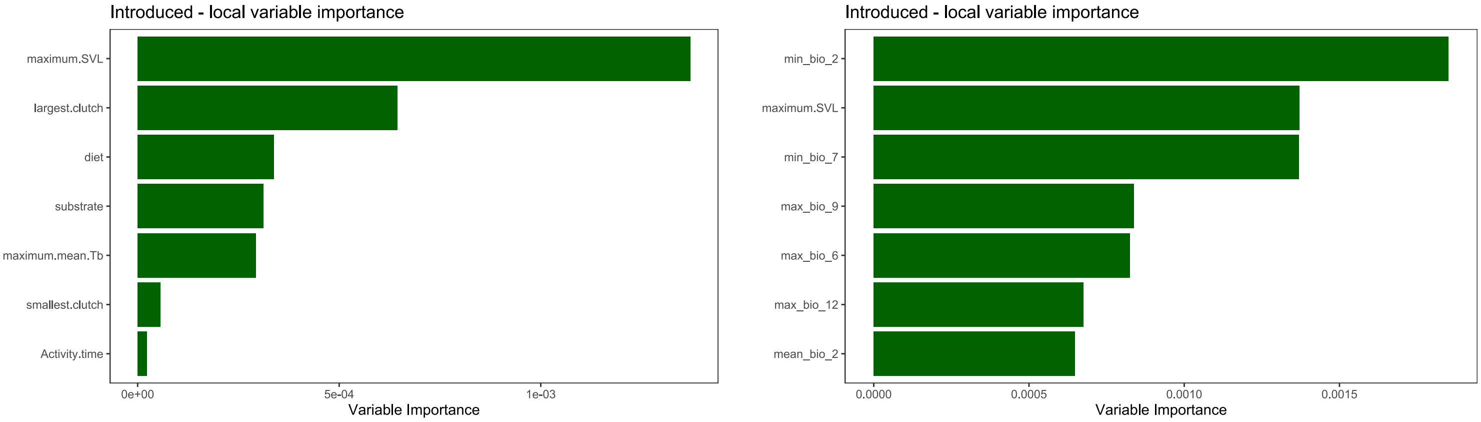


Figure 4.1: ROC plots for three different models all features, including phylogeny, trait and climate data (blue), excluding phylogeny (pink), and selected features from Boruta (yellow). Area under the curve is included in the legend, values closer to 1 indicate better performance.

We used permutation variable importance to determine which features were useful for splitting species into those that exist only in their native range and those that have successfully been introduced in new areas. We also estimated local variable importance, i.e. which features specifically best predicted whether a species was introduced.

Of the traits we incorporated body size (maximum snout-vent length or SVL), largest clutch size, diet, and substrate were the best predictors of introduced species (figure 4.2a). Our LIME analysis suggested that having a large body size and large clutch size made a species more likely to be classified as introduced (see figure 4.3 for an example). These results are consistent with a recent publication across reptiles and amphibians [8].

Among climate data, the features that were identified as the best predictors included minimum coldest month temperature (BIO6), mean diurnal temperature range (BIO2), and maximum annual precipitation (BIO12) for species classified as introduced (figure 4.2b). Our LIME analysis suggested that species with low mean temperature ranges (BIO2) were more likely to be introduced (see figure 4.3 for examples). It also suggested that low minimum precipitation (BIO16) predicts species are more likely to be classified as native. Species with low maximum coldest month temperature (BIO6) were more likely to be classified as native.



(a) Species Traits

(b) Climate and species traits

Figure 4.2: Local variable importance for introduced classification for species level traits (a) and species and climate traits (b)

Selected phylogenetic bipartitions can be used to identify clades with high invasiveness. Bipartitions that were classified as important highlighted clades containing agamids and chameleons, plated lizards, and basilisk lizards (figure 4.4).

4.5.3 Mapping invasion probabilities

Our results from mapping species invasion probabilities for species in their native ranges suggest that tropical and subtropical environments are often acting as sources of invasive species (figure 4.6). Islands are often sources of invasive species with many island areas having high invasion probabilities. Forested areas whether tropical or temperate, tend to have species with higher invasion probabilities. However, desert areas contain species with low invasion probabilities comparatively. This aligns well with our results from variable importance of climate data.

Plotting species invasion probabilities in their introduced ranges revealed higher vulnerability in tropical and subtropical environments (figure 4.7a). However, often the introduced species ranges were not included in range maps from Roll et al. [42], so we also plotted the

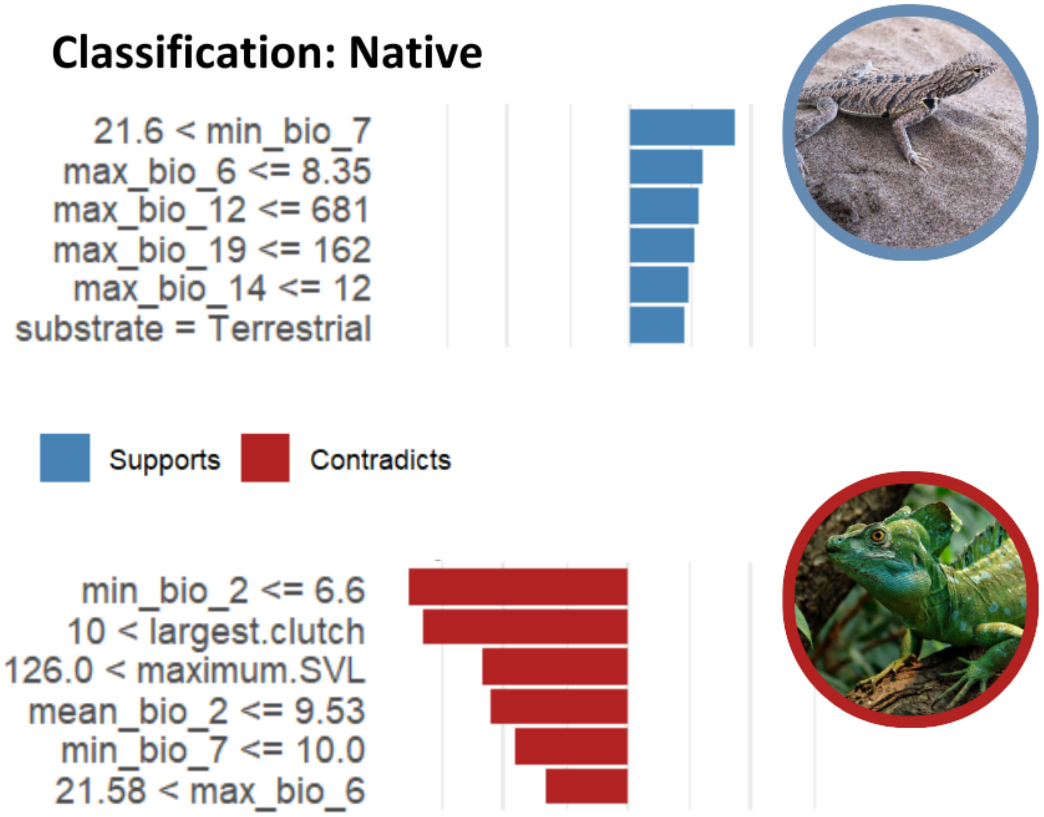


Figure 4.3: LIME results showing the traits that support or contradict species being classified as native for two representative species, *Uma exsul* an endemic species (photo from Claudio Cantú Muñiz on iNaturalist,
license: https://creativecommons.org/licenses/by/4.0/) and *Basiliscus plumifrons* an invasive species (photo from Joachim S. Müller on flickr,
license: https://creativecommons.org/licenses/by-nc-sa/2.0/).

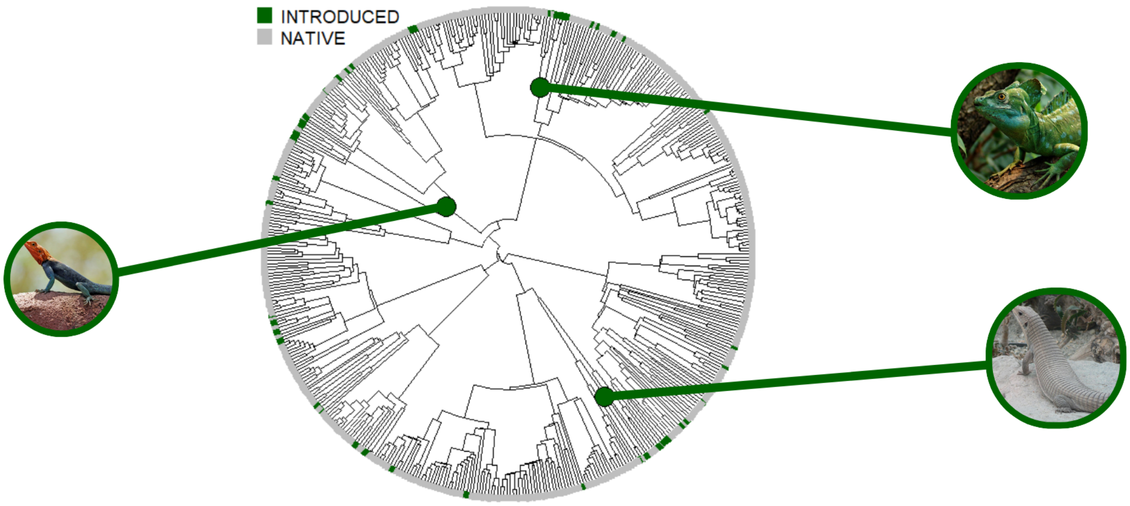


Figure 4.4: Phylogeny of species with important nodes (using selected features and variable importance) highlighted with representative species for each clade, *Agama agama* (left, photo from Peter Prokosch, https://www.grida.no/resources/2389, license: https://creativecommons.org/licenses/by-nc-sa/2.0/), *Basiliscus plumifrons* (top right, photo from Joachim S. Müller on flickr, license: https://creativecommons.org/licenses/by-nc-sa/2.0/), and *Broadleysaurus major* (bottom right, photo from Animalia.bio, license: https://creativecommons.org/licenses/by-sa/3.0/)

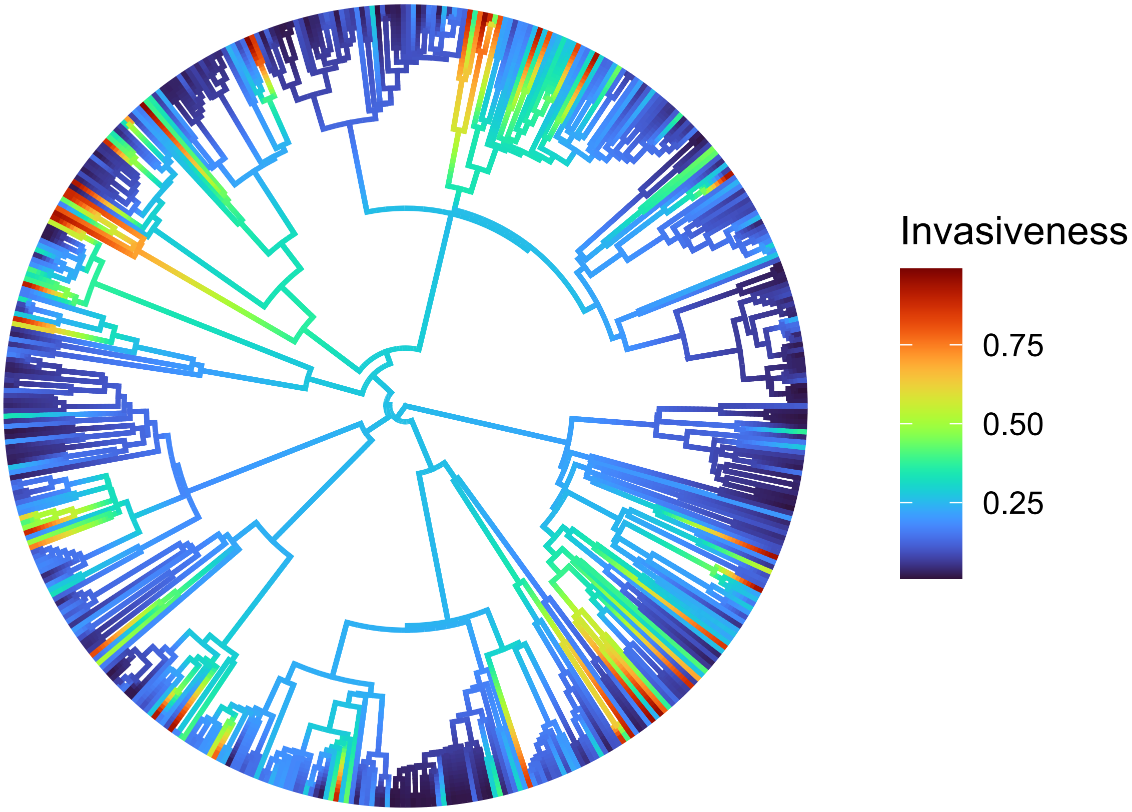


Figure 4.5: Invasion probabilities from the best model plotted on the phylogeny as a continuous trait (traits are only meaningful at tips as this is not a meaningful ancestral state reconstruction).

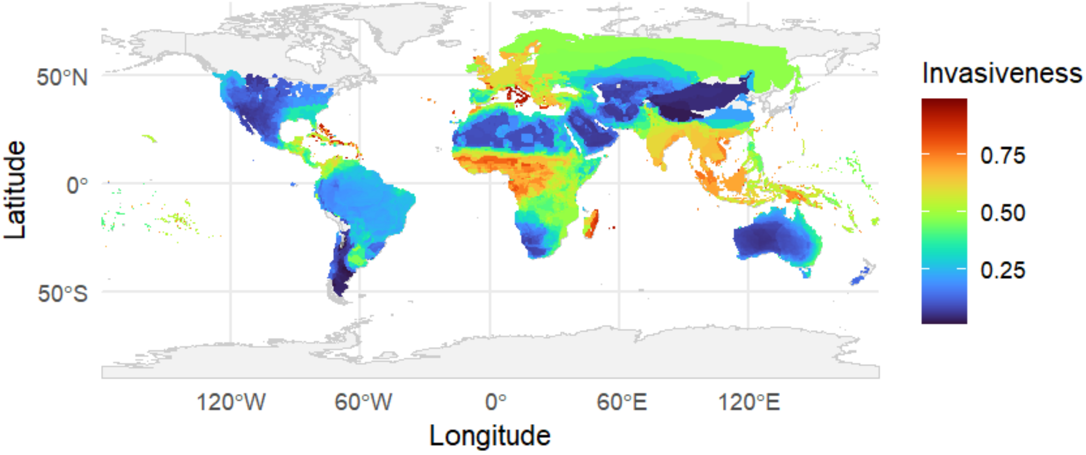
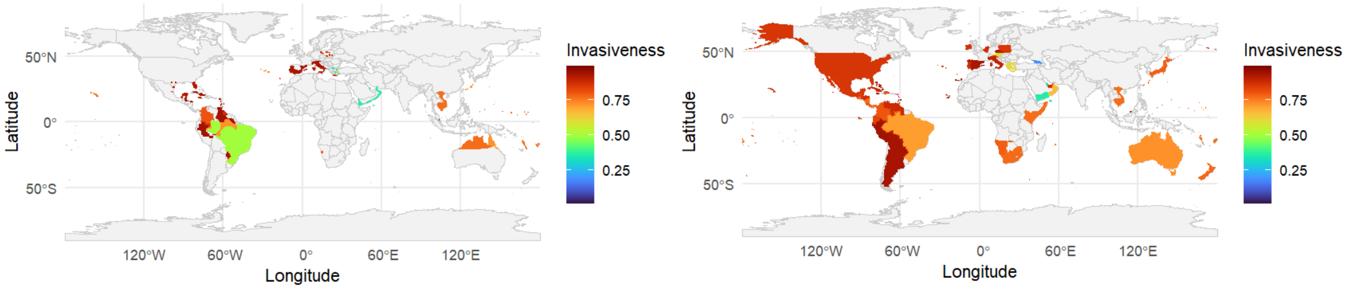


Figure 4.6: Invasion probability of species in their native range. These are reflective of areas that are acting as sources of invasive species. Areas in red represent higher invasion probabilities while blue is lower.



(a) Invasive ranges

(b) Invasive species by country

Figure 4.7: Invasion probabilities for species present in (a) invasive ranges based on Roll et al. and (b) countries that have been invaded by species in the GRIIS.

mean invasion probability of introduced species by country as well (figure 4.7b). While this information is less fine-scaled than true ranges, it does give a better idea of where species are often presently invading. This map provided similar results to that of our native plots, we saw that invasive species were not invading desert environments but that island areas are particularly vulnerable. However, this data does not necessarily mean that more invasive species are present in these areas, simply that the species which are present have traits which suggest they can invade more easily than other species. In order to account for the number of species present in each area we also plotted species richness (figure 4.8). Notably, although several desert environments exhibit high species richness, these areas still maintain low invasiveness scores.

4.6 Discussion

The traits species possess can help us determine how they will respond to new environments. A certain subset of traits may provide species with a pre-adaptation to a specific environment [43, 44] or they may make species more adaptable upon introduction [45, 46]. These traits are often dependent on phylogeny and past evolutionary events due to evolutionary contingency [47, 48]. In addition to knowing species traits, incorporating phylogenetic information can help improve our predictions by accounting for traits that may be difficult to measure or observe or be unequally phylogenetically distributed [49]. We sought to predict species invasion probabilities in lizards by incorporating both species traits and phylogenetic information using random forests.

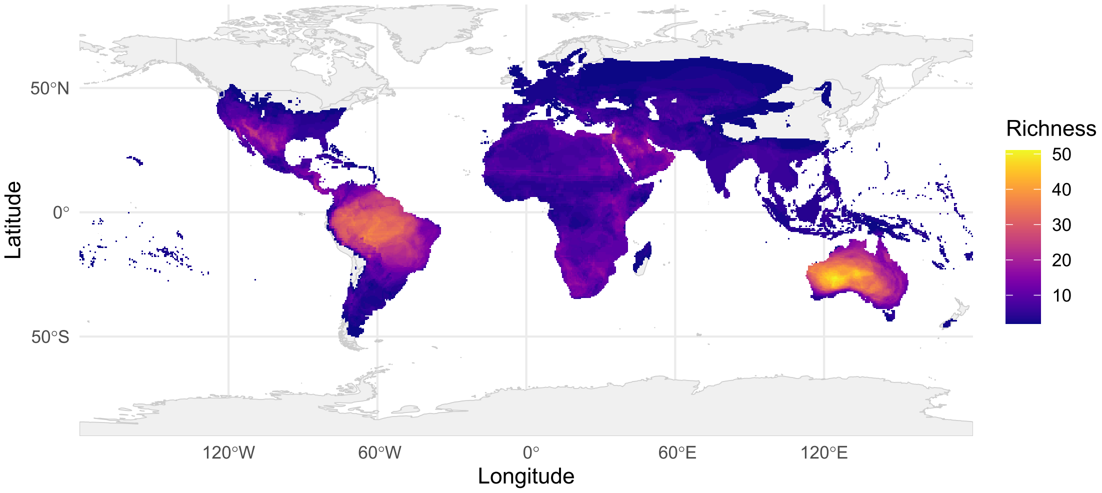


Figure 4.8: Species richness plotted globally where lighter colors correspond to higher species richness.

4.6.1 Invasion probabilities and feature importance

Our results suggest that including phylogenetic information did allow for improvements in our predictions. Our results when comparing different methods showed the retaining primarily the hierarchical structure of the phylogeny can provide excellent performance. This method is also beneficial as it allows for interpretability by allowing us to identify relevant nodes, even if information about branch lengths are excluded. However, only some of the phylogenetic features were selected as important through feature selection (see figure 4.4 and supplementary material). This result is not entirely unexpected, as many bipartitions carry largely redundant information given the nested nature of phylogenetic structure. Nevertheless, the inclusion of some bipartition splits suggests that including phylogenetic information can account for unobserved traits or traits that may be difficult to measure.

A variety of unmeasured traits or features may be represented by the phylogenetic clades identified as important for invasiveness potential. In our specific example, however, the particular clades identified seem to correspond to familiar and charismatic clades of lizards or species that are often included in the pet trade, with the important nodes that were identified highlighting agamids and chameleons, basilisk lizards, and plated lizards (see figure 4.4) [20]. All of these groups contain many species that are popular in the pet trade. However, what makes species desirable in the pet trade is highly variable. It could be because they have desirable colors or patterns, are easy to keep in captivity, are docile, extremely unique, have particular diets suitable for captive rearing, or any combination of these or other desirable traits [22, 50]. This makes it difficult to classify what makes a species "charismatic". However, incorporating phylogenetic information allows us to account for different combinations of these traits, since these desirable traits could have strong phylogenetic signal [51, 52].

In addition to phylogenetic features, many measured traits were useful for predicting species invasion probabilities. Some of these traits were body size, largest clutch size, diet, and substrate type (see figure 4.2a). Body size and life history are related to species' ability to invade and establish populations in new environments [53]. Our LIME results suggested that both small clutch sizes and large clutch sizes made species more likely to invade, demonstrating that sometimes extreme life history strategies increase dispersal potential although it is not clear how this impacts invasion success [54]. Additionally, many arboreal or pad-bearing species are transported to novel environments through their proximity to humans [44] or as hitchhikers on horticultural materials [55, 56], and many of these species also have small clutch sizes [57]. This could also be driving the invasion success of species with small clutch sizes in our dataset. There is also evidence within *Anolis sagrei*, a species with small clutch size, that invasive populations possess trait changes that lead to faster life histories than their native counterparts [58]. Substrate and diet appear to be related to generalism. Species with specialized diets are more likely to be classified as native and species that use multiple substrate types are more likely to be invasive. These results for feature importance were consistent with other similar studies also identifying body size, clutch size, and generalism as important species-level traits for invasion [6, 8].

Several climate features were also useful for predicting species invasion probabilities, including the minimum value of mean diurnal temperature range (BIO2) and the maximum value of precipitation of the wettest quarter (BIO16) for each species (see figure 4.2b). Our LIME results suggested that species that have low maximum precipitation in the wettest quarter were classified as native (see figure 4.3). This could be because species in areas with low precipitation may be physiologically adapted to desert conditions and therefore specialized with limited adaptability to different climatic conditions [59, 60]. It could also be driven by the lack of frequent transplantation or introductions due to humans. Our LIME results suggested species with low minimum diurnal temperature ranges are often invasive (see figure 4.3). Low minimum diurnal temperature range values are more common in tropical areas. This link between tropical locations and invasion could be because tropical species ranges are expanding due to climate change [1, 6], because their critical thermal temperatures are less canalized from selection [46], or because of transport to and from these areas.

Some examples of species which had relatively high invasion probabilities (i.e. would be classified as introduced under our model with a 0.5 threshold) were *Thecadactylus rapicauda*, *Varanus bengalensis*, and *Lampropholis guichenoti*. All of these species have wide geographic ranges. However, *Varanus bengalensis* is experiencing population declines in its range due to commercial exploitation, human interaction and habitat fragmentation [61], and while *Lampropholis guichenoti* seems to have the ideal set of traits to invade, some of this potential appears to be limited by behavior [62]. These results underscore the importance of chance events related to introduction and the role that species behavior plays in either hindering or driving these events.

Once we determined invasion probabilities for each species, we also wanted to map these invasion probabilities globally to determine areas that are acting as sinks and sources of invasive species. Our results suggested that many tropical and subtropical areas are acting as sources of invasive species. This could be because species living in these areas might be more adaptable, more desirable in the pet trade [63], because their ranges are expanding due to climate change [1] or primarily driven by transport to and from these areas where allowing species to hitchhike. It could also be because these are areas with high species richness since many species can thrive in the tropics. Meanwhile, despite having high species richness in many desert areas (see figure 4.8), these locations tend to have species with lower invasion probabilities (see figure 4.6). This could be because these species are physiologically adapted to their environments making them less adaptable to novel environments they might encounter or because there is less transport from these areas. Some areas which are frequently invaded are areas in northern South America and the southern United States as well as the northern part of Australia and parts of the Mediterranean (see figure 4.7a). It seems that while many of these areas also act as sources of invasive species, they are still vulnerable to being invaded by non-native species from other areas. This suggests a symmetry or reciprocity in the conditions that both promote invasiveness in native species as well as make them habitable for novel invaders from other areas.

4.6.2 Conclusions

Our results suggest that phylogenetic information can account for unobserved or difficult to measure traits in random forests. In our study, phylogenetic information seems to provide information about charismatic clades or those containing species that are often transported through the pet trade and significantly improved prediction accuracy. We identified several traits that were important for predicting invasion probabilities, such as body size, diet, substrate, and clutch size, and these results corroborate the findings from other studies. We also identified that tropical and subtropical countries are acting as sources of many invasive species and that they are similarly the most vulnerable to new invasions by other species. We further identified that species from harsher habitats, such as deserts and cold temperate regions, as less likely to become invasive—despite the high species richness of some of these areas. Overall, it is likely that vulnerable areas will continue to be impacted with ongoing climate change and human-mediated transportation of invasive species.

Bibliography

[1] Gian-Reto Walther, Alain Roques, Philip E. Hulme, Martin T. Sykes, Petr Pyšek, Ingolf Kühn, Martin Zobel, Sven Bacher, Zoltán Botta-Dukát, Harald Bugmann, Bálint Czúcz,

Jens Dauber, Thomas Hickler, Vojtěch Jarošík, Marc Kenis, Stefan Klotz, Dan Minchin, Mari Moora, Wolfgang Nentwig, Jürgen Ott, Vadim E. Panov, Björn Reineking, Christelle Robinet, Vitaliy Semenchenko, Wojciech Solarz, Wilfried Thuiller, Montserrat Vilà, Katrin Vohland, and Josef Settele. Alien species in a warmer world: risks and opportunities. *Trends in Ecology Evolution*, 24(12):686–693, December 2009.

[2] Stuart H. M. Butchart, Matt Walpole, Ben Collen, Arco van Strien, Jörn P. W. Scharlemann, Rosamunde E. A. Almond, Jonathan E. M. Baillie, Bastian Bomhard, Claire Brown, John Bruno, Kent E. Carpenter, Geneviève M. Carr, Janice Chanson, Anna M. Chenery, Jorge Csirke, Nick C. Davidson, Frank Dentener, Matt Foster, Alessandro Galli, James N. Galloway, Piero Genovesi, Richard D. Gregory, Marc Hockings, Valerie Kapos, Jean-Francois Lamarque, Fiona Leverington, Jonathan Loh, Melodie A. McGeoch, Louise McRae, Anahit Minasyan, Monica Hernández Morcillo, Thomasina E. E. Oldfield, Daniel Pauly, Suhel Quader, Carmen Revenga, John R. Sauer, Benjamin Skolnik, Dian Spear, Damon Stanwell-Smith, Simon N. Stuart, Andy Symes, Megan Tierney, Tristan D. Tyrrell, Jean-Christophe Vié, and Reg Watson. Global biodiversity: Indicators of recent declines. *Science*, 328(5982):1164–1168, May 2010.

[3] Céline Bellard, Phillip Cassey, and Tim M. Blackburn. Alien species as a driver of recent extinctions. *Biology Letters*, 12(2):20150623, February 2016.

[4] Robert A. Francis, , and Michael A. Chadwick. Urban invasions: non-native and invasive species in cities. *Geography*, 100(3):144–151, November 2015.

[5] Ashlyn L. Padayachee, Ulrike M. Irlich, Katelyn T. Faulkner, Mirijam Gaertner, Şerban Procheş, John R. U. Wilson, and Mathieu Rouget. How do invasive species travel to and through urban environments? *Biological Invasions*, 19(12):3557–3570, December 2017.

[6] Peter J. Mahoney, Karen H. Beard, Andrew M. Durso, Aimee G. Tallian, A. Lexine Long, Ryan J. Kindermann, Nicole E. Nolan, Daniel Kinka, and Harrison E. Mohn. Introduction effort, climate matching and species traits as predictors of global establishment success in non-native reptiles. *Diversity and Distributions*, 21(1):64–74, 2015.

[7] Alice Fournier, Caterina Penone, Maria Grazia Pennino, and Franck Courchamp. Predicting future invaders and future invasions. *Proceedings of the National Academy of Sciences*, 116(16):7905–7910, April 2019.

[8] Arman N. Pili, Boris Leroy, John G. Measey, Jules E. Farquhar, Adam Toomes, Phillip Cassey, Sebastian Chekunov, Matthias Grenié, Dylan van Winkel, Lisa Maria, Mae Lowe L. Diesmos, Arvin C. Diesmos, Damaris Zurell, Franck Courchamp, and David G. Chapple. Forecasting potential invaders to prevent future biological invasions worldwide. *Global Change Biology*, 30(7):e17399, 2024.

[9] Nicola J. Van Wilgen and David M. Richardson. The roles of climate, phylogenetic relatedness, introduction effort, and reproductive traits in the establishment of non-native reptiles and amphibians. *Conservation Biology*, 26(2):267–277, 2012.

[10] Lenka Moravcová, Petr Pyšek, Vojtěch Jarošík, and Jan Pergl. Getting the right traits: Reproductive and dispersal characteristics predict the invasiveness of herbaceous plant species. *PLOS ONE*, 10(4):e0123634, April 2015.

[11] Romana Aguiar Andrade and Daniel Cunha Passos. Low requirement on the nest site selection influencing the invasion success of house geckos. *Biotropica*, 57(2):e70018, 2025.

[12] Vinícius Soares Correa da Costa, Vivian de Mello Cionek, and Joaquim Olinto Branco. Feeding ecology of an invasive swimming crab: how feeding patterns favors invasion success. *Marine Biology*, 172(2):30, January 2025.

[13] Olivier Broennimann and Antoine Guisan. Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, 4(5):585–589, July 2008.

[14] Joseph H Connell. The influence of interspecific competition and other factors on the distribution of the barnacle chthamalus stellatus. *Ecology*, pages 710–723, 1961.

[15] Andrew C. Battles and Jason J. Kolbe. Miami heat: Urban heat islands influence the thermal suitability of habitats for ectotherms. *Global Change Biology*, 25(2):562–576, 2019.

[16] Shane C. Campbell-Staton, Kristin M. Winchell, Nicolas C. Rochette, Jason Fredette, Inbar Maayan, Rena M. Schweizer, and Julian Catchen. Parallel selection on thermal physiology facilitates repeated adaptation of city lizards to urban heat islands. *Nature Ecology Evolution*, 4(4):652–658, April 2020.

[17] William I. Lutterschmidt and Victor H. Hutchison. The critical thermal maximum: history and critique. *Canadian Journal of Zoology*, 75(10):1561–1574, October 1997.

[18] Susana Clusella-Trullas and Steven L. Chown. Lizard thermal trait variation at multiple scales: a review. *Journal of Comparative Physiology B*, 184(1):5–21, January 2014.

[19] Jeremy M Beaulieu and Brian C O'Meara. Hidden markov models for studying the evolution of binary morphological characters. *Modern phylogenetic comparative methods and their application in evolutionary biology: Concepts and practice*, pages 395–408, 2014.

[20] Julie L Lockwood, Dustin J Welbourne, Christina M Romagosa, Phillip Cassey, Nicholas E Mandrak, Angela Strecker, Brian Leung, Oliver C Stringham, Bradley Udell, Diane J Episcopio-Sturgeon, Michael F Tlusty, James Sinclair, Michael R Springborn,

Elizabeth F Pienaar, Andrew L Rhyne, and Reuben Keller. When pets become pests: the role of the exotic pet trade in producing invasive vertebrate animals. *Frontiers in Ecology and the Environment*, 17(6):323–330, 2019.

[21] Jérôme M. W. Gippet and Cleo Bertelsmeier. Invasiveness is linked to greater commercial success in the global pet trade. *Proceedings of the National Academy of Sciences*, 118(14):e2016337118, April 2021.

[22] Jose W. Valdez. Using google trends to determine current, past, and future trends in the reptile pet trade. *Animals*, 11(33):676, March 2021.

[23] Joseph W. Grigg and Lauren B. Buckley. Conservatism of lizard thermal tolerances and body temperatures across evolutionary history and geography. *Biology Letters*, 9(2):20121056, April 2013.

[24] Silvie Lišková and Daniel Frynta. What determines bird beauty in human eyes? *Anthrozoös*, 26(1):27–41, 2013.

[25] Shyama Pagad, Piero Genovesi, Lucilla Carnevali, Dmitry Schigel, and Melodie A. McGeoch. Introducing the global register of introduced and invasive species. *Scientific Data*, 5(1):170202, January 2018.

[26] GBIF.Org User. Occurrence download, 2024.

[27] Alexander Zizka, Daniele Silvestro, Tobias Andermann, Josué Azevedo, Camila Duarte Ritter, Daniel Edler, Harith Farooq, Andrei Herdean, María Ariza, Ruud Scharn, Sten Svantesson, Niklas Wengström, Vera Zizka, and Alexandre Antonelli. Coordinatecleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, 10(5):744–751, 2019.

[28] Stephen E. Fick and Robert J. Hijmans. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12):4302–4315, 2017.

[29] Shai Meiri. Traits of lizards of the world: Variation around a successful evolutionary design. *Global ecology and biogeography*, 27(10):1168–1172, 2018.

[30] João Filipe Riva Tonini, Karen H. Beard, Rodrigo Barbosa Ferreira, Walter Jetz, and R. Alexander Pyron. Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. *Biological Conservation*, 204:23–31, December 2016.

[31] Klaus Peter Schliep. phangorn: phylogenetic analysis in r. *Bioinformatics*, 27(4):592–593, February 2011.

[32] Emmanuel Paradis and Klaus Schliep. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in r. *Bioinformatics*, 35(3):526–528, February 2019.

[33] Guillaume Guénard, Pierre Legendre, and Pedro Peres-Neto. Phylogenetic eigenvector maps: a framework to model and predict species traits. *Methods in Ecology and Evolution*, 4:1120–1131, 2013.

[34] Marvin N. Wright and Andreas Ziegler. ranger: A fast implementation of random forests for high dimensional data in c++ and r. *Journal of Statistical Software*, 77:1–17, March 2017.

[35] J. D. Malley, J. Kruppa, A. Dasgupta, K. G. Malley, and A. Ziegler. Probability machines: consistent probability estimation using nonparametric learning machines. *Methods of Information in Medicine*, 51(1):74–81, 2012.

[36] André Altmann, Laura Toloşi, Oliver Sander, and Thomas Lengauer. Permutation importance: a corrected feature importance measure. *Bioinformatics*, 26(10):1340–1347, May 2010.

[37] Miron B. Kursa and Witold R. Rudnicki. Feature selection with the boruta package. *Journal of Statistical Software*, 36:1–13, September 2010.

[38] Joseph Felsenstein. Phylogenies and the comparative method. *The American Naturalist*, 125(1):1–15, 1985.

[39] Wayne P Maddison and Richard G FitzJohn. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic biology*, 64(1):127–136, 2015.

[40] Josef C Uyeda, Rosana Zenil-Ferguson, and Matthew W Pennell. Rethinking phylogenetic comparative methods. *Systematic Biology*, 67(6):1091–1109, 2018.

[41] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. “why should i trust you?”: Explaining the predictions of any classifier. (arXiv:1602.04938), August 2016. arXiv:1602.04938 [cs].

[42] Uri Roll, Anat Feldman, Maria Novosolov, Allen Allison, Aaron M. Bauer, Rodolphe Bernard, Monika Böhm, Fernando Castro-Herrera, Laurent Chirio, Ben Collen, Guarino R. Colli, Lital Dabool, Indraneil Das, Tiffany M. Doan, Lee L. Grismer, Marinus Hoogmoed, Yuval Itescu, Fred Kraus, Matthew LeBreton, Amir Lewin, Marcio Martins, Erez Maza, Danny Meirte, Zoltán T. Nagy, Cristiano de C. Nogueira, Olivier S. G. Pauwels, Daniel Pincheira-Donoso, Gary D. Powney, Roberto Sindaco, Oliver J. S. Tallowin, Omar Torres-Carvajal, Jean-François Trape, Enav Vidan, Peter Uetz, Philipp Wagner, Yuezhao Wang, C. David L. Orme, Richard Grenyer, and Shai Meiri. The global distribution of tetrapods reveals a need for targeted reptile conservation. *Nature Ecology Evolution*, 1(11):1677–1682, November 2017.

[43] Ruth A Hufbauer, Benoît Facon, Virginie Ravigné, Julie Turgeon, Julien Foucaud, Carol E Lee, Olivier Rey, and Arnaud Estoup. Anthropogenically induced adaptation

to invade (aiai): contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications*, 5(1):89–101, January 2012.

[44] Jesse B Borden and S Luke Flory. Urban evolution of invasive species. *Frontiers in Ecology and the Environment*, 19(3):184–191, 2021.

[45] Flavia M. Lanna, Guarino R. Colli, Frank T. Burbrink, and Bryan C. Carstens. Identifying traits that enable lizard adaptation to different habitats. *Journal of Biogeography*, 49(1):104–116, 2022.

[46] Juan G. Rubalcaba, Sidney F. Gouveia, Fabricio Villalobos, Miguel Á. Olalla-Tárraga, and Jennifer Sunday. Climate drives global functional trait variation in lizards. *Nature Ecology Evolution*, 7(4):524–534, April 2023.

[47] Stephen Jay Gould. *Wonderful Life: The Burgess Shale and the Nature of History*. W. W. Norton Company, September 1990. Google-Books-ID: q8BUAgAAQBAJ.

[48] Alan de Queiroz. Contingent predictability in evolution: Key traits and diversification. *Systematic Biology*, 51(6):917–929, December 2002.

[49] Charles G. Willis, Brad Ruhfel, Richard B. Primack, Abraham J. Miller-Rushing, and Charles C. Davis. Phylogenetic patterns of species loss in thoreau's woods are driven by climate change. *Proceedings of the National Academy of Sciences*, 105(44):17029–17033, November 2008.

[50] Sandra Altherr and Katharina Lameter. The rush for the rare: Reptiles and amphibians in the european pet trade. *Animals*, 10(1111):2085, November 2020.

[51] William L. Allen, Nickolas Moreno, Tony Gamble, and Ylenia Chiari. Ecological, behavioral, and phylogenetic influences on the evolution of dorsal color pattern in geckos*. *Evolution*, 74(6):1033–1047, June 2020.

[52] Jonathan Goldenberg, Karen Bisschop, Joshua W. Lambert, Michaël P. J. Nicolaï, Rampal S. Etienne, Liliana D'Alba, and Matthew D. Shawkey. Habitat openness and squamate color evolution over deep time. *Nature Communications*, 16(1):2625, March 2025.

[53] Sarah-Sophie Weil, Laure Gallien, Michaël P. J. Nicolaï, Sébastien Lavergne, Luca Börger, and William L. Allen. Body size and life history shape the historical biogeography of tetrapods. *Nature Ecology Evolution*, 7(9):1467–1479, September 2023.

[54] Sarah-Sophie Weil, Laure Gallien, Sébastien Lavergne, Luca Börger, Gabriel W. Hassler, Michaël P. J. Nicolaï, and William L. Allen. Chameleon biogeographic dispersal is associated with extreme life history strategies. *Ecography*, 2022(10):e06323, 2022.

[55] Diego F. Cisneros-Heredia. *The Hitchhiker Wave: Non-native Small Terrestrial Vertebrates in the Galapagos*, page 95–139. Springer International Publishing, Cham, 2018.

[56] Gad Perry, Charles R Knapp, Tandora D Grant, Stesha A Pasachnik, and Ioana Coman. From pets to threats: Invasive iguanas and other species cause significant harm to native iguanas. *Reptiles & Amphibians*, 28(2):213–217, 2021.

[57] Lukáš Kratochvíl and Lukáš Kubička. Why reduce clutch size to one or two eggs? reproductive allometries reveal different evolutionary causes of invariant clutch size in lizards. *Functional Ecology*, 21(1):171–177, 2007.

[58] Tamara L Fetters and Joel W McGlothlin. Life histories and invasions: accelerated laying rate and incubation time in an invasive lizard, anolis sagrei. *Biological Journal of the Linnean Society*, 122(3):635–642, October 2017.

[59] Eric R. Pianka. *Ecology and Natural History of Desert Lizards: Analyses of the Ecological Niche and Community Structure*. Princeton University Press, March 2017. Google-Books-ID: h4YrDgAAQBAJ.

[60] Raúl Araya-Donoso, Esteban San Juan, Ítalo Tamburrino, Madeleine Lamborot, Claudio Veloso, and David Véliz. Integrating genetics, physiology and morphology to study desert adaptation in a lizard species. *Journal of Animal Ecology*, 91(6):1148–1162, 2022.

[61] K. M. Rahman and Rakhimov Iliazovic. Consequences of habitat loss and habitat fragmentation on the survival of monitor lizard populations in bangladesh: a review and prospectus. *Journal of Biodiversity and Environmental Sciences*, 8:20166663–2222, 02 2016.

[62] David G Chapple, Sarah M Simmonds, and Bob BM Wong. Know when to run, know when to hide: can behavioral differences explain the divergent invasion success of two sympatric lizards? *Ecology and Evolution*, 1(3):278–289, November 2011.

[63] Emma R. Bush, Sandra E. Baker, and David W. Macdonald. Global trade in exotic pets 2006–2012. *Conservation Biology*, 28(3):663–676, 2014.

4.7 Supplementary Material

4.7.1 Phylogenetic methods

We tested multiple methods for incorporating phylogenetic information into our random forest model. All of the methods are described in the text. We identified close performance using all bipartitions and phylogenetic eigenvector maps (PEM) for predicting invasiveness, although bipartitions were preferred as they provided the best performance.

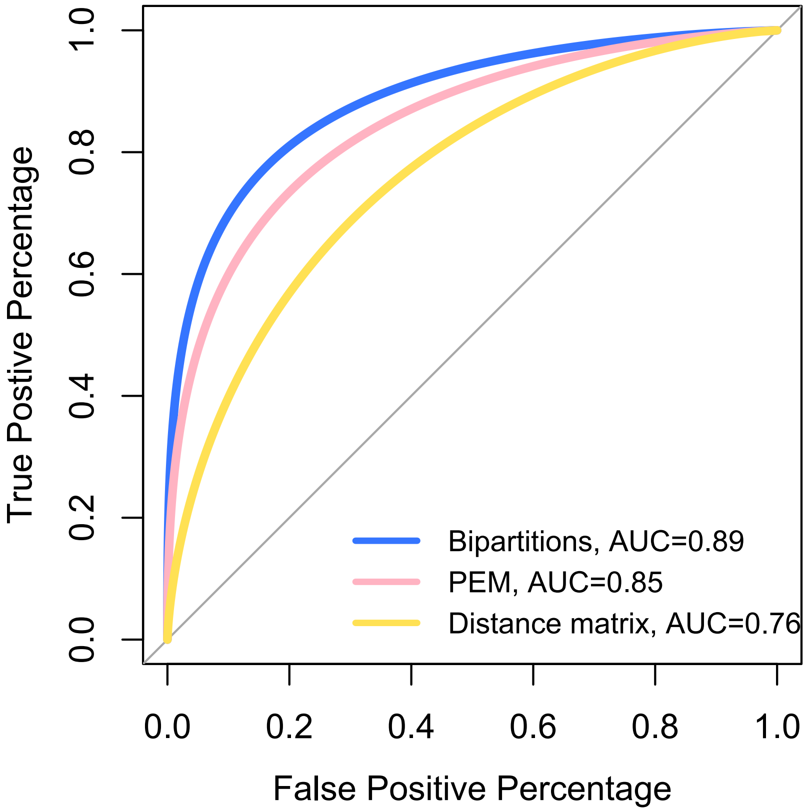


Figure 4.9: ROC plots for three different models incorporating phylogeny in different ways, all bipartitions (blue), phylogenetic eigenvector maps (PEM) (pink), and the phylogeny as a distance matrix (yellow). Area under the curve is included in the legend, values closer to 1 indicate better performance.

4.7.2 Feature selection

Of our trait data, largest clutch size, maximum body size, and diet were selected as important from our analysis with Boruta. From our climate data minimum BIO2, minimum BIO3, minimum BIO4, minimum BIO7, minimum BIO12, minimum BIO13, minimum BIO16, minimum BIO18, maximum BIO1, maximum BIO3, maximum BIO4, maximum BIO6, maximum BIO7, maximum BIO8, maximum BIO10, maximum BIO11, maximum BIO12, maximum BIO13, maximum BIO14, maximum BIO16, maximum BIO17, maximum BIO18, maximum BIO19, mean BIO2, mean BIO3, mean BIO4, mean BIO6, mean BIO7, mean BIO12, mean BIO13, mean BIO14, mean BIO16, mean BIO17, and mean BIO19 were selected as important. From our phylogenetic bipartitions V110, V156, V163, V203, V204, and V411 were selected as important.

Chapter 5

phyloBoruta: An R package to test if traits are more informative than phylogeny

5.1 Attributions

5.1.1 Co-Authors

Josef C. Uyeda, Professor, Department of Biological Sciences, Virginia Tech, Blacksburg, VA

5.1.2 Acknowledgments

5.2 Abstract

One of the primary uses of phylogeny is to leverage its broad power to predict a variety of traits among related species—from gene sequences to macroecological features of species. However, methods that rely on a large number of predictors are susceptible to multiple comparisons, overfitting, and—particularly in phylogenetic comparative studies—identifying features that improve prediction solely based on their shared phylogenetic signal. Identifying traits that carry additional information beyond phylogeny is a key goal for understanding why and how this data could be influencing predictions, and how these features may be related to causal biological processes. In this work, we extend the Boruta feature selection algorithm for random forests to test if biological traits are more informative than randomly simulated phylogenetic traits. We used this algorithm on trait data from lizards classifying them as native or invasive. We identified that maximum body size and largest clutch size will be important traits for predicting species' invasions—well beyond the information they carry about phylogeny. We present phyloBoruta as a general approach for identifying important traits in feature selection that carry predictive power beyond randomly evolved phylogenetic traits.

5.3 Introduction

Machine learning has rapidly grown in popularity for analyzing biological datasets [1]. Random forests are particularly popular as they allow users to identify and select important features for classification or regression problems[2, 3, 4] and due to their interpretability provided through feature importance metrics. Identifying relevant features is particularly important for biological or ecological datasets as biologists might want to determine the most relevant data for future collection and understanding why and how this data could be influencing predictions is particularly interesting to those studying biological processes.

Many algorithms have been developed that identify the minimal number of features to use based on classification accuracy [5, 6]. Rather than simply trying to maximize classification accuracy, the Boruta algorithm for feature selection prioritizes selecting all features that are most relevant for making a certain classification [7]. In particular, this approach is useful for retaining features that allow for greater model interpretability.

In biology, comparative datasets pose a particular challenge and opportunity for prediction tasks. This is because for datasets at the species level, all traits will be impacted by evolutionary history and have varying degrees of phylogenetic signal. Correlations between traits can emerge, and indeed be statistically significant, purely due to shared phylogenetic structure [8, 9, 10]. When using species traits (whether morphological, physiological, or even behavioral) to make predictions it is important to be aware that many traits likely carry with them some variation which can be explained by phylogenetic information, either through biogeography or genetic or developmental constraint [11, 12, 13, 14]. As the number of traits with phylogenetic signal increases, it becomes increasingly likely that some combination of these features could uniquely map to particular regions of the phylogenetic tree and explain clustered phylogenetic responses purely by chance [10]. This means that for biologists it is particularly valuable to identify traits that carry additional information beyond phylogeny, in order to identify how these traits are related to the trait of interest.

In this work, we sought to extend the Boruta algorithm to test if biological traits are more informative than randomly simulated phylogenetic traits to provide an understanding of traits that give additional insights into classification or regression. We validate the approach with an application to lizard invasivness prediction from phenotypic traits.

5.4 phyloBoruta

The traditional Boruta algorithm works as a wrapper of random forests by permuting values of features and comparing randomized feature importance to the importance of the non-randomized features [7]. This provides information about the relative importance of these features in comparison to random, presumably uninformative features. Our algorithm works similarly to the Boruta algorithm; however, rather than comparing true features to permuted features, we compare our features to parametrically simulated phylogenetic traits. If true features are more informative than our simulated features, this suggests they hold additional important information for making predictions which cannot be gleaned from the phylogeny alone.

To do this, we separate discrete and continuous features, as these features must be simulated using different methods. For simulating discrete characters, we use the method from the R package *ape* rTraitDisc, simulating two character states and using equal transition rates between states. We repeat this as many times as the user specifies, and we use these as our “shadow” attributes to compare the importance of our true discrete features against. For simulating continuous traits, we use the method fastBM, from the R package *phytools* [15]. We do this as many times as the user specifies and we compare the feature importance of these “shadow” features to the importance of our continuous traits.

From this point the algorithm behaves the same as the traditional Boruta algorithm, by comparing importance values to that of the maximum importance of shadow variables, and using two-sided tests of equality with this maximum importance value to assign whether true features are more informative than phylogeny or if phylogeny alone captures the information contained in the trait. We repeat this procedure until no tentative features remain or until the algorithm reaches the defined limit of random forest runs.

5.5 Implementation

Because the traditional Boruta algorithm relies purely on permutation there are no issues with changing the distributions of the data, as permutation preserves the distributional properties. However, permutation without respect to phylogeny will necessarily result in traits that have decreased phylogenetic signal. Thus, while our approach may be limited by differences in the distribution of the simulated traits, it benefits from preserving shared phylogenetic signal with the original traits but randomized with respect to the trait of interest. Consequently, our method adds features that contain phylogenetic information similar to the true traits. One consequence of this is that it splits feature importance across a greater number of features with shared phylogenetic information when increasing the number of simulations. To attempt to minimize this, we recommend only simulating the number of traits available in a dataset in order to avoid false confidence in trait selections. However, limiting the number of simulations to the number of traits in empirical data can lead to high variance across runs in the traits that may be selected since we are adding many presumably correlated features (a demonstration of this variance and uncertainty across runs can be seen in figures 5.1, 5.3). This is a limitation of our model in comparison with the traditional Boruta algorithm and to overcome this we recommend running multiple independent runs and identifying the traits which are most consistently identified.

5.6 Simulations

We tested phyloBoruta first on simulated data. To determine how strongly data must be co-varying with the response variable in order to be selected under our algorithm, we simulated random phylogenetic traits using fastBM from the R package *phytools* for 9 traits. We then simulated our response variable and one trait with varying levels of covariance between the two traits under multivariate Brownian motion using the function sim.corrs. We identified that across moderate to strong levels of covariance (0.4-1) the co-varying predictor was typically selected as important in at least 75% of runs, however when the covariance between traits dropped to less than 0.4, the predictor was no longer selected above this threshold. This indicates that levels of moderate to strong covariance should be identifiable under phyloBoruta, but weakly co-varying traits may be rejected under our model. We found relatively similar results when simulating biologically realistic data using sim.corrs, allowing all traits to co-vary with both the response variable and each other.

5.7 Empirical Example

We used the phyloBoruta algorithm on trait data from lizards classifying them as native or invasive. We used published trait data from Meiri 2018 [16], published invasion status from the Global Register of Introduced and Invasive Species (GRIIS) [17], and a published phylogenetic tree [18]. If a species was not introduced in any country in the GRIIS, it was classified as native, whereas if it was introduced in at least one country and registered in the GRIIS it was classified as introduced.

Using this dataset we ran the algorithm 100 times, specifying the number of simulated traits to be 11 (the same number of true traits). For continuous traits maximum body size was continuously identified as being more informative than phylogeny in 100% of runs, meanwhile maximum and minimum body temperature and latitude and longitude were rejected in at least 60% of runs. This was well-aligned with previous results from Chapter 4 which suggested maximum body size was most informative for predicting invasion probabilities. For discrete traits, largest clutch size, substrate, diel activity, reproductive mode, and diet were more informative than phylogeny in at least 85% of runs. Largest clutch size was also confirmed to be important in 100% of all runs. This result was also corroborated in Chapter 4 which showed this trait was important for predicting invasion probabilities. In comparison to the traditional Boruta algorithm which identified all traits as being important with the exception of foraging mode, our method excluded foraging mode in addition to maximum and minimum mean body temperatures as well as latitude and longitude. This result shows that our method is more conservative with feature selection, as expected. In particular, our approach highlights that while many traits may improve prediction accuracy in phylogenetically-structred random forests, these traits may not carry more predictive power than expected from sets of randomly evolving phylogenetic traits. This highlights

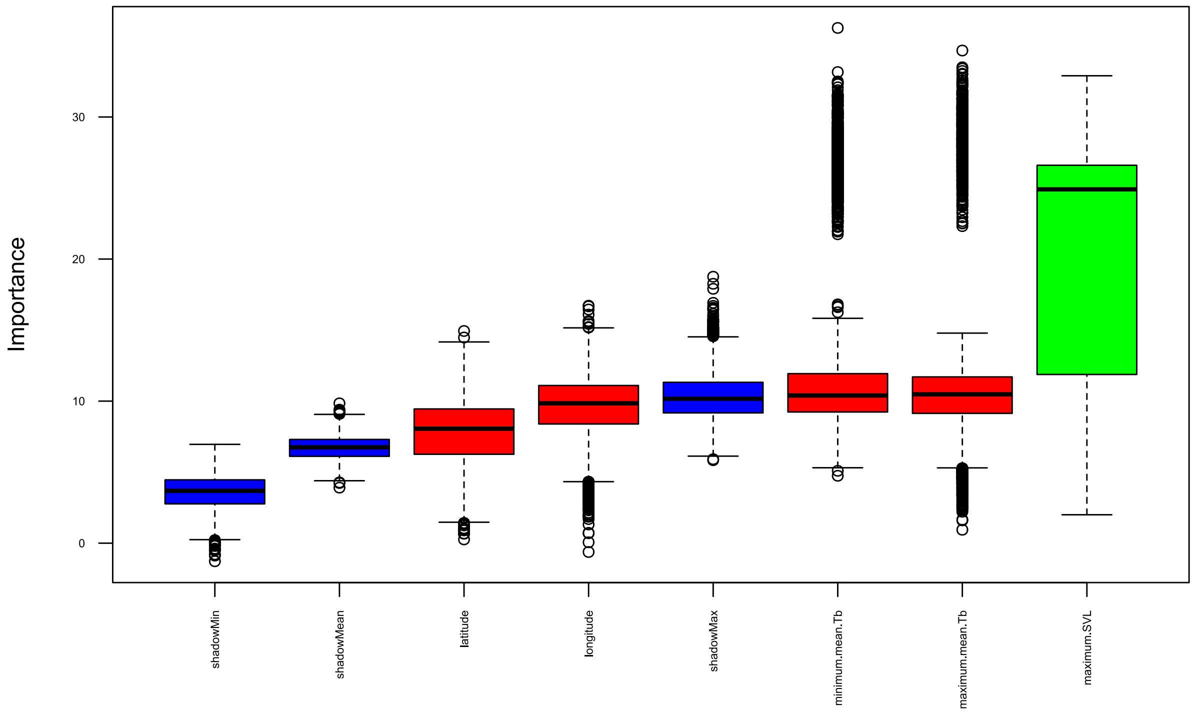


Figure 5.1: Boxplot showing the distributions of feature importances across 100 different runs of phyloBoruta for continuous features (green and red) compared against random phylogenetic features (blue). True features are colored green if they were selected as important across at least 75% of runs, otherwise they were rejected.

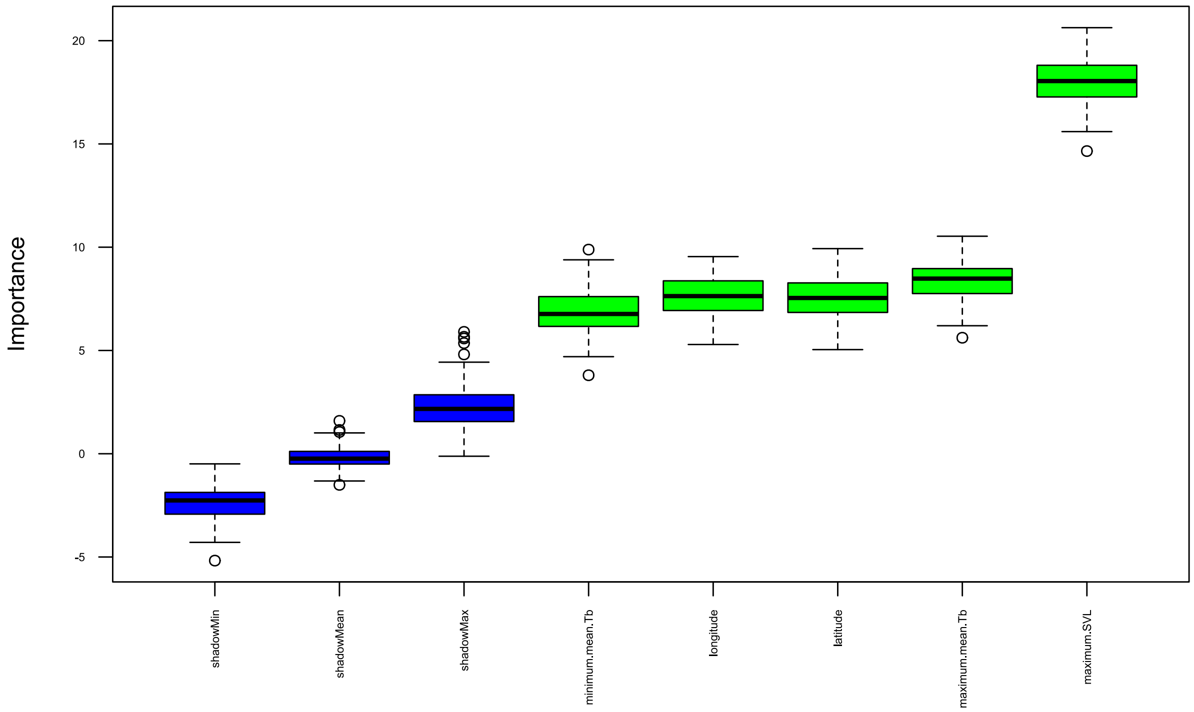


Figure 5.2: Plot of traditional Boruta algorithm feature importances for continuous features compared against permuted features in blue. True features are colored green if they were selected as important, red if they were rejected, and yellow if they were tentative features.

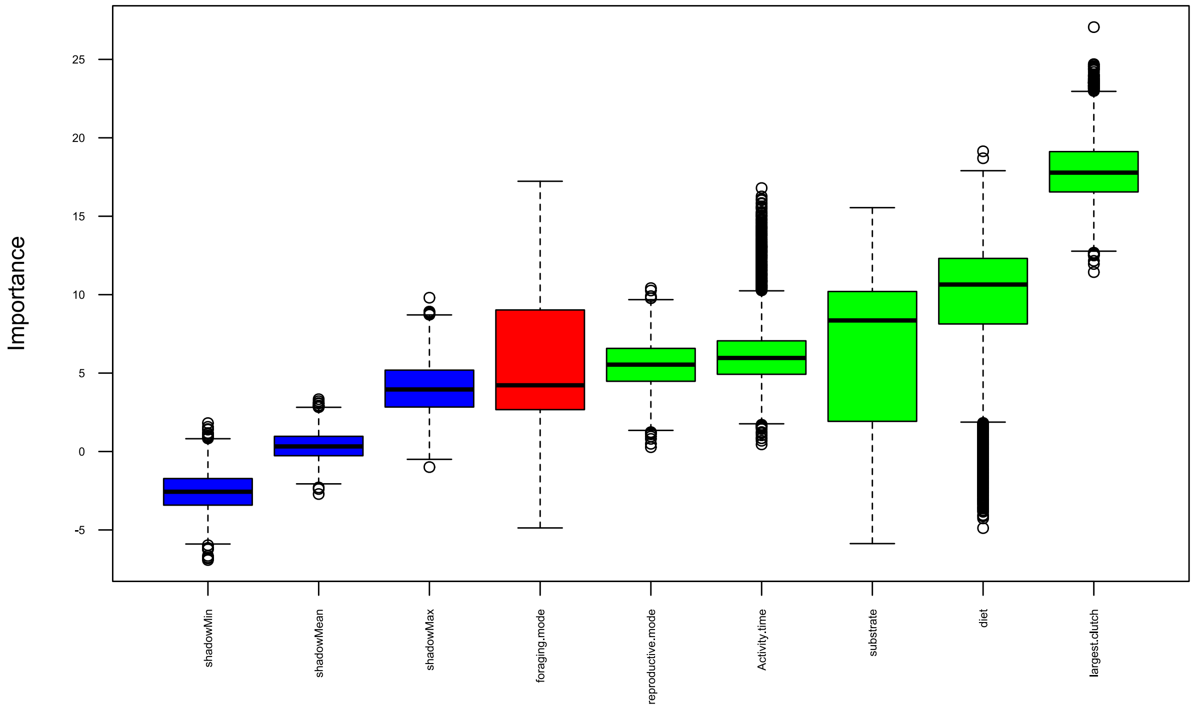


Figure 5.3: Boxplot showing the distributions of feature importances across 100 different runs of phyloBoruta for discrete features (green and red) compared against randomly simulated phylogenetic features (blue). True features are colored green if they were selected as important across at least 75% of runs and red if they were rejected.

that researchers should take caution when interpreting the meaningfulness of phylogenetic features identified in random forests algorithms.

5.8 Conclusions

Our phyloBoruta approach differed from feature selection in Boruta in our empirical example by selecting a smaller set of features. This highlights that when used in combination, phylogenetically-structured features may frequently carry predictive information for phylogenetically-structured responses—but that these features may occur by chance. While this method can certainly be used to narrow down the number of traits used in random forest analyses, biological trait datasets often are not as data rich as many other datasets. One way of viewing the effect of phylogeny structuring and phylogenetic signal is that it is a reduction in the effective sample size of the data, which is no longer independent [8]. Thus,

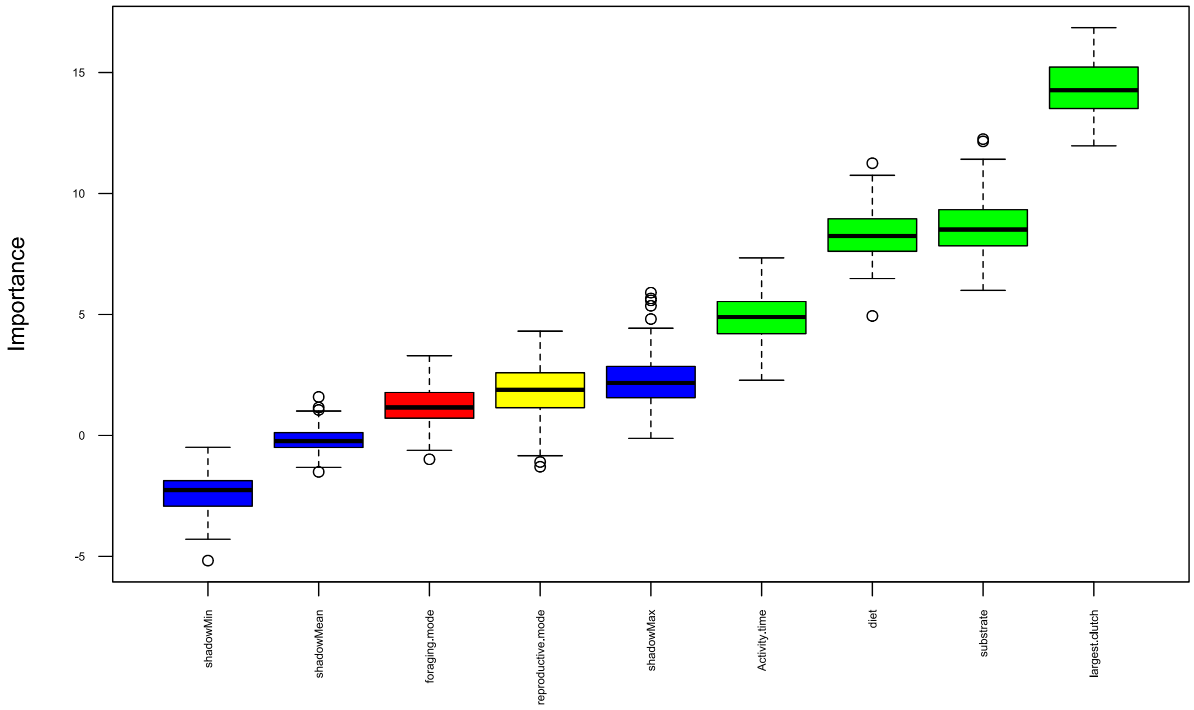


Figure 5.4: Plot of feature importance from the traditional Boruta algorithm for discrete features compared against permuted features in blue. True features are colored green if they were selected as important, red if they were rejected, and yellow if they are tentative features.

in comparative applications, these datasets may warrant the inclusion of a large number of possible features, but these features may contain many auto-correlated and shared features due to the nature of evolution and the phylogenetic structure of traits [9]. To this end, this method is also useful for identifying scenarios in which the phylogeny is able to capture and predict the structure of the trait itself, and that many alternative traits could—individually or in combination—achieve similar predictive accuracy as the focal traits. In such scenarios, more detailed analyses and data integrations may be required to identify the likely traits driving classification of species at the tips. In our example, we were able to identify traits that influence invasion which cannot fully be explained by phylogeny alone. We identified that maximum body size and largest clutch size will be important traits for predicting species' invasions, regardless of phylogenetic structure. Additionally, if the user plans to incorporate phylogenetic information in their analyses explicitly, this method can help the user to determine which traits are no longer necessary after accounting for phylogeny or they can help identify traits that must be included in addition to phylogenetic information. Identifying traits like this is certainly of interest to biologists as it aids in interpretability of the random forests predictions, allowing us to find new ways to approach studies of trait evolution and uncover new mysteries contained in these informative features.

Bibliography

[1] Joe G. Greener, Shaun M. Kandathil, Lewis Moffat, and David T. Jones. A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, 23(1):40–55, January 2022.

[2] Christian Kampichler, Ralf Wieland, Sophie Calmé, Holger Weissenberger, and Stefan Arriaga-Weiss. Classification in conservation biology: A comparison of five machine-learning methods. *Ecological Informatics*, 5(6):441–450, November 2010.

[3] Tara A. Pelletier, Bryan C. Carstens, David C. Tank, Jack Sullivan, and Anahí Espíndola. Predicting plant conservation priorities on a global scale. *Proceedings of the National Academy of Sciences*, 115(51):13027–13032, December 2018.

[4] Flavia M. Lanna, Guarino R. Colli, Frank T. Burbrink, and Bryan C. Carstens. Identifying traits that enable lizard adaptation to different habitats. *Journal of Biogeography*, 49(1):104–116, 2022.

[5] Xue-wen Chen and Jong Cheol Jeong. Minimum reference set based feature selection for small sample classifications. In *Proceedings of the 24th international conference on Machine learning*, ICML '07, page 153–160, New York, NY, USA, June 2007. Association for Computing Machinery.

[6] Roland Nilsson, José M. Peña, Johan Björkegren, and Jesper Tegnér. Consistent feature selection for pattern recognition in polynomial time. *Journal of Machine Learning Research*, 8(21):589–612, 2007.

[7] Miron B. Kursa and Witold R. Rudnicki. Feature selection with the boruta package. *Journal of Statistical Software*, 36:1–13, September 2010.

[8] Joseph Felsenstein. Phylogenies and the comparative method. *The American Naturalist*, 125(1):1–15, 1985.

[9] Wayne P Maddison and Richard G FitzJohn. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic biology*, 64(1):127–136, 2015.

[10] Josef C Uyeda, Rosana Zenil-Ferguson, and Matthew W Pennell. Rethinking phylogenetic comparative methods. *Systematic Biology*, 67(6):1091–1109, 2018.

[11] Thomas J. Sanger, Liam J. Revell, Jeremy J. Gibson-Brown, and Jonathan B. Losos. Repeated modification of early limb morphogenesis programmes underlies the convergence of relative limb length in anolis lizards. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729):739–748, August 2011.

[12] Joel W. McGlothlin, Megan E. Kobiela, Helen V. Wright, D. Luke Mahler, Jason J. Kolbe, Jonathan B. Losos, and Edmund D. III Brodie. Adaptive radiation along a deeply conserved genetic line of least resistance in anolis lizards. *Evolution Letters*, 2(4):310–322, August 2018.

[13] Arkadiusz Fr hlich, Dorota Kotowska, Rafał Martyka, and Matthew R. E. Symonds. Allometry reveals trade-offs between bergmann's and allen's rules, and different avian adaptive strategies for thermoregulation. *Nature Communications*, 14(1):1101, February 2023.

[14] Sarah-Sophie Weil, Laure Gallien, Michaël P. J. Nicolaï, Sébastien Lavergne, Luca Börger, and William L. Allen. Body size and life history shape the historical biogeography of tetrapods. *Nature Ecology Evolution*, 7(9):1467–1479, September 2023.

[15] Liam J. Revell. Phytools 2.0: An updated r ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, 12:e16505, January 2024.

[16] Shai Meiri. Traits of lizards of the world: Variation around a successful evolutionary design. *Global ecology and biogeography*, 27(10):1168–1172, 2018.

[17] Shyama Pagad, Piero Genovesi, Lucilla Carnevali, Dmitry Schigel, and Melodie A. McGeoch. Introducing the global register of introduced and invasive species. *Scientific Data*, 5(1):170202, January 2018.

[18] João Filipe Riva Tonini, Karen H. Beard, Rodrigo Barbosa Ferreira, Walter Jetz, and R. Alexander Pyron. Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. *Biological Conservation*, 204:23–31, December 2016.

Chapter 6

Synthesis

In this dissertation I sought to understand the relative contributions of constraint and contingency when studying macroevolution and to determine how to best account for evolutionary history when predicting species responses to novel environments. My first two chapters investigate the role of contingency and constraint across different scales of macroevolution by studying morphological change both broadly and in "model clades" containing species that are urban tolerant and invasive. My third and fourth chapters investigate different methods for accounting for phylogenetic information in random forest analyses and the utility of using phylogeny as a proxy for unobserved or difficult to measure traits. These chapters provide insights into important traits that will impact species responses to environmental change, particularly species-level invasion capacities.

In Chapter 2 I investigate constraint in evolutionary hindlimb allometry across lizards, identifying that across lizards the majority of variation is consistent with Brownian motion and displays a conserved evolutionary relationship between hindlimb length and body size across species. Additionally, this chapter suggests that within Scincidae there is a change in evolutionary dynamics that precedes the decoupling of limb length and body size which could prime species to explore morphospace and lead to the evolution of fossoriality in lizards. This speaks to the role of contingency in evolution, as it seems that belly-dragging in skinks acts as an intermediate between the traditional upright sprawling gait that most lizards exhibit and the snake-like movement of limbless, fossorial lizards [1, 2, 3, 4].

In Chapter 3 we look more closely at morphological evolution in "model clades" which contain urban tolerant and invasive species [5, 6, 7]. We studied two traits, hindlimb length and toepad area, that appear to be important for urban tolerance and invasion in two clades *Hemidactylus* and *Anolis* [8, 9, 10, 11, 12, 13]. We identified that the two clades have opposite patterns in their evolutionary dynamics of limb length and toepad area. This chapter highlights the importance of evolutionary history, even in species that are using similar microhabitats and are similarly successful. Despite possessing similar traits, these clades have completely different evolutionary dynamics. Future studies should investigate microevolution in these traits within species to determine if it is aligned with macroevolutionary patterns of trait change for their respective clades. This chapter underscores the importance of incorporating evolutionary history into analyses using species traits as predictors.

In Chapter 4 we highlight a path forward for and the utility of incorporating phylogenetic information in predictive modeling. Specifically we investigate different methods of incorporating phylogenetic information, species traits, climate and geographic data into random forest analyses to predict species invasion capacities. We found that incorporating all bipartitions in the phylogeny as data in our random forest provided the best performance. The highlighted clades using this method suggested that phylogenetic information was allowing us to account for complex traits like species “charisma” or things that are important for the pet trade which could encompass many traits such as rarity, docility, or ease of breeding [14, 15]. We also identified that maximum body size, largest cluth size, substrate type, and diet were good predictors of species invasion capacities. Maximum body size and largest clutch are related to species ability to spread and establish populations [16] while diet could be related to generalism. We also identified that mean monthly temperature range was a good predictor of species invasion capacities, which could be driven by the strength of selection on critical thermal limits [17, 18]. We also found that minimum precipitation was a good predictor of species’ invasion capacity which could because native species in these areas with little precipitation are physiologically adapted to desert conditions [19, 20]. This chapter demonstrated that phylogenetic information can be used to account for traits that may be difficult to measure or observe.

In Chapter 5 we took this a step further and developed an R package which allows users to test if the traits they are using in random forests are more informative than a randomly simulated phylogenetic trait. Comparison between this feature selection approach and a traditional feature selection approach [21] highlights which features are important simply because they carry phylogenetic signal (which is aligned with phylogenetic signal of the response) versus which features provide additional insight into traits that may be driving the response. Our package presents a strong tool to investigate species traits and their relationship to species responses to global change and could be extended to many other systems.

6.1 Conclusions

Throughout this dissertation I have learned to appreciate the complexity of scale and had the important realization that at no point will we ever have access to all of the information we need to make predictions about species responses to global change. Despite the evergrowing volume of scientific research, there will always be more to learn and more data to collect. The beauty of macroevolutionary data lies in its standardization across many species and its multifactorial nature. While it may not be perfect, it is often especially designed to be applied to analyses across many different species, making it easier to incorporate into analyses predicting species-level responses to global change. It also can carry signatures of many biological processes which while indiscernible at the macroevolutionary level are nonetheless extremely important for making accurate predictions about responses to global change.

I hope this dissertation has revealed to readers the intricate interplay of evolutionary constraint and contingency in shaping macroevolutionary patterns and underscored the critical role of accounting for evolutionary history when predicting species responses to novel environments. Through investigations spanning broad morphological trends in lizards to detailed analyses within specific "model clades," we have demonstrated that while certain evolutionary trajectories exhibit predictable constraint, contingent events can drive divergent evolutionary paths even among ecologically similar groups. Furthermore, our exploration of phylogenetic information within predictive modeling frameworks, particularly through the application of random forests, reveals its utility in capturing complex, unmeasured traits that significantly influence species invasion capacities. Finally, the development of a novel feature selection package provides a valuable tool for disentangling the predictive power of specific traits from the underlying phylogenetic signal. Taken together, these findings emphasize the necessity of integrating evolutionary perspectives into our understanding of species responses to global change, offering crucial insights for future conservation efforts and ecological forecasting. Future studies should continue to explore the utility of phylogenetic information in making predictions and attempt to validate these predictions over space and time.

Bibliography

[1] Massimo Bernardi and Marco Avanzini. Locomotor behavior in early reptiles: Insights from an unusual erpetopus trackway. *Journal of Paleontology*, 85(5):925–929, September 2011.

[2] Stefan Curth, Martin S. Fischer, and John A. Nyakatura. Ichnology of an extant belly-dragging lizard—analogies to early reptile locomotion? *Ichnos*, 21(1):32–43, January 2014.

[3] Gen Morinaga and Philip J. Bergmann. Evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards. *Proceedings of the Royal Society B: Biological Sciences*, 287(1923):20200192, March 2020.

[4] Edwin Dickinson, Christopher S. Hanna, Hannah M. Fischer, Elizabeth C. Davoli, Allen A. Currier III, and Michael C. Granatosky. Locomotor energetics in the indonesian blue-tongued skink (tiliqua gigas) with implications for the cost of belly-dragging in early tetrapods. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, 337(4):329–336, 2022.

[5] Robbie Weterings and Kai C Vetter. Invasive house geckos (hemidactylus spp.): their current, potential and future distribution. *Current Zoology*, 64(5):559–573, October 2018.

[6] Jesse B Borden and S Luke Flory. Urban evolution of invasive species. *Frontiers in Ecology and the Environment*, 19(3):184–191, 2021.

[7] Martha M. Muñoz, Luke O. Frishkoff, Jenna Pruett, and D. Luke Mahler. Evolution of a model system: New insights from the study of anolis lizards. *Annual Review of Ecology, Evolution, and Systematics*, 54(Volume 54, 2023):475–503, November 2023.

[8] Theodore Garland and Jonathan Losos. Ecological morphology of locomotor performance in squamate reptiles. In Peter C. Wainwright and Stephen M. Reilly, editors, *Ecological morphology: integrative organismal biology*, pages 240–302. University of Chicago Press, Chicago, IL, USA, 01 1994.

[9] Zani. The comparative evolution of lizard claw and toe morphology and clinging performance. *Journal of Evolutionary Biology*, 13(2):316–325, March 2000.

[10] Duncan J. Irschick, Anthony Herrel, and Bieke Vanhooydonck. Whole-organism studies of adhesion in pad-bearing lizards: creative evolutionary solutions to functional problems. *Journal of Comparative Physiology A*, 192(11):1169–1177, November 2006.

[11] B. A. Goodman. Nowhere to run: the role of habitat openness and refuge use in defining patterns of morphological and performance evolution in tropical lizards. *Journal of Evolutionary Biology*, 22(7):1535–1544, July 2009.

[12] Kristin M. Winchell, R. Graham Reynolds, Sofia R. Prado-Irwin, Alberto R. Puente-Rolón, and Liam J. Revell. Phenotypic shifts in urban areas in the tropical lizard anolis cristatellus. *Evolution*, 70(5):1009–1022, May 2016.

[13] Kristin M. Winchell, Inbar Maayan, Jason R. Fredette, and Liam J. Revell. Linking locomotor performance to morphological shifts in urban lizards. *Proceedings of the Royal Society B: Biological Sciences*, 285(1880):20180229, June 2018.

[14] Julie L Lockwood, Dustin J Welbourne, Christina M Romagosa, Phillip Cassey, Nicholas E Mandrak, Angela Strecker, Brian Leung, Oliver C Stringham, Bradley Udell, Diane J Episcopio-Sturgeon, Michael F Tlusty, James Sinclair, Michael R Springborn, Elizabeth F Pienaar, Andrew L Rhyne, and Reuben Keller. When pets become pests: the role of the exotic pet trade in producing invasive vertebrate animals. *Frontiers in Ecology and the Environment*, 17(6):323–330, 2019.

[15] Jose W. Valdez. Using google trends to determine current, past, and future trends in the reptile pet trade. *Animals*, 11(33):676, March 2021.

[16] Sarah-Sophie Weil, Laure Gallien, Michaël P. J. Nicolaï, Sébastien Lavergne, Luca Börger, and William L. Allen. Body size and life history shape the historical biogeography of tetrapods. *Nature Ecology Evolution*, 7(9):1467–1479, September 2023.

[17] Joanne M. Bennett, Jennifer Sunday, Piero Calosi, Fabricio Villalobos, Brezo Martínez, Rafael Molina-Venegas, Miguel B. Araújo, Adam C. Algar, Susana Clusella-Trullas, Bradford A. Hawkins, Sally A. Keith, Ingolf Kühn, Carsten Rahbek, Laura Rodríguez,

Alexander Singer, Ignacio Morales-Castilla, and Miguel Ángel Olalla-Tárraga. The evolution of critical thermal limits of life on earth. *Nature Communications*, 12(1):1198, February 2021.

[18] Juan G. Rubalcaba, Sidney F. Gouveia, Fabricio Villalobos, Miguel Á. Olalla-Tárraga, and Jennifer Sunday. Climate drives global functional trait variation in lizards. *Nature Ecology Evolution*, 7(4):524–534, April 2023.

[19] Eric R. Pianka. *Ecology and Natural History of Desert Lizards: Analyses of the Ecological Niche and Community Structure*. Princeton University Press, March 2017. Google-Books-ID: h4YrDgAAQBAJ.

[20] Raúl Araya-Donoso, Esteban San Juan, Ítalo Tamburrino, Madeleine Lamborot, Claudio Veloso, and David Véliz. Integrating genetics, physiology and morphology to study desert adaptation in a lizard species. *Journal of Animal Ecology*, 91(6):1148–1162, 2022.

[21] Miron B. Kursa and Witold R. Rudnicki. Feature selection with the boruta package. *Journal of Statistical Software*, 36:1–13, September 2010.