

Project Description

Introduction

Organismal phenotypes are the result of genetic *and* environmental forces, and the interaction between them. When explaining patterns of phenotypic evolution, we can account for some genetic influence by incorporating a phylogeny and drift, recognizing that phenotypes tend to be more similar between closely related species. Descent with modification thus provides us with some expectations as to how phenotypes should evolve. But how do we account for influence of extrinsic biotic factors on observed traits? In response to these questions, and in an attempt to explain the accumulation of organismal diversity, the field of phylogenetic comparative methods was developed. Unfortunately, nearly all commonly used probabilistic models for describing phenotypic evolution are ecologically neutral, meaning they fail to account for species interactions with one another, and their environment. This results in lifeless models that mathematically describe the *pattern* of phenotypic evolution, but lack a biological explanation of the *process*. **My overarching research goal is to extend our existing toolkit to include a more diverse methodological framework which accounts for macroecology in macroevolution.** This includes incorporating inter-lineage interactions, intra-lineage variation, and historical biogeography, to better explain observed patterns in trait diversity. To develop these methods, I intend to use Australian reptiles as an empirical example.

But, why Australian reptiles? As an island continent, Australia is a natural laboratory for studies of evolution. With more than 1,000 species of squamate reptiles (lizards and snakes), extensive museum collections, and curated occurrence records, Australia presents a prime opportunity for investigating macroecological models of phenotypic evolution. Since parting ways with Antarctica, the Australian continent has been isolated for around 35 million years. This means that with the exception of a single preexisting Gondwanan clade (Pygopodoid geckos), all other Australian reptile groups have arrived from Asia and radiated in situ. This includes iconic radiations of pythons, monitor lizards, elapid snakes, skinks, agamid lizards, and blind snakes. Australian institutions also provide extensive collections (**500,000+**), and carefully curated occurrence records, making it a prime location for investigating macroecological models of phenotypic evolution.

Research Plan

Research Objectives

The overarching goal of this project is to better account for macroecological processes in macroevolutionary methods. This includes designing new tools *and* implementing them to tackle empirical evolutionary questions. We can classify the broad aims of this study into four specific questions:

1. Does including interspecific interactions improve our inferences of phenotypic evolution?
2. How do interspecific interactions affect the spatial evolution and distribution of species?
3. Can we build an extensive ecological dataset of Australian reptiles on a continental scale?
4. Is competitive signal identifiable from phenotypic and spatial data of Australian reptiles?

Methods

This research depends directly on the development of novel methods and the collection and curation of a specimen-level eco-morphological database (of Australian reptiles). The ultimate goal is to integrate these to directly address our questions of interlineage interactions using empirical data.

I will use Australian museum specimens to collect existing and new data on reptiles to develop an ecological database that will be incorporated into the Atlas of Living Australia (ALA) [3]. The approaches used to collect and integrate the data will likely provide a benchmark for future ALA projects across other Australian organismal groups. The process of developing the comparative methods will also be of direct interest for those hoping to follow a similar methodology.

I anticipate that the incorporation of multiple factors influencing trait and spatial evolution may require complex, parameter-rich models. Methods for modelling univariate continuous traits have traditionally depended on the ability to directly estimate the likelihood of the data given the model. With increasing complexity, likelihood calculations become difficult or intractable, and so I suggest the development instead, of an **Approximate Bayesian Computation** (ABC) framework. ABC methods rely on the ability to simulate data under a designated model, then determine the similarity between simulated and observed summary statistics. In this way, ABC is a remarkably flexible and valuable tool for complex model-fitting, and will be a boon to this proposal.

Objective 1: Does including interspecific interactions improve our inferences of phenotypic evolution?

Interspecific interactions act as a strong selective force on organismal phenotypes. Interactions may be competitive (vying for resources), mutualistic (beneficial to both parties), commensal, parasitic, or predatory, and dictate a variety of dynamic responses. This has led to theories of character displacement, competitive exclusion, and phenotypic matching, among others. In a classic anecdote, Darwin and Wallace each considered a Madagascan orchid with a foot long nectary. They then independently determined there must be a moth pollinator with an equally long proboscis to match. The later discovery of just such a moth was ultimately less of a surprise than it was a fulfillment of the expectation that organismal interactions can dictate phenotypic evolution, even to absurd extremes.

Despite us witnessing these interactions every day, probabilistic evolutionary models still treat *all* lineages identically, and until just recently [1,4–6], ignored interactions altogether. Building off conceptual work by Nuismer & Harmon [7], Drury et al. [5] and Manceau et al. [6] elegantly integrated a system of ordinary differential equations for estimating the effect of competition on trait evolution in a maximum likelihood framework. We can thus estimate a parameter S which describes the strength of the interaction, as well as the polarity: negative values of S indicate competition or repulsion, positive values indicate attraction towards common values. In its most simplistic form, the S parameter interacts with the mean trait values of the interacting lineages (vector X_t), to reflect their relationship. To take into account changes through evolutionary time, the S parameter further interacts with the evolutionary rate (σ), and drift (d), to dictate the trajectory of trait evolution. As part of my PhD research, I have extended this method to account for historical biogeography (inferred separately), which attempts to capture the reality that not all lineages occur sympatrically (an assumption of the original model) (Figure 1). However, one could imagine that a number of factors dictate the interactions between lineages. Niches can be partitioned by ecological means (diel activity, microhabitat, etc.), and I think these aspects could be incorporated by including a number of interaction matrices, similar to how geography currently is (Figure 1). Alternatively, it is unlikely that all lineages are interacting with identical intensity, i.e. estimating a single S parameter. Instead, lineages may interact differently based upon the extent of their geographic (broad or subset sympatry) or active period (nocturnal, crepuscular, diurnal) overlap, or even some uncaptured aspect of their ecology (hidden states). The estimation of multiple interaction (S), rate (σ), and drift (d) parameters however, becomes increasingly difficult in a **maximum likelihood** framework, encouraging the development for methods which do not require full-likelihood calculations, such as **ABC** [8,9].

Objective 2: How do interspecific interactions affect the spatial evolution and distribution of species?

The distribution of species across a landscape can be dictated by both intrinsic and extrinsic influences. In maybe the most obvious examples, habitats may “filter” for species with similar ecologies, or instead, competition might result in exclusion from a given ecosystem. The strength of these interactions (be it competitive or mutualistic) therefore govern the extent to which species cohabitate. This suggests an intersection between geographic distributions and phenotypes. Unfortunately, no current biogeographic models incorporate trait values or competition in their

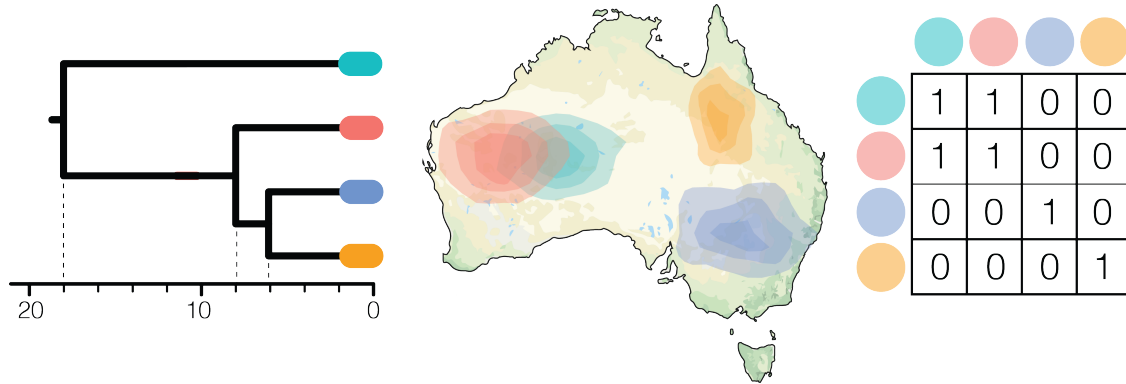


Figure 1: A partial schematic of the process for incorporating biogeographic histories and competition into trait evolution. Left to right: a phylogeny, the spatial distribution of extant taxa, and an interaction matrix for calculating the influence of lineages on one another. Note, it is a symmetric matrix (1 sympatric, 0 allopatric), which allows only sympatric taxa to interact.

reconstructions [10]. Of perhaps greater concern is that in nearly all biogeographic models, species distributions are simplified either to the centroid of their range, or discretized into island, continental, or biome bins. What is lost in translation is the biology of the organism: species ranges may be irregularly shaped, of inconsistent density, and/or vary greatly in size [11,12]. This may be a result of a number of influences such as available habitat and physiological limits, but is likely also dictated by interactions with related (and unrelated) species. I suggest the development of a spatially explicit forward-time simulation-based framework to approximate the processes resulting in current distribution patterns. This framework will incorporate ecology and species traits to inform how interactions affect the spatial evolution of species.

Objective 3: Can we build an extensive ecological dataset of Australian reptiles on a continental scale?

Existing eco-morphological databases cover much of the Australian vertebrate fauna, including mammals [13] and birds [14,15], but reptile taxa remain poorly represented [16]. This is surprising given the number of studies of Australian reptile phylogenetics, ecology, and morphology, and remarkable museum collections. In association with the **Atlas of Living Australia (ALA)** [3], I intend to collate existing, undatabased ecological and morphological information with newly collected data, to span the entirety of Australian reptile taxa (>1000 species). The ALA provides a portal for interacting with museum and citizen-scientist records of Australian organismal diversity, comprising 78,492,847 records of 123,364 species. Native reptiles account for over 1 million of these records, which include important spatial, temporal, and taxonomic information about these taxa. Noticeably absent however, are phenotypic data for any of the 500,000+ reptiles in museum collections. In a departure from existing databasing methods, I hope to integrate morphological and ecological data, along with metadata for all viewed specimens back into the ALA database. My hope is that instead of summarizing this information to species means, I can contribute a database which is more flexible to taxonomic changes and better describes intraspecific variation [17]. As a current research associate of the Western Australian Museum (130,000+ reptile records) and a proposed associate of the South Australian Museum (55,000+ reptile records), I would have access to sampling for the entirety of the continent, making this type of data collection feasible. A resource of this kind would be a world-first, providing invaluable information to future researchers, and returning inalienable data to museums about their own collections.

Objective 4: Is competitive signal identifiable from phenotypic and spatial data of Australian reptiles?

This question represents the synthesis of the three previous objectives. Broadly, my intention is to develop methods that may be of interest to other researchers studying macroevolution and macroecology. However, it is also important to ask these questions empirically, **by employing existing biological collections**. Australian reptiles are particularly well suited for these objectives because they represent a number of species rich, eco-morphologically diverse, endemic radiations. Diversification *in-situ* reduces the influence of immigration and emigration, and simplifies our research. Building on phylogenomic research from my PhD and others, we now have strong species-level time-calibrated phylogenetic hypotheses for nearly all continental reptile groups. This makes macroevolutionary inferences tractable. The Atlas of Living Australia also provides an unparalleled resource for species distributions based on museum occurrence records, for all described taxa. Given existing data and resources, and the access to material, I *honestly* believe that this is one of the best possible systems to focus this research on.

Significance and Broader Impacts

Since Joe Felsenstein [18] highlighted the nonindependence of species in comparative methods (owing to their underlying relatedness—phylogeny), Phylogenetic Comparative Methods (PCMs) have become commonplace throughout evolutionary and ecological research. This is likely the result of expanding phylogenetic and computational resources, and user-friendly implementations [19,20]. Nonetheless, comparative methods have exploded as a popular area of study and application, particularly in the last ten years. In their most basic form, PCMs can provide information about—and correct for phylogenetic signal present in data. Their flexible design however, means that they can be easily extended to incorporate diverse data types. Regardless, many of the inferences we can make from PCMs are intuitive, which makes them exciting and digestible for popular science media.

My proposed research builds on existing ecological theory and macroevolutionary methods, synthesizing the two. I anticipate that the output of this project will be valuable both empirically and methodologically. The construction of an ecomorphological database will provide a resource for a diverse array of biologists working from population to macroevolutionary scales. The comparative ABC method design will also make available a generalizable framework for creating evolutionary models. I aim to publish the data, methods, and empirical findings, and look forward to sharing our results with colleagues and the general public.

Training Objectives and Career Development

Evolutionary biology is increasingly intersecting with computational biology. To expand my versatility in phylogenetic comparative methods, I aim to first extend my abilities in statistical modelling. This means strengthening the connection between theoretical understanding and computational implementation of macroevolutionary processes. I have carefully chosen two sponsors who are experts in these fields, and in bridging the divide between them, bringing together the conceptual and empirical.

As a self-taught coder, there remain a number of shortcomings in my computational abilities, particularly with regards to efficiency. Dr. O'Meara is an experienced member of the new-wave of code-conscious evolutionary biologists, and is the author of popular comparative methods packages treePL, brownie, and OUwie [21,22]. I look forward to also developing a better understanding of the underlying mathematics behind current methods. Dr. Lee is a world-expert in Bayesian modelling, and I will receive training from him in directly in Bayesian computation. His experience in functional morphology will also be of benefit to my interest in lineage interactions.

I have always been interested in a career that merges the creativity of research with the outreach of science communication. I have been and will remain a strong activist for science education and literacy, and am committed to fulfilling this dream. I believe that the training provided by Drs. O'Meara and Lee will help to make me a rising star in the field of evolutionary biology, however this is not my exclusive goal. The South Australian Museum, and proposed workshops, will also provide the opportunity to continue my development as an educator and science communicator. I feel that

as a scientist, and a publicly funded researcher, I owe a debt to make my research digestible and accessible to the public.

Sponsoring Scientist and Institution Justification

I have chosen Drs. Lee and O’Meara specifically for their skill sets, research interests, and proven track records of pushing the evolutionary biology community forward. They are both creative and judicious thinkers, who I admire for their contributions to science and broader education.

Dr. Brian O’Meara (*Methods/Biology*) is an Associate Professor in the Department of Ecology & Evolutionary Biology at the University of Tennessee, Knoxville. His lab group work to build and apply comparative methods in phylogenetics. He is known for his expertise in comparative methods and phylogenetics, and his community-mindedness. Dr. O’Meara is a leader in the new-guard of biology, both a skilled researcher, and conscientious advocate.

Dr. Michael Lee (*Theory/Biology*) is the Leader of Research in Palaeontology at the South Australian Museum, and holds a joint appointment in the Department of Biological Sciences at Flinders University of South Australia. Dr. Lee is primarily an evolutionary biologist, who works at the interface of palaeontology, neontology, and computational biology.

I will spend the first year (Y1) of this fellowship at the **South Australian Museum** (SAM) in Adelaide. The SAM is an amazing resource for researchers across the natural sciences. Though I will be primarily mentored by Dr. Lee during this period, I will also be working remotely with Dr. O’Meara, and building relationships with SAM colleagues. I expect that sometime after the first year, I will transition my position from the data collection phase in Australia, to the methods development phase with Dr. O’Meara in the USA. Both institutions are world-recognized, and have state of the art facilities.

Timetable

The requested duration of this fellowship is **3 years**. The current timetable assumes Year 1 and part of Year 2 will be spent at the South Australian Museum in Adelaide, Australia. The remainder of Year 2 and Year 3 will be spent at the University of Tennessee—Knoxville, USA.

Year	Research	Training	Dissemination
Y1	Collect morphological data. Build reptile database, and integrate it into ALA ¹	Learn databasing methods. Work with the ALA	Release extended specimen data format for ALA, and present database
Y2	Build ABC ¹ framework for PCMs ¹ . Implement empirically to reptiles	Build software development skills	Present methods and results at scientific conferences. Publish project pre-prints
Y3	Package and release ABC framework for lineage interactions	Extend development skills. Train others on methods	Publish methods and empirical findings

¹ALA–Atlas of Living Australia; ABC–Approximate Bayesian Computation; PCM–Phylogenetic Comparative Model