

BACKGROUND:

Rapid climate change is resulting in global-scale alterations to ecosystems and biological communities¹. Among the most commonly reported responses to changing climates are shifts in species' geographic ranges². As climates warm, the suitable conditions constraining species' ranges move in space, with species expected to shift their distributions to track their climate niche optima³. These shifts are often due to differential population dynamic processes occurring across species' ranges⁴, with populations expanding at cold edges and contracting at warm edges⁵. The expected outcome of these processes is concerted range shifts⁶. However, increasing evidence shows that the majority of documented range shifts on land are occurring in a manner inconsistent with these expectations⁷, including range expansions and contractions occurring counter to climate gradients. Compounding evidence shows that speed and directionality of range shifts are ambiguous at best and counterintuitive at worst (Baecher co-author, *In review*)⁸. This has resulted in a call to focus on how species' key traits, such as dispersal, demography, and genetics, may be critical for uncovering a rules-based theory of range shift dynamics (Baecher co-author, *Submitted*)⁹.

Dispersal, demography, and genetics are perhaps the most critical factors linking individual- and population-level processes to range shift dynamics. In particular, the expansion of range edges under novel and potentially hazardous conditions requires sufficient dispersal (i.e., colonization), sustained demographic structure (i.e., reproduction and recruitment)¹⁰, and resilient phenotypes (i.e., higher critical thermal limits). Theory indicates that exposure to these conditions (e.g., unstable climate) selects for dispersive and environmentally tolerant phenotypes¹¹. Likewise, colonizing populations are sensitive to demographic processes (e.g., Allee effects), which impact their stability (e.g., stochastic extinction) and evolutionary trajectory (e.g., inbreeding, drift, or founder effects)¹². Therefore, expanding-edge populations often exhibit novel characteristics affecting range shift outcomes, including dispersive phenotypes¹³, unique demographic structure¹⁴, and adaptive genetic traits¹⁵. While the mechanics of these processes are well studied in some insects¹⁶ and plants¹⁷, studies of vertebrate range shift dynamics focusing on these factors are surprisingly rare.

These processes are further complicated by habitat disturbances, such as human land use modification, which can reduce individuals' ability to traverse landscapes and track suitable climates¹⁸. A recent study found that range shifts of terrestrial vertebrates are far less predictable than their marine counterparts¹⁹ due to land-use change. The impact of landscape alterations are both broad-reaching and varied, depending on the mode of change. For instance, habitat loss (i.e., reduction in the size of intact habitat) can extirpate populations in affected areas while isolating those in suitable areas. Alternatively, habitat fragmentation (i.e., division of intact habitat into smaller patches) can cause extirpation by isolating populations into discrete habitats, thereby altering gene flow^{20,21}. In sum, interactions between extrinsic factors (i.e., climate and habitat) and species' intrinsic traits (i.e., dispersal and demography) likely determine range shift outcomes⁹. While these processes are undoubtedly important for understanding range shift dynamics, the underlying eco-evolutionary dynamics are largely unknown. This highlights the critical need for both a process-oriented framework and comparative approaches that can scale our understanding from isolated single-species studies towards a comprehensive understanding.

This proposal addresses these inherent complexities by focusing on eco-evolutionary feedback between environmental disturbance and aspects of species' intrinsic traits. This complexity is rarely captured by phenomenological species distribution models, which use basic occurrence data and environmental layers to project species' range shifts into the future—perhaps explaining the current limited understanding of range shift dynamics on a changing planet. We believe that deriving Rules of Life governing range shift dynamics demands the integration of biological information, from genes to populations to environments, through first principles. This proposal seeks to develop the conceptual framework, data, and simulations necessary to improve our understanding of vertebrate range shift dynamics. To address these goals, we focus on North American birds and mammals—a data-rich group of terrestrial vertebrates representing divergent and varied dispersal modes, ecological characteristics, and both physiological and morphological attributes.

RESEARCH OBJECTIVES, METHODS, AND SIGNIFICANCE:

We lay out three research objectives (RO₁₋₃), representing independent yet complimentary activities with outputs (i.e., publications and databases; see timeline in Fig. 2) that contribute to a larger conceptual and empirical framework. We note that there are key broader impacts listed as separate objectives below. **RO₁ seeks to create phylogenetically-based predictions of dispersal capacity for North American birds and mammals.** This will be achieved by gathering empirical estimates of species' dispersal and associated phenotypes (e.g., morphology and physiology), and using phylogenetic comparative methods to impute dispersal capacities across poorly-studied species. **RO₂ classifies species into functional groupings based on emergent eco-evolutionary characteristics**, including dispersal capacity, phenotypes, demographics, and evolutionary distinctness. **RO₃ uncovers causal mechanisms of climate-induced range shift outcomes** by modeling spatial population dynamics of species groupings under demographic and environmental change and validating results with empirical range shift data. This approach uses agent-based eco-evolutionary models to establish mechanistic linkages between species' intrinsic traits (i.e., dispersal and demographics) and extrinsic climate and habitat change scenarios (habitat loss and fragmentation). **RO₂** and **RO₃** are flexible to data contingencies and can be further modularized by balancing taxonomic representation through stratified sampling, as well as standard sensitivity tests to explore model parameterization and starting conditions.

CREATE DATABASE OF DISPERSAL CAPACITY (RO₁):

Goal: Develop a comprehensive database of species dispersal and associated traits.

Question: Can we predict species' dispersal capacities using dispersal-associated phenotypes, biological traits, and phylogenetically-guided imputation?

Table 1: Data required to impute dispersal capacities of North American terrestrial birds and mammals.					Methods: Estimates of dispersal capacities (average distance per unit time; <i>sensu</i> ²²) will be collated through a structured literature search using PRISMA standards ²³ . Because the definition of dispersal varies widely by study and taxa, I	
Category	Variable		N _{species} (%coverage)			Source
	Birds	Mammals	Birds	Mammals		
Phylogeny	supertrees		993 (100%)	390 (100%)		VertLife
Physiology	metabolic rate		140 (14%)	69 (%18)		AnimalTraits
Morphology	mass, HWI mass		937 (94%)	290 (74%)		AVONET , PanTHERIA
Demography	reproduction, fecundity		892 (90%)	343 (88%)		Comadre
Ecology	homerange, range area		993 (100%)	390 (100%)		BirdLife, Marsh et al. 2022
Life History	diet, trophic position		937 (94%)	290 (74%)		AVONET , PanTHERIA
Migration	migratory behavior		993 (100%)	390 (100%)	IUCN Redlist	
N _{species} = estimated number of focal species in database; %coverage = per cent of total species.						

N_{species} = estimated number of focal species in database; %coverage = per cent of total species.

will develop a protocol for inclusion criteria, reporting standards, and data harmonization. To preserve the original context of estimates reported in the literature, the resulting database will include metadata and references to the literature (including specific passages). For species that lack dispersal estimates, we also query MoveBank²⁴ for raw movement data sets. These data sets vary by quantity (i.e., spatiotemporal scale and extent) and quality (i.e., bias, accuracy, and context)²⁵, therefore I will follow best practices to vet, clean, and filter movement data for high-throughput processing²⁶. Resulting data will be analyzed through a path segmentation framework²⁷ to identify discrete dispersal events and fit dispersal kernels (a model of dispersal distance) for our estimates²⁸.

To determine the phylogenetic basis of dispersal capacities, supertrees will be retrieved from VertLife for birds²⁹ and mammals³⁰. Next, phenotypes including morphology (e.g., body size) and physiology (e.g., metabolic rate) will be characterized using the AnimalTraits³¹ database. For birds, the morphometric measure, Hand-Wing Index (HWI), is a strong predictor of dispersal capacity³² and is available for all North American species through AVONET³³. Additionally, migratory species often have greater dispersal distances than non-migrants^{34,35}, and therefore migratory statuses (e.g., full migrant, nomadic, non-migrant) will be extracted from IUCN Redlist. Next, life history and demographic traits, such as fecundity and trophic position, have been shown to serve as proxies of dispersal for both birds and mammals²² (available through AVONET³³, PanTHERIA³⁶, and Comadre³⁷). Lastly, range size and latitudinal breadth are two key variables for predicting species' dispersal capacity²². These variables can be quantified by accessing species' range polygons for

birds (*BirdLife*³⁸) and mammals³⁹ and calculating the area and distance between poleward-most edges. Key to this effort will be taxonomic harmonization, which is tractable given high quality information available for North America.

A Phylogenetic Linear Mixed-Model (PLMM) will be fitted to species' known dispersal distances (response) and covariates described above to impute unknown dispersal distances. Predictions will be obtained for any species with sufficient clade-level dispersal estimates and covariate data. However, for species missing basic covariate information (e.g., body size and trophic position), multiple phylogenetic imputation⁴⁰ will be used to predict values under an assumed rate of evolutionary change (e.g., Brownian model⁴¹: $rate \propto \sqrt{branch\ length}$). To establish a threshold of acceptable data missingness and partiality, error and bias metrics will be used to determine which taxonomic groups to prune from final models^{42,43}. Thus this analysis does not strive for complete taxonomic coverage, but rather to achieve representation adequate to meet the needs of RO₂ and RO₃. Candidate model sets for birds and mammals will be created featuring combinations of above-mentioned variables and predictions will be gathered from the top-ranked model using Akaike information criterion⁴⁴. Prediction uncertainty will be estimated using maximum likelihood bootstrapping across alternative branch lengths and topologies⁴⁵. For species with imputed dispersal, model details will be included as metadata in the resulting database (i.e., model covariates and performance metrics). The resulting database and code necessary to reproduce analyses will be unveiled through the publication of an Open-Access data paper.

CLASSIFY SPECIES INTO GENERALIZED DISPERSAL GROUPINGS (RO₂):

Goal: Establish biological species groupings based on dispersal capacities and dispersal-associated characteristics (i.e., traits, demography, and genetics).

Question: Are there emergent groupings in species dispersal-associated intrinsic traits across the phylogeny of birds and mammals?

Methods: To further determine how species intrinsic traits influence range shift outcomes, a classification scheme will be created to identify generalized species groupings. These groupings will be determined by four main characteristics: (i) dispersal capacities; (ii) traits; (iii) demographic information; (iv) and the genetic architecture of dispersal. The first two categories will be informed by the database created through RO₁, including variables for dispersal distance, body size, metabolism, fecundity, and trophic position. The next category, demographic information, is largely focused on sex-limited dispersal behaviors and, where available, density-dependent dispersal within taxonomic groups. While these data are not extensively sampled across the >2800 species of mammals and birds in North America, the *Comadre* database provides data for these variables across a wide array of species via population matrix model predictions. These data will be used in addition to direct information from the literature. The final characteristic describes genetic linkages to dispersal phenotypes, including heredity, ploidy, and loci characteristics (neutral versus adaptive). These genetic mechanisms have been identified for a moderate number of taxonomically diverse species of birds and mammals⁴⁶. While these characteristics are not available broadly, simulations in RO₃ will explore a variety of known genetic architectures (e.g., single or multi-chromosome mapping; neutral or adaptive trait loci) through alternative model parameterizations.

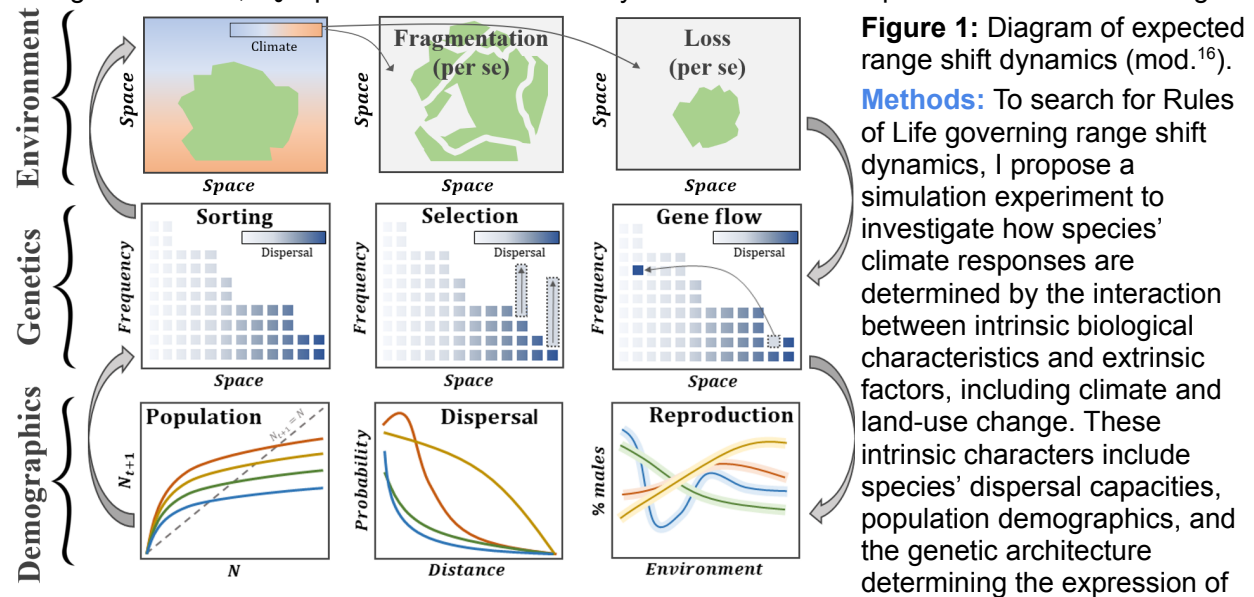
To quantify variation across these characteristics, we will calculate pairwise trait dissimilarities using Gower's distance⁴⁷ (robust to missing data). Next, we will map each species into a lower-dimensional "trait space" using a phylogenetic principal component analysis⁴⁸. To search for biological groupings, we will apply the Ward hierarchical agglomerative clustering algorithm⁴⁹ across a range of *a priori* groupings (e.g., 5 – 10). The result of this analysis will delineate groupings of birds and mammals who share unique and distinctive characteristics related to dispersal, allowing for generalizations of the intrinsic traits underlying range shift dynamics. The composite values of each of these variables will be further used to parameterize eco-evolutionary range shift models in RO₃.

MODEL MECHANISMS OF ECO-EVOLUTIONARY DYNAMICS (RO₃):

Question: How does dispersal, demography, and genetics influence species responses to environmental changes?

Goals: Simulate range shift dynamics using a process-based model.

Hypotheses: **H₁:** range shift rate will decrease under habitat loss and fragmentation when species' intrinsic traits are held equal; **H₂:** directionality of range shifts is dependent on the presence of contiguous habitat; **H₃:** species' intrinsic traits may alleviate the consequences of land-use change.



dispersal phenotypes. Dispersal capacities will be informed by species groupings identified in RO₂, while other characteristics will be explored through alternative model parameterizations based on their observed prevalence in each taxonomic group (e.g., density-dependent v. non-dependent dispersal; single v. multi-chromosome mapping, and neutral v. adaptive trait loci). In this framework, these characteristics are ascribed to a simulated population of animals traversing artificial landscapes under a user-defined climate gradient, with the goal of examining range shift outcomes under the factorial combination of intrinsic and extrinsic environmental factors. This experiment will utilize the *RangeShifter* software, an agent-based model that simulates spatially-explicit population dynamics in changing environments⁵⁰. *RangeShifter* differs from other software by providing a genetic module, which links a species' dispersal capacity to a custom genetic architecture and tracks each individuals' genome throughout the simulation. This provides a formal accounting of eco-evolutionary feedbacks between a species' dispersal capacity, phenotypes, and their environment, mediated by stochastic population processes.

The design of this experiment focuses on simulating hypothetical species under three simulated land-use change scenarios: intact habitat; habitat loss; and fragmentation. Therefore, the number of treatments in this simulation experiment will be determined by the factorial combination of extrinsic factors (3 land-use change scenarios) and intrinsic factors (quantity contingent upon the number of groupings generated in RO₂). The focal outputs of these simulations include range shift velocity, total area of occupancy, and evolution of dispersal capacity. Range shift velocity (RSV) will be calculated as the annual displacement of a species' range, measured from the range center and both leading and trailing edges (positive values indicate shifts toward cooler temperatures, negative values toward warmer temperature). Following simulations, these state processes will be compared across treatments to test above-described hypotheses.

This simulation experiment was designed to address core uncertainties in range shift science identified by the *BIOSHIFTS*⁵¹ working group—an international coalition of researchers led by Jonathon Lenoir (proposal co-sponsor) and Lise Comte (endorsed collaborator) with the goal of understanding the intrinsic and extrinsic factors driving contemporary range-shifts. Through my continued membership, I have access to the most up-to-date version of their range shift database⁵² (v.3), featuring geographically referenced range shift data (i.e., annual expansion rate and directionality) of birds (1,501 species) and mammals (68 species) across the globe. I will work in

concert with the 15 group members to create a model validation framework linking simulated range shift dynamics (range shift velocity and total occupied area) with empirical range shift outcomes in the *BIOSHIFTS* database. This goal will be accomplished by establishing analogs between the biological species groupings (created in RO₂) and species documented in the *BIOSHIFTS* database, including the habitat suitability within each species' range area (data currently available⁵³). This approach will allow for testing of above-described hypotheses (H₁₋₃) using both simulated and empirical data and relevant land-use change scenarios.

RESEARCH SIGNIFICANCE:

A broad understanding of the eco-evolutionary mechanics of range shifts is currently lacking in the literature. The majority of these pioneering studies have been limited primarily to insects¹⁶ and plants¹⁷, leaving the causes of vertebrate range shift dynamics poorly understood. This issue is underscored by recent meta-analyses showing that vertebrate's range shift dynamics largely departed from theoretical expectations¹⁹. This proposal describes three research objectives to address these limitations. The first and most critically-needed objective will be to collate and predict the dispersal capacities and associated phenotypes of North American birds and mammals. This folds into the second objective, which searches for emergent groupings in species' intrinsic traits that may underlie range shift outcomes. The final objective leverages the emergent groupings to conduct a series of simulation experiments, comparing range shift dynamics of each hypothetical species under alternative land-use scenarios, and validating results with empirical data. The goal of these experiments is to model the mechanisms which intertwine intrinsic and extrinsic factors through eco-evolutionary dynamics. These objectives will yield a conceptual framework, as well as empirical data, necessary to confront our limited understanding of vertebrate range shift dynamics.

CAREER DEVELOPMENT AND FUTURE RESEARCH:

This proposal represents a significant departure from my previous activities, which has centered on using data from targeted field work to address local management issues pertaining to reptiles and amphibians. This research proposes a continental-scale compilation of biodiversity information and comparative evolutionary approaches, using modern analytical tools and cloud computing, to develop the data and theory required to search for fundamental rules of biodiversity change. Additionally, I plan to mentor undergraduates, host a workshop on biodiversity informatics, design analytical tutorials, and develop cyberinfrastructure to promote cross-institutional collaborations. As an aspiring academic seeking a tenure-track faculty position in quantitative ecology, these activities will build towards my career goals of engaging with undergraduate and masters students through: (i) teaching spatial ecology and statistical modeling courses and (ii) developing a dual-research program focusing on both hands-on (field and museum) and computational (geospatial and statistical) projects. Future directions for my research include (i) investigating local to regional impacts of environmental change on biodiversity, (ii) using eco-evolutionary mechanisms, such as those proposed herein, to forecast range shift dynamics, and (iii) developing a framework for predicting global ecological connectivity.

TRAINING OBJECTIVES:

This fellowship will allow me to expand my scope of experience through three key Training Objectives (TO₁₋₃). **TO₁**: gain experience in modern biodiversity informatics techniques for integrating large databases, managing cross-platform workflows, creating structured simulations guided by theory, and computing on cloud or cluster infrastructure. My proposed research objectives require training in Google Earth Engine and Structured Query Language (SQL), which will be provided through the Data Carpentries Program and online JavaScript and Python programming tutorials. These objectives will fulfill my remaining requirements for instructorship with the Ecological Data Carpentries program at UF (BO₂). **TO₂**: practice analytical procedures for studying eco-evolutionary dynamics. Through the Guralnick lab, I will gain training in phylogenetic comparative methods for quantifying the evolutionary relationships of focal phenotypes. Next, I will enroll in a UF course on the use of Individual Based Models (IBMs) for simulating biological phenomena under dynamic environments. **TO₃**: work across taxonomic and disciplinary provinces. This proposal requires that I read broadly on the biology and natural history of birds and mammals, groups which I have limited

experience. As an aspiring academic with interests in global change biology, this goal will help expand my prospective research program to include broad subjects for study.

JUSTIFICATION OF SPONSORING SCIENTISTS AND HOST INSTITUTION:

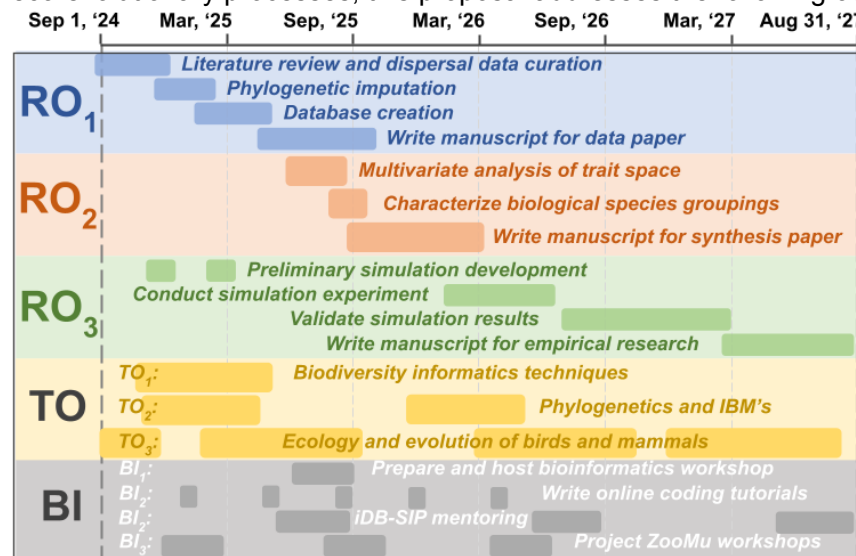
This research will be conducted under the sponsorship of Drs. Guralnick (UF and FLMNH) and Lenoir (CNRS, Université de Picardie Jules Verne). Dr. Guralnick is a leading biodiversity informatician, who will provide instruction with comparative phylogenetic methods, Individual Based Models, and database development. Dr. Lenoir is a prominent global change biologist who will guide the use of eco-evolutionary data for modeling and interpreting range shift dynamics. Through the FLMNH and Guralnick lab, I will be exposed to a diverse network of potential collaborators and resources within novel research areas. Furthermore, through Lenoir and our shared affiliation with the *BIOSHIFTS* working group⁵¹, I will expand my network of international collaborators, including organizations such as *iDiv*, the German Center for Integrative Biodiversity Research.

BROADER IMPACTS:

I have identified three broader-impact initiatives (BI_{1-3}) which underscore my existing skills and partnerships: (BI_1) highlighting the importance of transparency and accessibility in science; (BI_2) training students in modern biodiversity informatics; and (BI_3) promoting collaborations between natural history collections at zoos and museums. To address BI_1 , I will lead a workshop in 2025 on the use of APIs to access widely available databases (e.g., *IUCN*, *GBIF*, *NEON*) in the R programming language. I will also design a series of tutorials highlighting good-faith analytical procedures (i.e., publishing code and using open-source software) on my personal research blog (alexbaecher.com/#posts). With the support of the Guralnick Lab and FLMNH, I am uniquely poised to achieve BI_2 by providing support to undergraduate and early career scientists in guided research. I will train undergraduate researchers through a paid FLMNH internship program, with which the Guralnick lab regularly hosts students. I will also seek certification as a Data Carpentries instructor to educate students on programming fundamentals. Finally, I will lead the development of a data-sharing cyberinfrastructure for project *ZooMu*⁵⁴, an organization of researchers working to build collaboration between zoos and museums (BI_3), both of whom play key roles in biodiversity conservation. I will attend working groups to discuss three objectives: (i) incorporating metadata standards⁵⁵ in zoo archiving software; (ii) creating data standards, such as metadata and individual identifiers, for linking records between zoos (live data) and museums (preserved data); and (iii) increasing accessibility of zoo data, while respecting zoos' privacy and sensitive data.

ALIGNMENT WITH COMPETITIVE AREA 2:

This proposal seeks to identify Rules of Life governing climate-induced range shifts of North American birds and mammals. We hypothesize that range shift dynamics are a result of complex eco-evolutionary processes spanning multiple levels in the hierarchy of life: from the genetics of dispersal to population-level phenotype expression to species' distributions. To examine these eco-evolutionary processes, this proposal addresses the following objectives: (i) develop a database



of dispersal capacities and associated phenotypes using comparative phylogenetic methods; (ii) establish biological species groupings; and (iii) conduct computational experiments to simulate range shifts and validate results. This proposal will yield a predictive framework to understand how and why range shifts occur.

Figure 2: Timeline for completing proposal objectives.