The Difficulty of Predicting Evolutionary Change in Response to Novel Ecological Interactions: A Field Experiment with *Anolis* Lizards

Timothy J. Thurman,[1,2,*,†] Todd M. Palmer,[3] Jason J. Kolbe,[4] Arash M. Askary,[1] Kiyoko M. Gotanda,[1,5,‡] Oriol Lapiedra,[6] Tyler R. Kartzinel,[7] Naomi Man in't Veld,[8] Liam J. Revell,[9] Johanna E. Wegener,[4] Thomas W. Schoener,[10] David A. Spiller,[10] Jonathan B. Losos,[11] Robert M. Pringle,[12,*] and Rowan D. H. Barrett[1,*]

1. Redpath Museum and Department of Biology, McGill University, Montreal, Quebec, Canada;   2. Smithsonian Tropical Research Institute, Panama, Republic of Panama;   3. Department of Biology, University of Florida, Gainesville, Florida 32611;   4. Department of Biological Sciences, University of Rhode Island, Kingston, Rhode Island 02881;   5. Department of Zoology, University of Cambridge, Cambridge, United Kingdom;   6. Centre for Ecological Research and Applied Forestries, Cerdanyola del Vallès, Catalonia 08193, Spain;   7. Department of Ecology, Evolution, and Organismal Biology, Brown University, Providence, Rhode Island 02912; and Institute at Brown University for Environment and Society, Brown University, Providence, Rhode Island 02912;   8. Student Wellness Hub, McGill University, Montreal, Quebec, Canada;   9. Department of Biology, University of Massachusetts, Boston, Massachusetts 02125; and Facultad de Ciencias, Universidad Católica de la Santísima Concepción, Concepción, Chile;   10. Department of Evolution and Ecology, University of California, Davis, California 95616;   11. Department of Biology and Living Earth Collaborative, Washington University, St. Louis, Missouri 63130; 12. Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544

ABSTRACT: Determining whether and how evolution is predictable is an important goal, particularly as anthropogenic disturbances lead to novel species interactions that could modify selective pressures. Here, we use a multigeneration field experiment with brown anole lizards (*Anolis sagrei*) to test hypotheses about the predictability of evolution. We manipulated the presence/absence of predators and competitors of *A. sagrei* across 16 islands in the Bahamas that had preexisting brown anole populations. Before the experiment and again after roughly five generations, we measured traits related to locomotor performance and habitat use by brown anoles and used double-digest restriction enzyme–associated DNA sequencing to estimate genome-wide changes in allele frequencies. Although previous work showed that predators and competitors had characteristic effects on brown anole behavior, diet, and population sizes, we found that evolutionary change at both phenotypic and genomic levels was difficult to forecast. Phenotypic changes were contingent on sex and habitat use, whereas genetic change was unpredictable and not measurably correlated with phenotypic changes, experimental treatments, or other environmental factors. Our work shows how differences in ecological context can alter evolutionary outcomes over short timescales and underscores the difficulty of forecasting evolutionary responses to multispecies interactions in natural conditions, even in a well-studied system with ample supporting ecological information.

Keywords: experimental evolution, parallel evolution, species interactions, adaptation.

* Corresponding authors; email: timothy.thurman@mail.mcgill.ca, rpringle@princeton.edu, rowan.barrett@mcgill.ca.
† Present address: Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205.
‡ Present address: Department of Biological Sciences, Brock University, St. Catharines, Ontario, Canada.

ORCIDs: Thurman, https://orcid.org/0000-0002-9602-6226; Palmer, https://orcid.org/0000-0001-5710-9750; Kolbe, https://orcid.org/0000-0002-4236-9960; Gotanda, https://orcid.org/0000-0002-3666-0700; Lapiedra, https://orcid.org/0000-0003-0383-0061; Kartzinel, https://orcid.org/0000-0002-8488-0580; Revell, https://orcid.org/0000-0003-0767-4713; Wegener, https://orcid.org/0000-0002-3424-528X; Losos, https://orcid.org/0000-0003-4712-8258; Pringle, https://orcid.org/0000-0001-7487-5393; Barrett, https://orcid.org/0000-0003-3044-2531.

Introduction

In a famous thought experiment, Gould (1989) imagined "replaying life's tape": if the history of life on Earth started again, would evolution produce the same outcomes? This metaphor neatly encapsulates a fundamental debate in biology about the predictability of evolution and the relative role of deterministic versus stochastic forces (Blount et al.

2018). This debate has played out in many arenas, from the role of selection versus drift in structuring genetic variation within populations (Kreitman and Akashi 1995; Nei et al. 2010) to macroevolutionary studies of phenotypic convergence across distantly related lineages (Losos 2011). The predictability of evolution is not merely an academic concern. Anthropogenic disturbances are a growing threat to biodiversity, and predicting how species and ecosystems will respond to these challenges is crucial for effective conservation (Urban et al. 2016).

As the economic costs of invasive species increase and climate-change-induced range shifts create novel communities (Williams and Jackson 2007; Diagne et al. 2021), predicting how novel species interactions could alter evolutionary trajectories will be particularly important. Such novel interactions can have profound ecological and evolutionary consequences (Mooney and Cleland 2001; Sax et al. 2007; Simberloff et al. 2013). These outcomes depend on the nature of the interaction: novel competitors, for example, may be less likely to cause species extinctions than novel predators (Davis 2003; Doherty et al. 2016). The ecological effects of predation and competition frequently depend on environmental conditions (Chamberlain et al. 2014) and can interact with one another (Chase et al. 2002; Chesson and Kuang 2008). These dependencies and interactions may influence the degree to which evolution caused by novel species interactions is predictable.

Biologists have often studied "predictability" through a retrospective approach, by examining multiple lineages that appear to have faced the same adaptive challenges. Repeated evolution of the same genotypes or phenotypes in the past is evidence that similar changes could be predictable in the future. Studies taking this approach have generally found that evolution can be repeatable and thus retroactively predictable, although the degree of predictability varies across traits, timescales, and levels of biological organization (Oke et al. 2017; Bolnick et al. 2018). The chief drawback of this approach is that direct knowledge of the past is scarce. The prior states of both the lineages being studied and the ecological conditions under which they evolved must be assumed or reconstructed, introducing uncertainty into inferences about predictable evolutionary trajectories (Schluter et al. 1997). The difficulties of reconstructing the past multiply when considering complex scenarios, such as understanding how multiple species interactions might alter evolutionary trajectories in diverse assemblages.

In principle, experimental approaches are a powerful way to address this issue while also allowing for tests of "true" (i.e., prospective) predictability. In practice, experiments present their own challenges. Many evolutionary experiments are done in the laboratory using model organisms (e.g., Lenski et al. 1991; Segrè et al. 2006; Burke et al.

2010). Evolution in the laboratory can be predictable, depending on factors such as population size, experimental complexity, and the genetic diversity of starting populations (Bailey et al. 2018; Blount et al. 2018). However, a common critique of laboratory evolution studies is that they are too environmentally simple (Kawecki et al. 2012; Bailey and Bataillon 2016). Results from the laboratory may not apply to natural populations in which context-dependent behaviors or variable environments might complicate evolutionary predictions. Field experiments with natural populations are one way to address this issue. Unfortunately, the logistical difficulties of fieldwork often necessitate simple experimental designs and low levels of replication. Researchers often manipulate only one factor (e.g., Losos et al. 2004, 2006) or manipulate a slate of factors simultaneously, which complicates causal attribution (e.g., by transplanting organisms to a different habitat; Reznick et al. 1990). Although simple designs can be powerful, understanding the evolutionary outcomes of multispecies interactions or the independent effects of covarying environmental variables requires more complex experiments.

Here we present results from a large-scale, multigeneration field experiment with *Anolis* lizards in which we manipulated both predation and competition. Caribbean *Anolis* lizards are an oft-cited example of predictable evolution, as different species have repeatedly evolved into six ecomorphs during their adaptive radiation across the Greater Antillean islands (Williams 1972). Ecomorphs are adapted to different microhabitats. For example, trunk-ground species occupy broad perches near the ground and are stocky, long legged, and dusky colored; in contrast, trunk-crown anoles occur higher in the canopy, are shorter limbed and green, and have larger toepads (Losos 2009). Laboratory and field studies have shown how the traits that distinguish ecomorphs, such as skeletal (particularly limb) and toepad morphology, are linked to biomechanical performance in these different microhabitats (reviewed in Losos 2009).

The adaptive radiation of anoles is generally interpreted through the lens of competition: *Anolis* species from different ecomorphs spatially partition habitats to facilitate coexistence (Losos 2009), and this ecological displacement can drive rapid phenotypic adaptation (Stuart et al. 2014). However, previous field experiments have shown that predation can also have important impacts on anoles. Introduced predators can alter the strength and direction of natural selection on morphological traits and behavior (Losos et al. 2004, 2006; Lapiedra et al. 2018). Thus, there is strong reason to expect that both competition and predation cause predictable adaptive evolutionary changes in *Anolis* poulations, but there is less basis for predicting how these forces might interact.

We performed our experiment on 16 Bahamian islands already occupied by brown anoles (*Anolis sagrei*), a sexually dimorphic, trunk-ground ecomorph (Butler and Losos 2002). We manipulated predation and competition by introducing a terrestrial predator (the curly-tailed lizard, *Leiocephalus carinatus*) and a congeneric competitor (the green anole, *Anolis smaragdinus*, which is a trunk-crown anole) in a 2 × 2 factorial design. We have previously shown that these introductions have strong ecological effects (Pringle et al. 2019). Both predators and competitors suppressed population growth of brown anoles, and predators caused brown anoles to shift their habitat use to higher perches. On islands with all three species, this upward shift of brown anoles intensified competitive overlap, in terms of both space use and diet, with the more arboreal green anole. On two of the four experimental islands with all three species, this increased competition for predator-free space led to the extirpation of the introduced green anole population.

In the current study, we sought to predict the short-term evolutionary consequences of these strong ecological interactions in our experimental populations of brown anoles. On small islands in the Bahamas, low year-to-year survival rates make brown anoles effectively an annual species with one generation per year (Calsbeek and Smith 2007; Cox and Calsbeek 2010). We quantified phenotypic and genetic changes after 5 years (roughly five generations). Our research builds on a long tradition of laboratory and field experiments with *Anolis* lizards, which informs our predictions about how brown anoles might evolve in response to introduced predators and competitors. Three morphological traits in particular have been extensively studied in relation to how these ecological interactions affect survival, locomotion, and behavior.

First, body size has been shown to be under natural selection in anoles, as in many animals. Multiple studies have found that predators exert selection for larger body size in females but not in males (Losos et al. 2004; Calsbeek and Cox 2010). Larger lizards may be harder for predators to subdue, and female brown anoles, which are smaller and perch lower to the ground than males, may experience stronger predator-induced selection. We thus predicted that introduced predators would result in the evolution of larger females (as measured by snout-vent length [SVL]).

Second, hindlimb length has been extensively studied in relation to locomotor performance, perch use, and predation. Lizards with relatively longer hindlimbs are faster sprinters on broad surfaces, while lizards with shorter limbs are more agile on narrow perches (Losos and Sinervo 1989; Losos and Irschick 1996). Losos et al. (2004, 2006) introduced curly-tailed lizards to islands with brown anoles, which began perching on higher, narrower perches, leading to strong natural selection on relative hindlimb length. Thus, Losos et al. (2006) predicted that over multiple generations, brown anoles should evolve shorter hindlimbs better suited to moving on the high, thin perches that offer refuge from predators. Given the changes in perch behavior previously documented in our experiment (Pringle et al. 2019), we predicted that the relative hindlimb length of brown anoles would decrease on islands with predators relative to islands without predators.

Third, toepad morphology is related to perch use. Arboreal *Anolis* species tend to have larger toepads with more lamellae (subdigital scales; Glossip and Losos 1997; Macrini et al. 2003; Elstrott and Irschick 2004), conferring improved clinging performance (Elstrott and Irschick 2004). Stuart et al. (2014) found that native *Anolis carolinensis* perched higher in response to the introduction of brown anoles in Florida, and these populations evolved hindlimb toepads with more lamellae after 20 generations. Again, given the changes in perch height observed in our experiment, we expected brown anoles to evolve a greater number of hindlimb lamellae on islands with predators relative to controls.

For all three of these traits, the possible effects of interspecific competitors are less clear: most research on how competition leads to natural selection on brown anole traits has examined intraspecific competition (e.g., Calsbeek 2009; Calsbeek and Cox 2010; Calsbeek et al. 2010). However, given the increased competitive overlap on islands with all three species and the more pronounced effects on habitat use, diet, and population size on those islands (Pringle et al. 2019), we predicted that the presence of competitors would positively interact with the presence of predators, as intensified competition might lead to stronger selection on islands with all three species.

We measured a number of other traits associated with ecomorph type, but for which we did not have clear a priori predictions about how predators and competitors would affect phenotypic evolution. This was also the case at the genetic level, because little is known about the genetic architecture of most traits that define *Anolis* ecomorphs. For this reason, we tested whether predators and competitors drove predictable evolution by quantifying the consistency in multivariate phenotypic changes and by testing for consistent allele frequency changes across replicate populations. This is similar to the retrospective approach discussed above, with the advantage that it does not require assumptions about ancestral states (Bolnick et al. 2018).

In sum, our study examines how predation and competition influence the predictability of evolution in multiple ways and across multiple scales. First, we consider prospective predictability: do the data support our a priori predictions about the evolution of body size, relative hindlimb

length, and toepad lamellae? Next, we quantify whether predation, competition, and their interaction had consistent, and thus retroactively predictable, effects on the trajectories of multivariate phenotype. We ask similar questions at the genetic level, examining genome-wide changes in allele frequency for signals of parallelism. Finally, we consider these phenotypic and genetic changes jointly, across sexes, and in relation to data on population density to evaluate what forces modulate the predictability of evolutionary change in this system.

Material and Methods

Description of Field Experiment

Our work was conducted within a large-scale, long-term field experiment in the Bahamas. Established in 2011, this experiment uses island-level manipulations to test the independent and interactive impacts of interspecific competitors (green anoles) and predators (curly-tailed lizards) on brown anoles. We briefly describe the experimental design here; additional details are provided in the supplemental PDF and Pringle et al. (2019).

On 16 islands (average vegetated area, 1,634 m²; table S1) that have been continuously inhabited by brown anoles for decades, we translocated 10 or 11 green anoles and five to seven curly-tailed lizards from nearby islands in a 2 × 2 factorial design, creating four treatments that were randomly assigned to islands: unmanipulated control (CON) islands with no introduced species, competitor (COMP) islands with added green anoles, predator (PRED) islands with added curly-tailed lizards, and islands with all three species (ALL).

Each year from 2011 to 2016, we conducted a census in late April or early May to estimate the population sizes for each species of lizard. In 2011 this census was performed for brown anoles only and was completed immediately before we began our experimental translocations. In each survey, three to six researchers searched islands for all lizard species over three consecutive days, using a mark-resight procedure developed for Caribbean *Anolis* to estimate population size (Heckel and Roughgarden 1979; Pringle et al. 2015, 2019; see the supplemental PDF). We also recorded data on habitat use, including perch height and diameter, for each lizard we observed.

Phenotypic and Genetic Sampling

We collected phenotypic and genetic data from brown anoles at two time points. In May 2011, before the introduction of the green anoles and curly-tailed lizards, we sampled brown anoles from all 16 experimental islands.

In January 2016, after roughly five generations, we collected posttreatment samples from all 16 experimental islands. During sampling, we attempted to capture lizards as we encountered them on each island, although many individuals evaded capture. We used a portable X-ray to capture full-body X-ray images for measuring skeletal traits, scanned the undersides of lizards to count lamellae, and collected a small piece of tail tip as tissue for genetic analysis. For most islands, we phenotyped all captured lizards, but on some islands we collected additional tissue samples from brown anoles in the field without collecting phenotypic data. After sampling, all lizards were returned to the location where they were captured and marked with a small dot of water-soluble nail polish to prevent recapture.

Skeletal and Toepad Measurements

We measured 15 traits that are associated with ecomorph type: SVL (a measure of body size), head width and length, pectoral and pelvic width, forelimb length (total and separated into humerus, ulna, and toe length), hindlimb length (total and separated into femur, tibia, and toe length), and the number of lamellae on the toepad of the third toe of the fore foot and the fourth toe of the hind foot (fig. S3). We are unaware of estimates of heritability for toe length in *Anolis*, but all other skeletal traits we measured have been found to be heritable, sometimes highly so, in brown anoles (Calsbeek and Smith 2007; Calsbeek and Bonneaud 2008; McGlothlin et al. 2019, 2022). To our knowledge, the heritability of lamella number has not been quantified, but common-garden experiments also suggest a genetic basis for this trait (Thorpe et al. 2005; Stuart et al. 2014).

All skeletal measurements were made by a single person without knowledge of experimental treatment. Skeletal traits were measured from X-ray images in ImageJ version 1.49v (Rasband 1997–2018; Schneider et al. 2012) using the ObjectJ plug-in (Vischer and Nastase 2015). The number of toepad lamellae was independently counted twice by two researchers without knowledge of experimental treatment. We counted the number of lamellae directly from scans.

We measured a subset of X-rays and scans multiple times to estimate repeatability (the intraclass correlation). Repeatability was high (>0.95) for all skeletal traits and slightly lower (>0.85) for counts of toepad lamellae (table S2). For X-rays and scans measured multiple times, we took the mean of the repeated measurements. For toepad lamellae, we averaged counts of lamellae across measurers. For bilateral traits, we measured both the left and right sides whenever possible and used the mean of the two sides, unless only one side could be measured. After measuring all of the lizards, we excluded juvenile lizards

from further phenotypic analysis (see the supplemental PDF). After removing juveniles, we analyzed phenotypes from 967 adult brown anoles (mean $N = 15.1$/island/sex/year, range $= 3$–35; table S3).

Univariate Phenotypic Change

We analyzed univariate changes in SVL, relative hindlimb length, and number of lamellae on the hind foot (i.e., the traits for which we had a priori predictions) using linear mixed models (LMMs). We included presence/absence of the predator, presence/absence of the competitor, sampling year (as a factor), and all possible interactions of these three variables as fixed effects and included island as a random intercept. We ran separate LMMs for males and females for each trait. Because hindlimb length is strongly correlated with body size (Beuttell and Losos 1999), we included SVL as a covariate in the LMMs analyzing relative hindlimb length and used partial pooling to allow the effect of SVL to vary across islands. Lamella number is thought to be fixed at birth and independent of size (Hecht 1952) and is not significantly associated with size in our data (see table S5). Thus, we did not correct for body size in the LMM for lamella number. We tested for significant trait changes due to predators, competitors, and their interaction by examining the two- and three-way interactions of these effects with year. Although we focus our univariate analysis on these three traits for which we made a priori predictions, we also ran univariate models for all other traits (see the supplemental PDF).

We implemented LMMs in a Bayesian framework with the R package brms (Bürkner 2017, 2018). We used the Student's t distribution as our likelihood function for the residual error in the model. This is similar to a Gaussian likelihood but more robust to possible outliers (Gelman and Hill 2007). We used weakly informative priors and fit four independent chains with 1,000 iterations of warm-up and 1,000 iterations of sampling each, using the brms default control parameters (see the supplemental PDF for full details of priors and models). We assessed statistical significance for the coefficients (βs) of the interactions of interest by determining whether the 95% highest posterior density interval (HPDI) for that parameter contained zero.

We performed a post hoc power analysis with the R package simr to determine minimum effects that we could detect for our focal traits with 80% power (Green and MacLeod 2016; see fig. S4). We compared these with the amount of phenotypic change we might expect for SVL and relative hindlimb length given previously estimated strengths of selection and heritabilities (Losos et al. 2004, 2006; Calsbeek and Smith 2007; Calsbeek and Bonneaud 2008; McGlothlin et al. 2019).

Multivariate Phenotypic Change

We examined whether predators and competitors affected the predictability of multivariate phenotypic change by quantifying metrics of parallelism and convergence across our experimental populations. Following Bolnick et al. (2018), we define parallelism and convergence geometrically, in terms of the positions and trajectories of populations in multivariate morphospace. Under this framework, "parallelism" refers to the similarity of population trajectories in morphospace and can be broken down into two quantifiable components: parallelism in the magnitude of evolution (i.e., are trajectories of similar length?) and parallelism in direction of evolution (i.e., is the angle between population trajectories small?). The geometric definition of convergence is closely related: populations that are getting closer in multivariate morphospace are converging, and those getting further apart are diverging. Biologically, these quantities describe how the average phenotypes of populations change through time: magnitude considers the absolute value of the change, direction considers the sign of the change, and convergence considers whether differences in mean phenotype are increasing or decreasing. These definitions are useful because they decompose the related (and sometimes conflated) concepts of parallel evolution and convergent evolution into three independent, continuously varying metrics that quantify the degree to which populations are evolving in similar ways (Bolnick et al. 2018).

For our analysis, we treated each island as a population and calculated each population's position in morphospace at each of our two sampling points (in 2011 and in 2016). We considered 13 phenotypes, and thus a 13-dimensional morphospace, including the two counts of lamellae and all skeletal traits except total forelimb length and total hindlimb length (which we excluded to avoid redundancy with the component limb traits). To remove the effect of body size on skeletal traits other than SVL, we calculated residuals from sex-specific LMMs and used the modeling approach described above for relative hindlimb length but without the year, predator, and competitor effects. Then, within each sex, we standardized all trait values to have a mean of 0 and a standard deviation of 1, to remove possible scaling or sex ratio effects when averaging across traits. Finally, we determined each population's position in morphospace by taking the population average of each phenotypic trait at that time point.

We then quantified parallelism in the magnitude and direction of evolution using the phenotypic change vector (PCV) analysis developed by Collyer and Adams (2007) and Adams and Collyer (2009). The PCV is the vector between a population's position in morphospace across two time points. Importantly, the PCV quantifies both the

magnitude of change (the length of the vector, D) and the direction of change relative to another vector (θ). We calculated PCVs for each of our populations (islands) and then calculated the length of each PCV (D), all pairwise differences in length across islands (ΔD), and all pairwise differences in direction (θ). To quantify the degree of convergence or divergence, we calculated all possible pairwise Euclidean distances between each population at each of our two time points and took the difference of these pairwise distances. We then scaled this change in Euclidean distance by the starting distance so that all pairwise comparisons had the same maximum amount of convergence (-1). We call this scaled degree of convergence (when negative) or divergence (when positive) Ω. We used linear models to test whether competitors, predators, and their interaction had significant effects on the average D, ΔD, θ, and Ω within a treatment using linear models. We also used Welch's t-tests to examine whether the average pairwise ΔD, average pairwise θ, or average pairwise Ω varied depending on whether the islands being compared were in the same or different treatments.

Library Preparation and Sequencing

We extracted DNA from lizard tail tips using a standard phenol-chloroform-based method and cleaned the genomic DNA using NucleoSpin gDNA cleanup kits (Takara Bio). We next prepared sequencing libraries for double-digest restriction enzyme–associated DNA sequencing (Peterson et al. 2012). We used sequencing adaptors with 48 individual barcodes and 16 library barcodes, allowing us to multiplex up to 768 samples per lane of sequencing. In the P2 adaptor, we included a small degenerate base region (Schweyen et al. 2014), slightly modified from the design of Vendrami et al. (2017), that was designed to allow us to bioinformatically filter polymerase chain reaction (PCR) duplicates. Full details of library preparation are provided in the supplemental PDF. In total, we assembled 28 libraries for sequencing by Génome Québec (Montreal). We sequenced these libraries across two lanes (16 libraries in one lane, 12 in the other) of an Illumina NovaSeq 6000 S4 flow cell using paired-end 150-bp reads.

Bioinformatic Pipeline

We briefly summarize our bioinformatic pipeline here; see the supplemental PDF for full details of programs and options. We demultiplexed raw reads and filtered out PCR duplicates using the process_radtags and clone_filter tools in Stacks version 1.46 (Catchen et al. 2011, 2013). We used fastx_trimmer and cutadapt version 2.1 to remove extra bases and adaptor contamination and then mapped reads to the preliminary *Anolis sagrei* genome (A. Geneva,

personal communication) using the default settings of the BWA-MEM algorithm in bwa version 0.7.15, keeping only alignments with a mapping quality of at least 20 (Li and Durbin 2009; Martin 2011; Li 2013). We then used ANGSD version 0.918 and PCAngsd version 0.98 to perform a principal components analysis of genome-wide population structure across our experiment (Korneliussen et al. 2014; Meisner and Albrechtsen 2018). From this point, our analysis diverged according to the two methods we used to quantify genetic parallelism.

Our first approach measures correlations between changes in allele frequency (see below). We calculated allele frequencies for each island at each time point directly from the mapped reads using ANGSD version 0.918 (Korneliussen et al. 2014). Our second analysis uses the Cochran-Mantel-Haenszel (CMH) test to quantify parallel allele frequency change. We used the multiallelic single-nucleotide polymorphism (SNP) caller in the mpileup/call pipeline of samtools version 1.5 and bcftools version 1.5 to generate a VCF file of called genotypes (Li 2011; Danecek et al. 2014). We used vcftools version 0.1.14 (Danecek et al. 2011) to filter these raw variants. We retained only biallelic SNPs with an experiment-wide minor allele frequency greater than 5%, a minimum genotype quality score of 25, a maximum of 50% missing genotypes across the experiment, a minimum mean depth per sample of 5, and a maximum total depth of 42. After filtering, we retained 85,888 SNPs.

Parallelism of Genetic Change

We first examined parallel genetic change at a genome-wide scale. To do this we estimated allele frequencies for each population at each of the two time points (table S3). We calculated these estimates directly from the sequencing read data using the maximum likelihood model implemented in ANGSD (Kim et al. 2011). This model accounts for some of the special features of low-coverage, next-generation sequencing data, including read misalignment, base-calling errors, variation in read depth, and genotype uncertainty (Kim et al. 2011; Korneliussen et al. 2014). Importantly, this method allows us to accurately make use of read and SNP data that would be discarded in more traditional, quality-filtered approaches.

We removed SNPs with a minor allele frequency less than 5% at either time point and then calculated the change in allele frequency, Δp, for all SNPs within a given island. For each island, we used the changes in allele frequency to estimate the average variance effective population size, N_e, during the experiment (Jorde and Ryman 2007; see the supplemental PDF). Then, for all possible pairwise comparisons of islands, we calculated the correlation (Pearson's r) between Δp for variants shared between

the two populations, which we denote as $r_{\Delta p}$. As with phenotypes, we used linear models and Welch's t-tests to test whether correlations in allele frequency differed significantly across treatments.

We also tested for parallel changes in allele frequency at a per-SNP level using the CMH test. The CMH test is analogous to Fisher's exact test across replicate samples or populations: it is used to test for independence across $2 \times 2 \times K$ contingency tables, where K is the number of replicates and 2×2 refers to allele counts for a biallelic SNP across two groups (Orozco terWengel et al. 2012). This test is popular in evolve-and-resequence studies to test whether allele frequencies at a single locus have changed significantly across replicate populations (Schlotterer et al. 2015). We used PLINK version 1.9 to perform the CMH test on the four islands in each treatment (Chang et al. 2015; Purcell and Chang 2017). That is, we performed the CMH test four times: once as a $2 \times 2 \times 4$ test on the CON islands, once as a $2 \times 2 \times 4$ test on the PRED islands, and so on. At the individual-SNP level, we used the Benjamini-Hochberg method to correct for multiple testing, controlling the false discovery rate at $\alpha = .05$ for each treatment. We also tested for significant parallelism across larger genomic windows (genome-wide and in 100-, 50-, and 10-kb windows) using the harmonic mean P value (HMP) method for combining dependent tests of Wilson (2019). Last, we performed association mapping of our three focal traits (SVL, relative hindlimb length, and hindlimb lamellae) in an attempt to identify candidate SNPs for which we could examine possible parallel allele frequency changes; however, we did not find significantly associated SNPs (see the supplemental PDF for details) and thus did not pursue this analysis further.

Ecological Correlates of Predictability across Scales

We examined how measures of predictability at the phenotypic level and genetic level were correlated across sexes and whether they were related to differences in lizard density across islands. To examine whether sex-specific multivariate phenotypic changes might promote or inhibit population-level parallelism, we repeated our multivariate analysis separately for each sex. Treating the males and females from individual islands as belonging to separate populations, we calculated sex-specific PCVs and sex-specific metrics of parallelism and convergence (D, ΔD, θ, and Ω). For the density analysis, we used data on the yearly density of the competitor and predator species to calculate two pairwise distance matrices of density differences, one for differences in competitor density (Δ_{densCOMP}) and one for differences in predator density (Δ_{densPRED}). Thus, for all pairwise comparisons of islands, we have measures of both population-level and sex-

specific estimates of multivariate phenotypic parallelism and convergence (D, ΔD, θ, and Ω), a measure of genetic parallelism ($r_{\Delta p}$), and measures of the similarity of lizard density (Δ_{densCOMP} and Δ_{densPRED}). We tested for relationships between these measures either by using Mantel tests (when both measures were true distance or dissimilarity matrices) or by testing for significant correlations (when one measure was not a true dissimilarity matrix, as with $r_{\Delta p}$ and Ω).

Additional Software Used

In addition to the programs listed above, we used a variety of other R packages for data processing, analysis, and visualization: assertthat (Wickham 2019a), broom (Robinson et al. 2020), cowplot (Wilke 2020), doParallel (Microsoft and Weston 2020a), dplyr (Wickham et al. 2020), forcats (Wickham 2020a), foreach (Microsoft and Weston 2020b), ggmap (Kahle and Wickham 2013), ggokabeito (Barrett 2021), ggplot2 (Wickham 2016), ggrepel (Slowikowski 2021), ggsn (Santos Baquero 2019), IRanges (Lawrence et al. 2013), kableExtra (Zhu 2020), knitr (Xie 2014), lemon (Edwards 2020), lme4 (Bates et al. 2015), lmerTest (Kuznetsova et al. 2017), readr (Wickham and Hester 2020), renv (Ushey 2020), stringr (Wickham 2019b), tibble (Müller and Wickham 2020), tidyr (Wickham 2020b), and vegan (Oksanen et al. 2020).

Results

Univariate Trait Changes

Introduced predators both reduced brown anole population size and caused brown anoles to use higher perches (see Pringle et al. 2019; fig. 1A), consistent with previous work (Schoener et al. 2002, 2005). However, the effects of predators on phenotypic change did not occur as predicted. While we expected that predators would select for larger body size in female brown anoles, we found that predators had a negative, albeit not statistically significant, effect on female SVL (fig. 1B; posterior mean of $\beta_{\text{pred:year}} = -0.69$, 95% HPDI $= -1.48$ to 0.15) and a significant negative effect on male SVL (fig. 1B; posterior mean of $\beta_{\text{pred:year}} = -1.76$, 95% HPDI $= -3.02$ to -0.64). Likewise, our predictions that brown anoles would evolve shorter hindlimbs and a greater number of toepad lamellae were not supported. The effect of predators on female hindlimb length was near zero (fig. 1C; $\beta_{\text{pred:year}} = 0$, 95% HPDI $= -0.37$ to 0.39). In males, the effect of predators was positive but not statistically significant (fig. 1C; $\beta_{\text{pred:year}} = 0.39$, 95% HPDI $= -0.13$ to 0.82). Changes in lamella number were nonsignificant for females, while for males predators had a significantly
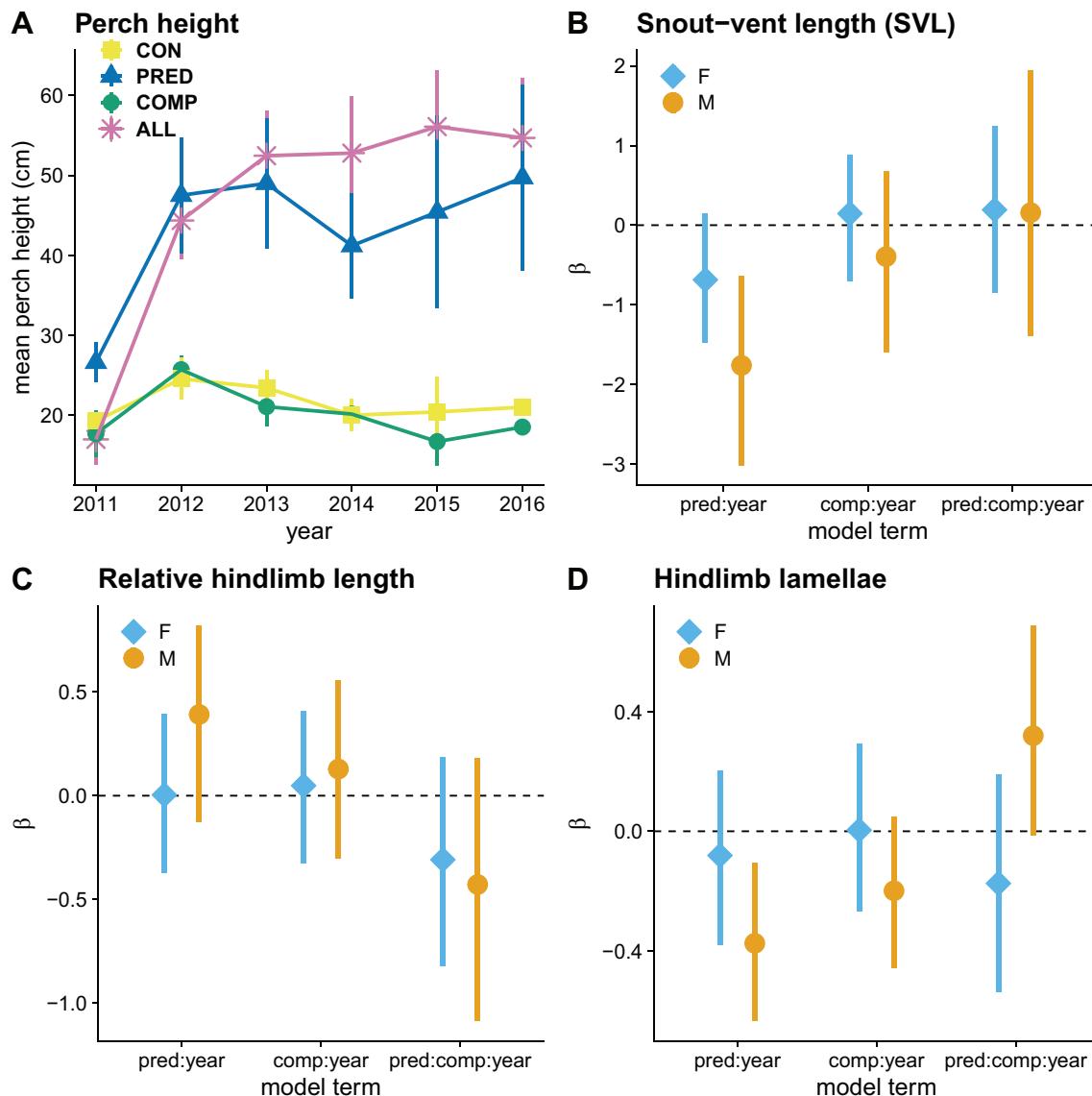
Figure 1: Outcomes of our predictions about the effects of predators and competitors on univariate phenotypes. *A,* Change in perch height through time (mean ± 1 SEM of island-wide averages for each treatment in each year). *B–D,* Effects of predators, competitors, and their interaction on change in snout-vent length (*B*), relative hindlimb length (*C*), and number of hindlimb lamellae (*D*). Each plot shows the posterior mean ± 95% highest posterior density interval of the coefficient (*β*) for the given interaction term in the linear mixed model. Coefficients quantify the effect, relative to controls, of the given interaction on the change in mean of the trait of interest. CON = unmanipulated control islands with no introduced species; COMP = competitor islands with added green anoles; PRED = predator islands with added curly-tailed lizards; ALL = islands with all three species.

negative, not positive as expected, effect on lamella number (fig. 1D; $\beta_{\text{pred:year}} = -0.37$, 95% HPDI = -0.64 to -0.11). For our three focal traits, the direct effects of interspecific competitors were nonsignificant, as were the interactive effects of predators and competitors (fig. 1; table S4). Considering all of the traits, the only significant effects of competitors were negative effects on male head width and pectoral width, and the only significant predator-

competitor interaction was a negative interaction for female femur length (table S6; see the supplemental PDF for discussion of these traits).

Our post hoc power analysis revealed that we had reasonable power to detect the phenotypic changes we expected, given the high heritabilities and strong selection gradients previously reported in brown anoles (Losos et al. 2004, 2006; Calsbeek and Smith 2007; Calsbeek

and Bonneaud 2008; McGlothlin et al. 2019; table S7; see the supplementary PDF for details). Thus, our nonsignificant results are not simply due to a lack of power but rather because univariate phenotypic changes did not occur as expected.

Multivariate Phenotypic Parallelism and Convergence

We found little evidence that introduced predators or competitors drove predictable changes in multivariate phenotype. Competitors and predators did not have significant effects on the magnitude of phenotypic change for each island, D (all $P > .05$; table S8). Considering only pairwise comparisons of islands in the same treatment (six pairwise comparisons within each treatment, 24 total), predators and competitors had little effect on differences in magnitude, ΔD (fig. 2A; all $P > .05$; table S9). In terms of the direction of phenotypic change, introduced competitors decreased parallelism (i.e., increased θ; fig. 2B; $\beta = 26.1$, $P = .002$), as did introduced predators ($\beta = 17.72$, $P = 0$; table S10). There was a significant predator-by-competitor interaction, such that evolution in the ALL treatment had a similar level of parallelism in direction as CON islands (fig. 2B; $\beta = -46.2$, $P = 0$). Overall, these within-treatment results suggest that phenotypic changes in response to predators and competitors on treatment islands were less parallel than the phenotypic changes on CON islands.

We further examined parallelism at the treatment level by considering all 120 possible pairwise comparisons between the 16 islands, testing for greater parallelism (smaller θ or smaller ΔD) when comparisons were between islands in the same treatment. We found no significant differences (θ: $t = 0.43$, $P = .67$; ΔD: $t = -0.41$, $P = .68$), and the distribution of θ and ΔD was similar for comparisons within and between treatments (fig. 2D, 2E).

Phenotypic convergence, Ω, differed significantly across treatments, but CON islands—not manipulated ones—showed the greatest convergence among pairwise comparisons within treatments (fig. 2C; table S11). Likewise, when considering all pairwise comparisons, the amount of convergence was not greater for islands from the same treatment (fig. 2F; $t = 1.07$, $P = .29$). Interestingly, there was a general trend toward phenotypic convergence in our experiment: of 120 pairwise comparisons between islands, 71 converged (binomial test, $P = .055$).

Parallelism of Genetic Changes

As at the phenotypic level, we found little evidence of parallel genetic change within treatments. Our principal component analysis found substantial initial population structure: the first six principal components explained roughly 35% of genome-wide variation and clustered individuals by island, with population structure changing little during the course of the experiment (fig. S5). Estimates of effective population sizes were generally small (range $N_e = 43$–145; table S13). This genetic distinctiveness led to wide variation in the number of shared SNPs for calculating $r_{\Delta p}$ (mean = 6,792 sites, range = 2,204–14,801). Across all pairwise comparisons, $r_{\Delta p}$ ranged from mild positive correlation ($r_{\Delta p} = 0.303$) to mild negative correlation ($r_{\Delta p} = -0.296$), but there were few clear patterns related to our experimental manipulations (fig. 3).

Considering only comparisons between islands of the same treatment (the dark-outlined triangles along the diagonal of fig. 3), correlations were weakest among the ALL islands, most strongly positive among the COMP islands, and most strongly negative among the PRED islands. Only predators had a statistically significant effect on the amount of correlation in allele frequency change among islands in the same treatment, tending to decrease the amount of genetic parallelism (table S12; $\beta = -0.11$, $P = .04$). Considering all pairwise tests, $r_{\Delta p}$ did not differ according to whether the islands were in the same or different treatments ($t = -0.72$, $P = .47$). This result again suggests that introduced predators and competitors did not promote parallel evolution at the molecular level in our experiment.

At the level of individual SNPs, no sites showed significant evidence of parallel changes in the CMH test after correcting P values for multiple testing. Using the HMP approach to test for signs of parallelism across larger genomic regions, we found a genome-wide signature of parallelism for the COMP islands (HMP = .04), although we could not narrow down this signal (all HMP > .05 across smaller windows). The genome-wide and window-level tests across the other treatments showed no significant parallelism (all HMP > .05).

Ecological Correlates of Predictability across Scales

Compared with each other, our metrics of phenotypic parallelism (ΔD, θ, and Ω) were generally not significantly correlated: ΔD and Ω were moderately negatively correlated ($r = -0.24$, $P = .01$). All other pairwise correlations between metrics of phenotypic parallelism (ΔD, θ, and Ω) were not significant ($P > .05$). The pairwise degree of genetic parallelism, $r_{\Delta p}$, was weakly and nonsignificantly correlated with all measures of pairwise phenotypic parallelism ($|r| < 0.15$, all $P > .05$). Considering the sex-specific PCVs, males and females showed little similarity in the direction of phenotypic change. The average angle between the male and female PCVs from the same island was large ($\bar{\theta} = 52.6°$) and not significantly different from the average
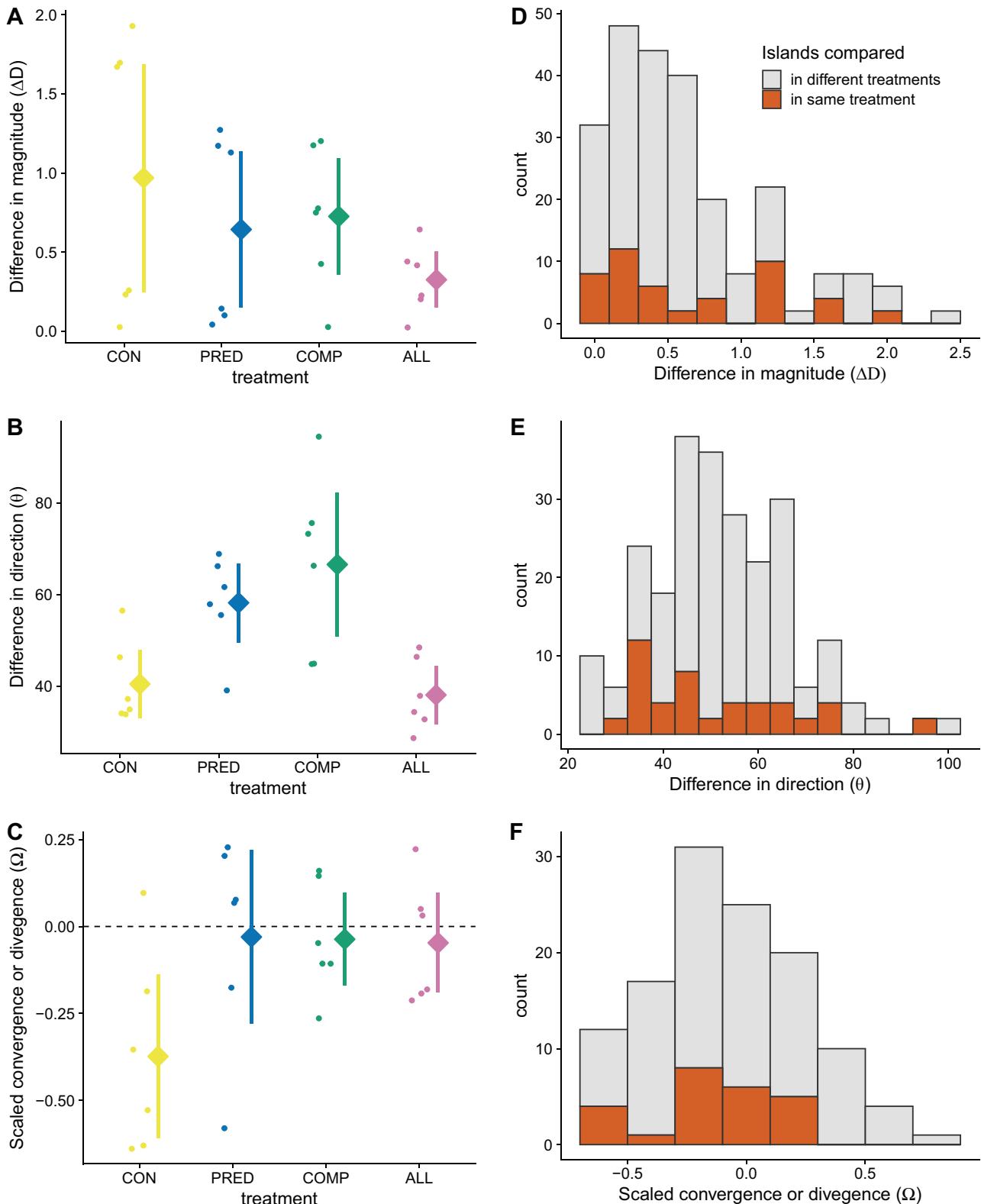
Figure 2: Parallelism and convergence of multivariate phenotypic change. *A–C*, Pairwise differences in magnitude (ΔD; *A*), differences in direction (θ; *B*), and degree of scaled convergence or divergence (Ω; *C*) in multivariate phenotypes for the six possible comparisons within each treatment. For ΔD and θ, larger values indicate less parallelism. For Ω, values less than zero indicate convergence, values greater than zero indicate divergence, and -1 is the maximum amount of convergence. Small points show individual comparisons, diamonds show the treatment mean ± 2 SE. *D–F*, Histograms of ΔD, θ, and Ω for all pairwise comparisons, by comparison type (within treatment or between treatments). CON = unmanipulated control islands with no introduced species; COMP = competitor islands with added green anoles; PRED = predator islands with added curly-tailed lizards; ALL = islands with all three species.
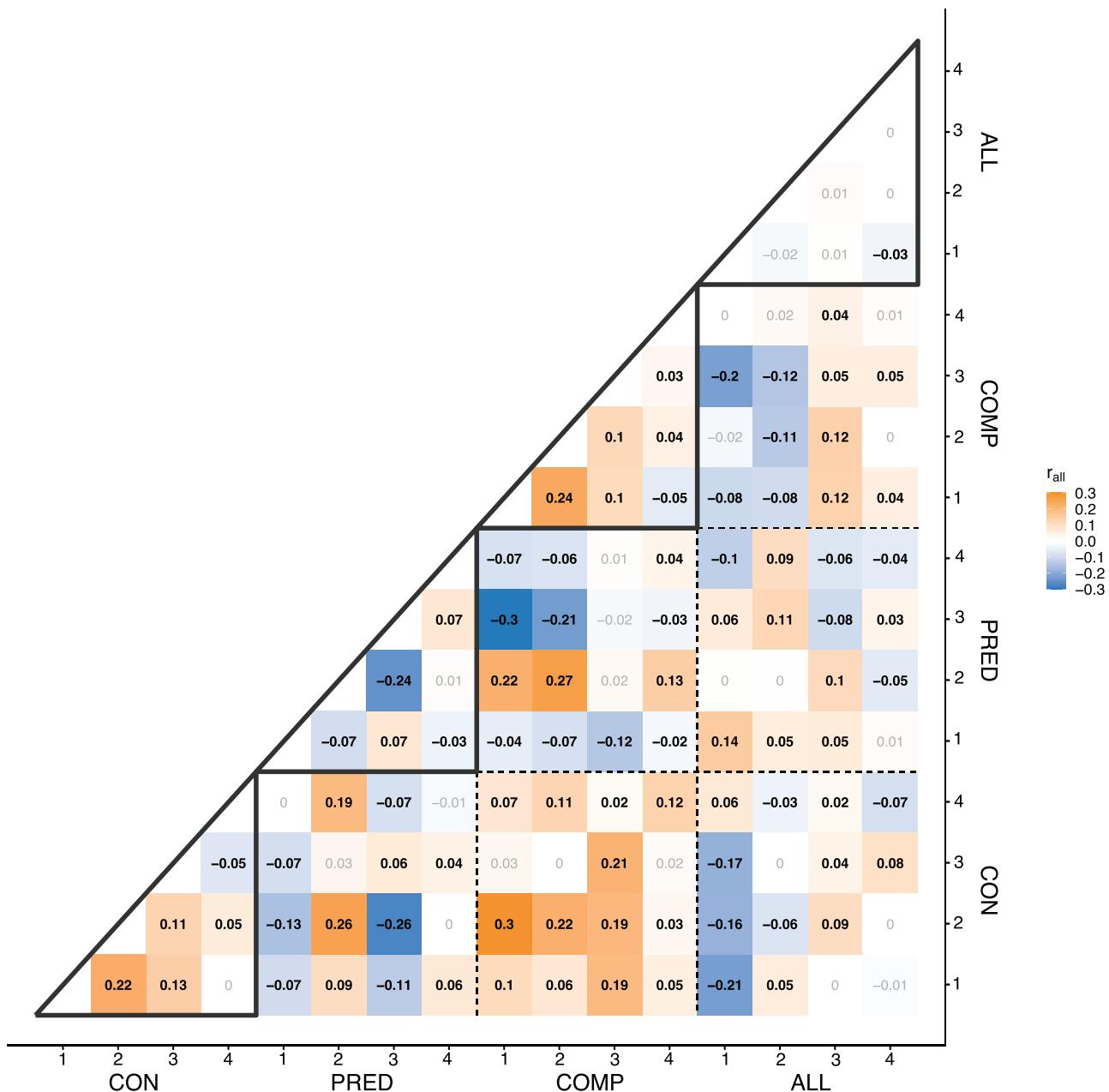
Figure 3: Heat map of the correlation in allele frequency change for all single-nucleotide polymorphisms (SNPs) across islands. Shaded colors indicate the strength and direction of correlation for each pairwise comparison. The Pearson correlation coefficient for all SNPs, $r_{\Delta p}$, is shown in each cell, printed in black if statistically significant and in gray if nonsignificant. Islands are grouped by treatment (shown with dotted lines), and the within-island comparisons are in the solid triangles along the diagonal. CON = unmanipulated control islands with no introduced species; COMP = competitor islands with added green anoles; PRED = predator islands with added curly-tailed lizards; ALL = islands with all three species.

difference in direction across all pairwise comparisons between populations ($t = 0.11$, $P = .913$). In other words, the amount of parallelism in direction for males and females from the same population was similar to the amount of parallelism (or lack thereof) at the population level. The lengths of sex-specific PCVs for each population were moderately

correlated (D: $r = 0.48$, $P = .059$), such that females and males within the same population had a similar amount of phenotypic change, albeit in different directions. We found that differences in the direction of phenotypic change were associated with differences in the density of the competitor, the green anole, such that islands with more similar

competitor densities evolved in more similar directions (fig. 4A; $r = 0.4$, $P = .02$). There was also a positive correlation between Ω and differences in density of curly-tailed lizards, such that islands with more similar predator densities showed more convergence (smaller Ω, $r = 0.22$, $P = .015$; fig. 4B). All other metrics of phenotypic and genetic parallelism were uncorrelated to predator and competitor density ($P > .05$).

Discussion

This study reports the evolutionary outcomes of a field experiment testing the responses of brown anole lizards to introduced predators and competitors. Although these novel species interactions had strong and consistent ecological effects on habitat use, diet, and population size of brown anoles (Pringle et al. 2019), they did not necessarily lead to predictable phenotypic and genotypic changes. Our a priori predictions about how three traits would evolve—informed not just by the ecological dynamics in our own experiment but also by a large body of previous work on selective dynamics (Losos et al. 2004, 2006; Kolbe

et al. 2012; Stuart et al. 2014; Lapiedra et al. 2018) and functional morphology (Losos and Sinervo 1989; Losos 1990a, 1990b; Losos and Irschick 1996) in brown anoles and related species—were not supported. At the treatment level, we found scant evidence that predators and competitors drove consistent changes in multivariate phenotype or genotype. There was little correlation between different metrics of parallelism or across biological scales. This general lack of parallelism was partially explained by sex-specific responses and by island-level differences in species' densities within treatments. Taken together, what do these results tell us about the predictability of evolution in response to novel species interactions?

Predicting Univariate Trait Changes

We first consider our predictions about evolution in body size, relative hindlimb length, and lamella number. Contrary to expectations, introduced predators did not lead to larger female brown anoles. Instead, predators had an overall negative effect on body size, especially in males. One can imagine selective reasons for this, drawn from
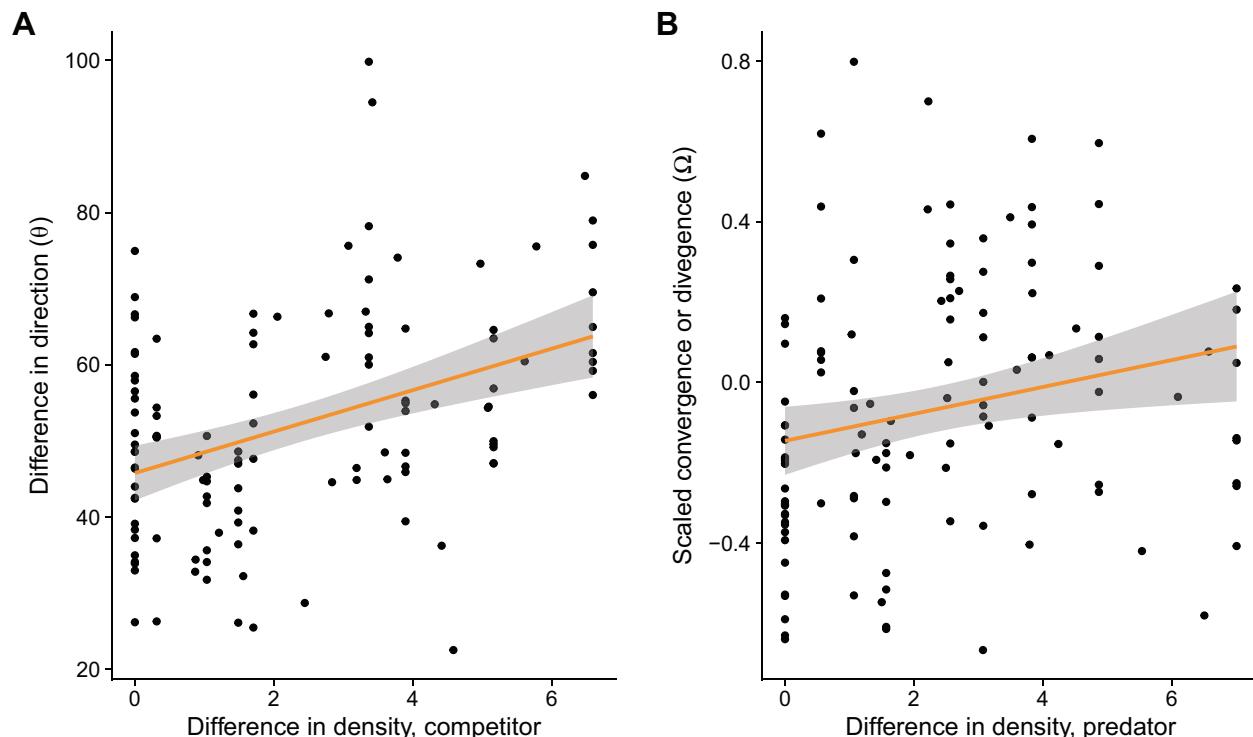

Figure 4: Relationships between metrics of phenotypic parallelism and differences in density of lizard species. *A*, Differences in density of the competitor species *Anolis smaragdinus* ($\Delta_{densCOMP}$) are related to pairwise differences in the direction of phenotypic change, θ. *B*, Differences in the density of the predator species *Leiocephalus carinatus* ($\Delta_{densPRED}$) are related to pairwise degree of convergence, Ω. Each point is a single pairwise comparison between islands. Lines show the linear trend between variables, and ribbons show the 95% confidence intervals of the trend, but these are for visualization purposes only: our statistical analysis examined correlations.

general theories of predation. Smaller lizards may experience lower mortality from a ground-dwelling predator if they are less conspicuous or can more easily shift to arboreal habitat or if larger prey are more profitable for predators (Pyke et al. 1977). In light of past *Anolis* research, however, these explanations seem unlikely. Previous within-generation studies of brown anoles have reported positive directional selection or no selection by predators on body size (Losos et al. 2004; Calsbeek 2009) rather than negative directional selection, while the size distributions of all brown anoles on islands with and without terrestrial predators suggest that small lizards likely face higher predation rates (Schoener et al. 2002). Two alternative ecological possibilities seem more plausible. First, although body size is highly heritable (Calsbeek and Smith 2007; Calsbeek and Bonneaud 2008), it is also related to food intake: lizards that eat less do not grow as large (Bonneaud et al. 2015). Many animals modify their foraging activity when predation risk is high (Sih 1980; Lima and Dill 1990), and brown anoles have been shown to reduce activity levels, especially on the ground, in the presence of curly-tailed lizards (Lopez-Darias et al. 2012; Lapiedra et al. 2018). It therefore seems probable that brown anoles on islands with predators forage less often and/or less productively, which could negatively affect mean body size in the population. This hypothesis is also consistent with data showing that brown anole diets in our experiment shifted dramatically in the PRED treatment, notably via reduced consumption of large, ground-dwelling cockroaches (Pringle et al. 2019). A second mutually compatible possibility is that predators alter the age structure of brown anole populations. Most reptiles have indeterminate growth, so a decrease in average body size could indicate a decrease in the proportion of old lizards in a population. The fact that the negative effect of predators on body size was strongest on PRED islands and not ALL islands (the positive predator-competitor interaction) could be explained by the fact that predator densities tended to be higher in the former treatment (Pringle et al. 2019). Phenotypic plasticity and changes in demography could together explain why within-generation selection, if it was occurring in our experiment, did not lead to the expected changes in mean phenotype over multiple generations.

We also predicted that the higher perch heights of brown anoles in the presence of predators would result in relatively shorter hindlimbs and more lamellae on the hindlimb toepads. If anything, the opposite occurred: predators had a positive (albeit nonsignificant) effect on relative hindlimb length in males and had negative effects on the number of lamellae in both sexes. These results are surprising, particularly for hindlimb length: comparative morphology within the *Anolis* genus (Losos and Irschick 1996; Beuttell and Losos 1999), performance trials with lizards (Losos and Sinervo 1989; Losos 1990*a*, 1990*b*; Losos and Irschick 1996), and field experiments (Losos et al. 1997, 2004, 2006; Kolbe et al. 2012) all point to relative hindlimb length as a target of natural selection related to habitat use. Another recent study, published after we began our experiment, also found a lack of hindlimb evolution in response to introduced *Leiocephalus carinatus*. Schoener et al. (2017) quantified changes in relative hindlimb length over 7 years and found no significant effects of predators, which they attributed to disturbances from multiple hurricanes. Hurricanes have not detectably impacted our study, however, suggesting the need for some other explanation.

A closer look at our census data on perch height and perch diameter reveals one possibility: although brown anoles moved to higher perches on islands with predators, they were rarely found on very narrow perches. Laboratory trials show that it is not perch height per se that impedes the performance of long-legged lizards in arboreal settings but rather perch diameter, which typically scales negatively with height from rocks and tree trunks up to branch tips (Losos and Sinervo 1989; Losos and Irschick 1996). In those trials, sprint performance markedly decreased on 0.6-cm perches (Losos and Irschick 1996). In our experiment, the average perch diameter used by brown anoles decreased on islands with predators, largely driven by a decrease in the number of brown anoles found on the ground (Pringle et al. 2019). However, even on islands with predators, brown anoles rarely perched on the very narrow perches that would impede performance. In 2011, before the introduction of predators, brown anoles on PRED and ALL islands were found on perches <1 cm in diameter in just 3.6% of observations. After introducing predators, that rate increased but remained low (5.5% of observations from 2012 to 2015). These patterns of perch behavior are consistent with results from laboratory trials, which show that brown anoles prefer to perch on broad surfaces when substrates with multiple diameters are available (Kolbe et al. 2021).

Our experimental islands were chosen, in part, because they were large enough to support the tall vegetation necessary for green anole populations to persist (Losos and Spiller 1999). This is in contrast to the smaller islands with shorter vegetation that have been used in the previous field experiments that informed our predictions (e.g., 104–324 m² in Losos et al. 2004; 487–3,320 m² in our study; table S1). The taller vegetation on our islands seems to have allowed brown anoles to move higher while still finding wide perches that did not impede their performance and thus select for shorter legs. Besides taller vegetation, larger islands may differ in other ways that could attenuate predator-induced selection: for example, increased environmental heterogeneity on large islands may lead to more predator-free refuges (Schoener et al.

2016). These hypotheses would not obviously explain the negative effect of predators on number of lamellae. Competitors also had a (marginal) negative effect on lamella number in males, but the predator-by-competitor interaction was strong and positive, making the results on lamellae particularly difficult to interpret. Given the links between lamella number and perch use, we have no coherent explanation for this puzzling combination of predator and competitor effects on behavior and morphology.

These patterns underscore the difficulty of forecasting evolution arising from novel species interactions. For body size, within-generation estimates of selection were not informative for predicting multigenerational changes. For hindlimb length and number of lamellae, a change in ecological context compared with the experiments on which we based our predictions (namely, our use of larger islands with taller vegetation) led to unexpected patterns of phenotypic change. In retrospect, our extensive behavioral and ecological data suggest a partial explanation for at least some of these results: the effects of predators are context dependent, and on islands with large vegetation, brown anoles can find perches that are high enough to avoid ground-dwelling predators yet sufficiently wide to not impose strong selection for more arboreal traits.

Predicting Multivariate Trait Changes

The observed patterns of multivariate phenotypic change reveal a similar story, in which our supporting ecological data suggest at least a partial explanation for the observed lack of parallelism. At the treatment level, introduced predators and competitors did not promote parallelism in the magnitude of multivariate phenotypic change and inhibited parallelism in the direction of change (fig. 2). The lack of parallelism could be the result of our experimental manipulations not imposing similar selection pressures across populations in the same treatment. On the other hand, similar selection pressures could lead to nonparallel trajectories of phenotypic change if populations start at different points in morphospace. For example, consider stabilizing selection on a single phenotype: populations starting with average phenotypes above and below the optimal trait value will evolve in different directions to reach the same adaptive peak. In that case, populations in the same treatment would tend to converge in multivariate morphospace. Although there was an experiment-wide trend toward convergence, this was unrelated to the presence of predators and competitors, which tended to decrease the degree of phenotypic convergence. Our analyses suggest two factors that might explain this overall lack of predictability.

First, we found clear differences in phenotypic change between sexes. This was most pronounced in the direction of phenotypic change, where the average angle between male and female trajectories within the same island ($\bar{\theta} = 52.6°$) was not significantly different from the average angle between population-level trajectories across islands ($\bar{\theta} = 50.9°$). Brown anoles are sexually dimorphic for many traits, including body size and shape (Butler and Losos 2002), dewlap size and color (Cox et al. 2017), perch behavior (fig. S2; Butler and Losos 2002), and defensive behaviors (Reedy et al. 2017). Previous comparative studies across the genus have found that dimorphism is likely caused by males and females adapting to independent niches and microhabitats, though sexual selection may also play a role (Schoener 1967; Schoener and Schoener 1980; Butler and Losos 2002; Butler 2007). Field experiments have also found differing patterns of natural selection on body size and hindlimb length for males and females in brown anoles (Calsbeek 2009; Cox and Calsbeek 2010). Our finding of sex-specific phenotypic responses to introduced predators and competitors reinforces the idea that selection due to habitat use is largely sex-specific in brown anoles. This complicates the prediction of population-level responses to novel species interactions, which will depend on the between-sex genetic correlation for the trait(s) under selection. Between-sex correlation for body size varies across some populations of brown anoles (McGlothlin et al. 2019), and it seems plausible that correlations among the shape traits we examine here could also vary across our experimental populations.

Second, we found that some differences in multivariate change were related to differences in the densities of the competitor and predator species: islands with more similar competitor densities had smaller differences in direction, while islands with more similar predator densities had greater convergence (fig. 4). These results further demonstrate the utility of ecological and behavioral data for parsing complex interactions, as they provide some explanation for why our treatment-level analysis, which did not consider density, found that competitors and predators tended to decrease parallelism. Categorical analysis is common in ecology and evolution: populations from discrete habitat types are often analyzed in this way (e.g., light and dark soils [Barrett et al. 2019] or high-predation vs. low-predation streams [Gotanda and Hendry 2014]), even when the underlying variables of interest (e.g., soil color, number of predators) are quantitative. Although a categorical, treatment-level analysis is appropriate given our replicated, factorial experimental design, discrete categories can also obscure important variation within groups. This might be particularly true in a field experiment like ours, in which the initial manipulation (absolute number of colonizing competitors) was constant but the amount of competition depended on a variety of factors, including island size and the growth rate of green anole populations.

In both our univariate and multivariate analyses of phenotypic change, predators and competitors rarely had the predicted significant interactive effects: only the direction of phenotypic change, θ, had a significant interaction term. This may partially reflect the decreased statistical power for interactions versus main effects (fig. S4), as in some cases the estimated interaction was large but uncertain. Overall, though, the interactive effect of predators and competitors on ecological outcomes in our experiment (e.g., decreased brown anole populations sizes, increased interspecific competition, and extirpation of green anole populations; Pringle et al. 2019) may not lead to strong interactions in evolutionary outcomes.

Taken together, these results suggest that predicting phenotypic change due to novel species interactions will continue to be challenging, requiring extensive ecological and behavioral data to account for the context dependence of species interactions and natural selection. The patterns of change that we see on our unmanipulated CON islands reinforce this. With our factorial experimental design, our conclusions about the effects of predators and competitors are drawn in comparison to our CON islands, which themselves displayed phenotypic changes and nonparallelism (fig. 2). This is not unexpected for a large-scale field experiment: our CON islands are not "controls" in the sense that we eliminated all other possible sources of variance, but instead they are simply unmanipulated reference points. This makes our experiment ecologically realistic: our CON islands provide context on the wide range of ecological and evolutionary forces that are at work outside those we manipulated. Of course, these many forces may lead to less tidy experimental outcomes, but attempts to control all of these sources of variance would make our study less relevant to its natural ecological context. This is the blessing and the curse of evolutionary field experiments. Although adding predators and competitors tended to decrease predictability even further, the CON islands revealed that these populations were not at an ecological or evolutionary equilibrium to begin with. This complicates prediction, but we expect that this is true for most if not all free-living populations and species that people might care to make predictions about. If further experiments can examine and disentangle some of these forces, then they can be incorporated into updated models, as can our new understanding of the context dependence of hindlimb length change, to improve future forecasts of phenotypic change (Urban et al. 2016).

Predicting Genetic Changes

At the genetic level, we found little evidence that evolutionary change could be explained, even post hoc, by these novel species interactions. Population structure changed little during our experiment (fig. S5), and population-level allele frequency changes were not correlated within treatments, not associated with phenotypic change, and not related to ecological differences between islands. These null results suggest that drift may be the main driver of the frequency changes in the SNPs we examined, which cover a small proportion of the genome. This would be consistent with the small effective population sizes estimated for these islands (table S13). Most previous field experiments with brown anoles have not had a genetic component, and the exceptions (e.g., Kolbe et al. 2012) have not calculated effective population size, so we do not know whether our populations have unusually small effective population size for island populations of brown anoles.

Another possible explanation is that genetic adaptation is occurring but through modes that we lacked the statistical power to detect in single-SNP tests. When traits under selection have a diffuse genetic architecture (i.e., hundreds or thousands of loci affecting the trait in small, difficult-to-detect ways), genetic adaptation occurs through subtle, coordinated shifts in the frequencies of multiple alleles. Of the phenotypic effects we focused on, relative hindlimb length is perhaps unlikely to experience highly polygenic selection: variation in this trait was recently mapped to a large-effect locus (Bock et al. 2021). We found no signal of this association in our populations. This could be biological: the genetic architecture of traits can vary across populations within a species. In rainbow trout (*Oncorhynchus mykiss*), for example, a large-effect locus underlying life history variation at lower latitudes is not associated with life history in an Alaskan population (Weinstein et al. 2019). Or we may simply lack statistical power to detect phenotype-associated SNPs, especially given the significant population structure across our islands (Santure and Garant 2018). We know little about the genetic architecture of the other traits we examined and thus can say little about whether selection at these traits could lead to polygenic adaptation. Detecting polygenic selection remains challenging, despite recent progress (e.g., Bourret et al. 2014; Gompert et al. 2017). One promising approach for detecting polygenic selection is to use time series data (e.g., Terhorst et al. 2015; Gompert 2016; Buffalo and Coop 2019). As our experiment continues and we accumulate more sampling points for allele frequencies, we may be able to uncover signatures of selection at the genetic level that we do not yet have the power to detect, if such selection is indeed occurring.

We can draw parallels between our experiment and similar studies to better understand some of the factors influencing the predictability of evolution in natural populations. Des Roches et al. (2017), for example, found unexpected patterns of survival for southwestern fence lizards (*Sceloporus cowlesi*) with seemingly mismatched genotypes, despite previously strong associations between

coloration, genotype, and survival. They attributed their unexpected results, in part, to a change in ecological context: their field experiment took place along an ecotone, whereas previous work had been performed in consistent habitat. As with our change in context from small to large islands, seemingly small changes in context can greatly alter evolutionary outcomes and complicate predictions. Other studies have found high repeatability and predictability. In threespine stickleback (*Gasterosteus aculeatus*), for example, multiple experiments have demonstrated consistent patterns of allele frequency change at *Eda*, a gene associated with armor plating, when marine-derived individuals colonize freshwater ponds (Barrett et al. 2008; Schluter et al. 2020). Such large environmental changes led to strong selection (i.e., $s \approx 0.5$), which increases predictability, especially when using genetically similar starting individuals (heterozygotes at *Eda* [Barrett et al. 2008]; F_2s from an intercross [Schluter et al. 2020]). However, this predictability decreases with the degree of ecological and genetic complexity being considered. Rennison et al. (2019) manipulated the presence/absence of predators in seminatural experimental ponds with freshwater stickleback and found less predictability: some armor traits and genes diverged as expected, while others were more variable. As we see in our experiment, adding further manipulations of ecological interactions can lead to an even bigger decrease in predictability, as does carrying out experiments in natural populations. Nevertheless, our results demonstrate the value of careful experimentation tied to extensive data collection informed by ecology, which led to explanations for some of the phenotypic changes we observed and which can in turn be incorporated into future attempts at prediction. In (one version) of Gould's thought experiment, the tape of life is rewound to an earlier point, some small change is made, and then life proceeds again. Our experiment is, in some ways, a large-scale instantiation of this thought experiment using wild populations of organisms. We found that these small changes can have important consequences for how the tape of life plays forward under complex natural conditions and that the limits to predictability can be surprisingly near.

Acknowledgments

Statement of Authorship

T.J.T., T.M.P., T.W.S., D.A.S., J.B.L., R.M.P., and R.D.H.B. designed the research; T.J.T., T.M.P., J.J.K., A.M.A., K.M.G., O.L., T.R.K., N.M.i.V., L.J.R., J.E.W., D.A.S., J.B.L., R.M.P., and R.D.H.B. performed fieldwork and collected data; T.J.T. performed lab work; T.J.T. analyzed data, with assistance from R.M.P. and R.D.H.B.; T.J.T., R.M.P., and R.D.H.B. wrote the initial manuscript; and all authors reviewed and edited the manuscript.

Data and Code Availability

All code and data needed to recreate this article are available in the Dryad Digital Repository (https://doi.org/10.5061/dryad.xd2547dhj; Thurman et al. 2022). Raw and summary data on lizard population sizes and perching behavior are also available in the supplementary information of Pringle et al. (2019; https://doi.org/10.1038/s41586-019-1264-6), and analysis code is also available on GitHub (https://github.com/tjthurman/anolis_pub). Raw sequencing reads are deposited in the National Center for Biotechnology Information Sequence Read Archive under BioProject accession number PRJNA874464.

Literature Cited

Adams, D. C., and M. L. Collyer. 2009. A general framework for the analysis of phenotypic trajectories in evolutionary studies. Evolution 63:1143–1154.

Bailey, S. F., and T. Bataillon. 2016. Can the experimental evolution programme help us elucidate the genetic basis of adaptation in nature? Molecular Ecology 25:203–218.

Bailey, S. F., Q. Guo, and T. Bataillon. 2018. Identifying drivers of parallel evolution: a regression model approach. Genome Biology and Evolution 10:2801–2812.

Barrett, M. 2021. ggokabeito: "Okabe-Ito" scales for "ggplot2" and "ggraph." R package version 0.1.0. https://CRAN.R-project.org/package=ggokabeito.

Barrett, R. D. H., S. Laurent, R. Mallarino, S. P. Pfeifer, C. C. Y. Xu, M. Foll, K. Wakamatsu, J. S. Duke-Cohan, J. D. Jensen, and H. E. Hoekstra. 2019. Linking a mutation to survival in wild mice. Science 363:499–504.

Barrett, R. D. H., S. M. Rogers, and D. Schluter. 2008. Natural selection on a major armor gene in threespine stickleback. Science 322:255–257.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1–48.

Beuttell, K., and J. B. Losos. 1999. Ecological morphology of Caribbean anoles. Herpetological Monographs 13:1–28.

Blount, Z. D., R. E. Lenski, and J. B. Losos. 2018. Contingency and determinism in evolution: replaying life's tape. Science 362: eaam5979.

Bock, D. G., S. Baeckens, J. N. Pita-Aquino, Z. A. Chejanovski, S. N. Michaelides, P. Muralidhar, O. Lapiedra, et al. 2021. Changes in selection pressure can facilitate hybridization during biological invasion in a Cuban lizard. Proceedings of the National Academy of Sciences of the USA 118:e2108638118.

Bolnick, D. I., R. D. H. Barrett, K. B. Oke, D. J. Rennison, and Y. E. Stuart. 2018. (Non)parallel evolution. Annual Review of Ecology, Evolution, and Systematics 49:303–330.

Bonneaud, C., E. Marnocha, A. Herrel, B. Vanhooydonck, D. J. Irschick, and T. B. Smith. 2015. Developmental plasticity affects sexual size dimorphism in an anole lizard. Functional Ecology 30:235–243.

Bourret, V., M. Dionne, and L. Bernatchez. 2014. Detecting genotypic changes associated with selective mortality at sea in Atlantic salmon: polygenic multilocus analysis surpasses genome scan. Molecular Ecology 23:4444–4457.

Buffalo, V., and G. Coop. 2019. The linked selection signature of rapid adaptation in temporal genomic data. Genetics 213:1007–1045.

Burke, M. K., J. P. Dunham, P. Shahrestani, K. R. Thornton, M. R. Rose, and A. D. Long. 2010. Genome-wide analysis of a long-term evolution experiment with *Drosophila*. Nature 467:587–590.

Bürkner, P.-C. 2017. brms: an R package for Bayesian multilevel models using Stan. Journal of Statistical Software 80:1–28.

———. 2018. Advanced Bayesian multilevel modeling with the R package brms. R Journal 10:395–411.

Butler, M. A. 2007. Vive le difference! sexual dimorphism and adaptive patterns in lizards of the genus *Anolis*. Integrative and Comparative Biology 47:272–284.

Butler, M. A., and J. B. Losos. 2002. Multivariate sexual dimorphism, sexual selection, and adaptation in Greater Antillean *Anolis* lizards. Ecological Monographs 72:541–559.

Calsbeek, R. 2009. Experimental evidence that competition and habitat use shape the individual fitness surface. Journal of Evolutionary Biology 22:97–108.

Calsbeek, R., and C. Bonneaud. 2008. Postcopulatory fertilization bias as a form of cryptic sexual selection. Evolution 62:1137–1148.

Calsbeek, R., L. Bonvini, and R. M. Cox. 2010. Geographic variation, frequency-dependent selection, and the maintenance of a female-limited polymorphism. Evolution 64:116–125.

Calsbeek, R., and R. M. Cox. 2010. Experimentally assessing the relative importance of predation and competition as agents of selection. Nature 465:613–616.

Calsbeek, R., and T. B. Smith. 2007. Probing the adaptive landscape using experimental islands: density-dependent natural selection on lizard body size. Evolution 61:1052–1061.

Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait. 2011. Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes, Genomes, Genetics 1:171–182.

Catchen, J. M., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22:3124–3140.

Chamberlain, S. A., J. L. Bronstein, and J. A. Rudgers. 2014. How context dependent are species interactions? Ecology Letters 17:881–890.

Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4:7.

Chase, J. M., P. A. Abrams, J. P. Grover, S. Diehl, P. Chesson, R. D. Holt, S. A. Richards, R. M. Nisbet, and T. J. Case. 2002. The interaction between predation and competition: a review and synthesis. Ecology Letters 5:302–315.

Chesson, P., and J. J. Kuang. 2008. The interaction between predation and competition. Nature 456:235–238.

Collyer, M. L., and D. C. Adams. 2007. Analysis of two-state multivariate phenotypic change in ecological studies. Ecology 88:683–692.

Cox, R. M., and R. Calsbeek. 2010. Sex-specific selection and intraspecific variation in sexual size dimorphism. Evolution 64:798–809.

Cox, R. M., R. A. Costello, B. E. Camber, and J. W. McGlothlin. 2017. Multivariate genetic architecture of the *Anolis* dewlap reveals both shared and sex-specific features of a sexually dimorphic ornament. Journal of Evolutionary Biology 30:1262–1275.

Danecek, P., A. Auton, G. R. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, et al. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Danecek, P., S. Schiffels, and R. Durbin. 2014. Multiallelic calling model in bcftools (-m). http://samtools.github.io/bcftools/call-m.pdf.

Davis, M. A. 2003. Biotic globalization: does competition from introduced species threaten biodiversity? BioScience 53:481.

Des Roches, S., R. Sollmann, K. Calhoun, A. P. Rothstein, and E. B. Rosenblum. 2017. Survival by genotype: patterns at *Mc1r* are not black and white at the White Sands ecotone. Molecular Ecology 26:320–329.

Diagne, C., B. Leroy, A.-C. Vaissière, R. E. Gozlan, D. Roiz, I. Jarić, J.-M. Salles, C. J. A. Bradshaw, and F. Courchamp. 2021. High and rising economic costs of biological invasions worldwide. Nature 592:571–576.

Doherty, T. S., A. S. Glen, D. G. Nimmo, E. G. Ritchie, and C. R. Dickman. 2016. Invasive predators and global biodiversity loss. Proceedings of the National Academy of Sciences of the USA 113:11261–11265. http://www.pnas.org/lookup/doi/10.1073/pnas.1602480113.

Edwards, S. M. 2020. lemon: freshing up your "ggplot2" plots. R package version 0.4.5. https://CRAN.R-project.org/package=lemon.

Elstrott, J., and D. J. Irschick. 2004. Evolutionary correlations among morphology, habitat use and clinging performance in Caribbean *Anolis* lizards. Biological Journal of the Linnean Society 83:389–398.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York.

Glossip, D., and J. B. Losos. 1997. Ecological correlates of number of subdigital lamellae in anoles. Herpetologica 53:192–199.

Gompert, Z. 2016. Bayesian inference of selection in a heterogeneous environment from genetic time-series data. Molecular Ecology 25:121–134.

Gompert, Z., S. P. Egan, R. D. H. Barrett, J. L. Feder, and P. Nosil. 2017. Multilocus approaches for the measurement of selection on correlated genetic loci. Molecular Ecology 26:365–382.

Gotanda, K. M., and A. P. Hendry. 2014. Using adaptive traits to consider potential consequences of temporal variation in selection: male guppy colour through time and space. Biological Journal of the Linnean Society 112:108–122.

Gould, S. J. 1989. Wonderful life: the Burgess Shale and the nature of history. Norton, New York.

Green, P., and C. J. MacLeod. 2016. SIMR: an R package for power analysis of generalized linear mixed models by simulation. Methods in Ecology and Evolution 7:493–498.

Hecht, M. K. 1952. Natural selection in the lizard genus *Aristelliger*. Evolution 6:112–114.

Heckel, D. G., and J. Roughgarden. 1979. A technique for estimating the size of lizard populations. Ecology 60:966–975.

Jorde, P. E., and N. Ryman. 2007. Unbiased estimator for genetic drift and effective population size. Genetics 177:927–935.

Kahle, D., and H. Wickham. 2013. ggmap: spatial visualization with ggplot2. R Journal 5:144–161. https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf.

Kawecki, T. J., R. E. Lenski, D. Ebert, B. Hollis, I. Olivieri, and M. C. Whitlock. 2012. Experimental evolution. Trends in Ecology and Evolution 27:547–560.

Kim, S. Y., K. E. Lohmueller, A. Albrechtsen, Y. Li, T. Korneliussen, G. Tian, N. Grarup, et al. 2011. Estimation of allele frequency and association mapping using next-generation sequencing data. BMC Bioinformatics 12:231.

Kolbe, J. J., N. Gilbert, J. T. Stroud, and Z. A. Chejanovski. 2021. An experimental analysis of perch diameter and substrate preferences of *Anolis* lizards from natural forest and urban habitats. Journal of Herpetology 55:215–221.

Kolbe, J. J., M. Leal, T. W. Schoener, D. A. Spiller, and J. B. Losos. 2012. Founder effects persist despite adaptive differentiation: a field experiment with lizards. Science 335:1086–1089.

Korneliussen, T. S., A. Albrechtsen, and R. Nielsen. 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinformatics 15:356.

Kreitman, M., and H. Akashi. 1995. Molecular evidence for natural selection. Annual Review of Ecology and Systematics 26:403–422.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. lmerTest package: tests in linear mixed effects models. Journal of Statistical Software 82:1–26.

Lapiedra, O., T. W. Schoener, M. Leal, J. B. Losos, and J. J. Kolbe. 2018. Predator-driven natural selection on risk-taking behavior in anole lizards. Science 360:1017–1020.

Lawrence, M., W. Huber, H. Pagès, P. Aboyoun, M. Carlson, R. Gentleman, M. Morgan, and V. Carey. 2013. Software for computing and annotating genomic ranges. PLoS Computational Biology 9:e1003118.

Lenski, R. E., M. R. Rose, S. C. Simpson, and S. C. Tadler. 1991. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. American Naturalist 138:1315–1341.

Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993.

———. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv, https://doi.org/10.48550/arXiv.1303.3997.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760.

Lima, S. L., and L. M. Dill. 1990. Behavioral decisions made under the risk of predation: a review and prospectus. Canadian Journal of Zoology 68:619–640.

Lopez-Darias, M., T. W. Schoener, D. A. Spiller, and J. B. Losos. 2012. Predators determine how weather affects the spatial niche of lizard prey: exploring niche dynamics at a fine scale. Ecology 93:2512–2518.

Losos, J. B. 1990a. Ecomorphology, performance capability, and scaling of West Indian *Anolis* lizards: an evolutionary analysis. Ecological Monographs 60:369–388.

———. 1990b. The evolution of form and function: morphology and locomotor performance in West Indian *Anolis* lizards. Evolution 44:1189–1203.

———. 2009. Lizards in an evolutionary tree: ecology and adaptive radiation of anoles. University of California Press, Berkeley.

———. 2011. Convergence, adaptation, and constraint. Evolution 65:1827–1840.

Losos, J. B., and D. J. Irschick. 1996. The effect of perch diameter on escape behaviour of *Anolis* lizards: laboratory predictions and field tests. Animal Behaviour 51:593–602.

Losos, J. B., T. W. Schoener, R. B. Langerhans, and D. A. Spiller. 2006. Rapid temporal reversal in predator-driven natural selection. Science 314:1111.

Losos, J. B., T. W. Schoener, and D. A. Spiller. 2004. Predator-induced behaviour shifts and natural selection in field-experimental lizard populations. Nature 432:505–508.

Losos, J. B., and B. Sinervo. 1989. The effects of morphology and perch diameter on sprint performance of *Anolis* lizards. Journal of Experimental Biology 145:23–30.

Losos, J. B., and D. A. Spiller. 1999. Differential colonization success and asymmetrical interactions between two lizard species. Ecology 80:252–258.

Losos, J. B., K. I. Warheit, and T. W. Schoener. 1997. Adaptive differentiation following experimental island colonization in *Anolis* lizards. Nature 387:70–73.

Macrini, T. E., D. J. Irschick, and J. B. Losos. 2003. Ecomorphological differences in toepad characteristics between mainland and island anoles. Journal of Herpetology 37:52–58.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17. https://doi.org/10.14806/ej.17.1.200.

McGlothlin, J. W., R. M. Cox, and E. D. Brodie. 2019. Sex-specific selection and the evolution of between-sex genetic covariance. Journal of Heredity 110:422–432.

McGlothlin, J. W., M. E. Kobiela, H. V. Wright, J. J. Kolbe, J. B. Losos, and E. D. Brodie III. 2022. Conservation and convergence of genetic architecture in the adaptive radiation of *Anolis* lizards. American Naturalist 200:621–737.

Meisner, J., and A. Albrechtsen. 2018. Inferring population structure and admixture proportions in low-depth NGS data. Genetics 210:719–731.

Microsoft, and S. Weston. 2020a. doParallel: foreach parallel adaptor for the "parallel" package. R package version 1.0.16. https://CRAN.R-project.org/package=doParallel.

———. 2020b. foreach: provides foreach looping construct. R package version 1.5.1. https://CRAN.R-project.org/package=foreach.

Mooney, H. A., and E. E. Cleland. 2001. The evolutionary impact of invasive species. Proceedings of the National Academy of Sciences of the USA 98:5446–5451.

Müller, K., and H. Wickham. 2020. tibble: simple data frames. R package version 3.0.4. https://CRAN.R-project.org/package=tibble.

Nei, M., Y. Suzuki, and M. Nozawa. 2010. The neutral theory of molecular evolution in the genomic era. Annual Review of Genomics and Human Genetics 11:265–289.

Oke, K. B., G. Rolshausen, C. LeBlond, and A. P. Hendry. 2017. How parallel is parallel evolution? a comparative analysis in fishes. American Naturalist 190:1–16.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, et al. 2020. vegan: community ecology package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan.

Orozco terWengel, P., M. Kapun, V. Nolte, R. Kofler, T. Flatt, and C. Schlotterer. 2012. Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. Molecular Ecology 21:4931–4941.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135.

Pringle, R. M., T. R. Kartzinel, T. M. Palmer, T. J. Thurman, K. Fox-Dobbs, C. C. Y. Xu, M. C. Hutchinson, et al. 2019. Predator-induced collapse of niche structure and species coexistence. Nature 570:58–64.

Pringle, R. M., D. M. Kimuyu, R. L. Sensenig, T. M. Palmer, C. Riginos, K. E. Veblen, and T. P. Young. 2015. Synergistic effects of fire and elephants on arboreal animals in an African savanna. Journal of Animal Ecology 84:1637–1645.

Purcell, S. M., and C. C. Chang. 2017. PLINK version 1.90b4.6. https://www.cog-genomics.org/plink/1.9/.

Pyke, G. H., H. R. Pulliam, and E. L. Charnov. 1977. Optimal foraging: a selective review of theory and tests. Quarterly Review of Biology 52:137–154.

Rasband, W. 1997–2018. ImageJ. US National Institutes of Health, Bethesda, MD. https://imagej.nih.gov/ij/.

Reedy, A. M., B. D. Pope, N. M. Kiriazis, C. L. Giordano, C. L. Sams, D. A. Warner, and R. M. Cox. 2017. Female anoles display less but attack more quickly than males in response to territorial intrusions. Behavioral Ecology 28:1323–1328.

Rennison, D. J., S. M. Rudman, and D. Schluter. 2019. Genetics of adaptation: experimental test of a biotic mechanism driving divergence in traits and genes. Evolution Letters 3:513–520.

Reznick, D., H. Bryga, and J. A. Endler. 1990. Experimentally induced life-history evolution in a natural population. Nature 346:357–359.

Robinson, D., A. Hayes, and S. Couch. 2020. broom: convert statistical objects into tidy tibbles. R package version 0.7.3. https://CRAN.R-project.org/package=broom.

Santos Baquero, O. 2019. ggsn: north symbols and scale bars for maps created with "ggplot2" or "ggmap." R package version 0.5.0. https://CRAN.R-project.org/package=ggsn.

Santure, A. W., and D. Garant. 2018. Wild GWAS-association mapping in natural populations. Molecular Ecology Resources 18:729–738.

Sax, D., J. Stachowicz, J. Brown, J. Bruno, M. Dawson, S. Gaines, R. Grosberg, A. Hastings, R. Holt, and M. Mayfield. 2007. Ecological and evolutionary insights from species invasions. Trends in Ecology and Evolution 22:465–471.

Schlotterer, C., R. Kofler, E. Versace, R. Tobler, and S. U. Franssen. 2015. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. Heredity 114:431–440.

Schluter, D., K. B. Marchinko, M. E. Arnegard, H. Zhang, S. D. Brady, F. C. Jones, M. A. Bell, and D. M. Kingsley. 2020. Fitness maps to a large-effect locus in introduced stickleback populations. Proceedings of the National Academy of Sciences of the USA 118:e1914889118.

Schluter, D., T. Price, A. O. Mooers, and D. Ludwig. 1997. Likelihood of ancestor states in adaptive radiation. Evolution 51:1699–1711.

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

Schoener, T. W. 1967. The ecological significance of sexual dimorphism in size in the lizard *Anolis conspersus*. Science 155:474–477.

Schoener, T. W., J. J. Kolbe, M. Leal, J. B. Losos, and D. A. Spiller. 2017. A multigenerational field experiment on eco-evolutionary dynamics of the influential lizard *Anolis sagrei*: a mid-term report. Copeia 105:543–549.

Schoener, T. W., J. B. Losos, and D. A. Spiller. 2005. Island biogeography of populations: an introduced species transforms survival patterns. Science 310:1807–1809.

Schoener, T. W., and A. Schoener. 1980. Densities, sex ratios, and population structure in four species of Bahamian *Anolis* lizards. Journal of Animal Ecology 49:19–53.

Schoener, T. W., D. A. Spiller, and J. B. Losos. 2002. Predation on a common *Anolis* lizard: can the food-web effects of a devastating predator be reversed? Ecological Monographs 72:383–407.

Schoener, T. W., D. A. Spiller, J. Piovia-Scott, and R. Ricklefs. 2016. Variation in ecological interaction strength with island area: theory and data from the Bahamian archipelago. Global Ecology and Biogeography 25:891–899.

Schweyen, H., A. Rozenberg, and F. Leese. 2014. Detection and removal of PCR duplicates in population genomic ddRAD studies by addition of a degenerate base region (DBR) in sequencing adapters. Biological Bulletin 227:146–160.

Segrè, A. V., A. W. Murray, and J.-Y. Leu. 2006. High-resolution mutation mapping reveals parallel experimental evolution in yeast. PLoS Biology 4:e256.

Sih, A. 1980. Optimal behavior: can foragers balance two conflicting demands? Science 210:1041–1043.

Simberloff, D., J.-L. Martin, P. Genovesi, V. Maris, D. A. Wardle, J. Aronson, F. Courchamp, et al. 2013. Impacts of biological invasions: what's what and the way forward. Trends in Ecology and Evolution 28:58–66.

Slowikowski, K. 2021. ggrepel: automatically position non-overlapping text labels with "ggplot2." R package version 0.9.1. https://CRAN.R-project.org/package=ggrepel.

Stuart, Y. E., T. S. Campbell, P. A. Hohenlohe, R. G. Reynolds, L. J. Revell, and J. B. Losos. 2014. Rapid evolution of a native species following invasion by a congener. Science 346:463–466.

Terhorst, J., C. Schlotterer, and Y. S. Song. 2015. Multi-locus analysis of genomic time series data from experimental evolution. PLoS Genetics 11:e1005069.

Thorpe, R. S., J. T. Reardon, and A. Malhotra. 2005. Common garden and natural selection experiments support ecotypic differentiation in the Dominican anole (*Anolis oculatus*). American Naturalist 165:495–504.

Thurman, T. J., T. M. Palmer, J. J. Kolbe, A. M. Askary, K. M. Gotanda, O. Lapiedra, T. R. Kartzinel, et al. 2022. Data from: The difficulty of predicting evolutionary change in response to

novel ecological interactions: a field experiment with *Anolis* lizards. American Naturalist, Dryad Digital Repository, https://doi.org/10.5061/dryad.xd2547dhj.

Urban, M. C., G. Bocedi, A. P. Hendry, J. B. Mihoub, G. Peer, A. Singer, J. R. Bridle, et al. 2016. Improving the forecast for biodiversity under climate change. Science 353:aad8466.

Ushey, K. 2020. renv: project environments. R package version 0.12.0. https://CRAN.R-project.org/package=renv.

Vendrami, D. L. J., L. Telesca, H. Weigand, M. Weiss, K. Fawcett, K. Lehman, M. S. Clark, et al. 2017. RAD sequencing resolves fine-scale population structure in a benthic invertebrate: implications for understanding phenotypic plasticity. Royal Society Open Science 4:160548.

Vischer, N., and S. Nastase. 2015. ObjectJ version 1.03w. https://sils.fnwi.uva.nl/bcb/objectj/index.html.

Weinstein, S. Y., F. P. Thrower, K. M. Nichols, and M. C. Hale. 2019. A large-scale chromosomal inversion is not associated with life history development in rainbow trout from southeast Alaska. PLoS ONE 14:e0223018.

Wickham, H. 2016. ggplot2: elegant graphics for data analysis. Springer, New York. https://ggplot2.tidyverse.org.

———. 2019*a*. assertthat: easy pre and post assertions. R package version 0.2.1. https://CRAN.R-project.org/package=assertthat.

———. 2019*b*. stringr: simple, consistent wrappers for common string operations. R package version 1.4.0. https://CRAN.R-project.org/package=stringr.

———. 2020*a*. forcats: tools for working with categorical variables (factors). R package version 0.5.0. https://CRAN.R-project.org/package=forcats.

———. 2020*b*. tidyr: tidy messy data. R package version 1.1.2. https://CRAN.R-project.org/package=tidyr.

Wickham, H., R. François, L. Henry, and K. Müller. 2020. dplyr: a grammar of data manipulation. R package version 1.0.2. https://CRAN.R-project.org/package=dplyr.

Wickham, H., and J. Hester. 2020. readr: read rectangular text data. R package version 1.4.0. https://CRAN.R-project.org/package=readr.

Wilke, C. O. 2020. cowplot: streamlined plot theme and plot annotations for "ggplot2." R package version 1.1.1. https://CRAN.R-project.org/package=cowplot.

Williams, E. E. 1972. The origin of faunas. Evolution of lizard congeners in a complex island fauna: a trial analysis. Pages 47–89 *in* T. Dobzhansky, M. K. Hecht, and W. C. Steere, eds. Evolutionary biology. Vol. 6. Springer, New York.

Williams, J. W., and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological surprises. Frontiers in Ecology and the Environment 5:475–482.

Wilson, D. J. 2019. The harmonic mean *p*-value for combining dependent tests. Proceedings of the National Academy of Sciences of the USA 104:201814092.

Xie, Y. 2014. knitr: a comprehensive tool for reproducible research in R. *In* V. Stodden, F. Leisch, and R. D. Peng, eds. Implementing reproducible computational research. CRC, Boca Raton, FL.

Zhu, H. 2020. kableExtra: construct complex table with "kable" and pipe syntax. R package version 1.3.1. https://CRAN.R-project.org/package=kableExtra.

References Cited Only in the Online Enhancements

Donihue, C. M., A. Herrel, A.-C., Fabre, A. Kamath, A. J. Geneva, T. W. Schoener, J. J. Kolbe, and J. B. Losos. 2018. Hurricane-induced selection on the morphology of an island lizard. Nature 560:88–91.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457–472.

Hereford, J., T. F. Hansen, and D. Houle. 2004. Comparing strengths of directional selection: how strong is strong? Evolution 58:2133–2143.

Herrel, A., R. Joachim, B. Vanhooydonck, and D. J. Irschick. 2006. Ecological consequences of ontogenetic changes in head shape and bite performance in the Jamaican lizard *Anolis lineatopus*. Biological Journal of the Linnean Society 89:443–454.

Herrel, A., L. D. McBrayer, and P. M. Larson. 2007. Functional basis for sexual differences in bite force in the lizard *Anolis carolinensis*. Biological Journal of the Linnean Society 91:111–119.

Les, A. M., and R. Powell. 2014. *Anolis smaragdinus*. Catalogue of American Amphibians and Reptiles 902:1–14.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. R. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

Quinn, G. P., and M. J. Keough. 2002. Experimental design and data analysis for biologists. Cambridge University Press, Cambridge.

Rohland, N., and D. Reich. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Research 22:939–946.

Schoener, T. W., and A. Schoener. 1982. Intraspecific variation in home-range size in some *Anolis* lizards. Ecology 63:809–823.

———. 1983. The time to extinction of a colonizing propagule of lizards increases with island area. Nature 332:332–334.

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

Whitlock, M. C., and D. Schluter. 2009. The analysis of biological data. Roberts, Greenwood Village, CO.

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

Zhou, X., P. Carbonetto, and M. Stephens. 2013. Polygenic modeling with Bayesian sparse linear mixed models. PLoS Genetics 9:e1003264.

Zhou, X., and M. Stephens. 2012. Genome-wide efficient mixed-model analysis for association studies. Nature Ecology and Evolution 44:821–824.