Phylogenetic conservation of behavioural variation and behavioural syndromes

Jeremy Dalos, Raphaël Royauté, Ann Hedrick, Ned Dochtermann

Phylogenetic conservation of behavioral variation and behavioral syndromes

Jeremy Dalos [a, b]
Raphael Royauté [c, d]
Ann Hedrick [e, f]
Ned A. Dochtermann [a,g]

[a] Department of Biological Sciences; North Dakota State University
[b] jeremy.dalos@ndsu.edu
[c] Movement Ecology Group, Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Frankfurt, Germany
[d] raphael.royaute@gmail.com
[e] Department of Neurobiology, Physiology, and Behavior; University of California, Davis
[f] avhedrick@ucdavis.edu
[g] **corresponding author**: ned.dochtermann@gmail.com

Author to whom correspondence should be addressed:
Ned Dochtermann
1340 Bolley Drive, 201 Stevens Hall
Department of Biological Sciences
North Dakota State University
Fargo, ND 58102

Running Title: Conservation of behavioral (co)variation

Keywords: behavioral syndrome, personality, field crickets, comparative method

Acknowledgements

We thank Jonathan Albers, Monica Berdal, Katelyn Cannon, Sarah Felde, Brady Klock, Ishan Joshi, Hannah Lambert, Jenna LaCoursiere, Alondra Neunsinger, Katie Pnewski, Maddi Rick, Teresa Tesarek, and Amanda Wilson for assistance in conducting behavioral trials and in rearing and care of the crickets and Martori Farms, David Lightfoot, Scott Bundy, Nico Franz, Sangmi Lee, Cameron Jones, Kenny Chapin, Ti Eriksson, Meranda Feagins, Charlotte Mulloney, Melody Martinez, Allyson Richins, Mauriel Rodriguez, Helen Vessels and David Wikman for assistance in collecting *Gryllus integer, Gryllus lineaticeps, and Gryllus assimilis*. We also thank Scott Sakaluk for providing us with populations of *Gryllodes sigillatus*. This work was supported by US NSF IOS grants 1557951 and 1558069 to N.A.D. and A.H. respectively.

Abstract

Individuals frequently differ consistently from one another in their average behaviors (i.e. "animal personality") and in correlated suites of consistent behavioral responses (i.e. "behavioral syndromes"). However, understanding the evolutionary basis of this (co)variation has lagged behind demonstrations of its presence. This lag partially stems from comparative methods rarely being used in the field. Consequently, much of the research on animal personality has relied on "adaptive stories" focused on single species and populations. Here we used a comparative approach to examine the role of phylogeny in shaping patterns of average behaviors, behavioral variation, and behavioral correlations. In comparing the behaviors and behavioral variation for five species of Gryllid crickets we found that phylogeny shaped average behaviors and behavioral (co)variation. Variation in average exploratory behavior and response to cues of predator presence attributable to phylogeny was greater or comparable to the magnitude of "personality variation". Likewise, magnitudes of variation were concordant with evolutionary relationships and behavioral correlations were consistent across species. These results suggest that phylogenetic constraints play an important role in the expression of animal personalities and behavioral syndromes and emphasize the importance of examining evolutionary explanations within a comparative framework.

Introduction

Behavioral syndromes, correlations between behaviors at the among-individual level (Dingemanse et al. 2012), have been documented across taxa (Brommer and Class 2017). Behavioral syndromes can conceptually be thought of as correlations between individual averages and stem from underlying genetic correlations and correlations due to developmental plasticity and other sources of permanent environmental covariance (Dingemanse et al. 2012, Dingemanse and Dochtermann 2014). Among-individual variation in behavior, often referred to as "personality variation", has been found to be similarly ubiquitous (Bell et al. 2009). Similar to behavioral syndromes, this personality variation can be thought of as variation across individuals in their average behaviors and likewise stem from genetic and permanent environmental variation (Dingemanse and Dochtermann

2013, Dochtermann et al. 2015). Attempts to infer whether general taxonomic patterns exist for both personality variation and behavioral syndromes have generally been conducted via literature reviews and meta-analyses (Bell et al. 2009, Dochtermann 2011, Garamszegi et al. 2012, 2013, Dochtermann et al. 2015, Brommer and Class 2017). These synthesis efforts have shown that among-individual variation is common (average repeatability ~ 0.37, Bell et al. 2009), that the magnitude of behavioral syndromes is generally weak (average r ~ 0.19, Garamszegi et al 2012, 2013), and that there is general alignment between nested patterns of correlations at the phenotypic, among-individual, within-individual, and genetic levels (Dochtermann 2011, Brommer and Class 2017).

Despite the observation that both among-individual variation and behavioral syndromes are common, we have a poor understanding of the evolution of either. This gap in our understanding is partly because comparative approaches have rarely been used in studies of among-individual behavioral variation and behavioral syndromes (White et al. 2020), despite having been the backbone of studies of morphological evolution. Such approaches allow for direct comparison across species of behavioral (co)variation and are necessary for a proper understanding of the importance of phylogeny in shaping "personality" and behavioral syndromes (Royauté et al. 2020, White et al. 2020).

Direct assessment of evolutionary hypotheses can also be extended to the study of personality and behavioral syndromes: both among-individual variation and behavioral syndromes have clear connections to quantitative genetic parameters; specifically, additive genetic variation, and additive genetic covariances (Dochtermann and Roff 2010, Dingemanse and Dochtermann 2014). The mathematical relationships between among-individual (co)variances and additive genetic (co)variances (Boake 1989, Dingemanse and Dochtermann 2014, Dochtermann et al. 2015) allows the extension of predictions from quantitative genetics to among-individual variation and behavioral syndromes.

One such prediction is that differences in the magnitude of variation present for a trait might be attributable to differences in selection between populations or species. Specifically, Mousseau and Roff (1987) argued that traits with low heritability might be indicative of strong selection having eroded genetic variation. Likewise, because among-individual variation represents the sum of additive genetic variation, dominance (and other epistatic) genetic variation, and permanent environmental variation (e.g. irreversible and

developmental plasticity), selection is expected to deplete this variation. Note, however, that drift is often also expected to reduce genetic and, therefore, among-individual variation.

Selection is likewise expected to shape additive genetic covariances and correlations, both by the loss of variation in single traits and changes to the magnitude and directions of covariances (Roff 1997). For example, correlational selection is expected to produce genetic correlations (Phillips and Arnold 1989, Armbruster and Schwaegerle 1996). As in the case of among-individual variances, these effects on genetic correlations are expected to carry over to behavioral syndromes. In other words, behavioral syndromes are expected to reflect the effects of selection on genetic correlations. Therefore, if behavioral syndromes differ across populations, species, or other groupings, then this suggests differences in genetic correlations and correlational selection (i.e. the "adaptive" hypothesis, Bell 2005). In contrast, if behavioral syndromes are conserved across groups then this would suggest that either behavioral syndromes stem from pleiotropic effects (i.e. the "constraints" hypothesis, Bell 2005) or that selection is similar across groups.

While these topics have been addressed for other types of traits, particularly morphological and chemical characteristics (Aguirre et al. 2014, Hine et al. 2014, McGlothlin et al. 2018), addressing them for behavior remains important for several reasons. First, considerable behavioral research assumes an adaptive framework for both among-individual variation and behavioral syndromes, thereby minimizing the importance of phylogeny and minimizing the potential role of phylogenetic constraints. Second, behaviors, life-history, and physiological traits exhibit substantially lower heritabilities than do morphological traits (Mousseau and Roff 1987, Stirling et al. 2002, Dochtermann et al. 2019). Consequently, the role of phylogeny and selection in constraining and shaping morphology may not generalize to traits with lower heritabilities and thus greater plasticity.

Here, we compared the behavior of five closely related cricket species: *Gryllus integer, Gryllus assimilis, Gryllus lineaticeps, Gryllodes sigillatus,* and *Acheta domesticus*. For each species we measured exploratory behavior and response to cues of predator presence. By working with the same behavioral assays in five closely related species we

were able to assess the importance of phylogeny for average behaviors and to evaluate predictions about trait (co)variation. Specifically, we addressed the following questions:

1. Does the average expression of behavior differ among species?
   We predicted that species would differ but do so in a manner constrained by phylogeny. Put another way, more closely related species will have more similar average behaviors.
2. Do among-individual variances differ among species?
   We did not have species level predictions but, because selection and drift should both reduce among-individual variance, we predicted that among-individual variation would differ across species independent of phylogeny.
3. Do within-individual variances differ among species? Within-individual variation, typically disregarded as residual variation, includes phenotypic plasticity—specifically reversible plasticity or "phenotypic flexibility" not captured by factors and covariates of a statistical model (Piersma and Drent 2003, Whitman and Agrawal 2009, Piersma and Van Gils 2011, Westneat et al. 2015, Berdal and Dochtermann 2019). Differences across groups in the magnitude of within-individual variation therefore are, in part, differences in the magnitude of plasticity. We did not have *a priori* expectations as to species differences or phylogenetic signal for within-individual variances.
4. Do behavioral syndromes differ among species?
   Because behavioral syndrome structure has been conserved at the genetic level across cricket populations of *G. integer* (Royauté et al. 2020), we predicted that syndromes would similarly be phylogenetically conserved and shared across species.

Methods

Cricket Acquisition, Housing, and Rearing Conditions

Data used in this study were originally collected for various studies investigating the effects of development on behavioral variation and the presence of behavioral constraints and behavioral syndromes (Royauté et al. 2019, Royauté et al. 2020). *A. domesticus* males

and females were obtained as nymphs (~ 1 mm in size) from a commercial supplier (Fluker's Cricket Farm, Port Allen, LA, U.S.A.) and were measured once mature. *G. integer* females were captured in Aguila, AZ, *G. lineaticeps* males and females were caught in Dunnigan, CA, and the *G. assimilis* males and females were caught in Maricopa County, AZ. These species were all captured during the summer of 2017. *G. sigillatus* individuals were taken from an outbred population established by S. Sakaluk with crickets collected from California and currently maintained in Fargo, ND. For *G. lineaticeps* and *G. assimilis*, the same individuals that were caught in the field were measured, while lab reared offspring of *G. integer* were measured. All species were reared under a 12:12 light: dark photoperiod at a temperature of 25-28˚C. All individuals were housed in 0.71-liter containers with transparent covers that included food, shelter, and water filled glass vials plugged with cotton balls. *A. domesticus* were exposed to a mixture of high and low quality diets described in Royauté et al. (2019), while all other species included in this study were fed ad libitum food (commercially purchased chicken feed).

Behavior Trials

To measure exploratory behavior and anti-predator responses we repeatedly recorded individuals' activity levels in an open field arena, followed by their responses to cues of predator presence created from diluted *Eublepharis macularius* excreta (see details below). *A. domesticus* were measured between March 2015 and October 2016, *G. lineaticeps* were measured from August 2017 to September 2017, *G. assimilis* were measured between September 2017 and October 2017, *G. integer* were measured between May 2018 and June 2018, and *G. sigillatus* were measured in May 2019. All trials were conducted in a plastic arena (60 cm x 60 cm and 15 cm high) with a Plexiglas lid. The arena was split into four 30 cm × 30 cm arenas separated by a divider, allowing up to four crickets to be tested at one time. Open field trials were always conducted first followed by antipredator response trials either immediately after or on another day to minimized potential carryover effects from exposure to cues of predator presence. After each behavioral assay, arenas were thoroughly cleaned with 70% ethanol wipes to avoid accumulation of any chemical traces of conspecifics. Mass at the time of behavioral trials was recorded to the nearest 1 mg. All individuals were measured in each assay for a maximum of three repetitions, with some

individuals measured fewer times due to escape or natural mortality (Table 1). In total, we conducted 2478 behavioral assays across a total of 460 individuals (Table 1).

Table 1. Number of individuals, by species, for which behavior was assayed in a first, second, and third repetition.

Species	Behavioral Assay	Repetition 1	Repetition 2	Repetition 3	Total Trials
Acheta domesticus	Open field	281	263	225	769
	Antipredator	262	235	220	717
Gryllus assimilis	Open field	16	16	16	48
	Antipredator	16	16	15	47
Gryllus integer	Open field	92	91	74	257
	Antipredator	88	88	72	248
Gryllus lineaticeps	Open field	21	17	11	49
	Antipredator	21	13	11	45
Gryllodes sigillatus	Open field	50	50	49	149
	Antipredator	50	50	49	149
Total		896	837	743	2478

Open field behavior

Individual crickets were left to rest for 30 seconds under a 5 cm diameter cup after being introduced into the lower right section of the arena (Figure S1). After these 30 seconds we allowed the individuals to move freely through the arena for 220 seconds. We measured each individual's exploratory propensity by calculating the number of *unique zones* visited (UZ) by the cricket with Ethovision X (Noldus Information Technology, Wageningen, The Netherlands). This behavioral protocol has previously been used with *A. domesticus* and *G. integer* to evaluate genetic and individual differences in activity and exploratory behaviors (Royauté et al. 2015, Royauté and Dochtermann 2017, Royauté et al. 2019, Royauté et al. 2020).

Predator cue response

To measure responses to cues of potential predator presence, we collected excreta from three adult leopard geckos, *Eublepharis macularius*, that were fed a mixed diet of *A. domesticus, G. sigillatus, G. lineaticeps*, *G. integer*, and *G. assimilus*. Leopard geckos were housed according to the standards of the Institutional Animal Care and Use Committee of

North Dakota State University (Protocol A14006, A17015, and A19067) and the Animal Behavior Society (2020). Collected excreta was frozen and then finely ground and diluted with deionized water (1 ml H2O: 5 mg of excreta). This solution was then applied to 15 cm diameter filter paper disks with a 5 cm diameter central cutout that allows crickets to be left to rest unexposed to the predator cues (Royauté and Dochtermann 2017, Royauté et al. 2019, Royauté et al. 2020). Each predator cue disk was left to dry for a minimum of 2 hours then stored at -23˚C until needed for trials. Predator cue disks were allowed to warm to room temperature before use in antipredator trials and discarded after a single use. Between each trial, cue disks were stored at 4˚C for a maximum of 14 days.

We placed the predator cue disk at the bottom of a 15 cm diameter arena and left the cricket to rest for a minimum of 30 seconds under a 5 cm diameter cup in the nontreated central cutout. We then removed the cup and allowed the cricket to move freely for 220 seconds and estimated the distance travelled in cm (AP distance) using Ethovision X (Figure S1). Previous studies with this protocol show that crickets had heightened activity levels in the presence of this diluted gecko excreta compared to water controls (Royauté and Dochtermann 2017). Consistent with this, *G. sigillatus* crickets have been found to increase their activity after direct exposure to predators (Bucklaew and Dochtermann 2020). Greater activity during these antipredator response assays, i.e. greater *AP distance*, was therefore interpreted as a greater responsiveness to predator cues.

Data Analysis: Univariate Models

To assess differences in behavioral responses between species for means and variances we analyzed behavioral data using separate univariate mixed-effects models for unique zones visited and AP distance (square root transformed). We included species, temperature (Celsius, mean centered), mass (using among- and within-individual centering (Van de Pol and Wright 2009)), and sex as fixed effects. Individual ID was included as a random effect. We compared the fit of four univariate mixed models structured as follows:

1) Model 1: V_i = & V_w = A null model where the among- (V_i) and within-individual (V_w) variances were kept constant between species.

2) Model 2: $V_i \neq$ & V_w = A model where the among-individual variance differed between species, but the within-individual variance was kept constant.
3) Model 3: $V_i =$ & $V_w \neq$ The within-individual variance differs between species, but the among-individual variance was kept constant.
4) Model 4: $V_i \neq$ & $V_w \neq$ Both the among and within-individual variances were allowed to vary between species.

These models were specified using the `MCMCglmm` package for Bayesian mixed models (Hadfield 2010) using Markov-chain Monte Carlo (MCMC) with 1.3 million iterations, 300,000 iteration burn-in, a thinning interval of 1000, and an inverse-Wishart prior. AP distance and unique zone models were fit with Gaussian and Poisson error distributions, respectively.

To determine whether species differed in average behavior, Models 1 and 4 were run with and without species as a fixed effect and compared based on deviance information criterion (DIC) values. If species differ in average behavior, models with species included as a fixed effect would be expected to have lower (DIC) values. Average behavioral differences among species reported in the Results section were then qualitatively assessed using posterior-modal estimates for each species (Congdon 2006).

We then compared DIC values among models 1 through 4 to determine whether either among- or within-individual variances differed among species following Royauté et al. (2019) and Royauté and Dochtermann (2020). The model with the lowest DIC value was considered the best model and models with Δ DIC>5 were considered to have a substantively poorer fit (Barnett et al. 2010). Models with ΔDIC<5 were considered as having comparable support relative to the best model (Barnett et al. 2010). All models were specified with the same fixed effect structure as specified above to prevent biased estimates of variance components and repeatability (Spiegelhalter et al. 2003, Nakagawa and Schielzeth 2010b, Westneat et al. 2011).

Data Analysis: Phylogenetic Signal

As our primary questions were about differences in behavioral averages and variances, our results and discussion focus on the above model comparisons. However, we also calculated the variation in behavior directly attributable to phylogeny. To do so, we fit

mixed effects models with the same fixed effects, prior structure, and chain parameters as above but omitting Species as a fixed effect. Species was instead incorporated as a random effect, along with individual ID, with the relationship among species modeled according to the current phylogeny (Figure 1, Weissman and Gray (2019)). From these models we then estimated the strength of phylogenetic signal as the proportion of variation attributable to the hierarchical pattern of relatedness among species (i.e. λ, Pagel 1999, Hadfield and Nakagawa 2010, Nakagawa and Santos 2012). From the same models we also estimated the proportion of variation attributable to among-individual differences (i.e. τ, repetability, Dingemanse and Dochtermann 2013). We estimated both phylogenetic signal and repeatability as unadjusted values; that is, we included the variation attributable to fixed effects in the ratio denominator (Nakagawa and Schielzeth 2010a).

Data Analysis: Bivariate Models

Behavioral syndromes were estimated using bivariate mixed-effects models with unique zones traveled and AP distance as response variables, also using the `MCMCglmm` library (Hadfield 2010), and analyzed separately for each individual species. We fit models using temperature (Celsius, mean centered), mass (using among- and within-individual centering on subjects (Van de Pol and Wright 2009)), and sex as fixed effects and individual ID was fit as a random effect. These models were fit with 2.6 million iterations, a 600,000 burn-in period, a thinning interval of 2000, and a prior that was flat for correlations. Among-individual correlations were estimated for all species, while within-individual correlations were only assessed when individuals were measured for unique zones traveled and antipredator activity during the same testing period (Dingemanse and Dochtermann 2013). Consequently, we were unable to assess within-individual covariation of *G. lineaticeps* and *G. assimilis* due to the fact that these species were not measured for each behavior in immediate succession. Because model comparisons as used above for single traits could not be conducted for correlations (due to software imposed model limitations), differences in behavioral correlations across species were assessed based on whether 95% HPD intervals overlapped. Overlap of 95% intervals is an over-conservative comparison metric (Royauté and Dochtermann 2020), but this did not affect our species comparison results here. All analyses were conducted in R 3.4.4 (Team 2018).

Results

Differences in average behavior among species

Species differed in average behaviors: the inclusion of species as a fixed effect substantially improved model fit for both behaviors (Table 2, Table S1). The monophyletic group of *G. assimilis*, *G. integer*, and *G. lineaticeps* exhibited the lowest number of unique zones visited (Figure 1a) but differences in AP distance were less obviously associated with phylogenetic structure (Figure 1b). Consistent with this, phylogenetic signal was stronger for unique zones visited (λ : 0.27) than for AP distance (λ : 0.16; Table S2).

Table 2. DIC values for statistical models with and without the inclusion of species as a fixed effect. The effect of species was evaluated in a model where variances did not (Model 1) or did (Model 4) differ by species. For both behaviors and both models, the inclusion of species substantially improved model fit, as indicated by the lower DIC values for models with species included as a fixed effect.

	Behavior	**DIC with species**	**DIC without species**	**DIC(without) – DIC(with)**
Model 1 (V_i = & V_w =)	AP Distance	8025.51	8058.71	32.2
	Unique Zones Visited	8982.97	8456.17	526.8
Model 4 ($V_i \neq$ & $V_w \neq$)	AP Distance	7763.82	7780.94	17.12
	Unique Zones Visited	8338.21	8344.88	6.67

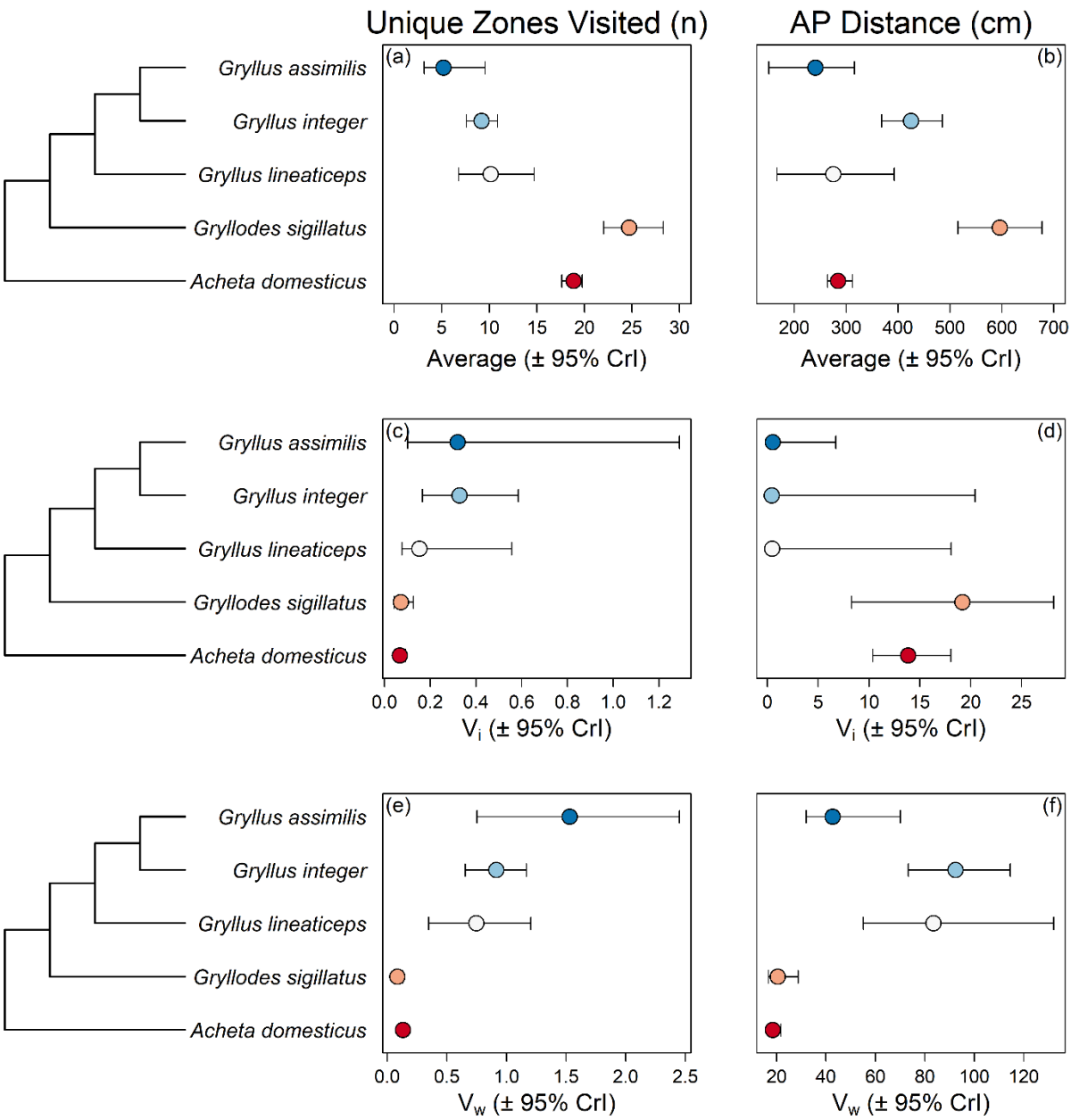

Figure 1. Species posterior-modal values with 95% HPD credibility intervals. (a) Average unique zones visited. (b) Average AP distance in centimeters. (c) Among-individual variances in unique zones traveled. (d) Among-individual variances in AP distance. (e) Within-individual variances in unique zones traveled. (f) Within-individual variances in AP distance.

Differences in variances among species

The best fit model for unique zones visited was Model 4, which allowed both among and within-individual variances to vary across species. All other models were poorly supported (ΔDIC>8; Table 3). This indicates that both among- and within-individual

variances differed among species in open field trials. For AP distance, Models 3 and 4 fit comparably well (Table 3). Both of these models support differences among species in within-individual variances for AP distance. The difference between the models therefore suggests mixed support for species differences in among-individual variances for AP distance.

The monophyletic group of *G. assimilis, G. integer,* and *G. lineaticeps* exhibited higher among-individual variation for unique zones visited and lower among-individual variation for AP distance (Figure 1c & d). This monophyletic group also exhibited higher within-individual variation for both unique zones visited and AP distance than observed for *A. domesticus* and *G. sigillatus* (Figure 1e & f).

Table 3. DIC and ΔDIC values of model fit for AP distance and unique zones visited.

Model (variance constraints)	Behavior	DIC	ΔDIC
Model 1 (V_i = & V_w =)	Unique Zones Visited	8982.97	644.76
Model 2 (V_i ≠ & V_w =)	Unique Zones Visited	8420.69	82.48
Model 3 (V_i = & V_w ≠)	Unique Zones Visited	8346.44	8.23
Model 4 (V_i ≠ & V_w ≠)	**Unique Zones Visited**	**8338.21**	**0**
Model 1 (V_i = & V_w =)	AP Distance	8025.51	263.31
Model 2 (V_i ≠ & V_w =)	AP Distance	8010.04	247.84
Model 3 (V_i = & V_w ≠)	**AP Distance**	**7762.20**	**0**
Model 4 (V_i ≠ & V_w ≠)	AP Distance	7763.82	1.62

Differences in behavioral correlations among species

Among-individual behavioral correlations were of similar magnitude for *A. domesticus, G. assimilis, G. lineaticeps*, and *G. sigillatus* (0.3 : 0.5, Figure 2a, Table S3) while the correlation for *G. integer* was estimated to be slightly higher (0.66, Figure 2a, Table S3). Importantly, the lower bounds of the HPD intervals for *G. assimilis, G. integer, G. lineaticeps*, and *G. sigillatus* also overlapped with 0 (Figure 2, Table S3). This is perhaps unsurprising given the small sample sizes for *G. assimilis* and *G. lineaticeps*.

Behavioral correlations at the within-individual level ranged from 0.1 to 0.35 for *A. domesticus*, *G. sigillatus*, and *G. integer*, with *G. integer* having the lower bound of its HPD interval overlapping with 0 (Figure 2b). The overlapping of 0 indicates that behavioral plasticity might not be integrated in this species. Behavioral correlations at either level did not show obvious patterns relative to phylogeny and were not significantly different across species (Figure 2).

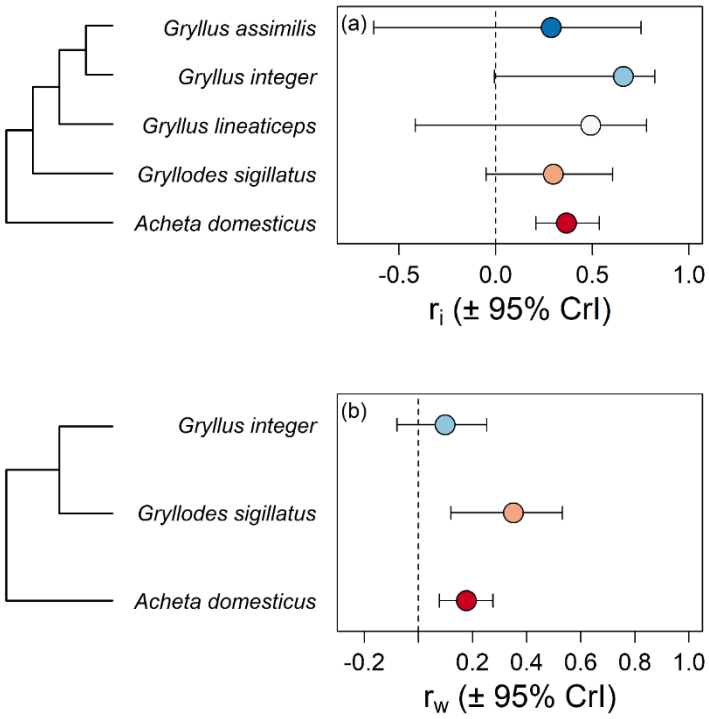

Figure 2. Species posterior-modal values with 95% HPD credibility intervals. (a) Among-individual behavioral correlations of unique zones visited and AP distance. (b) Within-individual differences of behavioral correlations of unique zones traveled and AP distance. Within-individual correlations for *G. assimilis* and *G. lineaticeps* were not calculated as behavior trials were not performed in close succession.

Discussion

Our results demonstrate that species differed in their exploratory behavior and response to cues of predator presence at all levels of variation but that behavioral syndromes were conserved across species. These results suggest an important influence of phylogenetic constraints on how behaviors evolve.

Species differed from one another in their average behaviors (Table 2), in a manner consistent with phylogenetic relationships. Specifically, the monophyletic group of *G. assimilis, G. integer,* and *G. lineaticeps* were generally similar in average unique zones visited (Figure 1a). In contrast, while average AP distance differed by species, it did not do so in a manner clearly concordant with phylogeny (Figure 1b). Indeed, phylogenetic signal, the proportion of variation attributable to the hierarchical pattern of relatedness among species, was higher for unique zones visited than for AP distance (Table S2). Interestingly, and relevant for future research, phylogeny explained considerably more variation in our measure of exploratory behavior—unique zones visited—than did among-individual variation, i.e. "animal personality" ($\tau = 0.15$ versus $\lambda = 0.27$; Table S2).

The species we examined also differed in among-individual variation in exploratory (unique zones visited) and predator response (AP distance) behaviors, again in a manner consistent with phylogenetic relationships (Figures 1c, d). Unfortunately, phylogenetic methods have been developed primarily with the goal of understanding differences in trait averages rather than trait (co)variances. Our comparisons of "personality" variation and syndromes among species are therefore based on the model comparison methods identifying the presence of species differences and subsequent qualitative comparisons of species level estimates. Nonetheless, the concordance between patterns of the magnitude of among-individual variation and the currently described phylogeny suggests phylogenetic constraints on the magnitude of "personality" variation. Of the five species, the monophyly of *G. assimilis, G. integer,* and *G. lineaticeps* exhibited the highest among-individual variation in unique zones visited and the lowest among-individual variation in AP distance (Figure 1c and d). While the expression of average behaviors and behavioral syndromes might be expected to exhibit phylogenetic signal, we did not expect this to be the case for among-individual variances. One possible explanation would be bottlenecks at more basal phylogenetic nodes leading to reduced genetic variation present in subsequent groupings. While this could explain the lower among-individual variation in AP distance for *G. assimilis, G. integer,* and *G. lineaticeps*, it does not explain that those same species exhibit higher within-individual variation for unique zones visited.

Differences observed in among-individual variation could also be attributable to selection differentially acting upon these species by reducing the additive genetic variation

present in a population or species (Mousseau and Roff 1987). Our results therefore suggest the possibility that exploratory behavior, for which unique zones visited is a proxy, has been under stronger selection for *A. domesticus* and *G. sigillatus* than for the other species. Importantly, because we do not know the strength and direction of selection acting on these phenotypes, the data presented here only suggests this possibility and cannot be used to distinguish between the effects of selection and drift for either behavior.

Alternative explanations for the observed differences in among-individual variances stem from differences across source populations and sampling of these populations. For example, individual *A. domesticus* used in this study were from a captive population where inbreeding could have reduced genetic variation over generations. This potentially explains the low among-individual variation the species shows for unique zones visited (Figure 1c) but is contradicted by the high among-individual variation in AP distance (Figure 1d). In contrast, *G. assimilis* and *G. lineaticeps* behavior was measured for field-caught individuals. If individuals of these species experienced different developmental environments from one another, we would predict higher among-individual variation in behavior because permanent environmental variation contributes, on average, 50% of the observed among-individual variation in behavior present in populations (Dochtermann et al. 2015). This explanation is not, however, supported: while *G. assimilis* and *G. lineaticeps* indeed showed high relative among-individual variation in unique zones visited, the same was not the case for AP distance (Figure 1c and d). Moreover, for both behaviors, *G. assimilis* and *G. lineaticeps* were very similar to *G. integer*, for which lab reared individuals were measured. To summarize, the conflicting patterns of among-individual variation observed between AP distance and unique zones visited prevents clear interpretation.

Estimated within-individual variances include variation from a variety of sources, including plasticity in response to short-term environmental variation and measurement error (Dingemanse et al. 2012, Berdal and Dochtermann 2019). In comparing species, if we assume measurement error is similar among species, differences in within-individual variation will primarily represent differences in plasticity. This short-term plasticity, also referred to as phenotypic flexibility (Piersma and Van Gils 2011), allows individuals to respond flexibly to an environment (Westneat et al. 2015). As was the case for among-individual variation and average unique zones visited, *G. assimilis, G. integer,* and *G.*

lineaticeps were grouped together and exhibited similar magnitudes of within-individual variation (Figure 1e and f). For both behaviors, this group exhibited considerably higher within-individual variation than observed for *A. domesticus* and *G. sigillatus*, differences supported by our model comparison results (Table 3). In other words, the *Gryllus* genus exhibited greater behavioral plasticity.

One possible explanation for this pattern is that our sample of *G. assimilis, G. integer,* and *G. lineaticeps* were of individuals either caught from the field or the direct offspring of field inseminated and subsequently captured individuals. In contrast, the population of *G. sigillatus* we sampled had been in captivity for around 75 generations and the population of *A. domesticus* was reared for production purposes for some undetermined but large number of generations. Consequently the differences in within-individual variation could be attributable to exposure to a frequently changing environment (Relyea 2001) in the case of *G. assimilis, G. integer,* and *G. lineaticeps* and the loss of plasticity in *A. domesticus* and *G. sigillatus*. This possibility could be assessed for crickets via experimental evolution with populations experiencing different levels of environmental heterogeneity.

With regard to behavioral correlations, Bell (2005) proposed two hypotheses for the expression of behavioral syndromes within a population relevant to the species level comparisons we performed. The first of these, the constraints hypothesis, chiefly attributes behavioral syndromes to the presence of pleiotropy, with the expression of genes affecting multiple behaviors. This hypothesis can be extended to other mechanistic connections constraining independent trait expression. Second, the adaptive hypothesis states that behavioral syndromes are the adaptive outcome of correlated selection. While pleiotropy and other mechanistic connections can evolve and be adaptive, syndromes attributable to the adaptive hypothesis are expected to respond more quickly to changes in selection (Roff 1997). Consequently, phylogenetic similarity in behavioral syndromes provides indirect support for the constraints hypothesis. Due to among-individual correlations not substantively differing among species (Figure 2), our results therefore support the constraints hypothesis, despite species differing in variances and average expressions of behaviors (Figure 1).

While a comparative approach has only rarely been used for examining behavioral variation, three particular studies are relevant to the interpretation of our results here.

First, Blankers et al. (2017) compared the phenotypic variances and (co)variances of seven calling traits of multiple cricket species (including *G. lineaticeps*, which was included in our study). These authors found that the phenotypic covariance matrices differed among cricket species. One of the major differences among species was in the magnitude of variation present in single traits (Blankers et al. 2017). This is consistent with our findings that variances of behaviors differed across species (Figure 1c-f). Unfortunately, these authors compared phenotypic (co)variances, which conflate among- and within-individual (co)variation (Dingemanse et al. 2012). Second, White et al. (2020) compared the among-individual covariance matrices of seven species of fish. Comparable to our results, these authors detected differences in the magnitude of among-individual behavioral variability and also found overall phylogenetic signal and similarity in how variation was expressed across multiple behaviors (White et al. 2020). Finally, Royauté et al. (2020) compared the expression of additive genetic (co)variance (i.e. **G** matrices) in behavior among four populations of *G. integer*. Similar to White et al. (2020) and our results presented here, Royauté et al. (2020) found differences in single trait variances and covariances but the overall structure of trait covariance was generally conserved across populations—indicating support for the constraints hypothesis.

More generally, our findings here suggest that behavioral correlations are phylogenetically conserved. Conserved trait correlations like those observed here constrain the divergence of populations and species (Schluter 1996). While the potential for such constraints has been speculated about for behaviors (Dochtermann and Dingemanse 2013), prior demonstrations of such have primarily focused on morphological traits (McGlothlin et al. 2018, Sztepanacz and Houle 2019) and chemical traits (Blows et al. 2004, Aguirre et al. 2014).

Jointly, our approach allowed us to determine whether there were differences in average behavior, "personality", behavioral plasticity, and behavioral syndromes among species. Our results demonstrate phylogenetic conservation of behavioral averages, behavioral variation, and behavioral syndromes. This finding is potentially surprising given that behavior is often assumed to be more flexible and labile than other types of traits (but see Zuk and Spencer 2020) and suggests an important role for phylogenetic constraints as

an alternative to the dominant adaptive explanations commonly employed when discussing animal personality and behavioral syndromes.

Literature Cited

2020. Guidelines for the treatment of animals in behavioural research and teaching. Animal Behaviour **159**:I-XI.

Aguirre, J., E. Hine, K. McGuigan, and M. Blows. 2014. Comparing G: multivariate analysis of genetic variation in multiple populations. Heredity **112**:21-29.

Armbruster, W., and K. Schwaegerle. 1996. Causes of covariation of phenotypic traits among populations. Journal of Evolutionary Biology **9**:261-276.

Barnett, A. G., N. Koper, A. J. Dobson, F. Schmiegelow, and M. Manseau. 2010. Using information criteria to select the correct variance–covariance structure for longitudinal data in ecology. Methods in Ecology and Evolution **1**:15-24.

Bell, A. M. 2005. Behavioural differences between individuals and two populations of stickleback (*Gasterosteus aculeatus*). Journal of Evolutionary Biology **18**:464-473.

Bell, A. M., S. J. Hankson, and K. L. Laskowski. 2009. The repeatability of behaviour: a meta-analysis. Animal Behaviour **77**:771-783.

Berdal, M. A., and N. A. Dochtermann. 2019. Adaptive alignment of plasticity with genetic variation and selection. Journal of Heredity **110**.

Blankers, T., D. A. Gray, and R. M. Hennig. 2017. Multivariate phenotypic evolution: Divergent acoustic signals and sexual selection in Gryllus field crickets. Evolutionary Biology **44**:43-55.

Blows, M. W., S. F. Chenoweth, and E. Hine. 2004. Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. American Naturalist **163**:329-340.

Boake, C. R. B. 1989. Repeatability - Its Role in Evolutionary Studies of Mating-Behavior. Evolutionary Ecology **3**:173-182.

Brommer, J. E., and B. Class. 2017. Phenotypic correlations capture between-individual correlations underlying behavioral syndromes. Behavioral Ecology and Sociobiology **71**:50.

Bucklaew, A., and N. Dochtermann. 2020. The effects of exposure to predators on personality and plasticity. Ethology.

Carter, A. J., W. E. Feeney, H. H. Marshall, G. Cowlishaw, and R. Heinsohn. 2013. Animal personality: what are behavioural ecologists measuring? Biological Reviews **88**:465-475.

Congdon, P. 2006. Bayesian model choice based on Monte Carlo estimates of posterior model probabilities. Computational statistics & data analysis **50**:346-357.

Dingemanse, N. J., and N. A. Dochtermann. 2013. Quantifying individual variation in behaviour: mixed-effect modelling approaches. Journal of Animal Ecology **82**:39-54.

Dingemanse, N. J., and N. A. Dochtermann. 2014. Individual behaviour: behavioural ecology meets quantitative genetics.*in* A. Charmantier, D. Garant, and L. E. B. Kruuk, editors. Quantitative genetics in the wild. Oxford University Press.

Dingemanse, N. J., N. A. Dochtermann, and S. Nakagawa. 2012. Defining behavioural syndromes and the role of "syndrome" deviation in understanding their evolution. Behavioral Ecology and Sociobiology **66**:1543-1548.

Dochtermann, N. A. 2011. Testing Cheverud's conjecture for behavioral correlations and behavioral syndromes. Evolution **65**:1814-1820.

Dochtermann, N. A., and N. J. Dingemanse. 2013. Behavioral syndromes as evolutionary constraints. Behavioral Ecology **24**:806-811.

Dochtermann, N. A., and D. A. Roff. 2010. Applying a quantitative genetics framework to behavioural syndrome research. Philosophical Transactions of the Royal Society B-Biological Sciences **365**:4013-4020.

Dochtermann, N. A., T. Schwab, M. A. Berdal, J. Dalos, and R. Royaute. 2019. The heritability of behaviour: a meta-analysis. Journal of Heredity **in press**.

Dochtermann, N. A., T. Schwab, and A. Sih. 2015. The contribution of additive genetic variation to personality variation: heritability of personality. Proceedings Of The Royal Society B-Biological Sciences **282**:20142201.

Garamszegi, L. Z., G. Marko, and G. Herczeg. 2012. A meta-analysis of correlated behaviours with implications for behavioural syndromes: mean effect size, publication bias, phylogenetic effects and the role of mediator variables. Evolutionary Ecology **26**:1213-1235.

Garamszegi, L. Z., G. Marko, and G. Herczeg. 2013. A meta-analysis of correlated behaviors with implications for behavioral syndromes: relationships between particular behavioral traits. Behavioral Ecology **24**:1068-1080.

Hadfield, J., and S. Nakagawa. 2010. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. Journal of Evolutionary Biology **23**:494-508.

Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. Journal of Statistical Software **33**:1-22.

Hine, E., K. McGuigan, and M. W. Blows. 2014. Evolutionary Constraints in High-Dimensional Trait Sets. The American Naturalist **184**:119-131.

McGlothlin, J. W., M. E. Kobiela, H. V. Wright, D. L. Mahler, J. J. Kolbe, J. B. Losos, and E. D. Brodie III. 2018. Adaptive radiation along a deeply conserved genetic line of least resistance in Anolis lizards. Evolution Letters **2**:310-322.

Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity **59**:181-197.

Nakagawa, S., and E. S. Santos. 2012. Methodological issues and advances in biological meta-analysis. Evolutionary Ecology **26**:1253-1274.

Nakagawa, S., and H. Schielzeth. 2010a. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biological Reviews **85**:935-956.

Nakagawa, S., and H. Schielzeth. 2010b. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biological Reviews **85**:935-956.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature **401**:877-884.

Phillips, P. C., and S. J. Arnold. 1989. Visualizing multivariate selection. Evolution **43**:1209-1222.

Piersma, T., and J. Drent. 2003. Phenotypic flexibility and the evolution of organismal design. Trends in Ecology & Evolution **18**:228-233.

Piersma, T., and J. A. Van Gils. 2011. The flexible phenotype: a body-centred integration of ecology, physiology, and behaviour. Oxford University Press.

Relyea, R. A. 2001. Morphological and behavioral plasticity of larval anurans in response to different predators. Ecology **82**:523-540.

Roff, D. A. 1997. Evolutionary Quantitative Genetics. Chapman and Hall, New York.

Royauté, R., and N. A. Dochtermann. 2017. When the mean no longer matters: developmental diet affects behavioral variation but not population averages in the house cricket (*Acheta domesticus*). Behavioral Ecology **28**:337-345.

Royauté, R., and N. A. Dochtermann. 2020. Comparing ecological and evolutionary variability within datasets. EcoEvoRxiv.

Royauté, R., C. Garrison, J. Dalos, M. A. Berdal, and N. A. Dochtermann. 2019. Current energy state interacts with the developmental environment to influence behavioural plasticity. Animal behaviour **148**:39-51.

Royauté, R., K. Greenlee, M. Baldwin, and N. A. Dochtermann. 2015. Behaviour, metabolism and size: phenotypic modularity or integration in *Acheta domesticus*? Animal Behaviour **110**:163-169.

Royauté, R., A. Hedrick, and N. A. Dochtermann. 2020. Behavioural syndromes shape evolutionary trajectories via conserved genetic architecture. Proceedings of the Royal Society B **287**:20200183.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution **50**:1766-1774.

Spiegelhalter, D., A. Thomas, N. Best, and D. Lunn. 2003. WinBUGS user manual. version.

Stirling, D. G., D. Reale, and D. A. Roff. 2002. Selection, structure and the heritability of behaviour. Journal of Evolutionary Biology **15**:277-289.

Sztepanacz, J. L., and D. Houle. 2019. Cross-sex genetic covariances limit the evolvability of wing-shape within and among species of Drosophila. Evolution **73**:1617-1633.

Team, R. C. 2018. R: A language and environment for statistical computing.

Van de Pol, M., and J. Wright. 2009. A simple method for distinguishing within-versus between-subject effects using mixed models. Animal behaviour **77**:753.

Weissman, D. B., and D. A. Gray. 2019. Crickets of the genus Gryllus in the United States (Orthoptera: Gryllidae: Gryllinae). Zootaxa **4705**:1-277.

Westneat, D. F., M. I. Hatch, D. P. Wetzel, and A. L. Ensminger. 2011. Individual variation in parental care reaction norms: integration of personality and plasticity. The American Naturalist **178**:652-667.

Westneat, D. F., J. Wright, and N. J. Dingemanse. 2015. The biology hidden inside residual within-individual phenotypic variation. Biological Reviews **90**:729-743.

White, S. J., D. J. Pascall, and A. J. Wilson. 2020. Towards a comparative approach to the structure of animal personality variation. Behavioral Ecology **31**:340-351.

Whitman, D., and A. Agrawal. 2009. What is phenotypic plasticity and why is it important? Phenotypic plasticity of insects: mechanisms and consequences:1-63.

Zuk, M., and H. G. Spencer. 2020. Killing the Behavioral Zombie: Genes, Evolution, and Why Behavior Isn't Special. Bioscience **70**:515-520.

Supplemental Materials

Table S1. Fixed effects coefficients for Model 4 (Table 3). The intercept estimate is for *Acheta domesticus* females (fixed effect coefficients are contrasts versus these values).

AP Distance (square-root transformed)

	posterior mean	95% credibility interval		effective sample size	pMCMC
		lower	upper		
Intercept	16.97	16.25	17.66	1000	<0.001
Gryllus assimilis	-1.44	-4.18	1.22	1000	0.298
Gryllus integer	3.67	2.31	5.18	1000	<0.001
Gryllus lineaticeps	-0.52	-4.00	2.95	1000	0.782
Gryllodes sigillatus	7.43	5.69	9.17	1000	<0.001
Temp2	0.65	0.34	0.96	1000	<0.001
SexM	-0.40	-1.44	0.72	1000	0.454
Mass (w/in individual centered)	0.72	-0.77	2.29	1000	0.344
Mass (b/w individual centered)	0.02	-0.76	0.98	800.8	0.972

Unique zones visited

	posterior mean	95% credibility interval		effective sample size	pMCMC
		lower	upper		
Intercept	2.929	2.869	2.984	1000	<0.001
Gryllus assimilis	-1.227	-1.757	-0.628	1000	<0.001
Gryllus integer	-0.728	-0.891	-0.526	1000	<0.001
Gryllus lineaticeps	-0.593	-0.979	-0.219	1000	0.004
Gryllodes sigillatus	0.282	0.160	0.411	1000	<0.001
Temp2	0.043	0.016	0.070	815.1	0.002
SexM	0.111	0.010	0.201	1000	0.022
Mass (w/in individual centered)	-0.022	-0.175	0.135	1000	0.792
Mass (b/w individual centered)	0.120	0.043	0.212	899.8	0.006

551

Table S2. Variance estimates (posterior modes with 95% credibility intervals) for models including phylogenetic structure as a random effect. Models were fit with temperature (centered), sex, and mass (within and between individual centered) as fixed effects. Phylogeny was modeled according to the trees shown in Figure 1 and with uniform branch lengths. Subject was also included as a random effect. Variance ratios are presented as unadjusted ratios; that is, variance due to fixed effects is included in the denominator. Ratios for unique zones include the distribution specific variance (DSV) in the denominator. λ and τ correspond to unadjusted phylogenetic signal and unadjusted repeatabilities respectively.

	Variance estimate (95% CrI)	Variance ratios* (95% CrI)
AP Distance		
Phylogeny	12.73 (2.10 : 106.61)	λ : 0.16 (0.05 : 0.68)
Subject	13.34 (10.21 : 17.66)	τ : 0.19 (0.07 : 0.27)
Fixed Effects	1.10 (0.32 : 2.60)	0.01 (0 : 0.04)
Residual	38.34 (33.96 : 41.50)	0.50 (0.24 : 0.71)
Unique Zones Visited		
Phylogeny	0.12 (0.03 : 1.07)	λ : 0.27 (0.09 : 0.72)
Subject	0.11 (0.07 : 0.13)	τ : 0.15 (0.05 : 0.23)
Fixed Effects	0.01 (0 : 0.02)	0.01 (0 : 0.03)
Residual	0.26 (0.23 : 0.31)	0.48 (0.16 : 0.58)
DSV**	0.06 (0.03 : 0.11)	*NA*

* while the ratios for any single MCMC estimate will sum to 1, the posterior modes can sum to other values due to uncertainty across the MCMC chain

** estimated as $ln\left(\frac{1}{exp(\beta_0)}+1\right)$ following Nakagawa & Schielzeth (2010)

552

Table S3. Among- and within-individual correlations by species. Correlation estimates are posterior modes and are presented along with 95% credibility intervals (CrI).

Species	Among-individual correlation (95% CrI)	Within-individual correlation (95% CrI)
Gryllus assimilis	0.37 (0.21 : 0.54)	*NA*
Gryllus integer	0.29 (-0.63 : 0.75)	0.10 (-0.08 : 0.25)
Gryllus lineaticeps	0.66 (-0.01 : 0.82)	*NA*
Gryllodes sigillatus	0.49 (-0.42 : 0.78)	0.35 (0.12 : 0.53)
Acheta domesticus	0.3 (-0.05 : 0.6)	0.18 (0.08 : 0.27)

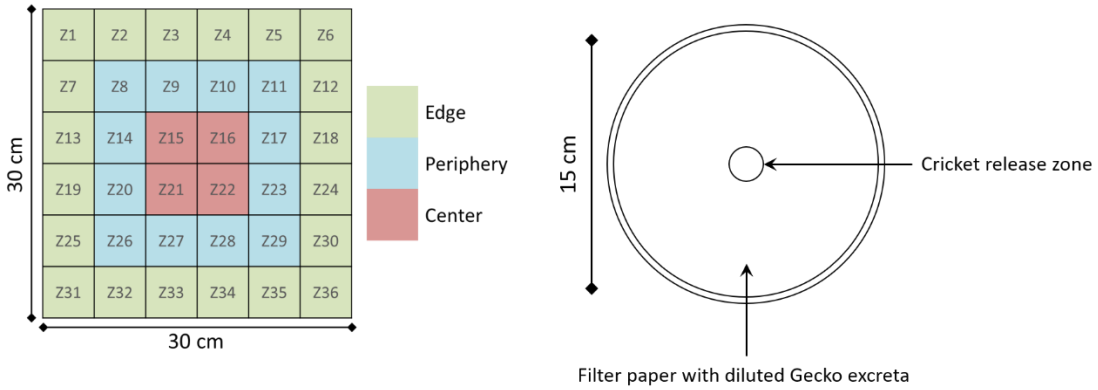

Figure S1. Schematics of the open field (left) and predator cue arenas (right). The open field arenas were subdivided into 36 unique "zones" during video processing. For the anti-predator response trials the cricket was introduced, under a container, to the center point. This cricket release zone did not have predator cues present.