



Functional Ecology

Multidimensional trait morphology predicts ecology across ant lineages

Christine E. Sosiak^{1*}, Phillip Barden^{1,2}

¹Federated Department of Biological Sciences, New Jersey Institute of Technology, Newark, NJ 07102.

²Division of Invertebrate Zoology, American Museum of Natural History, New York, NY 10024.

*Corresponding author: ces43@njit.edu

Acknowledgements

We thank Christine Johnson and Christine Lebeau as well as Morgan Hill and Andrew Smith and for facilitating access to specimens and imaging equipment, respectively, at the American Museum of Natural History (AMNH); Stefan Cover and David Lubertazzi for specimen access and hospitality at the Museum of Comparative Zoology (MCZ); Eugenia Okonski and Ted Schultz for their hospitality and flexibility in providing specimen access at the Smithsonian National Museum of Natural History (NMNH). We also thank Simon Garnier for thoughtful discussion on statistical approaches and Ian Hays for providing illustrations that contributed to figure preparation. A portion of this work was funded by an Arthur James Boucot Research Grant from the Paleontological Society and startup funds from the New Jersey Institute of Technology.

Author contributions

CS and PB conceived the ideas and designed methodology, analysed the data, drafted the manuscript; CS collected the data. Both authors contributed critically to the drafts and gave final approval for publication.

Data availability statement

All data are available in the Dryad Digital Repository (doi: <https://doi.org/10.5061/dryad.kh1893243>).

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/1365-2435.13697](https://doi.org/10.1111/1365-2435.13697)

This article is protected by copyright. All rights reserved

MS CHRISTINE SOSIAK (Orcid ID : 0000-0002-9057-8636)

Article type : Research Article

Section: Evolutionary Ecology

Editor: Dr Timothy Higham

Abstract

1. Understanding the link between ecology and morphology is a fundamental goal in biology. Ants are diverse terrestrial organisms, known to exhibit ecologically-driven morphological variation. While relationships between individual traits and ecologies have been identified, multidimensional interactions among traits and their cumulative predictive power remain unknown. Because selective pressures may generate convergent syndromes spanning multiple traits, we applied multivariate analyses across a wide sampling of taxa to assess ecomorphological variation in an integrative context.

2. How well does morphology predict ecology? Moreover, are there quantitatively-supported ant ecomorphs? We investigated the links between trait morphology and ecology by assembling a morphometric dataset spanning over 160 species within 110 genera. Because ants occupy a wide range of ecologies, we compiled natural history data on nesting microhabitat, foraging stratum, and functional role into 35 defined niche combinations. This tripartite ecological classification and our morphological dataset were optimized under dimension reduction techniques including Principal Component Analysis, Principal Coordinate Analysis, Linear Discriminant Analysis, and Random Forest supervised machine learning.

3. Our results describe ant ecomorphospace as comprising regions of shared, generalized morphology as well as unique phenotypic space associated with specialized ecologies. Dimension reduction and model-based approaches predict ecology with 77-85% accuracy and Random Forest analysis consistently outperforms LDA. While accounting for shared ancestry, we found eye, antennal scape, and leg morphology to be most informative in differentiating among ecologies. We also note some heterogeneity between trait significance in each ecological aspect (nesting niche, foraging niche, functional role). To increase the utility of ecomorphological classification we simplified our 35 observed niche combinations into 10 ecomorph syndromes, which were also predicted by morphology. The predictive power of these machine learning methods underscores the strong role that ecology has in convergently shaping overall body plan morphology across ant lineages. We include a pipeline for predictive ecomorphological modeling using morphometric data, which may be expanded with additional specimen-based and natural history data.

Keywords: ants, ecomorphology, machine learning, morphospace, trait morphology

Introduction

The relationship between form and function is a striking feature of organismal biology; across taxa, morphology may clearly reflect habitat or functional role within an ecosystem. The repeated evolution of trait morphology due to niche occupation is well established in disparate groups of fish, lizards, mammals and arthropods: more elongate caudal fins in fishes correlate with more pelagic lifestyles (Mihalitsis and Bellwood 2019); geckos that inhabit flatter terrains show reduced adhesive toe pads (Collins et al. 2015); ambush canid predators have more flexible forelimbs than pursuit canid predators (Janis and Figueirido 2014); and cave-dwelling amphipods are larger with longer legs than smaller, stubbier stream-dwelling amphipods (Trontelj et al. 2012). Morphological specialization may arise independently even within closely related taxa, indicating the powerful force of environmental selective pressures on species morphology (Losos et al. 1998).

Morphological adaptation to an ecological niche may represent an ecomorph: an overall form that is specialized for its ecology. Ecomorphs were formalized by Williams (1972) who found that *Anolis* lizard communities reflected assemblages of phenotypes that shared the same niche with similar morphology and behaviour. Since then, *Anolis* has typified the concept of the ecomorph, and the genus has become a model system for understanding convergence and repeatability in evolutionary patterns (Losos et al. 1998; Losos 1992; Langerhans et al. 2006; Dickson et al. 2017). This concept is widely applied to mammals (Saunders and Barclay 1992; Ghazali et al. 2017; Figueirido et al. 2019; Leonard et al. 2007); arthropods (Parker and Owens 2018; Barton et al. 2011; Malcicka et al. 2016; Mugleston et al. 2018); gastropods (Marquez et al. 2015); fish (Gerry et al. 2011); and other reptiles (Sanders et al. 2013). Ecomorphs are a valuable lens for understanding fundamental ecological and macroevolutionary concepts in biology. For example, comparisons at the ecomorph level have revealed insight into community assembly on islands (Losos 1992; Gillespie 2004), temporal stasis in vertebrates (Dickson et al. 2017) and iterative patterns of faunal associations (Figueirido et al. 2019).

With over 14,000 species (Bolton 2020), ants are highly diverse, globally ubiquitous, and fill most post-producer ecological niches, representing a significant opportunity for understanding ecomorphology. Their ecological breadth adds an important dimension for increased trophic and microhabitat granularity. Ants run the gambit from obligate leaf harvesting fungus growers and subterranean eyeless predators to nomadic army ants and arboreal omnivorous twig nesters (Figure 1) (Hölldobler & Wilson 1990).

Some individual ant traits are known to correlate with ecological niche within a given habitat or region. Ant leg length, in particular femur length, tends to be greater in species that live in open habitats compared to denser, more complex habitats (Kaspari and Weiser 1999; Yates and Andrew 2011; Weiser and Kaspari 2006). Leg elongation may be reflective of the need to move faster and more efficiently for resource acquisition or predator evasion, though the pattern is not universal. Femur length has also been shown to be invariant or even positively correlated with habitat complexity (Guilherme et al. 2019; Parr et al. 2003). Eye size is also correlated both with habitat and with trophic level; ants living in subterranean or leaf litter habitats exhibit smaller eyes than surface-foraging ants, and predators typically possess smaller eyes than omnivores (Weiser and Kaspari 2006; Gibb et al. 2015; Guilherme et al. 2019). Overall

body size reflects life history habits: ants in more complex habitats tend to be smaller than ants in open spaces (Yates et al. 2014; Weiser and Kaspari 2006). The relationship between functional role and body size is less clear; some findings suggest predators are on average larger than omnivores (Gibb et al. 2015), while others suggest they are typically smaller than omnivores (Weiser and Kaspari 2006), possibly as a function of taxon sampling. While some traits are known to correlate with certain ecologies, in this study we expand upon these findings by explicitly assessing multiple traits and aspects of ecology simultaneously using a global morphometric dataset of ants. Our goal is to elucidate how combinations of traits, or, syndromes, relate to ecology across ecosystems and lineages.

Here, we selected 17 morphological traits – 15 previously hypothesized or known to relate to microhabitat and trophic level – to evaluate links between form and function across ant lineages. We measured species from six continents and all major biomes, from virtually all subfamilies of ants to build a broad global multivariate dataset to address the following questions: Does morphology predict nesting, foraging, and functional role niche occupation across ants? What traits are most implicated in ecological niche occupations? Are there quantitatively supported ant ecomorphs?

Materials & Methods

Taxa, ecological niche binnings, and ecomorph syndromes

Our sampling includes at least one species from each accessible ant subfamily; we sampled all subfamilies except Apomyrminae and Martialinae, which are extremely rare in museum collections. In total we sampled 15 subfamilies, 113 genera, and 167 species, measuring 320 specimens in total (see Data Accessibility section, Dryad Digital Repository). When possible, we sampled three specimens per species. When insufficient specimens were available, we measured as many conspecific specimens as were present in collections. Many species are polymorphic or have specialized castes for specific ecological functions: in these situations, we either sampled the media caste for polymorphic species or sampled the nonspecialized workers for species with specialized castes. We included specimens from the New Jersey Institute of

Technology (NJIT), the American Museum of Natural History (AMNH), the Smithsonian National Museum of Natural History (NMNH), and the Harvard Museum of Comparative Zoology (MCZ).

All specimens were assigned a binning from each of three ecological niche aspect categories: functional role (six binnings), nesting niche (five binnings), and foraging niche (four binnings) (Figure 2, Table 1). We selected these niche aspects as they are broadly applicable: while other key abiotic or biotic factors such as climate or community composition may also influence morphology, our selected aspects are generalizable across widely disparate communities of species, allowing for global ecomorph definitions. Moreover, to achieve wide taxon sampling, our morphometric dataset was confined to historical museum collections, which frequently do not retain specimen-level information relating to community structures or precise localities. Binning determinations were based on surveys of the literature (Appendix 1, AntWiki). A list of natural history binning data sources used in the study are provided in the Supporting Information section. If an aspect of a species' niche was unclear, the species was assigned an "unknown" binning and excluded from all further analyses for that niche aspect. We attempted to sample representatives from every known ecological niche that ants inhabit; however, due to the extremely diverse nature of ants, the sampling is broadly comprehensive but likely not universal. In particular, given our reliance on museum specimens, we may be missing diversity from geographic areas that have been historically neglected as well as ecologies that are difficult to sample (e.g. subterranean taxa).

Additionally, we collapsed the 35 observed combinations of niche binnings across all niche aspects into 11 ecomorph syndromes (Figure 2, Table S1). These simplified syndromes were defined after preliminary agnostic analyses suggested extensive morphological overlap between some niche combinations. While we defined 11 ecomorph syndromes, one syndrome (subterranean omnivore) was represented by only one species and three specimens. This species, *Acropyga oreithauma*, represents a clearly defined ecomorph syndrome, but we did not have enough specimens for adequate statistical comparisons and so removed it from subsequent analyses.

Our analyses are constrained by the accuracy of natural history information; in some cases, species may be plastic in their behaviour, or their ecological niche may not be fully known. While our models do not explicitly incorporate this uncertainty, this may be explored in the future through sensitivity analyses with alternate niche assignments for ambiguous taxa. We do, however, incorporate some behavioral plasticity in our collapsed ecomorph syndromes; such as “ground or leaf litter-nesting epigaeic omnivores”, or the synonymization of specialist and generalist predators into a general predator class (Table S1).

Measurements

Morphometric sampling included linear measurements of 12 cephalic traits and 5 post-cephalic traits (Table S2). Fifteen selected traits have been previously correlated with ecology. These traits include metafemur length (Yates et al 2014), eye size (Weiser and Kaspari 2006; Gibb et al 2015), body size (Gibb et al 2015; Salas-Lopez 2017), and mandible length (Weiser and Kaspari 2006). Our initial sampling included five additional metrics (tarsal claw length, tarsal claw width, mandible articulation angle, ventral head length, and ventral pronotal length); however, these were excluded from data collection after preliminary analyses indicated overlap with other metrics, variation due to measurement artifacts, or lack of significance across all niche aspects.

All measurements were conducted on point-mounted specimens under stereo microscopy. Including raw measurements in dimension reduction techniques such as principal component analysis can result in body size driving the overwhelming majority of variation in a dataset, masking other potentially important contributors. To explore the impact of body size, we created two datasets for analyses: a dataset comprising raw measurements and a size-corrected dataset using only ratios (Table S3).

Data analysis

Our analyses spanned three broad categories: assessments of relationships between individual traits and ecologies; multivariate analyses of traits in the context of ecology; and multivariate predictive modeling of ecology from trait data. To agnostically determine whether or not individual traits were correlated with ecology, we performed analyses of variance

(ANOVAs) on trait data and functional role, nesting niche, and foraging niche binnings. We accounted for the potential effect of phylogenetic relatedness on ANOVAs by performing phylogenetic generalized least squares (PGLS) on our trait data using a published phylogeny of ant genera (Blanchard and Moreau 2017). To evaluate which traits drove maximum variation and to visualize ecomorphospaces (representations of taxa and their ecologies based on morphometric data) among groups, we ran principal component (PCAs) and principal coordinate analyses (PCoAs). While both PCAs and PCoAs are dimension reduction techniques, PCoA is unique in that it is derived from distance matrices and so can include inapplicable or missing data. To determine which traits drive maximum separation between groups, and whether traits can delineate distinct ecomorphospaces, we conducted linear discriminant analysis (LDAs), which reduces dimensions in the context of predefined binnings. We further tested the accuracy of LDA group separation by comparing posterior probabilities of classification. Finally, we evaluated the predictability of ecological niche from morphometric trait data using a supervised machine learning algorithm, Random Forest (RF) analysis. All analyses were performed in R v.3.6.3 (Team RC 2020).

Evaluating traits correlated with ecological niche

Our first aim was to test if traits differed significantly between ecological niche binnings and if the differences we observed were due to ecology or shared ancestry. We evaluated our dataset using analysis of variance (ANOVA), assessing variation of morphology across binnings in functional role, nesting niche, foraging niche, and ecomorph syndrome separately. We tested all data for normality of residuals and homoscedasticity of variances using Levene's test; if the data were nonnormal, we used a Kruskal-Wallis test in place of an ANOVA. If the ANOVA indicated a significant difference between groups, we used a post-hoc Tukey's test for pairwise comparisons to evaluate which groups differed.

Significant differences between groups may reflect trait similarity due to shared ancestry rather than shared ecology (Felsenstein 1985). We analyzed morphometric data under phylogenetic generalized least squares (PGLS) to account for the non-independence of data observations due to phylogenetic history (Martins and Hansen 1997). We evaluated both Brownian motion (BM) and Ornstein-Uhlenbeck (OU) evolutionary models; for each trait we

Accepted Article

chose the best-fitting model based on comparison of Akaike Information Criterion (AIC) values. Phylogenetic information for our analysis was derived from Blanchard and Moreau's comprehensive ant phylogeny (2017). A total of 90 taxa overlapped between Blanchard and Moreau's phylogeny and our dataset, including only species with known ecological niche binings. To perform PGLS, we pruned Blanchard and Moreau's phylogenetic tree and our dataset to these applicable matching taxa (see Data Accessibility). When trimming our morphological trait dataset, we matched species to the tree where possible and matched genera where there were not exact species matches, contingent on if the congeners occupied the same ecological niche; if the congeners did not occupy the same ecological niche, we did not include them. When there were multiple specimens per tree-matched species, we averaged the morphological trait measurements across all specimens to produce a single, averaged specimen.

We tested significant differences in both raw and size-corrected traits between groups using ANOVA and PGLS. In both cases, we used the taxonomically trimmed dataset (90 taxa matching Moreau & Blanchard 2017) to allow for more accurate comparisons between the results of the ANOVA and the results of the PGLS. ANOVA and PGLS were implemented in packages 'ape' (Paradis and Schliep 2019), 'phytools' (Revell 2012), 'nlme' (Pinheiro et al. 2018), and 'geiger' (Harmon et al. 2008).

Evaluating ecomorphospace structure

We used principal component analysis (PCA) to evaluate ecomorphospaces produced when considering maximum variation within the entire dataset, visualizing the variation amongst all species. However, because some species that lack certain traits (e.g. many doryline ants are eyeless, rendering eye position metrics impossible), we also performed principal coordinate analysis (PCoA), which illustrates maximum disparity within the dataset and can be performed with missing or inapplicable data. Thus, we were able to capture the full range of ecomorphospaces produced by the dataset. We extracted all principal components with associated eigenvalues to determine which traits drove the majority of variation within the dataset.

To evaluate between-group variation rather than total variation, we performed multi-group linear discriminant analysis (LDA). This analysis requires that specimens have group

assignments, and maximizes variation between groups, rather than variation in the dataset as a whole. To test whether the LDA is successful in separating groups, we split our data into a training (80%) and testing (20%) dataset, randomly assigned. The model is trained using the larger dataset, and with the testing dataset taxa are assigned to their most likely group using posterior probabilities based on the trait value's Mahalanobis distance to the group mean. We performed PCA, PCoA, and LDA using binnings from the functional role, foraging niche, nesting niche, and ecomorph syndromes using raw measurements, size-corrected measurements, and a subset of only measurements shown to be significant under PGLS. All dimension reduction techniques were implemented in R packages 'corrplot' (Wei et al. 2017), 'FactoMineR' (Lê, et al. 2008), 'vegan' (Oksanen et al. 2013), 'labdsv' (Roberts 2007), 'caret' (Kuhn 2015), 'MASS' (Ripley et al. 2019), and 'ade4' (Dray and Dufour 2007).

Quantitatively predicting ecological niche from morphology

To evaluate whether species can be clearly delimited into ecological niche based on morphology, we conducted Random Forest analysis, a supervised machine learning algorithm. Random forest models are based on a series of decision trees used to build a 'consensus tree', which partitions morphospace according to predefined ecological niche binnings. Random Forest was implemented in the R package 'randomForest' (Liaw and Wiener 2018). We constructed RF models for functional role, nesting niche, foraging niche, and ecomorph syndrome separately, using the raw, size-corrected, and phylogenetically-independent datasets. Our parameters were selected based on initial sensitivity tests: $mtry=4$ and $ntree=5000$. To visualize posterior probabilities of our Random Forest model classification, we constructed t-SNE plots for each model, based on model votes for each binning or syndrome. We used a perplexity of 30 and a maximum number of iterations of 500. t-SNE plots were constructed using the R package *Rtsne* (Krijthe 2015.).

To assess consistency and predictability of trait-niche matches, we constrained tree size by limiting the number of terminal nodes to 20, 50, 100, and 500. If predictive accuracy remains high despite limited terminal node number, relationships between traits and niches are less likely to be the result of model overfit. Conversely, if predictive accuracy is only high at larger tree sizes, it may be driven by a highly complicated relationship between trait form and ecological

function which is not necessarily consistent or predictive. To evaluate comparative predictive power of traits in our RF models, we tested for mean decrease in accuracy when a trait was omitted. We compiled predictions from our LDA and RF models into sensitivity tables to assess accuracy of predictions by method.

Results

Assessing measured trait relevance (ANOVAs and PGLS)

While nearly all raw measured traits were found to be significantly different across foraging niche, nesting niche, and ecomorph syndrome binnings, few traits appear to be implicated in functional role binnings (Figure 3A-C). When accounting for phylogeny, we find most measured traits remain implicated in niche aspects (Table S4). Comparatively, size-corrected ratios were less variable between binnings. Most ratios were still different between binnings before accounting for phylogeny, but the number of significantly different ratios was reduced afterwards (Table S5). Typically, procoxal-to-body-size proportion, eye-to-head-length proportion, pronotal expansion, pronotal flattening, scape-to-body-size proportion, and metafemur-to-body-size proportion were all significantly different between binnings even after accounting for phylogeny, regardless of the niche aspect examined (Figure S9).

Considering functional role, specialist predators typically exhibit smaller procoxae and metafemura than omnivores, reflecting overall shorter leg length in comparison with the body (Figures S2, S6). Predators also tend towards longer and straighter mandibles than phytophagivores and omnivores. Specialist predators typically have shorter antennae in comparison with the rest of the head, while omnivores, fungivores, and generalist predators have longer antennae; phytophagivores typically have extremely short antennae. Eye position on the anteroposterior axis of the head capsule is variable, though typically predators have their eyes positioned somewhat more anteriorly than omnivores. Looking at nesting niche, lignicolous and subterranean nesters all have much shorter legs, both procoxa and metafemur, than ground-nesting ants, with leaf litter ants more variable overall (Figures S3, S7). Arboreal ants, both lignicolous and carton-nesting, have much larger eyes than subterranean, leaf litter and ground-

Accepted Article

nesting ants. Lignicolous ants are also often distinguished by a flatter, broader mesosoma in comparison with ground-nesting ants, though this is not universal; and lignicolous ants have much shorter scapes than virtually all other groups, although subterranean ants typically possess shorter scapes as well. With respect to foraging niche, epigaeic ants overall have much longer legs, both procoxa and metafemur, than do leaf litter, arboreal, or column-raiding ants (Figures S1, S5). Scape length is also greater on average in epigaeic ants. Eye size is highly variable within groups but is greater among arboreal and epigaeic ants – arboreal ant eyes are the largest on average; column-raiding and leaf litter ants tend towards smaller compound eyes or eyelessness. Illustrating the overlap between lignicolous-nesting and arboreal-foraging species, arboreal-foraging ants are often distinguished by the broader, flatter mesosoma that typifies lignicolous ants.

Dimension reduction analyses

Agnostic visualization of niche structure

Both principal component analysis (PCA) and its corollary principal coordinate analysis (PCoA) produced ecomorphospace with a high degree of overlap, especially when using raw measurements alone (Figures 4, S10, S11). The typical pattern produced by raw measurements was an area of high density where most of the specimens plotted, with some ecomorphs radiating into their own unique ecomorphospace. When using only raw measurements, no ecomorph was completely distinct, though typically specialist predators and omnivores plotted further away from one another.

For our raw measurement dataset, principal component 1 was body size, comprising 87.2% of the variance; principal component 2 was the overall size of the mandibles, comprising 5.2% of the variance. The first five principal components explain over 97% of the variance, each explaining >1% of the variance (Table S6). The pattern was similar with principal coordinate analysis: the first five principal coordinates explain over 95% of the variance, with the first principal coordinate comprising 83.8% of the variance and the second principal coordinate comprising 5.0% (Table S8).

Using only size-corrected ratio measurements produced a slightly different pattern. There was no central area where most specimens plotted, but a greater degree of disparity throughout the ecomorphospace (Figures S12, S13). While again, no ecomorphospaces separated clearly and there was a high degree of overlap, specialist predators tended to separate more distinctly from omnivores. Using ratio measurements, individual principal components accounted for much less of the overall variance; principal component 1 comprised only 33% of the variance, while principal component 2 comprised 24.3%. Over 90% of the variance is explained by the first seven principal components (Table S7). Most of the variance is explained by ratios concerning the scape, metafemur, and mandibles; principal component 1 was driven mostly by scape and metafemur proportion, while principal component 2 was mostly mandibular proportions. Again, the pattern was similar using principal coordinate analysis with size-corrected ratio measurements: the first principal coordinate comprised 34.1% of the variance and the second comprised 23.1% (Table S9).

Visualization of discrete niche structure

Our linear discriminant analysis produced ecomorph clusters with a variable degree of overlap, contingent on the ecological niche aspect being plotted and raw vs. size-corrected ratio measurements (Figures 4, 5, S14). When analyzing functional role, we found that again specialist predators and omnivores plotted separately, with phytophagivores also plotting out distinctly; while lignicolous and ground-nesting ants were most distinct; and arboreal, epigaeic, and leaf litter-foraging ants all partially plotted in distinct ecomorphospaces. Similar to the PCAs and PCoAs, we recover a high-density area of ecomorphospace overlap, with some ecomorphospaces radiating out into unique topography.

Across ecological niche aspects, the most predictive traits for ecological niches are scape, metafemur, and mandible proportions, eye position, and raw trait measurements. Traits that did not strongly contribute to the major linear discriminants include overall head shape and mesosomal measurements, suggesting that these have a lower predictive value for discerning ecological niche categories. Generally, the first linear discriminant comprised 40-55% of the total variance, with the second linear discriminant comprising 20-35% of the variance, regardless of whether raw or size-corrected ratio measurements were considered (Tables S10-17).

We used the posterior probabilities of the linear discriminant analyses for each ecological niche aspect to evaluate how accurately LDA models separated specimens into their known binnings. The results were variable (Tables S18): overall the linear discriminant analyses predicted foraging niche and nesting niche from morphology with moderate accuracy when using the raw and size-corrected ratio measurement datasets. Functional role was poorly predicted, typically with only ~50% accuracy. Ecomorph syndrome predictions were variable depending on whether we used raw or size-corrected measurements: ecomorphs were predicted with moderate accuracy (70.83%) using raw measurements, but poorly (56.25%) using size-corrected ratios. When using datasets comprised of only PGLS-selected traits (traits shown to be significantly different between binnings even after accounting for shared ancestry), accuracy typically decreased, though this was contingent on niche aspect; for example, all traits were PGLS-selected for foraging and nesting niche.

Quantitatively predicting ecological niche from morphology

Random Forest analyses produced a range of accuracies, though consistently much higher than linear discriminant analyses. Both raw and size-corrected measurements predicted niche aspects and ecomorph syndrome with moderate to good accuracy, with out-of-bag (OOB) accuracy estimates from 77-85% (Table 2). Nesting niche and foraging niche were consistently predicted with greater accuracy than functional role or ecomorph syndrome, though these were not poorly predicted as with the LDAs. Again, functional role was most poorly predicted out of all niche aspects. There was no clear difference between the raw and size-corrected measurement datasets in terms of which dataset predicted niche aspects most accurately; this was variable from niche aspect to niche aspect, and differed in only a few percentage points. However, when using the PGLS-selected dataset – our most conservative set of traits – prediction accuracy rates typically decreased slightly (Table S19).

When visualizing the posterior probabilities of model classification using our t-SNE plots, we found that most binnings and ecomorph syndromes were well-separated, indicating that our Random Forest models performed well in discretely separating our binnings (Figures 4, 5, S15). In general, the greatest degree of overlap was typically seen in ecomorph syndromes, where leaf litter-nesting and foraging omnivores would cluster near ground/leaf litter-nesting

Accepted Article
epigaeic omnivores, and leaf litter-nesting epigaeic predators would cluster with ground-nesting epigaeic predators. This may reflect similar morphology adapted to overlapping nesting and foraging strata, with similar feeding mechanisms.

When using raw measurements, overall, the niche aspect models most frequently confused carton-nesting species with other binnings, particularly with ground-nesting species; fungivores with other binnings; and arboreal foragers with epigaeic foragers. The ecomorph syndrome model most frequently confused subterranean predators with column-raiding predators, leaf-litter omnivores with epigaeic ground-nesting omnivores; and carton-nesting arboreal omnivores with epigaeic ground-nesting omnivores. When using size-corrected ratio measurements, the niche aspect models most often confused leaf litter-nesters for ground-nesters; leaf litter foragers for epigaeic foragers; and granivores with other binnings. In terms of the ecomorph syndromes, model confusion was similar to models trained on raw measurements. These errors likely reflect similar morphologies for similar life habits, a flexibility in some life habits, or minimal data for some groups - in binnings with fewer representative species, it was more difficult for the model to be trained sufficiently.

Restricting the number of terminal nodes indicated that the model was not highly overfitted (Figure S16). We tested this using the raw morphological trait measurement dataset predicting ecomorph syndromes. When restricted to only 20 terminal nodes, the model's accuracy was 64.08%, which is approximately 15% below maximum accuracy. The model approached maximum predictive accuracy extremely rapidly; at 50 terminal nodes it was only slightly below maximum accuracy, and it converged on maximum accuracy around 100 nodes. This prediction accuracy despite a highly constrained number of predictive pathways indicates a strong correlation between our sampled morphological traits and ecomorph syndromes.

Our Random Forest analysis identified several key traits as the strongest predictors of ecological niche; fundamental divisions in the classification system were based on these traits (Figures 3D-G, S17, S18). For nesting niche, the strongest predictors were eye, scape, procoxa, and metafemur morphology both for raw measurements and size-corrected ratio data. Traits that were the strongest predictors of foraging niche were eyes, scapes, metafemura (again raw length and proportionate length), and thoracic shape. The strongest predictors of functional role were

eyes, scapes, mandibles, and metafemura (again raw length and proportionate length). In congruence with some LDA results, we found that traits such as mesosomal shape, pronotal shape, and head shape typically had little contribution to the model for discerning ecological binnings, and other traits, such as eye position and body size had variable contributions depending on the ecological niche aspect being examined.

Discussion

Functional morphology and ant ecomorphs

“These are, in fact, the most slender of ants and anyone who has seen colonies of them filling narrow twigs and stems like so many sardines packed in a box, will be sure to regard the lengthening of the body as an adaptation to life in small tubular cavities.”
Wheeler, 1910

There are several morphological traits that we find to be indisputably linked to ecology. Many of these have been identified through other comparative approaches, which informed our own sampling (Gibb et al. 2015, Guilherme et al. 2019, Kaspari and Weiser 1999, Parr et al. 2003, Weiser and Kaspari 2006, Yates and Andrew 2011, Yates et al. 2014). We build upon these previous key findings through multivariate analysis, which explicitly incorporates combinations of these traits to provide an integrative view of overall morphology as it relates to ecology. Collectively, these results quantitatively describe patterns in ant evolution that have been described by taxonomists and ecologists for well over a century (e.g. Wheeler 1910). Our results recover suites of traits that are linked to ecology, define simplified ecomorph syndromes, and provide relative assessment of individual trait relevance.

Arboreal ants, particularly lignicolous nesters, are typified by larger eyes, shorter and stubbier appendages (procoxae, scapes, etc.) and flatter, broader mesosomas. This is likely reflective of nesting in small cavities within trees and shrubs, and the advantage of being more appressed to foraging substrate. An exemplar is *Melissotarsus*, which cannot walk on flat surfaces outside of the wooden tunnels in which it lives (Khalife et al. 2018). Leaf litter ants are

similar in these traits – their scapes and legs are shorter and their mesosomas more compact – but their eyes are usually small, given that their habitat is dark interstitial spaces on the forest floor.

Conversely, epigaeic ground-nesting ants tend to be greater in size and longer-legged, with larger eyes. With less complex foraging surfaces (Kaspari and Weiser 1999), there is likely a selective pressure for faster-moving ants to improve rates of resource acquisition and predator avoidance. Here, *Cataglyphis* is an archetype with elongate legs and eyes in the context of its niche as an obligate epigaeic forager in the Saharan desert (Wehner 1983). Some ecomorphs are an amalgamation of conflicting traits: column-raiding predators may occupy both subterranean as well as epigaeic habitats and so can be long-legged and elongate relative to ants that nest in complex interstitial spaces. However, the reduction in eyes and typically shortened scapes are consistent with subterranean nesting. We find that predators trend toward longer and flatter mandibles, and omnivores toward shorter, curved mandibles. Prey capture may be easier with larger stouter mandibles, while shorter curved mandibles may be adapted to the acquisition and manipulation of diverse food items, from seeds to honeydew.

When considering global ant morphology in the context of ecology, we are able to make relative comparisons of ecomorphological proximity. Proximate ecomorphospace indicates similarity in morphology, while overlap reveals trait morphology that is applicable to multiple ecologies. Across our analyses, we find that ecomorphospaces overlap in a core region, where a large proportion of species are clustered. These ants reflect a generalized platonic ant morphology, where they are not strongly morphologically specialized relative to other taxa. The species occupying this common space are often omnivorous, ground-nesting, and epigaeic, suggesting that these niche binnings are associated with more generalized morphology.

Some ecologies straddle both this common space as well as unique topography, while others occupy primarily unique morphospace. Species that occupy distinct morphospace display a suite of traits that are strongly correlated with niche binning, such as *Pseudomyrmex* (lignicolous nesting, arboreal foraging omnivore), *Liometopum* (carton-nesting arboreal foraging omnivore), and *Strumigenys* (leaf litter nesting and foraging predator) species.

We find that multidimensionality greatly aids in the characterization of ecomorphs. Species displaying shorter and stubbier appendages could be arboreal, leaf litter, or subterranean

dwelling. It is only after we incorporate other traits, such as eye size and mesosomal shape, that we can more accurately predict its likely ecological niche. Collapsed or summarized ecomorph syndromes improve prediction by simultaneously considering multiple ecological niche aspects and accommodating some plasticity. Ecomorphs are often defined by integration of traits more so than individual measurements; for example, when using random forest analysis to predict only nesting niche, carton-nesters are often inaccurately classified as other niche binnings, and in dimension reduction analyses carton-nesters typically plot in a generalized platonic ecomorphospace. However, when the integrated ecomorph of the carton-nesting arboreal foraging omnivore is considered, species are consistently predicted accurately.

Our collapsed ecomorphs are convergently evolved in ants. Each of our 10 sampled ecomorph syndromes span multiple phylogenetic origins (Figure 6). Ecomorphs are typified by whole-body convergent morphology: arboreal dorylines strongly resemble arboreal pseudomyrmicines despite being separated by ~100 million years (Borowiec et al. 2019). Morphological convergence on shared ecomorphs between phylogenetically disparate taxa suggests predictability in the evolution of these ecomorphs and convergent selective pressures (Mahler et al. 2013). This predictability reinforces findings of repeated evolution of ecomorphs across lineages, where taxa evolve a suite of specific traits to occupy available niches (Cooper & Westneat 2009; Gillespie et al. 2018). Moreover, our analyses and previous assessments of trait variance in a phylogenetic context indicate that morphological similarity within niche binnings is not the result of shared ancestry (Figure 3).

Predicting ecological niche from morphology

We find that ecological niche aspects or ecomorphs can be predicted from morphology with 77-85% accuracy, contingent on raw measurement or size-corrected ratio input data. This accuracy is comparable to other studies incorporating RF analysis (Pigot et al. 2020; Rojas et al. 2012; Rossel and Arbizu 2018). Even as our number of binnings ranges between four (foraging niche) and ten (ecomorph), we recover high predictive accuracy for all our models, which suggests a high degree of trait similarity within ecological binnings. Importantly, RF prediction incorporates non-linear relationships of traits, which improves accuracy relative to techniques

limited to linear relationships such as LDA (Breiman 2001; Pigot et al. 2020). The predictive power of RF illustrates the strong correlation between organism-wide morphology and ecology.

While there is a core group of morphological traits that our models found were most important in classifying various binnings and syndromes, there were some distinct differences between niche aspects. Consistently, raw eye size and eye size in relation to body size were most important across all niche aspects and ecomorph syndromes: this reflects the extreme consistency of eye size within binnings, with subterranean ants bearing minute eyes or entirely eyeless; leaf litter ants with smaller eyes; epigaeic ants with moderately sized to large eyes, and arboreal ants with extremely large eyes. The importance of eye size is reflected even in functional role classification, perhaps reflecting correlation between certain feeding habits and foraging or nesting microhabitats. Models ranked mandible size in relation to body size and raw mandible size as a highly important variable in functional role binning predictions, possibly reflecting the importance of mandibles in feeding habits. Leg (procoxa and metafemur) length and scape length were consistently important in all niche aspect binning classification, reflecting the importance of appendage size in nesting microhabitat and foraging in interstitial spaces especially.

Model misclassification is also informative; our models typically misclassify niches that exhibit some degree of morphological overlap or ecological plasticity. One of the most frequently misclassified ecomorphs is the subterranean predator ecomorph, often misclassified as a column-raiding predator. Column-raiding predators have a mixed nesting niche where they will occasionally make temporary subterranean nests, in between periods of nomadism where they form bivouacs for nesting (Borowiec 2016). Morphological adaptations that are implicated in column-raiding could also relate to subterranean-nesting – column-raiders are typically highly dependent on chemical foraging trails and are often eyeless or bear minute eyes, another feature of subterranean dwellers.

Our models also generate misclassifications between species that forage and nest variably in epigaeic surficial and leaf litter habitats. This may be a result of plasticity in nesting strata: some species may nest in leaf litter structures, such as logs, and also construct ground nests. Indeed, these plastic ecologies may reflect the same pattern as overlapping ecomorphospaces

plotted by our LDAs. A “platonic ecomorph” will bear more generalized traits, and therefore is less easily classified using machine learning.

A pipeline for predictive ecomorphological modeling

We include in our supplementary material both an expandable morphometric dataset and template code to provide a pipeline for future ecomorphological modeling. For any specimen databank, workers can be measured according to traits sampled here. The predictive ecomorphological model is then trained on this expanded dataset, and can be used to evaluate hypotheses related to ecological diversity. We include both LDA model code and Random Forest model code, though Random Forest typically generates a higher predictive accuracy.

Increased taxonomic sampling in future datasets will further illuminate links between form and function in ants. Moreover, future intraspecific sampling of distinct castes may reveal more complex trends related to the capacity for lineages to become ecologically specialized. Here, we provide two options for predicting the ecomorph of a worker: a simplified set of ecomorph syndromes, or for more granular analysis, classification of worker functional role, foraging niche, and nesting niche. Additionally, new ecological dimensions may be incorporated including alternative ecomorph syndromes, abiotic features, or ethology.

References

- Barton PS, Gibb H, Manning AD, Lindenmayer DB, Cunningham SA. 2011. Morphological traits as predictors of diet and microhabitat use in a diverse beetle assemblage. *Biological Journal of the Linnaean Society* 102:301-310.
- Blanchard, B.D. and Moreau, C.S., 2017. Defensive traits exhibit an evolutionary trade-off and drive diversification in ants. *Evolution*, 71(2), pp.315-328.

Bolton, B., 2020. An online catalog of the ants of the world. Available from <http://antcat.org>. [accessed (March 2020)]

Borowiec ML. 2016. Generic revision of the ant subfamily Dorylinae (Hymenoptera, Formicidae). *Zookeys* 608:1-280.

Borowiec, M.L., Rabeling, C., Brady, S.G., Fisher, B.L., Schultz, T.R. and Ward, P.S., 2019. Compositional heterogeneity and outgroup choice influence the internal phylogeny of the ants. *Molecular Phylogenetics and Evolution*, 134, pp.111-121.

Breiman L. 2001. Random Forests. *Machine Learning* 45(1):5-32.

Collins CE, Russel AP, Higham TE. 2015. Subdigital adhesive pad morphology varies in relation to structural habitat use in the Namib Day Gecko. *Functional Ecology* 29:66-77.

Cooper, W.J. and Westneat, M.W., 2009. Form and function of damselfish skulls: rapid and repeated evolution into a limited number of trophic niches. *BMC Evolutionary Biology*, 9(1), p.24.

Dickson BV, Sherratt E, Losos JB, Pierce SE. 2017. Semicircular canals in *Anolis* lizards: ecomorphological convergence and ecomorph affinities of fossil species. *Royal Society Open Science* 4:170058.

Dray, S. and Dufour, A.B., 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of statistical software*, 22(4), pp.1-2

Felsenstein, J., 1985. Phylogenies and the comparative method. *The American Naturalist*, 125(1), pp.1-15.

Figueirido B, Palmqvist P, Perez-Claros JA, Janis CM. 2019. Sixty-six million years along the road of mammalian ecomorphological specialization. *Proceedings of the National Academy of Sciences of the United States of America* 116:12698-12703.

Gerry SP, Wang J, Ellerby DJ. 2011. A new approach to quantifying morphological variation in bluegill *Lepomis macrochirus*. *Journal of Fish Biology* 78:1023-1034.

Ghazali M, Moratelli R, Dzeverin I. 2017. Ecomorph evolution in *Myotis* (Vespertilionidae, Chiroptera). *Journal of Mammalian Evolution* 24:475-484.

Gibb H, Stoklosa J, Warton DI, Brown AM, Andrew NR, Cunningham SA. 2015. Does morphology predict trophic position and habitat use of ant species and assemblages? *Community Ecology* 177:519-531.

Gillespie, R., 2004. Community assembly through adaptive radiation in Hawaiian spiders. *Science*, 303(5656), pp.356-359.

Gillespie, R.G., Benjamin, S.P., Brewer, M.S., Rivera, M.A.J. and Roderick, G.K., 2018. Repeated diversification of ecomorphs in Hawaiian stick spiders. *Current Biology*, 28(6), pp.941-947.

Guilherme DR, Pereira Souza JL, Franklin E, Costa Lima Pequeno PA, das Chagas AC, Baccaro FB. 2019. Can environmental complexity predict functional trait composition of ground-dwelling ant assemblages? A test across the Amazon Basin. *Acta Oecologica*. 99:103434.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. and Challenger, W., 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics*, 24(1), pp.129-131.

Hölldobler, B, Wilson, E. O. 1990. *The Ants*. Harvard University Press, Cambridge, MA.

Janis CM, Figueirido B. 2014. Forelimb anatomy and the discrimination of the predatory behaviour of carnivorous mammals: the thylacine as a case study. *Journal of Morphology* 275:1321-1338.

Kaspari, M. and Weiser, M.D., 1999. The size–grain hypothesis and interspecific scaling in ants. *Functional Ecology*, 13(4), pp.530-538.

Khalife, A., Keller, R.A., Billen, J., Garcia, F.H., Economo, E.P. and Peeters, C., 2018. Skeletomuscular adaptations of head and legs of *Melissotarsus* ants for tunnelling through living wood. *Frontiers in Zoology*, 15(1), p.30.

Kuhn, M., 2015. Caret: classification and regression training. *Astrophysics Source Code Library*.

Krijthe, J.H., 2015. Rtsne: T-distributed stochastic neighbor embedding using Barnes-Hut implementation. R package version 0.13

Langerhans RB, Knouft JH, Losos JB. 2006. Shared and unique features of diversification in the greater Antillean *Anolis* ecomorphs. *Evolution* 60(2):362-369.

Lê, S., Josse, J. and Husson, F., 2008. FactoMineR: an R package for multivariate analysis. *Journal of statistical software*, 25(1), pp.1-18.

Leonard JA, Vila C, Fox-Dobbs K, Koch PL, Wayne RK, Van Valkenburg B. 2007. Megafaunal extinctions and the disappearance of a specialized wolf ecomorph. *Current Biology* 17:1146-1150.

Liaw, A. and Wiener, M., 2018. randomForest: Breiman and Cutler's random forests for classification and regression. R package version 4

Losos JB. 1992. The evolution of convergent structure in Caribbean *Anolis* communities. *Systematic Biology* 41(4):403-420.

Losos JB, Jackman TR, Larson A, de Queiroz K, Rodriguez-Schettino L. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279:2115-2118.

Mahler, D.L., Ingram, T., Revell, L.J. and Losos, J.B., 2013. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science*, 341(6143), pp.292-295.

Malcicka M, Berg MP, Ellers J. 2017. Ecomorphological adaptations in Collembola in relation to feeding strategies and microhabitat. *European Journal of Soil Biology* 78:82-91.

Marquez F, Nieto Vilela RA, Lozada M, Bigatti G. 2015. Morphological and behavioural differences in the gastropod *Trophon geversianus* associated to distinct environmental conditions, as revealed by a multidisciplinary approach. *Journal of Sea Research* 95:239-247.

Martins, E.P. and Hansen, T.F., 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *The American Naturalist*, 149(4), pp.646-667.

Mihalitsis M, Bellwood DR. 2019. Morphological and functional diversity of piscivorous fishes on coral reefs. *Coral Reefs* 38:945-954.

Mugleston JD, Naegle M, Song H, Whiting MF. 2018. A comprehensive phylogeny of Tettigoniidae (Orthoptera: Ensifera) reveals extensive ecomorph convergence and widespread taxonomic incongruence. *Molecular Phylogenetics, Phylogenomics and Phylogeography* 2(4):1-27.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H. and Oksanen, M.J., 2013. Package 'vegan'. Community ecology package, version, 2(9), pp.1-295.

Paradis, E. and Schliep, K., 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), pp.526-528.

Parker J, Owens B. 2018. *Batriscydaenus* Parker and Owens, new genus, and convergent evolution of a “reductive” ecomorph in socially symbiotic Pselaphinae (Coleoptera: Staphylinidae). *The Coleopterists Bulletin* 72(2):219-229.

Parr ZJE, Parr CL, Chown SL. 2003. The size-grain hypothesis: a phylogenetic and field test. *Ecological Entomology* 28:475-481.

Pigot AL, Sheard C, Miller ET, Bregman TP, Freeman BG, Roll U, Seddon N, Trisos CH, Weeks BC, Tobias JA. 2020. Macroevolutionary convergence connects morphological form to ecological function in birds. *Nature Ecology and Evolution* 4(2):230-239.

Pinheiro, J., Bates, D., DebRoy, S. and Sarkar, D., 2018. R Core Team. 2018. nlme: linear and nonlinear mixed effects models. R package version 3.1-137. URL: <https://CRAN.R-project.org/package=nlme>.

Rojas R, Morales MC, Rivandeneira MM, Thiel M. 2012. Male morphotypes in the Andean river shrimp *Cryphiops caementarius* (Decapoda: Caridea): morphology, coloration, and injuries. *Journal of Zoology* 288:21-32.

Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3(2), pp.217-223.

Ripley, B., Venables, B., Bates, D.M., Hornik, K., Gebhardt, A., Firth, D. and Ripley, M.B., 2019. Package 'MASS'.

Roberts, D.W., 2007. labdsv: Ordination and multivariate analysis for ecology. R package version, 1(1).

Rossel S, Arbizu PM. 2018. Automatic specimen identification of Harpacticoids (Crustacea: Copepoda) using Random Forest and MALDI-TOF mass spectra, including a post hoc test for false positive discovery. *Methods in Ecology and Evolution* 9:1421-1434.

Salas-López, A., 2017. Predicting resource use in ant species and entire communities by studying their morphological traits: Influence of habitat and subfamily. *Ecological Indicators*, 78, pp.183-191.

Sanders KL, Rasmussen AR, Elmberg J, de Silva A, Guinea ML, Lee MSY. 2013. Recent rapid speciation and ecomorph divergence in Indo-Australian sea snakes. *Molecular Ecology* 22:2742-2759.

Saunders MB, Barclay RMR. 1992. Ecomorphology of insectivorous bats: a test of predictions using two morphologically similar species. *Ecology* 73(4):1335-1345.

Sosiak C, Barden P. 2020. Data from: Multidimensional trait morphology predicts ecology across ant lineages. Dryad Digital Repository: <https://doi.org/10.5061/dryad.kh1893243> .

Team, R.C., 2020. R: A language and environment for statistical computing.

Trontelj P, Blejcek A, Fiser C. 2012. Ecomorphological convergence of cave communities. *Evolution* 66:3852-3865.

Wehner, R., 1983. Taxonomie, Funktionsmorphologie und Zoogeographie der saharischen Wüstenameise *Cataglyphis fortis* (Forel 1902) stat. nov. *Senckenbergiana biologica*, 64, pp.89-132.

Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y. and Zemla, J., 2017. Package 'corrplot'. *Statistician*, 56, pp.316-324.

Weiser MD, Kaspari M. 2006. Ecological morphospace of New World ants. *Ecological Entomology* 31:131- 142.

Wheeler, W.M., 1910. *Ants: their structure, development and behavior* (Vol. 9). Columbia University Press.

Williams, Ernest E. (1972). "3. The Origin of Faunas. Evolution of Lizard. Congeners in a Complex. Island Fauna: A Trial Analysis". In T. Dobzhansky; et al. (eds.). *Evolutionary Biology*. Meredith Corporation. pp. 47–89. refer page 72.

Yates M, Andrew NR. 2011. Comparison of ant community composition across different land-use types: assessing morphological traits with more common methods. *Australian Journal of Entomology* 50:118-124.

Yates ML, Andrew NR, Binns M, Gibb H. 2014. Morphological traits: predictable responses to macrohabitats across a 300 km scale. *PeerJ* 2:e271

Author contributions

CS and PB conceived the ideas and designed methodology, analysed the data, drafted the manuscript; CS collected the data. Both authors contributed critically to the drafts and gave final approval for publication.

Data Accessibility

All data are available on the Dryad Digital Repository: <https://doi.org/10.5061/dryad.kh1893243> (Sosiak and Barden 2020); and R code necessary for the reproduction of analyses are provided in the Supporting Information accompanying this manuscript.

Supporting Information

Additional supporting information may be found in the online version of this article.

Supporting Tables and Figures (FESosiakSTF): All supporting tables and figures listed in the main text.

Supporting Appendix 1 (FESosiakSA1): Natural history references for ecological niche binnings.

Supporting Appendix 2 (FESosiakSA2): Supplementary R code for all analyses.

Figures and Tables

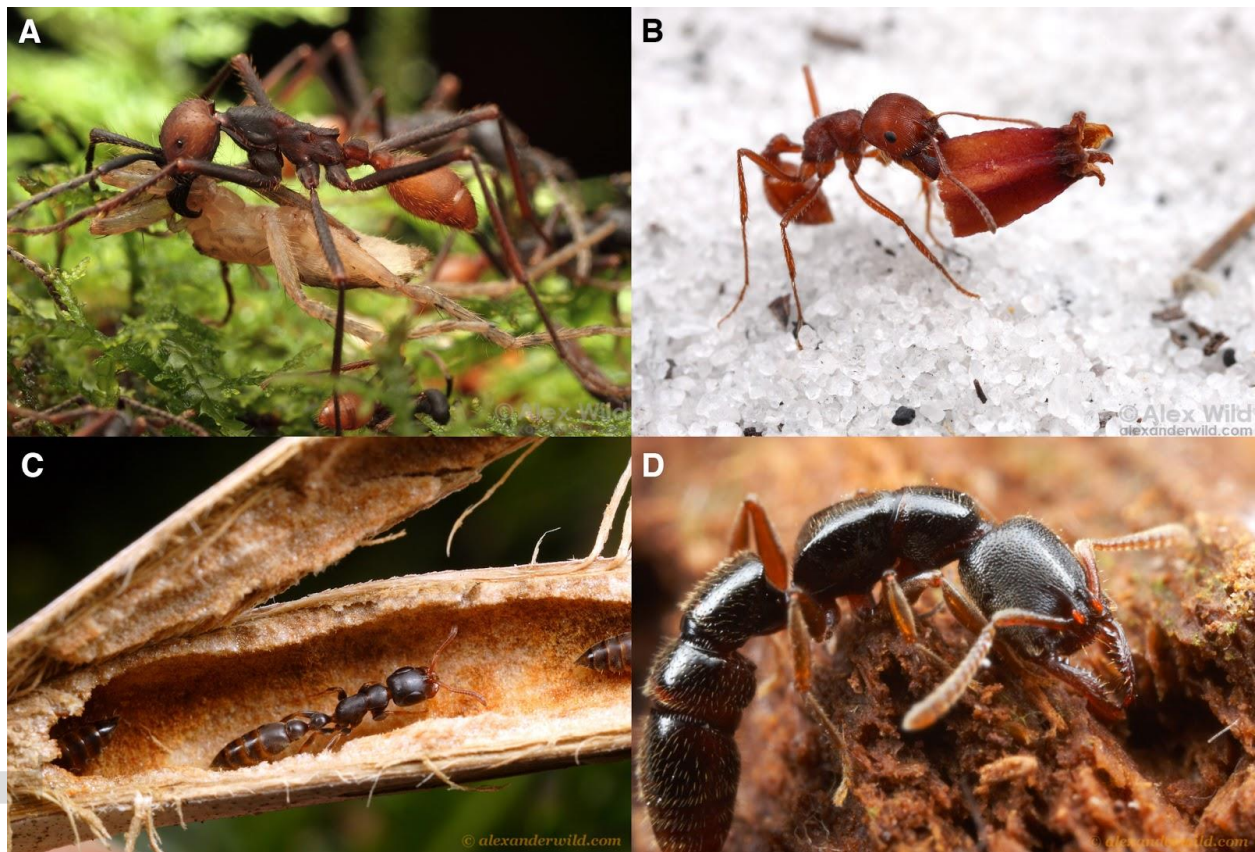


Figure 1. Ecological diversity among ant lineages. Ants inhabit a wide variety of ecological niches: column-raiding predators such as *Eciton burchellii* (Dorylinae) (A); granivorous epigeaic foragers such as *Pogonomyrmex badius* (Myrmicinae) (B); omnivorous arboreal twig-nesters such as *Tetraponera clypeata* (Pseudomyrmecinae) (C); and subterranean predators such as *Stigmatomma* sp. (Amblyoponinae) (D). Image ©Alex Wild, used by permission.

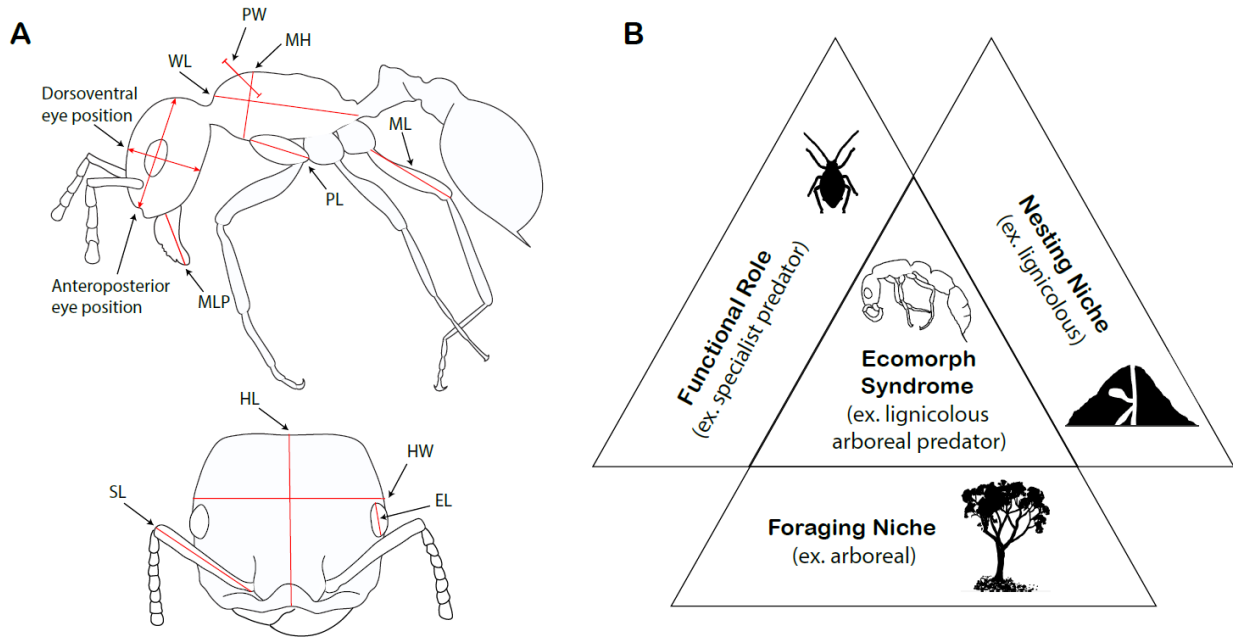


Figure 2. Morphological traits and ecological niche aspects. Diagram of morphometric measurements (A) and schematic illustrating ecological niche aspects and ecomorph syndromes (B) used in this study. All measurement abbreviations are listed in Table S2.

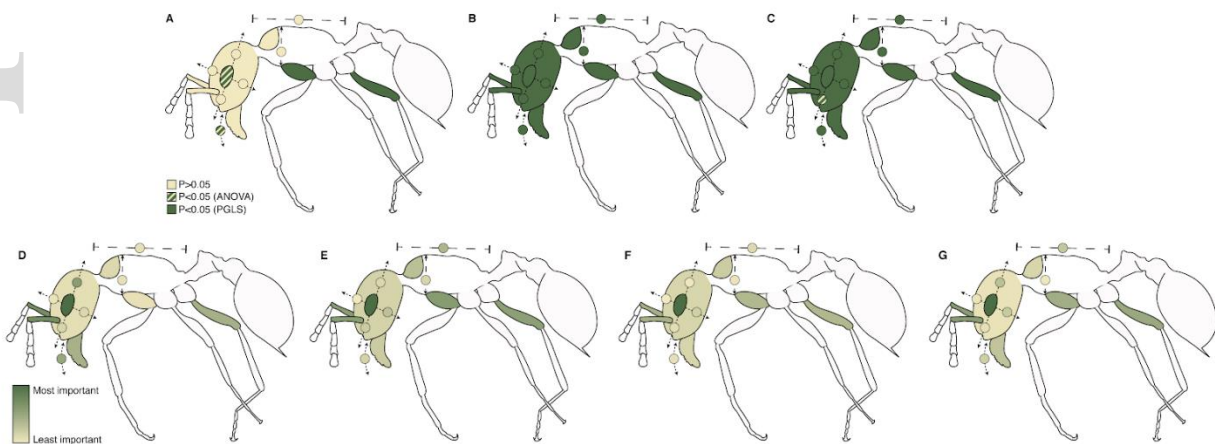


Figure 3. Significance, relative importance of raw trait measurements sampled by ecological niche aspect and ecomorph syndrome. First row: traits found to be significantly different between binnings using ANOVA/PGLS in functional role (A), nesting and foraging

niche, (B) and ecomorph syndrome (C). Second row: trait importance to Random Forest models in classifying functional role (D), foraging niche (E), nesting niche (F), and ecomorph syndrome (G). Head shape traits averaged in D-G. Size corrected ratio measurement results in Figure S9. ANOVA and PGLS results in Table S4-S5, relative importance results in Figure S17-S18.

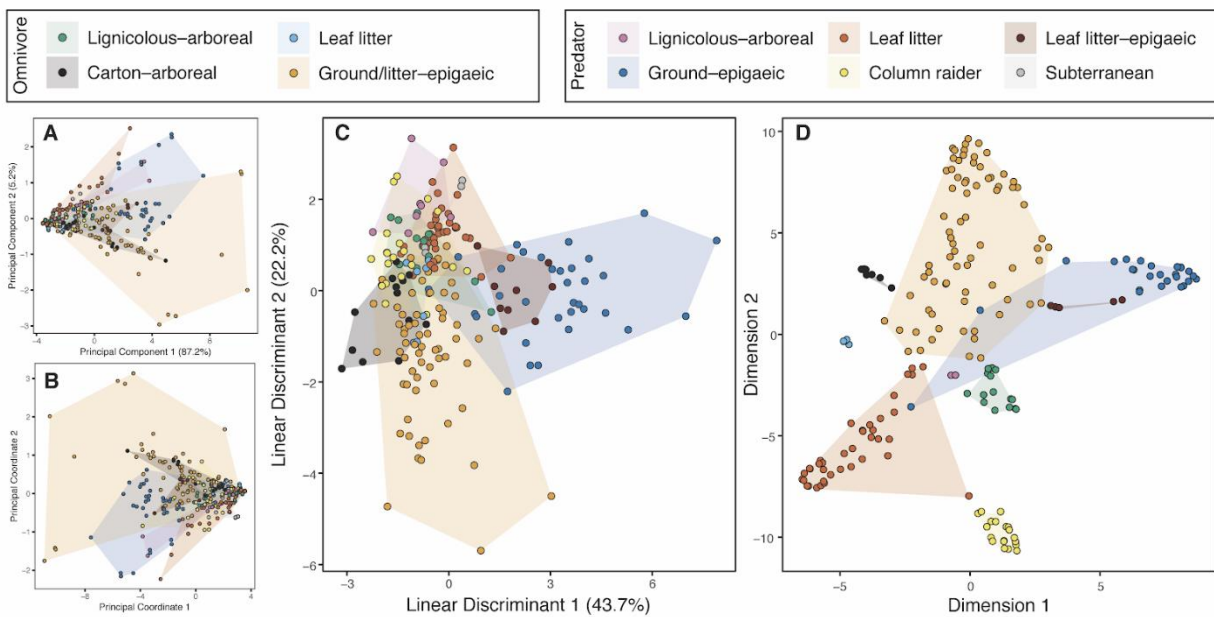


Figure 4. Ecomorphospace characterizing ecomorph syndromes across dimension reduction spaces. Ecomorphs are labeled by simplified functional role as well as nesting–foraging niches. A) principal component analysis (PCA); B) principal coordinate analysis (PCoA); C) linear discriminant analysis (LDA); D) t-distributed Stochastic Neighbour Embedding from Random Forest models (t-SNE RF). All taxa represented by raw trait measurements; size corrected ratio measurement results in Figures S11D, S13D-S15D. Summary statistics in Tables S6, S8, S13.

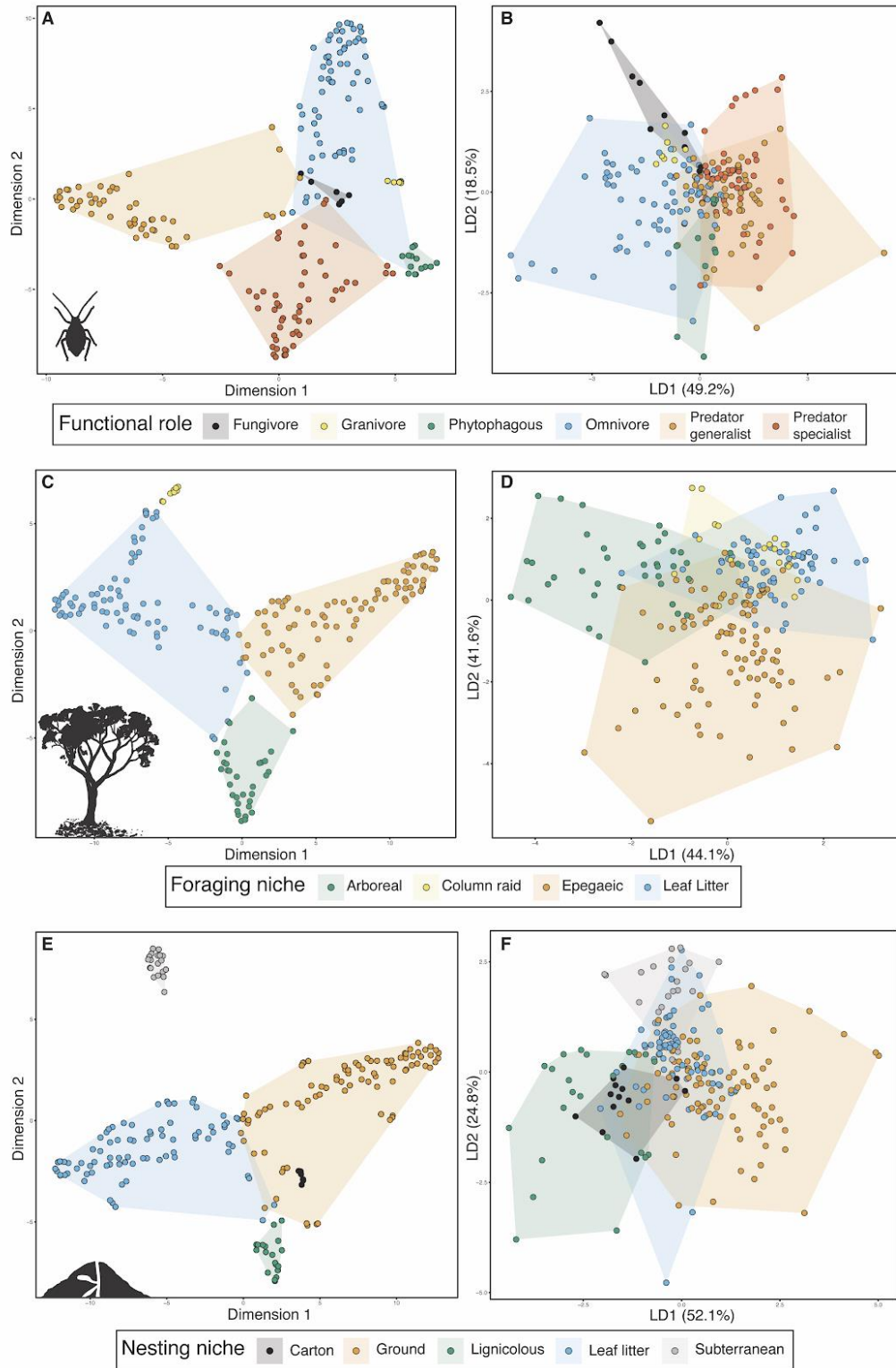


Figure 5. Ecomorphospace of niche aspects using dimension reduction and machine learning classification methods. Left) t-SNE plots; Right) LDA. A-B) functional role, C-D)

foraging niche, and E-F) nesting niche. All taxa represented by raw trait measurements; size corrected ratio measurement results in Figures S14A-C, S15A-C. Summary statistics in tables S10-S12.

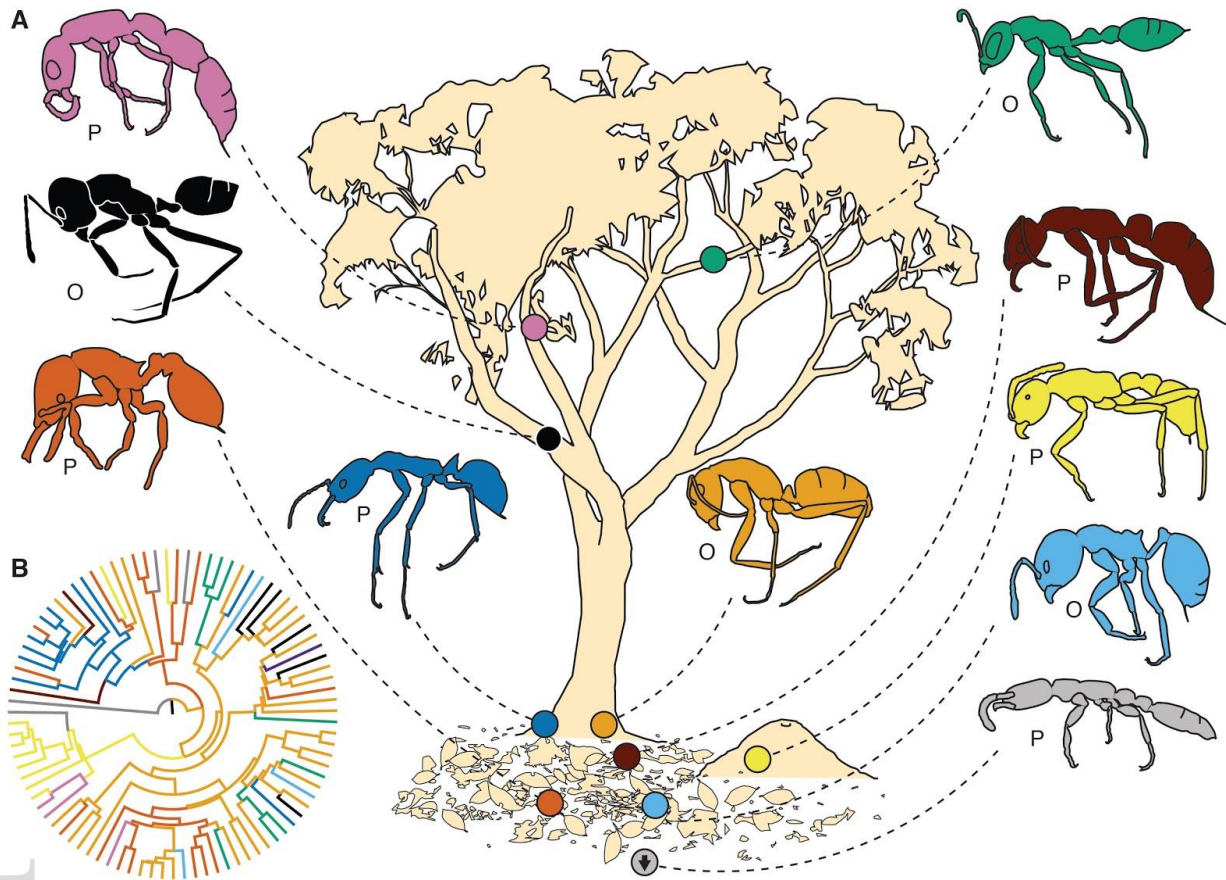


Figure 6. Ecological and phylogenetic distribution of ant ecomorphs. A) Schematic of niche occupation among ant ecomorphs (P: predatory, O: omnivory). Exemplar taxa in counterclockwise order: *Simopone* (pink); *Azteca*; *Strumigenys*; *Odontomachus*; *Camponotus*; *Leptanilla*; *Aneuretus*; *Neivamyrmex*; *Leptogenys*; *Pseudomyrmex*. Colors correspond with ecomorph syndromes denoted in Figure 4. B) Phylogenetic distribution of 90 sampled genera and associated ecomorphs, relationships from Blanchard & Moreau (2017); taxa noted in dataset available on Dryad, doi: <https://doi.org/10.5061/dryad.kh1893243>.

Functional Role		
<i>Binning designator</i>	<i>Definition</i>	<i>Exemplar taxa</i>
GP	Generalist predator – broad taxonomic diet	<i>Odontomachus; Diacamma; Harpegnathous</i>
SP	Specialist predator – obligate feeding on specific taxon (e.g. termites)	<i>Acanthostichus; Megaponera; Simopelta</i>
Om	Omnivorous – prey items, plant matter, etc.	<i>Paraponera; Camponotus; Iridomyrmex</i>
Py	Phytophagous – extrafloral nectaries, herbivory, etc.	<i>Pseudomyrmex; Tetraponera; Myrmelachista</i>
Fg	Fungus-growing	<i>Cyphomyrmex; Trachymyrmex; Atta</i>
Tr	Trophobiotic – symbiotic relationship with other insects (homopteran secretions, etc.)	<i>Acropyga; Melissotarsus; Rhopalomastix</i>
Gn	Granivorous – seed-harvesting	<i>Acanthomyrmex; Pogonomyrmex; Veromessor</i>
Mh	Mushroom-foraging	<i>Euprenolepis</i>
Nesting niche		
Cn	Carton-nesting – structured nests from plant material in trees and shrubs	<i>Oecophylla; Azteca; Liometopum</i>
Gr	Ground-nesting – nests in dirt mounds, under stones, rock cracks, etc.	<i>Platythyrea; Formica; Pheidole</i>
Lg	Lignicolous – nests in twig and tree cavities	<i>Pseudomyrmex; Simopone; Cylindromyrmex</i>
Ll	Leaf litter-nesting – nests in leaf litter interstitial space, rotten wood, etc.	<i>Strumigenys; Discothyrea; Typhlomyrmex</i>
Sb	Subterranean-nesting	<i>Leptanilloides; Leptanilla</i>
Foraging Niche		

Ab	Arboreal – in and on trees and shrubs	<i>Daceton; Tetraponera; Crematogaster</i>
CR	Column-raiding – cooperative, nomadic, or raiding predation	<i>Simopelta; Dorylus; Eciton</i>
Eg	Epigaeic – active foraging on the ground surface	<i>Leptomyrmex; Rhytidoponera; Myrmecocystus</i>
Ll	Leaf litter – within interstitial spaces in leaf litter	<i>Discothyrea; Amblyopone; Heteroponera</i>
Sb	Subterranean – underground	<i>Acropyga</i>

Table 1. Ecological niche aspect binning abbreviations, definitions, and exemplar taxa (abbreviations used in dataset available on Dryad, doi: <https://doi.org/10.5061/dryad.kh1893243>).

	Functional Role	Nesting Niche	Foraging Niche	Ecomorph
Raw measurements	79.43%	85.27%	82.14%	78.16%
Size-corrected ratios	77.51%	83.48%	84.38%	78.1%

Table 2. Classification accuracy of Random Forest based on out-of-bag (OOB) error rates. Models were optimized using both the raw morphological trait measurement dataset and the size-corrected ratio measurement dataset for each niche aspect and ecomorph syndrome. Classification accuracy for Random Forest models using only PGLS-selected traits in Table S19.