UNIVERSITY^{OF} BIRMINGHAM University of Birmingham Research at Birmingham

The species-area relationship in ant ecology

Ohyama, Leo; Holt, Robert D.; Matthews, Tom; Lucky, Andrea

DOI: 10.1111/jbi.14149

License: Other (please specify with Rights Statement)

Document Version Peer reviewed version

Citation for published version (Harvard):

Ohyama, L, Holt, RD, Matthews, T & Lucky, A 2021, 'The species-area relationship in ant ecology', *Journal of Biogeography*, vol. 48, no. 8, pp. 1824-1841. https://doi.org/10.1111/jbi.14149

Link to publication on Research at Birmingham portal

Publisher Rights Statement:

This is the peer reviewed version of the following article: Ohyama, L., Holt, R.D., Matthews, T.J. and Lucky, A. (2021), The species–area relationship in ant ecology. J Biogeogr., which has been published in final form at: https://doi.org/10.1111/jbi.14149. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.

Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)

•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

- 1 Strapline: The ant species–area relationship
- 2 The species-area relationship in ant ecology
- 3

12

- *Running title:* Species-area relationships in ant ecology
- Leo Ohyama^{1,2*}, Robert D. Holt³, Thomas J. Matthews^{4,5}, Andrea Lucky¹
- 8 *Corresponding Author, leoohyama12@gmail.com

¹ University of Florida, Entomology and Nematology Department, 1881 Natural Area Drive,
 Gainesville, Florida, USA, 32611

² University of Florida, Biodiversity Institute, 432 Newell Drive, Gainesville, Florida, USA,
 32611

- ³ University of Florida, Department of Biology, 876 Newell Drive, Gainesville, Florida, USA,
 32611
- ⁴ GEES (School of Geography, Earth and Environmental Sciences) and Birmingham Institute of
 Forest Research, University of Birmingham, Birmingham, B15 2TT
- 21

⁵ CE3C – Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity
 Group and Universidade. dos Açores – Depto de Ciências Agráriase Engenharia do Ambiente,

24 PT-9700-042, Angra do Heroísmo, Açores, Portugal.

- 25
- 26 Acknowledgments

We thank David G. Jenkins for consultation on analyses and discussion, Phil Hahn for feedbackon the working drafts, Juan D. Bogota for discussion, James Wetterer for sharing unpublished

29 and original data, and Heraldo Vasconcelos for sharing original data. We thank the editor,

30 Leonardo Dapporto, François Rigal, and one anonymous reviewer for providing constructive

- 31 feedback that helped improve this study. LO was supported by a University of Florida Graduate
- 32 School Funding Award and a University of Florida Biodiversity Institute Fellowship. RDH
- thanks the University of Florida Foundation for its support. This work is/was supported by the
- 34 USDA National Institute of Food and Agriculture, McIntire Stennis project 10111529. No
- 35 permits were required for this research.
- 36
- 37 *Author contributions*
- 38 LO conceived the idea, LO collected the data, LO and RDH were responsible for the
- development of the idea, LO and TJM analyzed the data, LO wrote the manuscript with input
- 40 from RDH, TJM, and AL.
- 41
- 42 *Statement of significance*
- 43 Understanding why the number of species in a given area increases across space is crucial to
- 44 determining the current and future states of biodiversity. We review and analyze data from the
- 45 literature on the ants, one of the best-studied invertebrate ecosystem engineers in insular systems.
- 46 We find that ant diversity scales faster across space in mainland versus insular systems. We also

- 47 find that precipitation and areas with unique evolutionary histories (e.g. the Neotropics)
- 48 influence the rate at which diversity scales across space. Finally, we identify knowledge gaps and
- 49 solutions, extendable to other taxa, to better understand the scaling of biodiversity.
- 50

51 Abstract

52 Aim: The positive relationship between species richness and area – the species–area relationship

(SAR) – is a key principle in ecology. Previous studies show large variation in the SAR across
 taxa, collectively indicating the necessity of a taxon-focused approach to accurately evaluate

55 biodiversity scaling patterns. Ants are ideal for this given their global distribution, and role in

56 ecosystem functioning. Using data from insular ant communities, this study quantified and

57 investigated various attributes of ant SARs and reviewed the SAR literature for ant faunas,

- 58 world-wide, to identify specific areas for improvement.
- 59

60 Location: Islands around the world

- 61
- 62 Taxon: Ants (Hymenoptera: Formicidae)

6364 Methods: We aggregated data on species richness and island characteristics from previous

65 studies on ant SARs to evaluate effects of climate, biogeographic realm and latitude on slope

values from these studies. A multi-model inference approach was used to determine the form of

67 the different SARs, and whether there were any differences between mainland and insular SARs.

68 We also assessed differences between mainland and insular SAR slopes and intercepts. To seek a

69 general slope coefficient for ants, we used a mixed-effect model. Finally, we tested for potential70 thresholds in the global insular SAR using piecewise regression models.

71

72 Results: There was a negative relationship between SAR slopes and precipitation in both

72 Results. There was a negative relationship between SAR slopes and precipitation in both
 73 mainland and insular SARs, while SAR slopes and intercept values were lower and higher,

respectively, in mainland compared to insular systems. Strong evidence of thresholds emerged in

the global insular SAR. Finally, a general slope of 0.16 was observed for insular systems, which

76 is lower than found in previous studies.

7778 Main conclusions: A taxon-focused approach proves to showcase unexpected patterns in the

79 SAR. Ant diversity increases faster across area in mainland areas compared to true islands. The

80 influences of climate and biogeographic realms on the ant SAR warrant deeper study. Our

- 81 review highlights knowledge gaps in the ant SAR that also extend to other taxa, such as the
- 82 effects of non-native species on the SAR.

8384 Key words

Ants, Biodiversity, Fragmentation, Invasive species, Islands, Island biogeography, Species–area
 relationship

87

88 Introduction

- 89 The Theory of Island Biogeography (MacArthur & Wilson, 1967), a foundational framework for
- 90 ecology and biogeography, was motivated in part by consideration of the species-area
- 91 relationship (SAR). This relationship was first noted by Forster (1778) and brought to light by
- 92 Wallace (1914), and has played a central role in biogeography up to the present (Matthews et al.,

93 2020, *In Press*). A SAR describes the tendency for species richness to increase with increasing

area (Mittelbach & McGill, 2019) and is one of the few patterns in ecology that has been called a
general "law" (Lomolino, 2000; Whittaker & Fernandez-Palacios, 2007). SARs have been

general "law" (Lomolino, 2000; Whittaker & Fernandez-Palacios, 2007). SARs have been
instrumental in the field of conservation biogeography (Whittaker et al., 2005), including in the

97 design of nature reserves, extinction forecasting, and gauging the effects of habitat fragmentation
98 (Halley et al., 2014; Suarez et al., 1998; Whittaker & Fernandez-Palacios, 2007).

99 In this study, we focus on the island species-area relationship (hereafter referred to as the 100 ISAR), which utilizes distinct islands or distinctly bounded patches within larger regions 101 (Matthews et al., 2016; Rosenzweig, 1995; Scheiner, 2003). ISARs in both mainland and insular 102 systems were investigated to better understand how biodiversity changes across different-sized independent areas and in relation to various environmental and biogeographical conditions. Here, 103 we define insular ISARs as those that focus on species richness and areas of true islands. 104 105 landforms completely surrounded by water. We define mainland ISARs as those that focus on "islands" (i.e. habitat islands) found within the boundaries of continents, such as discrete patches 106 107 created by geomorphological processes (e.g., outcrops of granite or patches of serpentine), or by 108 anthropogenic processes of habitat destruction and habitat fragmentation.

109 110

111

112

Most often, the ISAR is fit with a power function (Arrhenius, 1920b)(eq 1.1),

 $S = cA^z \tag{1.1}$

113 where S is species richness, A is area, and c and z are fitted constants. Different mathematical 114 models with different forms for describing ISARs have been proposed, and in some cases these 115 other forms (e.g. linear, sigmoidal) provide a better fit to the ISAR than does the convex power 116 law (Scheiner, 2003; Even Tjørve, 2003). When expression (1.1) is linearized by logarithmic 117 transformations of S and A (see (1.2.)) z is the slope of the resulting log-log relationship, and c 118 represents the intercept (eq 1.2) (Scheiner, 2003):

- 119
- 120 121

 $\log(S) = \log(c) + z \log(A) \tag{1.2}$

122 The meaning of z (called the z-value hereafter) holds a place of great interest as it captures the scaling of species richness with increasing area (Arrhenius, 1920a; Kylin, 1923). 123 The z-value has been observed to approach unity when the spatial areas under consideration are 124 125 considerably larger than entire species' ranges (Mittelbach & McGill, 2019), as seen in the triphasic ISAR where, at large provincial scales, the z-value increases compared to smaller scales 126 127 (Rosenzweig, 1995; Shmida & Wilson, 1985). The intercept $\log(c)$ (called the *c*-value hereafter) represents the species richness of an island of unit area and can be influenced by the study taxon 128 129 and biogeographic region (MacArthur & Wilson, 1967). It has also been used to indicate the 130 biotic richness of an insular system and hence the degree of island impoverishment (Whittaker & 131 Fernandez-Palacios, 2007).

There are several proposed explanations for the ISAR. Three of the more well-developed
hypotheses center on habitat diversity, sampling effects, and colonization-extinction dynamics.
The *habitat diversity hypothesis* assumes that habitat variety increases as area increases, resulting
in greater niche space and species richness (Lack, 1976; Williams, 1964). The *sampling effect*posits that larger areas include more individuals than smaller areas, hence representing a larger
'sample', in effect drawn from a regional species pool (Arrhenius, 1921; Connor & McCoy,

138 1979). Finally, the *colonization and extinction dynamics hypothesis* posits that a homogenous

area's increasing size supports larger populations with lower chance of extinction and higher
 chance of colonization (MacArthur & Wilson, 1967; Mittelbach & McGill, 2019).

One challenge to exploring the mechanisms underpinning ISARs is that multiple taxon 141 groups are often analyzed together. By contrast, using a single-taxon (e.g. mammals, birds, frogs, 142 or ants) approach to evaluating ISARs removes known variation in ISAR form and z-values due 143 144 to taxon effects (Nilsson et al., 1988). Previous studies have found that the z-value of the ISAR can vary between systems, across climatic (Kalmar & Currie, 2007) and latitudinal gradients 145 146 (Drakare et al., 2006), and by island type (Matthews et al., 2016). These studies document wide variation in the ISAR, possibly resulting from multiple mechanisms that vary depending upon 147 148 environmental context and ecologically heterogeneous taxa. Differences in z-values by taxa are widespread. For example, a study of ISARs in a mixture of both oceanic and continental islands 149 in the Lesser Antilles found birds to have a z-value of 0.21, bats 0.23, butterflies 0.27, and 150 151 reptiles and amphibians 0.17 (Ricklefs & Lovette, 1999). Even within lower taxa such as invertebrates, differences in both c and z-values are observed. For example, in the Mediterranean 152 153 islands centipedes, isopods, and tenebrionid beetles range in z-values (0.24-0.31, 0.20-0.26, 0.27, respectively) and *c*-values (0.24,-0.31, 0.59-0.80, 0.61-0.94) (Fattorini et al., 2017). Taxon-based 154 differences in both z and c values are likely to be influenced by each taxon's traits such as 155 156 dispersal capacity or life history strategies. This is observed in butterflies where higher z-values 157 are associated with butterflies with smaller ranges and lower reproductive potential (Franzen et 158 al., 2012). Therefore, an ideal taxon for studying ISARs would be a geographically widespread, 159 taxonomically resolved, species-rich group, with a wide range of well-defined traits and ecosystem roles (e.g. granivores, predators, omnivores) that can be separately analyzed. 160

161 According to these criteria, ants provide an ideal and important taxon to use for exploring ISARs. With over 13,800 species distributed globally, ants are one of the world's most dominant 162 163 arthropod groups (Bolton, 2020; Hölldobler & Wilson, 1990). They are also one of the most well-sampled invertebrates in insular systems. Their diversity is immense but well-studied; 164 authoritative taxonomic keys and identification materials permit accurate species identification 165 166 even in remote areas of the world. Ant trait variation has been well studied and defined (Gibb et al., 2015; Sosiak & Barden, 2020), allowing for the opportunity to disentangle variation in the 167 168 SAR due to trait diversity. Ants also have the ability to structure entire ecosystems (Lach et al., 2010) and can make up to 15-20% of terrestrial animal biomass (Schultz, 2000). Importantly, 169 ants show high levels of island endemicity. Ant island endemicity is known to vary greatly 170 across the globe (0% to 96% endemicity) and has been linked to the size, location, and isolation 171 of islands (Morrison, 2016). For example, over 70% (187 total species) of the ant fauna of the 172 173 island of Fiji is endemic (Sarnat & Economo, 2012), while in contrast, very few species are 174 endemic to the Bahamas (Morrison, 2003).

175 Based on the extensive literature on ant biogeography, a number of observations on ant 176 ISARs can be made. First, studies of biogeographical histories (Choi et al., 1993; Ranta et al., 1983; Trainor & Andersen, 2010) and speciation processes on islands illustrate the variability of 177 the ant ISAR across the globe (Economo et al., 2017; Economo & Sarnat, 2012; Wilson, 1961). 178 179 Second, habitat diversity exerts a persistent and significant influence (Boomsma et al., 1987; Goldstein, 1975; Torres & Snelling, 1997;. Wilson, 1961) as does elevation on insular (Morrison, 180 181 1997) or mainland (Sanders, 2002) ant diversity. Third, while dispersal is known to be an important driver of ISAR form (MacArthur & Wilson, 1967), its effect is less well studied in ants 182 compared to other taxa. Fourth, and linked to the third point, anthropogenic transport breaks 183 184 down natural barriers which has resulted in the colonization of many remote islands by nonnative ants (Espadaler & Bernal, 2003; Rica et al., 2005; Wetterer & O'Hara, 2002). The
resulting presence of non-native ants on many islands (Espadaler & Bernal, 2003; Rica et al.,
2005; Wetterer & O'Hara, 2002) can directly influence the ISAR (Wilson & Taylor, 1967).

Based on this previous literature, we can make certain predictions of how ant ISARs may 188 differ based on a series of extrinsic variables. Research across both vertebrates and invertebrates 189 190 show that mainland SARs often have lower z-values and higher c-values compared to insular systems (Matthews et al., 2016). This is attributed to the greater isolation of insular systems. 191 192 Given an ant's ability to take advantage of human-assisted dispersal (Pyšek et al., 2017) and to 193 establish in impoverished areas we expect more similar patterns in z and c-values between 194 mainland and insular systems than for other taxa. Ant ISARs are also heavily influenced by habitat structure and diversity (Goldstein, 1975; Morrison, 2016). As such, we expect habitat 195 diversity to be a bigger driver of ant richness than area per se. Therefore, we expect certain 196 197 climatic variables that track habitat diversity to negatively affect the rate of which ant diversity increases with area. Finally, competitive interactions (Wilson & Taylor, 1967) and microhabitat 198 199 diversity (Goldstein, 1975; Torres & Snelling, 1997) could constrain the species richness of ants 200 on small islands, while processes such as speciation are more likely to influence diversity on larger islands (Economo et al., 2017; Economo & Sarnat, 2012; Wilson, 1961). The shift in 201 202 importance of these different mechanisms of island community assembly along the area gradient 203 could result in points of inflection in the ant ISAR. While ant ISARs have been well investigated, 204 no research has synthesized all known information on ant ISARs and doing so would help highlight consistent findings and identify knowledge gaps. 205

This study aims to provide a general synthesis of ant ISARs by i) reviewing the ant ISAR 206 literature in order to identify significant data gaps, and ii) quantifying global patterns of ant 207 208 ISARs through the synthetic analysis of mainland and insular ant ISAR studies. We summarized 209 and compared ant ISAR *c*-values and *z*-values across multiple datasets to derive results that can be compared to previous research on ISARs. Specifically, we answered the questions: (1) how do 210 insular and mainland *c*-values and *z*-values differ for ants? (2) do climatic variables or isolation 211 212 influence the z-value of the ISAR for ants? (3) are z-values associated with latitude? (4) are there thresholds in ant ISARs? and (5) is there a general z-value for ants when accounting for variables 213 214 such as biogeographic realm?

- 215
- 216

217 Methods

218219 Data Collection

220 We acquired species-level datasets on ants from published articles, primarily species checklists, insular or mainland studies of the ISAR, and island biogeography studies. These data were used 221 222 to create two 'datasets' (datasets 1 and 2) that formed the basis of the analyses. Dataset 1 was constructed to assess the ant ISAR at the archipelago level which permits the analyses of z and c 223 values. Dataset 2 was constructed to include all individual islands from all datasets to seek 224 225 general patterns of the ant ISAR. For dataset 1 (see below), a total of 36 studies and 41 datasets were collected (Figure 1A). While dataset 2 comprised a total of 44 studies and 51 datasets. 226 227 Datasets were recovered using the Web of Science repository, FORMIS (A Master Bibliography of Ant Literature, USDA), scanning of supplementary data of review papers assessing the SAR 228 in multiple taxa (Drakare et al., 2006; Matthews et al., 2016), and unpublished manuscripts 229 230 (James Wetterer, pers comm.). The following variables were collected from each dataset when

available: species richness per island, total species richness of all islands, island/fragment area size, z-value of the SAR, standardized errors of the slope coefficient, R^2 values from the log-log

- 233 power model, locality of island/fragment, types of island/fragment, island/fragment area ranges
- 234 (m²), biogeographic realm, latitude, longitude, climate, percentage of non-native ants, and
- isolation (distance to the nearest continent for true island datasets) (Table 1). Percentages of non-
- antive ants were limited to 25 datasets that provided full species checklists or explicitly stated
- non-native ant richness in the publication. If a given true island dataset was comprised of both
- oceanic and continental islands, then the dataset was assigned the island type comprising the
 majority. If full datasets from publications were not available, the corresponding author(s) of the
- study were contacted for access to the original raw data. The supplementary material (Appendix
- 241 2 and 3) includes both datasets and a list of literature citations.
- 242

243 Species-Area Slopes and Intercepts

Whenever possible, z-values were recalculated using a study's raw data (natural logs were used). 244 245 Island/fragment areas and species richness values from the studies were used to calculate the zvalue using the log-log form of the power model (equation 1.2) with ordinary least squares 246 (OLS) regression. However, if raw data were unavailable then the reported z-value was used. In 247 248 total, 30 out of 41 z-values were recalculated. For non-recalculated z-values, only z-values of the 249 log-log model was used, as the majority of the data collected from the literature utilized that 250 model form. We used a paired t-test to assess the robustness of recalculated z-values compared to reported z-values (8 paired values). However, we also undertook multi-model analyses of 251 competing SAR models to discern the best-fitting SAR model (outlined below). None of the 252 253 studies contained zero values; as such, no added constants were necessary for log-254 transformations. If island/fragment areas were not provided, data were recovered from online 255 sources or manually calculated using the Google Earth engine. ISARs fitted with log-log axes 256 have intercepts that are not independent of the units in which area is measured. In order to mitigate this issue, c-values were recalculated after standardizing area across datasets to the same 257 258 units $(1-km^2)$, which resulted in 30 comparable intercepts.

- 259
- 260 *Abiotic data*
- 261 Climate data, including mean annual temperature (MAT) and mean annual precipitation (MAP),
- were extracted from coordinates of study locations using the Worldclim database (1 km^2)
- resolution) (Fick & Hijmans, 2017). If climate data were unavailable at specific coordinates then
- data were collected using island climate data from Weigelt et al. (2013). For studies that spanned
- islands across large latitudinal or longitudinal gradients (e.g. New Guinea to the Hawaiian
- 266 Islands), an averaged value of the MAT and MAP from the individual study islands available
- from Weigelt et al. (2013) was assigned. This approach was considered more accurate than
- 268 choosing coordinates at a midpoint and extracting the MAP/MAT using the Worldclim database.
- 269 Isolation for the true islands was calculated by measuring the distance of the island closest to a
- 270 continent from a dataset. Not all studies included island location or names and therefore isolation
- values were limited to 29 datasets.
- 272
- 273 Final Datasets for Analysis
- As stated above, two datasets were collated for the analyses: Dataset 1 was used to assess the
- best fitting model for ant ISARs using a multi-model inference approach, the *c* and *z*-values of
- each study, and the effects of isolation and climate on z-values. Dataset 2 was used to calculate a

277 general z-value from all islands and evaluate the potential for thresholds or points of inflection in the ant ISAR. Dataset 1 included c and z-values for each study, isolation, biogeographic realm, 278 279 difference in order of magnitude in island area range per study (calculated as: Log₁₀(Maximum Area / Minimum Area)), general latitude and longitude, MAP, MAT, standard errors of the 280 slope, R^2 of the log-log model, and ISAR type. The ISAR type variable specifies whether the 281 282 study was done on the mainland or in an insular system (true islands). Dataset 2 focused on individual islands and included data from each island in studies that provided areas and species 283 284 richness values for each island. Dataset 2 included studies not in dataset 1 as some studies only inventoried ants on a single island. If any duplicate islands appeared in dataset 2 (multiple 285 studies sampling the same island) then the most recent study and values were used. Variables 286 included with this dataset were: island area, species richness, SAR type, and biogeographic realm 287 288 of the island.

- 289
- 290 Analysis
- 291

292 Dataset 1

Dataset 1 comprised 41 datasets and their z-values. Only datasets with positive z-values were 293 294 considered. Four formats of dataset 1 were used for four different questions. The first format of 295 dataset 1 included the individual islands areas and species richness values. This format was used 296 to assess the robustness of the power model in fitting the datasets, and to determine whether 297 other model forms better represent ant ISARs. The second format used each dataset's z-value and abiotic data to determine differences between mainland and insular ISARs, as well as the effects 298 299 of abiotic variables on z-values. The third format used only true island (insular) studies with 300 isolation values to assess the effects of isolation on z-values. The final format used only the 301 datasets where c-values were recalculated using the same measurement units (km²). This dataset 302 was used to determine potential differences in *c*-values between mainland and insular studies.

To assess the robustness of the power model, a multi-model inference approach was 303 304 executed utilizing 20 different ISAR models (including the power model in its non-linear form) using the "sars" R package (Matthews et al., 2019) and following the approach outlined in 305 306 Matthews et al. (2016) and Triantis et al. (2012). The 20 models represented a range of ISAR forms including linear, convex, and sigmoidal. During maintenance of the 'sars' package while 307 preparing these analyses, it was discovered that the He & Legendre logistic model and the mmf 308 model were equivalent (see also Williams et al., 2009). Thus, as an alternative to the mmf model, 309 we used the standard logistic model (see Tjørve, 2003). Models were fitted to each study's 310 distinct dataset and compared using AIC_c (Akaike's Information Criterion corrected for small 311 sample sizes). The models were fitted twice to each dataset (areas were all first converted to 312 313 km²): once when implementing model residuals validation checks (for normality using a Lilliefors test, and homogeneity of variance using a Pearson's correlation of the squared 314 residuals with the area values) and once with no checks. For the former, if a model failed one or 315 both of the two checks it was removed from the model comparison for that dataset (see Matthews 316 317 et al., 2019). When fitting models, we used a grid search procedure (setting the grid search 318 argument to 'exhaustive' in the sar average function) to test a large number of starting parameter 319 values (5000 combinations for each model) in the non-linear regression optimization algorithm; this increases the likelihood that optimum parameter estimates are found. 320

The performance of individual models across datasets was calculated by comparing the mean AIC_c weight and the mean rank (i.e. a model's rank in the model comparison for each dataset) (Triantis et al., 2012). When model residuals tests were used, the mean rank and weight
of a model across datasets were calculated after removing the cases where the model did not pass
the tests. It was necessary to remove datasets with fewer than seven observations, the minimum
number required to calculate AIC_c for models with four parameters. This resulted in 25 datasets
being analyzed. All models were fitted using untransformed data.

To assess differences between the z-values of insular and mainland datasets a generalized 328 329 linear model (GLM) with a Gamma (link $= \log$) error distribution was run using the z-values as 330 the response variable, and study type as a predictor. The Gamma distribution was chosen as zvalues were not normally distributed and error distributions used for skewed data such as the 331 Poisson or negative binomial assume a discrete error distribution. In addition, because mainland 332 islands were smaller than the majority of true islands across the dataset, a secondary model (with 333 the same specifications) including a covariate of mean area of the datasets was run with a 334 335 reduced dataset of 30 observations (Number of data sets with a mean area). Model assumptions 336 were checked using residual diagnostic plots.

337 To assesses the potential effects of climate or latitude on z-values, GLMs with a Gamma (link = log) error distribution were used. Two model structures were used to incorporate two 338 different covariates. The first model used the order of magnitude in island size range per dataset 339 340 as a covariate, as potential variation in slope values corresponding to different ranges in areas 341 among different datasets should be considered (Martin, 1981). We included interaction terms 342 between the covariate and climate variables because of the non-linearity that is observed in how biodiversity increases across space (Keil & Chase, 2019). As such, the effect of climatic factors 343 on the rate at which species richness increases with area is not constant across scales. Seven of 344 345 the 41 datasets had no island area ranges (only z-values were provided with no island area data), 346 and as such the final modeling dataset here used a total sample size of 35.

347 The second model structure utilized the SAR type (insular or mainland) as a covariate, as z-values were observed to significantly vary by this category. Models used with this structure 348 had a sample size of 41 as all datasets had a specified SAR type. With this model structure, 349 350 interactions between the covariate and climate variables were not included because the effects of 351 climate were presumed to be the same in mainland and insular systems. Both covariates could 352 not be used in the same model because the low sample size would have resulted in lower 353 statistical power and an increased risk of model overfitting. Mixed-effect models were not used because the low sample size resulted in singularity in the models. 354

Other predictors in the two model structures included MAP, MAT, latitude, and 355 biogeographical realm. Models never included MAP and MAT variables together due to 356 357 collinearity. Specific model structures are detailed in Table 3. The maximum number of 358 predictors allowed in a model was three to avoid overfitting. All models were assessed using 359 generalized variance inflation factors (VIFs), with a particular focus on models with both biogeographical realm and climate variables. All continuous variables were scaled (to a mean of 360 zero and unit variance) prior to running models. Model comparison was conducted using AIC_c 361 for each set of models corresponding to the two different covariates. Pseudo-R² values were 362 363 calculated using the piecewiseSEM package in R.

Isolation values were measured for 29 insular datasets. To assess the impacts of isolation on *z*-values, a GLM with a Gamma (link = log) error distribution was used. Thirty of the 41 datasets had re-calculated *c*-values. OLS regression was used to test for differences in these *c*values between insular and mainland studies. The Gamma distribution was not implemented because it does not accommodate negative values. Data were not transformed because *c*-values 369 were already on a log-scale. Similar to the second GLM used to account for mean area when

assessing differences between the *z*-values of insular and mainland datasets, a second model was

run on *c*-values with a covariate of mean area per *c*-value dataset. To verify that all model

assumptions were met, model residuals were evaluated for normality and homogeneity ofvariance.

Finally, we re-ran models with datasets that contained 10 or more islands. This was done in an effort to reduce the impact of datasets that did not survey enough islands and to offer a potentially more biologically meaningful perspective on c and z-values. All analyses were

implemented in R (version 4.0.0) (R Development Core Team, 2018).

378

379 *Dataset 2*

380 Dataset 2 comprised all individual islands with area and species richness values. This dataset

permitted us to provide broad assessments of ISARs, utilizing information from singleton islands

not themselves part of an explicit SAR study. Linear mixed-effect models were used to seek a $\frac{282}{1000}$

383 general power model (based on the $log(SR) \sim log(Area)$ form) slope coefficient. Natural logs

were used to calculate both $\log(SR)$ and $\log(Area)$. Not all studies provided island-specific data,

but a total of 481 islands (both mainland and true islands) with species richness observations

were collected. There was a distinct difference in species richness and area between mainlandislands and true islands (Figure 2B); the two types of island are also likely affected, to different

islands and true islands (Figure 2B); the two types of island are also likely affected, to different
 degrees, by different biogeographical processes (e.g. dispersal, speciation). Furthermore,

mainland islands spanned a distinctly smaller area gradient compared to true islands. Therefore,

dataset 2 was further divided into mainland (n = 161) and true island (n = 320) subsets, which

391 were analyzed separately with two different mixed-effect models. In both models, log species

richness was assessed as a function of log area with biogeographic realm used as a randomeffect.

The model implemented with the mainland subset failed to meet model assumptions based on visual diagnostics of residuals and therefore results and discussion of this model are not included in this study. Failure to meet assumptions was likely due to an absence of a linear trend between log(SR) and log(Area).

All mixed-effects models were run using the lme4 package (Bates et al., 2015).
Visualization of fixed and random effects were conducted using the 'R' packages "ggeffects"
(Lüdecke, 2018) and "ggplot2" (Wickham, 2009). Pseudo-R² values (Marginal and conditional
R²) were calculated using the "MuMin" package (Barton, 2019). AIC_c values were calculated
using the package "bbmle" (Bolker & R Development Core Team, 2020).

403 In a separate analysis, threshold models were fitted to the insular and mainland subsets to identify any potential points of inflection in their ISARs. We fitted two piecewise models (the 404 405 continuous and left-horizontal one-threshold models) using new functionality in the "sars" R 406 package (Matthews & Rigal, 2021). Both models have two segments. The continuous model allows both segments to vary in slope, while the left-horizontal model assumes the first segment 407 to have a slope of zero, which is characteristic of the classic small island-effect (MacArthur & 408 409 Wilson, 1967; Whitehead & Jones, 1969), where there is a lack of relationship between area and 410 richness on smaller islands. The models were fitted in both log-log space (area and richness log 411 transformed; natural logarithms) and semi-log space (area log transformed) as the choice of log-412 transformation has been found to influence the fit of threshold models (Burns et al., 2009;

413 Matthews et al., 2014). In both cases, the two threshold models were compared (using AIC_c)

414 with a linear model (i.e. either the logarithmic or power SAR models depending on

415 transformation) and an interecept-only null model. Following Matthews et al. (2014), we

- 416 checked for influential observations in the threshold model fits using Cook's distance and a
- threshold of one. Threshold models were not fit to individual datasets (i.e. those in dataset 1) due
- 418 to the smaller sample sizes involved. Finally, to determine whether the observed piecewise
- relationships were driven by biome effects, we fitted a mixed-effects piecewise regression model
 using functions written by the author of the 'segmented' R package (Muggeo, 2016; Muggeo et
- 420 using functions written by the author of the segmented K package (Muggeo, 2010, Muggeo et 421 al., 2014). For ease, we focused on the continous one-threshold model (semi-log transformation)
- 422 for the insular subset. Log-transformed area was the fixed effect, and biogeographic realm was
- 423 used as a random effect. We fitted two variants of this model: i) a model with just a random
- 424 intercept for biogeographic realm, and ii) a model with both a random intercept and a random
- 425 breakpoint for biogeographic realm. An identity regression model for the changepoint was used
- 426 (Muggeo, 2016). As we were interested in the random effects and were not comparing models,
- 427 we fitted the models using restricted maximum likelihood.
- 428
- 429 **Results**

430

431 Dataset 1

432 Of the 41 total datasets examined for this study, z and c values were recalculated for 30 of them. 433 Of the 11 remaining datasets, all reported z-values (log-log power model slopes) in their 434 associated papers, none reported standard errors of z-values, and 7 reported R^2 values and cvalues (not used for analyses). The paired t-test showed no significant differences between 435 recalculated and reported z-values. In total, 30 datasets were from insular systems and 11 were 436 437 from mainland systems. The majority of the mainland datasets (7 of 11 datasets) were from 438 fragments due to anthropogenic activities. The biogeographical realm with the most studies was the Palearctic (13) and the realm with the fewest was the Afrotropics (2) (Table 2, Figure 2A). 439 440 Non-native ant percentage ranged from 0 to 87% and the average percentage of non-native ants in mainland and insular datasets was 2.82% and 29.90% respectively. Continental and oceanic 441 442 island datasets had an average non-native percentage of 18.7% and 34.6% respectively. The 443 biogeographical realms with the highest non-native ant percentage in mainland and insular 444 systems were the Palearctic (5%, one dataset) and Oceania (average 51%, 3 datasets) 445 respectively.

446 When no residual checks were used, the power model was the best overall model across 447 datasets, according to both mean model rank and mean AIC_c weight (see Fig. S1 in Appendix 1). The model provided the best fit (lowest AIC_c) to 11 of 25 datasets and was ranked in the top five 448 449 for 20. When model validation was implemented, the power model rank shifted slightly (Fig. 450 S1). It was the second-ranked model according to mean rank (behind the Kobayashi model; also a convex non-asymptotic model) but was still the top-ranked model according to mean AIC_c 451 452 weight. With model validation, the power model provided the best fit to eight datasets and was ranked in the top five for 15; for seven datasets it did not pass validation checks. Overall, these 453 results indicate that power model represents a generally robust model for characterizing ant 454 455 ISARs and allow us to confidently analyze power model z-values across our datasets.

456 Mainland islands had the highest variance in log–log power model *z*-values (Figure 2C). 457 The number of islands or fragments used to generate a slope value from a study ranged from 3 to 458 81. The range in difference in order of magnitude between the smallest and largest island areas 459 was 0.70 to 6.17. The minimum *z*-value was 0.04, the maximum was 0.89, and the average was 460 0.28. *z*-values varied by biogeographic realm and by ISAR type (insular or mainland). The mean 461 z-value for mainland studies (0.38 ± 0.08) was significantly higher than insular studies (0.25 ± 0.08) 0.02). This was verified with a GLM (effect size of mainland studies relative to insular studies: 462 0.43 ± 0.20 , p < 0.05, Pseudo-R² = 0.07). The second GLM run on the reduced dataset showed 463 no significant effects from either covariate (mean area) or SAR type on z-values. While the 464 covariate of mean area did not significantly impact z-values, its inclusion reduced the available 465 466 degrees of freedom. This coupled with the reduction of observations for this model (11 less observations) resulted in lower statistical power to potentially detect a difference between 467 468 mainland and insular z-values.

469 Nine competing GLMs, including a null model, were fitted to evaluate effects of climate 470 and latitude on z-values while accounting for the order of magnitude in island area range. Latitude effects were not observed in any of the models. The most plausible model based on 471 AIC_c was a model evaluating z-values as a function of the additive effects of order of magnitude 472 in island area (covariate) and MAP (Pseudo- $R^2 = 0.11$). The covariate had a non-significant 473 effect, but MAP had a significant negative effect on z-values (-0.24 ± 0.11). It's important to 474 475 note that the second ranked model (a model with only the covariate as the predictor) was within a 476 ΔAIC_{c} of 2 suggesting that both models were commensurate with one another. While the second 477 ranked model had one less parameter, the top-ranked model was still designated as most plausible given the higher pseudo- R^2 . However, we still exercised caution when interpreting this 478 479 model.

480 Seven models were run to assess the effects of climate and latitude on *z*-values while 481 accounting for ISAR type (Table 3). The most plausible model based on AIC_c was one that 482 evaluated *z*-values as a function of the additive effects of ISAR type (covariate) and MAP 483 (Pseudo- $R^2 = 0.18$). There was a non-significant effect from the covariate but a significant effect 484 of MAP on *z*-values (-0.26 ± 0.09).

485 Mainland ISARs had significantly higher *c*-values compared to insular ISARs (effect size of mainland studies relative to insular studies: 2.71 ± 0.66 , p < 0.05, Adjusted-R² = 0.35). 486 Average *c*-values for mainland and insular ISARs were 4.55 ± 0.86 and 1.44 ± 0.26 , 487 488 respectively. The second model that accounted for mean area when assessing differences 489 between insular and mainland datasets showed a significant effect of the covariate but no 490 significant effect of SAR type. Across biogeographical realms that had both insular and mainland ISAR types (Afrotropics, Nearctic, Neotropic, Palearctic), c-values were lower in mainland 491 492 systems only in the Nearctic (see Table S1 in Appendix 1). Finally, no significant effects of 493 isolation on z-values were observed.

494 Models re-run on datasets with 10 or more islands showed no difference between 495 mainland and insular z-values and no effect of isolation on z-values. There was a significant 496 difference between mainland and insular *c*-values with mainland *c*-values being on average higher than insular *c*-values (effect size of mainland studies relative to insular studies: 3.07 ± 0.9 , 497 p < 0.05, Adjusted-R² = 0.43). AIC_c rankings of GLMs assessing the effect of climate and 498 latitude with two different covariates showed the top ranked models both included a significant 499 and negative effect of MAT on z-values. However, both models were within a ΔAIC_c of 2 of the 500 501 null models (z-slope \sim 1). As such these models are not discussed further but their model outputs 502 and AIC_c rankings have been made available in the supplementary materials (see Tables S2-S4 503 in Appendix 1). 504

505 Dataset 2

506 Of the 481 islands in dataset 2, 161 were mainland islands, and 320 were true islands. True 507 islands types were made up of 160 oceanic and 160 continental islands. The linear-mixed effect 508 model evaluating the ISAR for insular data showed a significant effect of area on species 509 richness (model coefficient: 0.16 ± 0.03) (Figures 1B & 1C). Based on marginal and conditional 510 R^2 values, area explained 29% of the variation in species richness. However, when including the 511 random effects of biogeographic realm, the model explained 60% of the variation (difference = 31%). The lowest random intercepts were from Oceania and the Palearctic while the highest 512 513 intercepts were from Indomalaya and the Afrotropics (Figure 1C). The steepest random slopes were found in the Nearctic and Oceania while the lowest random slopes were in the Indomalaya 514 515 and Afrotropics.

Based on AIC_c, when using both the log–log and semi-log transformations, the two 516 threshold models both provided a better fit to the global insular dataset than the linear (log-log 517 518 power and logarithmic model, respectively) and intercept-only null models (Table 4), with the continuous one-threshold model providing the best fit in both cases. The R² value of the best 519 520 threshold model for the insular dataset ranged from 0.36 to 0.54, depending on the 521 transformation. Figure 5 shows the best threshold model fits to the insular dataset, using both the log-log and semi-log transformations. There was one threshold model fit (left-horizontal model 522 523 fit to the insular data in semi-log space) with a Cook's distance greater than one. However, we 524 left the data point in as it was only slightly greater than one (1.09) and removing it did not 525 change the overall results or general model fit. More generally, the insular semi-log threshold model fits should be viewed tentatively as their residuals deviated from normality and/or plots of 526 the fitted values against the residuals tended to show that the magnitude of the residuals was not 527 528 constant along the fitted values. For the global mainland dataset, the continuous one-threshold 529 model provided the best fit to the log-log data, while the left-horizontal model provided the best 530 fit to the semi-log data (Appendix 1 Table S5). However, the mainland threshold model fits were harder to interpret, with the R² value of the best threshold model ranging between 0.20 and 0.23, 531 some of the model residuals deviating from normality and exhibiting heteroscedasticity, and the 532 533 relationships seemingly driven by biome effects (see Figure S2). As such, these model fits are 534 not discussed further.

535 The mixed-effects piecewise models fitted to the insular subset revealed that the 536 population estimate (i.e., the fixed effect piecewise relationship) provided a good rough approximation for all the biogeographic realms except for the Nearctic, and to a lesser extent the 537 Afrotropics (the plots of these model fits are provided in Figure S3-S4 in Appendix 1). In the 538 539 case of the Nearctic, the lack of an obvious threshold relationship was likely due to the limited 540 range in island area, that is, there were no islands near or above the breakpoint value(s) observed for the other biogeographic realms (Fig. S3). For the model where the breakpoint varied by 541 biogeographic realm, the standard deviation of the random breakpoint was 2.6 (on the log-scale). 542 543 This variation was affected by the value for the Nearctic realm, where the breakpoint is much lower and is likely a statistical artefact (again due to the lack of large islands in our dataset for 544 this realm). In addition, the random breakpoints for the Afrotropics and Indomalaya were 545 546 considerably larger than the population estimate (Fig. S4). Overall, these results indicate that, although the intercept and the exact breakpoint value do vary to a degree between realms, the 547 548 observed threshold relationship in the insular ISAR is not simply driven by islands from one or 549 two realms.

- 550
- 551

552 Discussion

553 Studies of the ISAR in ants extend over the past 60 years, spanning all biogeographical realms, 554 except Antarctica. Here, we demonstrate that the scaling of ant species richness with area can 555 differ between mainland and insular systems, and along precipitation gradients. We also show 556 that incorporating biogeographical realm types into analyses provides a better explanation of this 557 scaling pattern than does area alone, suggesting a macroevolutionary perspective on species–area 558 relationships.

559

560 Dataset 1: Mainland vs. Insular Systems

We found that z-values, the rate at which log species richness increases with log area, vary 561 widely across studies of ant communities, highlighting the extent of variation in the z-value of 562 the ISAR within a single taxonomic group. An intriguing finding from our analyses is the 563 564 significant difference in average z-values between mainland and insular systems, a result which contrasts with results of an extensive meta-analysis which did not detect differences in z-values 565 566 (Drakare et al. 2006). Whereas past studies reported steeper z-values in insular systems than in mainland systems (MacArthur & Wilson, 1967; Matthews et al., 2016; Triantis et al., 2012; 567 Whittaker & Fernandez-Palacios, 2007), in this study mainland ant ISARs have steeper z-values 568 compared to insular systems (Figure 2B). While area effects can influence z-values given that the 569 570 majority of mainland islands were smaller than true islands in our data, we found a lack of 571 significant effect of mean area as a covariate, when analyzing a reduced dataset. Higher z-values in mainland systems may reflect a biological pattern or could be explained by sampling bias in 572 our dataset, as a higher proportion of mainland studies (50%) were conducted in the Neotropics, 573 574 a biogeographical realm considered to include the world's highest ant diversity (Hölldobler & 575 Wilson, 1990). If species geographical ranges are typically smaller in species rich faunas, one 576 would expect richness to tend to increase more strongly with area, for studies that span much of biogeographical provinces. However, mainland datasets in our study only focused on habitat 577 patches or fragments where dispersal is likely not as limited as in insular systems (Storch, 2016), 578 579 and this range size effect would not likely be evident at small spatial scales.

580 Another possible reason for higher z-values in mainland systems is the differences in ant 581 dispersal success and propagule pressure between mainland and insular systems. Insular systems often have higher levels of isolation for longer periods of time. As a result, when compared to 582 true islands, mainland islands often hold a larger portion of the regional species pool (Flantua et 583 al., 2020). Furthermore, ant species that successfully disperse and colonize insular systems are 584 usually generalists and are likely to be excellent dispersers (Morrison, 2016). Insular studies 585 586 often gauge dispersal possibilities of ants based on isolation distance between sampled islands (Woinarski et al., 1998); other studies view dispersal to be a hidden variable that cannot be 587 quantified accurately (Badano et al., 2005). While one study showed that increasing isolation 588 589 actually led to higher aboveground ant species richness in naturally fragmented landscapes (Cuissi et al., 2015), this relationship was not observed with arboreal ants, suggesting that 590 isolation may affect ants in various strata differently. Moreover, natural dispersal may be 591 592 irrelevant in some insular systems where human colonization of islands has led to humanassisted dispersal of ants and dominance of local ant communities by introduced species 593 594 (Holway et al., 2002; Pyšek et al., 2017; Wilson & Taylor, 1967). Although dispersal is likely a 595 key element underlying ISARs, the sparse literature on dispersal effects on ISARs points to the difficulty in quantifying this factor (Hakala et al., 2019). Clearly, more sampling in mainland 596

islands across different biogeographical realms is needed to fully understand the observeddifferences between mainland and insular slopes.

599 We observed lower *c*-values in insular systems than mainland systems indicating higher degrees of island impoverishment or lower biotic richness in local communities on true islands 600 (Whittaker & Fernandez-Palacios, 2007). This was also observed in datasets with only 10 or 601 602 more islands. When accounting for mean area of the islands per dataset, mean area significantly affected *c*-values while differences between mainland and insular studies became statistically 603 604 negligible. This suggests that area differences between mainland and true islands contributes to differences in *c*-values. Island impoverishment levels likely impact the colonization process of 605 606 insular ants compared to ant dispersal which often relies on the capitalization of habitat resources (Morrison, 2016). For example, impoverished true islands may have less habitat heterogeneity or 607 food resources specifically for ants due to lower plant biotic richness which can result in a lower 608 609 likelihood of successful colonization, less opportunity for speciation or higher extinction rate. Identifying the causal underpinnings of the differences in c (local richness) and z (spatial scaling 610 611 of richness) between continental and insular systems is an important challenge for future work.

Higher z and c-values found with the mostly anthropogenically created mainland 612 613 fragments from this study potentially opens exciting avenues of future research, such as the effect of fragment age on the mainland ISAR. Such research could develop a continental analog 614 615 of the general dynamic model of island equilibrium theory (Whittaker et al., 2008), which links 616 the temporal development of volcanic islands to the equilibrium theory of island biogeography. Previous research has already utilized landscape histories (e.g., how a landscape becomes 617 fragmented over time) to develop models capable of predicting species extinctions and 618 biodiversity in fragmented landscapes (Ewers et al., 2013). However, we still lack a general 619 620 framework for understanding how fragment age affects different biodiversity properties. Any 621 framework would need to account for extinction debts, especially those in smaller fragments, and the quality of habitats within different-sized fragments. Differences in how gradual 622 fragmentation versus pulse fragmentation events influence biodiversity patterns, similar to the 623 624 differences observed between continental and oceanic islands, could also be incorporated. 625 Perhaps, the integration of these different phenomena into a single framework could lead to the 626 development of a general dynamic model of anthropogenic landscapes.

627

628 Dataset 1: Climatic drivers of the ISAR

Another clear result from this study is the finding that higher precipitation levels lead to a

630 significant decrease in z-values. This was observed from the results of generalized linear models

631 that predicted z-values as a function of a covariate (either ISAR type or order of magnitude

632 island range) and precipitation (Figure 3). Precipitation is known to influence aboveground plant

biomass, habitat diversity (e.g., more vegetation strata and litter with increasing precipitation),

- and net primary productivity (NPP) (Yan et al., 2015; Zhu et al., 2014). Considering
- precipitation as a proxy for NPP, our results are similar to those reported in Storch et al. (2005), r_{2}^{2}
- 636 who identified a negative relationship between productivity and the ISAR z-value. Lower R^2 637 values are present at higher levels of precipitation (see Figure 4), illustrating that area effects
- explain less of the variation in ant species richness as precipitation increases. Our results also
- agree with other studies on insular ants that show the effect of habitat diversity as being as
- 640 important as area when predicting species richness (Goldstein, 1975; Torres & Snelling, 1997).
- 641 One mechanism through which increased precipitation could weaken area effects would 642 be a boost in overall abundance, per species, which could lower extinction rates or heighten

643 colonization rates in habitat patches or islands. Another mechanism would be that increased 644 precipitation can yield a higher diversity of microhabitats for ants at a local scale; the diversity 645 and available biomass of plants is crucial, providing a diversity of resources through both plant-646 ant interactions and heterogeneity in the range of suitable nesting sites (Hölldobler & Wilson, 647 1990; Lach et al., 2010). Unraveling the effects of precipitation on both macro- and microhabitat 648 diversity in insular systems that vary in spatial areas will help us understand the scales at which 649 abiotic conditions significantly impact biodiversity.

650 Morrison (2016) indicated that isolation may provide only weak effects on ant species 651 richness at small scales but may be more important at larger ones. This would be logical as dispersal distance among habitat islands in a continental setting is likely modest compared to the 652 distances relevant to (most) oceanic islands. While our analyses show that isolation had no 653 significant effect on the z-value despite isolation distances spanning from 0.348 to 4.820 654 655 kilometers, isolation cannot be discounted as an important component to ant community assembly in all true island systems. Further discussion on the potentially weakening effects of 656 657 isolation as a barrier to dispersal in ants are provided later in this discussion.

Overall, two clear results emerged from the analyses of dataset 1. First, z-values of
mainland ISARs appear to be significantly *higher* than insular ISARs. Second, increases in
precipitation lead to *lower z*-values in ant ISARs. Therefore, the extent of area effects on ant
ISARs is dependent on both abiotic conditions and whether or not one is considering 'true'
islands, or habitat islands (influences that have been noted in previous studies, e.g. Drakare et al.,
2006; Kalmar & Currie, 2007).

664

665 Dataset 2: A general species-area relationship in ants

666 To arrive at a global-scale z-value for ants, we implemented a linear mixed-effect model with 667 320 true islands and their respective ant species richness and found a global and general z-value of 0.16 for ants. This value is lower than what is expected from general ISARs observed in other 668 taxa, especially in true islands at regional scales (Rosenzweig, 1995). However, the random 669 670 slopes for the 7 biogeographical realms ranged from 0.06 (Indomalaya) to 0.25 (Oceania). This variation in random slopes across realms and the improvement of explanatory power when 671 672 accounting for biogeographical realms as a random effect, implies that the ant ISAR is not 673 constant, but varies according to different realms (Figure 1C). Differences between biogeographic realms are expected to influence the ant ISAR as each realm has a unique 674 geological and evolutionary history with different aged ant faunas and historical pattern of 675 speciation, which could lead to different ISAR patterns (Choi et al., 1993; Ranta et al., 1983; 676 677 Trainor & Andersen, 2010). The low random slope in the Afrotropics may be an artifact of small sample size as only 15 of the 320 islands come from this realm (this realm has relatively few true 678 679 islands), while lower random intercepts for Oceania and the Palearctic suggest systematically 680 lower alpha diversity in those regions. In Figure 2, the outlier among realms is Oceania, which has considerably steeper ISARs than the other realms. In the Nearctic, all islands that were 681 sampled were continental islands (n = 23) while 57% of islands (n = 65) in the Neotropics were 682 683 oceanic islands but both realms held similar random slope values (Nearctic 0.20, Neotropics 684 (0.19). However, the majority of islands from the Neotropics come from areas close to the 685 northern limits of this realm (e.g. Sea of Cortez, West Indies) which may explain the similarity in

686 slope values.

In regard to the threshold in the global true island ISAR (Fig. 5), the threshold (excluding
the left-horizontal model in log–log space) was identified at roughly 133-163 km². The mixed-

689 effects piecewise model fit revealed that the observed threshold did vary between biogeographic realms (e.g., being larger for Indomalaya), but, with the exception of the Nearctic and possibly 690 the Afrotropics, the general pattern was consistent. This finding does not mean such a threshold 691 does not exist for Nearctic islands, but simply that our data did not include a wide enough range 692 of island areas to test this proposition. These thresholds, or points of inflection, can indicate 693 694 specific spatial scales at which species richness accumulates at different rates across areas, 695 reflecting processes such as the "small-island effect" (MacArthur and Wilson 1967; a 696 distinctiveness of habitats on small islands, e.g. edge effects in habitat fragments). The high 697 threshold observed for insular systems here could indicate an important role for speciation in 698 driving ant diversity on large true islands. Speciation events can strongly influence the ISAR in larger true islands compared to smaller-sized island systems in ants (Economo et al., 2017; 699 Economo & Sarnat, 2012; Wilson, 1961). Speciation rates are positively correlated with island 700 701 area and, above a certain threshold, in-situ speciation can be a richer source of new species within an island than establishment from colonizing species (Losos & Schluter, 2000). 702 703 Speciation in natural true insular systems likely affects ant ISAR patterns because ecoevolutionary specialization in colonizing generalists can lead to in situ speciation, a pattern 704 705 consistent with the taxon cycle hypothesis (Economo & Sarnat, 2012; Wilson, 1961). These 706 speciation events within island systems can yield high levels of endemism in ants, levels that in 707 some cases are higher than those observed in birds or vascular plants (Andersen et al., 2013).

708

710

709 Gaps in the ant species–area relationship

711 Sampling gaps and data deficiencies

712 The rich history of research in ant biodiversity has resulted in a significant body of species-level 713 data on ant communities. Our analyses of these data clearly illustrate that different island systems exhibit different species-area relationships. However, as Figure 1A illustrates, the 714 studies are unevenly distributed across the globe, with scant studies from Africa and tropical 715 716 mainland Asia. These gaps highlight the regions of the world from which additional ant 717 biodiversity data would be especially valuable. It would be valuable to expand the range of 718 studies in anthropogenically generated islands (resulting from habitat fragmentation). Especially now that forests in particular are becoming increasingly fragmented around the globe, creating 719 multiple new continental 'islands' (Haddad et al., 2015) for examination of species-area 720 721 relationships.

722

723 Incorporation of habitat diversity & other variables

724 Future research aiming to understand the drivers of variation in the ant ISAR will need to 725 incorporate variables beyond area, a point extensively discussed in the literature (Chase et al., 726 2019; Triantis et al., 2003; Whittaker & Fernandez-Palacios, 2007). One clear way forward with ants is to consider variables such as habitat diversity in the context of competing hypotheses. It is 727 unarguable that habitat diversity is an influential factor affecting ant ISARs (Torres & Snelling, 728 729 1997; Wilson, 1961) and, in some cases, habitat diversity measures have been incorporated as 730 covariates within the log-log ISAR model (e.g. as variables quantifying soil clay content) 731 (Goldstein, 1975). Future investigations of ant ISARs should use ISAR models that implement 732 habitat diversity, such as the choros model, which utilizes habitat diversity indices within the power function by multiplying a habitat diversity index with area. This model has led to a better 733 734 fit of the ISAR when compared with conventional models (Triantis et al., 2003). Moreover, our

finding that higher precipitation leads to a shallower ISAR (see also Storch et al. 2005) suggests
broad-scale ecosystem drivers of the spatial scaling of species richness, a pattern that warrants a
clear mechanistic interpretation.

738

739 Sample size

The average number of islands used per study from this paper in insular and mainland systems 740 was 16 and 15 respectively. However, recent work indicates that a minimum of sample size of 25 741 742 is often necessary to identify significant trends of variation when there is high variance in the data (Jenkins & Quintana-Ascencio, 2020). That same study reported that 79% of 217 island 743 744 biogeography studies constructed species-area curves with fewer than 25 observations. While 745 acquiring insular data can be difficult, it is important that future studies aim to sample enough islands to reach a minimum sample size that can effectively quantify the ISAR. Although this 746 747 can be problematic in studies of true archipelagos that contain few islands (i.e. the small sample size is intrinsic to the system, not a sampling issue), it should not be a problem in continental 748 749 regions, where anthropogenic habitat fragmentation is creating an abundance of fragments 750 varying greatly in area and isolation.

751

752 Impacts of non-native ants on the SAR

753 Our ability to understand and predict the effects of non-native species on the ISAR is limited by 754 a paucity of knowledge of the interactions between non-native and native ants within insular 755 systems. However, observational evidence does exist of specific non-native and invasive ants, called "plague ants", completely displacing native ant populations in insular systems (Morrison, 756 2016; Wilson, 2005). Multiple studies have previously demonstrated that slope values of ISARs 757 758 can be different when native or non-native status of species are taken into account (Rica et al., 759 2005; Wilson & Taylor, 1967). Well-sampled areas like the Florida keys show specific non-760 native species as being dominant on certain islands while native species remain dominant on others (Wetterer & O'Hara, 2002). Such systems raise questions as to why certain islands may be 761 762 more susceptible to the establishment of non-native ants and how non-native ants influence the 763 ISAR both empirically and theoretically.

764 In our study, we show that non-native ants can comprise anywhere from 0% (Boomsma et al., 1987; Dean & Bond, 1990; Leal et al., 2012; Vasconcelos et al., 2006) to 87% (Sugiura, 765 2010) of the ant fauna. We also observed higher percentages of non-native ants in insular 766 systems compared to mainland systems, likely as a result of human-assisted dispersal often 767 favoring non-native species with generalist tendencies (Morrison, 2016). Generalist ant species 768 769 often have traits such as a broad range of nesting habitats (Fournier et al., 2019) that aid in surviving human-assisted dispersal and subsequent island colonization. In fact, Rizali et al. 770 771 (2010) clearly shows that non-native species track human settlements on the islands of West 772 Java, Indonesia. Human-assisted dispersal allows species to move over much greater distances than feasible by natural dispersal, resulting in colonization of more remote islands (often distant 773 oceanic islands) which may have impoverished native faunas. This is reflected in our data where 774 775 we observe almost double the non-native ant percentage in oceanic islands compared to 776 continental islands. While many ants can produce thousands of reproductive alates for dispersal, 777 the chance of any one successful colonization is extremely low. For example, Levins et al. 778 (1973) estimated the rate of successful colonization by queen ants to islands near Puerto Rico to 779 be $\sim 4\%$. In terrestrial mainland systems, the red-imported fire ant is reported to have queen 780 survival rates as low as 0.5% in field conditions (Tschinkel & King, 2017). However, this may

be offset by higher propagule pressure from human-assisted dispersal. Ultimately, a better
understanding of the trade-offs in ant dispersal and colonization will shed more light on the
effects of non-native ants on the SAR.

784

785 Conclusions

Our study presents exciting and unexpected results of analyses on ISARs for ants and offers the
most definitive evidence to date that ant communities on mainland islands tend to have steeper
relationships, with diversity increasing more rapidly over area, than do true islands. We also

- demonstrate that precipitation is a major factor influencing ISAR *z*-values, leading to a
 weakening of area effects on the scaling of biodiversity in localities with likely high primary
- weakening of area effects on the scaling of biodiversity in localities with likely high primary
 productivity and/or habitat diversity. Further research is needed to broaden our understanding of
- the ISAR in ants. Based on a review of ~ 60 years of literature we have identified major gaps in
- 793 ISAR knowledge as it pertains to ants, highlighted priority areas for future research and
- suggested approaches to fill these gaps of knowledge. Immediate areas of potential improvement
- 795 lie in further developing testable models based on habitat diversity hypotheses and further efforts
- into better understanding the trade-offs in dispersal and colonization in ants.

797 While our study focuses on ants, the implications and suggestions from this study extend 798 well beyond a single taxon. This is especially important at a time where human impacts on 799 biodiversity patterns from local to global scales show conflicting trends in the scientific literature 800 (Primack et al., 2018). The remaining pockets of biodiversity in minimally disturbed systems are rapidly disappearing. Therefore, now more than ever, there is a need to further understand the 801 scaling properties of biodiversity. Research conducted on taxonomically well-resolved and 802 803 globally distributed organisms, like ants, can provide this much needed and time-sensitive 804 information.

805

806 Data Availability Statement

B07 Datasets 1 and 2 which were used for all formal analyses from this study are available in the
supplementary materials (Appendix 2 and 3).

810 **References**

- Andersen, A. N., Kohout, R. J., & Trainor, C. R. (2013). Biogeography of timor and surrounding
 wallacean islands: Endemism in ants of the genus polyrhachis fr. Smith. *Diversity*, 5(1),
 139–148. https://doi.org/10.3390/d5010139
- Arrhenius, O. (1920a). Öcologisher Studien in den Stockholmer Scären. University of Stolkholm,
 Stockholm.
- 816 Arrhenius, O. (1920b). Yta och arter. I. Svensk Botansik Tidsskrift, 14, 327–329.
- 817 Arrhenius, O. (1921). Species and area. *Journal of Ecology*, 9, 95–99.
- Badano, E. I., Regidor, H. A., Núñez, H. A., Acosta, R., & Gianoli, E. (2005). Species richness
 and structure of ant communities in a dynamic archipelago: Effects of island area and age.
- *Journal of Biogeography*, *32*(2), 221–227. https://doi.org/10.1111/j.1365 2699.2004.01174.x
- Baroni-Urbani, C. (1968). Studi sulla mirmecofauna d'Italia. IV. La fauna mirmecologica delle
 isole Maltesi ed il suo significato ecologico e biogeografico. *Ann. Mus. Civ. Stor. Nat. "Giacomo Doria," 77*, 408–559.
- Baroni-Urbani, C. (1971). Studien zur Ameisenfauna Italiens XI. Die Ameisen des Toskanischen
 Archipels. Betrachtungen zur Herkunft der Inselfaunen. *Rev. Suisse Zool.*, 7(78), 1037–

- 827 1067.
- Baroni-Urbani, C. (1974). Studi sulla mirmecofauna d'Italia. XII. Le Isole Pontine. *Fragm. Entomol.*, 9, 225–252.
- Baroni-Urbani, C. (1976). Le formiche dell'arcipelago della Galita (Tunisia). *Redia*, 59, 207–
 223.
- Barton, K. (2019). *MuMIn: Multi-Model Inference. R apakcage version 1.43.6.*
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models
 Using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- 835 https://doi.org/10.18637/jss.v067.i01.
- Bolker, B., & R Development Core Team. (2020). *bbmle: Tools for General Maximum Likelihood Estimation*. https://cran.r-project.org/package=bbmle
- Bolton, B. (2020). *An online catalog of the ants of the world*. https://antcat.org
- Boomsma, J. J., Mabelis, A. A., Verbeek, M. G. M., & Los, E. C. (1987). Insular Biogeography
 and Distribution Ecology of Ants on the Frisian Islands. *Journal of Biogeography*, 14(1),

841 21. https://doi.org/10.2307/2844784

- Borowiec, L., & Salata, S. (2018a). Notes on ants (Hymenoptera: Formicidae) of Samos Island,
 Greece. *Annals of the Upper Silesian Museum, Entomology*, 27(003), 1–13.
- Borowiec, L., & Salata, S. (2018b). Notes on ants (Hymenoptera: Formicidae) of Zakynthos
 Island, Greece. *Annals of the Upper Silesian Museum, Entomology*, 27(004), 1–13.
- Boulton, A. M., & Ward, P. S. (2002). Ants. In T. J. Case, M. L. Cody, & E. Exequiel (Eds.), *Island Biogeography of the Sea of Cortez* (pp. 112–126). Oxford University Press.
- Burns, K. C., Paul Mchardy, R., & Pledger, S. (2009). The small-island effect: Fact or artefact?
 Ecography, 32(2), 269–276. https://doi.org/10.1111/j.1600-0587.2008.05565.x
- Chase, J. M., Gooriah, L., May, F., Ryberg, W. A., Schuler, M. S., Craven, D., & Knight, T. M.
 (2019). A framework for disentangling ecological mechanisms underlying the island
 species-area relationship. *Frontiers of Biogeography*, 11(1), 0–17.
 https://doi.org/10.21425/F5FBG40844
- Choi, B. M., Ogata, K., & Terayama, M. (1993). Comparative studies of ant faunas of Korea and
 Japan I. Faunal comparison among islands of Southern Korea and Northern Kyushu, Japan. *Bulletin of the Biogeographic Society of Japan*, 48(1), 37–49.
- Clark, A. T., Rykken, J. J., & Farrell, B. D. (2011). The effects of biogeography on ant diversity
 and activity on the Boston Harbor islands, Massachusetts, U.S.A. *PLoS ONE*, 6(11).
 https://doi.org/10.1371/journal.pone.0028045
- Cole, B. J. (1983). Assembly of mangrove ant communities: Patterns of geographical distribution. *Journal of Animal Ecology*, *52*(2), 339–347.
- Collingwood, C. (1993). A comparative study of the ant fauna of five Greek islands. *Biologia Gallo-Hellenica*, 20, 191–197.
- Connor, E. F., & McCoy, E. D. (1979). The statistics and biology of the species-area
 relationship. *The American Naturalist*, *113*(6), 791–833.
- Cuissi, R. G., Lasmar, C. J., Moretti, T. S., Schmidt, F. A., Fernandes, W. D., Falleiros, A. B.,
 Schoereder, J. H., & Ribas, C. R. (2015). Ant community in natural fragments of the
 Brazilian wetland: species–area relation and isolation. *Journal of Insect Conservation*, *19*(3), 531–537. https://doi.org/10.1007/s10841-015-9774-5
- Bro Dean, W. R. J., & Bond, W. J. (1990). Evidence for rapid faunal changes on islands in a manmade lake. *Oecologia*, 83(3), 388–391. https://doi.org/10.1007/BF00317565
- 872 Drakare, S., Lennon, J. J., & Hillebrand, H. (2006). The imprint of the geographical,

- evolutionary and ecological context on species-area relationships. *Ecology Letters*, 9(2),
 215–227. https://doi.org/10.1111/j.1461-0248.2005.00848.x
- Economo, E. P., Janda, M., Guénard, B., & Sarnat, E. (2017). Assembling a species–area curve
 through colonization, speciation and human-mediated introduction. *Journal of*
- 877 *Biogeography*, 44(5), 1088–1097. https://doi.org/10.1111/jbi.12884
- Economo, E. P., & Sarnat, E. M. (2012). Revisiting the ants of melanesia and the taxon cycle:
 Historical and human-mediated invasions of a tropical archipelago. *American Naturalist*, *180*(1), 1–16. https://doi.org/10.1086/665996
- Espadaler, X., & Bernal, V. (2003). Exotic ants in the Canary Islands (Hymenoptera,
 Formicidae). *Vieraea: Folia Scientarum Biologicarum Canariensium*, 31(31), 1–8.
- Ewers, R. M., Didham, R. K., Pearse, W. D., Lefebvre, V., Rosa, I. M. D., Carreiras, J. M. B.,
 Lucas, R. M., & Reuman, D. C. (2013). Using landscape history to predict biodiversity
 patterns in fragmented landscapes. *Ecology Letters*, *16*(10), 1221–1233.
 https://doi.org/10.1111/ele.12160
- Fattorini, S., Borges, P. A. V., Dapporto, L., & Strona, G. (2017). What can the parameters of the
 species–area relationship (SAR) tell us? Insights from Mediterranean islands. *Journal of Biogeography*, 44(5), 1018–1028. https://doi.org/10.1111/jbi.12874
- Fick, S. E., & Hijmans, R. J. (2017). New 1-km spatial resolution climate surfaces for global
 land areas. *International Journal of Climatology*, *37*(12), 4302–4315.
- Flantua, S. G. A., Payne, D., Borregaard, M. K., Beierkuhnlein, C., Steinbauer, M. J., Dullinger,
 S., Essl, F., Irl, S. D. H., Kienle, D., Kreft, H., Lenzner, B., Norder, S. J., Rijsdijk, K. F.,
 Rumpf, S. B., Weigelt, P., & Field, R. (2020). Snapshot isolation and isolation history
 challenge the analogy between mountains and islands used to understand endemism. *Global Ecology and Biogeography*, 29(10), 1651–1673. https://doi.org/10.1111/geb.13155
- Forster, J. R. (1778). *Observations made during a voyage round the world*. Printed for G.
 Robinson, London.
- Fournier, A., Penone, C., Pennino, M. G., & Courchamp, F. (2019). Predicting future invaders
 and future invasions. *Proceedings of the National Academy of Sciences of the United States*of America, 116(16), 7905–7910. https://doi.org/10.1073/pnas.1803456116
- Franzen, M., Schweiger, O., & Betzholtz, P.-E. (2012). Species-area relationships are controlled
 by species traits. *PloS One*, 7(5), e37359. https://doi.org/10.1371/Citation
- Gibb, H., Stoklosa, J., Warton, D. I., Brown, A. M., Andrew, N. R., & Cunningham, S. A.
 (2015). Does morphology predict trophic position and habitat use of ant species and assemblages? *Oecologia*, 177(2), 519–531. https://doi.org/10.1007/s00442-014-3101-9
- 907 Goldstein, E. L. (1975). Island biogeography of ants. *Evolution*, 29(4), 750–762.
- Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., Lovejoy, T.
 E., Sexton, J. O., Austin, M. P., Collins, C. D., Cook, W. M., Damschen, E. I., Ewers, R.
- 910 M., Foster, B. L., Jenkins, C. N., King, A. J., Laurance, W. F., Levey, D. J., Margules, C.
- R., ... Townshend, J. R. (2015). Habitat fragmentation and its lasting impact on Earth's
 ecosystems. *Science Advances*, 1(2), e1500052–e1500052.
- 913 https://doi.org/10.1126/sciadv.1500052
- Hakala, S. M., Seppa, P., & Helantera, H. (2019). Evolution of dispersal in ants (Hymenoptera:
 Formicidae): a review on the dispersal strategies of sessile superogranisms. *Myrmecological News*, 29, 35–55. https://doi.org/10.25849/myrmecol.news
- Halley, J. M., Sgardeli, V., & Triantis, K. A. (2014). Extinction debt and the species-area
 relationship: A neutral perspective. *Global Ecology and Biogeography*, 23(1), 113–123.

- 919 https://doi.org/10.1111/geb.12098
- 920 Hölldobler, B., & Wilson, E. O. (1990). The Ants. Harvard University Press.
- Holway, D. A., Lach, L., Suarez, A. V., Tsutsui, N. D., & Case, T. J. (2002). The causes and consequences of ant invasions. *Annual Review of Ecology and Systematics*, 33(Redford
- 923 1987), 181–233. https://doi.org/10.1146/annurev.ecolsys.33.010802.150444
- Jenkins, D. G., & Quintana-Ascencio, P. F. (2020). A solution to minimum sample size for regressions. *PLoS ONE*, *15*(2), 1–15. https://doi.org/10.1371/journal.pone.0229345
- Kalmar, A., & Currie, D. J. (2007). A unified model of avian species richness on islands and
 continents. *Ecology*, 88(5), 1309–1321. https://doi.org/10.1890/06-1368
- Keil, P., & Chase, J. M. (2019). Global patterns and drivers of tree diversity integrated across a continuum of spatial grains. *Nature Ecology and Evolution*, 3(3), 390–399.
 https://doi.org/10.1038/s41559-019-0799-0
- 931 Kylin, H. (1923). Växtsociologiska randanmärkningar. *Botaniska Notiser*, 1161–1234.
- 932 Lach, L., Parr, C., & Abbott, K. (2010). Ant Ecology. Oxford University Press.
- Bard Biology Ilustrated by the Land Birds of Jamaica. University of
 California Press.
- Leal, I. R., Filgueiras, B. K. C., Gomes, J. P., Iannuzzi, L., & Andersen, A. N. (2012). Effects of
 habitat fragmentation on ant richness and functional composition in Brazilian Atlantic
 forest. *Biodiversity and Conservation*, 21(7), 1687–1701. https://doi.org/10.1007/s10531012-0271-9
- Begakis, A. (2011). Annotated list of the ants (Hymenoptera, Formicidae) of Greece. *Hellenic Zoological Archives*, 7, 1–55.
- 941 Lomolino, M. V. (2000). Ecology's most general, yet protean pattern: The species-area
 942 relationship. *Journal of Biogeography*, 27(1), 17–26. https://doi.org/10.1046/j.1365943 2699.2000.00377.x
- Losos, J. B., & Schluter, D. (2000). Analysis of an evolutionary species-area relationship.
 Nature, 408(6814), 847–850. https://doi.org/10.1038/35048558
- 946 Lüdecke, D. (2018). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models.
 947 *Journal of Open Source Software*, 3(26), 772.
- MacArthur, R. H., & Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton
 University Press.
- Martin, T. E. (1981). Species-area slopes and coefficients: A caution on their interpretation. *The American Naturalist*, 118(6), 823–837.
- Matthews, T. J., Guilhaumon, F., Triantis, K. A., Borregaard, M. K., & Whittaker, R. J. (2016).
 On the form of species–area relationships in habitat islands and true islands. *Global Ecology and Biogeography*, 25(7), 847–858. https://doi.org/10.1111/geb.12269
- Matthews, T. J., & Rigal, F. (2021). Thresholds and the species–area relationship: a set of
 functions for fitting, evaluating and plotting a range of commonly used piecewise models in
 R. *Frontiers of Biogeography*, 0(0), 0–9. https://doi.org/10.21425/f5fbg49404
- Matthews, T. J., Steinbauer, M. J., Tzirkalli, E., Triantis, K. A., & Whittaker, R. J. (2014).
 Thresholds and the species-area relationship: A synthetic analysis of habitat island datasets. *Journal of Biogeography*, 41(5), 1018–1028. https://doi.org/10.1111/jbi.12286
- Matthews, T. J., Triantis, K. A., & Whittaker, R. J. (Eds.). (2020). *The Species–Area Relationship: Theory and Application*. Cambridge University Press.
- Matthews, T. J., Triantis, K. A., Whittaker, R. J., & Guilhaumon, F. (2019). sars: an R package
 for fitting, evaluating and comparing species–area relationship models. *Ecography*, 42(8),

- 965 1446–1455. https://doi.org/10.1111/ecog.04271
- 966 Menozzi, C. (1936). Nuovi contributi alla conoscenza della fauna delle Isole italiane dell'Egeo.
 967 VI. Hymenoptera Formicidae. *Bollettino Del Laboratorio Di Zoologia Generale e Agraria* 968 *Della Reale Scuola Superiore d'Agricoltura. Portici, 29*, 261–311.
- Mittelbach, G. G., & McGill, B. J. (2019). Patterns of biological diversity. In *Community Ecology* (2nd ed., pp. 11–36). Oxford University Press.
- Morrison, L. W. (1997). Polynesian ant (Hymenoptera: Formicidae) species richness and
 distribution: A regional survey. *Acta Oecologica*, 18(6), 685–695.
- 973 https://doi.org/10.1016/S1146-609X(97)80051-3
- 974 Morrison, Lloyd W. (2003). A review of Bahamian ant (Hymenoptera: Formicidae)
 975 biogeography. *Journal of Biogeography*, 25, 561–571. https://doi.org/10.1016/B978-0-323976 60984-5.00062-7
- 977 Morrison, Lloyd W. (2016). The ecology of ants (Hymenoptera: Formicidae) on islands.
 978 *Myrmecological News*, 23(September), 1–14.
- 979 Muggeo, V. M. (2016). Segmented mixed models with random changepoints in R. R Working
 980 Paper.
- Muggeo, V. M., Atkins, D. C., Gallop, R. J., & Dimidjian, S. (2014). Segmented mixed models
 with random changepoints: a maximum likelihood approach with application to treatment
 for depression study. *Statistical Modelling*, *14*, 293–313.
- Mühlenberg, A. M., Leipold, D., Mader, H. J., & Steinhauer, B. (2016). Island ecology of
 arthropods. II. Niches and relative abundances of Seychelles Ants (Formicidae) in different
 habitats. *Oecologia*, 29(2), 135–144.
- Nilsson, S. G., Bengtsson, J. A. N., & Ast, S. (1988). Habitat diversity or area per se? Species
 richness of woody plants, Carabid beetles, and land snails on islands. *Journal of Animal Ecology*, 57(2), 685–704.
- Poldi, B., Mei, M., & Rigato, F. (1995). Hymenoptera Formicidae. *Checklist Delle Specie Della Fauna Italiana*, 102, 1–10.
- Primack, R. B., Miller-Rushing, A. J., Corlett, R. T., Devictor, V., Johns, D. M., Loyola, R.,
 Maas, B., Pakeman, R. J., & Pejchar, L. (2018). Biodiversity gains? The debate on changes
 in local- vs global-scale species richness. *Biological Conservation*, 219, A1–A3.
 https://doi.org/10.1016/j.biocon.2017.12.023
- 996 Pyšek, P., Blackburn, T. M., García-Berthou, E., Perglová, I., & Rabitsch, W. (2017).
 997 Displacement and local extinction of native and endemic species. In *Impact of biological*998 *invasions on ecosystem services* (pp. 157–175). Springer.
- 999 R Development Core Team. (2018). *R: A language and environment for statistical computing*1000 (3.4.1).
- 1001 Ranta, E., Vepsalainen, K., As, S., Haila, Y., Pisarski, B., & Tiainen, J. (1983). Island
 1002 biogeography of ants (Hymenoptera, Formicidae) in four Fennoscandian archipelagoes.
 1003 Acta Entomologica Fennica, 42(August 2014), 64.
- Rica, C., Solomon, S. E., & Mikheyev, A. S. (2005). The ant (Hymenoptera : Formicidae) fauna of Cocos Island, Costa Rica. *Florida Entomologist*, 88(4), 415–423.
- Ricklefs, R. E., & Lovette, I. J. (1999). The roles of island area per se and habitat diversity in the
 species-area relationships of four Lesser Antillean faunal groups. *Journal of Animal Ecology*, 68(6), 1142–1160. https://doi.org/10.1046/j.1365-2656.1999.00358.x
- 1009 Rizali, A., Lohman, D. J., Buchori, D., Prasetyo, L. B., Triwidodo, H., Bos, M. M., Yamane, S.,
- 1010 & Schulze, C. H. (2010). Ant communities on small tropical islands: Effects of island size

- 1011 and isolation are obscured by habitat disturbance and "tramp" ant species. Journal of 1012 *Biogeography*, 37(2), 229–236. https://doi.org/10.1111/j.1365-2699.2009.02194.x 1013 Rosenzweig, M. L. (1995). Species Diversity in Space and Time. Cambridge University Press. 1014 Sanders, N. J. (2002). Elevational gradients in ant species richness: Area, geometry, and Rapoport's rule. Ecography, 25(1), 25-32. https://doi.org/10.1034/j.1600-1015 1016 0587.2002.250104.x 1017 Sarnat, E., Blanchard, B., Guénard, B., Fasi, J., & Economo, E. (2013). Checklist of the ants 1018 (Hymenoptera, Formicidae) of the Solomon Islands and a new survey of Makira Island. ZooKeys, 257, 47-88. https://doi.org/10.3897/zookeys.257.4156 1019 1020 Sarnat, E. M., & Economo, E. P. (2012). Ants of Fiji. University of California Press. Scheiner, S. M. (2003). Six types of species-area curves. *Global Ecology and Biogeography*. 1021 12(6), 441–447. https://doi.org/10.1046/j.1466-822X.2003.00061.x 1022 1023 Schoereder, J. H., Galbiati, C., Ribas, C. R., Sobrinho, T. G., Sperber, C. F., DeSouza, O., & Lopes-Andrade, C. (2004). Should we use proportional sampling for species-area studies? 1024 1025 Journal of Biogeography, 31(8), 1219-1226. https://doi.org/10.1111/j.1365-2699.2004.01113.x 1026 Schultz, T. R. (2000). In search of ant ancestors. Proceedings of the National Academy of 1027 1028 Sciences of the United States of America, 97(26), 14028–14029. 1029 https://doi.org/10.1073/pnas.011513798 Shmida, A., & Wilson, M. V. (1985). Biological determinants of species diversity. Journal of 1030 *Biogeography*, *12*(1), 1–20. 1031 Sosiak, C. E., & Barden, P. (2020). Multidimensional trait morphology predicts ecology across 1032 1033 ant lineages. Functional Ecology, October 2020, 139-152. https://doi.org/10.1111/1365-1034 2435.13697 Storch, D. (2016). The theory of the nested species-area relationship: geometric foundations of 1035 biodiversity scaling. Journal of Vegetation Science, 27(5), 880-891. 1036 https://doi.org/10.1111/jvs.12428 1037 1038 Storch, D., Evans, K. L., & Gaston, K. J. (2005). The species-area-energy relationship. Ecology 1039 Letters, 8(5), 487–492. https://doi.org/10.1111/j.1461-0248.2005.00740.x 1040 Suarez, A. V., Bolger, D. T., & Case, T. J. (1998). Effects of fragmentation and invasion on native ant communities in coastal southern California. Ecology, 79(6), 2041–2056. 1041 1042 https://doi.org/10.1890/0012-9658(1998)079[2041:EOFAIO]2.0.CO;2 Sugiura, S. (2010). Species interactions-area relationships: Biological invasions and network 1043 structure in relation to island area. Proceedings of the Royal Society B: Biological Sciences, 1044 1045 277(1689), 1807–1815. https://doi.org/10.1098/rspb.2009.2086 Terayama, M. (1992). Structure of ant communities in East Asia I. Regional differences and 1046 1047 species richness. Bulletin of the Biogeographic Society of Japan, 41, 1–31. 1048 Tjørve, E. (2003). Shapes and functions of species-area curves: a review of possible models. Journal of Biogeography, 30, 827–835. 1049 Tjørve, Even. (2003). Shapes and functions of species-area curves: A review of possible models. 1050 1051 Journal of Biogeography, 30(6), 827-835. https://doi.org/10.1046/j.1365-1052 2699.2003.00877.x
- Torres, J. A., & Snelling, R. R. (1997). Biogeography of Puerto Rican ants: A Non-equilibrium case? *Biodiversity and Conservation*, 6(8), 1103–1121.
 https://doi.org/10.1023/A:1018332117719
- 1056 Trainor, C. R., & Andersen, A. N. (2010). The ant fauna of Timor and neighbouring islands:

- Potential bridges between the disjunct faunas of South East Asia and Australia. *Australian Journal of Zoology*, 58(3), 133–144. https://doi.org/10.1071/ZO09113
- Triantis, K A, Mylonas, M., Lika, K., Natural, K. V., & Museum, H. (2003). A model for the
 species area habitat relationship. *Journal of Biogeography*, *30*, 19–27.
- Triantis, Kostas A, Guilhaumon, F., & Whittaker, R. J. (2012). The island species-area
 relationship: Biology and statistics. *Journal of Biogeography*, *39*(2), 215–231.
 https://doi.org/10.1111/j.1365-2699.2011.02652.x
- Tschinkel, W. R., & King, J. R. (2017). Ant community and habitat limit colony establishment
 by the fire ant, Solenopsis invicta. *Functional Ecology*, *31*(4), 955–964.
 https://doi.org/10.1111/1365-2435.12794
- 1067 Vasconcelos, H. L., Vilhena, J. M. S., Magnusson, W. E., & Albernaz, A. L. K. M. (2006).
 1068 Long-term effects of forest fragmentation on Amazonian ant communities. *Journal of Biogeography*, *33*(8), 1348–1356. https://doi.org/10.1111/j.1365-2699.2006.01516.x
- Wallace, A. R. (1914). The World of life: a manifestation of creative power, directive mind and
 ultimate purposeNo Title. Chapman and Hall.
- Weigelt, P., Jetz, W., & Kreft, H. (2013). Bioclimatic and physical characterization of the
 world's islands. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), 15307–15312. https://doi.org/10.1073/pnas.1306309110
- 1075 Wetterer, J. K. (2002). Ants of Tonga. *Pacific Science*, 56(2), 125–135.
 1076 https://doi.org/10.1353/psc.2002.0019
- Wetterer, J. K. (2006). Ants (Hymenoptera: Formicidae) of Niue, Polynesia. *Pacific Science*, 60, 413–416.
- Wetterer, J. K., Espadaler, X., Wetterer, A. L., Aguin-Pombo, D., & Franquinho-Aguiar, A. M.
 (2007). Ants (Hymenoptera: Formicidae) of the Madeiran archipelago. *Sociobiology*, 49(3), 265–297.
- Wetterer, J. K., Espadaler, X., Wetterer, A. L., & Cabral, S. G. M. (2004). Native and exotic ants
 of the Azores (Hymenoptera: Formicidae). *Sociobiology*, 44(1), 1–20.
- Wetterer, J. K., & O'Hara, B. C. (2002). Ants (Hymenoptera: Formicidae) of the Dry Tortugas,
 the outermost Florida Keys. *Florida Entomologist*, *85*(2), 303–307.
- 1086 https://doi.org/10.1653/0015-4040(2002)085[0303:ahfotd]2.0.co;2
- Whitehead, D. R., & Jones, C. E. (1969). Small Islands and the Equilibrium Theory of Insular
 Biogeography. *Evolution*, 23(1), 171. https://doi.org/10.2307/2406492
- Whittaker, R. J., Araújo, M. B., Jepson, P., Ladle, R. J., Watson, J. E. M., & Willis, K. J. (2005).
 Conservation biogeography: Assessment and prospect. *Diversity and Distributions*, 11(1),
 3–23. https://doi.org/10.1111/j.1366-9516.2005.00143.x
- Whittaker, R. J., & Fernandez-Palacios, J. (2007). *Island Biogeography: Ecology, Evolution, and Conservation*. Oxford University Press.
- Whittaker, R. J., Triantis, K. A., & Ladle, R. J. (2008). A general dynamic theory of oceanic
 island biogeography. *Journal of Biogeography*, *35*(6), 977–994.
 https://doi.org/10.1111/j.1365-2699.2008.01892.x
- 1097 Wickham, H. (2009). ggplot2: Elegant graphics for data analysis. Springer-Verlag.
- 1098 Williams, C. B. (1964). Patterns in the balance of nature. Academic Press.
- Williams, M. R., Lamont, B. B., & Henstridge, J. D. (2009). Species–area functions revisited.
 Journal of Biogeography, *36*, 1994–2004.
- Wilson, E. (1961). The nature of the taxon cycle in the Melanesian ant fauna. *The American Naturalist*, 95(882), 169–193.

1103	Wilson, E. O. (2005). Early ant plagues in the New World. Nature, 433(7021), E1.
1104	https://doi.org/10.1038/nature03256
1105	Wilson, Edward O. (1988). The Biogeography of the West Indian Ants
1106	(Hymenoptera:Formicidae). In Zoogeography of Carribean Insects (pp. 214–230).
1107	Wilson, Edward O., & Taylor, R. W. (1967). An estimate of the potential evolutionary increase
1108	in species density in the Polynesian ant fauna. Evolution, 21(1), 1.
1109	https://doi.org/10.2307/2406735
1110	Woinarski, J. C, Reichel, H., & Andersen, A. N. (1998). The distribution of ants on the Wessel
1111	and English Company Islands, in the seasonal tropics of Australia's Northern Territory.
1112	Australian Journal of Zoology, 46, 557–578.
1113	Yan, H., Liang, C., Li, Z., Liu, Z., Miao, B., He, C., & Sheng, L. (2015). Impact of precipitation
1114	patterns on biomass and species richness of annuals in a dry steppe. <i>PLoS ONE</i> , 10(4), 1–
1115	14. https://doi.org/10.1371/journal.pone.0125300
1116	Zhao, Y., Dunn, R. R., Zhou, H., Si, X., & Ding, P. (2020). Island area, not isolation, drives
1117	taxonomic, phylogenetic and functional diversity of ants on land-bridge islands. Journal of
1118	Biogeography, August 2019, 1–11. https://doi.org/10.1111/jbi.13860
1119	Zhu, H., Wang, D., Wang, L., Fang, J., Sun, W., & Ren, B. (2014). Effects of altered
1120	precipitation on insect community composition and structure in a meadow steppe.
1121	<i>Ecological Entomology</i> , 39(4), 453–461. https://doi.org/10.1111/een.12120
1122	
1123	
1124	
1125	Biosketches
1126	
1127	Leo Ohyama is a PhD candidate and Biodiversity Institute Fellow at the University of Florida
1128	with main interests in macroecology, biogeography, and ant ecology. He has previously worked
1129	with ants at local community scales through running large-scale field studies
1130	The area of the second second and age realized and a second from second s
1131	Editor: Spyros Sfenthourakis
1132	
1133	
1134	
1135	
1136	
1137	
1138	
1139	
1140	
1140	
11 <u>/</u> 7	
11/2	
11//	
11/15	
11/6	
1140	



1120	
1157	Figure 1. (A) A global map showcasing the variety of datasets used in modeling ant species
1158	richness as a function of area across all biogeographical realms. Size of each circle
1159	represents species richness from each individual study/dataset. Biogeographical realms are
1160	represented by different colors with the same color scheme applied to (B) and (C). Both (B)
1161	and (C) show the fixed and random effects of a linear mixed-effect model for only true
1162	islands. (B) Fixed effect (solid black line) of area on species richness from the linear mixed-
1163	effect model predicting log species richness as a function of log area with a random slope
1164	for biogeographical realm. Grey shading represents the 95% confidence intervals of the
1165	fixed effect. Points represent the data color coded by biogeographic realm. (C) Random
1166	effects (colored-lines) from the same linear-mixed effect model from (B). Points represent
1167	the data color coded by biogeographic realm. Natural logarithms are used and the area unit
1168	is km ² . Map uses Mollweide projection.



Figure 2. Boxplots showing the distribution of slope values (Z) across studies (n = 41) for ants across all biogeographic realms. (A) Boxplots of z by biogeographic realms. (B) Boxplots of z by insular or mainland studies. (C) Boxplots of z-values by the types of islands in the study.



Figure 3. Predictions of the most plausible model of ant z-values across all biogeographical
realms when accounting for the covariate ISAR type (insular or mainland studies). The yaxis represents slope values (Z) and the x-axis represents mean annual precipitation
(scaled). Lines represent model predictions. Shaded regions represent 95% confidence
intervals. Points represent actual data. Green points and shading indicate mainland studies
while blue indicates insular studies.





Figure 4. Plot showing both R^2 values and slope values (Z) of ant studies across all

biogeographical realms (both mainland and insular) used in the analysis. The y-axis
 represents precipitation (in mm). The x-axis represents both R² values and z-values from 0
 to 1 going in both left and right directions. Colors represent the biogeographic realm of the
 study.



Figure 5. Plot showing the best threshold model for ants across biogeographical realms
(continuous one-threshold model; black lines) fitted to dataset 2 for true islands (n = 320) in
(A) log-log and (B) semi-log space. Points are coloured by biogeographic realm, and the
colours are consistent across plots (i.e. the legend in (b) applies to both plots). Natural

logarithms are used and the area unit is km².

1210 Tables 1211

1212 Table 1. All variables and a description of each, collected from studies to build the datasets used1213 in the analyses.

Variables collected from studies	Description
Slope value (z)	Slope coefficient of the log-log species area relationship
Standard error of slope value	Error attributed to the slope coefficient estimate
Coefficient of correlation (R^2)	Amount of variation in richness explained by area
Biogeographic realm of study	Afrotropic, Australasia, Indomalaya, Nearctic,
	Neotropic, Oceania, Palearctic
Island types used in the study	Habitat-patches, Continental islands, Oceanic islands
Number of sampling methods	The number of unique sampling methods used in
	collecting ants for the study

Coordinates of study	Latitude and longitude of study. If study was done across large spatial areas, a midpoint coordinate was calculated.
Island area	Areas for islands/fragments used in the studies if disclosed
Species richness per island	Species richness values for each island/fragment if disclosed
Non-native ant percentage	Calculated as the species richness of non-native ants divided by total species richness for each dataset

Table 2. A summary table showing the number of datasets used in datatset 1, by biogeographic realm, species–area relationship type, island type, and citation. Second column indicates specific citation by number-code in parentheses. Note datasets focused on only single islands used in dataset 2 are cite at the bottom but not listed in the table.
Table 2. A summary table showing the number of datasets used in datatset 1, by biogeographic realm, species–area relationship type, island type, and citation. Second column indicates specific citation by number-code in parentheses. Note datasets focused on only single
Table 2. A summary table showing the number of datasets used in datatset 1, by biogeographic realm. species—area relationship type, island type, and citation. Second column indicates
Table 2 A summary table showing the number of datasets used in datatset 1 by biogeographic

Biogeographic Realm	and Citation	Species-area relationship type	# of datasets	Island type	# of datasets
		Incular	1	Continental:	0
Afrotropic	2 (1)	Insular	1	Mainland Islands:	1
		Mainland	1	Oceanic:	1
Australasia	4 (2)	Insular	4	Continental:	2

				Mainland Islands:	0	
		Mainland	0	Oceanic:	2	
			i	Continental:	1	
Indomalaya	2 (3)	Insular	2	Mainland Islands:	0	
		Mainland	0	Oceanic:	1	
		Ingular	2	Continental:	2	
Nearctic	6 (5)	Insular	Z	Mainland Islands:	4	
		Mainland	4	Oceanic:	0	
		Insular	5	Continental:	3	
Neotropic	10 (6)		3	Mainland Islands:	5	
		Mainland	5	Oceanic:	2	
Oceania	4 (7)	Insular	4	Continental:	0	
				Mainland Islands:	0	
		Mainland	0	Oceanic:	4	
		To see 1 or	10	Continental:	5	
Palearctic	13 (8)	Insular	12	Mainland Islands:	1	
		Mainland	1	Oceanic:	7	
		Total Insulant	20	Total Continental:	13	
Total:	41	1 otal Insular: 30		I otal <u>Mainland</u> I <u>Islands</u> :	11	
		Total Mainland:	11	Total Mixed:	0	
		i ovar i rannunut		Total Oceanic:	17	

Citations: 1. (Dean & Bond, 1990, Wetterer et al., *unpublished*), 2. (E. Sarnat et al., 2013; E. Wilson, 1961; Woinarski et al., 1998), 3. (Rizali et al., 2010; Trainor & Andersen, 2010), 4. (Clark et al., 2011; Goldstein, 1975), 5. (Clark et al., 2011; Goldstein, 1975; Sanders, 2002; Suarez et al., 1998), 6. (Badano et al., 2005; Boulton & Ward, 2002; Cole, 1983; Cuissi et al., 2015; Leal et al., 2012; Schoereder et al., 2004; Vasconcelos et al., 2006; Wetterer & O'Hara, 2002; Edward O. Wilson, 1988) 7. (L. W. Morrison, 1997; Wetterer, 2002; Edward O. Wilson & Taylor, 1967), 8. (Baroni-Urbani, 1968, 1971, 1974; Boomsma et al., 1987; Choi et al., 1993; Collingwood, 1993; Menozzi, 1936; Sugiura, 2010; Wetterer et al., 2004, 2007; Zhao et al., 2020)
Papers with single island records used in dataset 2: (Baroni-Urbani, 1976; Borowiec & Salata, 2018b, 2018a; Legakis, 2011; Mühlenberg et al., 2016; Poldi et al., 1995; Terayama, 1992;

Wetterer, 2006, Wetterer, *unpublished*)

1243 1244

1245Table 3. A table showing the competing models under two different covariates, island range

1246 (order of magnitude) (n = 35) and species-area relationship type (mainland or insular) (n = 42),

1247 assessing slope values (z) as a function of abiotic variables based on AIC_c (Akaike Information

1248 Criterion with correction for small sample sizes) rankings. Predictor variables for each model are

1249 shown along with each model's AIC_c score, the change in AIC_c for every lower ranked model,

1250	AIC _c weig	nts, and th	e adjusted R ² .
------	-----------------------	-------------	-----------------------------

Model with island order of magnitude range (OMR) covariate	ΔAICc	Weight (wi)	Pseudo-R ²
$z \sim \text{OMR} + \text{Precipitation}$	0.0	0.415	0.11
$z \sim OMR$	1.2	0.223	0.03
$z \sim \text{OMR} + \text{Temperature}$	2.5	0.122	0.05
$z \sim \text{OMR} * \text{Precipitation}$	2.6	0.115	0.11
$z \sim \text{OMR} + \text{Latitude}$	3.7	0.066	0.03
$z \sim \text{OMR} * \text{Temperature}$	4.9	0.036	0.05
Null Model	5.9	0.021	0
$z \sim \text{OMR} + \text{Biogeographic realm} + \text{Precipitation}$	11.1	0.0016	0.26
$z \sim \text{OMR} + \text{Biogeographic realm} + \text{Temperature}$	20.9	< 0.001	0.07
Model with ISAR type covariate	AAIC _c	Weight (wi)	Pseudo-R ²
$z \sim ISAR$ type + Precipitation	0.0	0.631	0.18
$z \sim ISAR$ type + Temperature	2.3	0.200	0.12
$z \sim ISAR$ type	4.0	0.085	0.07
$z \sim ISAR$ type + Latitude	4.1	0.083	0.10
$z \sim ISAR$ type + Biogeographic realm + Precipitation	14.3	< 0.001	0.12
Null Model	18.3	< 0.001	0
$z \sim ISAR$ type + Biogeographic realm + Temperature	19.4	< 0.001	0.12

Table 4. Threshold models comparison summary. Results are presented for the insular global dataset, for both the log–log and semi-log transformations. For each model, the AIC_c and R^2 values are provided, and for the threshold models the area value (km²) where the inflexion point is located (on a log scale) is provided (Th1). For each comparison, the intercept-only model is not included to save space as it was always the worst model. ContOne is the continuous one-threshold model, and ZslopeOne the left-horizontal one-threshold model. In log-log space the linear model is the power model, and in semi-log space it is the logarithmic model.

Model	AICc	R ²	Th1
Insular: log–log			
ContOne	867.31	0.36	4.895
ZslopeOne	882.15	0.32	1.695

	Linear	889.79	0.30	NA
	Insular: semi-log			
	ContOne	2751.14	0.54	5.095
	ZslopeOne	2753.08	0.54	4.995
	Linear	2887.31	0.29	NA
1274				
1275				
1276				
1277				
1278				
1279				
1280				
1281				
1282				