Mainland size variation informs predictive models of exceptional insular body size change in rodents

Paul A. P. Durst and V. Louise Roth

Department of Biology, Duke University, Durham, NC 27708-0338, USA

The tendency for island populations of mammalian taxa to diverge in body size from their mainland counterparts consistently in particular directions is both impressive for its regularity and, especially among rodents, troublesome for its exceptions. However, previous studies have largely ignored mainland body size variation, treating size differences of any magnitude as equally noteworthy. Here, we use distributions of mainland population body sizes to identify island populations as ‘extremely’ big or small, and we compare traits of extreme populations and their islands with those of island populations more typical in body size. We find that although insular rodents vary in the directions of body size change, ‘extreme’ populations tend towards gigantism. With classification tree methods, we develop a predictive model, which points to resource limitations as major drivers in the few cases of insular dwarfism. Highly successful in classifying our dataset, our model also successfully predicts change in untested cases.

1. Introduction

The ‘island rule’ for mammals, a pattern of divergence in body size between insular populations of mammals and their mainland counterparts, was once deemed ‘an extraordinary phenomenon which seems to have fewer exceptions than any other ecotypic rule in animals’ [1, p. 35]. Yet, as examples have proliferated—from anecdotal observations of insular mammals of unusual size [2,3], to tabulations and tallies of taxa showing insular gigantism or dwarfism [4], to regressions on body mass of the continuous variable ‘size ratio’ (average island/average mainland body mass of a species; [5])—so too have the exceptions. Shifts in the average body size of a population of insular mammals have been observed to differ in direction and in degree for different populations on different islands. And in recognition of this, the number of factors used to predict the changes has also grown. The simple scenario that Foster [4] described, in which size increase or decrease predominates within a given order of mammals, has been replaced with the description of a generally monotonic trend of decrease in size ratio with increasing body size across species [5]. This, in turn, has been refined and replaced by models of size change that apply differently in different cases and are contingent both on the phylogenetic affinity of the population and on an array of ecological, environmental, geographical and species-specific attributes of the island + population pair ([6–12]; for more thorough treatments of island rule literature and theory, see [7,13,14]).

Given the importance of body size in all aspects of an organism’s biology [15,16], evolutionary changes in size are expected to be multifactorial and subject to contingency. And size, as a composite trait that can be obtained readily for a large and broad sampling of cases, therefore serves as a proxy for suites of other traits on which selection acts more directly. While these can eventually be the subjects of more detailed and mechanistic studies, our approach here is intentionally broad and synthetic, on the premise that important clues to causal factors emerge when predictive models identify broader patterns. Up to this point, however, broad patterns have been elusive.

In this paper, we focus on rodents, the most species-rich order of mammals (with nearly half of all extant species [17]). Reports of genetic differentiation of
insular populations of rodents are common (e.g. [18–20]) and experiments have shown divergence in insular body size to have a genetic component [21], yet this order has presented the greatest obstacle to formulation of a general ‘rule’ for mammalian body size change on islands. A widespread but phylogenetically circumscribed group of mammals, rodents share a distinctive and well-defined set of morphological traits. Yet despite their morphological coherence, rodents vary widely in body size, habitat and diet. Extant insular populations of rodents exhibit higher variance in and a wider range of size ratios than any other order, and both the direction and the magnitude of size change can vary among different island populations within a single rodent species [8,9,11,13,22].

To develop a model for body size change on islands, we must first ask, what is the main variable of interest? Is it the direction of size change on an island—smaller or larger, regardless of degree? If the goal is to explain body size divergence of island populations from their mainland counterparts, size differences of a magnitude not expected on the mainland are clearly of interest. When the time of first colonization of an island and the genetic variance of the founding population are unknown, minor body size differences can reflect either conditions that do not differ from those on the mainland or insufficient time or variation for those influences to take effect. Setting a size-difference threshold informed by body size variation among mainland populations allows us to focus on cases in which the underlying signal is not obscured by such limitations.

Here, we examine body size divergence between conspecific island and mainland populations and address the following questions:

(i) for the purpose of discerning ‘rules’ that govern body size evolution in insular rodents, what constitutes a meaningful difference in body sizes between populations—that is, a size difference that exceeds that commonly found among populations on the mainland?

(ii) what factors are associated with such extreme—as opposed to more moderate—cases of insular body size change?

(iii) within the set of cases that have shown extreme change, what factors differentiate extremely small- and extremely large-bodied island populations?

To address question (i), we compare the distribution of island : mainland size ratios for insular rodents with the distribution of analogous ratios from the mainland and we develop criteria for labelling island populations as ‘extreme’. In doing this, we define a threshold for the magnitude of body size difference we will consider when we develop a predictive model, excluding, in the process, populations that are undergoing minor short-term fluctuations in body size or have had inadequate time (or genetic variance) for dramatic size differences to arise. For question (ii), we examine differences in the distributions of 19 organismal, climatic and geographical variables between moderate and extreme insular rodent populations. Finally, for question (iii), we use classification trees and random forest techniques to develop a predictive model describing factors associated with extremely large and extremely small body sizes in insular populations. We then test this model on a set of cases independent from the sample used to create it.

2. Material and methods

(a) Island data collection

Records of insular rodent populations were taken from the database of McClain et al. [11,23], which is an expansion of the database from Meiri et al. [8]. The size ratio for each population was calculated as the mean island body size (either mass or the cube of linear measurements when mass was unavailable [3,8]) divided by that same measure of body size from the closest mainland population (see [8] for description of original dataset). Additional information for each population included organismal traits (mass, dietary data and substrate preference), six climatic variables (temperature and precipitation, each averaged spatially across the island, their standard deviations (SD) and their temporal variation), four measures of island productivity (the spatial mean, maximum, minimum and standard deviation of net primary productivity, NPP) and six geographical variables (measures of island area, elevation and isolation). Although several previous studies examining the island rule have suggested that changes in predation, interspecific competition and population density are responsible for insular body size change (see electronic supplementary material, appendix S1), accurate accounts of island population densities and community composition are not available from primary literature for as broad a sampling of cases as we wished to include. Details on data collection can be found in the electronic supplementary material for McClain et al. [11]. As island area and elevation can both be indicators of island terrain heterogeneity [10], we combined measures of the two in a principal components analysis and used PC1 as a measure of island heterogeneity (electronic supplementary material, table S1). A total of 306 populations of insular rodents (67 species on 182 islands across the world) were included in the analyses.

(b) Mainland population collection

We obtained mainland rodent records from the databases at the Natural History Museum of Los Angeles County and the Smithsonian, US National Museum of Natural History [24]. We discarded all records that did not include body mass or locality information, and then grouped the records together by species and population. Populations were defined as more than one record of the same species obtained from exactly the same location as noted in the museum records. This process resulted in a total of 1076 rodent populations (mean = 5.4 records/population, median = 3 records/population) from across the world, including species from every family that is represented in the island database. We took the mean mass for each population and divided it by the mean mass for the species as defined in the database from Meiri et al. [8]. The size ratio for each population analogous to (albeit distinct from) the size ratios between island and mainland populations (each standardized, because sample sizes differed, to total 100% in a probability density function), we calculated as the mean island body size (either mass or the cube of linear measurements when mass was unavailable [3,8]) for a median of 3 records/population (n = 3—or n = 4—for a median of n = 6—leaves our interpretations and conclusions unchanged.)

(c) Defining ‘meaningful’ size change

To test for a difference between the size-ratio distributions of mainland and island populations (each standardized, because sample sizes differed, to total 100% in a probability density function), we used the Matching package in R [25,26] to perform a bootstrapped Kolmogorov–Smirnov (K–S) test on the two groups.
We then defined ‘normal’ size differences for the mainland as any size ratio that fell between the lowest and highest 2.5% of all mainland size ratios. Anything falling outside those boundaries (in the tails) was defined as ‘extreme’ size difference. The same numerical size-ratio cut-offs (values of the ratios, as opposed to percentage) were then applied to the island populations to identify insular body size ratios that were ‘normal’ and ‘extreme’ with respect to variation on the mainland. The extreme island populations were further categorized as either extremely big (‘Big’) or extremely small (‘Small’).

(d) Comparing conditions associated with ‘normal’ versus ‘extreme’ island populations
To examine the factors influencing extreme insular body size change, we plotted as probability density functions the frequency distributions for the ‘normal’ and ‘extreme’ cases with respect to each of the geographical, climatic and productivity variables, and then compared these distributions using bootstrapped K–S tests. The pairwise comparisons included ‘normal’ versus Big, ‘normal’ versus Small, ‘normal’ versus Big + Small combined, and Big versus Small. In all instances, the threshold for significance was taken as $p < 0.05$. No correction was made for multiple comparisons, as we viewed these comparisons as descriptive and exploratory.

(e) Classification tree analyses
In addition to comparing frequency distributions for each variable, we used classification trees to create a predictive model for the directionality of extreme insular body size change (Big or Small). CART (classification and regression tree, jointly called ‘decision tree’) methods identify a succession of predictor variables to group observations (figure 1 a) Size ratios for mainland (black) and island (purple) populations; vertical lines delimit tails of the distributions showing extreme size reduction (left) or increase (right). (b–i) Various attributes whose distributions differ significantly (K–S test, $p < 0.05$) between ‘normal’ (solid black), Small (dashed blue) and/or Big (dotted red) island populations. (Online version in colour.)

to obtain the optimal tree (cross-validated error rate within 1 s.e. of minimum error tree; as discussed in [9]). To test the stability of our classification tree, we also built a random forest model consisting of 5000 trees built from bootstrapped subsets of the data using the randomForest package in R [28]. A relatively stable classification tree will have a predictive accuracy similar to that of the related random forest model.

Earlier studies that applied CART methods to identify variables associated with insular body size change [9,10] were largely descriptive and inferential, constructing the decision trees and examining the chosen variables for their biological relevance. However, CART methods, especially classification trees, are widely used in other disciplines not only to describe the data used to build the tree, but to build a model to predict values of the response variable for samples independent of those used to construct the tree [29]. Recent work by Lomolino et al. [12] provided size ratios and locality data for almost 100 insular rodents populations for which we had island information (obtained earlier for purposes of studying insular mammals other than rodents; [11]) but had not included in our ‘model-training’ data-set. Of those populations, 39 of them qualified as extreme according to our definition. We ran these 39 samples through our classification tree and random forest to assess the efficacy of our models as predictive tools.

3. Results
(a) Defining ‘meaningful’ size change
According to the K–S test, size-ratio distributions for island and mainland rodent populations differed significantly ($p < 0.001$; figure 1a). On the mainland, the lower 2.5% quantile for size ratios was 0.80, while the upper 97.5% quantile was 1.09. We took these as the boundaries for ‘normal’ body size, and the 5% of mainland samples falling outside these boundaries were considered ‘extreme’ mainland populations (figure 1a). Applying the same boundaries to the island samples, 54% of the samples were classified as ‘extreme’ insular populations. Of these extreme populations, 78% (deemed ‘Big’) had increased in size, while 22% (deemed ‘Small’) decreased in size. Although there were no consistent patterns within species or more inclusive clades (see electronic supplementary material, appendix S3) in the

Figure 1. Density plots comparing distributions of numbers of island populations (y-axis) for various characteristics (x-axis). (a) Size ratios for mainland (black) and island (purple) populations; vertical lines delimit tails of the distributions showing extreme size reduction (left) or increase (right). (b–i) Various attributes whose distributions differ significantly (K–S test, $p < 0.05$) between ‘normal’ (solid black), Small (dashed blue) and/or Big (dotted red) island populations. (Online version in colour.)
tendency to produce insular populations of extreme body size, for species that produced extreme populations on multiple islands the tendency was strong for those extreme changes to occur in the same direction (Big or Small). The lack of a consistent tendency for changes within species to be extreme or not could reflect the genetic or time constraints mentioned above; the intraspecific consistency in direction of extreme changes echoes earlier findings from analyses of insular body size change of any magnitude [9]. There was also some inclination for more than one species of rodent on a single island to be either normal or extreme, but the tendency was slight and the extreme populations had not necessarily changed in the same direction. The extent to which size changes documented in our dataset are associated with particular islands or particular species is an additional indication that characteristics of species and of islands are relevant for predicting when and in which direction extreme body size change should occur.

(b) Comparing normal and extreme islands
The distributions of seven variables were found to differ significantly between insular populations of Big and ‘normal’-sized rodents (figure 1b–g). Among these seven variables, values were generally lower for heterogeneity of island terrain (a composite measure of area and elevation), mean temperature (averaged spatially) and the spatial standard deviation in annual precipitation for the Big populations, whereas distance to continent, minimum NPP and mean annual precipitation tended to be higher. Three variables were distributed significantly differently between island populations of ‘normal’-sized rodents and Small insular populations, which, by contrast, tended to experience high mean annual temperature, a high standard deviation of NPP and lower mean precipitation.

Big and Small insular populations differed significantly in four variables. Mean precipitation was higher for Big populations, mean temperature and the spatial standard deviation of NPP were both higher for Small populations, and Big populations were found at higher latitudes than Small ones.

(c) Classification tree analyses
The classification tree based on our original dataset (figure 2a) correctly classified 90% of its extreme cases of insular body size changes (98% of the Big and 62% of the Small populations correctly predicted, Cohen’s \( \kappa = 0.671, p < 0.001 \)), whereas the random forest correctly predicted 87% of the size changes (97% of the Big and 52% of the Small, Cohen’s \( \kappa = 0.559, p < 0.001 \)). The first split on the classification tree distinguishes between different dietary categories. While populations in most dietary categories overwhelmingly tend to become Big on islands (102 populations becoming Big, 10 becoming...
Small), frugivores and granivores vary (30 Big, 27 Small). Size change in frugivores and granivores appears to depend on the mean productivity of the islands they inhabit, the tendency being to become Small on low-productivity islands (one Big, 13 Small populations). Frugivore and granivore populations on high-productivity islands are subdivided a final time on the basis of the heterogeneity of the productivity landscape on the island. Islands that vary in productivity spatially (i.e. with high spatial productivity standard deviations) tended to produce Small body sizes (two Big, 10 Small populations), whereas those with low standard deviations produced Big body sizes (27 Big, four Small).

When applied to a new dataset derived from [12], the classification tree was able to correctly classify 33 of the 39 ‘extreme’ populations (85%: 94% of Big populations, 20% of Small populations correctly predicted), while the random forest correctly predicted 90% of the new populations (100% of Big populations, 20% of Small populations correctly predicted).

4. Discussion

From this examination of extreme insular body size evolution, it is clear that island rodents exhibit some striking differences from their mainland counterparts. Not only do the distributions of size ratios differ between island and mainland populations, but over half of the insular rodent populations fall outside the boundaries encompassing 95% of the size variation ordinarily found among mainland populations—a pattern unlikely to have arisen through founder effects alone.

We found a greater than expected proportion of insular rodent populations of extremely small body size (12% of all island populations, 4.8 times the frequency observed on the mainland), but the great majority of extreme sized island rodents were extremely big (42% of all island populations, 16.8 times the frequency on the mainland). The predominance of extremes of larger size could suggest that a form of ‘immigrant selection’ was at work [30], with greater success of large-bodied individuals surviving dispersal over water; however, these percentages stand in stark contrast to the broader pool encompassing all island rodents in our dataset (including non-extreme and extreme body size), where it is a virtual coin toss whether a population will increase or decrease in body size relative to the closest mainland populations—a pattern unlikely to have arisen through founder effects alone.

The idiosyncratic nature of insular rodent populations of extremely reduced body size is also reflected in the groups identified by the classification tree. Figure 2b shows the differences in body mass between sets of populations predicted to be extremely big and others predicted to be extremely small. The two terminal nodes in which body size was predicted to be extremely big comprise a large number of populations (112 and 31, respectively, together representing 84.7% of all extreme populations) and have a wide range of masses, with each encompassing almost the entire range of rodent body sizes considered in this analysis (6–304 g; populations for which the predictions of size increase were accurate spanned the same range). By contrast, the terminal nodes for which size was predicted to have decreased consist of fewer samples (12 and 14, 15.3% of extreme populations) and have much narrower body size ranges, and these narrow ranges lie at opposite ends of the size spectrum for rodents, separated by an order of magnitude. In one set, and noting the cases in which the prediction of size decrease was accurate, all but two of the populations were large rodents (mean = 357 g, median = 322 g) that had become extremely smaller (comparatively ‘Small’) on islands; in the other set, 13 out of 14 populations were small rodents (mean = 24 g, median = 16 g) that had become extremely small. Based on these patterns, it appears that rodents on islands are generally extremely big, that extreme insular body size decreases represent a deviation from the more typical pattern, and that each group of size-reduced insular populations requires a distinct set of variables to explain its deviation. On the classification tree, populations of extremely reduced rodents were dispersed among several terminal nodes, suggesting that different insular circumstances lead distinct subgroups to evolve substantially smaller size. Yet despite differences among these subgroups in the specific sets of relevant variables, resource availability may be an underlying factor that influences all of them. All groups that were successfully predicted to be extremely reduced in size are fruit and grain specialists found on islands with productivity that is either low or spatially heterogeneous (i.e. ‘complex’)—a finding aligned with predictions from previous studies [31,32]—whereas the fruit and grain specialists that were predictably greatly enlarged are found on islands with homogeneous productivity landscapes. As resource availability declines much faster with decreasing area for fruit/grain specialists than for more generalist species [32,33], small frugivore/granivores on low-productivity islands may not be able acquire the resources necessary for maintaining populations of a ‘normal’ body size. Likewise, large frugivore/granivores on islands with a heterogeneous productivity landscape may be similarly limited owing to their inability to access patchy resources across an island.

The factors identified as significant by the K–S tests largely agree with the classification tree, suggesting that resource-related variables are major drivers of extreme size change. The concordance between the classification tree and the K–S tests is particularly noteworthy, given that some island populations classified as ‘normal’ might, with additional time or genetic variation, have diverged significantly in size. Also noteworthy is the transitivity of the K–S results—differences between ‘Small’ and ‘normal’ were in the same direction as differences between ‘Small’ and ‘Big’—which need not have been the case. Terrain heterogeneity, temperature, precipitation and productivity differ in their distributions for islands that produce ‘extreme’ rather than ‘normal’ body sizes in their rodent inhabitants: Greatly enlarged rodents tend to be found on small, homogeneous islands with above-average precipitation and high, relatively consistent resource availability; extremely reduced rodents are typically found on hot, dry islands where availability of resources is spatially variable.

While the classification tree and random forest both excelled in classifying the data used to build the models, their success at predicting independent data is perhaps even more impressive. Of the 34 cases of insular rodent populations in the independent dataset that exhibited extreme size
increase, the classification tree successfully predicted all but one while the random forest correctly classified all of them. Although the classification rate for populations of small body size from that dataset was less successful (one of five correctly classified), all four of the misclassified populations reached the same terminal node as the largest group of misclassified populations from the original data: extremely reduced non-frugivorous/granivorous rodents. As measurements such as NPP and diet are imprecise proxies for factors like resource availability, we might anticipate that predictions would improve with the inclusion of more detailed diet and resource availability data.

The results of our analyses underscore the importance of taking variation that is commonly found on the mainland into account when identifying factors that produce unusual body sizes on islands. Striking differences exist between mainland and island populations, but many of the factors associated with changes in body size in insular rodents come into focus only when the statistical noise produced by variation ordinarily found also on the mainland is removed. Specifically, it becomes clear that some version of the island rule holds true for rodents, with the great majority of insular populations that undergo extreme changes increasing in size. While exceptions to this rule still exist, and while additional data not currently available for all populations we considered might clarify trends and further enhance our understanding of them (e.g. electronic supplementary material, Appendix S1), we have developed a simple yet highly successful predictive model of insular body size change based largely on diet and resource availability that correctly classifies over 90% of extreme populations, remains robust when classifying novel samples and suggests future avenues for more mechanistic studies of this phenomenon. Although previous authors have despaired of finding generalities that could predict the response of populations to novel insular environments, using models designed to handle the very factors and contingencies that frustrated previous work, we find that significant patterns do emerge. These patterns and the resulting predictability of size change found in this study represent a major step in the description and understanding of an ecotypic ‘rule’ half a century in the making, and they open the way for further work at the population level to examine the mechanisms and processes that affect body size and its correlates.

Data accessibility. Both the island and mainland body size databases will be accessible on Dryad following publication of this manuscript. Island database: Dryad (http://dx.doi.org/10.5061/dryad.1b736). Mainland body sizes: Dryad (http://dx.doi.org/10.5061/dryad.sdnj).

Authors’ contributions. P.A.P.D. initiated the project, assembled the data, conducted analyses, prepared the figures and drafted the manuscript; V.L.R. participated in project design, interpretation of findings and writing the paper.

Competing interests. We declare we have no competing interests.

Funding. This work was supported in part by a graduate fellowship from NESCent, National Science Foundation EF-0905606 (P.A.P.D.) and a National Science Foundation DDIG no. 1311597 (V.L.R., P.A.P.D.).

Acknowledgements. We thank the Natural History Museum of Los Angeles County (especially Jim Dines) and the Smithsonian, US National Museum of Natural History for making their collection databases available for this study.

References

22. Millen V, Damuth J. 2004 Climate change and size evolution in an island rodent species: new


