The urban tree of life: quantifying relationships between body size and urban tolerance for more than 30,000 plant and animal species

Corey T. Callaghan[1*], Diana E. Bowler[2], Vaughn Shirey[3,4], Brittany M. Mason[1], Laura H. Antão[5], Ingmar Staude[6,7], John H. Wilshire[8,9], Thomas Merckx[10]

[1]Department of Wildlife Ecology and Conservation, Fort Lauderdale Research and Education Center, University of Florida; Florida, United States of America.

[2]UK Centre for Ecology & Hydrology; Wallingford, United Kingdom.

[3]Department of Biology, Georgetown University – Washington; District of Columbia, United States of America.

[4]Marine and Environmental Biology Section, Department of Biological Sciences, University of Southern California - Los Angeles; California, United States of America.

[5]Research Centre for Ecological Change, Faculty of Biological and Environmental Sciences; University of Helsinki, Helsinki, Finland.

[6]Leipzig University; Leipzig, Germany.

[7]German Centre of Integrative Biodiversity Research (iDiv); Halle-Jena-Leipzig, Germany.

[8]Department of Ecology and Evolutionary Biology, Yale University; Connecticut, United States of America.

[9]Center for Biodiversity and Global Change, Yale University; Connecticut, United States of America

[10]WILD, Department of Biology, Vrije Universiteit Brussel; Brussels, Belgium.

*Corresponding author. Email: c.callaghan@ufl.edu

Statement of authorship:

CTC, DEB, TM conceptualized the project and developed project methodology. CTC, TM, VS, JHW, BMM conducted analyses. CTC conducted project administration and supervision. CTC, DEB, TM wrote the original draft, and CTC, DEB, VS, BMM, LA, IS, JHW, TM were involved in review and editing.

Data accessibility:

All data analysis was conducted in R statistical software and relied heavily on tidyverse (Wickham et al. 2019). The doi representing our download is: https://doi.org/10.15468/dl.4dcbgt. Our processing of the GBIF data to quantify urban scores is available here: https://github.com/coreytcallaghan/body_size_vs_urban_tolerance. Not all body size data can be openly shared due to copyright restrictions, however the freely available dataset, cleaned and processed for our analysis is available here: https://github.com/coreytcallaghan/body_size_vs_urban_tolerance. In this repository we did

provide reproducible data and code, albeit not for the entire dataset of body sizes. We will archive a cleaned version of this repository in Zenodo upon acceptance of this manuscript.

Competing interests:

Authors declare that they have no competing interests.

Abstract:

Urbanization is a major global driver of biodiversity change, with species responses to urban settings ranging from avoidance to exploitation. To better understand these responses, we conducted a global analysis of urban tolerance across more than 30,000 animal and plant species, with a focus on body size—an ecological trait fundamental to space use, mobility, metabolism, and environmental sensitivity. Our synthesis showed a consistent pattern across taxa and biogeographic regions: many species are urban avoiders, while few thrive as urban exploiters—a pattern we coin "Species Urbanness Distribution". We further found that larger body sizes were generally associated with greater urban tolerance in plants compared to animals, though these size-tolerance relationships showed considerable variability among families. Our findings highlight the heterogeneous relationship between body size and urban tolerance across the tree of life, underscoring the importance of tailored strategies to support urban biodiversity. This research advances ecological understanding of urban filtering and provides a framework for guiding biodiversity-sensitive urban planning amid accelerating global urbanization.

Keywords: animals; body size; plants; Species Urbanness Distribution (SUD); traits; urban ecology; urban avoiders; urban exploiters; urban filtering; urban tolerance

Main

Cities represent novel anthropogenic environments, leading to significant mismatches with the evolutionary history of most organisms[1]. As urban areas are set to expand two- to sixfold over the 21st century[2], urbanization poses a substantial and accelerating threat to global biodiversity[3]. Yet species vary widely in their responses: some are completely extirpated ('urban avoiders'), others persist through behavioral or ecological acclimation ('urban adapters') and/or rapid adaptive evolution to cities[4,5], and some even thrive in urban environments[6] ('urban exploiters'). Quantifying these divergent responses is key for understanding how urbanization will continue to shape global biodiversity and urban ecosystem functioning (e.g., urbanization reshapes a food web[7,8]). However, the young field of urban ecology has typically relied on small-scale studies, and/or has often focused on relatively few species constrained to particular taxonomic groups such as birds[9] or mammals[10]. Thus, whether general patterns of urban preference or avoidance emerge across species within and across regional pools remains untested.

Species likely do not respond randomly to urbanization as traits may modulate species' responses. Trait-based approaches are powerful for identifying generalizable patterns across diverse taxa, particularly at broad spatial and phylogenetic scales[11]. Although many traits (e.g., diet, reproductive strategy, dispersal ability) may influence species-specific responses to urbanization, body size stands out as both the most widely measured and most taxonomically integrative. It is a key trait of any organism, relating to space use, life-history, and metabolic rate across the tree of life[12,13,14], making it a plausible direct or indirect predictor of species' responses to key stressor gradients in urban landscapes, including resource fragmentation, heat stress from urban heat islands (UHI), and anthropogenic disturbance (Fig. 1).

42

43 Body size correlates with urban mobility, but in complex ways[15]. For instance, while higher
44 mobility allows some animal species to better cope with the fragmented nature of urban habitat
45 resources, reduced mobility might be advantageous for exploiting localized resources and
46 avoiding risks associated with the urban matrix[16,17]. In plants, body size—often indexed by
47 height—correlates with competitive ability, light acquisition, and reproductive strategy, with
48 larger species generally having greater resource needs but potentially higher resilience to urban
49 stressors[18]. Notably, the UHI may drive shifts to smaller-bodied animal species due to elevated
50 metabolic costs[14,15], echoing similar patterns observed under global warming[19,20]. However, these
51 shifts toward smaller size can be overruled by requirements for increased urban mobility in taxa
52 where mobility increases with body size[15]. All in all, despite body size's fundamental role in
53 modulating species responses to their environment, a systematic, cross-taxon assessment of how
54 urbanization filters body size distributions is lacking.

55

56 Here, we present a global synthesis of urban tolerance across more than 30,000 animal and plant
57 species, using occurrence records from the Global Biodiversity Information Facility (GBIF) and
58 remotely sensed night-time light intensity as a proxy for urbanization. We define a continuous
59 metric of urban tolerance as a species' relative affinity to urban areas, with negative values
60 indicating avoidance and positive values indicating preference[21]. To account for regional
61 variation in both species pools and in species' urban tolerance across their range[22,23], we
62 calculated urban tolerance within 52 geographic subrealms (Fig. S1). Our dataset spans 47
63 taxonomic classes—from flowering plants (8,980 species) to insects (8,674 species) to mammals
64 (648 species)—and represents a uniquely broad assessment across the tree of life. We first

introduce the concept of "Species Urbanness Distribution" (SUD) to characterize the composition of a regional community in terms of its species' urban affinities. We then use Bayesian hierarchical modeling to test whether body size predicts urban tolerance across 17,722 species from 371 families, across multiple subrealms. These analyses allow us to quantify both consistency and heterogeneity of body size effects on urban tolerance across multiple taxonomic levels, and across regions. This synthesis addresses two core questions: (1) What is the shape of species' urban tolerance distributions within regional communities? (2) Does body size consistently correlate with species' urban tolerance across taxonomic groups and biogeographic contexts?

Results and Discussion

Species Urbanness Distributions (SUDs)

We identified a consistent pattern in the extent to which regional communities—defined as all species within a given biogeographic subrealm—consist of urban avoiders and exploiters. We term this the Species Urbanness Distribution (SUD), which captures the full distribution of urban tolerance values across species in a community. SUDs exhibited a characteristic shape: a pronounced peak at negative values and a long right-skewed tail toward high positive values, indicating that most species are 'urban avoiders', while only a few are 'urban exploiters'. This pattern was consistent across subrealms and several taxonomic levels (Fig. 2). Conceptually, SUDs parallel the well-known Species Abundance Distributions[24] (SADs), which describe the general law of communities being composed of many rare and few common species[24,25]. Similarly, much like the skewed distributions observed in SADs[24,26], the skewed shape of SUDs indicates that only a select few species seem to possess traits that enable them to thrive in urban

environments. While previous studies have reported a similar pattern at smaller spatial scales[27], our synthesis demonstrates this applies across the globe.

The skewed shape of SUDs suggests that traits enabling species to tolerate urban environments are both rare and specific—an indication that adaptation (*sensu lato*) to urban settings is uncommon, yet occurs geographically widespread across taxa. These traits may be partly found in human-commensal species, which are found globally and often include invasive non-native species[28]. Such traits "pre-adapted" to urban conditions allow for some species to not only persist but thrive in urban environments where most species cannot. Indeed, the median urban tolerance is most often below one, indicating widespread avoidance among species. This is expected, given that urbanization typically involves severe loss and fragmentation of habitat resources for most organisms[1], which negatively impacts numerous taxa, including birds[29], beetles[30], and moths[31]. Additionally, the UHI—typical of most cities[32]—imposes thermal stress, especially during extreme heat events[33], which may drive community shifts toward heat-tolerant species, as observed in ants[34], bees[35], and plants[36]. Yet, urban environments also frequently contain green infrastructure and remnants of natural habitats that can act as sanctuaries for rare and endangered species[37], though these spaces are often also dominated by widespread generalist species. The multitude of environmental stressors, combined with competition dynamics between urban-tolerant and intolerant species, likely influence the shape of SUDs (Fig. 2). Future ecological research should aim to disentangle these contrasting drivers across taxa and regions to better understand the ecological filters at play.

The relationship between body size and urban tolerance

To test whether there is a consistent pattern of body size filtering associated with urban tolerance across the tree of life, we integrated our species-level urban tolerance values with body size measures collated from the literature. We performed this analysis for 17,722 species spanning both plants (i.e., plant height) and animals (e.g., body length, mass, wingspan). We first evaluated the overall relationship between body size and urban tolerance across all taxa, fitting two models (see Methods for details). We found that the effect of body size on urban tolerance is stronger in plants than in animals (0.64 vs. 0.21), although uncertainty was substantial in both groups, with zero-overlapping 95% credible intervals (plants: -0.68 to 1.84; animals: -0.25 to 0.66). Variation in urban tolerance was most pronounced at the family level (sd = 2.43), followed by class and order level (sd = 2.34 and 1.75, respectively). Similarly, the effect of body size varied most among families (sd = 1.24), compared to orders (sd = 0.73) and classes (sd = 0.27), suggesting that body size predicts urban tolerance most consistently within the finer family-level taxonomic scale, rather than at higher scales where other factors may obscure relationships. Based on these patterns, we focused subsequent analyses at the family level, fitting Bayesian models for families with data on at least ten species to independently explore the relationships between urban tolerance and body size.

Across 371 families (93 plant and 278 animal families), we found no universal relationship between body size and urban tolerance; instead, patterns were heterogeneous (Fig. 3). Using 80%, 90%, and 95% credible intervals to infer weak, moderate, and strong effects, respectively, we found evidence of body size filtering in 83 families (29 strong, 19 moderate, and 35 weak). These included 48 animal families (17% of animal families) and 35 plant families (38% of plant families), indicating that for most families, body size does not consistently predict urban

tolerance. This suggests that other traits or ecological factors likely underpin the emergence of urban exploiters and the skewed shape of SUDs. Nonetheless, among plant families, the effect of body size on urban tolerance was predominantly (N=77; 83%) positive, with 23 families showing strong or moderate effects. In contrast, animal families (N=278) exhibited more variable responses: 58% showed positive effects and 42% negative, with only a subset showing strong or moderate effects (positive: 18 families; negative: 6 families). Notable families with the strongest positive associations included Salicaceae (willows) and Columbidae (doves and pigeons), where larger species are more urban-tolerant. In contrast, Dipsadidae (snakes) and Accipitridae (birds of prey) showed the strongest negative associations. Further discussion of notable order-level effects is provided in the Supporting Information. While the overall patterns for family-level effects are visible in Fig. 3, we also developed an interactive online visualization of family-level effect sizes, available here.

The variability in both the direction and strength of urban body size filtering across taxa likely reflects the diversity of urban stressors and the contrasting taxon-specific pathways through which body size mediates sensitivity to them. In raptors, for example, the negative association between body size and urban tolerance—consistent with previous findings[38,39]—likely stems from large species' need for extensive hunting territories, whose availability and access are constrained by urbanization-driven habitat loss and fragmentation. Similarly, the strong negative effect observed in Coleoptera supports the hypothesis that the UHI drives shifts to smaller body sizes in urban-tolerant ectotherms, consistent with Atkinson's temperature-size rule[40]. Larger beetles may also be more vulnerable to urban disturbance due to lower reproductive output and reduced dispersal capacity[41]. In contrast, urban communities of Lepidoptera (moths and

157 butterflies) typically shift toward larger—and hence more mobile—species than rural
158 communities[15], likely because greater mobility facilitates persistence in typically fragmented
159 urban landscapes[42].
160
161 Order-level analyses (see Supplementary Text) offer further insight into the mechanisms shaping
162 body size responses to urbanization. In plants, excluding grasses and ferns, we consistently
163 observed a positive association between body size and urban tolerance. Taller species may gain a
164 competitive edge in urban settings[43], by exploiting higher temperatures, increased CO_2 and
165 nitrogen availability[44], and typically reduced O_3 concentrations[45]. In contrast, grasses and ferns
166 favored smaller species, likely stemming from high disturbance regimes in urban grasslands and
167 ruderal environments. Among ectotherm animals—including Diptera, Coleoptera, Hemiptera,
168 and Squamata—the general expectation is a shift toward smaller species, consistent with
169 metabolic constraints imposed by the UHI (Fig. 1, Fig. 3). However, exceptions occur when
170 urban fragmentation requires greater mobility, favoring larger and more mobile species[15]. Urban
171 spiders also exhibit body size reductions, driven by both elevated temperatures and reduced prey
172 size[46]. For endotherms such as rodents and birds, we typically anticipate a shift toward larger
173 species because of increased habitat fragmentation, especially when body size correlates
174 positively with mobility. While this pattern held for most bird families, notable exceptions—like
175 Piciformes and Accipitridae—demonstrate negative associations, possibly due to constraints
176 related to large home range requirements and prey size[47,48]. Future research should further
177 explore such trophic consequences of urban size filtering, as these effects can either align across
178 trophic levels or become mismatched. For instance, urban environments may favor larger
179 predators—such as insectivorous birds and mammals—while simultaneously selecting for

180 smaller insect prey species. Such mismatches can lead to nutritional stress, reduced reproductive
181 success, and ultimately a more severe homogenization of urban predator guilds[49,50]. Another
182 ecological consequence of size mismatches may occur when urban pollinator assemblages
183 become disproportionally large or small relative to the floral traits of native plants, potentially
184 reducing pollination efficiency[51]—although rapid evolutionary shifts in floral morphology could
185 help mitigate these effects[52].
186
187 **Current limitations and future directions**
188 Our analysis, incorporating over 17,000 species, represents the most extensive assessment to date
189 of how body size relates to urban tolerance across the tree of life. Such scale was enabled by the
190 growing mobilization of biodiversity data through GBIF, particularly from citizen science
191 platforms like iNaturalist[53]. Still, our analysis largely focused on common species, as we applied
192 a cutoff of at least 100 observations per species per subrealm. As a result, rare or less frequently
193 observed species are underrepresented, and their responses to urbanization remain an important
194 avenue for future study. This is especially important given the growing recognition of cities as
195 sanctuaries for various species[37]. Nevertheless, our findings highlight the relevance of common
196 species to understanding macro-ecological patterns (sensu[54]). We focused here on broad-scale
197 patterns, but future work should further explore the context-dependency of body size filtering—
198 for instance, how it varies with city size, urban green infrastructure configuration, regional
199 climate, and human cultural preferences. For example, one possible explanation for the observed
200 trend of larger, urban-tolerant plants may lie in the widespread use of non-native woody
201 ornamentals, such as alien trees and tall shrubs, which reflect human aesthetic preferences[55,56]—
202 suggesting human preferences are an additional filter on plant size. In parallel, while our study

203 examined interspecific variation, growing evidence also points to important intraspecific
204 responses, where body size shifts within species can modulate urban tolerance[31,57,58].
205
206 We anticipate that the urban tolerance scores presented here will form a valuable foundation for
207 future research and local-scale planning efforts, in particular those that use citizen science to
208 track restoration progress. For example, these tolerance scores could inform urban greenspace
209 surveys aimed at calculating integrity indices—providing a repeatable, quantitative tool to assess
210 and monitor the long-term success of urban ecological restoration by measuring the 'urbanness'
211 of local species communities in urban greenspaces[59]. In this context, incorporating taxon-specific
212 body size-urban affinity relationships can enhance conservation outcomes by tailoring strategies
213 to the size dynamics of assemblages. For instance, in assemblages where larger species are
214 disproportionally filtered out (i.e., a negative size-tolerance link), efforts should focus on
215 increasing the size and quality of habitat patches and mitigating the UHI. Conversely, in
216 assemblages where smaller, less-mobile species are more vulnerable, improving functional
217 connectivity between habitat patches should be prioritized[60]. For butterflies in particular, Pla-
218 Narbona *et al.*[60] advocate for such context-specific management strategies—balancing habitat
219 patch quality and connectivity—to promote more diverse urban butterfly communities. Future
220 research should also test the role of other key functional traits, like seed size (Fig. S2), dispersal
221 capacity (Fig. S3), or thermoregulatory ability[61], as analyses involving multiple functional traits
222 will likely offer a better understanding of the mechanisms underlying species' urban tolerance
223 responses.
224

Conclusions

Our global analysis of over 17,000 common species provides new insight into how diverse taxa respond to urbanization, revealing a consistent skew in urban tolerance—characterized by many urban avoiders and few exploiters—across biogeographic regions and taxonomic groups. We introduce the concept of Species Urbanness Distributions (SUDs) as a novel framework to quantify and compare the impact of urbanization on species communities. Much like Species Abundance Distributions (SADs) are essential for ecology and biodiversity research, SUDs offer a generalizable lens through which the community structure and ecological resilience of urban biotas can be assessed. Understanding the processes that generate SUDs, and understanding the ecological impacts of differently shaped SUDs, represent key avenues for future research. Although body size emerged as a predictor of urban tolerance, we found not only substantial heterogeneity across families and orders, but also that body size filtering alone is unlikely to explain the consistently skewed SUD shape. Nevertheless, trait-based approaches—especially those integrating multiple functional traits—hold strong potential for uncovering the processes driving the diversity of species' urban responses and for interpreting the shape and skew of SUDs. Moreover, trait-based predictions of species' vulnerability could be used to formulate effective strategies to promote biodiversity in urban environments[62]. Looking ahead, the continued growth of citizen science data platforms, such as iNaturalist, will play an increasingly critical role in tracking and understanding the mechanisms driving biodiversity responses to urbanization. These data streams will enable broader quantification of species' urban tolerance, including rare and currently underrepresented taxa. Combined with trait-based modeling, this growing dataset could be leveraged to identify urban-vulnerable species and hence guide urban

habitat restoration initiatives. In this way, SUDs and trait-informed predictions offer powerful tools for more effective urban strategies to conserve biodiversity on an increasingly urban planet.

Methods

Our methodological approach can be broken down into three key steps: (1) quantifying species-specific urban tolerance scores, stratified by subrealm (i.e., biogeographic region); (2) collating measures of body size for as many species as possible for which we were able to quantify urban tolerance; and (3) quantifying the relationship between urban tolerance and body size at various taxonomic levels.

Quantifying urban tolerance

Our goal was to derive a continuous metric of the extent to which species vary in their responses to urban areas (hereafter urban tolerance), ranking species from those that are largely intolerant (i.e., urban avoiders) to those that thrive in urban areas (i.e., urban exploiters). To quantify species-specific urban tolerances, we followed the approach described in detail in Callaghan et al.[21,63,64].

Following Callaghan et al.[63], we downloaded data from the Global Biodiversity Information Facility (hereafter GBIF) (https://doi.org/10.15468/dl.4dcbgt) on February 4th, 2021, including ~1.4 billion biodiversity records from >24,000 datasets. GBIF data were filtered to only include observations that were recorded as species, and an additional step was taken to ensure that the listed genus in GBIF matched the genus of the species name, ensuring validity of the taxonomic nomenclature of GBIF. Due to uncertainty in matching observations with remotely-sensed

products, any GBIF observation with a coordinate uncertainty > 1 km was removed. We only included species from a taxonomic Class (i.e., the rank between Phylum and Order) that had at least 10 species reported to GBIF, focusing on the most common classes for downstream analyses. We then overlaid GBIF species occurrence observations with a remotely-sensed layer representing a continuous proxy of urbanization—Visible Infrared Imaging Radiometer Suite (VIIRS) night-time lights[65]. We used this proxy as it is a continuous measure of urbanization commonly used to represent urban extent[66,67,68]. Previous work has shown that VIIRS night-time lights is negatively correlated with greenness measured through the Enhanced Vegetation Index (EVI) and positively correlated with population density[63,65]. Further, previous work has also shown that measures of urban tolerance calculated from VIIRS night-time lights, population density, and global human modification index are all strongly correlated[59]. We used Google Earth Engine[69] for our geospatial processing and data were obtained for the VIIRS Stray Light Corrected Nighttime Day/Night Band Composites product, representing monthly composites, (i.e., this dataset in Google Earth Engine: NOAA/VIIRS/DNB/MONTHLY_V1/VCMSLCFG). We took the median of all monthly composites for each pixel to calculate a pixel-level urbanization value, measured in average radiance, and used imagery from January 2015 to January 2021 to calculate this median. GBIF data were filtered from January 1st, 2010 to February 4th, 2021. We acknowledge that these two temporal scales do not precisely match, but because urban conversion at the landscape/regional scale happens relatively slowly, we assume that the level of urbanization between 2015 and 2021 accurately captures the relative differences between regions with low and high urban land cover.

To avoid computational overload, we used geohash encoding to assign every record in GBIF to a "pixel" representing the VIIRS average radiance. We used geohash7 encoding to divide the geographic area into grid cells (referred to as blocks), each representing an approximate spatial area of 150 m^2. Each GBIF record was assigned to a corresponding geohash7 block based on its geographic coordinates. The VIIRS night-time lights data, with a native resolution of ~500 m^2, was then matched to these blocks by assigning each geohash7 block the average VIIRS radiance value that intersects it. While the geohash resolution is finer than the VIIRS data, this approach allows us to efficiently associate highly precise GBIF occurrence points with nearby VIIRS pixels using spatial indexing. This avoids unnecessary redundancy in the data while maintaining the original VIIRS resolution.

After each record from GBIF was assigned a measure of VIIRS, described above, for each species, we took the mean of the distribution of VIIRS values as a relative species-specific measure of a response to urban land cover (hereafter referred to as *urban score*). This urban score represents the relative ranking of that species to other species in response to urban land cover. To account for potential intraspecific differences in species' response to urbanization throughout their global range, we stratified our analysis to a total of 52 subrealms[70] (Fig. S1). Therefore, a species could be represented across multiple subrealms with differing measures of urban tolerance (Fig. S4). Importantly, this continuous metric of urban tolerance is a relative measure of a species' preference, or affinity, to urban areas: it should be interpreted only within each subrealm (i.e., urban tolerance scores are not interpretable across subrealms), providing a comparative or relative ranking of how a species responds to urbanization compared with other species within that subrealm. Within each subrealm, we further accounted for the potential of

315 different levels of urbanization by scaling each species' urban score by subtracting the mean
316 VIIRS of all observations in the subrealm (this value is hereafter referred to as *urban tolerance*).
317 This 'urban tolerance' (Fig. S5) value can be negative—when species under-occupy urban areas
318 suggesting they actively avoid them—or positive—when species over-occupy urban areas
319 suggesting they prefer them (i.e., ranging from urban avoiders to urban exploiters, respectively).
320
321 We applied the above procedure to any species that had at least 100 observations in a subrealm.
322 Previous work has shown that this cutoff approximates the variability of a species' response to
323 urbanization and ensures that enough of a species urban habitat use has been captured[21,71]. After
324 this filtering we were left with a total of 56,181 data points—i.e., unique urban tolerance
325 measures of a species in a subrealm—for potential inclusion in our analysis. These records
326 represented 30,373 species that had at least one urban tolerance measure, from 47 classes and
327 across 51 subrealms (Fig. S6; Data S1). The number of subrealms a species was found in was
328 mostly (64%) 1 but ranged from 1 to 34 (Fig. S7). All urban tolerance measures (i.e., urban
329 tolerance by subrealm measures) which we produced as part of our workflow are made available,
330 as we hope to advance further testing of urban tolerance among different taxonomic groups and
331 regions. But correct interpretation of these urban tolerance values is essential before use.
332
333 Sampling biases are common in citizen science data. In GBIF, urban areas are better sampled
334 compared to remote regions, as contributors typically concentrate around areas with high human
335 density. However, sampling bias across the urban modification gradient was accounted for
336 because each species' urban tolerance value was calculated relative to the observations across its
337 geographic range (i.e., within the subrealm), compared with other species within the same

338 geographic range. Importantly, while disproportionate sampling in human occupied areas may
339 inflate the urban tolerance value, each species in the dataset was subject to similar biases,
340 enabling comparisons across species. And hence, the focus here should be on comparisons
341 among species. Previous work has found that this approach of assigning urban tolerance is
342 strongly correlated with occupancy-detection models[21]. As illustrative examples, these data can
343 be used to better understand the differences in urban tolerance among Hymenoptera in Northeast
344 American Forests (Fig. S8), Lepidoptera in Southeast Asian Forests (Fig. S9), or Asterales in
345 Scandinavia & West Boreal Forests (Fig. S10).
346
347 Species urbanness distributions
348 Our analysis of species urbanness distributions was descriptive in nature, relying on data
349 visualization using ggplot2 to make histograms and density histograms of urban tolerance values
350 for each subrealm. In these analyses at the subrealm level, species are treated equally without
351 any differentiation taxonomically, compared with the taxonomically-stratified analyses where
352 species within each taxonomic rank are treated as similar.
353
354 Aggregating measures of body size
355 Our objective was to compile a dataset of species-specific measures of 'body size', or 'organism
356 size' more broadly to integrate with our measures of species-specific urban tolerance described
357 above. We aimed to find measures for as many species as possible, maximizing our sample size
358 for modeling. For plants, we used 'plant height' as a proxy for body size, and data were
359 downloaded from the TRY database—a database compilation of plant traits from many different
360 datasets[72]. For animals, body size can be measured in different ways, depending on the

361 taxonomic group of interest (e.g., body length, wingspan, biomass, shell size, radius, forewing
362 length, body mass, or biovolume). Therefore, our dataset was purposefully broad in its
363 composition of body size estimates. Because 'body size' is the predominant form referred to in
364 the literature, we use this throughout, but take it to broadly mean 'organism size' as discussed
365 above.
366
367 To develop this dataset, we first performed a semi-systematic broad literature search, using Web
368 of Science, Google Scholar, Zenodo, Dryad, and Figshare with variations of the key words:
369 "body size", "organism size", "body length", "body mass" and "interspecific", "global patterns",
370 "animals", and higher-level groups such as classes (e.g., "Aves", "Birds", "Mammals",
371 "Mammalia"). We started by downloading known compilations for various taxa, for example
372 mammals[73] or Odonata[74]. In these instances, we spent minimal further time searching for specific
373 body size data for these larger taxonomic groups, assuming that these compilations had already
374 aggregated the majority of data available. Following this, we conducted more detailed literature
375 searches, predominantly using Google Scholar and Web of Science, which focused on our *a*
376 *priori* list of species for which we had a measure of urban tolerance (Data S1). These were
377 conducted at class, order, or family taxonomic levels, but not at taxonomic levels below family.
378 Multiple keywords were used for each group while searching when it was obvious to do so. For
379 example, when searching for body size measures in Coccinellidae both "ladybirds", "ladybugs",
380 and "Coccinellidae" were used in search strings, as well as "wing length", "body length", or
381 "body size". Search effort was qualitatively proportional to the number of species we had in each
382 taxonomic level. For example, because we only had 43 species with urban scores from the class
383 Diplopoda but 3,515 species from the order Lepidoptera, we spent proportionately more time and

384 effort searching for body size data for Lepidoptera. Similarly, we spent more time on hard-to-
385 find taxa, namely groups from Insecta. In our searches we generally ignored studies that were
386 focused on intraspecific variation in body size for one or a few species, unless we found that the
387 species was included in our urban tolerance list. In addition to literature searches, we emailed
388 networks of colleagues, looked at reference lists in papers that had data, and emailed
389 corresponding authors of papers who had potentially relevant data but for which data were not
390 accessible.
391
392 After this semi-systematic searching, we re-assessed the remaining species for which we had an
393 urban tolerance measure but not a measure of body size. We then performed a second-level
394 targeted search (i.e., more systematic) for any family with at least 10 species remaining for
395 which we had a measure of urban tolerance. First, we searched Web of Science and Google
396 Scholar one family at a time and looked at the first 100 hits, with the family name 'and' "body
397 size" used as search terms. Second, we used Google search engine to search for these remaining
398 species individually (N~3,100). We focused on including species for which Google immediately
399 returned an estimate of body size (e.g., through Wikipedia, or an aggregating website), and did
400 not go to original species descriptions. This detailed searching, and the aggregation of body size
401 measures, concluded in June 2022.
402
403 *Measurements of body size*
404 In our searching for body size, we were agnostic to the type of body size, given that many
405 different measures are used for quantifying body size, or proxies for body size. For each dataset
406 we downloaded or processed, we kept track of the following key metadata properties: 'type',

‘units’, and ‘measurement detail’. Type refers to the different types of body size measurement, ranging from body length and body mass to Weber’s length in ants. Table S1 shows a list of the different body size types encompassed in our analysis (see statistical details below for how different body size types were accounted for in our analyses). For each dataset, we also kept track of the units (e.g., mm, g, mg cm^-2, or mm^3). And lastly, wherever possible, we noted the measurement detail of the body size measures, whether the mean, median, or maximum was used, or whether it was for males or females only. Many times, however, this information was not readily available, and thus this metadata field was not required in our dataset composition.

Taxonomic harmonization and dataset integration

Because body size data were aggregated across many potentially different datasets, we performed a taxonomic harmonization step to regain some species that did not match due to typos, spelling mistakes, or synonyms. For this, we treated Plantae and Animalia separately and used the taxize package[75] in R to search through the list of species and find any synonyms or fuzzy matches of other taxonomic entities, always matching back to the GBIF taxonomic backbone, given that our urban tolerance measures follow this taxonomy (Data S1). With the possible synonyms and matches, we then integrated the urban tolerance measures with the body size dataset using either direct matches or matches of synonyms. The taxonomic harmonization step regained a total of 831 (see Fig. S11) species that were integrated in the dataset for analysis. Data S3 provides details on the 223 different ‘datasets’ (i.e., downloaded data, or manually added data) that were processed as measures of body size. The most prevalent class in our final dataset (Fig. S12) was Insecta (N=7,526), followed by Magnoliopsida (N=4,872) and Aves (N=3,313).

Our final dataset for potential analysis and modeling (Data S2) included a total of 94,087 observations (i.e. unique combination of species urban tolerance values, subrealm, and body size measure) for 20,957 species that had at least one measure of urban tolerance and at least one measure of body size.

Quantifying the relationship between urban tolerance and body size

Our objective here was to understand the relationship between urban tolerance (the response variable) and body size (the predictor variable). Our dataset had potentially multiple measures of body size for an individual species, and differing levels of metadata. For example, some datasets may have specified whether the mean was used, or others might have specified for females only (see Table S1). Therefore, before modeling, our goal was to minimize the undue effect of having too many random effect levels. First, because an individual species could have multiple measures of body size, we only selected the body size measure that was most common among the species in each family. For example, if a species had both a measure of body length and of body mass, we selected the measurement type that was most commonly available among the species in each family and the largest source of data. Second, we further reduced the potential number of levels and variation in body size measurement by creating a synthetic "metadata" variable that was the combination of 'type', 'units', and 'measurement detail'. By doing this, we decreased the potential number of levels of a random effect from 223 (i.e., the number of data sources) to 71 (i.e., the combination of type, unit, and measurement detail). Third, several papers/data sources could have all reported the same metadata (e.g., maximum body length for females in mm), and instead of being treated as individual random effects because they were from different data

sources, we assumed that these represented similar data and collapsed them into one random effect level (here, for maximum body length for females in mm).

Overall kingdom-level model construction

We used Bayesian hierarchical models to test for the effect of body size on urban tolerance. In our models, urban tolerance was the response variable and body size was the predictor variable of interest. To evaluate the overall relationship between body size and urban tolerance, we fitted two comprehensive models that incorporated hierarchical taxonomic levels and the relationship with body size. The first model included fixed effects for kingdom and log-transformed body size due to its skewed distribution, along with their interaction, to investigate differences in the body size effect on urban tolerance between plants and animals. The random effects accounted for hierarchical taxonomic levels, incorporating random intercepts and slopes for class, order, and family to capture variation at each level. Additional random effects for metadata and subrealm were included as random intercepts to control for study-specific differences in data and regional differences in urban tolerance. The model formula was: urban tolerance ~ body_size_scaled_log10 * kingdom + (1 + body_size_scaled_log10 | class/order/family) + (1 + body_size_scaled_log10 | metadata) + (1 | subrealm). The second model built upon this structure but only included an interaction term between log-transformed body size and kingdom as well as kingdom as a fixed effect to determine if the effect of body size on urban tolerance varied consistently across the plant and animal kingdom. It maintained the same hierarchical random effects structure as the first model, with class, order, and family nested to account for hierarchical variation, and metadata and subrealm included as additional grouping factors. The formula was: urban tolerance ~ kingdom + body_size_scaled_log10:kingdom + (1 +

body_size_scaled_log10 | class/order/family) + (1 + body_size_scaled_log10 | metadata) + (1 | subrealm).

Family-specific model construction

However, for our main analyses we chose to employ a 'many models' approach, where we fit individual models for each family that met a set of minimum criteria because many families do not have a known association or published information on the extent to which urban tolerance associates with body size. A family was only included if there were at least 10 unique species with an urban tolerance measure and body size measure, and a subrealm was only included if it had at least two species included to specify a random slope. For example, if there was only one species of a particular family included in our dataset, this family would not have received an independent model. We decided to focus our main analysis at the family level for three key reasons. First, this was done because for most of the taxonomic families included in our analyses there is little *a priori* expectation about the effect of body size on urban tolerance (i.e., little to no literature on the subject), and therefore we wanted each model to represent an advance in knowledge of that relationship for that family. Second, because there were large imbalances in the number of data points available for each taxonomic group, for instance some families having hundreds of species with available data and others with only a handful of species, we wanted to avoid shrinkage where the model results were primarily driven by the taxonomic groups with a larger number of data points represented. Third, we wanted to avoid an instance where a family (or other taxonomic grouping) could have little or no relationship between body size and urban tolerance, but if included together in one model could result in a positive effect (see Fig. S13 for a simulated example). This was a smaller concern at the family level, which we present as the

main analysis, because of the smaller imbalance in data availability across the taxa within each family.

When we fit the family-specific models, there were different requirements for the random effect structure depending on how many subrealms and metadata sources (i.e., types of body size measures). In the simplest case, all species in a family were just from one subrealm and had just one data source for body size. But this was not always the case, and therefore we developed four different 'types of models' which varied in their treatment of random effect structure. Each 'type' of model, however, had the same overall objective—to quantify the relationship between body size and urban tolerance. The four model types are as follows. Type 1: had at least two subrealms and at least two types of synthetic metadata (see above) sources included. Type 2: had data from only one subrealm and only one type of synthetic metadata source included (but could have data from more than one initial data source). Type 3: had at least two subrealms and only one type of synthetic metadata source included. Type 4: had data from only one subrealm and at least two types of synthetic metadata sources included. For all types of model fits, the response variable was the urban tolerance measure, described above, and the fixed effect of interest was body size which was log10-transformed and scaled and centered by the synthetic metadata source. Type 1 had a random intercept and slope for subrealm, allowing the intercept and slope to vary across different levels of subrealm as well as a random intercept for the synthetic metadata source. Type 2 had no random effects. Type 3 had only a random intercept and slope for subrealm, allowing the intercept and slope to vary across different levels of subrealm. Type 4 had only a random intercept for the synthetic metadata source. Modeling was done using the 'brm()' function in R from the brms package[76,77,78]. For all model types we specified a standard

522 normal prior, with a mean of 0 and a standard deviation of 1 applied to the fixed effects of the
523 model. Models were fit with 1000 warmup iterations, 6000 iterations for the Markov Chain
524 Monte Carlo sampling, 4 chains, and the adaptation target acceptance probability was set to 0.99.
525
526 For models of type 1 and type 4, where we combined metadata for body size measurements
527 among different species, we also explored more complex model structures that included a
528 random slope for body size by metadata source, in addition to subrealm. Such a model structure
529 would allow the relationship between body size and urban tolerance to vary across measurement
530 types. However, these models often substantially reduced interpretability and inflated uncertainty
531 around slope estimates, especially when metadata groups were sparsely represented, which was
532 common. Additionally, many of the families only had 2 types of metadata, limiting our statistical
533 power to fit random slopes. Comparative analyses showed that such models produced weaker or
534 more uncertain estimates than our main model structure (i.e., random intercept for metadata of
535 body size). We also note that to account for heterogeneity in body size measurement types, all
536 body size values were log10-transformed, scaled, and centered by their metadata source prior to
537 analysis, following standard practices in allometric studies. We use two illustrative examples
538 (Figures S14 and S15) to illustrate the influence of including a random slope for metadata
539 source.
540
541 *Additional analyses beyond family level*
542 Although our main analysis was focused on the family level, one potential drawback is that
543 species could be excluded from any of the analyses for a family that has a small number of
544 species. Plus, because some analyses have previously been conducted at higher taxonomic levels

(e.g., Lepidoptera), we repeated the above procedure described for family at the order (see Supplementary Text) and kingdom level.

Acknowledgements

We thank Tim Robertson and the GBIF support team in supplying avro files of the GBIF download (https://doi.org/10.15468/dl.9ufars) for upload to our personal Google Big Query. Henrique M. Pereira provided helpful discussion in the early stages of this project. We thank the following people for help in qualitatively assessing our proxy for urban tolerance: Luis Aguirre, Timothy Bonebrake, Jason Bried, Alex Córdoba-Aguilar, Michael Crossley, Wouter Deconinck, Geert De Knijf, Tim Gardiner, Michael Greenfield, Ralf Gyselings, Frederik Hendrickx, Axel Hochkirch, Roy Kleukers, Anton Kristin, Paulo Lemos, André Victor Lucci Freitas, Onildo João Marini Filho, Daniele Matenaar, Elisa Montes-Fuentemayor, Tara Murray, Howon Rhee, Leslie Ries, Szabolcs Sáfián, Ann Swengel, Scott Swengel, Pieter Vantieghem, Wade Worthen. We thank the following people for supplying unpublished data of body sizes: Roel van Klink, Michelle Tseng, Erik Öckinger, Philip Barton, Werner Ulrich, Axel Höchkirch, Thomas Sherratt, Michiel Wallis de Vries, Lea Heidrich, Jörg Müller, Robert Tropek, Maldwyn Evans, Martin Gossner, Bernhard Hausdorf, Stano Pekar, Carlo Seifert, Onildo João Marini Filho, Sei-Woong Choi, Dominik Rabl, Markus Franzén, Elena Piano, Derek Hennen. CTC, DEB, IS acknowledge funding of iDiv via the German Research Foundation grant DFG FZT 118. CTC was supported in part by a Marie Skłodowska-Curie Individual Fellowship grant 891052 and was supported in part by the intramural research program of the U.S. Department of Agriculture, National Institute of Food and Agriculture, Hatch FLA-FTL-006297. VS was supported by Georgetown University, a U.S. National Science Foundation Graduate Research Fellowship grant 1937959,

and a David H. Smith Conservation Research Fellowship. LHA was supported by the Academy of Finland grant 340280.

References

1. Johnson, M. T., Munshi-South, J. Evolution of life in urban environments. *Science*, **358**, 6363, eaam8327 (2017).
2. J. Gao, B. O'Neill. Mapping global urban land for the 21st century with data-driven simulations and shared socioeconomic pathways. *Nature Communications* **11**, 3202 (2020).
3. K. C. Seto, B. Güneralp, L. R. Hutyra. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences* **109,** 16083–16088 (2012).
4. M. R. Lambert, K. I. Brans, S. Des Roches, C. M. Donihue, S. E. Diamond. Adaptive evolution in cities: Progress and misconceptions. *Trends in Ecology & Evolution* **36**, 239–257 (2021).
5. J. S. Santangelo, R. W. Ness, B. Cohan, C. R. Fitzpatrick, S. G. Innes, S. Koch, L. S. Miles, S. Munim, P. R. Peres-Neto, et al. Global urban environmental change drives adaptation in white clover. *Science* **375**, 1275–1281 (2022).
6. R. B. Blair. Land use and avian species diversity along an urban gradient. *Ecological Applications,* **6**, 506–519 (1996).
7. D. Start, M. A. Barbour, C. Bonner. Urbanization reshapes a food web. *Journal of Animal Ecology* **89**, 808–816 (2019).

8. E. C. Lahr, R. R. Dunn, S. D. Frank. Getting ahead of the curve: cities as surrogates for global change. *Proceedings of the Royal Society B* **285**, 20180643 (2018).
9. J. Richardson, A. C. Lees, E. T. Miller, S. J. Marsden. Avian diversity and function across the world's most populous cities. *Ecology Letters* **26**, 1301–1313 (2023).
10. J. D. Haight, S. J. Hall, M. Fidino, S. A. Adalsteinsson, A. A. Ahlers, J. Angstmann, W. J. B. Anthonysamy, et al. Urbanization, climate and species traits shape mammal communities from local to continental scales. *Nature Ecology & Evolution* **7**, 1654–1666 (2023).
11. M. F. J. Aronson, C. H. Nilon, C. A. Lepczyk, T. S. Parker, P. S. Warren, S. S. Cilliers, M. A. Goddard, A. K. Hahs, C. Herzog, et al. Hierarchical filters determine community assembly of urban species pools. *Ecology* **97,** 2962–2963 (2016).
12. S. L. Chown, K. J. Gaston. Body size variation in insects: a macroecological perspective. *Biological Reviews* **85**, 139–169 (2010).
13. J. F. Gillooly, J. H. Brown, G. B. West, V. M. Savage, E. L. Charnov. Effects of size and temperature on metabolic rate. *Science* **293**, 2248–2251 (2001).
14. J. H. Brown, J. F. Gillooly, A. P. Allen, V. M. Savage, G. B. West. Toward a metabolic theory of ecology. *Ecology* **85**, 1771–1789 (2004).
15. T. Merckx, C. Souffreau, A. Kaiser, L. F. Baardsen, T. Backeljau, D. Bonte, et al. Body-size shifts in aquatic and terrestrial urban communities. *Nature* **558**, 113–116 (2018).
16. P. O. Cheptou, O. Carrue, S. Rouifed, A. Cantarel. Rapid evolution of seed dispersal in an urban environment in the weed *Crepis sancta*. *Proceedings of the National Academy of Sciences* **105**, 3796–3799 (2008).

17. A. K. Hahs, B. Fournier, M. F. J. Aronson, C. H. Nilon, A. Herrera-Montes, A. B. Salisbury, et al. Urbanisation generates multiple trait syndromes for terrestrial animal taxa worldwide. *Nature Communications* **14**, 4751 (2023).

18. N. S. G. Williams, A. K. Hahs, P. A. Vesk. Urbanisation, plant traits, and the composition of urban floras. *Perspectives in Plant Ecology, Evolution and Systematics* **17**, 78–86 (2015).

19. J. Forster, A. G. Hirst, D. Atkinson. Warming-induced reductions in body size are greater in aquatic than terrestrial species. *Ecology* **109**, 19310–19314 (2012).

20. B. R. Scheffers, L. De Meester, T. C. L. Bridge, A. A. Hoffmann, J. M. Pandolfi, R. T. Corlett, S. H. M. Butchart, P. Pearce-Kelly, et al. The broad footprint of climate change from genes to biomes to people. *Science* **354**, aaf7671 (2016).

21. C. T. Callaghan, F. Sayol, Y. Benedetti, F. Morelli, D. Sol. Validation of globally-applicable method to measure urban tolerance of birds using citizen science data. *Ecological Indicators* **120**, 106905 (2021).

22. C. T. Callaghan, F. X. Palacio, Y. Benedetti, F. Morelli, D. E. Bowler. Large-scale spatial variability in urban tolerance of birds. *Journal of Animal Ecology* **92**, 403–416 (2023).

23. M. H. C. Neate-Clegg, B. A. Tonelli, C. Youngflesh, J. X. Wu, G. A. Montgomery, C. H. Şekercioğlu, M. W. Tingley. Traits shaping urban tolerance in birds differ around the world. *Current Biology* **33**, 1677–1688 (2023).

24. B. J. McGill, R. S. Etienne, J. S. Gray, D. Alonso, M. J. Anderson, H. K. Benecha, M. Dornelas, B. J. Enquist, et al. Species abundance distributions: Moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* **10**, 995–1015 (2007).

25. F. W. Preston. The commonness, and rarity, of species. *Ecology* **29**, 254–283 (1948).

26. C. T. Callaghan, L. Borda-de-Água, R. van Klink, R. Rozzi, H. M. Pereira. Unveiling global species abundance distributions. *Nature Ecology & Evolution* **7**, 1600–1609 (2023).

27. J. E. Humphrey, A. Haslem, A. F. Bennett. Avoid, adapt or exploit: Re-visiting bird responses to urbanization using a novel landscape approach. *Global Ecology and Conservation* **48**, e02735 (2023).

28. D. Sol, C. González-Lagos, O. Lapiedra, M. Díaz. Why are exotic birds so successful in urbanized environments? in *Ecology and Conservation of Birds in Urban Environments* (Springer Cham, 2017), pp. 75–89.

29. K. L. Evans, S. E. Newson, K. J. Gaston. Habitat influences on urban avian assemblages. *Ibis* **151**, 19–39 (2009).

30. E. Piano, K. De Wolf, F. Bona, D. Bonte, D. E. Bowler, M. Isaia, L. Lens, T. Merckx, D. Mertens, M. van Kerckvoorde, L. De Meester, F. Hendrickx. Urbanization drives community shifts towards thermophilic and dispersive species at local and landscape scales. *Global Change Biology* **23**, 2554–2564 (2017).

31. T. Merckx, A. Kaiser, H. Van Dyck. Increased body size along urbanization gradients at both community and intraspecific level in macro-moths. *Global Change Biology* **24**, 3837–3848 (2018).

32. G. Manoli, S. Fatichi, M. Schläpfer, K. Yu, T. W. Crowther, N. Meili, P. Burlando, G. G. Katul, E. Bou-Zeid. Magnitude of urban heat islands largely explained by climate and population. *Nature* **573**, 55–60 (2019).

33. L. Zhao, M. Oppenheimer, Q. Zhu, J. W. Baldwin, K. L. Ebi, E. Bou-Zeid, K. Guan, X. Liu. Interactions between urban heat islands and heat waves. *Environmental Research Letters* **13**, 034003 (2018).

34. S. B. Menke, B. Guénard, J. O. Sexton, M. D. Weiser, R. R. Dunn, J. Silvermann. Urban areas may serve as habitat and corridors for dry-adapted, head tolerant species; an example from ants. *Urban Ecosystems* **14**, 135–163 (2011).

35. A. L. Hamblin, E. Youngsteadt, M. M. López-Uribe, S. D. Frank. Physiological thermal limits predict differential responses of bees to urban heat-island effects. *Biology Letters* **13**, 20170125 (2017).

36. N. Jogan, F. Küzmič, U. Šilc. Urban structure and environment impact plant species richness and floristic composition a Central European city. *Urban Ecosystems* **25**, 149–163 (2022).

37. C. A. Lepczyk, M. F. J. Aronson, F. A. La Sorte. Cities as sanctuaries. *Frontiers in Ecology and the Environment* **21**, 251–259 (2023).

38. D. S. Cooper, A. J. Shultz, Ç. H. Şekercioğlu, F. M. Osborn, D. T. Blumstein. Community science data suggest the most common raptors (Accipitridae) in urban centres are smaller, habitat-generalist species. *IBIS* **164**, 771–784 (2022).

39. T. Headland, D. Colombelli-Négrel, C. T. Callaghan, S. C. Sumasgutner, S. Kleindorfer, P. Sumasgutner. Smaller Australian raptors have greater urban tolerance. *Scientific Reports* **13**, 11559 (2023).

40. D. Atkinson. Temperature and organism size-A biological law for ectotherms? *Advances in Ecological Research* **25**, 1–58 (1994).

41. T. Magura, G. L. Lövei. Consequences of urban living: Urbanization and ground beetles. *Current Landscape Ecology Reports* **6**, 9–21 (2021).

42. T. Merckx, H. Van Dyck. Urbanization-driven homogenization is more pronounced and happens at wider spatial scale sin nocturnal and mobile flying insects. *Global Ecology and Biogeography* **28**, 1440–1455 (2019).

43. R. P. Duncan, S. E. Clemants, R. T. Corlett, A. K. Hahs, M. A. McCarthy, M. J. McDonnell, M. W. Schwartz, K. Thompson, P. A. Vesk, N. S. G. Williams. Plant traits and extinction in urban areas: A meta-analysis of 11 cities. *Global Ecology and Biogeography* **20**, 509–219 (2011).

44. R. O'Riordan, J. Davies, C. Stevens, J. N. Quinton, C. Boyko. The ecosystem services of urban soils: A review. *Geoderma* **395**, 115076 (2021).

45. D. Li, B. J. Stucky, J. Deck, B. Baiser, R. P. Guralnick. The effects of urbanization on plant phenology depends on regional temperatures. *Nature Ecology & Evolution* **3**, 1661–1667 (2019).

46. M. Dahirel, M. De Cock, P. Vantieghem, D. Bonte. Urbanization-driven changes in web building and body size in an orb web spider. *Journal of Animal Ecology* **88**, 79–91 (2019).

47. C. Sheard, M. H. C. Neate-Clegg, N. Alioravainen, S. E. I. Jones, C. Vincent, H. E. A. MacGregor, T. P. Bregman, S. Clarmunt, J. A. Tobias. Ecological drivers of global gradients in avian dispersal inferred from wing morphology. *Nature Communications* **11**, 2463 (2020).

48. A. Arango, J. Pinto-Ledezma, O. Rojas-Soto, A. M. Lindsay, C. D. Mendenhall, F. Villalobos. Hand-wing index as a surrogate for dispersal ability: The case of the

Emberizoidae (Aves: Passeriformes) radiation. *Biological Journal of the Linnean Society* **137**, 137–144 (2022).

49. C. Sinkovics, G. Seress, I. Pipoly, E. Vincze, A. Liker. Great tits feed their nestlings with more but smaller prey items and fewer caterpillars in cities than in forests. *Scientific Reports* **11**, 24161 (2021).
50. L. Stidsholt, C. Scholz, U. Hermanns, T. Teige, M. Post, B. Stapelfeldt, et al. Low foraging rates drive large insectivorous bats away from urban areas. *Global Change Biology* **30**, e17063 (2024).
51. M. Stang, P. G. Klinkhamer, N. M. Waser, I. Stang, E. Van der Meijden. Size-specific interaction patterns and size matching in a plant–pollinator interaction web. *Annals of Botany* **103**, 1459-1469 (2009).
52. G. Lendvai, D. A. Levin. Rapid response to artificial selection on flower size in Phlox. *Heredity* **90**, 336-342 (2003).
53. C. T. Callaghan, M. Thompson, A. Woods, A. G. B. Poore, D. E. Bowler, F. Samonte, et al. Experimental evidence that behavioral nudges in citizen science projects can improve biodiversity data. *BioScience* **74**, 302–313 (2023).
54. E. H. van Nes, D. G. F. Pujoni, S. A. Shetty, G. Straatsma, W. M. de Vos, M. Scheffer. A tiny fraction of all species forms most of nature: Rarity as a sticky state. *Proceedings of the National Academy of Sciences* **121**, e2221791120 (2024).
55. L. C. Grapow, C. Blasi. A comparison of the urban flora of different phytoclimatic regions in Italy. *Global Ecology and Biogeography Letters* **7**, 367–378 (1998).

56. N. S. G. Williams, M. W. Schwartz, P. A. Vesk, M. A. McArthy, A. K. Hahs, S. E. Clemants, et al. A conceptual framework for predicting the effects of urban environments on flora. *Journal of Ecology* **97**, 4–9 (2009).

57. W. Liao, H. Lin. Urbanisation drives inter- and intraspecific variation in flight-related morphological traits of aquatic insects at different landscape scales. *Insect Conservation and Diversity* **17**, 287–303 (2023).

58. G. Schmitz, A. Linstädter, A. S. K. Frank, H. Dittberner, J. Thome, A. Schrader, K. H. L. von Ber, et al. Environmental filtering of life-history trait diversity in urban populations of *Arabidopsis thaliana*. *Journal of Ecology* **112**, 14–27 (2024).

59. C. T. Callaghan, R. E. Major, M. B. Lyons, J. M. Martin, J. H. Wilshire, R. T. Kingsford, W. K. Cornwell. Using citizen science data to define and track restoration targets in urban areas. *Journal of Applied Ecology* **56**, 1998–2006 (2019).

60. C. Pla-Narbona, C. Stefanescu, J. Pino, F. J. Cabrero-Sañudo, E. García-Barros, M. L. Munguira, Y. Melero. Butterfly biodiversity in the city is driven by the interaction of the urban landscape and species traits: a call for contextualised management. *Landscape Ecology* **37**, 1-12 (2022).

61. A. J. Bladon, M. Lewis, E. K. Bladon, S. J. Buckton, S. Corbett, S. R. Ewing, et al. How butterflies keep their cool: Physical and ecological traits influence thermoregulatory ability and population trends. *Journal of Animal Ecology* **89**, 2440-2450 (2020).

62. A. K. Hahs, B. Fournier, M. F. J. Arson, C. H. Nilon, A. Herrera-Montes, A. B. Salisbury, et al. Urbanisation generates multiple trait syndromes for terrestrial animal taxa worldwide. *Nature Communications* **14**, 4751 (2023).

63. C. T. Callaghan, D. E. Bowler, H. M. Pereira. Thermal flexibility and a generalist life history promote urban affinity in butterflies. *Global Change Biology* **27**, 3532–3546 (2021).

64. C. T. Callaghan, Y. Benedetti, J. H. Wilshire, F. Morelli. Avian trait specialization is negatively associated with urban tolerance. *Oikos* **129**, 1541–1551 (2010).

65. C. D. Elvidge, K. Baugh, M. Zhizhin, F. C. Hsu, T. Ghosh. VIIRS night-time lights. *International Journal of Remote Sensing* **38**, 5860–5879 (2017).

66. B. Pandey, P. K. Joshi, K. C. Seto. Monitoring urbanization dynamics in India using DMSP/OLS night time lights and SPOT-VGT data. *International Journal of Applied Earth Observation and Geoinformation* **23**, 49–61 (2013).

67. Q. Zhang, K. C. Seto. Can night-time light data identify typologies of urbanization? A global assessment of successes and failures. *Remote Sensing* **5**, 3476–3494 (2013).

68. C. D. Elvidge, H. Kalita, U. Choudhury, S. Rehman, B. R. Tripathy, P. Kumar. Inter-calibration and urban light index of DMSP-OLS night-time data for evaluating the urbanization process in Australian capital territory. *Applications and Challenges of Geospatial Technology* (Spring Cham, 2018), pp. 163–180.

69. N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, R. Moore. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* **202**, 18–27 (2017).

70. One Earth. Bioregions: Nature's map of the earth. *One Earth* (2023).

71. G. Liu, J. J. L. Rowley, R. T. Kingsford, C. T. Callaghan. Species' traits drive amphibian tolerance to anthropogenic habitat modification. *Global Change Biology* **27**, 3120–3132 (2021).

72. J. Kattge, G. Bönisch, S. Díaz, S. Lavorel, I. C. Prentice, P. Leadley, S. Tautenhahn, et al. TRY plant trait database – enhanced coverage and open access. *Global Change Biology* **26**, 119–188 (2020).

73. C. D. Soria, M. Pacifici, M. Di Marco, S. M. Stephen, C. Rondinini. COMBINE: a coalesced mammal database of intrinsic and extrinsic traits. *Ecology* **102**, e03344 (2021).

74. J. T. Waller, B. Willink, M. Tschol, E. I. Svensson. The odonate phenotypic database, a new open data resource for comparative studies of an old insect order. *Scientific Data* **6**, 1–6 (2019).

75. S. A. Chamberlain, E. Szöcs. taxize: taxonomic search and retrieval in R, version 2, R language (2013); https://doi.org/10.12688/f1000research.2-191.v2.

76. P. C. Bürkner. brms: An R package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software* **80**, 1–28 (2017).

77. P. C. Bürkner. Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal* **10**, 395–411 (2018).

78. P. C. Bürkner. Bayesian Item Repsonse Modeling in R with brms and Stan. *Journal of Statistical Software* **100**, 1–54 (2021).

79. K. Sand-Jensen, H. H. Bruun, T. F. Nielsen, D. M. Christiansen, P. Hartvig, J. C. Schou, L. Baastrup-Spohr. The dangers of being a small, oligotrophic and light demanding freshwater plant across a spatial and historical eutrophication gradient in southern Scandinavia. *Frontiers in Plant Science* **9** (2018).

80. S. E. Everingham, F. Hemmings, A. T. Moles. Inverted invasions: Native plants can frequently colonise urban and highly disturbed habitats. *Austral Ecology* **44**, 702–712 (2019).

List of Supplementary Materials:

Supplementary Text
Supplementary Figures S1 to S15
Supplementary Tables S1
Data S1 to S3

Figures

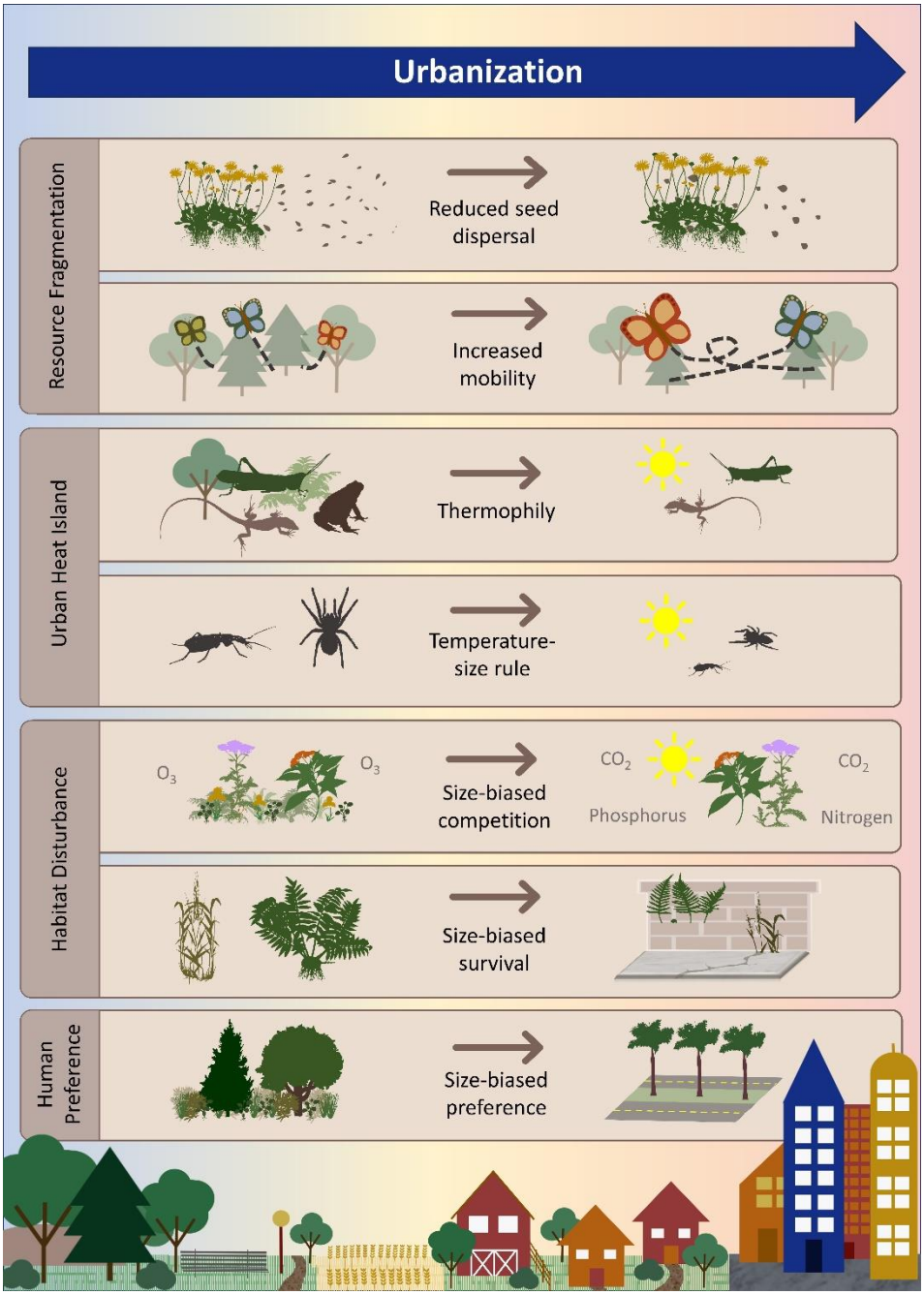

Fig. 1: We hypothesize a series of mechanisms that could explain interspecific shifts in body size as a function of urban tolerance. These hypotheses include reduced seed dispersal[59], increased mobility[42], thermophily[30], the temperature-size rule[15], size-biased competition[79], size-biased survival[80], and size-biased human preference[56].

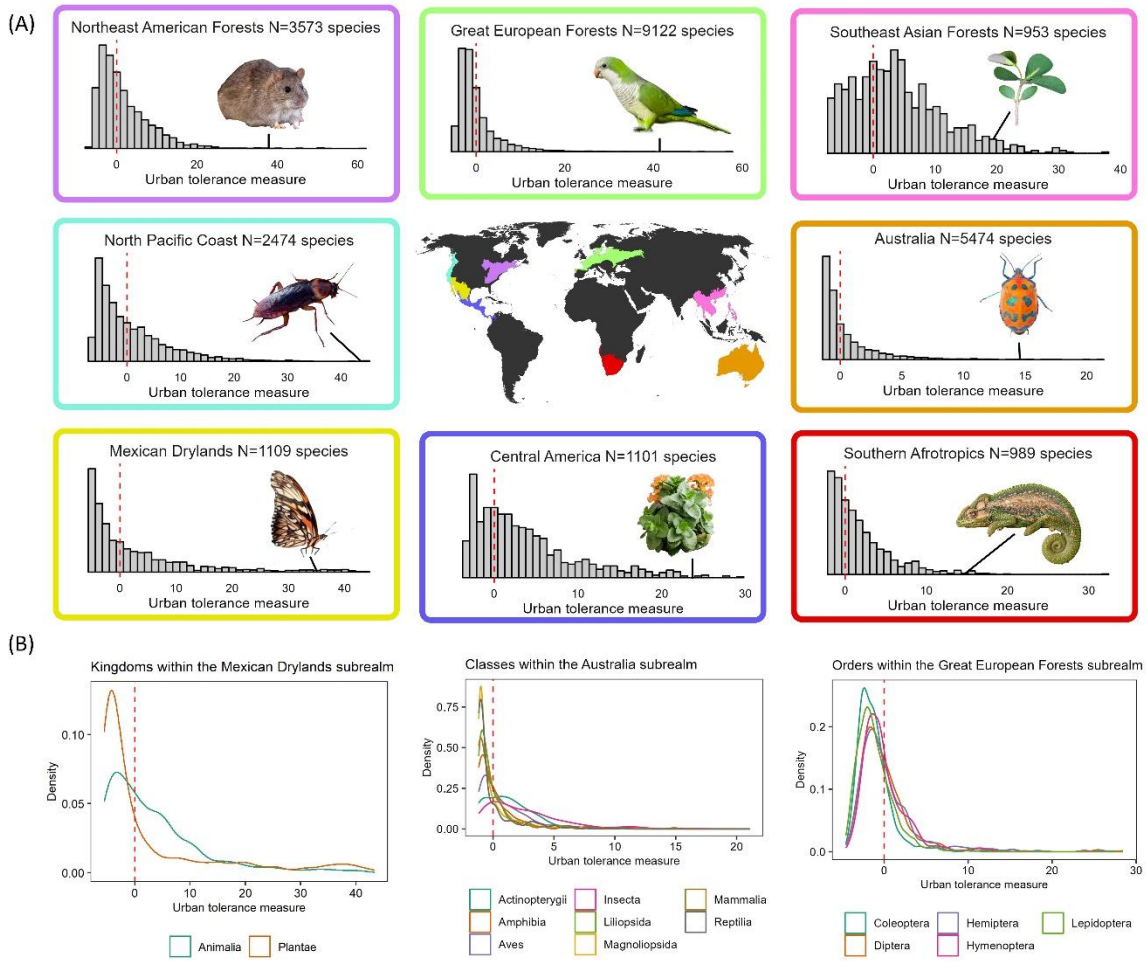

Fig. 2: Species urbanness distributions (SUDs) exemplified for eight subrealms. Plotted are all species per subrealm (**A**), with the images highlighting an example 'hyper-exploiter' species from each of these subrealms (i.e., with a high urban tolerance score). The x-axis shows the urban tolerance measure whereas the y-axis is the number of species within that bin. There were consistent patterns for kingdoms, classes, and orders (**B**) as shown by generally similar density histogram shapes for each of these. The vertical dashed line represents where species are neutral towards urban land cover. Photos acquired from iNaturalist CC BY-NC and background was removed by authors: Brown rat (© Ouwesok), Monk parakeet (© Juan Emilio), Cape dwarf chameleon (© Berkeley Lumb), American cockroach (© Len Worthington), Flaming kay (© Lyubo Gadzhev), Juno silverspot (© Rigoberto Ramírez Cortés), Hibiscus harlequin bug (© Sam Fraser-Smith), and Mascarene island leaf-flower (© Douglas Goldman).

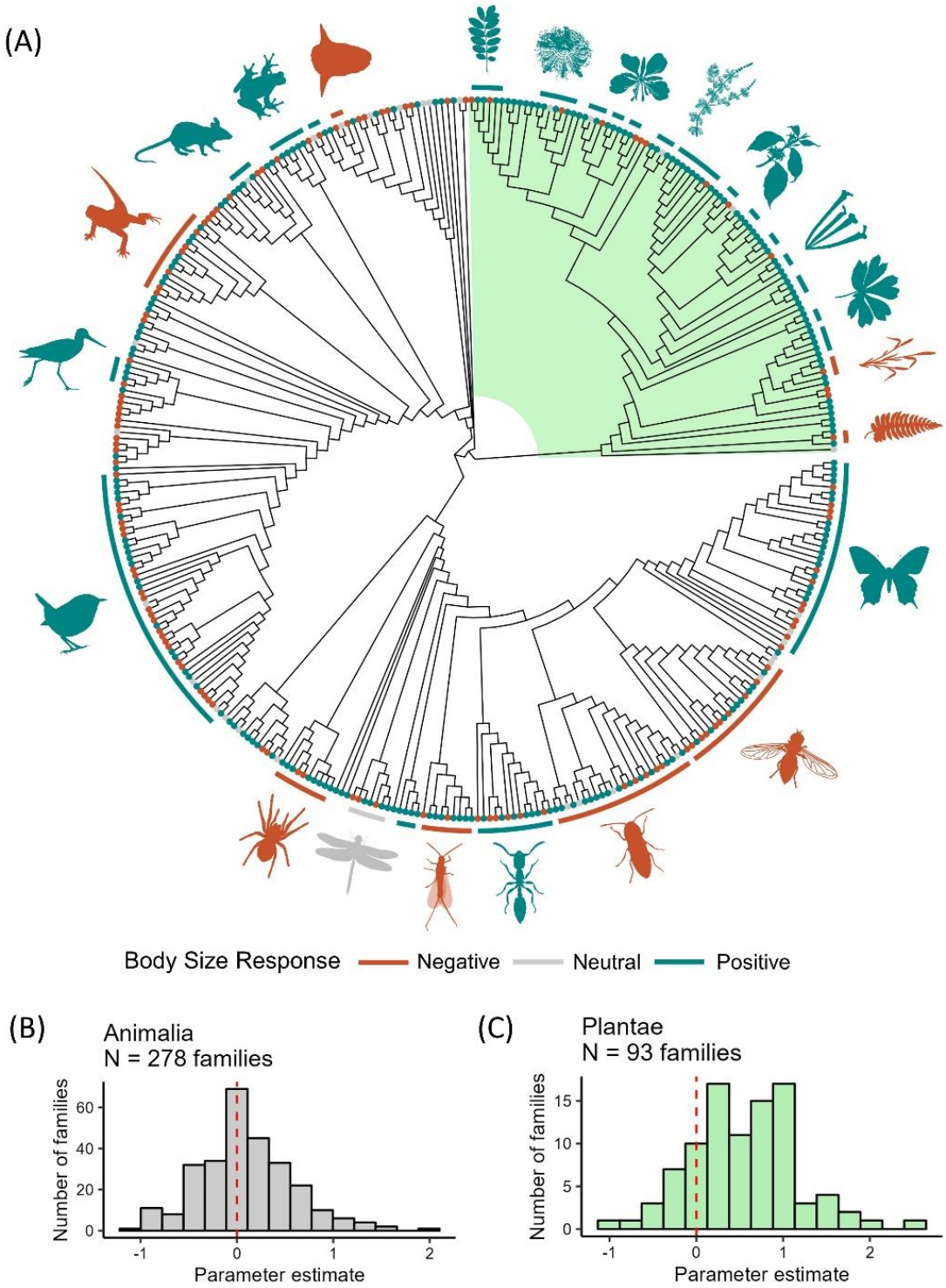

Fig. 3: Effect sizes between body size and urban tolerance across the tree of life and individual effect sizes for animals and plants families. (**A**) Effect sizes of the relationship between body size

and urban tolerance for the 371 families included in our analysis, plotted along a phylogenetic tree of life (see Methods); Plantae are highlighted and shaded in green. Colors indicate the direction of the effect: orange indicates negative, petrol indicates positive, grey indicates neutral (i.e. any effect sizes between -0.05 and 0.05). (**B**) and (**C**) histograms of individual effect sizes for each family, for animals (B) and plants (C). Orders are shown along the outside edge of the phylogenetic tree, each with a bar and icon, for any order with more than 3 families. An interactive version for full exploration of our results at both family and order level is available here.

Supplementary Materials for

The urban tree of life: Species assemblages show variable links between body size and urban tolerance across the globe

Corey T. Callaghan, Diana E. Bowler, Vaughn Shirey, Brittany M. Mason, Laura Antão, Ingmar Staude, John H. Wilshire, Thomas Merckx

Corresponding author: c.callaghan@ufl.edu

Supplementary Text

In addition to family-level effect sizes, we performed the analysis at the order level, and found similar heterogeneous results. Among the 128 orders, a total of 45 had some evidence of an effect of body size on urban tolerance (14 weak, 6 moderate, and 25 strong). Some of the orders with the strongest positive effect size included Sapindales (dicots, soapberry plants) and Columbiformes (birds, doves and pigeons), and orders with the strongest negative effect sizes included Coleoptera (insects, beetles) and Piciformes (birds, woodpeckers). Of note are the differential responses down the taxonomic guilds, where some families can show contrasting results, but there may be a strong effect when looking at the order level, or higher (e.g., Fig. S2).

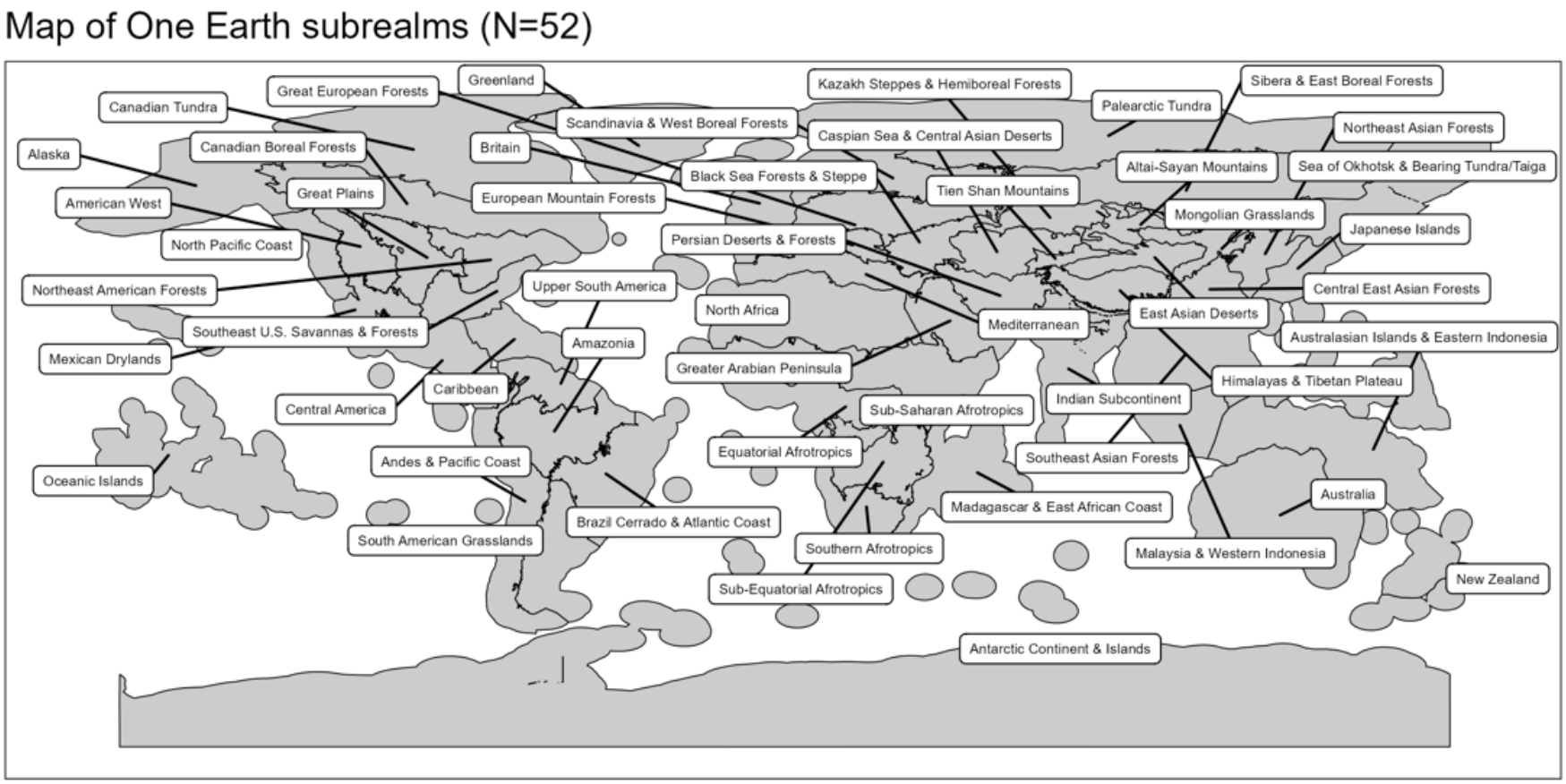

Fig. S1: We used subrealms as our geographical aggregation. Subrealms were quantified from aggregating bioregions as identified by One Earth (see more here: https://www.oneearth.org/bioregions/). The subrealms level was chosen after exploring the tradeoff between accounting for geographic differences in urban tolerance and the number of species that could be included.

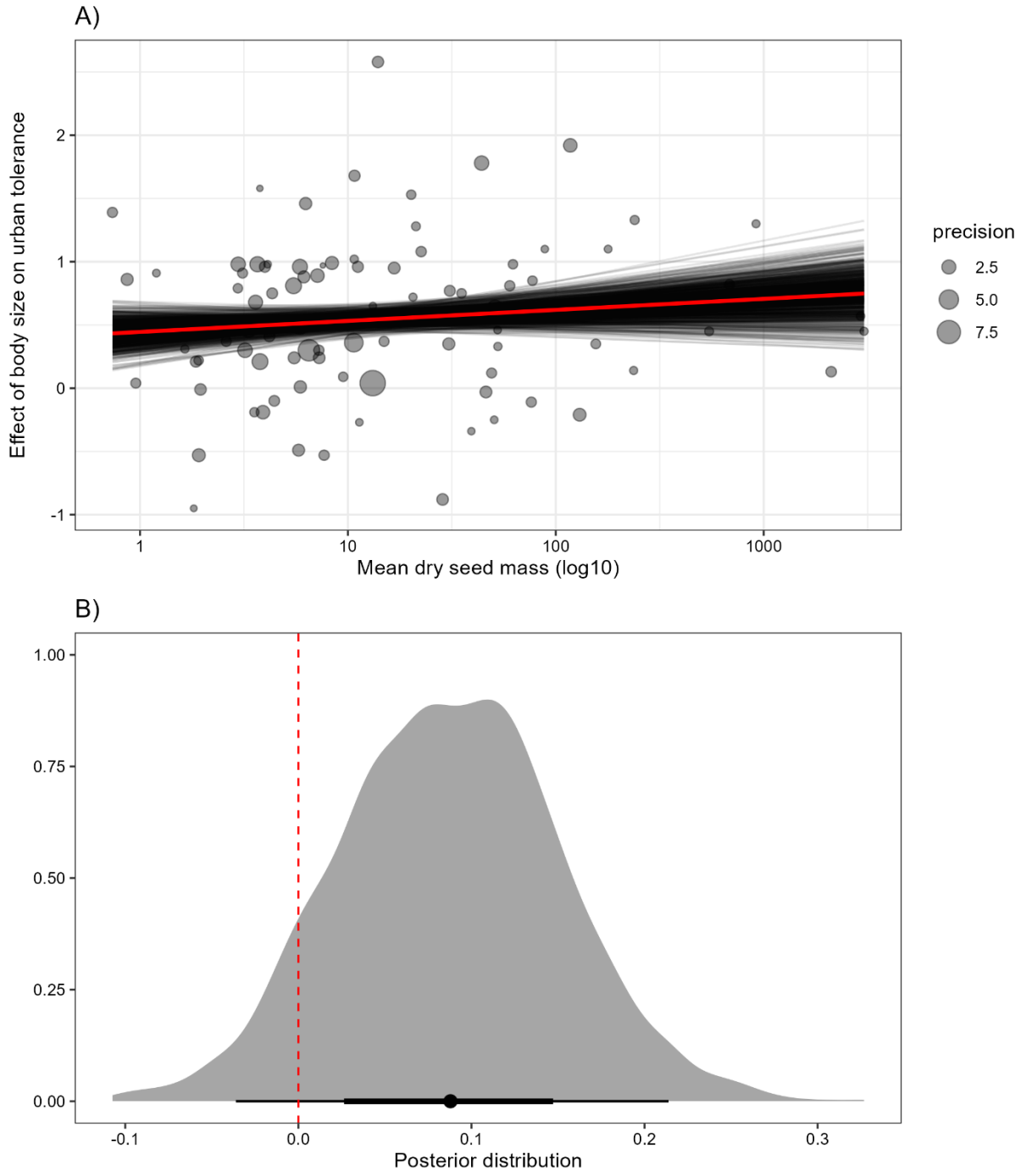

Fig. S2: (**A**) Exploratory analyses show some correlation between the effect of body size on urban tolerance for plants and the mean dry seed mass of those plants. (**B**) A Bayesian model finds moderate evidence that this is a positive relationship. This indicates that dispersal may be an underlying mechanism influencing the effect of body size (i.e., plant height) on urban tolerance.

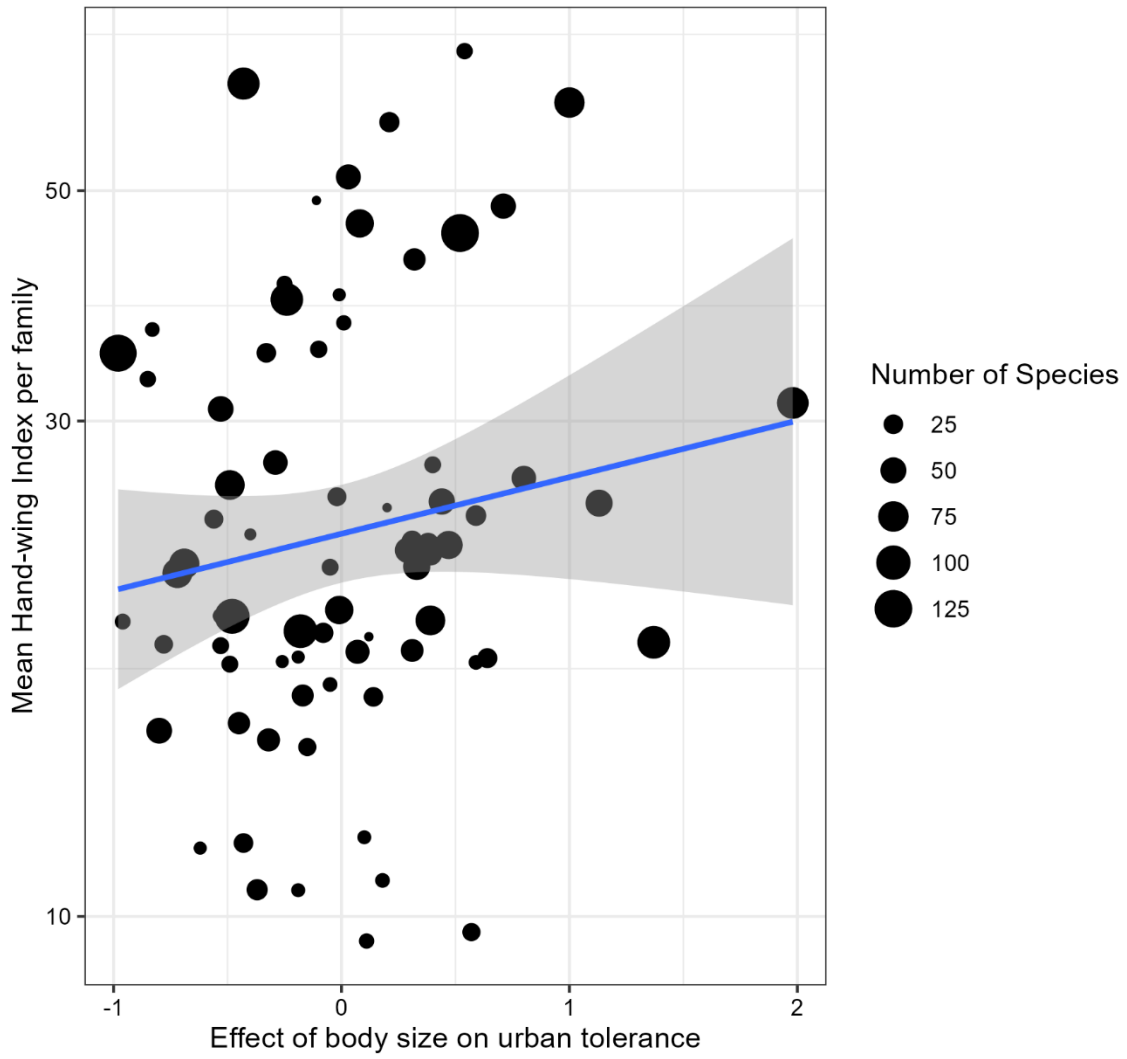

Fig. S3: Exploratory analysis for 74 bird families, showing a potential relationship between dispersal (measured using the mean hand-wing index per family) and the effect of body size on urban tolerance. This indicates that dispersal may be an underlying mechanism influencing the effect of body size on urban tolerance. The blue line is the linear model trend line and the grey shading represents the 95% confidence interval.

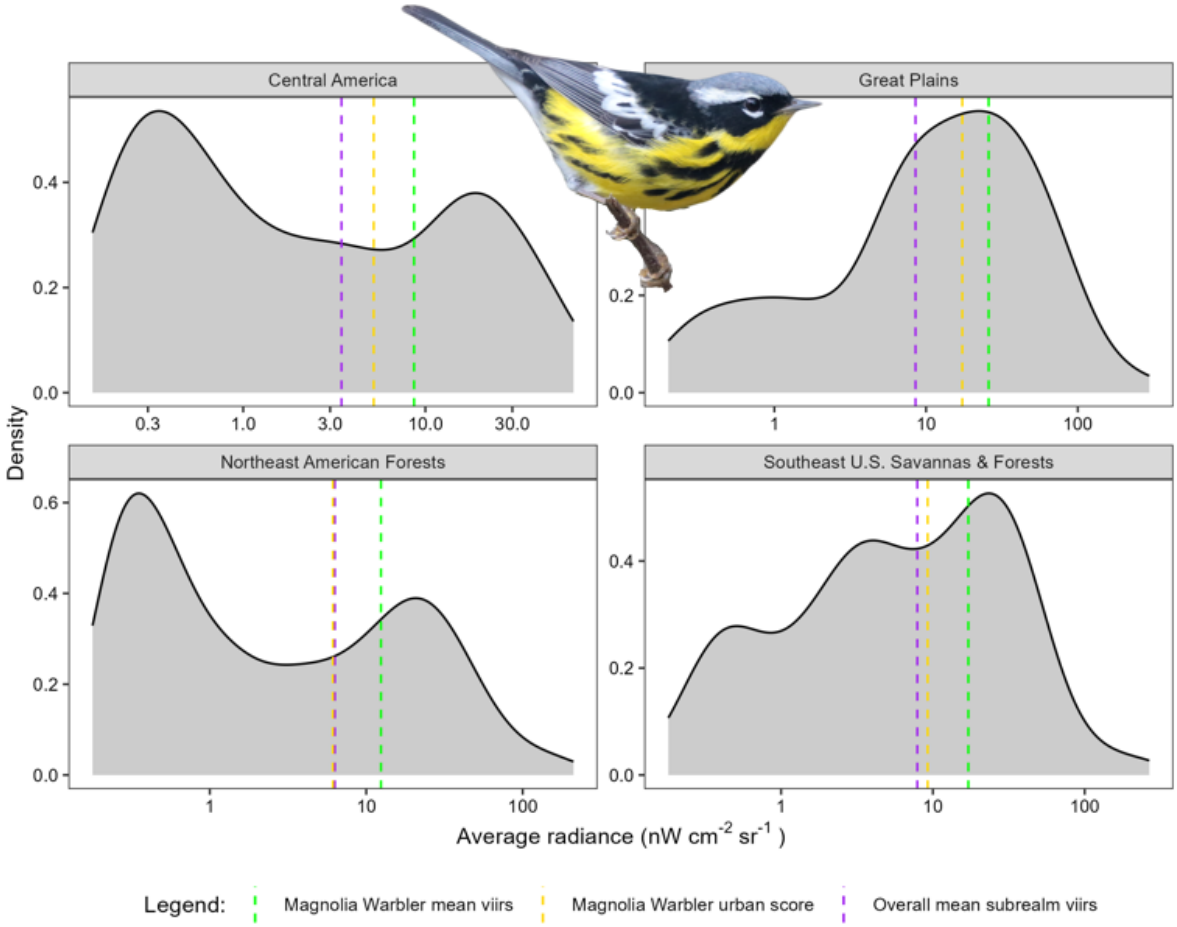

Fig. S4: The distribution of Magnolia warbler *Setophaga magnolia* observations and the VIIRS values, in average radiance, of those observations in four different subrealms. Shown using dashed lines are the species-specific mean (green), the subrealm-specific mean of VIIRS (violet), and the resulting urban tolerance measure (yellow).

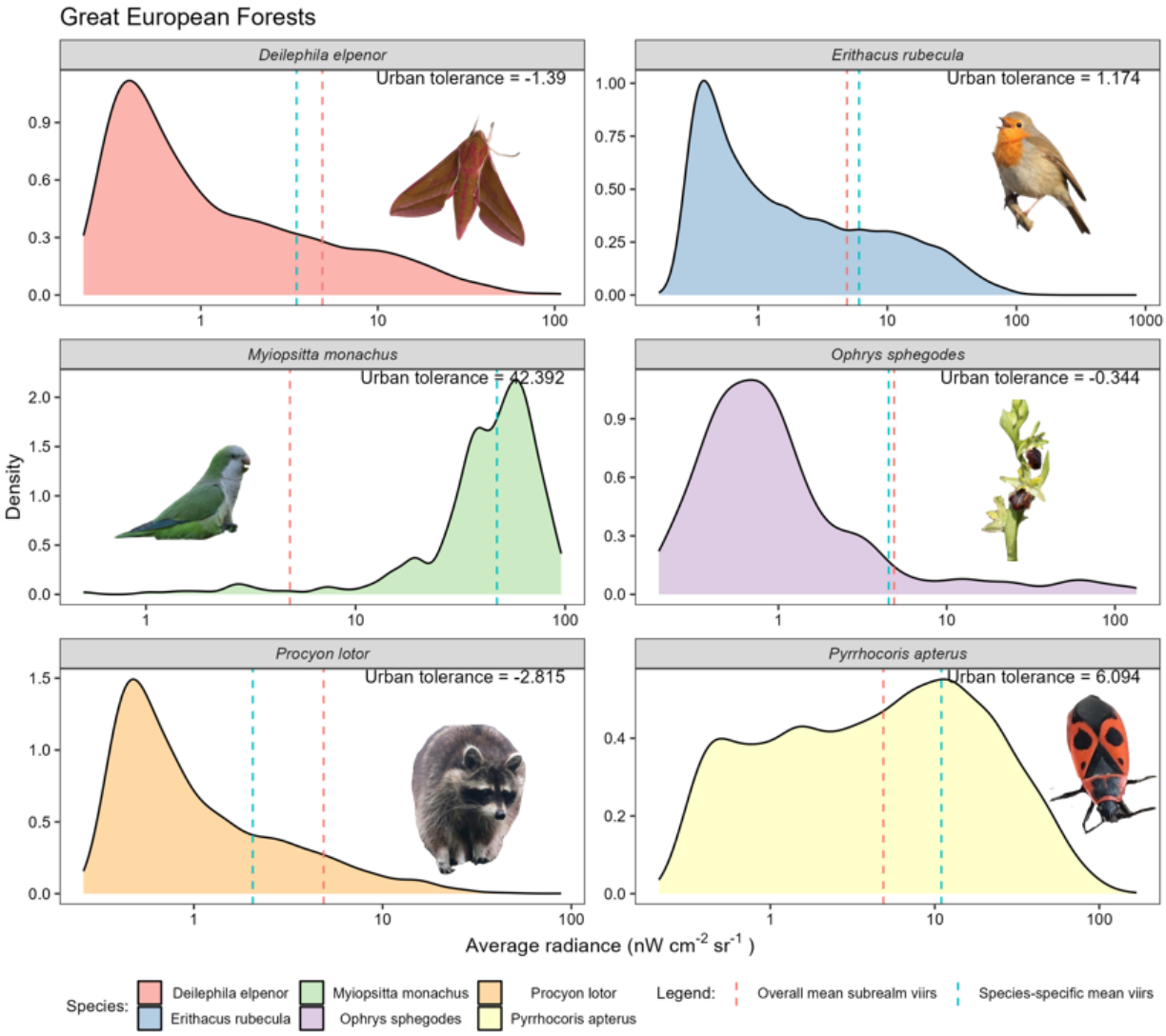

Fig. S5: An illustration of the species-specific distribution of observations and the VIIRS values, in average radiance, of those observations for six different species within the Great European Forests subrealm. Of note is that the urban tolerance is not shown, as some are negative values, but these values are shown in the text.

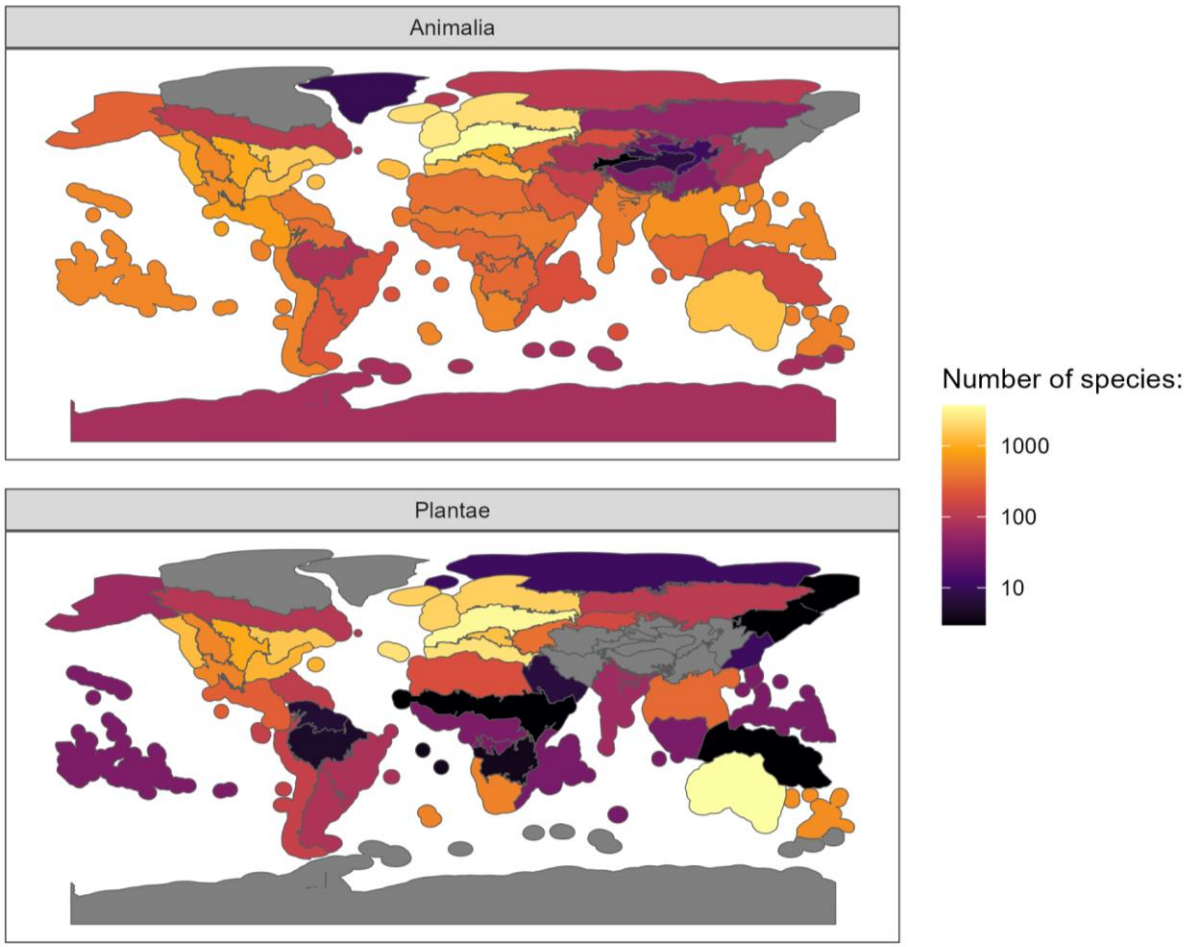

Fig. S6: The total number of species for which we calculated urban tolerance scores, stratified by subrealm and shown separately for Animalia (top) and Plantae (bottom). The subrealms with their names are shown in Figure S1. The dataset of species' urban tolerance scores is provided in Table S1.

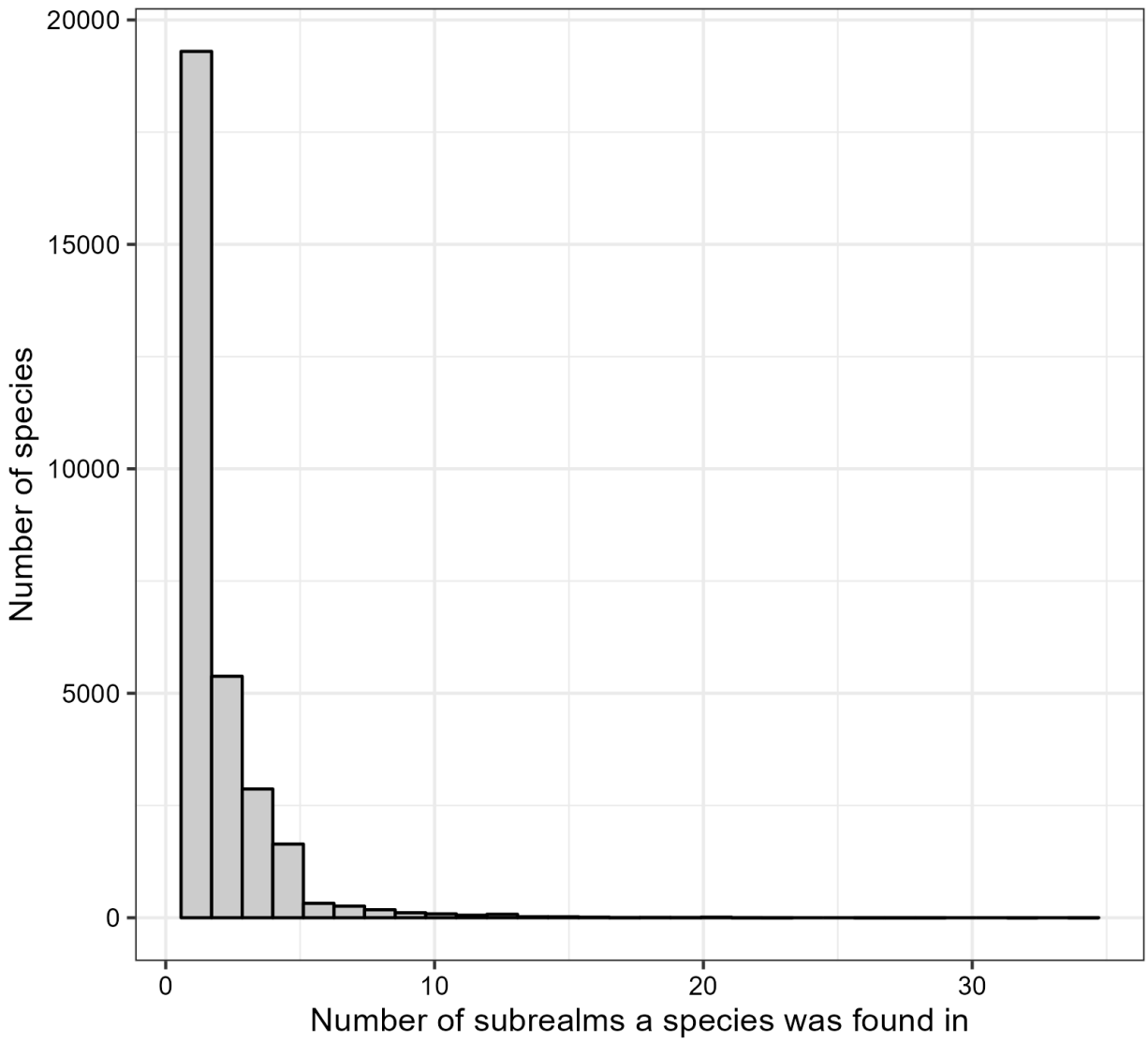

Fig. S7: The total number of subrealms for which a species had an urban tolerance score. The majority of species (64%) only had an urban tolerance score from one subrealm, but the range was from 1 to 34.

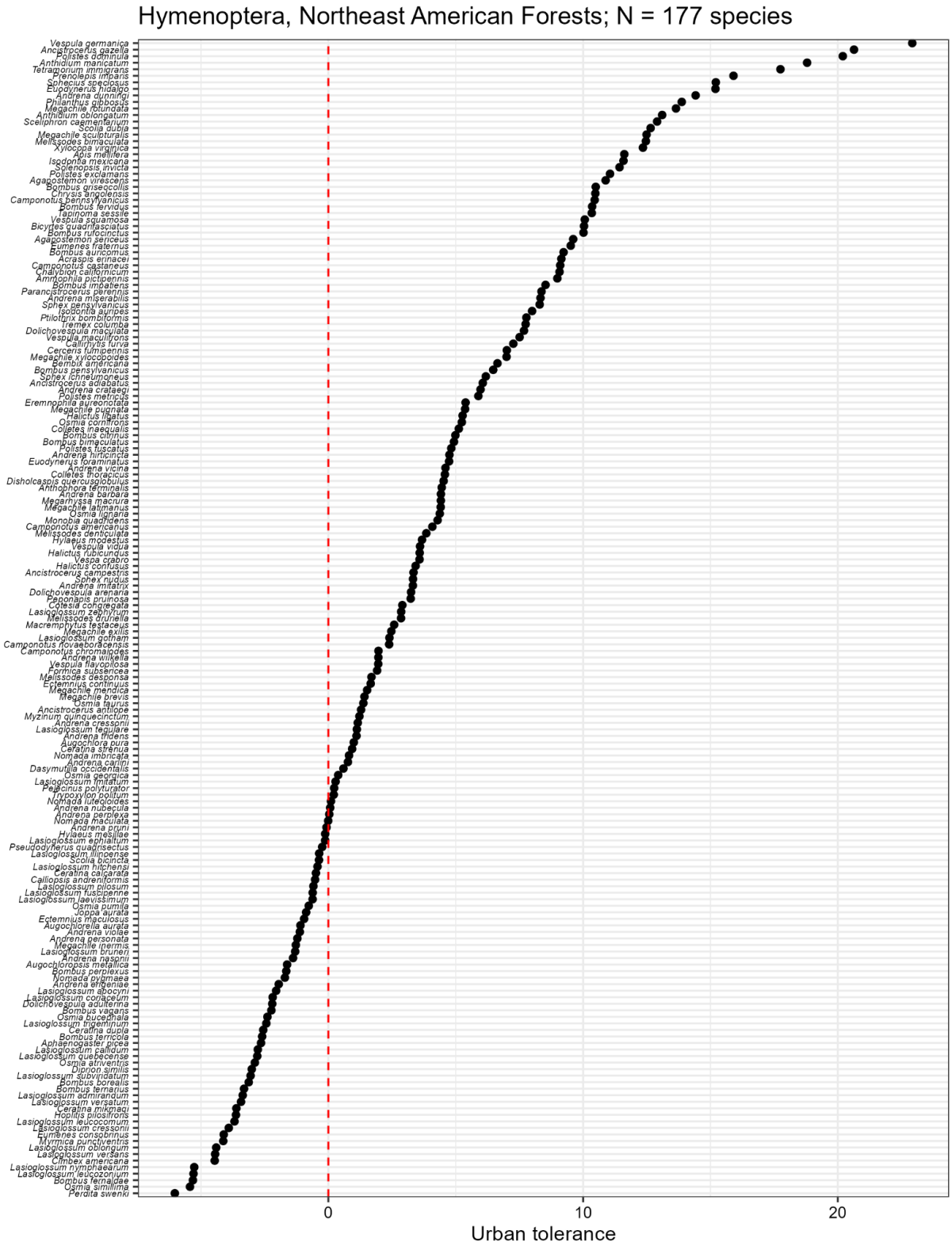

Fig. S8: An illustrative example, showing the relative urban tolerance scores for Hymenoptera in the Northeast American Forests subrealm. The plot is for illustrative purposes and the values for each species can be found in Data S1.

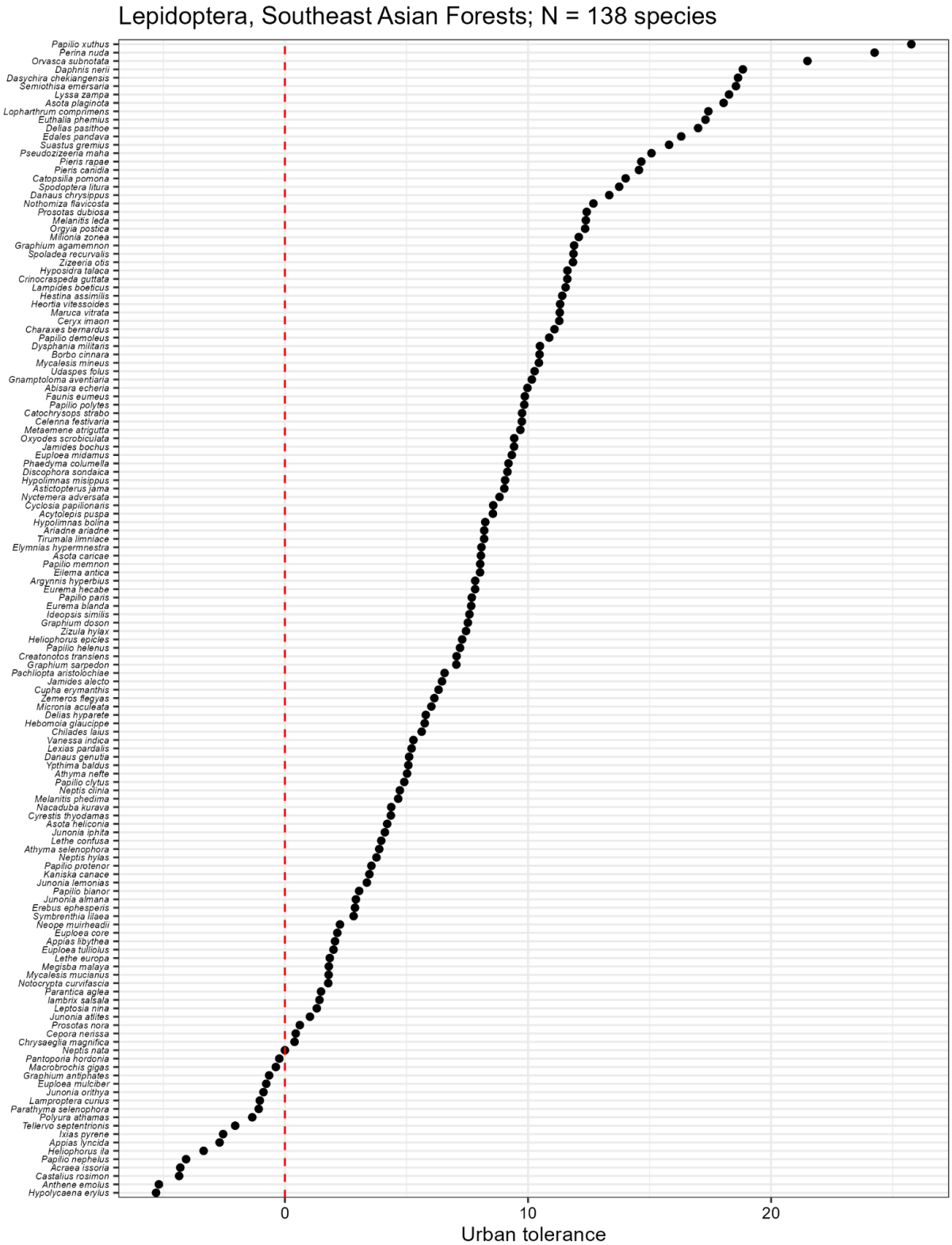

Fig. S9: An illustrative example, showing the relative urban tolerance scores for Lepidoptera in the Southeast Asian Forests subrealm. The plot is for illustrative purposes and the values for each species can be found in Data S1.

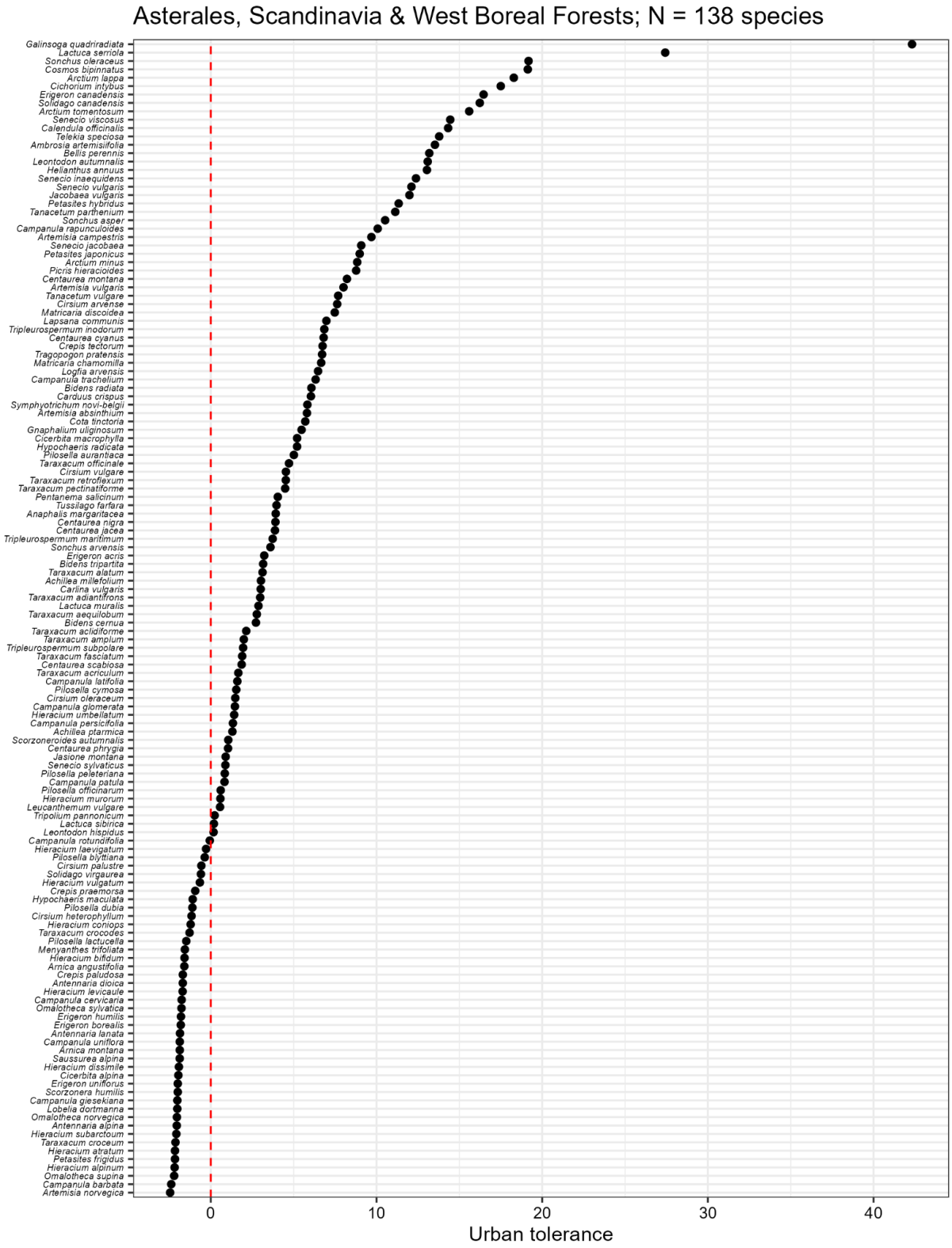

Fig. S10: An illustrative example, showing the relative urban tolerance scores for Asterales in the Scandinavia & West Boreal Forests subrealm. The plot is for illustrative purposes and the values for each species can be found in Data S1.

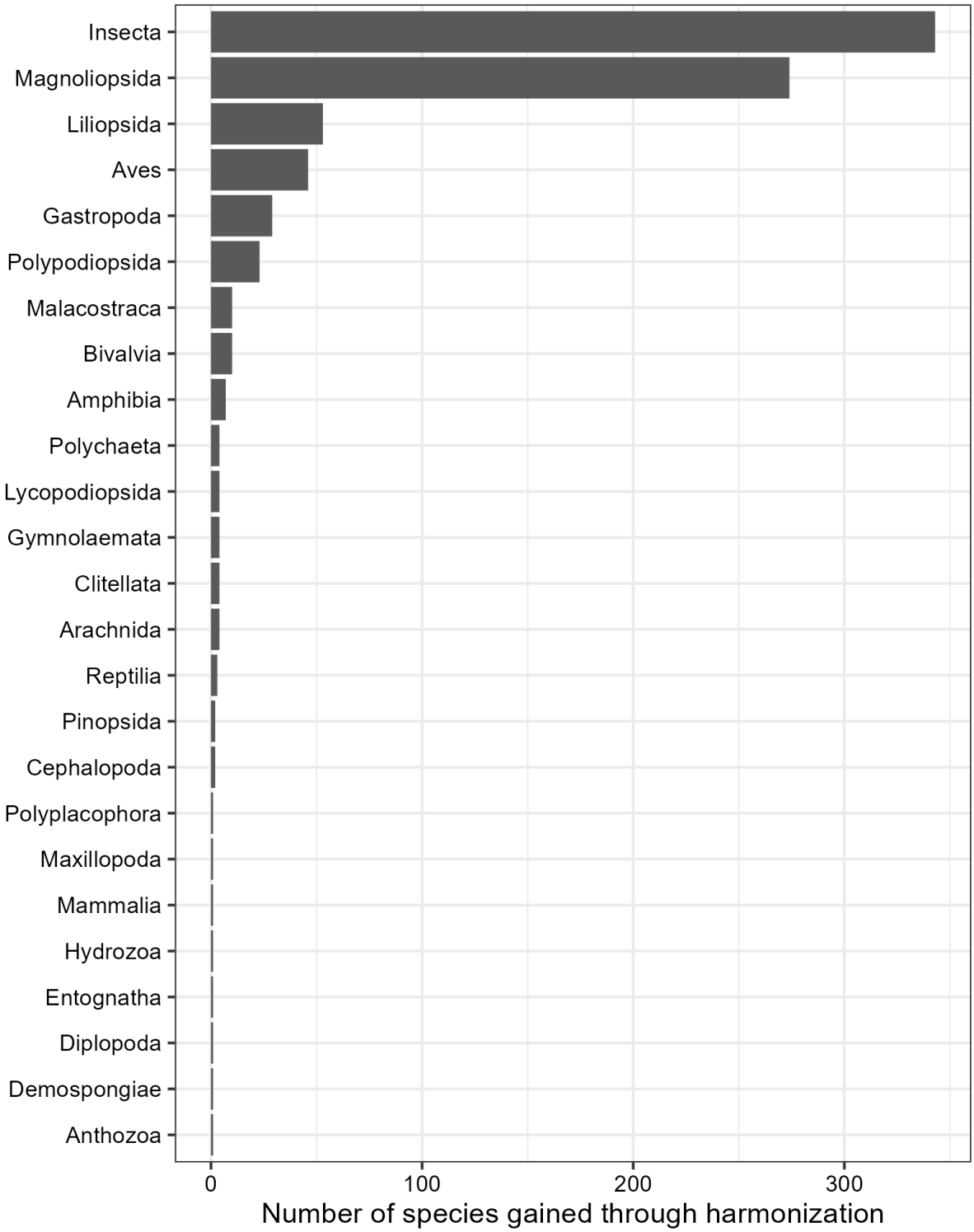

Fig. S11: A summary of the number of species, shown per class, gained through taxonomic harmonization, and therefore included in the analysis dataset.

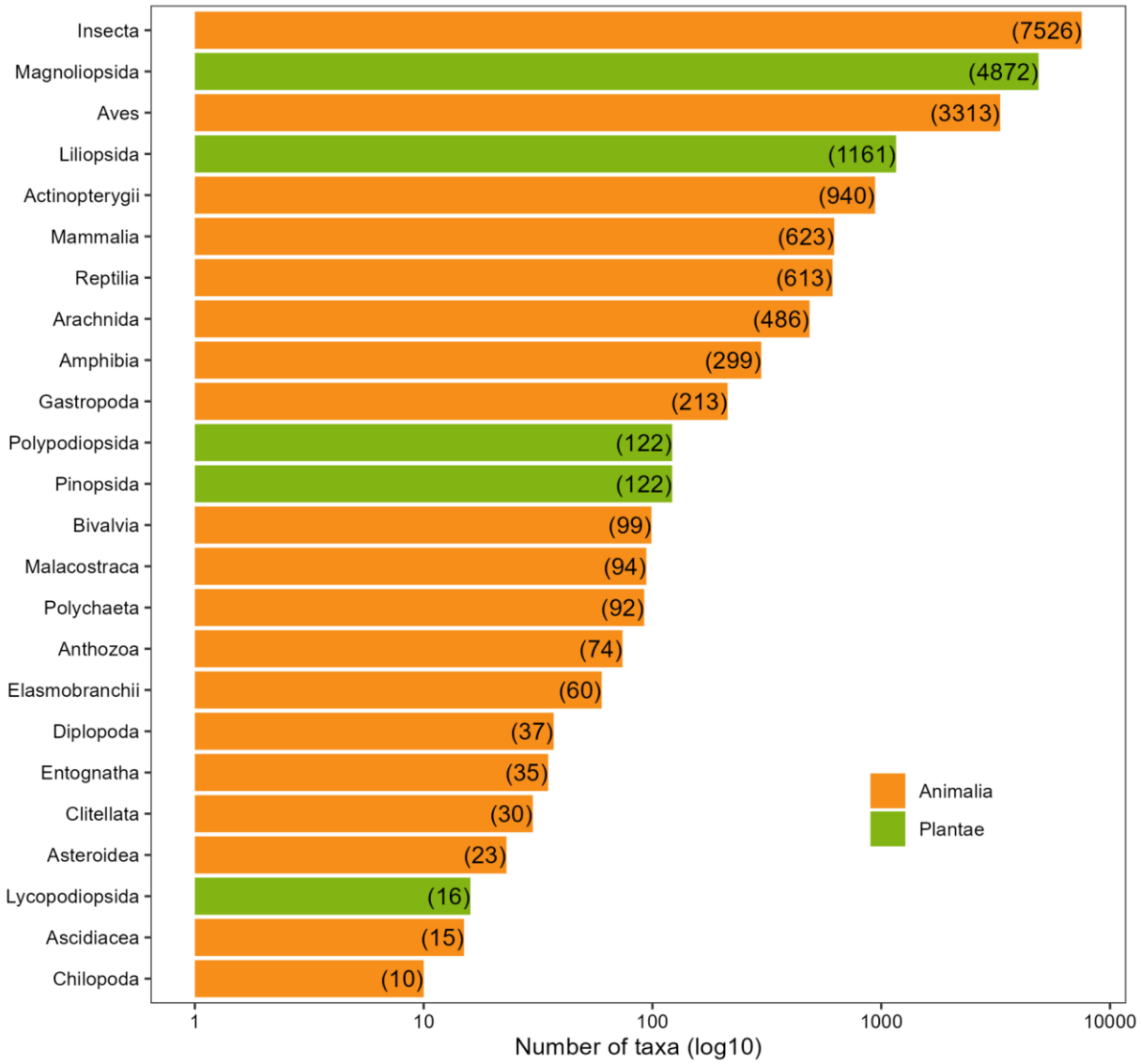

Fig. S12: The number of taxa, on a log10-transformed scale. These can be found in Table S3.

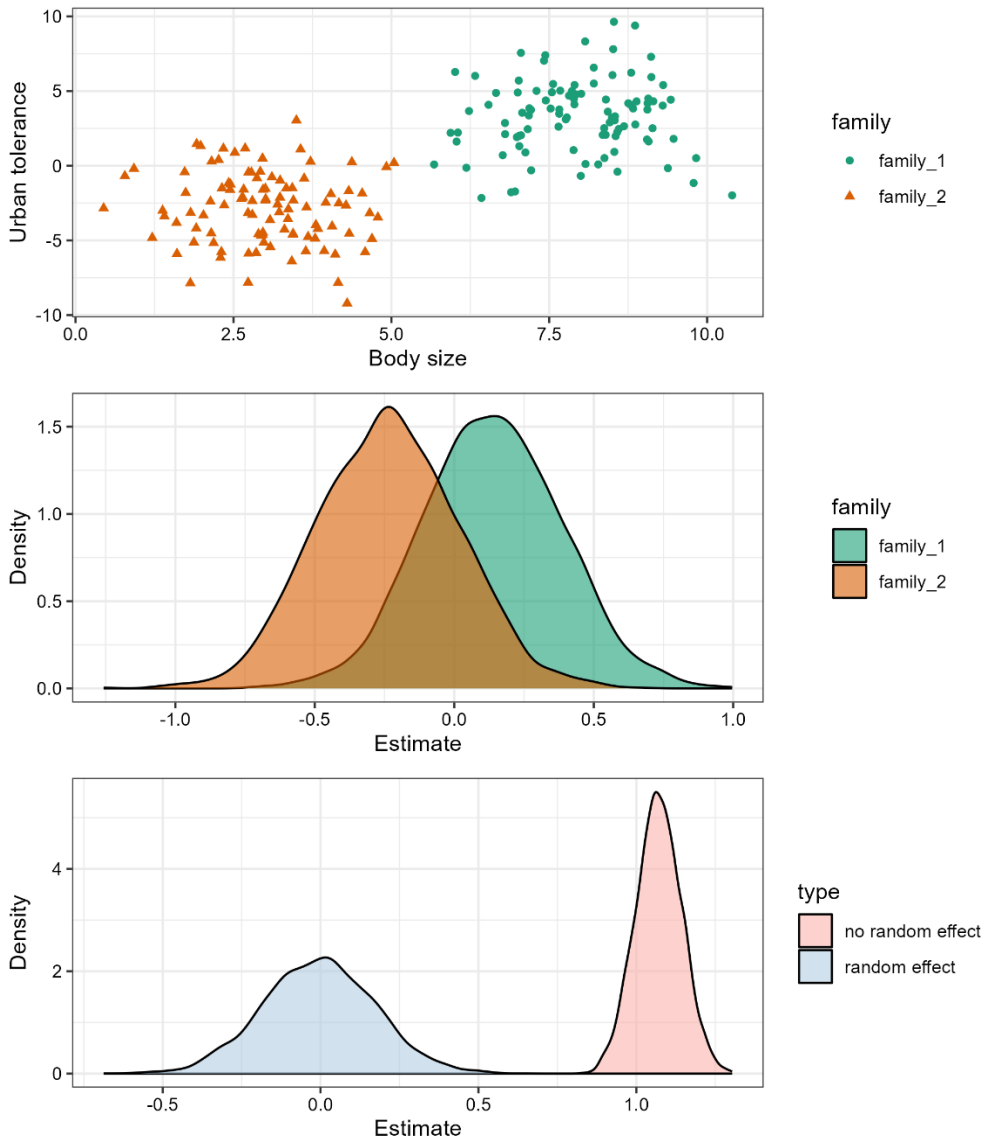

Fig. S13: We chose to model each taxonomic group (i.e., family, order, class, phylum, and kingdom) independently of one another to avoid the influence of where individual families could have no effect but result in a positive effect if modeled jointly. The top panel shows simulated raw data for two families; the middle panel shows the posterior distribution of a bayesian model fit for each of these families separately; the bottom panel shows the posterior distribution of a bayesian model fit for a model including family as a random effect and for a model not including family as a random effect.

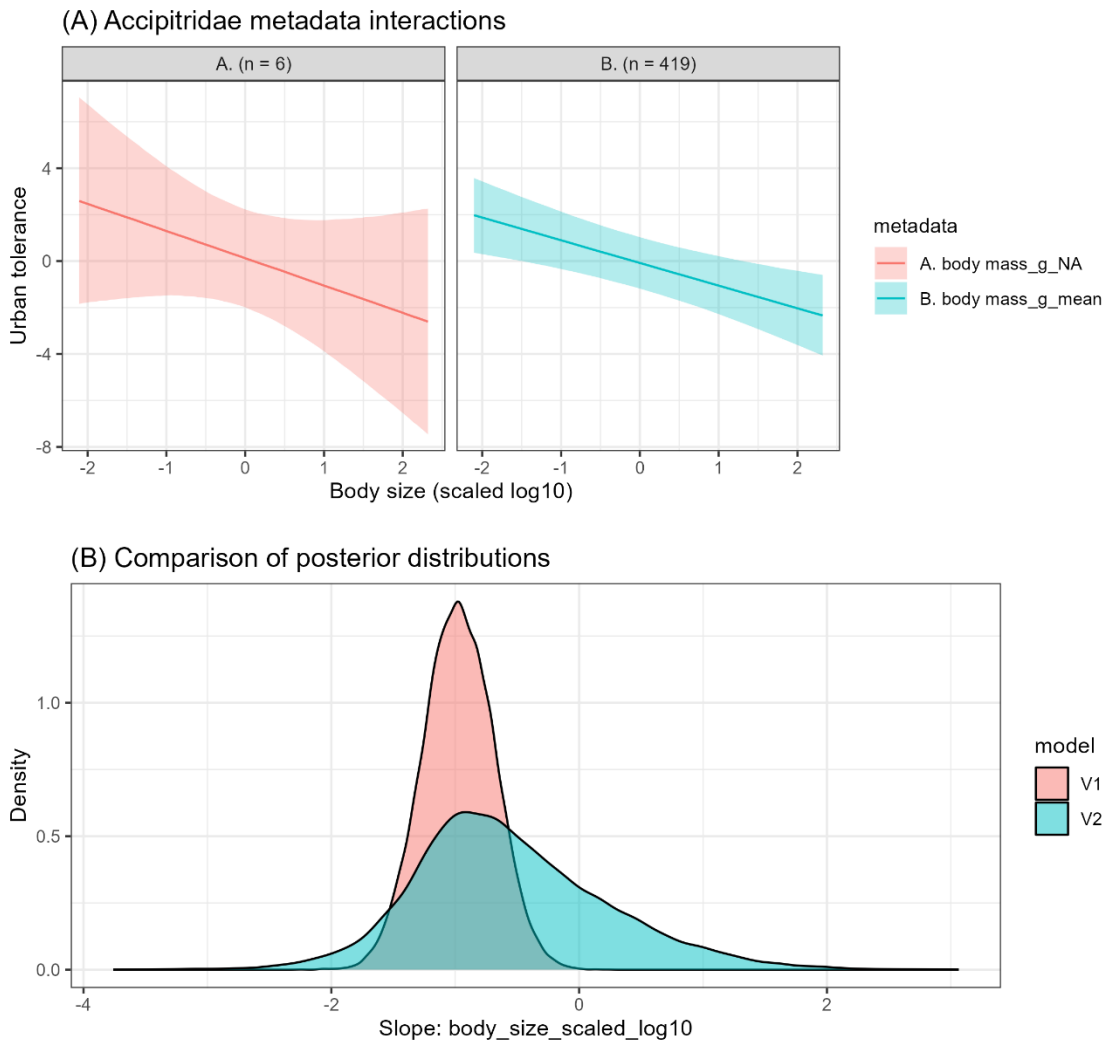

Fig. S14: An illustrative example showing the influence of using a random slope for body size, by showing the (A) Interaction between body size (scaled and log10 transformed) and body size measurement type (metadata; see Methods for details) for the family Accipitridae. Facets represent different types of body size metrics used in the dataset, labeled with a letter identifier and corresponding sample size (n). Lines represent predicted urban tolerance based on body size, with 95% credible intervals shown as shaded ribbons. These are extracted from a model fit with an interaction between metadata and body size to purposefully investigate the influence of metadata and the relationship with body size. (B) Posterior distributions of the slope for body size (scaled and log10 transformed) under two model structures: V1 includes a random slope for body size by subrealm and a random intercept for metadata (as presented in main results), while V2 adds a random slope for body size by metadata. The inclusion of this additional random slope in V2 increases uncertainty and pulls slope estimates toward zero, particularly when data are sparse within body size measurement types as illustrated here. Also see Fig. S15.

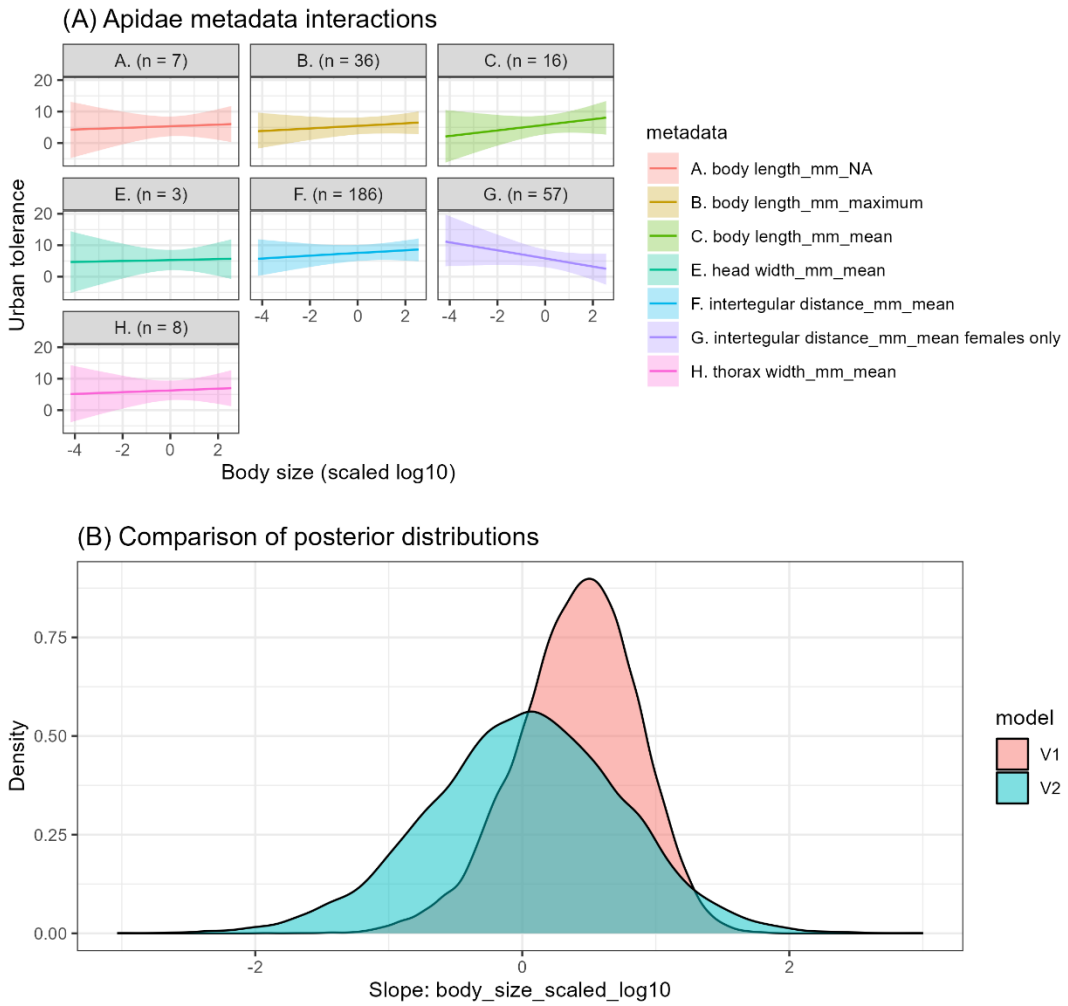

Fig. S15: An illustrative example showing the influence of using a random slope for body size, by showing the (A) Interaction between body size (scaled and log10 transformed) and body size measurement type (metadata; see Methods for details) for the family Apidae. Facets represent different types of body size metrics used in the dataset, labeled with a letter identifier and corresponding sample size (n). Lines represent predicted urban tolerance based on body size, with 95% credible intervals shown as shaded ribbons. These are extracted from a model fit with an interaction between metadata and body size to purposefully investigate the influence of metadata and the relationship with body size. (B) Posterior distributions of the slope for body size (scaled and log10 transformed) under two model structures: V1 includes a random slope for body size by subrealm and a random intercept for metadata (as presented in main results), while V2 adds a random slope for body size by metadata. The inclusion of this additional random slope in V2 increases uncertainty and pulls slope estimates toward zero, particularly when data are sparse within body size measurement types as illustrated here. Also see Fig. S14.

Table S1: A list of 41 potential 'types' of body size that were used for potential inclusion in our body size dataset. We aimed to incorporate as many types of body size measures as possible and were not restrictive in our searching for body size measures.

Type of body size			
Weber's length	abdominal length	biovolume	body length
body mass	body mass female only	carapace width	cephalothorax width
colony height	column diameter	coral diameter	diameter
dorsal mantle length	dry mass	elytron length	femur length
forewing length	fork length	head height	head length
head width	height	hind tibia length	hindwing length
intertegular distance	longest length	plant height	pronotal width
radius	shell size	shell volume	snout vent length
standard length	thorax length	thorax width	total biomass
total length	wet mass	width	wing length
wingspan			

Data S1: (separate file) The urban tolerance values (N=56,181 total values (i.e., unique urban tolerance score in a subrealm)) for potential inclusion in our analysis. This table is uploaded as a separate CSV file.

Data S2: (separate file) A final dataset for potential analysis and modeling, including a total of 94,087 observations (unique combination of species urban tolerance values, subrealm, and body size measure) of 20,957 species that had at least one measure of urban tolerance and at least one measure of body size. This table is uploaded as a separate CSV file. Note, however, that not every body size measure is made available due to a lack of permissions to share some datasets.

Data S3: (separate file) A list of the final potential 'datasets' that were used to aggregate measures of body size. Metadata refer to our naming scheme employed in our workflow; citation is a descriptor of either the paper citation or dataset citation or a descriptor of manual data aggregated by us; URL is the potential URL if applicable, and number of data points is the number of potential data points that could be used for the analysis. This table is uploaded as a separate CSV file.