How has the environment shaped geographical patterns of insect body sizes? A test of hypotheses using sphingid moths

Nicolas Beerli¹, Florian Bärtschi¹,², Liliana Ballesteros-Mejía³, Ian. J. Kitching⁴ & Jan Beck¹,⁵*

¹) University of Basel, Department of Environmental Science, Basel, Switzerland
²) Umweltbüro Bärtschi, Basel, Switzerland
³) Muséum national d'Histoire naturelle, Sorbonne Université, Institut de Systématique, Évolution, Biodiversité (ISYEB). UMR 7205 - CNRS, MNHN, UPCM, EPHE, Paris, France
⁴) Department of Life Sciences, Natural History Museum London, London, UK
⁵) University of Colorado, Museum of Natural History, Boulder, USA

*) Correspondence: Jan Beck, University of Colorado, Museum of Natural History, 265 UCB, Boulder, CO 80309, USA. E-mail: jan.beck@colorado.edu

Manuscript for Journal of Biogeography, Humboldt special issue

Research paper

Short title: Geographical patterns of moth body sizes

Keywords: Hawkmoths; Lepidoptera; Bergmann's rule; comparative; ectotherms; phylogeny;

Word count: Abstract 300 words; main text 5404; 4 figures, no tables; 56 references;

Acknowledgements

We thank Walter Jetz for fruitful discussions on the subject, and Meret Hornstein, Katherine Winkel, Jonas Egli and Gunnar Brehm for constructive comments on earlier drafts. Blanca Huertas supported our measuring of museum specimens. We received a travel grant from the Synthesys program of the European Community Research Infrastructure Action (FP7 "Capacities" program). This study was funded in part by NERC grant number NE/P003915/1 to IJK, the modelling of species ranges by Swiss National Science Foundation grant 3100AO_119879 to JB.

Biosketch

Nicolas Beerli completed his MSc in the research team of JB, which focusses on insect macroecology and the environmental impacts on biodiversity. Author contributions: NB and JB designed the study and analysed data, supported by FB; JB, NB, LB-M and IJK provided sphingid moth data; NB, JB and IJK wrote the manuscript, with input from all authors.
ABSTRACT

Aim: We mapped the geographical pattern of body sizes in sphingid moths and investigated latitudinal clines. We tested hypotheses concerning their possible environmental control, i.e., effects of temperature (negative: temperature size rule or Bergmann’s rule; positive: converse Bergmann rule), food availability, agility in densely vegetated habitats, robustness to starvation during extreme weather, and seasonality.

Location: Old World and Australia/Pacific region

Methods: Body size data of 950 sphingid species were compiled and related to their distribution maps. Focusing on body length, we mapped the median and maximum size of all species occurring in 100 km grid cells. In a comparative approach we tested the predictions from explanatory hypotheses by correlating species’ size to the average environmental conditions encountered throughout their range, under univariate and multivariate models. We accounted for phylogeny by stepwise inclusion of phylogenetically informed taxonomic classifications into hierarchical random-intercept mixed models.

Results: Median body sizes showed a distinctive geographical pattern, with large species in the Middle East and the Asian tropics, and smaller species in temperate regions and the Afrotropics. Absolute latitude explained very little body size variation, but there was a latitudinal cline of maximum size. Species’ median size was correlated to net primary productivity, supporting the food availability hypothesis, whereas support for other hypotheses was weak. Environmental correlations contributed much less (i.e., <10%) to explaining overall size variation than phylogeny (inclusion of which led to models explaining >70% of variability).

Main conclusion: The intuitive impression of larger species in the tropics is shaped by larger size maxima. Median body sizes are only very weakly related to latitude. Most of the geographic variation in body size in sphingid moths is explained by their phylogenetic past. NPP and forest cover correlate positively with the body size, which supports the idea that food availability allowed the evolution of larger sizes.
INTRODUCTION

Tropical insects of impressively large body size compared to their temperate counterparts are a common observation of field naturalists, as well as those admiring museum collections – starting with A. v. Humboldt’s scientific explorations of the Neotropics around 1800, on which he collected insects of appreciable size (Barragán et al., 2009). 50 years later his successor, Amazonian explorer H.W. Bates (1864, p.115) reported accidentally shooting hawkmoths so large he mistook them for small birds. However, it remains unclear whether such observations necessarily imply a larger size for the average tropical insect species – it just may be that their maxima are more extreme in regions of higher species richness, and it is these that capture our attention (cf. coloration in tropical birds; Bailey, 1978). Furthermore, it must be expected that body sizes, like species richness or other trait variation, exhibit more complex geographical patterns than simply a latitudinal gradient (Hawkins & Diniz-Filho, 2004). Here, using sphingid moths (hawkmoths) as the exemplar, we map for the first time the body size distribution of a group of insects at near global spatial extent intercontinental geographic scale (i.e., global excluding only the Americas) and at high spatial resolution. In a comparative approach we investigate how environmental conditions and phylogenetic inertia may have shaped these patterns.

Body size is a prime example of geographic trait variation, with important links to physiological functioning, population-level processes and biodiversity as a whole (Gaston & Blackburn, 2000; Blanckenhorn, 2000; Brown et al., 2004). Consequently, its relationships to environmental variables are of major interest (Gaston & Chown, 2013), yet conclusive, large-scale studies, intra- or interspecific, have mostly been carried out on endotherms (Blackburn & Hawkins, 2004; Freckleton et al., 2003; Olson et al., 2009). Environmental gradient studies on ectotherm body sizes are scarcer, less firmly rooted in theory, and provide more conflicting results at an intraspecific (Ashton et al., 2003; Puzin et al., 2014) as well as an interspecific level (Entling et al., 2010; Fattorini et al., 2013; Slavenko & Meiri, 2015; Zeuss et al., 2017; Brehm et al., 2019; Pallares et al., 2019).

The formulation of eco-geographical rules has been the cornerstone of attempts to describe patterns and understand mechanisms of trait variation, and the most well-known of these, Bergmann’s rule (Meiri, 2011) concerns body size. Originally primarily addressing endothermic vertebrates restricted to mammals, it proposes an interspecific body size increase towards higher latitudes if restricted to closely related taxa. While the pattern itself is well-supported (Meiri & Dayan, 2003), fewer studies have tested and supported the underlying thermoregulatory mechanism (Blanckenhorn & Demont, 2004; Watt et al., 2010; Meiri, 2011; Fattorini et al., 2013; Beck et al., 2016). Other mechanistic, environmental effects on body size, both evolutionary and developmental, have been suggested for insects and other ectotherms (Chown & Gaston, 2010, 2013; Chown & Nicolson, 2013), and these may not act in a mutually exclusive manner. Furthermore, body sizes are strongly affected by phylogeny, which complicates analyses and the inference on environmental causes of the observed variation (Freckleton & Jetz, 2009).

Using data on size, high-resolution maps of species’ geographical ranges, and a phylogenetically-informed taxonomic classification, we mapped the geographical body size pattern of sphingid moths and tested its consistency with the following hypotheses while accounting for phylogenetic effects.

(1) Bergmann’s rule expects a negative relationship between body size and environmental temperature. It assumes that thermoregulation, through the ratio of heat-exchanging surface and heat-producing volume, favors larger animals in colder climates. Although it was originally focused on endotherms, Zamora-Camacho et al.
(2014) argued for a Bergmann-type heat preservation mechanism in an ectotherm lizard. Some studies have also investigated Bergmann’s rule in Lepidoptera, with mixed results (e.g., Beck et al., 2016; Brehm et al., 2019). Furthermore, adult sphingid moths are known for pre-flight thoracic muscle temperature regulation (Heinrich 1993), hence temperature efficiency may be not be entirely irrelevant.

Alternatively, an intraspecific effect known as the ‘temperature size rule’ (TSR; Kingsolver & Huey, 2008; Chown & Gaston, 2010) predicts the same pattern. Laboratory experiments with many ectotherms have shown that individuals develop slower yet mature into a larger body size under colder temperatures (Atkinson, 1994). This may be either an effect of adaptive plasticity or it may be due to non-adaptive constraints of temperature on growth rates (e.g., the discrepancy between oxygen supply and demands; see Makarieva et al., 2005 for an interspecific approach). The TSR has also been shown to apply to a sphingid moth from North America (Manduca sexta; Davidowitz & Nijhout, 2004).

(2) The ‘converse Bergmann pattern’ expects a positive relationship of body size and temperature. The reasoning here is that lower temperatures lead to lower growth rates in ectotherms, favoring species that become adult (i.e., reproductive) at smaller size (Meiri, 2011; Shelomi, 2012).

(3) The resource availability hypothesis postulates an increase of body size with increased food availability (Blackburn et al., 1999; Watt et al., 2010). Intraspecifically, lower food availability is known to result in smaller specimens (Slansky & Scriber, 1985). If this mechanism was relevant interspecifically, we can predict a body size increase with net primary productivity, which we use as a proxy for food availability in our herbivorous study taxon.

(4) The agility hypothesis postulates that smaller species have an advantage in agility in dense, forest-covered habitats. It predicts that body size decreases with increasing forest cover. Supporting data are so far restricted to large mammals (Bro-Jørgensen, 2008), where larger species are found in open habitats than in forests. Adult sphingids are known for their very fast, nocturnal flight that could be obstructed by dense vegetation, hence we considered the idea here.

(5) The starvation hypothesis postulates that larger species occur in regions with a higher risk of unforeseeable starvation. With increasing body size fat storage increases faster than metabolic rate, hence survival of such catastrophic events becomes more likely (Chown & Gaston, 2010). This hypothesis assumes that insect species in variable climates need more physiological tolerances than those in zones of stable climate (Addo-Bediako et al., 2000). We predict that body size increases with increasing temperature extremes during their active period.

(6) The seasonality hypothesis (Mousseau, 1997) assumes that in seasonal habitats, where a part of the year is unsuitable for growth, smaller species profit from a faster completion of their reproductive cycle. We expect a negative link of seasonality and body size.

Phylogeny plays an important role in the current spatial distribution of species’ traits, due to prior adaptation to past environmental conditions and the constraints of evolutionary history (Gaston & Chown, 2013). It can be assumed that a part of the variation in body size is explained by phylogenetic inertia. For interspecific comparisons, such phylogenetic signals within the data must be accounted for to avoid spurious effects on the analysis of environmental predictor variables (Freckleton & Jetz, 2009, and references therein).

Here, we first map geographical patterns of size in sphingid moths at large geographical scale by combining detailed distribution data and body size measurements. We test-explored whether hawkmoths are larger or smaller towards the equator using a
simple latitudinal cline. We then tested the above hypotheses by investigating the link between body size and environmental variables, such as temperature, seasonality, forest cover, temperature extremes and net primary productivity (NPP). We used a comparative, ‘species-focused’ approach where every species counts equally (Chown & Gaston, 2010), which is evolutionarily more informative than a ‘geography-focused’ approach (i.e., comparing grid cells). The latter would be overly impacted by widespread species that occur in many cells (Jetz & Rahbek, 2002). Phylogenetic information was included into the analyses at different taxonomic classification levels, which helped to identify the importance of phylogenetic history in comparison to environmental effects.

**METHODS**

Hawkmoths or sphingids are a family of Lepidoptera that have attracted the attention of professional as well as amateur insect collectors for centuries (Kitching & Cadiou, 2000). Consequently, they are more extensively studied, taxonomically and biogeographically, than most other insects. Recently, Ballesteros-Mejia et al. (2017) published and analyzed detailed geographic ranges for all species occurring outside the Americas (data available at www.mol.org). These maps stem from expert-edited species-distribution modelling (relating published and unpublished occurrence records to climate and vegetation variables, at 5 x 5 km resolution). Details on the procedures of modelling and validation of geographic range maps are found in Ballesteros-Mejia et al. (2017). Of 981 species occurring in the research region, we considered 950 species for the present analysis. The reasons for excluding some taxa were phylogenetic uncertainty (eleven species represented taxa of unreliable species status), missing environmental data (eight small-island endemics of the far eastern Pacific, longitude beyond 180°E) or because no male specimens (or pictures) were available for body size measurement (see below for a further reduction to 938 species for body length measures) (14 species).

Theoretically, body mass is the physiologically most meaningful size metric (Gaston & Blackburn, 2000), but it is exceedingly difficult to measure when relying on dried museum specimens or images. Many length measurements have been shown to correlate with body mass and such linear measurements are commonly used as a proxy for body size in insects (Chown & Gaston, 2010; Brehm et al., 2019). We compiled data on body length (head to tip of abdomen), thorax width, and forewing length (base to tip). Measurements were taken from scaled pictures published in d’Abrera ([1987]; 471-459 species), from pictures that are publicly available (28 species), and unpublished pictures-colour photographs (38 species, mostly from IJK’s personal photographic print collection). For the The remaining 422 species, which were not available in scaled illustrations, were measured from pinned specimens in the collection of the Natural History Museum, London. Details of the measured specimens can be found in Appendix ES1. All measurements were taken with a digital caliper, rounded to 0.1 mm. Since there is sexual size dimorphism in some sphingid groups (e.g., tribes Smerinthini, Ambulycini, where males tend to be smaller than females; Kitching & Cadiou, 2000) we measured and compared only males (the more common sex in collections) for consistency. In other Lepidoptera taxa, size measures for males and females are highly correlated (data from Brehm et al., 2019, see there also for further discussion). Furthermore, we would note that any resulting size underestimates within those clades’ size would be controlled for by fitted random effects models for phylogenetic association (in mixed models, see below).

As in other macro-studies (e.g., Zeuss et al., 2017) we were only able to measure one specimen per species (often only one picture or specimen was available).
Although body size can be variable within species, such a sample (n = 1) is an unbiased estimate of the mean. While a single-specimen measure would be unsuitable for intraspecific studies, it should yield near-identical patterns as to averages based on many specimens if studied across many species. This is because randomly-distributed errors will even out, and correlation analyses will provide reliable results. We confirmed this theoretical expectation with subsampling simulations based on moth body size data from Brehm et al. (2019; J. Beck & G. Brehm, unpubl.). Furthermore, we assessed the size variability in eight abundant sphingid species, where the body lengths of 208 specimens (14–34 per species) were measured from scans of collection drawers (Johnson et al., 2013; Trueman & Yeates, 2015; data in ES1). The absolute difference between mean body lengths of these specimens and the (independent) single-specimen measure was on average 3.4 mm, which we find small when considering an interspecific body length range of 55.8 mm (from 12.5 to 68.3 mm) in our dataset. Thus, while intraspecific effects may not be entirely negligible, they are probably relatively small irrelevant in comparison to interspecific variability in a dataset with many hundreds of species. Over our large sample of 93850 species they may have just added some additional random noise. Furthermore, intraspecific variability in Lepidoptera often appears large due to occasional ‘dwarf specimens’, but these do not tend to be illustrated in pictures whenever ‘normal’ specimens are available. Measuring tens of thousands of specimens to fully appreciate species’ intraspecific variation was simply beyond the scope of this research project. Data for the size measures for all species in analysis are available in ES1.

Environmental predictors

To test our hypotheses on the environmental control of body size, we considered the mean annual temperature of frost-free months, net primary production, forest cover, extreme temperatures and seasonality. Additionally, absolute latitude was used as a non-environmental predictor.

- **Mean annual temperature** was calculated for those months with an average monthly temperature above 0°C. Data were based on monthly temperature data from the WorldClim database (resolution 30 arc-second, time period 1950–2000; Hijmans et al., 2005). Excluding cold winter months is reasonable because sub-zero winter temperatures are quite irrelevant for diapausing organisms while including them into averages would bias relevant temperatures downward.

- **Net primary production** (NPP) data was taken from remotely sensed normalized differential vegetation index values (NDVI), corrected for modelled periods of water limitation (Running et al., 2004; time period 2000–2014). In this dataset ‘No Data’ values caused by lack of green reflection were set to zero. In this dataset ‘No Data’ values caused by lack of green reflection were set to zero where they occurred on land, as vegetation-free landscapes (e.g., deserts, glaciers) must be interpreted as unproductive for our purposes, rather than being excluded from analysis.

- **Forest cover** (in percent) was based on Tuanmu & Jetz (2014), who provided 12 generalized land cover maps based on a consensus land cover dataset. Four of these land cover categories were forests (evergreen deciduous needleleaf trees; evergreen broadleaf trees; deciduous broadleaf trees, mixed other trees). Percentages from these four forest types were summed (and corrected to 100% where this value was exceeded due to integer rounding).

- **Temperature extremes** data were based on temperature data from Smith et al. (2008) at a resolution of 5°. For every grid cell, data included monthly temperature deviation of a long-term average (128 years: 1880–2008). In the northern hemisphere, we...
excluded data north of 35°N for November, December, January and February to account for diapause; analogously we excluded data south of 35°S for May, June, July and August in the southern hemisphere. We then counted within each grid cell the months with temperature values that were below one standard deviation (SD) from the grid cell mean (referring to SD of all grid cells in analysis). The number of months was divided by the estimated length of the hawkmoths’ activity period (nine months above 35° S/N and 12 in the other areas) to render them comparable. This provided a metric of extreme negative temperature events (i.e., months
  temp. < mean − SD months of activity per year in 128 years of data). Finally, the data were interpolated to a finer resolution using inverse distance weighting (search radius 6 points), to make data comparable with other environmental data. The higher our metric, the more months with extremely low temperatures occurred.

- Temperature seasonality data were used from the WorldClim database (Hijmans et al., 2005). It is the standard deviation of the annual temperature seasonality, multiplied by 100.

Predictor data were re-projected to a Mollweide World equal area projection (bilinear resampling) to match moth data. Pixels in oceans, seas or lakes >1 km² were excluded from all raster datasets. We carried out the handling and extraction of spatial data in ArcGIS 10.3 (http://desktop.arcgis.com), subsequent data compilation and all further analyses were done in R 3.3 (https://cran.r-project.org/; packages ade4, ape, caper, data.table, ecodist, lme4, matrix, nlme, xtable).

We used two different approaches to map interspecific body size clines. In an assemblage-based (Chown & Gaston, 2010) (or grid cell-based) approach we mapped the median body length of all species occurring in 100 x 100 km grid cells. Grid-cells containing ≤5 species were removed to reduce random noise. We measured the variability of body length with the interquartile range, as well as the maximum body length (i.e., largest species per cell). We used this approach for map visualizations and for assessing body size variation with (absolute) latitude (using adequate techniques to account for spatial autocorrelation in statistical tests, details in ES3). However, our main analyses did not follow this grid cell-based approach as it is weakened by not accounting for phylogenetic effects, and by pseudo-replication due to one species occurring in several (or many) grid cells (Jetz & Rahbek, 2002; Entling et al., 2010).

In our comparative approach (individual-species focus) we treated each species as one data point. This also allowed inclusion of phylogenetic effects into models. To associate environmental predictors with each species, for all environmental variables (except forest cover) we calculated the mean across those 5 x 5 km pixels where the species occur (according to our range maps). For forest cover, we calculated the median at point localities of raw distribution records of high spatial accuracy (≤1 km uncertainty) to minimize error on habitat association, excluding ten species with ≤5 distribution records.

Data analyses

We replaced 24 ‘No Data’ values for forest cover by the means of the variable so as not to lose these species from multivariate models (Zuur et al., 2010). Furthermore, all data were standardized with a z-transformation \( \frac{x - \mu}{\sigma} \), where \( \mu \) is the mean and \( \sigma \) is the standard deviation) to make model coefficients comparable across different analyses. Body sizes, temperature extremes and seasonality were log₁₀-transformed after visual inspection of histograms to reach normality prior to the z-transformations, whereas
temperature data was $x^2$-transformed. NPP data (already normally distributed) and forest cover data (percent values) were standardized without transformation. We assured the fulfillment of model assumptions by checking visually for residual normality, outliers and heteroscedasticity (using R diagnostic plots).

Statistical analyses were conducted for all three body size measures (forewing length, thorax width, body length) as well as the product of body length $\times$ thorax width as response variables, but we present only body length analyses in the main text. This variable is less affected by different body shapes among subfamilies (e.g., systematic differences in thorax/wing length ratios). Models with body length as a response also received the highest pseudo-$r^2$'s in environmental models. We report results for the other metrics in the Electronic Supplement (ES5). In some specimens, no body lengths could be measured due to missing abdomens (a common problem issue in Lepidoptera collections where abdomens may be removed for genitalia dissection or simply be lost over time), which reduced sample size for these analyses to 938 species.

Phylogenetic effects

It must be assumed that a substantial part of body size variation is due to phylogenetic inertia – i.e., closely related species are similar in size. Such phylogenetic signal must be accounted for in interspecific comparisons to avoid spurious effects of non-independent data. We employed a stepwise, hierarchical approach to do so, accounting for the fact that we had a reliable, phylogenetically informed taxonomic classification, but not a true phylogeny with branch lengths, available for analyses (this would be required for many other approaches). Our classification was based on the Sphingidae Taxonomic Inventory (http://sphingidae.myspecies.info), which builds on the tribal-level molecular phylogeny of Kawahara et al. (2009) and integrates most recent findings of phylogenetic and taxonomic studies, molecular and morphology-based. The classification is available in ES2.

We started with environmental models containing no phylogenetic information, using a generalized least squares (gls) model. In three further models, we added information on subfamily, tribe and genus associations stepwise into hierarchical linear mixed effect models (lme) as random intercept effects, whereas links with the environment were modelled as fixed effects (Zuur, 2009). We see Stone et al. (2011) for mixed models as a method to account for phylogenetic effects. To assess the effect of phylogeny onto body size variability, we compared model Akaike information criteria (AICs) and pseudo-$r^2$ values (i.e., correlation of predicted vs. observed values), as well as the standardized coefficients fitted for environmental effects. For the model without random effects, we had to use a different algorithm (gls) to that used for models with random effects (lme). However, the gls and lme model coefficients are comparable (Pinheiro & Bates, 2006).

Univariate models

We correlated body length in separate models with every environmental predictor variable, using the stepwise approach described above to account for phylogeny. All univariate lme analyses were conducted with restricted maximum likelihood estimation (REML), since this is required for a comparison between models with differently nested random structure (Zuur, 2009).
Multivariate models

Multivariate analyses were conducted to investigate the independent influence of each predictor variable, using all predictors except latitude. Unlike for univariate \textit{lme} models, we used maximum likelihood (ML) fitting for multivariate \textit{lme} models because AICs from models with a different structure of fixed effects can only be compared when based on ML fits (even though REML fits are less biased; Zuur, 2009). With large sample sizes, differences in models fitted with the two different methods diminish (Zuur, 2009).

To detect and account for collinearity among the predictor variables, the variance inflation factor (VIF) of the multivariate model with all predictor variables was calculated. Zuur \textit{et al.} (2010) recommend dropping the predictor variable with the highest VIF, then recalculating the VIF values with the new model, repeating this until the VIF values are <3.

Results

There is no clear latitudinal cline in median body lengths (Fig. 1A). Rather, species in the Arabian Peninsula, in South Asia and in the Indo-Australian tropics are distinctively larger than temperate as well as Afrotropical taxa. A grid cell-based correlation of median BL with distance from the equator yields \( r^2 = 0.19 \), but it is non-significant due to high spatial non-independence of data (ES3, Fig. S3.4). However, when comparing temperate Eurasia and Africa it becomes clear that the variability in body lengths is larger in the Afrotropics, i.e., there are smaller and larger species than in temperate regions, leading to similar medians. Variability (Fig. 1B) is also high in desert regions of Africa and Asia, despite relatively low species richness (not shown; Ballesteros \textit{et al.}, 2017). Interquartile ranges are significantly negatively correlated with distance from the equator (\( r^2 = 0.47 \); ES3, Fig. S3.5), and maximum body length per grid cell is strongly and significantly negatively correlated to distance from the equator (\( r^2 = 0.54 \), ES3, Fig. S3.6). A map of minimum BL (ES3, Fig. S3.3) shows that the tropics also have smaller species than temperate regions, which supports the assessment that the latitudinal variation of maximum body size is mainly a function of greater variability in more species-rich regions (i.e., larger sample sizes).

Plotting median data in a comparative approach (Fig. 2) revealed a weak latitudinal pattern; species occurring at higher absolute latitudes are smaller, on average, than those at low latitudes. This relationship is significant but has almost no explanatory value (\( r^2 = 0.024 \); ES4, Table S4.1).

Fig. 2 shows raw data relationships of body length with latitude and the environmental predictors; Fig. 3 highlights the main characteristics of univariate correlations (i.e., model coefficients and explained variance; further test statistics in ES4, Table S4.1). Across all models, explained variance is low (pseudo-\( r^2 < 0.2 \)) if no or only higher-level phylogenetic classifications are integrated into models, while adding genus-level classification (hierarchically within higher-level classifications) yields pseudo-\( r^2 > 0.7 \). Integrating phylogenetic information weakens all fitted environmental effects, but it does not affect the rank order of their strengths, their sign, or their significance (all \( p < 0.001 \)) except for mean annual temperature (\( p < 0.05 \) when genus-level phylogeny is included).

Univariate correlations best support the resource availability hypothesis (positive correlation of body length and NPP). More weakly, data were also compatible with the ‘converse Bergmann-hypothesis’ (positive correlation of body length and mean...
temperature; Fig. 3) and the seasonality hypothesis (negative correlation of seasonality and size). Univariate correlations are inconsistent in sign with predictions from the TSR and Bergmann rule (both expecting a negative link with mean temperature), the agility hypothesis (expecting a negative link with forest cover), and the starvation hypothesis (expecting positive links with temperature extremes).

Multivariate modelling (Fig. 4; details in ES4, Table S4.2), as in the univariate models, features low pseudo-$r^2$’s unless genus-level classifications are integrated into the models. They also support the link of body length with NPP (i.e., resource availability hypothesis) whereas the positive link with temperature (converse Bergmann hypothesis) is weaker, in particular when phylogenetic data are included. Seasonality effects are weak and inconsistent depending on how much phylogenetic information is included. As in the univariate analyses, multivariate models indicated larger, not smaller species in forested habitat, contrary to the agility hypothesis. This pattern is not because forests feature higher NPP (which is accounted for in models). Unlike with the univariate models, temperature extremes show positive coefficients with body length in multivariate models (as expected by the starvation hypothesis), but links are weak, non-significant, and VIF analysis indicated predictor collinearity issues.

Based on VIF we simplified this full multivariate model by first dropping seasonality, then temperature extremes. In a final step we also dropped mean annual temperature although it had a very low VIF value because univariate analyses had suggested that it was non-significant and yielded higher AIC values than a model consisting only of forest cover and NPP (ES4, Table S4.2). Finally, a comparison of NPP and forest cover showed that they truly express quite independent aspects of the environment (correlation: $r^2 = 0.26$). Notably, multivariate models do not explain substantially more variance overall than univariate models if phylogeny is fully included. Models using other body size metrics as responses, whether univariate or multivariate, lead to the same conclusions (ES5).

**DISCUSSION**

For 950 sphingid species, occurring from the northern temperate through the African and Asian tropics to the southern temperate zone, we observed distinct geographical patterns of average body size (Fig. 1) that were only weakly related to absolute latitude. Rather, they featured larger species in the Middle East through South- and Southeast-Asia to Melanesia, and smaller species elsewhere, including the African tropics. However, species richness in the Middle East is low (Ballesteros-Mejia *et al.*, 2017), which increases chance effects (i.e., very high values on the Arabian Peninsula are based on few species per pixel). A comparative, species-focused analysis revealed landscape productivity (NPP) as the strongest correlate of body size, supporting the ‘resource availability’ hypothesis. Other hypotheses of environmental control of body size were weakly (i.e., ‘converse Bergmann’; seasonality) or not at all supported; for many-some hypotheses (i.e., TSR, agility, starvation) even the sign of the expected relationship was not met by data. These conclusions held for univariate and multivariate modelling approaches, with or without the inclusion of phylogenetic relatedness of taxa.

However, broad environmental conditions, even NPP, generally explained relatively little of the variability of body sizes (i.e., <10%) whereas hierarchically including phylogeny to genus-level consistently produced sound predictions of the global variability in sizes (i.e., >70% of size variability explained). This implies that phylogenetic inertia in body sizes, combined with the unknown, past evolutionary events
that shaped the sizes of today’s higher-level taxonomic groups’ ancestors are much more relevant for predicting observed body size distributions than current environments.

In light of this it is not surprising that, consistent with other invertebrate studies (Shelomi, 2012), the much-discussed latitudinal gradient of (average) body sizes, although statistically observable in data (Fig. 2), is buried under random noise and explains very little of the data variability (Table S3.1). Because this scatter is wider with higher species richness towards the tropics (Fig. 2), maximum body sizes feature a much stronger latitudinal pattern, which shapes the intuitive perception of the pattern (Fig. 1).

Resource availability effects

Our data best supported the idea that body size increases with food availability, which we approximated with productivity (NPP; Blackburn et al., 1999). However, a mechanism whereby plant productivity constrains consumers’ body size appears in conflict with the idea of ‘energy equivalence’ (Damuth, 2007; Brown et al., 2004), which postulates that the variability of species richness, not body size, offsets variability in energy availability. In light of these theoretical uncertainties, we caution against interpreting our statistical support, or that of other studies, as unequivocal proof for the postulated mechanism of the food availability hypothesis. However, we see additional post-hoc support for interspecific food limitation effects on body size in the positive correlation of forest and size. We had originally included/considered forests in preliminary analyses to test the agility hypothesis (dense forest benefitting small, agile species in large mammals; Bro-Jørgensen, 2008), but removed this idea from our presentation as quite unreasonable for flying insects (data rejecting-rejected it comprehensively, not shown it because we found positive, not negative links with size).

Although forest data appear statistically independent of NPP (Table S3.2), the recovered positive correlation may be due to the fact that a sizable portion of NPP in some regions derives from grassland habitats. However, sphingid caterpillars are folivorous and only a single genus, Leucophlebia, is confirmed as feeding on Poaceae (Diehl, [1982]). The absence of forests may therefore pinpoint where there is little food for sphingids even if NPP may not be particularly low (i.e., in grasslands).

Other hypotheses

We found only weak coefficients for the ‘converse Bergmann’ and seasonality hypotheses, and multivariate analyses suggest predictor collinearity issues or other artefacts. Recently, Zeuss et al. (2017) presented similar data on geographic patterns of size in European Lepidoptera (and Odonata). While our study went beyond Zeuss et al. (2017) in some important aspects (i.e., larger geographic coverage, more fully including global environmental gradients; comparative analyses accounting for phylogenetic effects), they had data for, and highlighted, a key variable to at least potentially disentangle some mechanistic effects, i.e., voltinism (the number of generations per year). Several hypothetical mechanisms assume, naively, a constant and consistent number of generations per year, which is known to be untrue for many well-studied insect taxa (Zeuss et al., 2017; and references therein). If longer favorable growing seasons or warmer temperatures lead to more generations, this may offset any potential body size increases as postulated by ‘converse Bergmann’ or seasonality hypotheses.

For the other tested hypotheses (i.e., Bergmann/TSR, agility) we did not even observe correlations of the correct sign, which clearly rules them out within the following limitations. The starvation hypothesis postulates that larger species have a lower risk of
death due to unexpected climate events, such as starvation or desiccation (Chown & Gaston, 2010). We did not find any support for this idea from negative temperature extremes (which would limit movement and hence foraging), while we could not test, due to data limitations, effects of unusual drought events. However, most growth in sphingid larvae happens during the last instar (e.g., 90% in *Manduca sexta*; D’Amico et al., 2001), so adult body sizes may not provide a suitable test of postulated starvation effects on the size of earlier instars (i.e., phenology may be more relevant than final size). However, in other insects there has also been little support for this hypothesis except for ants, which are a special case because of their eusocial behavior (Kaspari & Vargo, 1995). Support for the TSR in the literature is mainly from intraspecific studies (including a sphingid; Davidowitz & Nijhout, 2004), but its role is apparently small when it comes to explaining interspecific patterns.

**Our results are, at least in parts, in conflict with some recent, more localized studies on the body sizes of Lepidoptera and other taxa.** Brehm et al. (2019) reported a size increase with elevation in Costa Rica moths, which was best explained by a negative effect of temperature (cf. Bergmann, TSR) but not by an effect of productivity. Beck et al. (2016) also found increasing moth sizes with higher elevations, although they could not link this to flight-time temperatures in their highly seasonal study region, the Swiss Alps. While inconsistent results on from ectotherm body size patterns and their causes are common in the scientific literature so far (e.g., Entling et al., 2010; Fattorini et al., 2013; Slavenko & Meiri, 2015; Zeuss et al., 2017; Pallares et al., 2019), it is particularly surprising that sphingids, well-known for their (partial) thermoregulation as imago adults (Heinrich, 1993), did not follow a pattern interpretable towards Bergmann’s rule and heat-preservation whereas other moths groups showed did show such trends in elevational gradient studies. We cannot assess whether methodological differences, peculiarities of elevational studies, a generally high potential for spurious results when dealing with small effect sizes, or other factors may have caused these discrepancies.

**Limitations of the study**

This study tested, and rejected, some of the global hypotheses regarding environment-body size relationships, based on currently available data on adult sphingid moth sizes. However, our database is necessarily imperfect, and we need to briefly discuss how this may have biased the conclusions of our analyses. First, we used an up-to-date, but nevertheless coarse and incomplete taxonomic classification as a proxy for a complete, time-calibrated phylogeny with branch length data. Our step-wise inclusion of lower-taxon information gradually increased the explanatory value of statistical models while reducing the tested environmental effects (Fig. 3). We therefore assume that the predominance of phylogeny, as opposed to current environment, in explaining size patterns would be even more pronounced with a more detailed, finely-resolved phylogeny. Second, we only had adult size data available, but many of the mechanistic explanations proposed may apply more strongly to the larval stage, where growth occurs. Thus, concluding that a given hypothesis does not explain observed adult sizes remains valid, but this does not rule out that the proposed mechanisms might play some role in shaping the life histories of juvenile stages, which can differ in important aspects (among them, in sphingids, thermoregulation; Heinrich, 1993; Kingsolver et al., 2015). Third, ignoring intraspecific size variation essentially means that we had to view our measurements as a (small) statistical sample around an unknown per-species mean. We do not see any obvious directional bias in this, but it implies that our data may contain considerable random noise simply due to occasionally measuring unusually large or small specimens of a species. As a consequence, the tested effects may be underestimated to an
unknown degree. However, the relatively large sample size (i.e., 950 species) would help
to counter such chance effects. Programs of computer-aided photography and
measurement of specimens in museum collections are under way but they may not
directly provide reliable data for solving this issue (Johnson et al., 2013; Trueman &
Yeates, 2015). Fourth, our analytical setup involved two potentially confounding effects,
phylogenetic inertia and spatial autocorrelation. It is exceedingly complex to control
analyses for both effects (Freckleton & Jetz, 2009), and we choose to account in our
models for the more important of the two, phylogeny. Strong phylogenetic effects on
body size have been repeatedly reported in the literature, and they were shown on our
data. Spatial autocorrelation occurs in most geographic data and has the potential to
affect significance assessments and possibly also coefficient estimates (Bini et al., 2010).
However, while we do not wish to give the impression of taking this lightly, it is our
assessment that statistically strong patterns rarely lead to changing conclusions when
applying spatially explicit modelling, whereas the phylogenetic inertia of body sizes most
probably has the potential to affect conclusions in empirical studies (Diniz-Filho &
Torres, 2002).

Conclusions

Average sphingid moth body sizes show distinctive geographic patterns, but they
vary only very weakly with absolute latitude. However, maximum body sizes per cell,
which are affected by species richness, indicate an increase towards the tropics, which
creates the impression of an overall size increase towards the tropics. Among various
hypotheses regarding how the environment is shaping such patterns globally, we found
strongest support for the food availability hypothesis, as sizes increase with net primary
productivity. There was no support at all (i.e., opposite sign of effect) for Bergmann’s
rule or the temperature size rule, the agility hypothesis, or the starvation hypothesis,
while support for the seasonality hypothesis or the converse Bergmann pattern was weak
and potentially unreliable. However, phylogenetic effects were much more relevant than
any of the tested environmental factors in shaping the observed size data. This implies
that past environmental factors in the regions of occurrence of today’s higher taxon’s
ancestors may have shaped observed size patterns, which will be exceedingly challenging
to test.

References

variability and latitude. Proceedings of the Royal Society (B), 267, 739–745.
reptiles: turtles follow it, lizards and snakes reverse it. Evolution, 57, 1151–1163.
Advances in Ecological Research, 25, 1-58.
map: Near-global variation in sphingid moth richness along spatial and


**Data Accessibility Statement**

As electronic supplements to this article we publish our data on sphingid body sizes (ES1), a taxonomic classification of species (ES2) as used in analysis, and GIS-compatible data of Fig. 1 (ES6). Environmental data used in our analysis are from the public domain. Raw distribution maps of sphingid species can be accessed at Map of Life, [www.mol.org](http://www.mol.org). *

*) currently only for browsing, by the time of acceptance of this paper they will be available for download in GIS format.

**Appendices: Electronic Supplements**

- **ES1** Body size data and specimen sources (spreadsheet in csv-format)
- **ES2** Taxonomic classification as used for analyses
- **ES3** Supplementary maps and figures
- **ES4** Supplementary model statistics
- **ES5** Model output for other body size metrics
- **ES6** Body size distribution maps (median, interquartile, maxima; ASCI-format)
FIGURE CAPTIONS

Figure 1 Geographical pattern of the median body length (A), its interquartile range (B) and maximum body length found in each cell (C). Grid cells with ≤5 species were removed (map resolution: 100 x 100 km, Mollweide World geographical projection). See ES3 for additional maps; data for Fig. 1 are available in GIS-compatible format (ES6).

Figure 2 Body length plotted against all predictor variables. LOESS (locally weighted scatterplot smoothing) is fitted to indicate main data trends.

Figure 3 Summary of univariate model results (response: body length) with stepwise hierarchical inclusion of phylogenetic information. Bars (right y-axis) indicate the variance explained by the models, line plots (left axis) the fitted coefficients for the different predictors. Because data are standardized, coefficients can be compared across models as a measure of effect size.

Figure 4 Summary of multivariate model results (response: body length) with stepwise hierarchical inclusion of phylogenetic information. Bars (right y-axis) indicate the overall variance explained by the multivariate model models, line plots (left axis) the fitted coefficients for the different predictors.