Phylogenetic analyses suggest that diversification and body size evolution are independent in insects

James L. Rainford¹, Michael Hofreiter² and Peter J. Mayhew¹*

Abstract

Background: Skewed body size distributions and the high relative richness of small-bodied taxa are a fundamental property of a wide range of animal clades. The evolutionary processes responsible for generating these distributions are well described in vertebrate model systems but have yet to be explored in detail for other major terrestrial clades. In this study, we explore the macro-evolutionary patterns of body size variation across families of Hexapoda (insects and their close relatives), using recent advances in phylogenetic understanding, with an aim to investigate the link between size and diversity within this ancient and highly diverse lineage.

Results: The maximum, minimum and mean-log body lengths of hexapod families are all approximately log-normally distributed, consistent with previous studies at lower taxonomic levels, and contrasting with skewed distributions typical of vertebrate groups. After taking phylogeny and within-tip variation into account, we find no evidence for a negative relationship between diversification rate and body size, suggesting decoupling of the forces controlling these two traits. Likelihood-based modeling of the log-mean body size identifies distinct processes operating within Holometabola and Diptera compared with other hexapod groups, consistent with accelerating rates of size evolution within these clades, while as a whole, hexapod body size evolution is found to be dominated by neutral processes including significant phylogenetic conservatism.

Conclusions: Based on our findings we suggest that the use of models derived from well-studied but atypical clades, such as vertebrates may lead to misleading conclusions when applied to other major terrestrial lineages. Our results indicate that within hexapods, and within the limits of current systematic and phylogenetic knowledge, insect diversification is generally unfettered by size-biased macro-evolutionary processes, and that these processes over large timescales tend to converge on apparently neutral evolutionary processes. We also identify limitations on available data within the clade and modeling approaches for the resolution of trees of higher taxa, the resolution of which may collectively enhance our understanding of this key component of terrestrial ecosystems.

Keywords: Body size, Diversification, Hexapoda, Insects, Phylogeny

Background

One of the most prevalent patterns observed in natural systems is the overrepresentation of small-bodied taxa [1]. The observation of right skew in body size distributions, following transformation to the log scale, has been made for a variety of vertebrate clades [2–4] and provides the basis for a variety of size-selective diversification mechanisms that have been previously proposed as general models for the macroevolution of animals (reviewed in [1, 4]). Despite widespread interest in these patterns, comparatively little effort has been spent in examining whether such relationships are truly universal and there is limited evidence for their presence across major non-vertebrate lineages [5–7]. In this study, we explore the relationship between species richness and body size, and the universality of size biased diversification, in one of the largest terrestrial invertebrate clades, the six-legged arthropods or Hexapoda.

Interest in body size distributions relates to the importance of size in impacting on an organism’s ecology and thus potential evolution and diversification. Body
size determines the scale of an organism’s interactions within the fractal structure of natural environments [8, 9], the relative strength of gravitational (i.e. body weight) vs. viscous and inertial forces [10] and, via surface area to volume ratios and the scaling of exchange networks, controls the rates of metabolic processes such as temperature response [11] and gas diffusion [12]. As a consequence, body size impacts on almost every major life history trait including: growth, parental investment, range size, dispersal and degree of host specificity (see [13–15], and references therein, for reviews of Hexapoda).

Based on these observations a number of size-dependent mechanisms linked to clade diversification have been proposed (reviewed in [3, 4]). These include; hard limits on minimum size, which restrict random character change [16], energetic models emphasizing the relative efficiency of small body sizes in the production of offspring [11, 17], and fractal environmental models, exploring the capacity for small-bodied taxa to more finely subdivide a given environmental landscape [8]. The relationship of these processes to macro-evolutionary diversification remains incompletely understood including, for example, the relative contributions of size-biased cladogenesis (i.e. small taxa being more prone to speciation) [2], directional bias in size evolution within lineages; e.g. “Copes rule” [18], and size-biased extinction [19], on the generation of observed size distributions. Testing the predictions of these models, e.g. the presence of a relationship between clade richness and body size, as well as more generally exploring the processes that may underlie size evolution, requires that we extend our perspectives to encompass other major lineages that may show differences from our vertebrate model systems [20].

The extreme species richness of hexapod clades, which collectively account for over half of all described species, is one of the most well-known features of terrestrial biomes [21]. Hexapoda are also morphologically diverse, including body lengths ranging over four orders of magnitude, comparable with the range of well-studied mammal and bird radiations [13]. The longest known hexapods are females of the phasmid (stick-insect) Phobaeticus chani with specimens up to 357 mm long in body length. By contrast, the smallest recognized adult insect, the male of the mymarid wasp Dicopomorpha echmepterygis has a total body length of merely 139 μm (or 0.139 mm) [13] (see [22] for further examples of extreme miniaturization in hexapods). Evidence to suggest that processes in hexapod size evolution may be distinct from larger vertebrate groups includes taxonomic compilations (e.g. [23]), regional faunal data (e.g. [24, 25]) and broad-scale continental surveys [26], all of which suggest that compared with vertebrates hexapods exhibit relatively little right skew in the distribution of log body size [13, 15]. Likewise, where formal phylogenetic tests of association between clade richness and body size have been conducted for hexapod sub-clades, they have generally failed to recover evidence for small size promoting richness within the group (e.g. [27]), with one study even identifying the opposite pattern with respect to Anisoptera (dragonflies) [28].

In addition to these apparent divergences from size-structured models there are also potential interactions between size evolution and other hexapod traits, several of which have been previously explored as correlates of species richness including complete metamorphosis, and dietary substrate [21, 29, 30]. Metamorphosis has the potential to structure size evolution via the promotion of modularization of life history stages, and the separation of selection pressures on larval and adult stages [13, 31]. This process is taken to extremes in Holometabola, where during metamorphosis there is a fundamental reorganization of the body plan [32], and as a result various authors have suggested divergent processes of size evolution associated with this clade (it should be noted, however, that the manifestation of these effects in terms of models of trait evolution remains poorly understood [13, 33]).

The recent and growing consensus with regard to hexapod higher taxonomic relationships from molecular markers e.g. [30, 34, 35] provides us, for the first time, with a framework for exploring large scale patterns of trait evolution within the group. In this study, we combine a published phylogeny of insect higher taxa [30] with comprehensive descriptive information regarding size variation within the clade to explore patterns of body size evolution and its relationship with clade diversification. Hypotheses we test include: a) if the apparent lack of skew in body size distributions (on the log scale) identified for regional faunas can be identified in a global phylogenetic perspective on hexapod body size, b) if consistent relationships between clade richness and body size occur after accounting for phylogeny and size variation within terminal groups. In addition, we explore the probable evolutionary process that may underpin size evolution in hexapods, and whether different major clades (e.g. Holometabola or major orders) are associated with divergent evolutionary processes, as has previously demonstrated in mammals [36], with an aim to explore the possible roles of key innovations such as complete metamorphosis [30].

Results
Frequency distribution of body sizes
Body length range data were gathered for 774 higher taxa of insects (resolved primarily to the family level; Additional file 1: Table S1). The frequency distributions of the observed values of mean-of-logs (mean of the logged values of the size range limits for each higher taxon), log maximum and log minimum body length for terminal
taxa are shown in Fig. 1. In all three cases the overall distributions are approximately normal (two-sided Agostino test, log minimum: skew = 0.3333, z = 2.455, p-value = 0.0141, log maximum: skew = 0.0752, z = 0.567, p-value = 0.5706, mean-of-logs: skew = 0.210, z = 1.572, p = 0.116), although the distribution of minimum sizes shows a small secondary peak associated with an over-prevalence of taxa reported as bounded at 1 mm (commonly used for convenience in descriptions of small taxa). When mean values are weighted according to their species richness, the resulting distribution shows a significant skew towards larger body sizes (skew = -0.0290, z = -7.91, p-value = <0.001) running contrary to the expectations of the paradigm described above.

Comparing major clades we can identify pronounced differences in typical size distributions observed among groups. As Holometabola, the most diverse clade (more than 75% of all extant hexapods) [32] account for the majority of the terminals included in this study (508 out of 775), it is unsurprising that the size distribution of Holometabola (insects with complete metamorphosis) mirrors that of hexapods as a whole, with similar average size to the global mean (Hexapoda; (log) mean = 1.946 ln (mm), sd = 0.9491 ln (mm), Holometabola; (log) mean = 1.8032 ln (mm), sd = 0.8078 ln (mm)). By contrast both the clades Entognatha (non-insect hexapods including springtails; mean =0.8879 ln (mm), sd = 1.061 ln (mm) and Paraneoptera (true bugs and their relatives; mean = 1.5506 ln (mm), sd = 0.7755 ln (mm)) are predominantly composed of groups falling at the small end of the size spectrum, the latter particularly with respect to minimum sizes, while large insects include disproportionate representation of Polyneoptera (mean = 3.045 ln (mm), sd = 0.7455 ln (mm)) and Palaeoptera (particularly large bodied Odonata (dragonflies)); mean = 3.060 ln (mm), sd = 0.8825 ln (mm)).

The value of the inferred standard deviation of the terminal distributions shows a rather different phylogenetic
pattern from that of the mean size values, although after taking phylogeny into account the two are strongly correlated (PGLS [37] assuming a Brownian covariance structure: Estimate = 0.4219, SE = 0.1830, t = 2.3049, p = 0.0214). Clades associated with particularly low values of standard deviation (implying relatively little size variation after accounting for species richness within terminal groups) include Trichoptera, Neuroptera (lacewings and relatives), Psocodea and Odonata while the largest values occur in Coleoptera and advanced Lepidoptera (Fig. 2), with the single largest value occurring in the morphologically diverse (4-39 mm) but species poor Lepidoptera family Aididae (6 species).

**Phylogenetic distribution of body size and ancestral states**

The above patterns are reinforced on the phylogenetic ancestral reconstruction plots for the group (Fig. 2, Additional file 1: Figure S1), in which the following clades show strong deviations from the average size dynamics: Odonata (with respect to larger than average minimum body size), Psocodea (booklice andlice; small maximum sizes), micro-hymenoptera (the smallest members of Holometabola with particularly small minimum size bounds) and various polyneopteran clades, notably Phasmatodea and Orthoptera. Beyond these limited examples, the majority of hexapod higher taxa log-means lie close to global average size, and ancestral reconstruction of internal nodes rapidly approaches this value as an approximation of the global ancestral state.

Evidence of phylogenetic signal was recovered in both the full dataset and in all the major sub-clades (Table 1) with very strong support, with the exception of Entognatha, where evidence of structuring is present but support is much lower (likely due to the small number of tips on this subtree: 12). Blomberg’s K values indicate that Hexapoda as a whole demonstrate somewhat lower values of K than would be expected under a Brownian motion (BM) process, consistent with related species resembling one another less than under the expected BM distribution (see further discussion below). Similar patterns are also identified in Holometabola and Polyneoptera. By contrast, Paraneoptera and Palaeoptera show strong tendencies towards higher-than-expected values of K, indicating differences in the size evolution process among major clades.

**Body size and species richness**

The standardized contrasts in body size and relative rate difference (RRD; defined as, ln (N₁/N₂), where N₁ = richness of descendant clade with larger body size, and N₂ = the richness of the other descendant clade [38–40]) across major clades are plotted in Fig. 3. The estimated relationships through the origin were calculated on the observed mean-of-log sizes and confidence intervals were based on the parametric bootstrap samples as drawn from the estimated terminal distributions for both observed (colored) and randomized (black) data (parameter values in Table 2).
Overall, the data for Hexapoda support the presence of a weak positive relationship between richness and body size within the clade, although following the parametric bootstrap this relationship is not significant once the uncertainty of terminal states is taken into account. Similar patterns of null relationships once tip variance is taken into consideration occur in all of the major sub-clades examined, although in the case of Palaeoptera the direction of the relationship observed is negative. When these statistics were recalculated based on PDI (Additional file 1: Table S2) no significant relationships were observed between mean size and richness, rendering further parametric bootstrapping redundant.

**Process of body size evolution**

Considering the potential processes responsible for generating observed patterns of size evolution (see Methods), our data suggest that, of our process based models; the majority of hexapod clades favor simple Brownian motion, with the exception of Holometabola, where the favored process is an single stationary peak (SSP/OU) model with convergence on a single global optimum or elevated diversification at distant tips (Table 3, Additional file 1: Table S3). However, when models without an explicit generating process are considered (i.e. lambda and white noise (WN)), this picture changes, such that for Hexapoda as a whole and Holometabola, there is evidence for considerable non-phylogenetic signal in body size, resulting in lambda values that significantly diverge from the expectations of BM (although in all cases the WN model with no phylogenetic signal is strongly rejected, see also Table 1). Similar patterns are observed when the major holometabolan orders are examined individually, with Hymenoptera (bees, wasps and ants), Coleoptera (beetles) and Lepidoptera (moths and butterflies) all favoring BM processes, while Diptera (flies) shows strong evidence for non-phylogenetic signal (thus favoring the lambda model). The implications of these differences for our understanding of size evolution in hexapods, and particularly within Holometabola and Diptera, will be explored below.

The findings of Bayesian Analysis of Macro-evolutionary Mixtures (BAMM) further support the idea that the process of size evolution behaves differently in holometabolan and non-holometabolan groups (Fig. 4). A single shift in the rate model associated with the origins of Holometabola is recovered with a marginal probability of 0.988, i.e. it is found in >95 % of all sampled models from the post burn-in chain. The single most sampled configuration recovers only this shift (with a relative frequency of 0.5; Additional file 1: Figure S2), suggesting that the impact of other events on size evolution within the group is comparatively marginal. This regime shift in Holometabola is associated with a reversal in the rate of size evolution, such that within this clade rates appear to accelerate through time, contrasting with the weak deceleration observed across the remaining hexapods (potentially consistent with the BM process described above). The only other nodes found to significantly contribute to heterogeneity in size evolution within hexapods are associated with decelerations in size evolution within Trichoptera, both when including (relative frequency 0.17) and when excluding (relative frequency 0.18) the basal family Hydropsyphidae.

**Discussion**

The findings of this study corroborate previous taxonomic surveys at continental scales (e.g. [24–26]) suggesting that the distribution of body lengths in hexapod families does not show a strong skew towards an over-abundance of small sized taxa on the log scale. We also demonstrate that, while size does show phylogenetic structuring with respect to different hexapod groups, after accounting for these relationships and the variances observed within tip groups, there is no global negative association between body length and diversification across the studied taxa. Finally, our survey of possible evolutionary models suggests that the pattern and processes of size evolution in Holometabola, and possibly Diptera, are distinct from those of other hexapod groups. In both cases evidence for non-phylogenetic signal suggests that these differences cannot be adequately accounted for in single parameter extensions of Brownian motion, although for other groups, body size evolution looks approximately Brownian.

The recognition that body length distributions in Hexapoda show relatively little bias on a log scale, and that diversification rates within the group are approximately independent of size, supports the idea that concepts derived from the study of vertebrate groups [1, 3] may be inappropriate when discussing other taxonomic groups [5, 6], and hexapods in particular [13, 15, 26]. Possible explanations for these differences focus on the potential for small absolute body size to alter the link between body-size and clade diversification. For example, small-bodied organisms experience distinct flow conditions where viscous forces, such as surface tension

| Table 1 Tests of phylogenetic signal within major clades incorporating within-terminal standard error |
|-----------------------------------------------|
| Taxa         | Blomberg’s K | Sigma² rate parameter | Model log likelihood | P randomization test |
|--------------------|---------------|-----------------------|----------------------|----------------------|
| Hexapoda           | 0.8870        | 0.002368              | −778.95              | <0.001               |
| Holometabola       | 0.6864        | 0.002694              | −515.43              | <0.001               |
| Paraneoptera       | 1.3166        | 0.001436              | −117.07              | <0.001               |
| Polyneoptera       | 0.8144        | 0.002122              | −66.26               | <0.001               |
| Palaeoptera        | 1.7806        | 0.001467              | −40.192              | <0.001               |
| Entognatha         | 1.1244        | 0.002574              | −15.711              | 0.0247               |

The findings of Bayesian Analysis of Macro-evolutionary Mixtures (BAMM) further support the idea that the process of size evolution behaves differently in holometabolan and non-holometabolan groups (Fig. 4). A single shift in the rate model associated with the origins of Holometabola is recovered with a marginal probability of 0.988, i.e. it is found in >95 % of all sampled models from the post burn-in chain. The single most sampled configuration recovers only this shift (with a relative frequency of 0.5; Additional file 1: Figure S2), suggesting that the impact of other events on size evolution within the group is comparatively marginal. This regime shift in Holometabola is associated with a reversal in the rate of size evolution, such that within this clade rates appear to accelerate through time, contrasting with the weak deceleration observed across the remaining hexapods (potentially consistent with the BM process described above). The only other nodes found to significantly contribute to heterogeneity in size evolution within hexapods are associated with decelerations in size evolution within Trichoptera, both when including (relative frequency 0.17) and when excluding (relative frequency 0.18) the basal family Hydropsyphidae.

**Discussion**

The findings of this study corroborate previous taxonomic surveys at continental scales (e.g. [24–26]) suggesting that the distribution of body lengths in hexapod families does not show a strong skew towards an over-abundance of small sized taxa on the log scale. We also demonstrate that, while size does show phylogenetic structuring with respect to different hexapod groups, after accounting for these relationships and the variances observed within tip groups, there is no global negative association between body length and diversification across the studied taxa. Finally, our survey of possible evolutionary models suggests that the pattern and processes of size evolution in Holometabola, and possibly Diptera, are distinct from those of other hexapod groups. In both cases evidence for non-phylogenetic signal suggests that these differences cannot be adequately accounted for in single parameter extensions of Brownian motion, although for other groups, body size evolution looks approximately Brownian.

The recognition that body length distributions in Hexapoda show relatively little bias on a log scale, and that diversification rates within the group are approximately independent of size, supports the idea that concepts derived from the study of vertebrate groups [1, 3] may be inappropriate when discussing other taxonomic groups [5, 6], and hexapods in particular [13, 15, 26]. Possible explanations for these differences focus on the potential for small absolute body size to alter the link between body-size and clade diversification. For example, small-bodied organisms experience distinct flow conditions where viscous forces, such as surface tension
Fig. 3 (See legend on next page.)
and air resistance, have the potential to overwhelm the effect of the gravitational forces (i.e. body weight) that are responsible for structuring body size changes at larger spatial scales [10, 41]. Likewise, fractal environmental models, which postulate the existence of a higher number of niches at small body sizes [8, 9], may become inapplicable below a certain scale, particularly with respect to “parasitic” taxa, which live on the surface of larger host organisms (typical of the majority of hexapods), and are therefore subject to local homogeneity in the composition of their environment across a range of spatial scales [23, 42, 43]. In addition with respect to hexapods, despite a general trend towards larger-bodied organisms showing greater reproductive output, there is evidence from well-studied systems to suggest that this pattern is not universal across the group [17, 44, 45]. Thus, several of the mechanisms typically invoked to account for size-biased diversification in vertebrates may not be applicable to Hexapoda, reflecting a potential danger of extrapolation from well-studied, but atypical clades to describe global evolutionary processes [6]. There is a need to further investigate processes of size evolution across a broader range of invertebrate groups for comparative purposes (e.g. [46]), which, when taken together, may provide us with new insights into underlying mechanisms controlling the size structuring of natural environments [47].

Despite the presence of non-phylogenetic signal in some specific groups, there is considerable evidence that the majority of hexapod clades are strongly phylogenetic structured with respect to body size, and hence size evolution within Hexapoda is broadly described by a BM process on the log scale. However, many specific clades appear, within the limits of available data, to be constrained to a particular subset of possible sizes. The mechanisms underlying such constraint are likely to be variable across different lineages. For example, the absence of small body sizes within Odonata may be attributed to limitations on the minimum size required for the group (e.g. the absence of large bodied members of the order Psocodea (booklice; even after accounting for the parasitic and small-bodied Pthiraptera), which may reflect constraints of a cryptic and concealed lifestyle in a group that has received comparatively little detailed study. The effect of such constraints at the super-ordinal scale appears to be marginal, as all of the major lineages demonstrated a wide variation in size as well as homogeneity of process within clades (and across clades, with the exception of Holometabola and Diptera). The overriding impression therefore is that, within the limitations imposed by restricted phylogenetic resolution, size evolution within hexapods is dominated by comparatively localized factors operating at the sub-ordinal or super-familial level.

The reconstruction of estimated standard deviation in body size within Hexapoda generated here bears a strong qualitative resemblance to previously recovered patterns of diversification rate shifts across the clade [30]. This is particularly striking in that clades previously recovered as downsized with respect to diversification rate, e.g. Psocodea, Neuroptera and Trichoptera, are here recovered as having comparatively low standard deviation in body size, suggesting a link between the diversification process and radiation into novel morphospace [49]. Similar ideas have been previously proposed with respect to bird families, [49], but formalized testing via multiple regression has been shown to be statistically problematic, due to an inability to distinguish time-dependent and speciation-

Table 2 Macrocaic analysis of contrasts in RRD and vs. mean-of-logs size for major hexapod clades (Fig. 3)

| Taxa               | N  | Estimate | (Adj) R²  | SE | t     | p     | Obs. QR | NULL QR |
|--------------------|----|----------|-----------|----|-------|-------|---------|---------|
|                    |    |          |           |    |       |       | 2.5 %   | 97.5 %  |
|                   |    |          |           |    |       |       |         |         |
| Hexapoda           | 773| 4.538    | 0.004203  | 2.219| 2.045| 0.0412 | 1.886   | 5.383   |
| Holometabola       | 507| 4.415    | 0.003232  | 2.715| 1.626| 0.105  | 1.246   | 5.580   |
| Non-Holometabola   | 265| 5.416    | 0.003874  | 3.801| 1.425| 0.155  | 1.927   | 7.304   |
| Paraneoptera       | 126| 11.759   | 0.02523   | 5.696| 2.064| 0.0415 | 5.495   | 14.35   |
| Polyneoptera       | 64 | 9.135    | 0.009866  | 7.139| 1.28  | 0.205  | 1.256   | 14.02   |
| Palaeoptera        | 57 | –8.866   | –0.00021  | 8.919| –0.994| 0.325  | –12.63  | –2.987  |
| Ectognatha         | 11 | 12.43    | –0.04417  | 17.00| 0.731| 0.481  | 5.118   | 17.94   |

Data shown are the results of parametric bootstrap, with 50,000 replicates. Shown are the observed quartile ranges (Obs. QR) and those of the Null tip-randomized data (NULL QR). * indicates a significant relationship prior to parametric bootstrap (but not after)
Table 3  Parameter estimates and relative likelihoods for models of mean-of-logs body size incorporating within-terminal standard error  

| Clade       | Model | Sigma squared | z₀              | LnLik  | k  | AICc  | Delta AICc from optimal model | Akaike weights |
|-------------|-------|---------------|-----------------|--------|----|-------|-------------------------------|----------------|
| Hexapoda    | BM    | 0.002403      | 1.749           | −779.4 | 2  | 1562.7 | 21.031                        | 0.00003        |
|             | EB    | 0.002404      | 1.748           | −779.4 | 3  | 1564.7 | 23.051                        | 0.00001        |
|             | delta | 0.002196      | 1.766           | −791.1 | 3  | 1564.3 | 22.627                        | 0.00001        |
|             | SSP   | 0.002666      | 1.764           | −780.0 | 3  | 1562.1 | 20.434                        | 0.00004        |
|             | lambda | 0.001957     | 1.759 0.92093   | −767.8 | 3  | 1541.7 | 0                             | 0.9991         |
|             | WN    | 0.8985        | 1.946           | −1057.3 | 2  | 2118.7 | 576.99                       | 0.0000         |
| Holometabola| BM    | 0.002726      | 1.846           | −515.4 | 2  | 1034.8 | 17.571                        | 0.0002         |
|             | EB    | 0.002727      | 1.846           | −515.4 | 3  | 1036.9 | 19.600                        | 0.0001         |
|             | delta | 0.001787      | 1.802 1.881     | −511.2 | 3  | 1028.5 | 11.265                        | 0.0035         |
|             | SSP   | 0.003613      | 1.830           | −510.7 | 3  | 1027.4 | 10.170                        | 0.0061         |
|             | lambda | 0.002138     | 1.845 0.89028   | −505.6 | 3  | 1017.3 | 0                             | 0.9901         |
|             | WN    | 0.6498        | 1.803           | −611.9 | 2  | 1227.8 | 210.52                       | 0.0000         |
| Paraneoptera| BM    | 0.001469      | 1.132           | −117.0 | 2  | 238.2  | 0                             | 0.3939         |
|             | EB    | 0.001518      | 1.130           | −0.00111 | 3  | 240.3  | 2.094                         | 0.1382         |
|             | delta | 0.001559      | 1.119           | 0.9931 | 3  | 240.1  | 1.9781                        | 0.1465         |
|             | SSP   | 0.001469      | 1.132           | 0.00   | 3  | 240.3  | 2.0983                        | 0.1379         |
|             | lambda | 0.001368     | 1.139 0.9343   | −116.7 | 3  | 239.7  | 1.5276                        | 0.1835         |
|             | WN    | 0.5961        | 1.531           | −147.4 | 2  | 299.0  | 60.78                         | 0.0000         |
| Polynoeoptera| BM   | 0.002121      | 2.759           | −66.26 | 2  | 136.7  | 0.1955                        | 0.2922         |
|             | EB    | 0.002121      | 2.759           | −66.26 | 3  | 138.9  | 2.3961                        | 0.0972         |
|             | delta | 0.001389      | 2.822           | 2.186  | 3  | 136.5  | 0                             | 0.3221         |
|             | SSP   | 0.003247      | 2.812           | 0.00286 | 3  | 137.6  | 1.081                         | 0.1876         |
|             | lambda | 0.002005     | 2.765 0.9636   | −66.22 | 3  | 138.8  | 2.334                         | 0.1003         |
|             | WN    | 0.5465        | 3.045           | −72.66 | 2  | 149.5  | 12.99                         | 0.0005         |
| Palaeoptera | BM    | 0.001485      | 2.918           | −40.18 | 2  | 84.58  | 0                             | 0.3195         |
|             | EB    | 0.002088      | 2.917           | −40.06 | 3  | 86.57  | 1.991                         | 0.1181         |
|             | delta | 0.002322      | 2.938 0.5462   | −39.51 | 3  | 85.46  | 0.8857                        | 0.2052         |
|             | SSP   | 0.001485      | 2.918           | 0.00   | 3  | 86.80  | 2.226                         | 0.1050         |
|             | lambda | 0.00119      | 2.928 0.8993   | −39.30 | 3  | 85.05  | 0.4729                        | 0.2522         |
|             | WN    | 0.7646        | 3.060           | −74.55 | 2  | 153.3  | 68.73                         | 0.0000         |
| Entognatha  | BM    | 0.002414      | 1.074           | −15.71 | 2  | 36.75  | 0                             | 0.5003         |
|             | EB    | 0.01257       | 1.048           | −0.006225 | 15.16 | 3  | 39.31  | 2.561                         | 0.1390         |
|             | delta | 0.002921      | 1.070 0.6378   | −15.58 | 3  | 40.16  | 3.407                         | 0.0911         |
|             | SSP   | 0.002414      | 1.074           | 0.00   | 3  | 40.42  | 3.667                         | 0.0800         |
|             | lambda | 0.002414     | 1.074 1       | −15.71 | 3  | 40.42  | 3.667                         | 0.0800         |
|             | WN    | 1.0335        | 0.888           | −17.23 | 2  | 39.79  | 3.035                         | 0.1097         |

Models and relevant parameters are denoted as follows: BM: Brownian motion (Sigma squared: ML estimate of rate of the underlying size evolution, z₀: ML estimate of value for the root state); EB: Early burst model (a: exponential rate scale for relationship through time); Delta: Pagel’s delta rate change through time model (delta: tree scaling parameter); SSP: Single stable peak Ornstein-Uhlenbeck model with centralizing tendency towards an optimum (alpha: strength of central attraction); lambda: Pagel's lambda measuring deviation of inter-tip covariance matrix from expectations of BM (lambda: multiplication factor applied to the off-diagonal covariance matrix elements maximizing similarity to BM); WN: white noise non-phylogenetic model with all data drawn from a common distribution. Also given are log likelihood values of the observed data (LnLik), number of parameters (k) and AICc values, deviation from optimal model (Delta AICc), and Akaike weights. Models in bold are the favoured models, either by virtue of lowest AICc scores or are those with fewest parameters within 2 AICc units of the lowest AICc scores. *denotes parameters estimated at the bounds placed on the optimization procedure i.e. their actual values may be smaller than given.
dependent generation of variance [50, 51]. This, in combination with the data abstraction required to treat higher taxonomic groups here (see below; [52]), and the fact that our approaches to estimate standard deviation are confounded with clade richness (see methods; [53]), meant that we did not feel secure in pursuing this line of investigation within the current study. However, in the presence of better data, particularly for within clade body size distributions, this is an intriguing concept and one that merits further investigation.

When considering the processes that may underlie the evolution of hexapod body size, our analyses identify Holometabola and in particular Diptera, as having undergone divergent evolutionary processes when compared with the remaining Hexapoda (the latter being dominated by an overall Brownian drift across the phylogeny). None of the explicit process models explored here were recovered as adequate descriptors of what this divergent process may be, although the BAMM analysis of rate heterogeneity suggests a rate acceleration through time may be involved. The (favored; Table 3) lambda model is not in itself a process description, hence this parameter is most commonly described as a test of phylogenetic signal (e.g. [54]). Despite this limitation, we can conceptually distinguish three possible sources of non-phylogenetic signal that may individually or collectively explain the deviation from BM within these clades: random noise in the dataset (e.g. from inadequate descriptive data), phylogenetic error in taxon assignments, and the presence of complex evolutionary processes that are inadequately accommodated within the single parameter extensions of BM examined above.

Focusing on Diptera as the extreme case of divergence from BM (Additional file 1: Table S3), it can be noted that, in comparison with e.g. Lepidoptera, where the majority of large bodied members are restricted to two derived clades (Macroheterocera; "macro-moths", and Rhopalocera; butterflies [55]), large bodied flies occur in basal, (e.g. Tipulidae; crane flies), intermediate (e.g. Asilidae and Mydidae; robber and Mydas flies), and highly derived, phylogenetic positions (e.g. Oestridae; bot flies). Likewise, miniaturization also occurs in a range of unrelated families, e.g. Braulidae (bee lice; approximated mean length = 1.30 mm), Corethrellidae (mean =1.22 mm) and Phoridae (mean =1.75 mm), which collectively may further skew size distributions across the order [56]. Thus, there is the potential for divergent processes of size evolution within the clade that are not fully captured by the simplistic evolutionary models implemented here. However, noise in the dataset e.g. from the use of regional taxonomic descriptions (North and Central America [57–59]) as proxies for global size distributions, and phylogenetic uncertainty in relationships, e.g. within Schizophora [30, 60, 61], mean that we should be cautious of over-interpreting these patterns and await better comparative information, preferably incorporating developmental and larval data [13]. It should also be noted that Diptera, and to a lesser extent all Holometabola are, in terms of proportion of probable species described, less well-known than comparable groups (e.g. Coleoptera, Odonata) [62], and thus may be more strongly impacted by collection and modeling biases outlined below.

The apparent association of Holometabola with accelerating rates of size evolution through time (even if we cannot define the specific underlying model) is interesting given that complete metamorphosis has previously been identified as a key innovation in hexapod diversification.
Plausible mechanisms for a different process of size evolution within the clade include: modularization of life history stages decoupling adult body-size from larval ecology and so permitting greater adaptive flexibility [13, 31], and historical factors relating to the differential extinction of large bodied non-holometabolan groups [19, 63]. There have been various suggestions, based on the small size of early fossil representatives [33], that patterns within Holometabola may follow the widely acknowledged principal known as Cope’s rule, which postulates that increased niche specialization tends to lead to increased body sizes within a clade over evolutionary time [18] (although in hexapods extreme miniaturization is just as much associated with specialization [14, 22]). However, the lack of a joint systematic framework for extant and fossil taxa has restricted formal testing of this assertion in recent fossil compilations (e.g. [64]).

Unlike well-studied vertebrate clades, there is currently no universal reference source for comparative data within Hexapoda, nor of the demographic or ecological information that may aid in interpreting models of size evolution [36, 65]. As a result, the information used here is derived from a mix of global and regional scale data-sets collected at the level of individual clades (Additional file 1: Table S1). This imposes additional assumptions beyond the selection of phylogenetic framework (see discussion of the tree used in [30]) and the use of described species as proxies for total clade richness [66]. There are two major sources of error that may impinge on this analysis and whose extents are problematic to test in the absence of more finely resolved taxonomic data. The first relates to the representative nature of the compiled size limits as accurately reflecting the true size range of studied terminal groups. Due to a lack of data for tropical faunas, the information used here includes an over-reliance on North American, Australian and European taxa, which, due to the presence of a well-known latitudinal cline in insect body size [13], has the potential to bias the raw data on which our findings are based. While acknowledging that such a bias is difficult to explicitly test, we note that previous work has found evidence that regional data for taxonomic groups is predictive of global patterns with respect to hexapod body size [26] and that by combining multiple regional sets we at least attempt to consolidate our size ranges across the known taxonomic range.

A second subtle source of bias originates from the conversion of raw size range data into lognormal distributions that are the source of the parameters used in our modeling procedure. An implicit assumption of using lognormal distributions is that on the logged scale the data is symmetrical around the mean (allowing us to use the observed mean-of-logs as our estimate of average size). However, faunal body size compilations suggest that, with increasing species richness, size distributions becomes increasingly right skewed on the log scale [15], although individual sub-ta-xa often vary in skew independently of the overall fauna [67].

For the global family distributions considered here, available data on size-distributional skew is insufficiently resolved to contribute to the models considered here, and as a result we have elected to retain the explicit linkage between raw observations and parametric descriptors provided by the assumption of log-normality.

Another difficult-to-test but implicit assumption in our work is that the probability of species description within terminal taxa is not itself biased by body size [68–70] or, to put this another way, that the estimates of described species richness for terminal groups are unbiased approximations of their true extant diversity [66]. The problem of acquiring estimates of “true” species richness based on incomplete records of described species is one of the most profound challenges facing work on any diverse clade (see discussions in [66, 71] and references therein). Of the work conducted here, the observed pattern, i.e. a weak and statistically nonsignificant positive correlation is potentially consistent with systematic under description of small bodied species; however, this effect would have to be large in-order to mask any “real” negative relationship present within the group. As with many issues relating to unknowns in the richness of large clades, efforts to integrate global taxonomic databases together with associated rates of species description, synonymy resolution and meta-data such as body size, will go a long way towards characterizing what it is that we still do not know regarding hexapod diversity [21].

In addition to description bias, there are also issues relating to the appropriate partitioning of within tip variance, which here we have treated as arising entirely from taxonomic under-sampling. Thus, the effect that novel species description would have on the estimate of the mean body size of a given clade depends on the number of described species in this clade (hence why the estimate of variance is clade-richness dependent [53]), whereas in reality, such estimates also encompass other sources of error such as length variation among individual specimens [72] and sexual dimorphism [73], which may contribute to variation observed across lineages. Dealing within tip variance in trait measurements is perhaps the greatest outstanding challenge in modeling of trait evolution at deep phylogenetic levels [74]. The methods used here, based on [75, 76], were originally developed with the aim to incorporate measurement error in tip values, with the result that they contain assumptions regarding the distribution of such variance that may not be appropriate for all of the contributing sources of variance present within this dataset. Alternative approaches exist, e.g. “MECCA” [77]; however, these involve simulating multiple species-complete trees (computationally unfeasible on the scale of Hexapoda) and also make strong assumptions regarding variance.
structure within tip taxa. Further work on partitioning variance within phylogenetic models [74], as well as improved understanding in how such variance is structured in groups where there is good phylogenetic information, represents an area of great potential in understanding how trait evolution may be modeled across very large taxonomic groups.

Conclusions
Within the limits of the available data and the neontological approach, our analyses suggest that the evolutionary forces structuring macro-evolutionary patterns of body size within Hexapoda are not simply and directly related to those responsible for structuring the diversity of the group. The overall pattern of body size evolution within the group, based on its extant representatives appears to be broadly driven by essentially neutral forces (at a log scale) with the exception of the poorly defined processes operating within Holometabola and Diptera. This conclusion differs from that of fossil based surveys of the group, which have emphasized constraints in shaping size evolution in hexapods, such as oxygen limitation (e.g. [12, 64]) and the evolution of vertebrate predators (notably birds) [78]. These differences reflect differences in the underlying data, including a focus on the evolution of mean body size within clades as opposed to the limits of its maximum value [64], the inability of analyses based on extant data to take account of no-longer existing diversity [79] and impacts of phylogenetic non-independence, which are often neglected in fossil analyses of hexapods [32].

The consequences of these findings for the standard size paradigm (e.g. [1]), with its emphasis on vertebrates, in which size and richness show a strong degree of coupling [2, 3], are significant in that they attack the universality of these findings to other terrestrial clades [6]. As with any macro-evolutionary study involving incompletely described taxonomic groups, we must pay special attention to the role of missing data and interpolation in defining the observed pattern. Hence here we have attempted at a basic level to incorporate within-tip variance into our discussion of body size and diversification. Great challenges remain in trying to tease apart ecological and evolutionary processes in groups operating on temporal and spatial scales profoundly different from our own. The analysis presented here thus should be taken as a step on the road towards a broader understanding of the processes of size evolution and its consequences for an invertebrate perspective of the natural world.

Methods
An ideal analysis of body size evolution would comprehensively explore patterns and processes at the species level. However, because of the enormous richness of Hexapoda, phylogenetic and trait data are currently too sparse to support a comprehensive species-level analysis. Therefore, for practical reasons we restrict our discussion to the family level, based on recently proposed phylogenetic relationships [30].

All size data for this study is based on family-level estimates of minimum and maximum body length collected from global, regional and taxonomic datasets ([57–59, 80–202], Additional file 1: Table S1). The use of length as a proxy for size is common in Hexapoda due to difficulties in estimating mass from dried museum specimens [13, 15]. Taxon-specific length to mass conversion factors (e.g. [203]) were explored for use in this study and produced qualitatively similar results; however, due to the large amount of uncertainty associated with these values, the presented analyses are restricted to raw length data. Body length was taken as from the anterior margin of the head to the termination of the abdomen, discounting wing cases, abdominal limbs, antennae or cerci where such resolution was available. For taxa such as Lepidoptera (moths) where data-sources record body-size via an alternative metric (e.g. wingspan), average measurements of accompanying illustrations (between one and eight per terminal; selected to encompass the observed diversity) were used to convert these values to body length (examples listed in Additional file 1: Table S1). For Trichoptera (caddis flies), which are typically not illustrated so as to make both the wingspan and body length visible, conversion for the whole order was based on specimens of the various families illustrated in [81].

Estimates of clade richness follow [30]. Resolution of taxonomic conflict is described in Additional file 1: Table S1. In order to avoid issues associated with estimating standard deviation for mono-specific clades (see below) all richness estimates were increased by two for the purposes of modeling relationships. This process is recognized as ad-hoc but regarded as preferable to the loss of phylogenetic information resulting from the exclusion of such lineages. In total, the dataset consisted of 774 terminal taxa spanning all major hexapod lineages (Additional file 1: Table S1).

For modeling purposes, we assumed that, within terminal groups, species conform to a lognormal size-distribution, the parameters of which are estimated from the observed minimum, maximum and richness data. This is a strong assumption, but one conforming to available data regarding hexapod size distributions at the family level [204, 205], and can therefore be regarded as the obvious default in the absence of data to the contrary. The mean of the approximated distributions (henceforth treated on a log scale) was taken as the mean of the log values of the minimum and maximum
size estimates (henceforth mean-of-logs). The standard deviation of approximated distributions was estimated using meta-analysis statistics that assume a sample-size dependent relationship between the estimated sd and the observed range [53]. Thus, for very small clades (<15 taxa) sd was calculated using Equation [16] of [53], for moderately diverse groups (16–70 taxa) sd was estimated as range over four, and for large clades (>70 taxa) sd was estimated as range over six [53]. These procedures assume that the mean values for species rich groups are known with greater accuracy (i.e. have smaller associated variance) than species poor groups with the same size-range, reflecting the fact that the former are less likely to be perturbed by further species description (see Discussion). Given that our estimates of standard deviation are thus dependent on corrected clade richness it is appropriate that we maintain this assumption into the derived estimates of standard error (SE) around the clade specific mean-of-logs values. Hence our SE estimates for modeling evolutionary processes [75] were calculated, under the assumption that sample size was equivalent to corrected clade richness.

Descriptive plots of the observed frequency distribution of size were generated for hexapods as a whole and for the major super-ordinal sub-clades [30, 34, 35]. The normality of the overall mean distributions, both at the level of terminal taxa, and with taxa weighted by their observed species richness (Fig. 1), was assessed using an Agostino test [206] (implemented in R [207]; package moments [208]). The phylogenetic distribution of minimum, maximum and mean body length, as well as the estimates of terminal standard deviation (Fig. 2, Additional file 1: Figure S1) were plotted using a Brownian motion (BM) ancestral reconstruction [209] implemented in the package phytools [210].

The degree of phylogenetic signal present in the data with respect to mean-of-logs size was assessed using Blomberg’s K statistic [211], and by comparing the observed variance among the phylogenetically independent contrasts (PICs) with 1000 randomized data replications, applying the correction of [75] to account for within-group variance (implemented in the package phytools) (Table 1). Blomberg’s K can be visualised as measuring the degree to which an observed dataset converges on the expectations of BM (producing an expected value of 1) [211]. Data with no phylogenetic signal will produce a K value of 0 and values less or greater than 1 should be interpreted as lower or higher than expected similarity among terminal taxa, which can be a manifestation of more complex trait evolutionary processes (see below).

To explore the relationship between diversification and body size, we used an adaptation of the PIC derived “macrocaic” method implemented in the package caper [212], which is optimized to explore associations of traits values and species richness at the level of higher taxa [38–40]. Richness contrasts at each node were standardized using two metrics: relative rate difference (RRD; Table 2, Fig. 3) and proportion dominance index (PDI; \( \frac{(N_1^2 + N_2^2)}{2(N_1 + N_2)} \))-0.5, Additional file 1: Table S2). Size was modeled as the mean-of-logs estimate and the relationship between the two sets of independent contrasts assessed using regression through the origin [39]. To incorporate within-tip variance in size we used a parametric bootstrap, where across 50,000 pseudo-replicated datasets the values of terminal groups were taken as random draws from the estimated terminal distributions (see above) and the 95 % bounds on the relationship between contrasts were estimated. This distribution was compared with that of an identical number of replicated null data samples where terminal size-values were randomized across the tree. Significance was judged on whether the 95 % confidence intervals on the bootstrapped data excluded those of the randomized null data.

To explore the processes responsible for generating the observed size distribution we used a model testing framework: fitContinuous, in the package geiger [213, 214]. Candidate models fitted were: a simple BM process; the early burst model (EB/ACDC), [20, 211] where rates of evolution through time exponentially increase or decrease; the delta model [54], which scales the phylogeny so as to bias the distribution of rates of trait evolution towards either the root or tips; and the SSP model (single stationary peak; modeled as an Ornstein-Uhlenbeck process) [215], which assumes that trait evolution convergences on a single global optimum value (Table 3, Additional file 1: Table S3). All of these models are capable of expressing BM as a special case, resulting from near-zero estimates of the relevant scaling parameters.

In addition, we also fitted two models without an explicit generating process, in order to measure the role of noise and non-phylogenetic signal in the structure of our dataset. The lambda model [54] calculates a global statistic measuring the extent of deviation in the inter-tip covariance matrix from the assumptions of BM (which corresponds to a lambda value of 1). The white noise model (WN) corresponds to a lambda value of 0, and reflects the result that would be obtained in the absence of any phylogenetic structure (star tree) with tip states being drawn from a single underlying normal distribution (Table 3, Additional file 1: Table S3). All fitted models incorporated estimates of standard error around the mean-of-logs, using the methodology of [75] (see above for how these are calculated). Model selection was performed on the basis of AICc values and Akaike weights, see discussion in [20].
Finally, we conducted an exploration of the homogeneity of the process of size evolution within hexapods using the shift-based reversible jump Markov Chain Monte Carlo framework BAMM [216]. As implemented here, the analysis fits EB/ACDC models of size evolution to nodes within the tree signifying regime changes among descendant clades based on an underlying Poisson proposal mechanism. This allows the identification of potential breakpoints in the underlying process of size evolution without the imposition of an explicit prior model. Note that this procedure in its current form is unable to accommodate error in the tip value estimates, thus only the mean-of-log size values for terminal clades were modeled.

Starting values for BAMM were calculated as a homogenous BM process in fitContinuous (betaInit = 0.002424, betaShiftInit = 0), and prior distributions calculated using the package BAMMtools (poissonRatePrior = 1, betaInitPrior = 412.47 betaShiftPrior = 0.002408). We set informative priors on the rate of regime change favoring a homogenous diversification process in order to maximize the credibility of any shifts recovered. Chains were run for 500 million generations with sampling conducted every 5 million generations. Burn-in was estimated based on the stabilization of the inferred likelihood measurements at 10% of the total sample. Adequate sampling of the stable distribution was assessed on the convergence of two independent runs from divergent starting parameters, based on complete overlap of the credible shift set of models accounting for 70% of the overall described likelihood. The results presented here are taken only from the first chain, based on the estimated homogenous BM parameters.

Availability of supporting data
The dataset supporting the results of this article is available in Additional file 1: Table S1.

Additional file

Additional file 1: Figure S1. Phylogenetic plot of (log) size traits. A) log maximum body length; B) log minimum body length. Ancestral reconstruction of internal nodes based on a BM process (ancML) (Revel [209]). Lower bars denote the minimum and maximum values of observed traits (ln (mm)); coloration on a red to blue scale. Terminal bars denote membership of major clades; colors as in Fig. 1. Figure S2. Maximum credible model set from Bayesian Analysis of Macroevolutionary Mixtures (BAMM) corresponding to 95% of the overall model likelihood. Models are listed in order of frequency (f) of obtaining model in the post burnin set corresponding to their inferred probability (listed from top, left to right). Coloration and tree orientation are as in Fig. 4. Table S1. Compiled body length data for included terminal groups with references. Species richness estimates taken from (Rainford et al., [30]); SI. Where multiple references are given they refer respectively to the minimum maximum values. Taxonomic alterations from (Rainford et al., [30]) are listed in notes. Table S2. Outputs of Macroecological analysis of relationship between PIC of diversification rate (measured as PDI) and mean log size for major clades. Table S3. Parameter estimates and relative likelihoods of alternative models of mean body size for major orders of Holometabola (including terminal standard error). Models and parameters denoted as in Table 3. (DOCX 2358 kb)

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
JR conceived and designed the study, collected the data, performed the analysis and wrote the manuscript. MH helped conceive the study, provided technical oversight and resources and helped to draft the manuscript. PM helped conceive the study, participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements
The work was funded by NERC (http://www.nerc.ac.uk) grant NE/I500197/1. The authors would like to acknowledge Dr. Colin Beal for advice and assistance in the implementation of the parametric bootstrap and for helpful discussion of findings.

Author details
1Department of Biology, University of York, Heslington, York YO10 5DD, UK.
2Institute of Biochemistry and Biology, Faculty of Mathematics and Natural Sciences, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany.

Received: 8 September 2015 Accepted: 15 December 2015
Published online: 08 January 2016

References
1. Kozlowski J, Gawelczyk AT. Why are species’ body size distributions usually skewed to the right? Funct Ecol. 2002;16:419–32.
2. Maurer BA. The evolution of body size in birds. I. Evidence for non-random diversification. Evol Ecol. 1998;12:925–34.
3. Gardezi T, da Silva J. Diversity in relation to body size in mammals: A comparative study. Am Nat. 1999;153:110–23.
4. Allen C, Garretstani AS, Havlicek TD, Marquet PA, Peterson GD. Patterns in body size distributions: sifting among alternative hypotheses. Ecol Lett. 2006;9:630–43.
5. Orme CDL, Isaac NJB, Purvis A. Are most species small? Not within species—level phylogenies. Proc R Soc Lond B Biol Sci. 2002;269:1279–87.
6. Orme CDL, Quicke DLJ, Cook JM, Purvis A. Body size does not predict species richness among the metazoan phyla. J Evol Biol. 2002;15:335–47.
7. McClain CR, Boyer AG. Biodiversity and body size are linked across metazoans. Proc R Soc Lond B Biol Sci. 2009;276:2209–15.
8. Hutchinson GE, MacArthur RH. A theoretical ecological model of size distributions among species of animals. Am Nat. 1959;153:117–25.
9. Morse DR, Lawton JH, Dodson MM, Williamson MH. Fractal dimension of vegetation and the distribution of arthropod body lengths. Nature. 1985; 314:731–3.
10. Vogel S. Life in moving fluids: The physical biology of flow. Princeton: Princeton University Press; 1994.
11. Brown JH, Gillooly JF, Allen AP, Savage VM, West GB. Towards a metabolic theory of ecology. Ecology. 2004;85:1771–89.
12. Harrison JF, Kaiser A, VanderBrooks JM. Atmospheric oxygen level and the evolution of insect body size. Proc R Soc B Biol Sci. 2010;277:1937–46.
13. Chown SL, Gaston KJ. Body size variation in insects: a macroecological perspective. Biol Rev. 2010;85:139–69.
14. Davis RB, Óunap E, Javoiowski J, Gawelczyk AT. Why are species’ body size distributions usually skewed to the right? Funct Ecol. 2002;16:419–32.
15. Davis RB, Bessey T, Javoiowski J, Gawelczyk AT. Why are species’ body size distributions usually skewed to the right? Funct Ecol. 2002;16:419–32.
17. Sokolovska N, Rowe L, Johannesson F. Fitness and body size in mature odonates. Ecol Entomol. 2000;25:239–48.
18. Hone DWE, Benton MJ. The evolution of large size: how does Cope’s Rule work? Trends Ecol Evol. 2005;20:4–6.
19. Monroe MJ, Bokma F. Mass extinctions do not explain skew in interspecific body size distributions. J Zool Syst Evol Res. 2013;51:13–8.
20. Harman LI, Losos JB, Jonathan Davies T, Gillespie RG, Gittleman JL, Bryan Jennings W, et al. Early bursts of body size and shape evolution are rare in comparative data. Evolution. 2010;64:2385–96.
21. Mayhew PJ. Why are there so many insect species? Perspectives from fossils and phylegogenies. Biol Rev. 2007;82:425–54.
22. Poliolv AA. Small is beautiful: features of the smallest insects and limits to miniaturization. Annu Rev Entomol. 2015;60:103–21.
23. Poulin R, Morand S. Parasite body size distributions: interpreting patterns of skewness. Int J Parasitol. 1997;27:959–64.
24. Finlay BJ, Thomas JA, McGavin GC, Fenchel T, Clarke RT. Self-similar patterns of insect diversity at local to global scales. Proc R Soc B Biol Sci. 2006;273:1935–41.
25. Katzourakis A, Purvis A, Azmeh S, Rotheray G, Gilbert F. Macroevolution of hoberflies (Diptera: Syrphidae): the effect of using higher-level taxa in studies of biodiversity, and correlates of species richness. J Evol Biol. 2001; 14:219–27.
26. Mifsud B. Diversity of Anisoptera (Odonata): Inferring speciation processes from patterns of morphological diversity. Zool. 2002;105:355–65.
27. Mitter C, Farrell B, Wiegmann B. The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? Am Nat. 1988;132:107–28.
28. Misof B. Diversity of Anisoptera (Odonata): Inferring speciation processes from patterns of morphological diversity. Zool. 2002;105:355–65.
29. Mitter C, Farrell B, Wiegmann B. The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? Am Nat. 1988;132:107–28.
30. Ulrich W. Body weight distributions of European Hymenoptera. Oikos. 2006;114:318–28.
31. Ulrich W. Body weight distributions of central European Coleoptera. Eur J Entomol. 2007;104:75–76.
32. Grimaldi DA, Engel MS. Evolution of the Insects. Cambridge: Cambridge University Press; 2005.
33. Nel A, Roques P, Nel P, Prokin AA, Bourgain T, Prokop J, et al. The earliest known holometabolous insects. Nature. 2013;503:257–61.
34. Trautwein MD, Wiegmann BM, Beutel RG, Kjer KM, Yeates DK. Advances in insect phylogeny at the dawn of the postgenomic era. Annu Rev Entomol. 2012;57:449–68.
35. Mifsud B, Liu S, Meusmann K, Peters RS, Donat A, Mayer C, et al. Phylogenomics resolves the timing and pattern of insect evolution. Science. 2014;346:763–7.
36. Cooper NJ, Purvis A. Body size evolution in mammals: complexity in tempo and mode. Am Nat. 2010;175:272–38.
37. Pagel M. Inferring evolutionary processes from phylogenies. Zool Scr. 1997;26:331–48.
38. Agapow P-M, Isaac NJB. MacroCAIC: revealing correlates of species richness by comparative analysis. Divers Distib. 2002;8:41–3.
39. Isaac NJB, Agapow P-M, Harvey PH, Purvis A. Phylogenetically nested comparisons for testing correlates of species richness: a simulation study of continuous variables. Evolution. 2003;57:178–26.
40. Freckleton RP, Phillimore AB, Pagel M. Relating traits to diversification: a comparative approach. Philos Trans R Soc B Biol Sci. 2001;359:602–7.
41. Pagel M. Inferring the historical patterns of biological evolution. Nature. 1999;401:877–84.
42. Regier JC, Mitter C, Zwick A, Bazinet AL, Cummins MP, Kowahara AY, et al. A large-scale, higher-level, molecular phylogenetic study of the insect order Lepidoptera (moths and butterflies). PLoS One. 2013;8:e58568.
43. Marshall S. Flies: the natural history and diversity of Diptera. New York: Firefly books; 2012.
44. Mc Alpine JF, Peterson BV, Shewell GE, Teskey HJ, Vockerot JR, Wood DM. Manual of nearctic Diptera. Volume 1. Ottawa: Research Branch, Agriculture Canada; 1981.
45. McAlpine JF, Peterson BV, Shewell GE, Teskey HJ, Vockerot JR, Wood DM. Manual of nearctic Diptera. Volume 2. Ottawa: Research Branch, Agriculture Canada; 1987.
46. Brown BV, Borkent A, Cumming JM, Wood DM, Woodley NE, Zumbado M, editors. Manual of Central American Diptera, vol. 1. Ottawa: NRC Research Press; 2009.
47. Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim J-W, Lambkin C, et al. Episodic radiations in the fly tree of life. Proc Natl Acad Sci. 2011;108:5690–5.
48. Atkins J, Friedrich M. Shaking the Diptera tree of life: performance analysis of nuclear and mitochondrial sequence data partitions. Syst Entomol. 2013;38:93–103.
49. Gaston KJ. The magnitude of global insect species richness. Conserv Biol. 1991;5:283–96.
50. Nicholson DB, Ross AJ, Mayhew PJ. Fossil evidence for key innovations in the evolution of insect diversity. Proc R Soc Lond B Biol Sci. 2014;381.
51. Clapham ME, Karr JA. Environmental and biotic controls on the evolutionary history of insect body size. Proc Natl Acad Sci U S A. 2012;109:10927–30.
52. Ingram T, Harmon LJ, Shurin JB. When should we expect early bursts of trait evolution in comparative data? Predictions from an evolutionary food web model. J Evol Biol. 2012;25:1902–70.
53. Costello MJ, Wilson S, Houlding B. Predicting total global species richness using rates of species description and estimates of taxonomic effort. Syst Biol. 2012;61:871–83.
54. Leder N. Insect species body size distributions. PhD dissertation, University of Sheffield, 1997.
55. Gaston KJ. Body size and probability of description: the beetle fauna of Britain. Ecol Entomol. 1991;16:505–8.
56. Blackburn TM, Gaston KJ. The distribution of body sizes of the world’s bird species. Oikos. 1994;70:127–30.
57. Gaston KJ, Blackburn TM. Are newly described bird species small-bodied? Biodivers Lett. 1994;2:16–20.
58. Poulin R. Parasite biodiversity revisited: frontiers and constraints. Int J Parasitol. 2014;44:581–9.
59. Gouves EJ, Gaston KJ, Chown SL. Intraspecific body size frequency distributions of insects. PLoS One. 2011;6:e16660.
60. Cohen JE, Jonsson T, Müller CB, Godfray HCJ, Savage VM. Body sizes of hosts and parasitoids in individual feeding relationships. Proc Natl Acad Sci. 2005;102:6268–9.
61. Revell LJ, Reynolds RG. A new bayesian method for fitting evolutionary models to comparative data with intraspecific variation. Evolution. 2012;66:2697–707.
62. Ives AR, Midford PE, Garland T. Within-species variation and measurement error in phylogenetic comparative methods. Syst Biol. 2007;56:252–70.
63. Felsenstein J. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. Am Nat. 2008;171:713–25.
64. Woodward G, Ebenman B, Emmerson M, Montoya JM, Olesten JM, Valido A, et al. Body size in ecological networks. Trends Ecol Evol. 2005;20:402–9.
65. Dudley R. The biomechanics of insect flight: form, function, evolution. Princeton: Princeton University Press; 2002.
66. Ricklefs RE. Cladogenesis and morphological diversification in passerine birds. Nature. 2004;430:338–41.
67. Purvis A. Evolution: how do characters evolve? Nature. 2004;432.
68. Ricklefs RE. Time, species, and the generation of trait variance in clades. Syst Biol. 2006;55:151–9.
69. Bokma F. Time, species, and separating their effects on trait variance in clades. Syst Biol. 2010;59:602–7.
70. Hozz SP, Dujbergovic B, Hozz I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol. 2005;5:13.
71. Poulin R, Morand S. Parasite body size distributions: interpreting patterns of skewness. Int J Parasitol. 1997;27:959–64.
192. Whiting MF, Whiting AS, Hastriter MW, Dittmar K. A molecular phylogeny of fleas (Insecta: Siphonaptera): origins and host associations. Cladistics. 2008;24:677–707.

193. Wiegmann BM. A phylogenetic revision of the family Atelestidae (Diptera: Empidoidea) and its implications for the origin of the Cyclorrhaphous Diptera, MSc Thesis, University of Maryland; 1989.

194. Wiggins GB. The caddisfly family Phryganeidae (Trichoptera). Toronto: University of Toronto Press; 1998.

195. Wiggins GB. The Kitagamiidae, a family of caddisflies new to North America (Trichoptera). Contrib R ON Mus Zool Palaeontol. 1956;44:1–10.

196. Wilson KD. Dragonfly groups of conservation interest confined to the Oriental Region. https://www.yumpu.com/en/document/view/39638885/dragonfly-groups-of-conservation-interest-confined-asia-dragonfly. Accessed 1st July 2014.

197. Zborowski P, Edwards T. A guide to australian moths. Collingwood: CSIRO Publishing; 2007.

198. Zloty J, Pritchard G. Larvae and adults of Ameletus mayflies (Ephemeroptera Ameletidae) from Alberta. Can Entomol. 1997;129:251–89.

199. Zompro O. Microphasma, a new stick insect genus from Sri Lanka. (Phasmatodea: Pachymorphinae). Entomol J. 1999;109:124–7.

200. Zompro O. New phasmids from Venezuela and Ecuador Phasmatodea Neue Phasmiden aus Venezuela und Ecuador Phasmatodea. Entomol Z. 1998;10811:456–9.

201. Zompro O. Revision of the genera of the Areolatae, including the status of Timema and Agathamera (Insecta, Phasmatodea). Keltern-Weiler: Goecke & Evers; 2004.

202. Wardhaugh CW. Estimation of biomass from body length and width for tropical rainforest canopy invertebrates: invertebrate biomass measures. Aust J Entomol. 2013;52:291–8.

203. Novotny V, Kindlmann P. Distribution of body sizes in arthropod taxa and communities. Oikos. 1996;75:75–82.

204. Hodkinson ID, Casson DS. Patterns within patterns: abundance-size relationships within the Hemiptera of tropical rain forest or why phylogeny matters. Oikos. 2000;88:509–14.

205. Butler MA, King AA. Phylogenetic comparative analysis: a modelling approach for adaptive evolution. Am Nat. 2004;164:683–95.

206. R Development Core Team. R: A Language and Environment for Statistical Computing. 2011.

207. R Development Core Team. R: A Language and Environment for Statistical Computing. 2011.

208. Komsta L, Novomestky F. moments: Moments, cumulants, skewness, kurtosis and related tests. R package version 0.13. 2012.

209. Revell LJ. Two new graphical methods for mapping trait evolution on phylogenies. Methods Ecol Evol. 2013;4:754–9.

210. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 2012;3:217–23.

211. Blomberg SP, Garland T, Ives AR. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution. 2003;57:717–45.

212. Oame CDI, Feebleton RP, Thomas GH, Petzoldt T, Fritz SA, Isaac NB, Pears W. caper: Comparative Analyses of Phylogenetics and Evolution in R. 2012.

213. Harmon LJ, Weir JT, Brock CD, Gloor RE, Challenger W. GEIGER: investigating evolutionary radiations. Bioinformatics. 2008;24:129–31.

214. Pennell MW, Eastman JM, Slater GJ, Brown JW, Uyeda JC, Fitzjohn RG, et al. geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. Bioinformatics. 2014;30:2116–8.

215. Butler MA, King AA. Phylogenetic comparative analysis: a modelling approach for adaptive evolution. An. Nat. 2004;164:683–95.

216. Rabosky DL. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. PloS One. 2014;9:e89543.