Diversification in the Hexapoda: A molecular phylogenetic perspective

James Lewis Rainford

Doctor of Philosophy

University of York

Biology

February 2015

Abstract

Hexapoda (insects and their relatives) comprise over half of all described species, and demonstrates large variation in species richness among major sub-clades. This has led to various hypothesized controls responsible for structuring diversification within hexapod lineages, including morphological key innovations, dietary shifts (in particular plant feeding) and small body size. This thesis explores these ideas in the context of an explicit phylogenetic hypothesis for the group, constructed from published sequence data and literature derived constraints, and dated using a fossil calibrated relaxed molecular clock (Chapter 2). Based on this framework, models of the diversification process identify complete metamorphosis as a likely key innovation in the hexapod radiation, in addition to further up and down shifts in diversification rate responsible for the observed richness distribution (Chapter 3). Analysis also suggests that ideas regarding the role of plant feeding in diversification are related to restricted clade sampling, and a more comprehensive approach recovers no consistent association between particular diets and net diversification rates, in addition to heterogeneity in the age of dietary groups and in transition rates among dietary categories (Chapter 4). Our data also suggests body size evolution in hexapods occurs independently of clade richness, and is broadly dominated by neutral evolution on a log scale (Chapter 5). Thus, this thesis supports some hypotheses regarding controls on insect richness, whilst conflicting with other, well established ideas. It also provides a novel dated phylogenetic framework for further studies of hexapod evolution and identifies several novel directions for research into the origins and development of this diverse and important radiation.

.

List of Tables

List of Figures

“The sciences, each straining in its own direction, have hitherto harmed us little; but some day the piecing together of dissociated knowledge will open up such terrifying vistas of reality, and of our frightful position therein, that we shall either go mad from the revelation or flee from the light into the peace and safety of a new dark age.”

H.P. Lovecraft. The Call of Cthulhu 1928

"It is interesting to contemplate a tangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent upon each other in so complex a manner, have all been produced by laws acting around us. These laws, taken in the largest sense, being Growth with reproduction; Inheritance which is almost implied by reproduction; Variability from the indirect and direct action of the conditions of life, and from use and disuse; a Ratio of Increase so high as to lead to a Struggle for Life, and as a consequence to Natural Selection, entailing Divergence of Character and the Extinction of less improved forms. Thus, from the war of nature, from famine and death, the most exalted object which we are capable of conceiving, namely, the production of the higher animals, directly follows. There is grandeur in this view of life, with its several powers, having been originally breathed by the Creator into a few forms or into one; and that, whilst this planet has gone circling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being evolved."

C. Darwin. On the Origin of Species, 6th Edition, 1872.

Acknowledgments

The work presented in this thesis would not have been possible without the support and assistance of a great many people. First and foremost, thanks must go to my supervisors; Dr. Peter Mayhew, who chaired the project and provided a great deal of intellectual support and oversight to the presented work, and Prof. Michi Hofreiter, who gave invaluable technical support, particularly during the phylogenetic phases of this project. Other people in need of thanks include: the members of my training advisory panel; Prof. Peter Young and Dr. Julia Ferrari, my collaborator in defining fossil calibrations; Dr David Nicholson, my examiners; Dr. Colin Beele and Dr. Gavin Thomas, Nancy Irwin, and the members of the Whole-Organism Biology and Evolution lab discussion-groups, all of whom have provided much helpful commentary and discussion over the last three and a bit years. I would also like to thank the various authors who have provided me with copies of their publications and/or addressed queries regarding software, as well as the editors and anonymous reviewers who contributed to discuss and improve the published sections of this work. A huge thank you to my long-suffering girlfriend Penny Faulkner, who has put up with my antics, proof-read much of this work and provided the love and support I have needed to complete my studies. Similar thanks also go to my parents, sisters, family and friends, as well as to anyone who has in any way helped me to maintain my sanity, through what has, at times, been something of a trying ordeal. Finally I would like to thank you, the reader, for taking the time look through this work, and thereby validating the effort put in to producing this document. This work was funded by a NERC studentship (http://www.nerc.ac.uk) grant NE/J500197/1.

Author's declaration

I herby declare that the results contained within represent my own work and have not been submitted for examination elsewhere. Work from Chapters 2 and 3 has been previously published in PLoS One (Rainford, J.L. et al., 2014. Phylogenetic Distribution of Extant Richness Suggests Metamorphosis Is a Key Innovation Driving Diversification in Insects. *PLoS ONE*, 9(10), p.e109085.). Work from Chapters 4 has been accepted for publication in the American Naturalist under the title "Diet evolution and clade richness in Hexapoda: A phylogenetic study of higher taxa". At the time of writing work from Chapter 5 is in preparation for publication.

1. Introduction

1.1. Why study diversification? Concepts and tools for the study of species richness.

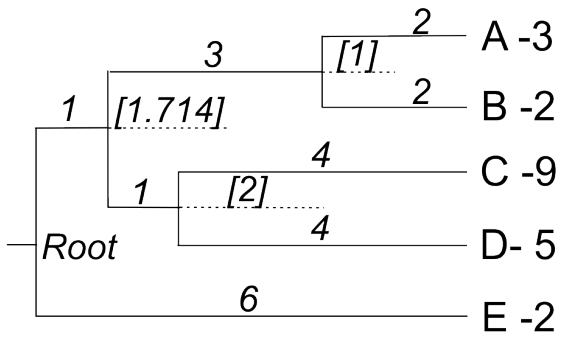
There is perhaps no more fundamental question in evolutionary biology than the issue of why some groups seem to flourish in vast numbers and variety while others languish in obscurity or apparent decline (Coyne & Orr 2004; Benton & Emerson 2007; Losos 2010). It is this question, of what controls observed patterns of species richness across time and space, that has been at the heart of macro-evolutionary studies and as such has shaped much of our views regarding the development of natural systems (Simpson 1967; Stanley 1998; Gaston & Blackburn 2000). One of the most pervasive ideas in macro-evolutionary theory is that the average rates of speciation and extinction among clades can be causally attributed to traits possessed by lineages (Mayhew 2006; Fritz et al. 2013; Morlon 2014). Within this paradigm a distinction is often made between so called key innovations, which may represent unique evolutionary developments in an organism history (Coyne & Orr 2004), and more continuous properties that may have originated multiple times (thus given replicated samples for comparative study) (Mitter et al. 1988; Farrell et al. 1991) or whose effects may be dependent on the trait magnitude (Bokma et al. 2014). In this thesis I will be exploring a number of hypotheses regarding both of these types of trait dependent diversification in the context of the largest terrestrial clade of organisms, the Hexapoda or six legged arthropods.

1.1.1. Sister group comparisons and the role of phylogeny

It has long been recognized that common descent through phylogeny represents a non-independence within the data when comparing lineages, leading to the violation of the assumptions that underpin the majority of statistical tests, including those used in modeling controls on diversification (Felsenstein 1985; Bokma et al. 2014). Comparisons of sister taxa, i.e. between taxa that are each other's closest relative and thus have had identical time for divergence and diversification from their common ancestor are, by definition, statistically independent with respect to phylogeny, assuming of course that such relationships are correctly identified (Felsenstein 1985). Such relationships therefore underpinned early attempts to deal with phylogenetic correlation.

Early sister group comparison methods focused on exploring statistical regularities between the occurrence of traits and relative patterns of extant richness, i.e. whether the presence of a trait is more commonly associated with a given richness pattern than would expected from chance e.g. (Mitter et al. 1988; Farrell et al. 1991; de Queiroz 1998). Such approaches have subsequently been refined by the incorporation of explicit underlying models of the diversification process (Slowinski & Guyer 1989; Paradis 2012)(although see (de Queiroz 1998; Vamosi & Vamosi 2005) for statistical limitations on these approaches) or by the use of non-parametric tests that explicitly incorporate the magnitude of richness divergence into the hypothesis test, e.g. (Wiegmann et al. 1993; Barraclough et al. 1995; Barraclough et al. 1996).

Extending this concept of phylogenetic independence to whole tree problems allows for a more generalized hypothesis-testing framework. Two basic approaches exist to do this, which recent analyses have shown to be statistically equivalent (Blomberg et al. 2012), termed phylogenetic generalized least squares (PGLS) (Martins & Hansen 1997) and phylogenetically independent contrasts (PICs) (Felsenstein 1985; Martins & Hansen 1996). Independent contrasts are a direct extension of standardized sister group comparisons, incorporating correction factors to account for variance in the context of more deeply nested clades (Figure 1) (Garland et al. 1992; Gittleman & Purvis 1998). PGLS, by contrast, refers to the use of phylogenetic distance as a weighting function in the context of a generalized least squares (GLS) linear model (Rohlf 2001). Both methods implicitly assume an underlying model of trait evolution; typically the Brownian motion random walk process (Section 1.1.3.) and both have been widely used to study both the co-evolution of continuous traits on trees and the relationship of these traits to species richness (e.g. (Isaac et al. 2005; Phillimore et al. 2006)).



Contrast	Values	Raw Contrast	SD	Standardised Contrast	Corrected Branch Length	
					Left	Right
A - B	3 - 2	1	2	0.5	2.0	2.0
C - D	9 - 5	4	2.828	1.414	4.0	4.0
AB - CD	2.5 - 7	-4.5	2.646	1.701	[4.0]	[3.0]
ABCD - E	5.071 - 2	3.071	2.952	1.040	[2.714]	[6]

Figure 1 Illustration of Phylogenetically Independent Contrasts (PIC's). Worked example taken from (Garland et al. 2005). Numbers at terminals represent tip states. Numbers above branches give branch lengths and bracket numbers give additional branch lengths added at the relevant nodes to account for deviation within contrasts.

One of the issues with modeling species richness using these techniques is that the generating processes of speciation and extinction are not themselves expected to conform to a BM process. This necessitates the use of descriptive metrics that approximate the overall rate of diversification (see below for further discussion), common examples of which include the total diversification (defined as log species richness) and PIC based measures such as the relative rate difference (RRD), defined as; $ln(N_1/N_2)$ where N_1 is the richness of the descendant clade with larger value of the reference variable and N_2 is the richness of the sister group and the proportion dominance index (PDI) defined as; $(N_1/(N_1 + N_2))-0.5$ (Agapow & Isaac 2002; Isaac et al. 2003).

One property of whole tree comparisons is non-independence between nodes due to nested increases in diversification rate, i.e. that shifts in the diversification process occurring at a particular node may cascade down the topology to produce spurious significant contrasts on more deeply nested nodes (Figure 2) (Davies et al. 2004). This is a particular problem if we are interested in the localization of changes in diversification and its correspondence to particular events inferred in the phylogeny (e.g. the development of key innovations (Bokma et al. 2014)). One way of dealing with this issue is the "trickle

down method" (Davies et al. 2004), whereby significant shifts in richness (based on the Slowinski-Guyer test (Slowinski & Guyer 1989)) are corrected for in the context of more deeply nested comparisons (Figure 2; (Davies et al. 2004)). This approach renders comparisons at each node partially independent; although the result is comparatively crude relative to explicit birth-death models discussed below.

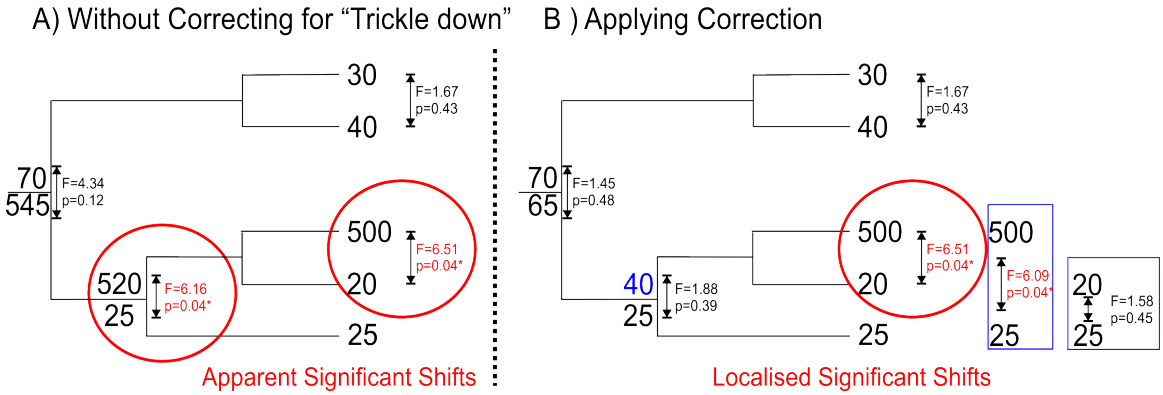


Figure 2 Illustration of the "Trickle down problem" (Davies et al. 2004) and the impact of correction. Tip values are number of species within clades. F and p values for each node calculated using the Slowinski & Guyer (1989) test implemented in ape (Paradis et al. 2004)

1.1.2. Stochastic modeling of diversification rates

As speciation and extinction are essentially random processes in the context of phylogeny they have the potential to be explicitly modeled using a stochastic framework, where the probability of an event on a given branch occurring within a given time interval is defined by the instantaneous rate of each of the process (Nee 2006; Morlon 2014). Of course speciation and extinction also define the shape of the topology, which has led to the development of a number of metrics that characterize the shape of phylogenies and look for imbalance associated with changes in the underlying rate parameters (Mooers & Heard 1997; Chan & Moore 2002; Moore et al. 2004; Freckleton et al. 2008) including the well known, and widely implemented, gamma statistic of (Pybus & Harvey 2000). In the context of the work discussed here, which focuses on patterns observable within topologies of higher taxa (See Section 1.2), these metrics, which are reliant on species complete data, are unsuitable and interested readers are referred to the referenced studies, and to (Quental & Marshall 2010) with respect to the limitations of some of these methods.

Underlying all stochastic models of diversification is the exponential birth process first described by Yule (Yule 1925) sometimes termed the pure birth model (e.g. (Nee 2000; Nee 2006). Under this process species numbers within a tree are expected to grow exponentially through time under the control of a single parameter birth rate (commonly denoted b or λ) that can be interpreted as the instantaneous probability of a given lineage undergoing a speciation event. The value of b can be estimated from the slope of a plot of the (log) lineage accumulation through time (LTT plot) or based on the exponential accumulation of species through time such that the richness of a clade N after a time interval t, $N_{(t)} = N_{(0)}e^{bt}$ where $N_{(0)}$ is the starting number of species within the clade (typically 1)(Nee 2006).

In order to incorporate extinction this model is extended to postulate an instantaneous probability of extinction (d) and, because most diversification questions focus exclusively on extant taxa, involves conditioning on the survival of the observed number of lineages (Nee et al. 1994; Nee 2006) (usually referred to as the "birth death" or "bd" process). For statistical reasons, related to the ease and speed of implementation (see below), parameterization of bd models is usually undertaken using a pair of composite parameters that combine b and d, as these provide a more natural framework for mathematical interpretation (Nee 2006). The first of these (r) represents the difference between the speciation and extinction rates (i.e. b-d) and is commonly referred to as the "net diversification rate". The value of r associated with a particular clade can be estimated from its age and species richness using the method-of-moments approach (Magallon & Sanderson 2001). The second parameter, epsilon (ε), is a dimensionless scaling parameter giving the ratio of extinction to speciation (i.e. d/b) and is various referred to as the "extinction fraction" or the "turnover parameter". Note that because the value of epsilon is more responsive to minor changes in the value of d, the precision on its estimates will usually be lower than that of r (Foote 1988; Alfaro et al. 2009) and can be unreliable in cases of extreme extinction rates (Rabosky 2010).

Given these parameter estimates the likelihood of a clade leaving n descendants after time t (time since divergence of stem group on dated topology), conditional on that the clade does not go extinct (which it does with a probability $1-\varepsilon$), is $Pr(n|t,r,\varepsilon,n>1)=(1-\beta)\beta^{n-1}$ where $\beta=(e^{rt}-1)/(e^{rt}-\varepsilon)$ (Nee et al. 1994). The product of these probabilities across the T terminals of the phylogeny gives the log likelihood of the taxonomic data: $log\ L_T=\Sigma^{T}_{i=1} log(1-\beta_i)+\Sigma^{T}_{i=1}(n_i-1)log\beta_i$ (Rabosky et al. 2007).

What makes this rate-based framework so attractive for macro-evolutionary study is that it can potentially be extended into an almost infinite array of more complex models simply by postulating mechanisms that change the local values of parameters r and ε. As such extensions are defined on relatively simple parameters they remain computationally tractable for typically sized datasets (see (Pyron & Burbrink 2013) and (Morlon 2014) for a reviews of recent methods). Such complex models can potentially include time dependence in diversification rate (Stadler 2011; Morlon et al. 2011; Hallinan 2012; Condamine et al. 2013), hypothesis testing of differential rates in particular clades (Ricklefs 2007; Morlon et al. 2011), trait dependent diversification rates (Maddison et al. 2007; FitzJohn et al. 2009; FitzJohn 2010; Goldberg et al. 2011; Magnuson-Ford & Otto 2012) and optimization of patterns of diversification rates to be maximally explanatory of the observed data without an a-priori hypothesis for the location of rate changes (Alfaro et al. 2009; Rabosky 2014).

As the last methods play a major role in the conclusions of this study I will make a brief review of one of the major algorithms used (the MEDUSA algorithm of (Alfaro et al. 2009)). This procedure, in its simplest form, is a greedy comparative algorithm that progressively compares increasingly complex models to explain the observed pattern of richness within the framework of (corrected) Akaike Information Criterion (AICc). At the first time step the global values for the diversification rate parameters for the observed topology and richness data are estimated (Magallon & Sanderson 2001) and then preceding node by node the optimal position of a break in the diversification model, i.e. where all the descendants of that clade diversify under a different parameter set is identified on the basis of the joint model AICc (Burnham & Anderson 2002)). This procedure is repeated iteratively, adding further process breaks until the global AICc of the joint model fails to increase significantly according to some preset criteria. At this point the joint set of inferred node shifts is outputted as the optimal ML description for diversification within the clade. As a method for the study of diversification MEDUSA (and similar algorithms such as the reversible-jump MCMC approach BAMM (Rabosky 2014)) primarily serve as descriptors of patterns within a given phylogeny, as opposed to testing a preconceived hypothesis, a role in which they have successfully furthered our understanding of many major radiations including vertebrates (Alfaro et al. 2009), plants (Fiz-Palacios et al. 2011) and birds (Jetz et al. 2012).

While providing a wide array of computationally tractable tools the rate based approach to diversification is not without its critics, see discussion in (Rabosky 2009a; Quental & Marshall 2010; Rabosky et al. 2012). The central theme of such criticisms relates to the way in which a naïve rate model, with a positive *r* parameter, will always lead to an exponential increase in clade richness through time, which runs contrary to patterns observed in fossil record (Allen & Gillooly 2006; Alroy et al. 2008; Quental & Marshall 2010) and is incompatible with the idea of niche mediated limits on clade diversity (Ricklefs 2007; McPeek & Brown 2007; Rabosky 2009a). As a result (Rabosky 2009b) postulated an alternative paradigm termed "ecological limits" where clades were defined as demonstrating diversity dependent rates of diversification leading to logistic patterns of clade growth.

In truth however, these apparently divergent world views share many fundamental underlying similarities (Wiens 2011) and it is possible to incorporate the central tenants of "ecological limits", within the rate framework through extension of the *bd* model to incorporate changes in the rates of diversification through time, e.g. (Etienne et al. 2011; Rabosky 2014). Methods that build on these ideas remain in their infancy (and are strongly dependent on the availability of species complete trees (Etienne et al. 2011)) and will not be discussed further in this thesis (see review in (Condamine et al. 2013) and (Morlon 2014)).

1.1.3. Modeling trait evolution

Analogous to the modeling of diversification rate described above there has also been related work in characterizing rates of the evolution of traits that potentially act as co-variants in diversification models (e.g. (Adams et al. 2009)). For discrete trait data common implementations employ a variation of the continuous-time finite-state Markov model (O'Meara 2012) to generate maximum likelihood estimates of the transition rates between different states, e.g. (Pagel 1994; Maddison et al. 2007; Beaulieu et al. 2013), although more complex approaches such as threshold models have recently been developed (Felsenstein 2012), and many studies still make use of traditional parsimony based approaches e.g. (Maddison & Maddison 2011).

By contrast, continuous trait evolution models are generally based around extensions of the multivariate normal distribution model commonly referred to as Brownian motion (BM) (O'Meara 2012), where numerous small, independent and

randomly directed stochastic changes along branches result in an overall effect where the mean expectation of the generated process is invariant through time (i.e. the mean expectation of change is 0) but the associated variance among terminal groups scales with increased the branch length (Felsenstein 1985).

Numerous models have been developed that extend this process to add directionality to trait evolution (Harmon et al. 2010). Some of the most important include: the single stable peak (or SSP) mode; which uses an Ornstein-Uhlenbeck process to model convergence of stochastic process on single trait optima (Butler & King 2004), the early burst model (EB/ACDC); in which the net rate of evolution slows (or increases) exponentially through time as the radiation proceeds (modeled as a BM process with a time dependent dispersion parameter) (Blomberg et al. 2003; Freckleton & Harvey 2006; Harmon et al. 2010), tree scaling procedures; which modify the branch lengths in a topology in-order to make the rescaled tree compatible with BM (e.g. Pagel's delta; which reweights the relative branch lengths of terminal and internal branches in order to model rate change through time, and Pagel's lambda; which accesses convergence with BM without the assumption of a particular generating process (Pagel 1994; Pagel 1999)); and linear trend models; where fossil derived trends are implemented alongside the BM component (Finarelli 2007).

Some of these derivatives, notably OU and EB models have themselves been used as the basis for more complex shift based models where processes are allowed to vary at different regions in the topology (analogous to the MEDUSA algorithm described above) for example (Ingram & Mahler 2013; Thomas & Freckleton 2012; Rabosky 2014). A consequence of this have been a number of surveys interested in joint processes of diversification and trait evolution e.g. (Adams et al. 2009; FitzJohn 2010; Slater et al. 2012) that collectively are beginning to explore the possible linkages between these two aspects of diversity.

1.2. Hexapoda as a model system and hypotheses for diversity in the group

If all mankind were to disappear, the world would regenerate back to the rich state of equilibrium that existed ten thousand years ago. If insects were to vanish, the environment would collapse into chaos. (Common paraphrasing from (Wilson 2010)).

The above sentiment neatly encapsulates the central role played by hexapods in the structure of terrestrial and freshwater ecosystems. As the most species rich invertebrate lineage, hexapods provide much of the vital linkage between the macro-scale world inhabited by higher plants and vertebrates, including ourselves, and the micro-world of fungi and bacteria and other groups that provide the nutrient basis on which all life on land depends (Grimaldi & Engel 2005). Here at the heart of terrestrial ecosystems hexapods have radiated into a truly astonishing array of forms and ecologies that collectively outstrip almost any other major organismal group (Mayhew 2007).

So vast is the hexapod radiation that its limits can only be guessed at, as a huge proportion of the total richness of the group (perhaps as much as 70%) remains un-described (May 1988; Hamilton et al. 2010; Mora et al. 2011). Estimates of the total richness of the clade have included numbers up to 30 million species globally (Erwin 1982), although subsequent refinements of the assumptions underpinning such estimates suggest something closer to modal estimates of 3.7 million *(*90% *CIs;* 2.0-7.4 million) or 2.5 million (1.1-5.4 million) depending on the precise model used (Hamilton et al. 2010). Estimates of total number of described insect species are also variable with typical values including 855,000 (May 2000), 926,400 (Grimaldi & Engel 2005) and 1,049,000 (Compiled estimates used in this study; Appendix 7.1, includes 15,231 species belong to clades not present on the tree described in Chapter 2). These differences reflect a combination of novel description, revision of synonymous names in available catalogues and improved estimation techniques.

What is of great macro-evolutionary interest is the manner in which these species are distributed among the major clades of the group (Mayhew 2007). Across the hexapod orders there exist differences in extant richness spanning four orders of magnitude, for example the orders Zoraptera (“angel insects”) and Coleoptera (beetles) have respectively 32 and 350,000 extant described species (Grimaldi & Engel 2005). Coleoptera are one of

four huge orders that collectively comprise almost 80% of all hexapod species (Grimaldi & Engel 2005) all of which are members of the vast clade Holometabola (the other three are Diptera (flies), Hymenoptera (wasps) and Lepidoptera (moths and butterflies)) (see Section 1.4). Given such major differences in richness among the higher taxonomic groups it is perhaps unsurprising that numerous causal explanations for richness patterns within Hexapoda have been proposed throughout the study of the group (reviewed in (Mayhew 2007)). Sets of these ideas can be categorized into three forms; key morphological innovations, ecological co-evolution and other traits that are not necessarily specific to hexapods (for example body size, see below). A fourth category, the role of historical contingency and the impact of past events such as mass extinctions, will be touched upon but is not explicitly tested here (interested readers are referred to (Nicholson 2012) and references therein).

1.2.1. Key morphological innovations

Of the various hypotheses accounting for the observed species richness across hexapod groups perhaps the most attention has been paid to the idea that there are particular morphological transitions within the group's history that have ultimately acted as drivers for species richness within the group (Mayhew 2007). Four innovations seen as of particular importance include the origins of: the insect body plan, flight, the capacity to fold the wings when not in use, and complete metamorphosis (Figure 3)(Carpenter 1953; de Queiroz 1998; Yang 2001; Dudley 2002; Mayhew 2002; Mayhew 2003; Grimaldi & Engel 2005; Mayhew 2007; Davis et al. 2010a). These ideas have rarely been explicitly linked with causal mechanisms on how traits might impact diversity. However some discussion has been made regarding the linkage between flight and dispersal (Mitterboeck 2012; Ikeda et al. 2012), and there are widely discussed ideas relating (complete) metamorphosis to increased ecological partitioning via the separation of larval and adult ecological niches (Yang 2001; Grimaldi & Engel 2005; Mayhew 2007).

The importance of these ideas in shaping ideas on hexapod richness has led to their explicit testing in the context of the then current ordinal hexapod phylogeny (Mayhew 2002; Mayhew 2003; Davis et al. 2010a). These studies reveal support for the idea that for the capacity the fold the wings flat across the abdomen when at rest, synapomorphic of the clade Neoptera, resulted in a major upshift in diversification rate followed by downshifts on particularly species deficient groups including the relic order Neuropterida and

Zoraptera (Davis et al. 2010a). In the years since this work was undertaken our understanding of the hexapod phylogeny has advanced dramatically (Section 1.4) and as a consequence one of the goals of this work was thus to expand on these ideas in the context of an improved phylogeny and provide a novel test of which events are the key innovations controlling richness within the group (Chapter 3).

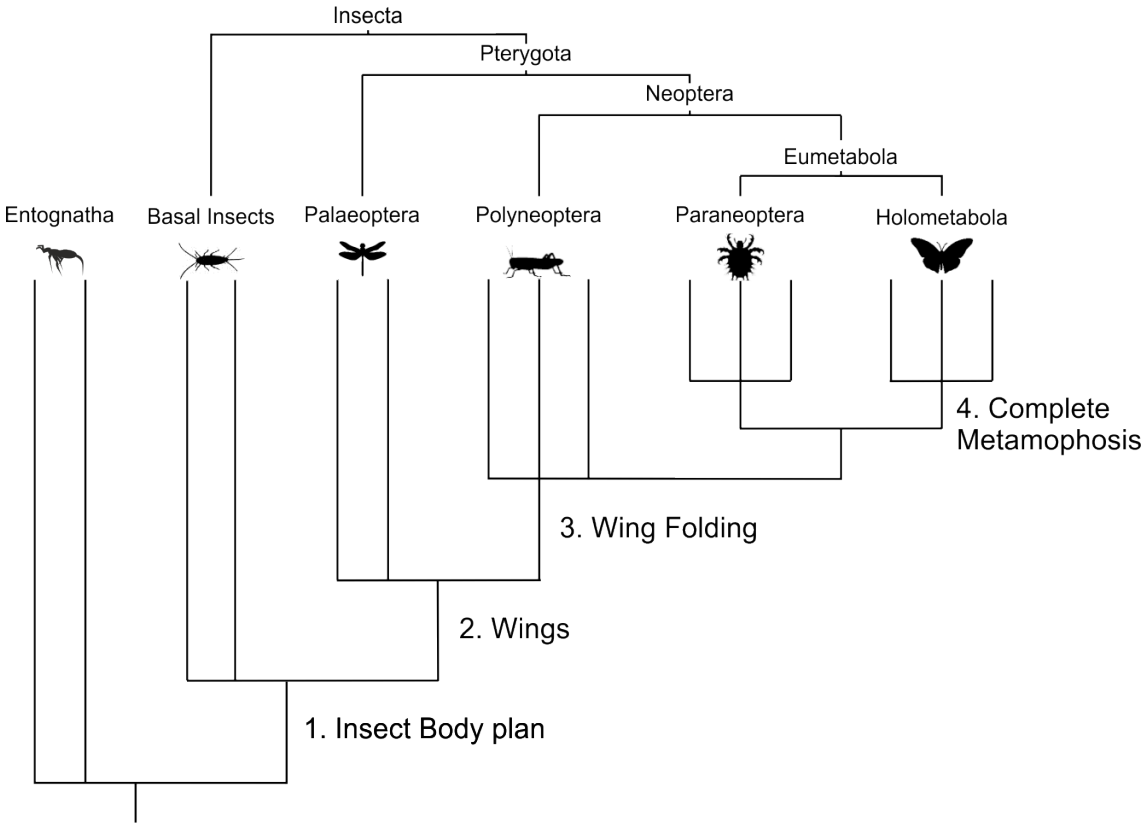


Figure 3 Key morphological innovations proposed to have shaped hexapod diversification (Mayhew 2007; Davis et al 2010a). Redrawn from (Nicholson et al. 2014)

1.2.2. Ecological innovations and the role of co-evolution

While morphological key innovations have been dominant in discussions regarding the species richness patterns across hexapod orders, within these clades ecological opportunity and associated patterns of co-radiation are generally thought to be responsible for shaping observed patterns of species richness. Ideas regarding co-evolution as a driving force behind hexapod species richness originate in an influential paper by Ehrlich & Raven (1964), where the authors postulate that observed relationships between species richness of caterpillars and their host plants could be attributable to antagonistic co-evolution between these groups. Since then co-evolution, particularly with angiosperms, has featured

prominently in various hypotheses regarding controls on diversity both across Hexapoda as a whole (Mitter et al. 1988; Winkler & Mitter 2008) and within Coleoptera (Farrell 1998; Hunt et al. 2007) (the latter including a number of hyper-diverse plant feeding lineages, see Section 1.4.7.4).

The idea that angiosperm diversification drove that of hexapods, many of which are highly specialized in terms of their host selection, has great intuitive appeal (Grimaldi & Engel 2005). However actual evidence that plant feeding promotes richness within the clade remains thin on the ground (see Chapter 4). Some authors have taken the next logical step and questioned as to whether co-evolution with a host represents a more general property in species-rich clades (Futuyma & Moreno 1988) although to date the groups in which this has been explored, notably carnivorous parasitoids, do not appear to show consistent patterns of host-driven diversification (Wiegmann et al. 1993; Stireman 2005). Testing these ideas in the context of our current understanding of hexapod ecology and phylogeny represents the second major aspect of this work.

1.2.3. Broad ecological trends and opportunities for radiation

The third theme to be explored in thesis is not centrally focused on Hexapoda but instead reflects more general patterns in macroevolution that I aim to test in the context of my study system. Body size is of fundamental importance in understanding an organism's ecology in that it is correlated with a huge number of other life history traits (Chown & Gaston 2010; Gaston & Chown 2013). Based on studies of vertebrate clades, many of which show a pronounced skew in their overall size distribution even on a log scale (Maurer 1998; Gardezi & Silva 1999; Albert & Johnson 2012), a number of general mechanisms linking size and diversification have been proposed (see review in (Gardezi & Silva 1999; Allen et al. 2006)). The universality of such mechanisms and their application to non-vertebrate clades has however been challenged, e.g. (Orme, Isaac, et al. 2002; Orme, Quicke, et al. 2002). In Chapter 5 I take a novel compilation of body size data for hexapod groups in combination with the inferred tree (Chapter 2) to explore the relationships with diversification in Hexapoda and to describe the patterns that may be responsible for structuring trait evolution in the group.

1.3. Molecular phylogenies for the study of diversification: advantages and pitfalls

A phylogeny describes the pattern of relationships among a set of taxa and, in modern forms, is based on a specific hypothesis of character distribution. In principle the origins of phylogenetic characters are of no importance, phylogenies can be constructed from morphological, molecular and even ecological datasets. However in the context of the described work molecular systematics presents a distinctive set of opportunities and challenges which will now be briefly reviewed (see also (Sanderson & Shaffer 2002; Baldauf 2003)).

The greatest advantage that molecular data provides for phylogenetic analysis is the relative ease with which large character sets can be collected. As well as the ever-declining cost of sequencing technologies (Mardis 2013) a major part of this ease is the availability of public databases, such as Genbank and its mirror servers (Benson et al. 2013), that store available sequence data. This means that assembling datasets even for large clades such as Hexapoda is a relatively simple matter of defining appropriate search criteria and extracting suitable sequences for the taxa of interest (Altschul et al. 1990). However drawing, as is done here (Chapter 2), entirely on such databases entails certain problems, in particular that the availability of markers across taxa is dependent on the patterns of sequencing within groups, and as such different patterns of sequencing across different lineages can lead to the addition of considerable quantities of missing data into the combined dataset (Sanderson & Driskell 2003) (Section 1.3.4). In addition, in conducting broad taxonomic surveys one is often restricted to a limited set of widely sampled markers, for example the barcoding gene COI and 18S rRNA (see Chapter 2). Some of these have un-desirable properties such as alignment difficulties or susceptibility to saturation (Sections 1.3.1 and 1.3.2). In the following section I describe the basic protocols underpinning the construction, from molecular data, of phylogenies for large clades, focusing on the particular challenges offered by this data source and the potential impacts of missing data in the context of species rich groups.

1.3.1. Alignment and defining homology

The price one pays for the ease with which molecular data can be collected is a level of ambiguity with respect to the quality of the hypotheses of character distributions

that arise from such datasets (Wheeler 2008). A fundamental property of characters used for phylogenetic inference is homology, that is the idea that examined states at a given character, or position within a sequence, are comparable because they reflect a common evolutionary origin (Phillips et al. 2000; Lee 2001). When dealing with morphological data, a strong statement of homology can be obtained from detailed examination of the fine structure of the feature in question and its ontogeny during the course of growth (Beutel et al. 2011; Friedrich et al. 2014). Molecular data, being comprised of strings of otherwise identical chemical bases, is altogether more ambiguous and requires alternative approaches to defining homologous states in cases where there is no direct matching among taxa, collectively referred to as alignment procedures (Wheeler 2008).

The most common approach to homology in molecular data is to treat the issue as a mathematical optimization problem i.e. defining a criteria or score (usually based on the number of base matches and the distribution of gap states- "indels" (Giribet & Wheeler 1999)) against which the quality of alignment is measured and then using iterative computation functions to minimize this cost function. For a given alignment and gap cost function, there exists an exact optimal solution for a multiple sequence alignment defined by dynamic programing; however generating such a solution is an NP-complete problem (i.e. one which cannot be solved in polynomial time) and scales exponentially in time on the length and number of sequences in the combined alignment (Needleman & Wunsch 1970; Lee et al. 2002). As a result most commonly implemented alignment procedures such as the well-known Clustal algorithm (Chenna et al. 2003; Larkin et al. 2007), and its improvements such as Muscle (Edgar 2004) or MAFFT (Katoh et al. 2002) represent heuristic approximations of this exact solution on which they converge to varying degrees, with a key factor being the potential for iterative improvements in estimation of alignment parameters (Edgar & Batzoglou 2006).

Alternative alignment procedures rely on the use of a reference database to identify conserved features of particular sequences and use these as landmarks to align novel targets against this set. Such approaches are obviously only suitable for molecules where structure is highly conserved across deep phylogenetic splits, for example the stem-loop structure of rRNA (although see (Letsch et al. 2010)), as implemented in SILVA (Pruesse et al. 2007). Alignments can also be defined using global properties of the estimated molecule, for example its minimum energy state, e.g. RNAalifold (Bernhart et al. 2008),

although such approaches are inherently dependent on the quality of the molecular model used, which can be problematic for highly divergent sequences.

Finally there is concept of "direct optimization" (Wheeler 1995; Terry 2003; Kjer et al. 2007) wherein the alignment of a data matrix and the resulting phylogenetic tree are jointly optimized with reciprocal feedback acting to improve each step (Ogden & Rosenberg 2007; Wheeler et al. 2006). Such approaches are rarely implemented in recent studies due to their reliance on explicit models of gap evolution, which may fail to adequately characterize the complexities involved of deletion and replication of sequence information, and the fact that available implementations remain restricted to parsimony (Kjer et al. 2007; Simmons et al. 2010). However such procedures have played an important historic role in our understanding of hexapod relationships, e.g. (Terry & Whiting 2005; Yoshizawa 2010), the consequences of which are discussed below (Section 1.4.5.1).

1.3.2. Phylogenetic inference and model selection

The establishment of alignment represents only the first stage of phylogenetic inference from molecular data. The fundamental procedures of phylogenetic inference are identical regardless of the source of character information and entail defining a criterion against which to optimize a given character set and then using heuristic tree search algorithms to identify the optimal topology describing the relationships among included taxa (Swofford et al. 1996). Common optimization criterion include minimizing distances within a similarity matrix (as in neighbor joining (NJ) and other distance based methods), minimizing the number of inferred character transitions (maximum parsimony) or maximizing the joint likelihood of an explicit model of character evolution (maximum likelihood or ML methods) (see (Swofford et al. 1996) and (Holder & Lewis 2003) for reviews). ML approaches are strongly associated with molecular data due to the fact that, unlike morphological state transitions, which usually result from strong directional and idiosyncratic selection (and therefore requires complex, parameter rich models), it is intuitively reasonable to treat, essentially random, molecular base substitutions in a stochastic probabilistic framework that lends itself to explicit models of sequence evolution (introductory discussion in (Swofford et al. 1996) and (O'Meara 2012)).

In some cases model based frameworks have been extended to consider not only the optimal parameter values defining the ML solution but to also sample the shape and

structure of the underlying probability distributions. This concept of quantifying the uncertainty in parameter estimates underpins the Bayesian approches to phylogenetics, which typically rely on the use of Markov Chain Monte Carlo (MCMC) samplers to define the shape of the probability distributions underlying the parameters involved in modeling sequence evolution in order to give an overall distibution of likelihood states associated with the data (Huelsenbeck et al. 2001; Holder & Lewis 2003). Numerous attempts have been made compare ML and Bayesian approches to phylogenetic inference with key areas of dispute including; the suitability of Bayesian partition frequencies (approximating the posterior probability of the inferred model) as measures of clade support as opposed to the more conventional use of bootstrap pseudo-replication (Alfaro et al. 2003; Douady et al. 2003; Simmons et al. 2004), the role of the prior probability distribution set by the user in determining the data outputs (Pickett & Randle 2005; Alfaro & Holder 2006) and the problems of determining if adequate sampling of the tree space has occurred via the MCMC algorithm (Randle et al. 2005). Dealing with the various issues arising from such comparisons is beyond the scope of this review (see (Randle et al. 2005) for a summary), and instead we acknowledge the various strengths of these different procedures for conducting particular tasks. For example, molecular clock dating (Section 1.3.3) where uncertainty in the ages of calibration points is a key component of the analysis is most well suited to the Bayesian framework, while ML approaches may be more computationally efficient in producing point estimates for particular topologies.

Given the reliance on explicit models the assertion of model appropriateness of a given dataset is of key importance in molecular phylogenetic analyses, with the standard approach (as implemented in programs such as PAML (Yang 2007) or jModeltest (Posada 2008)) being to construct an estimated working tree (often the NJ tree) and then fit different models of character evolution to the tree and dataset comparing the inferred likelihoods and selecting the optimal model for more in depth analysis (Sullivan & Joyce 2005). While this procedure has many advantages in terms of computation efficiency it does open up issues of model misspecification which has been shown to be a source of error in phylogenetic inference (Lemmon & Moriarty 2004; Brown & Lemmon 2007).

One of the ways in which model misspecification can manifest itself is associated with the issue of sequence saturation (Philippe & Forterre 1999; Philippe et al. 2011). As noted previously the models for the process of base transition in molecular data are probabilistic process of random mutation, as a simplification from the processes of

selection, drift and mutation acting within real populations (Swofford et al. 1996). Saturation occurs where, over the divergence time scales being considered in a phylogeny, this random process of inferred mutation is sufficiently rapid as to obscure any signal of relationships present within the dataset (Figure 4) (Ho & Jermiin 2004; Jeffroy et al. 2006). Modeling hidden substitution events is one of the major reasons for favoring ML approaches for molecular data (Swofford et al. 1996); however even allowing for this the majority of implemented models will produce misleading results in the face of truly saturated sequence data (Suzuki et al. 2002), leading to errors analogous to the well-known issue of long branch attraction (Swofford et al. 1996; Bergsten 2005).

The identification of partitions subject to sequence saturation relies on the calculation of indices of which the most well-known is the entropy based index of (Xia et al. 2003), which in effect measures in the information content of sequence data relative to a randomized series (see also (Xia & Xie 2001)). Given that, by definition, sequences subject to significant saturation contain no useable phylogenetic information, (although they can inform parameter values within the estimated model), common practice is to remove such partitions from consideration within the analysis, for example by dropping the less constrained third codon position from the consideration of protein coding sequences. Some authors have suggested that such draconian methods are overly harsh in terms of the loss of information and have instead advocated compromises, such as RY coding to remove the most common forms of base transition (Phillips et al. 2004; Ishikawa et al. 2012), although these have seen little use due to limited implementation in widely used phylogenetic software.

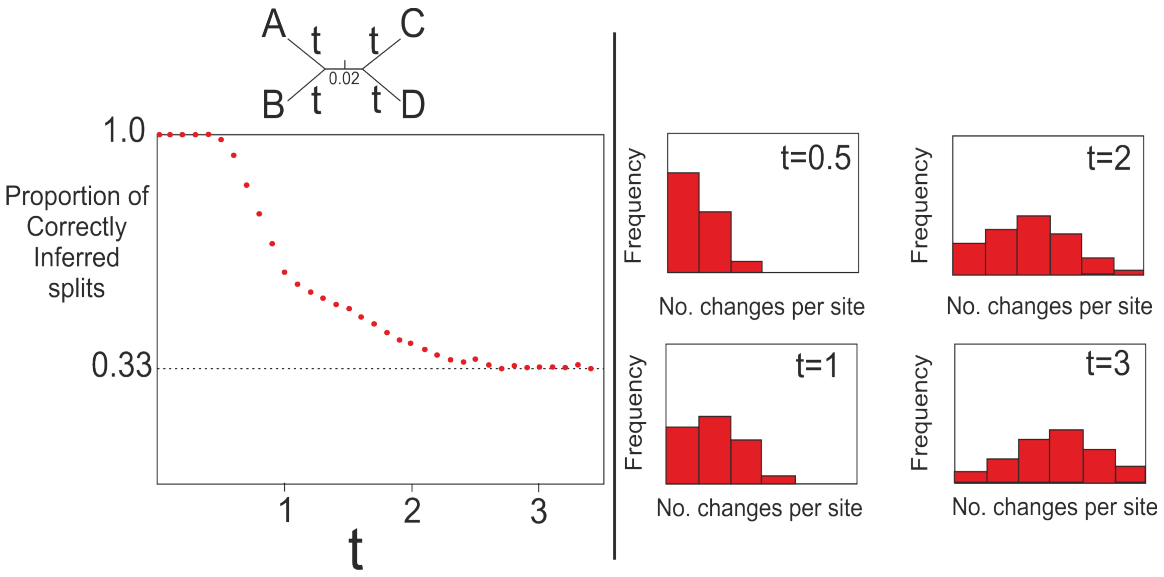


Figure 4 Figurative description of the problem of loss of historical signal (sequence saturation) based on the simulation analysis of Ho & Jermiin (2004). Left panel; shows the reduction in probability of the correct inference of the tree ((A,B),(C,D)); under random site mutation with increasing terminal branch length (*t*-measured in average no. of substitutions). Depicted curve is based on the ML case described in Ho & Jermiin (2004). Dashed line at 0.33 demonstrates the point where split inference among the three possible trees is equivalent to random choice. Right panel; shows that the failure to recover the correct tree at high branch lengths is driven by an increasing tendency for multiple changes per site which are problematic to reconstruct from extant data due to the masking effect of latter substitutions. Discussion of these findings and other sources of model misspecification in (Ho & Jermiin 2004).

1.3.3. Dating using the molecular clock

The inherent randomness of the underlying process of sequence evolution is also harnessed in one of the most important uses of phylogenetic molecular data: the molecular clock and its derivitives. The fundamental concept underpining molecular clock divergence time estimates is the recognition that random stochastic base mutations can be appropriately described by an average rate over time (Zuckerkandl & Pauling 1962; Bromham & Penny 2003) and that by calibrating this rate on known splits we can infer divergence times among taxa beyond the bounds imposed by an incomplete fossil record (Donoghue & Benton 2007).

In reality of course, this model of a simple universal clock has been found to be inadequate, as taxa undergo different intrinsic rates of molecular evolution that must be accounted for in any attempt to model the ages of divergence within the tree (Li & Tanimura 1987; Bromham & Penny 2003). This conundrum has resulted in the development of so called relaxed clock methods (Drummond et al. 2006) that impose a rate evolution process onto the estimate of the molecular clock. Common implementations of such ideas include allowing rates to undergo a Brownian motion like drift from the ancestral rate across each node in the tree (Thorne & Kishino 2002), imposing a process of stochastic rate shifts defined by an underlying compound Poisson process (Huelsenbeck et al. 2000; Drummond & Suchard 2010), or simply assuming that rates are drawn from an underlying mathematical distribution (typically a gamma distribution due to its combination of flexibility and computation tractability) (Lepage et al. 2007). In each case the hyper-parameters of rate evolution process must themselves be estimated from the data, a situation that can in some cases lead to problematic interactions with transition rate matrices and partition scaling factors within highly parameterized Bayesian models (Ronquist et al. 2012).

Very recently there has been a move to further extend relaxed clock approaches to directly incorporate morphological evolution (and thus fossil data –see below) an idea known as the total evidence clock (Ronquist et al. 2012). Despite their increasing popularity there remains some need for caution in interpreting these approaches, as they tend to be extremely parameter rich and thus particularly vulnerable to parameter interactions within the complex modeling framework.

The use of fossil data to calibrate molecular clock studies remains one of the most contentious issues in dating groups with an inadequate fossil record (Donoghue & Benton 2007). Leaving aside, for the moment, issues of data quality and assuring an explicit link between a numerical calibration and the specimen on which it is based (Gandolfo et al. 2008; Ksepka et al. 2011; Parham et al. 2012), most of the contentious issues in calibration relate to the manner in which such data are used to constrain the ages of nodes within the tree (Ho & Phillips 2009). With the exception of total evidence dating described above, where the process of morphological character evolution is inherently built into the estimation of rates, procedures for incorporating fossil calibrations can be generalized as the imposition of a mathematic distribution on the age of one or more nodes within the

topology, the structure of which is informed by the available fossil data (Ho & Phillips 2009) (examples in Figure 5).

The simplest node calibration procedure is to treat the age of a given fossil as the age of divergence of a particular node, the so called point calibration approach (Ho & Phillips 2009). Such protocols formed the basis for the earliest molecular clock analyses, but are now rarely used due to the recognition that the true divergence of lineages must always exceed the earliest fossil representative (Graur & Martin 2004). However, see (Graur & Martin 2004) and (Ho 2007) for commentary on the equally problematic but widespread use of secondary calibrations from broad scale clock studies as point calibrations within nested clades.

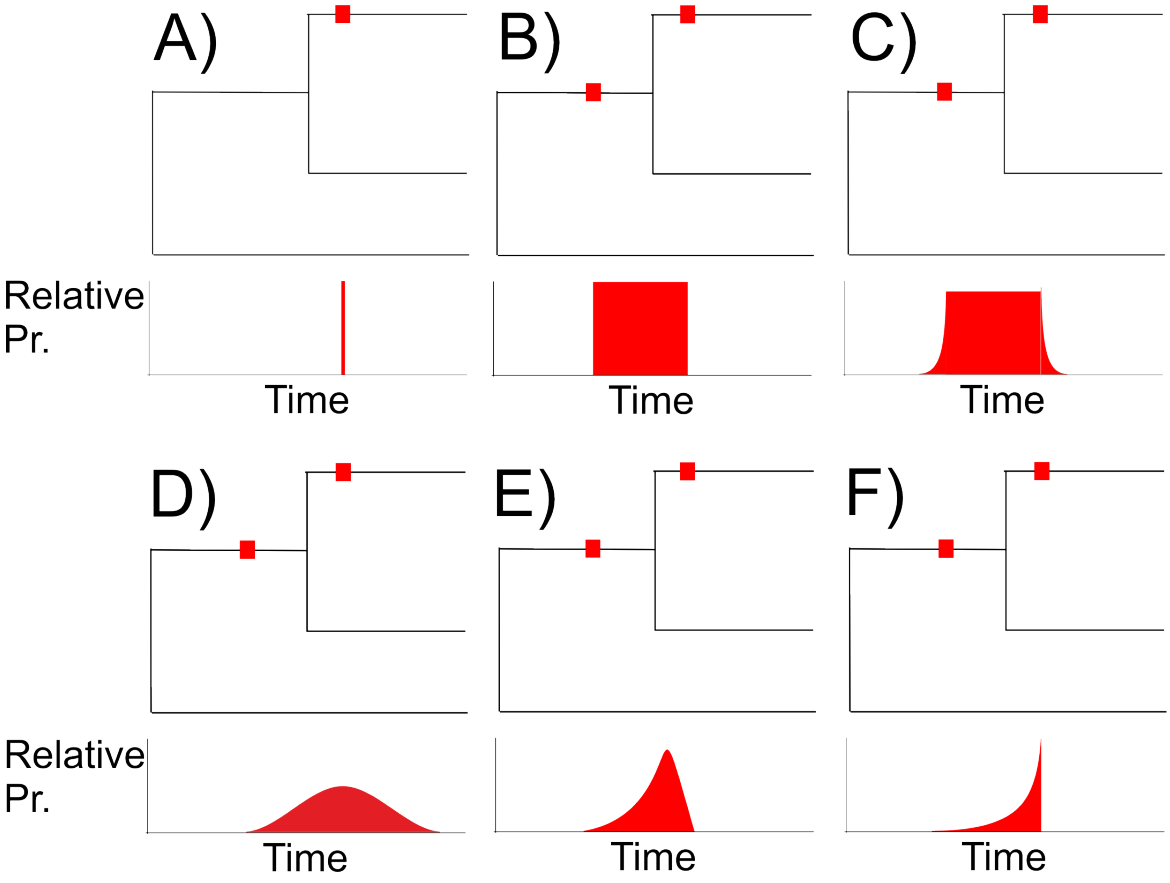


Figure 5 Probability densities on age of focal node given different implemented priors for fossil data (illustrated by red boxes on branches); A) Point calibration, B) Hard maximum and minimum bounds, C) "Soft bounds", D) Normal distribution, E) Log-normal distribution, F) Exponential distribution. See Section 1.3.3. for discussion.

Other approaches to calibration can be thought of as extensions of such bounded intervals to incorporate explicit probability distributions. For example 'soft bounds' work by enforcing a 95% probability that the age of a given node lies within a bounded interval, with the remaining probability treated as exponential functions associated with the interval margins (Figure 5) (Yang & Rannala 2006). Alternatively one can use conventional probability distributions such as the normal, gamma and log normal distributions (Ho & Phillips 2009), although in all cases there is a reliance on good fossil data to justify the explicit bounds placed on the node distributions (Nowak et al. 2013). As yet, while there is growing consensus regarding best practice and increased appreciation of the role calibration can play in determining the outcome of molecular clock analyses, e.g. (Sauquet et al. 2012), the implementation of calibration in a given analyses remain determined by the nature of the data at hand and are subject to the limitations in implementation within available software packages.

1.3.4. Missing data

All stages of phylogenetic analysis, including use of the molecular clock, are potentially impacted by the presence of missing data within the studied matrix. As noted above missing data in phylogenetics can arise from several sources, including sampling incongruence across the sequenced markers, alignment issues and the procedures used to generate sequence data (e.g. expressed sequence tags; EST data) (Hartmann & Vision 2008). Missing data has the potential to impact on phylogenetic inference in a number of ways including; statistically compromising of the underlying models, contributing to model misspecification and, via lack of overlap between taxa within the data matrix, contributing to ambiguous placements within the topology (Hartmann & Vision 2008).

The precise role of missing data, and philosophies on how to deal with its presence, has a long history in phylogenetic studies (reviewed in (Wiens 2003a; Wiens 2006)) and recently there have been a number of simulation studies that have explored its effects in the context of multi-gene molecular datasets e.g. (Wiens 2003b; Wiens 2005; Sanderson et al. 2010). The broad consensus of this work has been that the proportion of missing data present within the data matrix has a relatively limited effect on the quality of inference so long as the total amount of data available is sufficiently large (Wiens 2003a; Wiens 2003b; Philippe et al. 2004) (although see (Lemmon et al. 2009; Wiens & Morrill 2011; Roure et al. 2012) regarding the role of introduced non-parsimony informative states). The

structuring of missing data within the matrix is however recognized as significant, with randomized additions of missing data (e.g. due to alignment issues) being regarded as less problematic than missing gene partitions, which are themselves less of an issue than the missing data associated with the partial overlapping fragments typical of EST data (Hartmann & Vision 2008; Sanderson et al. 2010). Partly as a result of these findings numerous studies have now been conducted on large taxonomic groups with high proportions of missing data with little apparent loss of phylogenetic accuracy e.g. (Wiens et al. 2005; Burleigh et al. 2009; Cho et al. 2011).

The impact of missing data on molecular clock estimates is less clearly understood and again appears to be strongly dependent on the structure of the missing information (Douzery et al. 2004), and its interactions with model misspecification and sequence saturation (Soubrier et al. 2012; Zheng et al. 2011). There is evidence to suggest that relaxed Bayesian procedures are relatively insensitive to issues of data composition compared with alternative dating protocols (Mulcahy et al. 2012), although it is unclear how this fits within the generally accepted sensitivity of Bayesian procedures to model misspecification (Lemmon & Moriarty 2004).

1.4. An introduction to the phylogeny of Hexapoda

This section documents the current understanding of the hexapod phylogeny and outstanding issues in relationships within the group (based, in part, on (Trautwein et al. 2012) and (Yeates et al. 2012)).

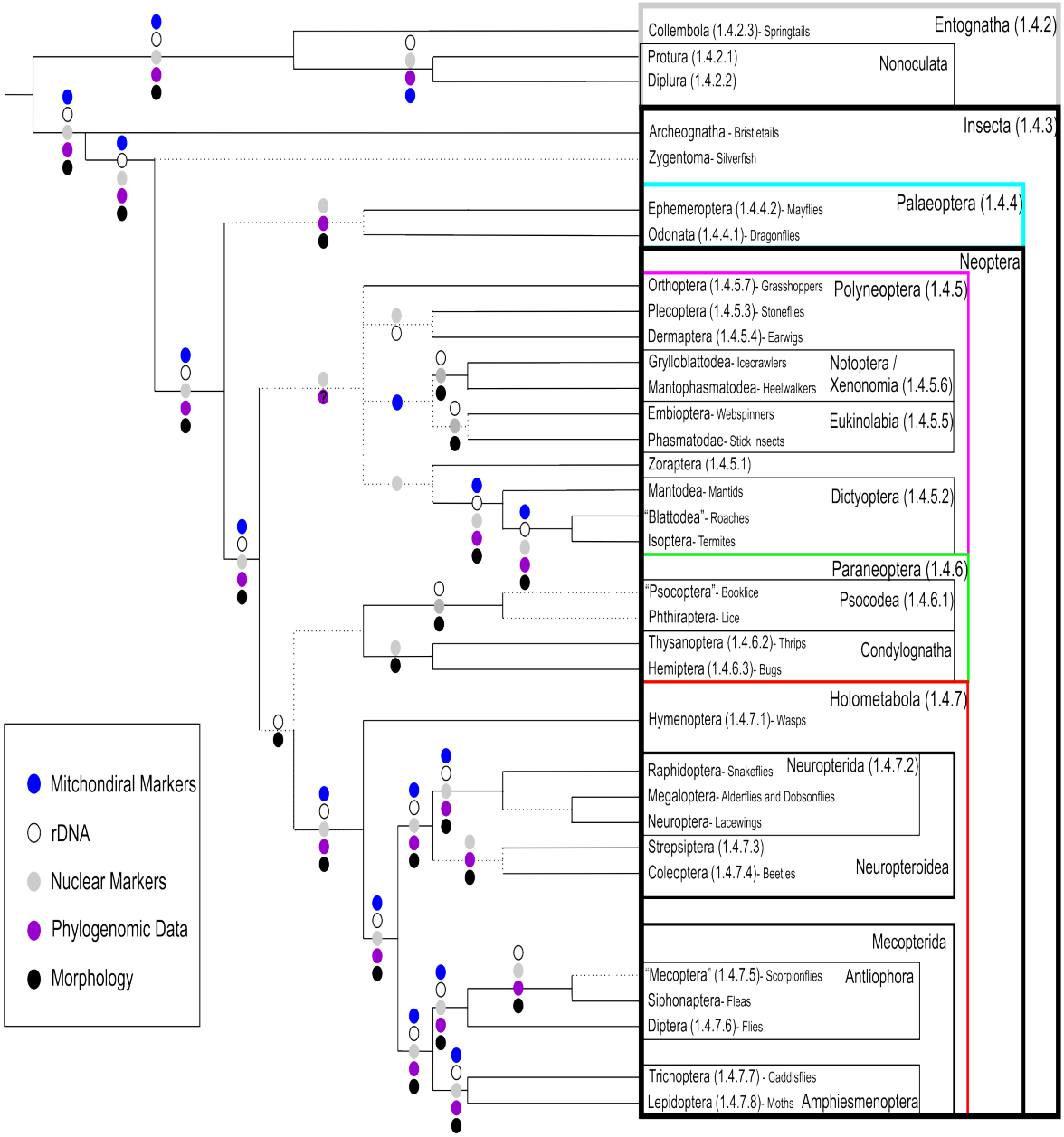


Figure 6: Summarized consensus relationships among hexapod orders with relevant evidence supporting groupings. Redrawn with modification from (Trautwein et al 2012). Major clades dicussed in subsequent chapters are highlighted. Nodes subject to phylogentic uncertainty are denoted by dashed lines, see dicussion below.

1.4.1. Hexapod origins and monophyly

The placement of Hexapoda within the wider phylogeny of arthropods has traditionally been controversial, with the majority of early workers favoring a close relationship with Myriapoda (centipedes, millipedes and allies), with whom hexapods share a number of morphological features, e.g. the loss of the second pair of antennae and the development of the tracheal system, which respectively give this grouping the interchangeable names "Atelocerata" and "Tracheata" (Giribet & Edgecombe 2012). Following the development of molecular techniques and associated advances in developmental biology the majority of workers now recognize that hexapods are in fact derived from within "Crustacea" (the so called Pancrustacea hypothesis) (Giribet & Edgecombe 2012) and that the features shared with Myriapoda were derived in parallel during the movement onto land in both lineages (Grimaldi 2010).

While early molecular studies based on mitochondrial genomes challenged the monophyly of Hexapoda (Nardi et al. 2003), subsequent work using larger datasets have generally favored a single origin of the group within Pancrustacea (Timmermans et al. 2008; Regier et al. 2010). The sister group to Hexapoda within crustaceans remains controversial, with recent analyses favoring a close relationship either with; remipedes (Von Reumont et al. 2012; Rota-Stabelli et al. 2013), remipedes + cephalocarids (Xenocarida) (Regier et al. 2010) or Branchiopoda (Meusemann et al. 2010; Andrew 2011) (although in the last remipedes are not represented). Extant Remipedea are a small group of unusual crustaceans associated with marine flooded (anchialine) cave systems, whose placement with respect to other crustacean groups is highly unstable in recent studies (Von Reumont et al. 2012). Beyond genomic data, few characters link hexapods and remipedes, although the some studies have reported shared features of brain anatomy (Harzsch 2006) and conserved structures of hemocyanin compounds used to transport oxygen (Ertas et al. 2009).

As inhabitants of transitional saline environments the placement of remipedes as potential sister groups to Hexapoda is intriguing from the perspective of hypotheses regarding the transition of the group to terrestrial environments (Von Reumont et al. 2012). It has long been recognized that hexapods have a marine origin (Giribet & Edgecombe 2012) (as opposed to the freshwater origin ascribed to tetrapods (Clack 2012)); however no marine fossil has thus far been attributed to the stem of the group (the former stem

hexapod *Devonohexapodus* is now considered a synonym for *Wingertshellicus backesi* and excluded from crown Pancrustacea (Kühl & Rust 2009)). Molecular clock studies have provided increasingly strong evidence that the transition to land happened very early in the development of terrestrial ecosystems, with estimates typically falling in the late Cambrian or early Ordovician (Section 2.4.2) (Sanders & Lee 2010; Rehm et al. 2011; Rota-Stabelli et al. 2013). Such estimates precede the earliest definitive fossils of land plants (so called "cryptospores" typical of the Middle Ordovician) by up to 50Ma, and the earliest land arthropod trace fossils (*Diplichnites* and *Diplopodichnus* from the Late Ordovician of England (Johnson et al. 1994)), by up to 80Ma (Kenrick et al. 2012). The early Paleozoic terrestrial record is poor (Grimaldi 2010; Kenrick et al. 2012) and as result we know almost nothing about the environments into which the earliest terrestrial arthropods emerged, although it seems likely at this early stage communities were restricted to marginal semi-aquatic habitats prior to the development of true land plants (Clarke et al. 2011).

1.4.2. The basal hexapods- Entognatha

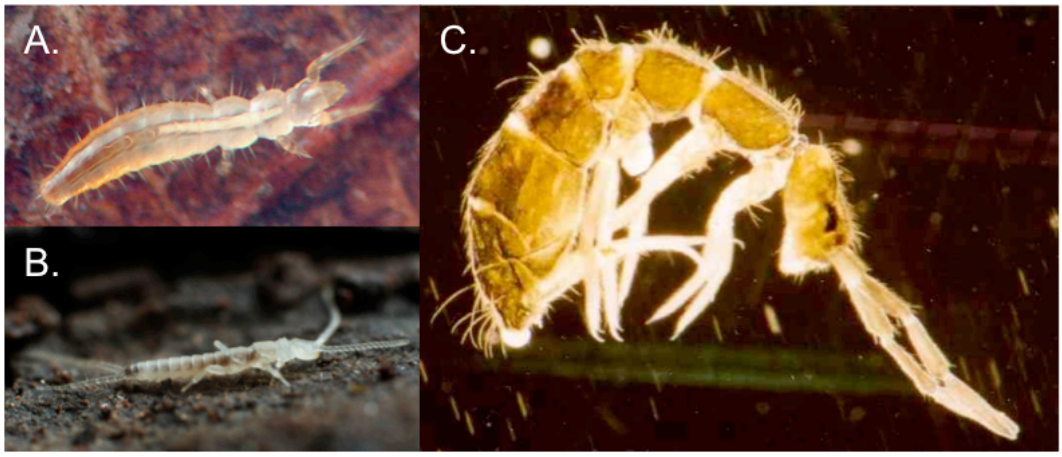


Figure 7: Examples of entognathan orders; A) Protura (1.4.2.1), B) Diplura (1.4.2.2), C) Collembola (1.4.2.3)

Photo Attribution- All pictures used under a Creative Commons license, and sourced from http://commons.wikimedia.org:

A. **David R. Maddison 2004- http://commons.wikimedia.org/wiki/File:Protura_from_Durham,_NC,_USA.jpg**

B. **Andy Murray 2013; http://commons.wikimedia.org/wiki/File:Campodeidae_sp._%2811499938054%29.jpg**

C. **U. Burkhardt 2006; http://commons.wikimedia.org/wiki/File:Isotoma_Habitus.jpg**

At the base of Hexapoda are three orders of small soil living organisms typically united due to the structure of their mouthparts. In true insects (Ectognatha) the various limbs that comprise the mouthparts are external to the head capsule, while in these three groups (Protura, Diplura and Collembola) the mouthparts are recessed into a space known as the gnathal pouch; hence the collective name Entognatha (Grimaldi & Engel 2005). Whether Entognatha represents a monophyletic lineage or a grade within basal hexapods been controversial, with several morphologists suggesting that some members of Diplura may have close affinities with true insects e.g. (Kukalova-Peck 1987). However, studies including representatives from all three orders have typically supported the group's monophyly (Meusemann et al. 2010; Regier et al. 2010; Von Reumont et al. 2012) (although see alternative placement in (Misof et al. 2014)). Relationships between the three orders remain controversial, with some molecular studies favoring a sister grouping of Protura and Diplura (termed Nonoculata) (Giribet et al. 2004; Luan et al. 2005; Meusemann et al. 2010), as opposed to the more traditional grouping of Protura with Collembola (Ellipura) (Von Reumont et al. 2012). Morphological support for the former grouping is scant, with the name being derived from the absence of eyes in both lineages, whereas Ellipura is favored by details of the structure of the gnathal pouch (Grimaldi & Engel 2005).

1.4.2.1. Protura- proturans

Among the least familiar hexapod groups, Protura are tiny, functionally tetrapod, soil dwelling organisms whose global diversity is poorly known (Pass & Szucsich 2011). The group has received little phylogenetic attention with the only substantial molecular study to date, (Dell'Ampio et al. 2011) supporting the monophyly of the major sub clades Acerentomata (Protentomidae, Acerentomidae and Hesperentomidae) and Eosentomata (Eosentomidae, Antelientomidae) but suggesting that "Sinentomata" (Fujientomidae, Sinentomidae) represents a paraphyletic grade at the base of the order. Inevitably, taxonomic sampling is sparse and the reciprocal monophyly of the various families remains to be demonstrated (Dell'Ampio et al. 2011). What little is known about proturan ecology suggests that fungal hyphae may be an important component of the diet although few ecological studies have been conducted for the group (Pass & Szucsich 2011). Given their small size and fragile structure it is perhaps unsurprising that proturans lack any fossil representatives and nothing is known regarding the group's origins and history (Grimaldi & Engel 2005).

1.4.2.2. Diplura- diplurans

Diplura are a low diversity, soil-living order, whose most characteristic feature are large paired cerci that, in some super-families are used in prey capture. The group is conventionally divided into several lineages, the occasionally large sized (up to 50mm) predatory Japygoidae, and the smaller, generalist feeding Campodeomorpha and Projapygidae (the latter grouped with japygids as the suborder Japygomorpha). Due to a number of apparently shared features between Japygomorpha and true insects, e.g. the presence of accessory tubules within the sperm axonome (Carapelli et al. 2006), the monophyly of Diplura (and thus Entognatha) has been subject to a number of challenges (Kukalova-Peck 1987; Carapelli et al. 2006), although most molecular studies to date have consistently supported the order e.g. (Mallatt & Giribet 2006; Gao et al. 2008; Regier et al. 2010), with the exception of the genome level study of (Misof et al. 2014). The earliest definitive Diplura is a well preserved japygid from the Cretaceous of Brazil (Wilson & Martill 2001), and members of both suborders are known from Cenozoic ambers (Grimaldi & Engel 2005).

1.4.2.3. Collembola- springtails

Collembola are the most species rich of the Entognthan orders, with typical members being extremely abundant microbial grazers associated with soil and other detritus-rich environments, although a number of more specialized ecologies are known to occur among the various families (reviewed in (Hopkin 1997)). The most distinctive feature of the group is the furculum spring mechanism, derived from the fused appendages of the fourth abdominal segment and used in predator avoidance. Recent systematic surveys favor the existence of five major lineages; Poduromorpha, Tomoceroidea, Entomobryomorpha, Symphypleona and Neelipleona although the relationships between these groups remain poorly understood (D'Haese 2003; Xiong et al. 2008). Outstanding problems include the removal of Tomoceroidea from the Entomobryomorpha (conflicting with traditional taxonomy) and the placement of the highly derived miniaturized family Neelidae (Xiong et al. 2008).

Compared with other Entognatha the fossil record of Collembola is surprisingly rich with most extant families having stem representatives in Cretaceous ambers e.g. (Christiansen & Nascimbene 2006). The most well-known fossil Collembola is also the

earliest definitive representative of Hexapoda, *Rhyniella praecursor* (Hirst & Maulik 1926) from the Middle Devonian; Rhynie chert formation of Scotland (Whalley & Jarzembowski 1981), whose modern appearance has led to classification within the extant family Isotomidae (Greenslade & Whalley 1996), although uncertainties in character polarity have resulted in many authors, particularly in the context of calibrating molecular clocks, favoring treatment as a stem member of the order e.g. (Rehm et al. 2011).

1.4.3. Insecta- basal members (bristletails and silverfish)

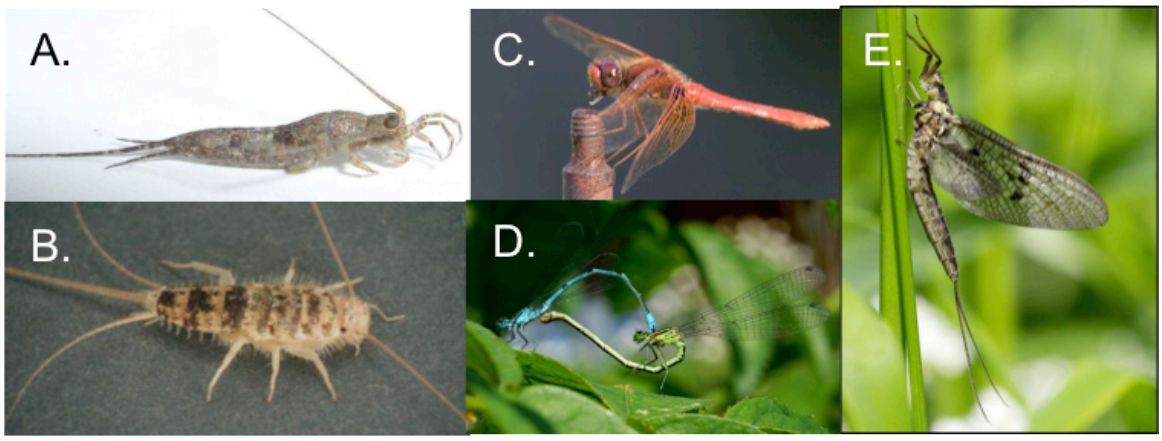


Figure 8 Examples of Basal Insects (1.4.3); A) Archaeognatha, B) Zygentoma; and Palaeoptera (1.4.4); Odonata (1.4.4.1) C) Anisoptera, D) Zygentoma; E) Ephemeroptera (1.4.4.2)

Photo Attribution- All pictures used under a Creative Commons license, and sourced from http://commons.wikimedia.org

- **A. User:Stemonitis 2006- http://commons.wikimedia.org/wiki/File:Archaeognatha.jpg**
- **B. David R. Madison 2003-http://commons.wikimedia.org/wiki/File:Thermobia_domestica1.jpg**
- **C. Ingrid Taylar 2010- http://commons.wikimedia.org/wiki/File:Dragonfly%27s_Meal.jpg**
- **D. Friedrich Böhringer 2006- http://commons.wikimedia.org/wiki/File:Coenagrion_puella_Paarung1.JPG**
- **E. Cameraman 2008- http://commons.wikimedia.org/wiki/File:Mayfly_resting_on_the_river_bank_at_Thornborough_Bridge._-_geograph.org.uk_-_1503506.jpg**

In many traditional discussions of the hexapod phylogeny Entognatha are linked with two small lineages of primitively wingless insects forming the grade “Apterygota”, e.g. (Hennig 1969). Occasionally both lineages are grouped as the paraphyletic order “Thysanura”, however most workers recognize two distinct orders; Archaeognatha (jumping bristletails) and Zygentoma (silverfish and firebrats), with the latter considered

sister to the winged insects; forming the clade Dicondylia (Grimaldi & Engel 2005; Trautwein et al. 2012). The key feature uniting Zygentoma and winged insects is the presence of a second point of articulation (or "condyle") on the mandible, restricting its movement to a single plane and increasing the power of the "bite" (Grimaldi & Engel 2005).

Both bristletails and silverfish are generally small, drab and dorsally flattened insects, usually feeding as nocturnal detritivores in damp environments. Neither group has received much systematic attention, e.g. (Comandi et al. 2009). An outstanding controversy is the placement of the mono-specific zygentoman family Lepdiotrichidae, which unlike other members of the order retain occeli (Grimaldi & Engel 2005). On this basis, the latter has been proposed as sister to Pterygota (winged insects) (Beutel & Gorb 2001; Engel 2006) although this placement conflicts with the majority of available molecular data (Giribet et al. 2004; Trautwein et al. 2012).

1.4.4. Winged insects and the Palaeoptera problem

When talking about the insects as a terrestrial group it is impossible to overstate the importance of flight (Dudley 2002; Mayhew 2007). Flight defines so much about the ecology of insects, and their importance in terrestrial and freshwater ecosystems, that it is perhaps unsurprising that the origins and basal relationships within the winged insects should be one of the most active and contentious phylogenetic issues in the group (Hovmöller et al. 2002; Whitfield & Kjer 2008; Trautwein et al. 2012; Thomas et al. 2013).

Pterygota, the clade that contains winged insects, is almost universally recognized as monophyletic (Trautwein et al. 2012) and contains four major lineages: Odonata (the dragonflies and damselflies), Ephemeroptera (mayflies), Neoptera (insects able to fold the wings flat across the abdomen, including all other living winged insects) and the entirely extinct superorder Palaeodictyoptera (Grimaldi & Engel 2005). The relationships between the three extant winged groups has come to be referred to as the "Palaeoptera problem" and represents one of the best known examples of an ancient divergence with very short internode differences leading to phylogenetic uncertainty (Whitfield & Kjer 2008; Thomas et al. 2013).

The term Palaeoptera refers to a hypothesis linking Odonata and Ephmeroptera (Hennig 1969) which until very recently had dropped out of favor in the entomological community (Thomas et al. 2013). Instead authors have focused on alternatives such as: "Metapterygota"; uniting Odonata and Neoptera, based on loss of the non-reproductive flying stage ("sub-imago") and the fixation of mandibular articulation (Wheeler et al. 2001; Ogden & Whiting 2003; Grimaldi & Engel 2005; Beutel & Gorb 2006), and "Chiastomyaria"; linking Ephmeroptera with Neoptera by indirect sperm transfer and ribosomal markers (Kjer 2004; Mallatt & Giribet 2006; Misof et al. 2007; Wang et al. 2013)(Mallatt & Giribet 2006), as well as genomic (Simon et al. 2009; Meusemann et al. 2010) and mitochondrial datasets (Li et al. 2014).

The revival of Palaeoptera as a hypothesis rests on recent high resolution phylogenomic studies (Thomas et al. 2013; Misof et al. 2014), evidence from slow evolving nuclear markers (Regier et al. 2010; Ishiwata et al. 2011) and reanalysis of character states in the head (Blanke et al. 2012; Blanke, Greve, Wipfler, et al. 2013). This is the currently favored hypothesis, although most reviews prefer to denote the node as unresolved, e.g. (Trautwein et al. 2012; Yeates et al. 2012). As with many controversial phylogenetic issues the impacts of taxon sampling and marker choice on the outcomes of studies remains unclear and it will be some years before the application of new genomic data begins to resolve these long standing issues.

1.4.4.1. Odonata- dragonflies and damselflies

Thanks to their often-large size and metallic coloration Odonata are among the most charismatic of insect groups and as a result have been thoroughly studied in terms of their ecology and distribution (e.g. (Silsby 2001)). All members of the order develop as predatory nymphs in freshwater and the adults are typically active aerial hunters. The difference between dragonflies (sub-order Epiprocta) and damselflies (Zygoptera) lies in the strongly oblique thorax structure of the latter, allowing the wings to be held upright as opposed to flat resting position typical of dragonflies. Traditionally a third sub-order Anisozygoptera was recognized for the extant family Epiophlebiidae, however these are now usually grouped within Epiprocta, as basal to the other dragonflies (Anisoptera) (Bybee et al. 2008; Dumont et al. 2010).

Both Zygoptera and Anisoptera are subject to phylogenetic issues in the placement of included families. Zygoptera is now recognized as comprising (at least) two major

clades; with the superfamily Lestoidea (Davis et al. 2011; Dijkstra et al. 2014) (also termed Lestomorpha (Dumont et al. 2010)) being recognized as sister to the remaining taxa, however support for relationships above the family level is low, and the monophyly of some traditional families, e.g. Coenagrionidae, Megapodagrionidae and Amphipterygidae, are considered suspect (Dumont et al. 2010; Dijkstra et al. 2014). Within Anisoptera, molecular analysis place the superfamily Aeshnoidea as basal and generally support most of the traditional groupings, although the large heterogeneous family Corduliidae is often found to be a polyphyletic assemblage, with groups spread across different parts of the tree (Fleck et al. 2008; Bybee et al. 2008; Dumont et al. 2010; Blanke, Greve, Mokso, et al. 2013).

Crown Odonata are strictly a post-Permian radiation (Grimaldi & Engel 2005; Davis et al. 2011). However stem members of the group, including the famous giant Protodonata (Grimaldi & Engel 2005), are important components of early hexapod faunas and include the largest known winged insects, e.g. *Meganeuropsis permiana,* with wingspans approaching 70cm (Grimaldi & Engel 2005; Clapham & Karr 2012).

1.4.4.2. Ephemeroptera - mayflies

Best known for their mass emergences and proverbially brief adult lifespans (the adults never feed and most live only a few hours) mayflies are another major hexapod group that conduct the majority of their lifecycle in freshwater. In most families the nymphs are detritivores, although a few are carnivorous in late instars, and members are common in both fluvial and lacustrine environments (Barber-James et al. 2007). Mayfly systematics are currently poorly resolved, with the most recent review challenging the traditional suborders (Pisciforma and Setisura), although the major clades Carapacea (including taxa with a notal shield) and Furcatergalia (defined by adaptions to burrowing) were recovered (Ogden et al. 2009). Like Odonata, extant Ephemeroptera are a Mesozoic radiation with diverse Palaeozoic stem groups, although the timings of major transitions and the origins of extant families remain poorly defined (Grimaldi & Engel 2005).

1.4.5. Polyneoptera and basal Neoptera

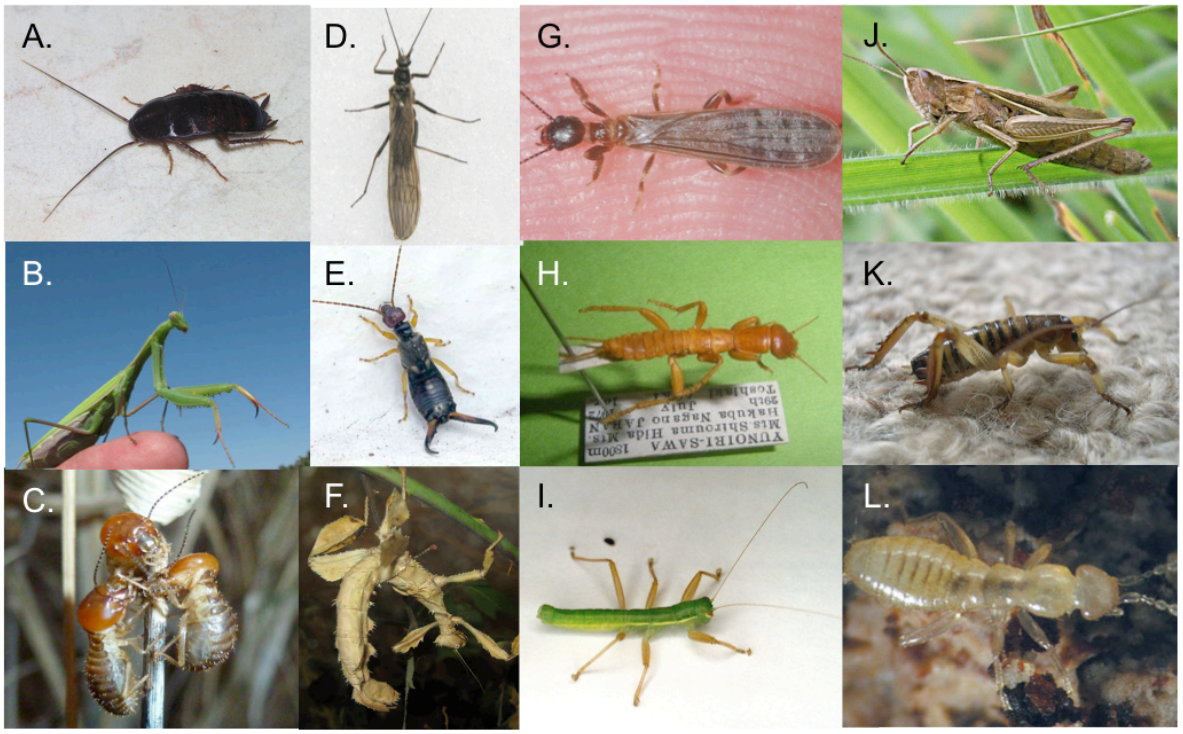


Figure 9 Examples of Polyenopteran orders; Dictyoptera (1.4.5.2); A) "Blattodea", B) Mantodea, C) Isoptera; D) Plecoptera (1.4.5.3); E) Dermaptera (1.4.5.4); Eukinolabia (1.4.5.5); F) Phasmatodea, G) Embioptera; Notoptera/Xenonomia (1.4.5.6); H) Grylloblatodea, I) Mantophasmatodea; Orthoptera (1.4.5.7); J) Caelifera, K) Ensifera; L) Zoraptera (1.4.5.1)

Photo Attribution- All pictures used under a Creative Commons or Free Art License and sourced from http://commons.wikimedia.org and http://tolweb.org/tree/

A. **Alvesgaspar 2007- http://commons.wikimedia.org/wiki/File:Cockroach_May_2007-1.jpg**

B. **Friedl Foelsche 2012- http://commons.wikimedia.org/wiki/File:Gottesanbeterin_1.jpg**

C. **JMK 2014- http://commons.wikimedia.org/wiki/File:Hodotermes_mossambicus,_op_dro%C3%AB_gras,_d,_Voortrekkerbad.jpg**

D. **Algirdas 2006- http://commons.wikimedia.org/wiki/File:Taeniopteryx_nebulosa2.jpg**

E. **Luis Miguel Bugallo Sánchez 2005- http://commons.wikimedia.org/wiki/File:Cadeladefrade_045eue.jpg**

F. **Adrian Pingstone 2008-http://commons.wikimedia.org/wiki/File:Macleays_spectre_stick_insect_arp.jpg**

G. **S. Dean Rider, Jr. 2010- http://commons.wikimedia.org/wiki/File:Osaundersii_Male.jpg**

H. **OpenCage 2007 -http://commons.wikimedia.org/wiki/File:Galloisiana_nipponensis.jpg**

I. **P.E. Bragg 2002 - http://commons.wikimedia.org/wiki/File:Mantophasma_zephyra_Zompro_et_al_2002.jpg**

J. **Francisco Welter-Schultes 2011 http://commons.wikimedia.org/wiki/File:Caelifera-7303r.JPG**

K. **Andrew McMillan 2013 http://commons.wikimedia.org/wiki/File:New_Zealand_weta.jpg**

L. **David R. Maddison, 2004 ; http://tolweb.org/Zoraptera**

Neoptera, insects able to fold the wings flat across the abdomen, is generally considered to comprise three major lineages. Of these two; Holometabola (insects with complete metamorphosis) and Paraneoptera (bugs and their relatives), are generally regarded as secure monophyletic groups (although see Section 1.4.6), while the third, Polyneoptera, refers to a disparate assemblage of orders, whose relationships remain one of the greatest outstanding challenges in hexapod phylogenetics (Trautwein et al. 2012). From a morphological perspective the orders making up Polyneoptera show few uniting features (Grimaldi & Engel 2005), although the group has been recovered in analysis of wing base structures (Yoshizawa 2011) and attachment devices (Beutel & Gorb 2006); the latter excluding Zoraptera (See Section 1.4.5.1). However molecular analyses show remarkable consistency in uniting this clade (Trautwein et al. 2012), although both the internal relationships and the sampling of lineages are highly inconsistent across different studies, e.g. (Letsch et al. 2012; Simon et al. 2012; Misof et al. 2014).

For convenience the discussion here is structured around the phylogeny presented in Chapter 2 (broadly consistent with the consensus view shown in (Trautwein et al. 2012) and Figure 6), and concerns five major lineages the interrelationships of which remain unclear. Of these, only the Dictyoptera (Section 1.4.2.2) (Mantodea + "Blattodae" + Isoptera) is universally recognized and strongly supported by their reduced ovipositor, gizzard like proventriculus and the deposition of eggs in specialized pods known as oothecae (although within Isoptera, the last occurs only in the most basal taxa) (Grimaldi & Engel 2005).

By contrast with Dictyoptera, the sister grouping of Plecoptera (Section 1.4.5.3) and Dermaptera (Section 1.4.5.4) has little to recommend it from a morphological perspective, linking as it does one of the most primitive looking and one of the most derived neopteran orders with respect to traditional wing-vein characters (Haas & Kukalová-Peck 2001). Nevertheless the grouping has received consistent support across a range of molecular markers, including 18S (Misof et al. 2007) and 28S (Wang et al. 2013) rRNA, nuclear protein coding genes (Ishiwata et al. 2011), mitochondrial genomes (Wan et al. 2012), genome data (Simon et al. 2012; Misof et al. 2014) and supermatrix studies e.g. (Kjer et al. 2006).

Another grouping that appears robust on recent analyses but which challenges traditional views of hexapod relationships is the concept of Eukinolabia; uniting the orders Phasmatodea and Embioptera (Section 1.4.5.5). This grouping was originally defined in the supermatrix study of Terry & Whiting (2005), and has since been recovered by (Kjer et al. 2006; Ishiwata et al. 2011; Wang et al. 2013; Letsch et al. 2012) and (Misof et al. 2014) (although see, (Misof et al. 2007) for a counter example). The clade receives some morphological support based on the flexor of the paralossae (Friedemann et al. 2012) and the presence of an operculum on the egg (Bradler 2009). Mitochondrial genome studies by contrast, tend to reject this concept in favor of uniting Phasmatodea (in particular the basal genus *Timema*) with the recently described order Mantophasmatodea (see below) (Klass et al. 2002; Damgaard et al. 2008; Cameron et al. 2006; Plazzi et al. 2011), although this may be due to the lack of suitable data for Embioptera (Kômoto et al. 2012).

Mantophasmatodea is conventionally placed in the clade variously referred to as Notoptera (Arillo & Engel 2006), Chimaeraptera (Uchifune & Machida 2005) or Xenonomia (Terry & Whiting 2005), the other members of which are the bizarre cryophilic relic taxon Grylloblattodea (Section 1.4.5.6). Discounting the mitochondrial data discussed above, the close affinities of these orders have been supported by a wide range of markers (Terry & Whiting 2005; Kjer et al. 2006; Ishiwata et al. 2011; Wang et al. 2013; Misof et al. 2014) and various features of the head (Baum et al. 2007; Wipfler et al. 2011), egg (Uchifune & Machida 2005), and attachment structures (Beutel & Gorb 2006; Beutel & Gorb 2008)(Beutel & Gorb 2008).

By far the most species rich Polyneopteran order, Orthoptera (Section 1.4.5.7) stands somewhat apart in recent studies partially as a consequence of incomplete taxonomic sampling. The nuclear protein gene analysis of (Ishiwata et al. 2011) placed the group as sister to the Plecoptera +Dermaptera + Dictyoptera, while EST data has variously placed orthopterans as sister to Dictyoptera (Simon et al. 2012) or as basal to Polyneoptera as a whole (Letsch et al. 2012). (Terry & Whiting 2005) suggested a sister group with Eukinolabia, while mitochondrial genomes for the most part support Orthoptera as sister to Dictyoptera/Eukinolabia/Xenomia (Plazzi et al. 2011; Wan et al. 2012) (also found with weak support in (Misof et al. 2014)). The 18S study of (Misof et al. 2007) supports the traditional view of an Orthoptera/Phasmatodea sister relationship ("Orthopterida"). Perhaps unsurprisingly morphology does little to resolve these issues, with the most comprehensive datasets to date supporting the traditional "Orthopterida" grouping

(Yoshizawa 2011) or a sister group to Phasmatodea + Xenonomia (Beutel & Gorb 2008). At the present time therefore the placement of Orthoptera with respect to other polyneopteran groups remains poorly defined.

1.4.5.1. The Zoraptera problem

Of all the Polyneopteran orders, none has presented such a substantial classification problem as Zoraptera. This tiny and obscure group of 32 species, which form semi-social fungus feeding colonies in dead wood, were originally classified in Paraneoptera (e.g. (Hennig 1969)) although they have also appeared as basal Eumetabola (the clade uniting Paraneoptera and Holometabola) (Wheeler et al. 2001) or even sister to the Holometabola (reviewed in (Beutel & Weide 2005)). While there is broad consensus regarding a placement within Polyneoptera (reviewed in (Beutel & Weide 2005; Grimaldi & Engel 2005; Yoshizawa 2007)), Zoraptera's affinities within the clade now presents the most severe outstanding problem in insect ordinal systematics (Yoshizawa 2011; Trautwein et al. 2012).

A major part of the "Zoraptera problem" is lack of data. The cryptic lifestyle and relative rarity of zorapteran taxa have led to them being underrepresented with respect to otherwise widely sampled molecular markers, e.g. at the time of writing, Zoraptera are the only insect order to lack a representative mitochondrial genome sequence (Cameron 2014). In addition both the 18S and 28S rRNAs are structurally highly modified in Zoraptera relative to other insect groups leading to challenges in alignment and the identification of homologous sites (Yoshizawa & Johnson 2005; Wang et al. 2013). Contamination of sequences within molecular studies has also played a role in generating ambiguity, as at least one study, that of (Terry & Whiting 2005), contains sequences that, while labeled as zorapterans, appear to in fact be of dermapteran origin, which has resulted in problematic alignment of markers under direct optimisation (Yoshizawa 2010).

There are four major hypotheses regarding the position of Zoraptera within Polyneoptera. The first, as sister group to Dermaptera ("Haplocercata" (Terry & Whiting 2005)) is called into question by the contamination issues noted above (although it reappears with weak support in the recent genome study of (Misof et al. 2014)). A second view, commonly identified in morphological data but lacking molecular support, places Zoraptera with Embiodea ("Mystroptera") (Yoshizawa 2007; Friedrich & Beutel 2008; Yoshizawa 2011). Comprehensive studies of 18S suggest links with Plecoptera, although

as noted there are issues in the alignment of this marker across groups (Misof et al. 2007). Finally, single copy nuclear markers support placement close to the Dictyoptera, which is our favored hypothesis as it is derived from data sources independent of known erroneous markers and explains unique shared features of ribosomal structure seen between the groups (Yoshizawa & Johnson 2005; Ishiwata et al. 2011; Wang et al. 2013).

1.4.5.2. Dictyoptera- roaches, mantids and termites:

The Dictyoptera are an undisputed monophyletic group comprising three lineages with differing ecologies, the carnivorous mantids, the eusocial termites and the generalist detritivorous cockroaches (Grimaldi & Engel 2005). The focus of recent phylogenetic efforts within the group has been on the interrelationships of these lineages, and in particular establishing the appropriate placement of termites with respect to roaches. Traditionally recognized as their own order (Isoptera), recent studies have increasingly converged on placing termites within the roach clade (formerly the order "Blattodea") as sister to the wood roach family Cryptoceridae (Lo et al. 2000; Inward et al. 2007; Ware et al. 2008; Davis et al. 2009; Djernæs et al. 2012). Such studies have however differed in their treatment of the basal roach families Polyphagidae and Nocticolidae, and the placement of Mantodea, which have been shown to be sensitive to out-group choice (Ware et al. 2008). The most recent study identified a monophyletic Mantodea as sister to the combined "Blattodea"/ Isoptera clade (Djernæs et al. 2012).

Within each of the "orders", family relationships remain unsettled. In roaches the respective monophyly and composition of the large families Blaberidae and "Blattellidae" (=Ectobiidae) remains an outstanding challenge. Termites face questions regarding the (traditional) basal placement of the monospecific Mastotermittidae and the division between the "lower"; harvester, damp and dry wood termites vs. the more derived "higher termite" families (Ware et al. 2008; Djernæs et al. 2012). The phylogeny of mantids highly problematic, with recent analysis supporting traditional basal lineages (Chaeteessidae, Mantoididae and possibly Metallyticidae) (Svenson & Whiting 2004), but calling into question many of the more derived families and indicating that morphological convergence is rife among unrelated clades within the group (Svenson & Whiting 2009).

Dictyoptera have one of the richest fossil records of any polyneopteran group (Grimaldi & Engel 2005). The earliest division among the crown groups is the Permian stem mantid *Mesoptilus dolloi* (Béthoux & Wieland 2009), although the modern clades

appear to have a late Mesozoic origin, based on both fossil evidence (Vršanský 2008; Grimaldi 2003) and patterns of vicariance observed in phylogenies (Svenson & Whiting 2009). The earliest definitive crown members of the Blattodea/ Isoptera clade are of Cretaceous origin in the Bassia deposits of central Russia (Vršanský 2005; Engel et al. 2007).

1.4.5.3. Plecoptera – Stoneflies

Stoneflies are traditionally (and misleadingly) regarded as the most primitive neopteran lineage and have long been used a models for the origins of the group, and of insect flight in general (Hennig 1969; Dudley 2002; Engel et al. 2013). Uniquely among Polyneoptera, plecopteran nymphs are aquatic, and the order as a whole is principally associated with running or cool water environments with relatively high oxygen concentrations. Most Plecoptera nymphs are shredders of decaying leaf material, although some families in the superfamily Perliodae are strictly predatory (Zwick 2000), and the short-lived adults appear to feed very little, if at all (Hynes 1942).

Plecoptera are another polyneopteran group in need of systematic review based on molecular data and improved out-group selection (Thomas et al. 2000). The traditional taxonomy, based on the informal cladistics study of (Zwick 2000) distinguishes between a southern hemisphere lineage Antarctoperlaria and a predominantly northern hemisphere group ("Arctoperlaria") encompassing the suborders Systellognatha and Euholognatha (Grimaldi & Engel 2005). Molecular data shows only limited support for this disjoint pattern, with the 18S study of Thomas et al. (2000) recovering a paraphyletic "Arctoperlaria" with Nemouroidea at the base (although see (Terry 2003) for an alternative view).

Despite a presumed ancient, origin, the fossil record of Plecoptera is rather poor, likely a reflection of the groups preference for environments with low preservation potential (Grimaldi & Engel 2005). The earliest known stem representative is the Carboniferous *Gulou carpenter* (Béthoux et al. 2011), although well preserved members of extant families are restricted to the Jurassic of Germany and China (Zhao et al. 2010).

1.4.5.4. Dermaptera –earwigs

Earwigs are an often overlooked, homogenous group of nocturnal omnivorous insects, whose most distinctive features are the paired forceps on the terminal

of the abdomen and the modification of the forewing to form a rigid tegminious cover analogous to the elytra of beetles (Grimaldi & Engel 2005). This description overlooks two remarkable small families Arixeniina and Hemimerina, which are the only ecto-parasitic polyneopterans (on molossid bats and two genera of African rodents respectively) (Kocarek et al. 2013). The placement of these lineages, which until recently had been allotted sub-ordinal status (i.e. regarded as sister to all free-living earwigs; Forficulina), provides the context for recent phylogenetic study within the group (Kocarek et al. 2013). Molecular investigations have, unsurprisingly, found that both parasitic groups nest within Forficulina, and have reaffirmed the important superfamily Forficuloidae (Jarvis et al. 2005; Kocarek et al. 2013). Dermaptera are rare in the fossil record and most Mesozoic taxa are placed within the extinct suborders "Archidermaptera" and Eodermaptera (Grimaldi & Engel 2005; Zhao et al. 2010; Nel, Aria, et al. 2012).

1.4.5.5. Phasmatodea and Embiodae- stick insects and webspinners

On the surface one would be hard pressed to find two more distinct clades than the, often large, cryptic, herbivores of Phasmatodea and the comparatively small, web-spinning detritivores of Embiodae. However as noted above the two groups share a number of features strongly suggesting a common origin (Eukinolabia) and I will treat them together here. Stick and leaf insects are perhaps best known from their extraordinary body form and camouflage abilities, and include some of the longest of all extant insects (Hennemann & Conle 2008). Within the order there exists a fundamental division between the small (21 species) relict genus *Timema* native to the western USA, and the remaining taxa collectively known as Euphasmatodea, although within the latter the status of a number of recognized families and subfamilies remains uncertain (Buckley et al. 2009; Kômoto et al. 2011; Bradler et al. 2014).

Embids are a predominantly tropical or sub-tropical forest group, where they live within communal galleries spun from silk produced in their uniquely modified forelimbs. All female Embiids are wingless and males have unique collapsible wings that are inflated for use by pumping haemolymph into "blood sinuses" that are the group's most distinctive synapomorphy. Surprisingly, given their cryptic nature, the phylogeny of Embioptera is comparatively well understood, with supermatrix studies confirming the monophyly of four of the eight recognized families (in total around 400 species) and redefining the remainder into monophyletic lineages (Miller et al. 2012). The internal phylogeny of the

group remains poorly supported in these studies with either Clothodidae or Australembiidae being considered the basal members of the group and Oligotomidae and Tetatembiidae being highly supported as sister taxa (Miller et al. 2012). The family Embiidae remains a taxonomic problem and further redefinition and combination with other groups is required (Szumik et al. 2008; Miller et al. 2012).

In keeping with their arboreal habits the fossil records of both Phasmatodea and Embioptera are exceptionally poor. Traditionally a wide array of Mesozoic forms have been considered as stem group Phasmatodea (e.g. (A. Nel et al. 2010; Nel & Delfosse 2011)). However no unambiguous characters link these with either of the two extant lineages and thus their ability to inform us regarding the group's origin is extremely limited (Bradler & Buckley 2011). Fossil embiids showing all of the group's major characters are known from late Jurassic (Huang & Nel 2009) and late Cretaceous ambers (Engel & Grimaldi 2006); however as yet we have only a limited picture of how these findings fit within the wider context of Polyneopteran diversification.

1.4.5.6. Notoptera/Xenonomia- ice crawlers and heelwalkers

The discovery of Mantophasmatodea stands out as one of the most unexpected and significant events in entomology in recent decades (Klass et al. 2002). This small group of nocturnal carnivores from southern Africa had apparently been overlooked due to their remarkable similarity to nymphs of other Polyneopteran groups, notably Phasmatodea and Orthoptera (Klass et al. 2002). Nevertheless their anatomy is highly distinct, not least for their characteristic tarsal morphology, whereby the terminal segments are held clear of the substrate, a feature which has earned them one of their many common names, the "heelwalkers" (others include "rock-crawlers" or "gladiator insects"). The other members of Xenomia are likewise unusual, the 27 species of Grylloblatodea (ice crawlers) are almost completely restricted to mountain tops and other cool environments in the Northern hemisphere where they live as slow moving predator-scavengers. Given the low diversity and comparatively homogenous form of these orders, both are typically treated as single families (although see (Damgaard et al. 2008) for an alternative treatment of Mantophasmatodea) and some schemes treat them as suborders of a united Notoptera.

While the living fauna of Notopotera are clearly relictual the group has traditionally played host to a large array of difficult-to-place fossil taxa, most of which have been described in Grylloblattodea. The linking of these, typically isolated, wing fragments to the

wingless extant forms is often tenuous (Huang, Nel & Petrulevičius 2008) and the several authors have recognized that the group has acquired a wastebasket status and is in need to review (Grimaldi & Engel 2005). It seems likely that among these various Paleozoic groups the stem lineages of Notopera and possibly of other extant groups (notably Eukinolabia) are present but as yet unidentified. Definitive members of both Mantophasmatodea and Grylloblatodea are known from the Jurassic of China (Huang, Nel, Zompro, et al. 2008; Huang, Nel & Petrulevičius 2008).

1.4.5.7. Orthoptera- grasshoppers and crickets

Grasshoppers and crickets are the most numerous (20,000 species; (Grimaldi & Engel 2005)) and familiar of the Polyneopteran orders. The monophyly of the order is defined by a number of features of which most obvious are the presence of a cryptopleuron, the saltatorial (jumping) hind limbs and the widespread presence of acoustic capabilities (Grimaldi & Engel 2005). The order is divided into two natural groupings: the typically nocturnal and omnivorous crickets and katydids (Ensifera), and the almost universally herbivorous, "shorthorn grasshoppers" and locusts (Caelifera). Most molecular treatments have tended to focus on only one of these suborders with the result that they differ considerably in their degree of phylogenetic resolution.

Of the two suborders, Ensifera has proved to be the most problematic from a phylogenetic perspective, with many issues centering on a collection of large wingless taxa, formerly grouped in the superfamily Stenopelmatoidae (cave crickets; Rhaphidophoridae, Jerusalem crickets; Stenopelmatidae, king crickets and weta; Anostostomatidae and raspy crickets; Gryllacrididae) the monophyly of which, and their respective placement with respect to other major clades such as the true crickets (Grylloidea), the splay footed crickets (Schizodactylidae) and the, often herbivorous, grigs and katydids (Prophalangopsidae and Tettigoniidae) remains highly unstable in recent studies (Legendre et al. 2010).

By contrast Caelifera shows a fair degree of consensus among analyses, including a basal division between Tridactyloidea (a detriverious groups with subterranean habits) and the remaining taxa comprising the pigmy grasshoppers (Tetrigidae) and the true locusts and their allies; Acridoidea (Song 2010; Leavitt et al. 2013). Within Acridoidea the limits of some of the families remains unclear; however recent taxonomic treatments have begun to resolve some of the problematic areas (Leavitt et al. 2013).

Orthoptera are recognized as has having a diverse fossil record (Béthoux & Nel 2002; Béthoux 2007) and a stem member of this clade is considered the earliest definitive Neoptera (Prokop, Nel, and Hoch 2005). Given the taxonomic uncertainties, the relationships among fossil and extant taxa remain unclear, particularly with respect to Ensifera and as such the origin of modern familes remains poorly understood.

1.4.6. Paraneoptera

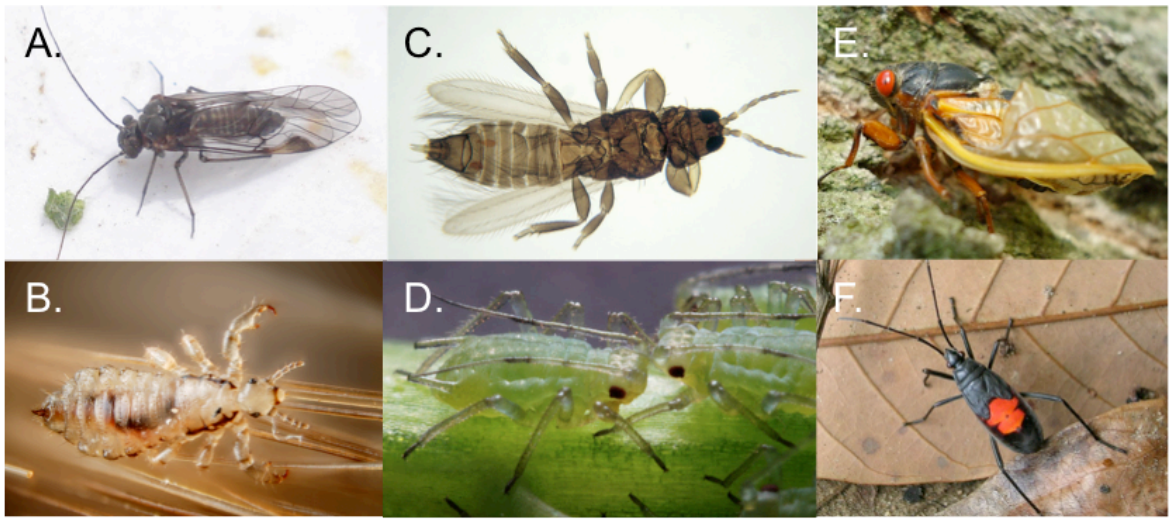


Figure 10 Examples of Parneopteran orders; Psocodea (1.4.6.1); A) "Psocoptera", B) Pthiraptera; C) Thysanoptera (1.4.6.2); Hemiptera (1.4.6.3); D) Sternorrhyncha, E) Auchenorrhyncha, F) Heteroptera.

Photo Attribution for previous page- All pictures used under a Creative Commons license, and sourced from http://commons.wikimedia.org

- **A. Mick Talbot, 2009, http://commons.wikimedia.org/wiki/File:Amphigerontia_contaminata_%283630028268%29.jpg**
- **B. Gilles San Martin, 2010, http://commons.wikimedia.org/wiki/File:Male_human_head_louse_%284900867458%29.jpg**
- **C. Luis Fernández García, 2007, -http://commons.wikimedia.org/wiki/File:Thysanoptera1.jpg**
- **D. Drc406, 2009-http://commons.wikimedia.org/wiki/File:Greenflies_or_Aphids.jpg**
- **E. Tomwsulcer, 2013, http://commons.wikimedia.org/wiki/File:Cicada_on_tree_ready_to_party.jpg**
- **F. Shyamal, 2006,- http://commons.wikimedia.org/wiki/File:Unidentified_Largid_or_Pyrrhocorid_nymph.jpg**

When compared with Polyneoptera the three orders that comprise Paraneoptera (also referred to as Acercaria) have a number of robust features that suggest they form a monophyletic group including; the reduction in the tarsomeres to three or less, the large

postclypeus, and the detachment of the stylet-like lacinia from the stipes (Grimaldi & Engel 2005). It is therefore somewhat surprising that a number of recent phylogenetic studies have failed to recover this relationship with robust support, e.g. (Yoshizawa & Johnson 2005; Meusemann et al. 2010; Ishiwata et al. 2011; Simon et al. 2012; Misof et al. 2014), although single marker studies e.g. (Misof et al. 2007; Wang et al. 2013) have successfully recovered the group, as has at least one genomic study under some analytical procedures (Letsch et al. 2012). Complete mitochondrial genomes have been highly inconsistent in Paraneopteran relationships, which may reflect high degrees of modification and compositional bias in certain lineages (Cameron et al. 2011). A key issue appears to be that the clade Psocodea, typically represented in such studies by parasitic groups such as the human louse *Pediculus humanusm,* whose highly modified and compositionally biased genome has proved problematic to compare with that of other hexapods (Letsch et al. 2012).

Within Paraneoptera, most classifications treat Thysanoptera (thrips) and Hemiptera (bugs) as sister groups (forming the clade Condylognatha) which is supported by the structure of the mouthparts and forewings (Yoshizawa & Saigusa 2001; Nel, Prokop, et al. 2012), but which has received only modest support in molecular studies (Johnson et al. 2004; Misof et al. 2007; Wang et al. 2013) and is contradicted by others (e.g. (Wheeler et al. 2001; Talavera & Vila 2011)). Given the currently available data, and the lack of sequence data for non-parasitic Psocodea, it is perhaps prudent to consider the ordinal relationships within Paraneoptera as presently unresolved (Trautwein et al. 2012).

1.4.6.1. Psocodea –lice and booklice

Most traditional classifications divide the members of Psocodea into two distinct orders; the cryptic detritivorous booklice and barklice (“Pscoptera”) and the parasitic lice (Pthiraptera) (Grimaldi & Engel 2005). In most sources Pthiraptera is considered monophyletic, and either sister to or derived from within “Pscoptera”, with the family Liposcelididae often mentioned as an intermediate form (Grimaldi & Engel 2005). The first major test of the placement of lice originated from molecular data (Yoshizawa & Johnson 2003; Johnson et al. 2004; Murrell & Barker 2005) and confirmed that that they originated from within “Pscoptera” (rendering the later paraphyletic). However, these studies also implied that Pthiraptera, at the time regarded as one of the most robust insect

orders, was itself paraphyletic, i.e. that the specialist ecto-parasitic life style had originated several times independently in different booklice taxa (Yoshizawa & Johnson 2006).

Subsequent analysis revealed considerable conflict among markers, with the strongest support for pthirapteran polyphyly coming from 18S rRNA (Yoshizawa & Johnson 2010) and the issue remains unresolved. As currently understood, lice are divided into two broad assemblages, the traditional suborder Amblycera vs. the suborders "Ischnocera", Anoplura and Rhychophthirina, which may or may not be sister taxa, and nest somewhere within "Pscoptera", with the families Liposcelididae, Pachytroctidae and Sphaeropsidae regarded as potential sisters to the louse clades (Yoshizawa & Johnson 2010). The other three major clades of "Pscoptera" (Trogiomorpha, Psocomorpha and Amphientometae) are all regarded as monophyletic although their branching order and relationships differ across different markers (Yoshizawa & Johnson 2010; Yoshizawa & Johnson 2014). Accelerated rates of substitution in lice, possibly associated with their parasitic lifestyle, mean that this is a difficult region of the tree to resolve (Yoshizawa & Johnson 2003; Cameron et al. 2011) and no doubt further sampling of "Pscoptera" lineages will lead to fresh insights into the relevant relationships. All known fossil Psocodea are of the "Pscoptera" type, with numerous stem forms known from Paleozoic deposits (Grimaldi & Engel 2005) and the earliest undisputed crown taxa occurring in early Cretaceous ambers (Grimaldi & Engel 2006).

1.4.6.2. Thysanoptera- thrips

Thrips are distinctive tiny insects easily defined as monophyletic by their hair fringed wings, eversible pretarsal bladder and the reduction of the right mandible (Grimaldi & Engel 2005). Despite their often tiny size they exhibit a surprising ecological diversity with the majority being phytophagous or fungivorious, and the occasional predatory lineage, notably in the family Aeolothripidae (Mound 2010). The most important distinction in thrip systematics is between the suborder Tubulifera, which includes only the diverse family Phlaeothripidae, and the remaining seven extant families that collectively form Terebrantia. Traditionally Tubulifera were seen as having been derived from within Terebrantia (Mound & Morris 2007); however a recent molecular study concluded that both groups are monophyletic and also confirmed the monophyly of a number of the Terebrantia families (Buckman et al. 2013). Stem group thrips have been described from the late Carboniferous of France (Nel, Azar, et al. 2012). However, secure members of the

crown group are first recovered in mudstone (Shmakov 2009) and amber deposits of the early Cretaceous (P. Nel et al. 2010).

1.4.6.3. Hemiptera- bugs

The true bugs are the most species rich and ecologically important group of non-holometabolan insects. Traditional classifications divide the group into the phytophagous "Homoptera" and the (primitively) predatory Heteroptera, although more modern schemes divide "Homoptera" into several lineages: Sternorrhyncha (including aphids and scales), Auchenorrhyncha (including froghoppers, cicadas and planthoppers) and Coleorrhyncha containing the single family Peloridiidae, with the last being recognized as the sister to Heteroptera (Resh & Cardé 2009). While the monophyly of Sternorrhyncha and Heteroptera have never been seriously challenged Auchenorrhyncha remains an outstanding problem, as the links between its two constituent infraorders Fulgoromorpha and Cicadomorpha have never been robustly established in molecular studies (Campbell et al. 1995; Xie et al. 2008). The most recent treatment based on seven gene regions supports a monophyletic Auchenorrhyncha with the overall topology being (Sternorrhyncha, (Auchenorrhyncha, (Coleorrhyncha, Heteroptera))) (Cryan & Urban 2012) although this arrangement is disputed in mitochondrial genome study of (Song et al. 2012).

Within Sternorrhyncha arrangement of the five super-families are fairly well known with Phylloxeroidea (pine aphids) sometimes being grouped within Aphidoidea (true aphids) and the whole clade considered sister to scales (Coccoidea) with whitefly (Aleyrodidae) and plant lice (Psyllidae) as successive out-groups (Cryan & Urban 2012).

Cicadomorpha is divided into three super families, Membracoidea (leaf and treehoppers), Cicadoidea (Cicadas) and Cercopoidea (froghoppers) with the last two being sister taxa (Cryan 2005; Cryan & Urban 2012). Relationships within Fulgoromorpha (plant hoppers) have only recently been established, with these results resolving a number of outstanding questions regarding the monophyly of a number of the included families, as well as identifying Cixiidae + Delphacidae as the outgroup to the remaining taxa (note that this topology implies a paraphyletic Auchenorrhyncha) (Song & Liang 2013).

Heteroptera are the most morphologically diverse Hemipteran lineages with seven infra-orders traditionally recognized. Which of these groups is basal to the suborder remains an outstanding problem, with some authors favoring the terrestrial Enicocephalomorpha (gnat bugs) (Xie et al. 2008) and others the aquatic Nepomorpha

(water bugs) (Li et al. 2012), or the aberrant nepomorph family *Pleidae* (pigmy backswimmers), which in mitogenome studies is placed within its own infra-order (Hua et al. 2009). The placements of three of the other major lineages, Gerromorpha (semi-aquatic bugs) Leptopodomorpha (shore bugs) and Dipsocoromorpha (litter bugs), remain variable among studies, resulting in much confusion regarding the direction of character evolution within the clade (Li et al. 2012). The final two infra-orders, the mostly predatory Cimicomorpha (which includes the large families Miridae; flower bugs and Reduviidae; (Hwang & Weirauch 2012; Jung & Lee 2012)) and the plant feeding Pentatomomorpha (shield bugs, stink bugs and their relatives (Li et al. 2005; Hua et al. 2008)), are almost universally regarded as sister taxa and comprise the land bugs or Geocorisae (Xie et al. 2008; Li et al. 2012).

Hemiptera are an ancient group with stem Heteroptera present in the Permian (Grimaldi & Engel 2005) and a number of extant families recorded from Triassic deposits (Fraser et al. 1996; Yao et al. 2012). A recent dating study indicated that Heteroptera represented a post Permian radiation; however the calibrations used did not incorporate these early fossils (Li et al. 2012). Less is known regarding the history of Sternorrhyncha and Auchenorrhyncha although both groups are present and diverse in late Mesozoic deposits (Grimaldi & Engel 2005).

1.4.7. Holometabola

Collectively the Holometabolan orders, defined as those insects which conduct complete metamorphosis, represent over three quarters of all insect species and close to half of all described species on earth (Grimaldi & Engel 2005). The monophyly of the group has never been seriously challenged and it is the growing consensus regarding its internal relationships from both molecular e.g. (Castro & Dowton 2005; Savard et al. 2006; Zdobnov & Bork 2007; Cameron et al. 2009; Wiegmann et al. 2009; McKenna & Farrell 2010; Longhorn et al. 2010; Meusemann et al. 2010; Niehuis et al. 2012; Peters et al. 2014; Misof et al. 2014), and morphological data (Beutel et al. 2011), that represents one of the great triumphs of modern hexapod systematics.

Based on these, and related studies, the ordinal phylogeny of Holometabola is now more-or-less settled (Trautwein et al. 2012; Yeates et al. 2012) with Hymenoptera being recognized as basal and the remaining orders clustering in two large assemblages; Neuropteroidea (Neuropterida, Coleoptera, Strepsiptera), and Mecopterida, with the later

including two traditional groupings, Amphiesmenoptera (Trichoptera and Lepidoptera) and Antliophora which includes Diptera, "Mecoptera" and Siphonaptera. Outstanding issues, discussed in detail below, include the monophyly and association of "Mecoptera" within Antliophora (Whiting 2002; Wiegmann et al. 2009; McKenna & Farrell 2010), the ordinal relationships within Neuropterida (Winterton et al. 2010; Peters et al. 2014), and the precise placement of Strepsiptera with respect to Coleoptera (McKenna & Farrell 2010; Niehuis et al. 2012; Boussau et al. 2014).

1.4.7.1. Hymenoptera- wasps, ants and bees

Figure 11 Examples of major lineages of Hymenoptera (1.4.7.1); A) "Symphyta", B) "Parasitica", C) Aculeata

Photo Attribution- All pictures used under a Creative Commons license, and sourced from http://commons.wikimedia.org

- **A. Marshal Hedin 2008- http://commons.wikimedia.org/wiki/File:Common_sawfly_%28F._Tenthredinidae%29_%282657533824%29.jpg**
- **B. pompilid 2006- http://commons.wikimedia.org/wiki/File:Buprestid_parasite.jpg**
- **C. Karumar 2013-http://commons.wikimedia.org/wiki/File:Abeja-004.jpg**

Hymenoptera are the smallest of the "big four" Holometabolan orders (based on described species (Grimaldi & Engel 2005)) and are now recognized as the basal member of the lineage. The key division within Hymenoptera is ecological, separating a primitive paraphyletic assemblage of basal taxa, the majority of whom have foliage feeding or wood boring larvae ("Symphyta" or sawflies), from the more derived groups, which are overwhelmingly parasitoids of other invertebrate taxa (Apocrita). From within the latter are derived the stinging wasps (Aculeata), a large and important clade that includes such ecologically key taxa as the "yellowjackets" (social wasps), bees and ants (Grimaldi & Engel 2005). Within this broad scheme recent phylogenetic work has focused on

establishing infra-ordinal and superfamily relationships, with a particular focus on the origination of the key Aculeata groups.

Within "Symphyta" phylogenetic relationships in recent studies more or less mirror traditional opinions such as the position of Xyeloidea as sister to the rest of the order (Sharkey et al. 2012) (although alignment dependent in (Heraty et al. 2011)), the monophyly of Tenthredinoidea (the "true sawflies") (Heraty et al. 2011; Sharkey et al. 2012) and the recognition of the clade Unicalcarida uniting the various plant boring sawfly groups Cephoidae, Siricoidae and Xiphydrioidea, with the parasitic Vespina. Within Unicalcarida relationships are unstable, with morphology favoring the traditional view of the wood boring Xiphydrioidea as sister to Vespina (Vilhelmsen et al. 2010; Sharkey et al. 2012), while recent molecular evidence favors the herbaceous boring Cephoidae (Heraty et al. 2011). Traditionally also treated within "Symphyta", the status of Orussidae as the basal member of Vespina has been challenged in some recent molecular studies (e.g. (Heraty et al. 2011)) a view that would render paraphyletic the traditional clade Apocrita defined, by the distinctive "wasp waist" contraction of the first and second abdominal segment, although the clade was retained in the total evidence study of (Sharkey et al. 2012).

The majority of Apocrita families comprise the problematic grade "Parasitica", whose major features, particularly above the superfamily level, are the key issue facing hymenopteran systematics (Rasnitsyn 2010; Ronquist et al. 2012). The most recent studies have begun to suggest the presence of two broad assemblages: Proctotrupomorpha (Rasnitsyn 1988); which includes the "micro-hymenoptera" traditionally placed in the superfamilies' Platygastroidea, Cynipoidea, Diaprioidea, Mymarommatoidea, Chalcidoidea and "Proctotrupoidea" (Heraty et al. 2011; Vilhelmsen et al. 2010; Munro et al. 2011; Sharkey et al. 2012), and Evaniomorpha; which includes Megalyroidea, Ceraphronoidea, Trigonaloidea, Evanioidea and Aculeata (Rasnitsyn 1988). Support for the later (and for the monophyly of Aculeata) is very low in most datasets (Heraty et al. 2011; Sharkey et al. 2012). Within this scheme the placement of the large lineage Ichneumonoidea remains unclear ((Sharkey et al. 2012) favour a placement close to Proctotrupomorpha), as are the positions of the basal families Stephaniidae and Megaspilidae (linked with Orussidae in (Heraty et al. 2011)).

Despite low support in some recent molecular datasets the monophyly of Aculeata remains widely accepted among hymenopteran systematics, which may reflect limits in the available data (Sharkey et al. 2012). Key questions within the group concern the origin of

the two ecologically important groups of eusocial insects, the predatory ants (Formicidae) and the pollinating bees (Anthophila), which in their own way have each played a role in revolutionizing post-Cretaceous ecosystems (Grimaldi & Engel 2005). One of the biggest insights in recent years has been to recognize that the traditional superfamily "Vespoidea", appears to be paraphyletic (Pilgrim et al. 2008; Debevec et al. 2012; Johnson et al. 2013) which has led to various attempt to reexamine the placement of its former components (which includes the ants). Current theories regarding ant origins place them as sister to a loose assemblage of taxa which include Scoliidae, Bradynobaenidae (in part) and Apoidea (the last superfamily including bees) (Pilgrim et al. 2008; Debevec et al. 2012), or as sister to bees and their immediate relatives (Crabronidae) (Johnson et al. 2013).

Adding to the ambiguities regarding the early evolution of Hymenoptera, the fossil record of the group shows a remarkable rapid radiation of forms beginning in the Late Triassic, without obvious precursors among the Palaeozoic fauna (Rasnitsyn & Quicke 2002; Grimaldi & Engel 2005; Rasnitsyn & Zhang 2010; Warnock et al. 2011; Ronquist et al. 2012). Given that phylogeny implies the group must have been present since at least the Carboniferous (Nel et al. 2007; Nel et al. 2013) this apparent gap in the record is puzzling, leading some authors to revisit ideas regarding links to the mysterious stem Holometabolan order Miomoptera (Ronquist et al. 2012). The group has also been subject to molecular clock studies, including a total evidence analysis supporting a Permian diversification (well before the earliest known fossil evidence) (Ronquist et al. 2012) and genomic studies looking at the late Jurassic divergences within Aculeata (Wilson et al. 2013).

1.4.7.2. Neuropterida- lacewings, alderflies, and snakeflies

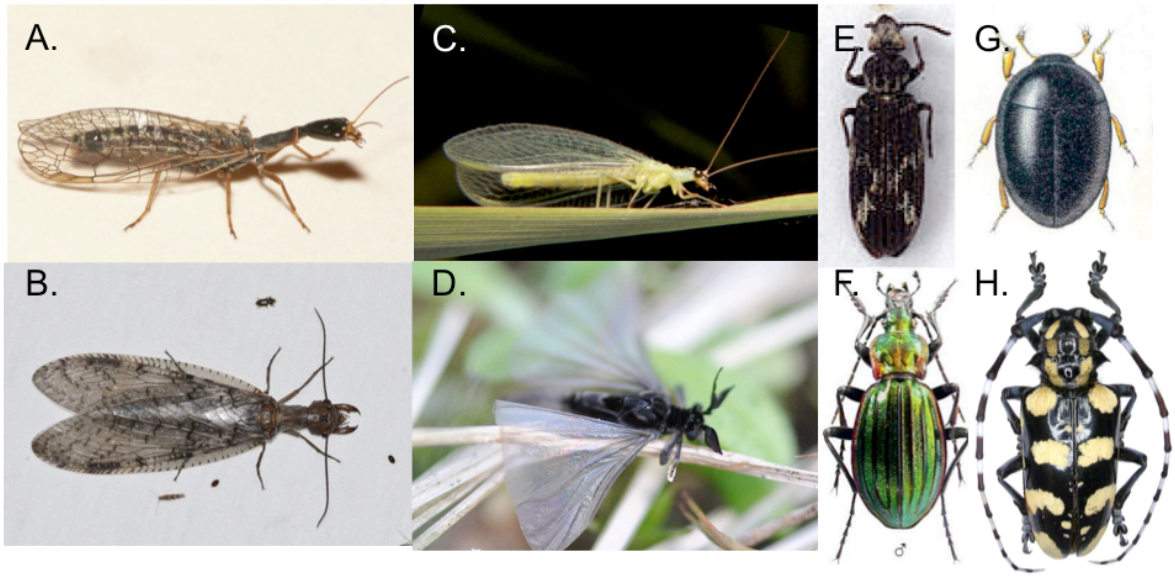


Figure 12 Examples of neuropteroid orders; Neuropterida (1.4.7.2), A) Raphidoptera, B) Megaloptera, C) Neuroptera; D) Strepsiptera (1.4.7.3); Coleoptera (1.4.7.4); E) Archostemata, F) Adephaga, G) Myxophaga, H) Polyphaga

Photo Attribution- All pictures used under a Creative Commons license, and sourced from http://commons.wikimedia.org

- **A. Dick Belgers 2012- http://commons.wikimedia.org/wiki/File:Xanthostigma_xanthostigma.jpg**
- **B. Andy Reago & Chrissy McClarren 2014- http://commons.wikimedia.org/wiki/File:Dobsonfly_%2814586103566%29.jpg**
- **C. Bruce Marlin 2005- http://commons.wikimedia.org/wiki/File:Green_lacewing.jpg**
- **D. Aiwok 2011- http://commons.wikimedia.org/wiki/File:Stylops_melittae_m1.jpg**
- **E. David R. Maddison 1996 -http://commons.wikimedia.org/wiki/File:Omma_sagitta.jpg**
- **F. M.Tarrier 1965 http://commons.wikimedia.org/wiki/File:Soltar.jpg**
- **G. Reitter 1908 - commons.wikimedia.org/wiki/File:Sphaerius.acaroides.Reitter.tafel64.jpg**
- **H. Udo Schmidt 2003 - http://commons.wikimedia.org/wiki/File:Cyriocrates_horsfieldi_%28Hope,_1842%29_%283266354577%29.jpg**

Neuropterida comprise a small group of relictual Holometabola, conventionally divided into three orders Raphidoptera (snakeflies), Megaloptera (alderflies, and dobsonflies) and Neuroptera (lacewings). All three groups comprise relatively large bodied Holometabola with, for the most part, predatory larvae. While most Neuropterida are terrestrial, Megaloptera and some of the basal Neuroptera (Nevrorthidae, Sisyridae and

some Osmylidae) are aquatic during the larval stage, with the primitive ecology for the group as a whole being highly controversial (Aspöck et al. 2012).

Relationships between the three orders are controversial (Winterton et al. 2010; Beutel et al. 2011; Trautwein et al. 2012; Peters et al. 2014) with some studies going so far as to challenge the monophyly of Megaloptera, although this appears to be associated with highly modified, or mis-sequenced rRNA attributed to one of the families (Winterton et al. 2010). The most recent study, which uses comprehensive genomic data, supports Raphidoptera as the basal taxon, with a monophyletic Megaloptera as sister to Neuroptera (Peters et al. 2014), a placement that is also consistent with morphological analyses (Beutel et al. 2011). Within Neuroptera the monophyly of the traditional suborder Hemerobiiformia has been challenged, and various families have subsequently been proposed as basal within the group including Coniopterygidae (Winterton et al. 2010), Nevrorthidae (Haring & Aspöck 2004) and Sisyridae (Zimmermann et al. 2011) (reviewed in (Aspöck et al. 2012)). In contrast with many other Holometabolan groups the fossil diversity of Neuropterida exceeds its extant richness (Labandeira & Sepkoski Jr 1993; Grimaldi & Engel 2005), leading to an interpretation of the group as relictual (see (Nicholson 2012) for review).

1.4.7.3. Strepsiptera

Strepsiptera are miniaturized and highly modified parasitoids that were until recently the most phylogenetically problematic holometabolan order. The group is characterized by a number of remarkable features including the extreme reduction of forewings (recalling the halters of true flies), larviform females that never (except in the most basal family) leave their invertebrate hosts, and most importantly for understanding their phylogenetic history, a highly modified structure for rRNA molecules (Xie et al. 2009) and a genome showing strong compositional bias (A/T richness) when compared with other hexapod groups (Gillespie et al. 2005). It was on the basis of these rRNA markers that early analyses e.g. (Whiting & Wheeler 1994; Whiting et al. 1997; Whiting 1998) linked the group with Diptera (forming the clade “Haltaria” (Whiting et al. 1997)), and the issue has become one of the most well known examples of the widespread error in phylogenetic analysis known as long branch attraction (Huelsenbeck 1998).

More recently there has developed a strong consensus from both morphological (Friedrich & Beutel 2010a; Beutel et al. 2011; Koeth et al. 2012; Peters et al. 2014), and

molecular studies (Wiegmann et al. 2009; McKenna & Farrell 2010; Longhorn et al. 2010)(McKenna & Farrell 2010)(Longhorn et al. 2010) placing Strepsiptera as close to Coleoptera (Niehuis et al. 2012; Trautwein et al. 2012; Peters et al. 2014). Problematically the monophyly of Coleoptera with respect to Strepsiptera has not been robustly established on all molecular markers (McKenna & Farrell 2010), although the most recent studies which have included genomic data from the basal coleopteran suborder Archostema (see below) have rejected nesting Strepsiptera within the beetles (Niehuis et al. 2012; Boussau et al. 2014; Misof et al. 2014). In the interests of simplicity, and in order to reduce problems of long branches associated with this group, I here treat Strepsiptera as a single terminal taxon and refer readers interested in the relationships between the seven included families to (McMahon et al. 2011).

1.4.7.4. Coleoptera- beetles

Subject to a well-known "inordinate fondness" as well as an astonishing array of ecological diversity, Coleoptera is, in terms of described species, the largest of the holometabolan orders and was the first to undergo systematic study using molecular data at the family level (Hunt et al. 2007). This topology, the raw data for which is partially used here, remains the standard framework for discussing the group, with the majority of other recent discussion focusing on either deep relationships among the four suborders (Beutel & Haas 2000; Hughes et al. 2006; Friedrich et al. 2009; Lawrence et al. 2011) or on relationships within particular superfamily groups (e.g. (Kundrata et al. 2014)). The four suborders of Coleoptera differ substantially in their ecologies, with the small suborder Archostema and Myxophaga being wood-boring fungus feeders and aquatic algal scrapers respectively (Arnett & Thomas 2000), while the larger predominately predatory Adephaga are split among aquatic ("Hydradephaga") and terrestrial forms (Geoadephaga) (Maddison et al. 2009). The vast majority of the group's diversity falls within the largely terrestrial subfamily Polyphaga, which includes five major series: Staphyliniformia, Scarabaeiformia (=Scarabaeoidea), Elateriformia, Bostrichiformia and Cucujiformia, all of which show substantial ecological diversity.

The relationships of the suborders remain controversial, with traditional views, placing Archostema as sister to the remaining sub-orders (with Myxophaga sister to Polyphaga) (Beutel & Haas 2000; Friedrich et al. 2009; Hughes et al. 2006) conflicting with more recent morphological studies favoring a topology of ((Archostema,

Adephaga),(Myxophaga, Polyphaga)) (Lawrence et al. 2011), and with molecular data, which supports Adephaga and Polyphaga as sister groups with Archostema being nested within Myxophaga (Hunt et al. 2007) (although see (Maddison et al. 2009) for an alternative opinion).

Within Adephaga phylogenetic questions focus on whether the groups shows a single (Ribera et al. 2002; Hunt et al. 2007) or multiple (Maddison et al. 2009; Dressler & Beutel 2010; Beutel et al. 2013) transitions from water to land (or vis-versa) (reviewed in (Jardine 2010)). In Polyphaga the deep phylogenetic structure is uncertain e.g. the respective monophylies of Staphyliniformia, Elateriformia and Bostrichiformia, (Hunt et al. 2007; Lawrence et al. 2011). One major clade that has been robustly supported on molecular and morphological data is Cucujiformia (Hunt et al. 2007; Lawrence et al. 2011), containing over half of all beetle families, and of particular interest due to containing the spectacularly rich plant feeding lineages Curculionoidae (weevils or snout beetles) and Chrysomeloidae (leaf and longhorn beetles) (often grouped as phytophaga (Lawrence et al. 2011)). Due to their ecological and economic importance, these last groups have received further phylogenetic attention, e.g. (Gómez-Zurita et al. 2007; McKenna et al. 2009), although their derivation and relationships with the other Cucujiformia, most of which are small bodied and cryptic fungus feeding beetles, remains poorly understood.

Until recently the earliest recognized stem Coleoptera was also the earliest known Holometabolan, *Adiphlebia lacoana* from the Carboniferous Mason Creek formation of the USA (Béthoux 2009). However this placement has been challenged, which renders the earliest known coleopterans as various Permian Protocoleoptera, primarily known from Russian deposits (Kukalová-Peck & Beutel 2012). Extant families of three of the suborders Archostema (Cupediae) (Martins-Neto et al. 2006), Adephaga (Carabidae) (Grimaldi & Engel 2005) and Polyphaga (Staphylinidae) (Chatzimanolis et al. 2012) are known deposits from the Middle to Late Triassic, with the earliest Myxophaga reported from Cretaceous mud-stones (Cai et al. 2012) and ambers (Kirejtshuk 2009).

1.4.7.5. Mecoptera and Siphonaptera –scorpion flies and fleas

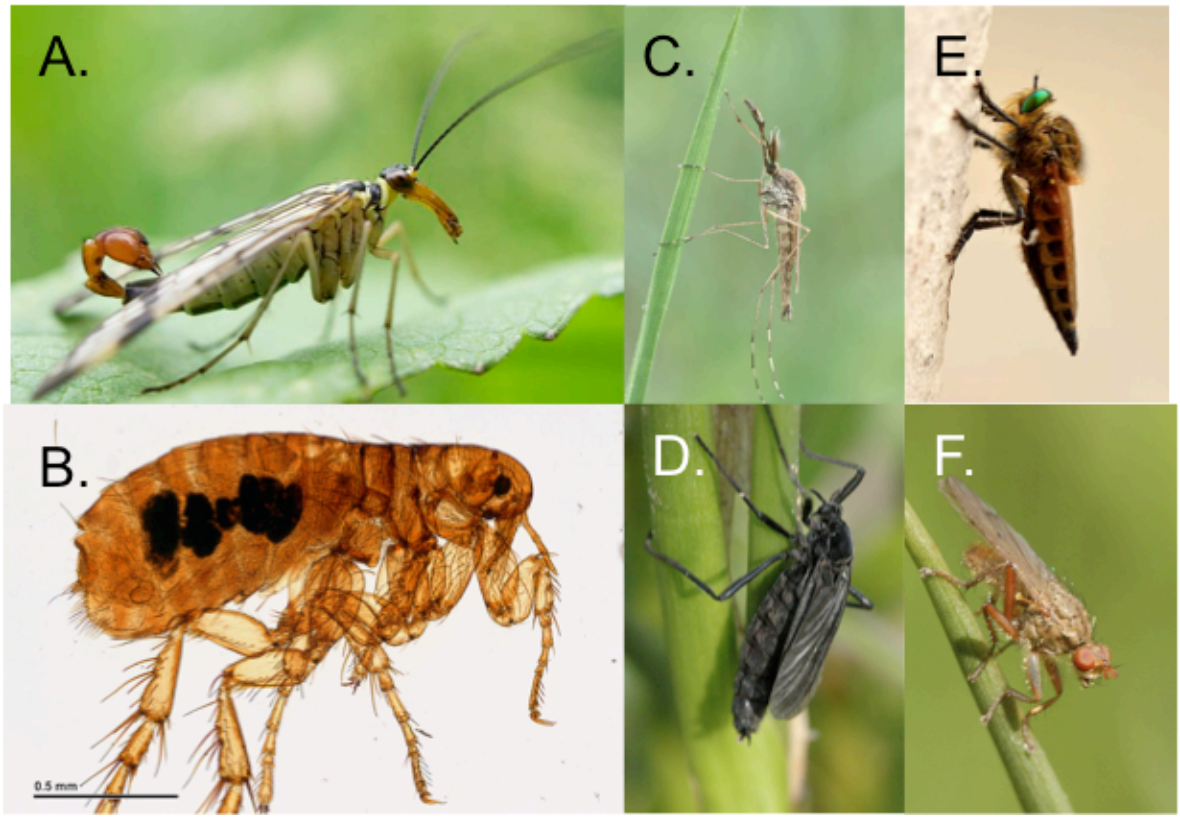


Figure 13 Examples of major lineages of Antliophora; "Mecoptera" (1.4.7.5); A) Pistillifera, B) Siphonaptera; Diptera (1.4.7.6); "Nematocera"; C) Culicomorpha, D) Bibionomorpha; Brachycera; E) "Orthorrhapha", F) Schizophora

Photo Attribution- All pictures used under a Creative Commons license, and sourced from http://commons.wikimedia.org

- **A. Ian Kirk 2013- http://commons.wikimedia.org/wiki/File:Scorpion_Fly_%28male%29_%289057482915%29.jpg**
- **B. Josef Reischig 2014- http://commons.wikimedia.org/wiki/File:Flea_%28251_01%29_Aphaniptera;_total_preparation.jpg**
- **C. Mathias Krumbholz 2011- http://commons.wikimedia.org/wiki/File:Culicidae_03_%28MK%29.jpg**
- **D. James Lindsey 2005- http://commons.wikimedia.org/wiki/File:Macropelopia.nebulosa.jpg**
- **E. Grassjewel 2014- http://commons.wikimedia.org/wiki/File:Asilidae_robber_fly.jpg**
- **F. James Lindsey 2006- http://commons.wikimedia.org/wiki/File:Scathophaga.furcata.jpg**

Mecoptera are another cryptic and possibly relictual group, whose placements with respect to other Antliophora remain one of the major outstanding problems in Holometabolan phylogenetics (Trautwein et al. 2012; Yeates et al. 2012). Most treatments recognize four basic divisions, the placements of which, with respect to each other and to

Diptera, remain highly contentious. One of these divisions includes the parasitic fleas, which were formerly treated as their own order (Siphonaptera). Evidence placing fleas within a paraphyletic Mecoptera has come from a number of molecular studies, proposing links with the flightless northern hemisphere family Boreidae (Whiting 2002; Kjer 2004; Kjer et al. 2006), although this view is less favored in studies using nuclear molecular markers (Wiegmann et al. 2009; McKenna & Farrell 2010; Ishiwata et al. 2011; Misof et al. 2014) (see also (Beutel et al. 2011)). Also problematic is the placement of the obscure Gondwanan family Nannochoristidae, which has unique aquatic larvae, and has traditionally been seen as sister to the rest of the order (Friedrich & Beutel 2010b; Fraulob et al. 2012).

All other families comprise the infra-order Pistillifera that is dominated by two relatively large families, the scavenging Panorpidae ("true" scorpionflies) and the predatory Bittacidae ("hangingflies") and whose phylogeny above the family level remains poorly established (Whiting 2002). Within fleas there has been some attempt to resolve the phylogeny with recent efforts revealing Tungidae as sister to the rest of the order, and also confirms the monophyly of most of the recognized families, with exceptions of Hystrichopsyllidae, Leptopsyllida and Ctenophthalmidae, all be it with low support above the family level (Whiting et al. 2008).

The fossil record of Mecoptera and Siphonaptera are active areas of research, in part with an aim to resolve the phylogenetic ambiguities of the group. A major recent advance has been the recovery from Cretaceous deposits in China of stem group fleas, the study of which is providing fresh insights into the origins of ecto-parasitism within the group (Huang et al. 2012; Gao et al. 2013).

1.4.7.6. Diptera- Flies

Of the four "mega diverse" holometabolan orders, Diptera is the most ecologically varied and contains taxa feeding on almost any conceivable organic resource (see (Marshall 2012)). Progress in establishing the phylogeny of Diptera has been rapid in recent years with the publication of major new morphological (Lambkin et al. 2013) and molecular datasets (Bertone et al. 2008; Trautwein et al. 2010; Wiegmann et al. 2011; Caravas & Friedrich 2013).

The most basal Diptera are now recognized as highly modified aquatic groups associated with fast flowing mountain streams in the Northern hemisphere

(Deuterophlebidae, and possibly also Nymphomyiidae) (Wiegmann et al. 2011; Lambkin et al. 2013) and aquatic or semi-aquatic larvae are common among a number of the basal dipteran lineages including: Tipulomorpha (crane flies), Culicomorpha (mosquitoes and midges), Psychodomorpha (sand flies and scavenger flies) and Blephariceromorpha (net winged midges). The order of branching in this part of the tree is highly uncertain and receives low support in both morphological (Lambkin et al. 2013) and molecular studies (Bertone et al. 2008; Caravas & Friedrich 2013). The remaining, more terrestrialised, Diptera comprise the clade Neodiptera whose basal members form a monophyletic group including the fungus midges, gall midges and march flies known as Bibionomorpha (Wiegmann et al. 2011; Caravas & Friedrich 2013).

All of the flies described above belong the paraphyletic suborder "Nematocera". Flies with reduced antenna, as well as a variety of other morphological features, comprise the robustly supported monophyletic suborder Brachycera. Relationships among the basal Brachycera remain contentious with (Wiegmann et al. 2011) grouping many of these taxa into a modestly supported clade "Orthorrhapha", most of whose members (with the exception of soldier flies- Stratiomyomorpha) have predatory larvae and are divided in traditional classifications into three infra-orders "Tabanomorpha", "Asilomorpha" (sensu stricto- excluding Empidoidae) and "Stratiomyomorpha"). The status of "Orthorrhapha" remains ambiguous as the group lacks morphological support (Lambkin et al. 2013) and was not recovered in the molecular sensitivity analysis of (Caravas & Friedrich 2013). By contrast the sister group relationship between Empidoidae (dance flies and long-legged flies) with remaining Diptera (the clade Cyclorrhapha) is robustly supported in all datasets confirming traditional ideas regarding the placement of these taxa (Eremoneura; (Wiegmann et al. 2011; Caravas & Friedrich 2013; Lambkin et al. 2013)).

The Cyclorrhapha includes three (possibly four) clades, the relationships between which have important implications for the history of larval development in Diptera, and which were investigated with micro-RNA data by (Wiegmann et al. 2011). Phoridae and its close relatives (the superfamily Platypezoidae) share with the more derived Cyclorrhapha (Schizophora) a number of important developmental innovations including patterns of gene expression in the early embryo (Wiegmann et al. 2011). Despite this, both micro-RNA work and multi-gene datasets support Syrphoidea (hoverflies) as sister to Schizophora (Wiegmann et al. 2011), with the status of Pipunculidae, usually treated within Syrphoidea but sister to Schizophora in (Wiegmann et al. 2011), remaining

uncertain (Lambkin et al. 2013). This topology is also robustly supported in the molecular sensitivity analysis of (Caravas & Friedrich 2013).

The phylogeny of Schizophora is the major outstanding problem in Dipteran systematics with over half of fly families being grouped among vaguely defined super-families and a large number systematic questions regarding isolated and rare genera (Marshall 2012). A notable member of this messy grade is the model "fruit fly" *Drosophila*, the potential sister groups of which are discussed in (Wiegmann et al. 2011). The most well-resolved clade within Schizophora is Calyptratae which includes at its base the blood feeding, ecto-parasitic superfamily Hippoboscoidea, and then a series of familiar groups such as house flies (Muscidae), dung flies (Scathophagidae), blow flies (Calliphorida) and the parasitoid tachinids (Tachinidae) whose relationships and reciprocal monophyly remains an area of active contention (Marinho et al. 2012; Caravas & Friedrich 2013).

As with Hymenoptera, Diptera first appear in the fossil record as a diverse radiation (in this case in the middle Triassic) with early members linked to many of the major clades including basal Brachycera (Krzemiński & Krzemiński 2003; Blagoderov et al. 2007). While the fossil record of basal Diptera is rich, due to their small size and delicate form, Schizophora are extremely rare, with the earliest example being a leaf mine trace from the early Paleocene of Montana, USA (Winkler et al. 2010).

1.4.7.7. Trichoptera- Caddisflies

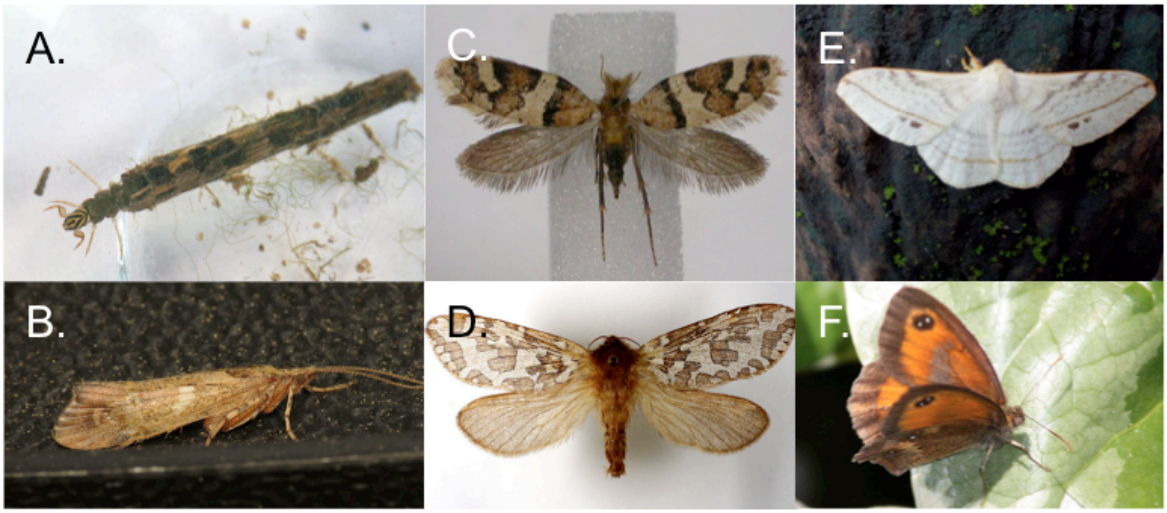


Figure 14 Major lineages of Amphiesmenoptera; Trichoptera (1.4.7.7); A) Larva, B) Adult; Lepidoptera (1.4.7.8); C) Micropterigidae, D) basal Glossata, E) Macroheterocera, F) Papilionoidea

Photo Attribution- All pictures used under a Creative Commons license, and sourced from http://commons.wikimedia.org

- **A. MyForest 2001- http://commons.wikimedia.org/wiki/File:Caddisfly_Larva.jpg**
- **B. Donald Hobern 2014 - http://commons.wikimedia.org/wiki/File:Pending_%2814253583265%29.jpg**
- **C. Neville Hudson 2009-http://commons.wikimedia.org/wiki/File:Sabatinca_lucilia.jpg**
- **D. Virtala 2006 -http://commons.wikimedia.org/wiki/File:Hepialus_fuscoargenteus.jpg**
- **E. Zuhairali 2010- http://commons.wikimedia.org/wiki/File:Moth_Kerala.jpg**
- **F. Charlesjsharp 2013- http://commons.wikimedia.org/wiki/File:Gatekeeper_hedge_brown.jpg**

The long-recognized sister group to Lepidoptera, adult Trichoptera are cryptic, and short lived insects that resemble small drab moths with wings covered by fine hairs (Holzenthal et al. 2007). The larvae are aquatic and in the most familiar examples constructing protective cases from organic or inorganic debris. Such larvae play key roles in a wide variety of aquatic ecosystems, with the majority acting as detritivores or filter feeding, with a limited number of derived predatory forms (reviewed in (Holzenthal et al. 2007))

Case-building behavior provides the basis for traditional systematics in Trichoptera with three suborders being recognized; the “retreat” forming Annulipalpia, the mobile case building Integripalpia and the free living “Spicipalpia”, the last being almost certainly

paraphyletic and basal to the clade as whole (Holzenthal et al. 2007; Malm et al. 2013) (although see (Wiggins 2004)). The most recent phylogenetic treatment suggested that Annulipalpia and Integripalpia are sister groups and recovers Rhyacophilidae as basal among the "Spicipalpia" (Malm et al. 2013), although support within the latter is low and other recent publications have reported different arrangements of the families involved, e.g. (Holzenthal et al. 2007).

The earliest definitive Trichoptera are Jurassic and a recent molecular clock study placed the divergence between Trichoptera and Leipdoptera as around 234 MA during the late Triassic (Malm et al. 2013). As is typical of aquatic groups, the record of Trichopteran families is comparatively rich (Grimaldi & Engel 2005), and the earliest examples of case building are known from the early Jurassic of Siberia, consistent with a rapid early divergence of the groups in question (Malm et al. 2013).

1.4.7.8. Lepidoptera- moths and butterflies

By contrast with the other mega-diverse orders, Lepidoptera are, with only a few isolated exceptions, exclusively plant feeders, making the group one of the most important global consumers of angiosperm tissue and, as a result, subject to close coevolution with the flowering plants (Ehrlich & Raven 1964; Powell et al. 1998). Recent phylogenetic analyses have challenged some of the long held views from traditional morphological phylogenies of the group e.g. (Kristensen 1999) particularly with respect to the derived clade "Macrolepidoptera" (Minet 1990); traditionally uniting the butterflies- Papilionoidea (which contrary to some recent analyses e.g. (Cho et al. 2011) appear to be monophyletic (Heikkilä et al. 2011; Kawahara & Breinholt 2014)) with the relatively large bodied "macro-moths" of the super-families Bombycoidea, Geometroidea, Lasiocampoidea and Noctuoidea (now united as the clade Macroheterocera (Regier et al. 2013)). "Macrolepidoptera" is now recognized as paraphyletic with butterflies instead falling at the base of the more inclusive group Obtectoma (Regier et al. 2013; Kawahara & Breinholt 2014) which also includes the small bodied Gelechioidea (Kaila et al. 2011) and Pyraloidea.

Towards the base of the tree the status of the mandibulate moths (Micropterigidae) as sister to the rest of the order (Wiegmann et al. 2000; Wiegmann et al. 2002) has recently been challenged, with evidence linking them to another lineage of basal moths, the Agthiphagidae (Regier et al. 2013). Other areas of contention and low support include the

phylogeny of the basal Glossata, moths with a proboscis, (Regier et al. 2013) and in particular the identification of the sister group to Ditrysia (moths with partitioned female reproductive tracts including >90% of all species) with (Regier et al. 2013) favoring part of the paraphyletic "Palaephatidae", as well as the monophyly of the traditional superfamily Tineoidae (Regier et al. 2013; Mutanen et al. 2010). When compared with other large insect groups the fossil record of Lepidoptera is exceptionally poor (see (Sohn et al. 2012) for a review of the known taxa). Like their sister group the earliest definitive Lepidoptera are Jurassic in age, e.g. (Huang et al. 2010).

1.5. Overall thesis aims

The main hypotheses that I aim to test within this thesis and the relevant background information are summarized above (Section 1.2). To reiterate my goals are:

1. To construct a dated phylogenetic framework based on our current understanding of hexapod relationships, that can form the basis for addressing issue of taxonomic diversification within the clade, within the limits of current data (Chapter 2)
2. To test which, if any, of the following morphological innovations (origin of the insect body plan, flight, wing folding and complete metamorphosis) correspond with shifts in the diversification rate of hexapods and to examine other patterns of rate shifts responsible for the extant variation in richness among different sub-clades (Chapter 3)
3. To formally test, in the context of an explicit phylogeny, ideas relating to co-evolution and plant feeding as drivers of hexapod diversity and to extended this framework to looking at patterns associated with dietary ecology more generally (Chapter 4)
4. To test the association between body size and diversity among hexapod families and examine the macro-evolutionary mechanisms underpinning size evolution in the group (Chapter 5).

2. Constructing and Dating the Hexapod Tree

2.1. Abstract

This section describes the data and protocols used in the construction of the dated topology for Hexapoda used in subsequent chapters to infer diversification patterns within the group. A global phylogenetic framework for Hexapoda was inferred from eight widely sampled molecular markers, both nuclear and protein coding regions and rRNA sequences from public databases, in combination with literature-defined constraints to control the placement of unstable taxa. The final data-matrix had a length of 7021bp and was 50.69% complete at the nucleotide level for the 874 included terminal taxa. Topology was inferred using the ML algorithm RAxML, and the resulting tree dated using an independent gamma rates clock implemented in MrBayes 3.2. Calibration of the clock was based on 86 fossil constraints implemented as hard minimum bounds. Recovered ordinal relationships corresponded closely with the consensus hexapod phylogeny presented in Chapter 1. However, as in previous studies using molecular markers, our methods struggled to resolve diverse and rapidly evolving groups, such as Apocrita (Hym.), "backbone" relationships of Polyphaga (Col.), Schizophora (Dip.) and Apoditrysia (Lep.). With respect to dating, our results show greatest similarity with previous studies sharing a common philosophy regarding implementation of calibration fossils, although there are outstanding issues related to differences in the phylogenetic resolution of different datasets. Our findings support a crown radiation of hexapod taxa well before the earliest fossils of the group, potentially back as early as the Ordovician, with a Middle Paleozoic ordinal diversification followed by family level divergences in the Late Permian and early Mesozoic.

2.2. Introduction

Understanding patterns of diversification requires a phylogenetic perspective on the group in question in order that the location of shifts in diversification rate and their relationship to potential traits of interest can be correctly identified (Stanley 1998). The phylogenic relationships among the Hexapoda have traditionally been difficult to infer due to limitations on informative morphological characters and inadequacies in the fossil record in documenting transitional forms (Whitfield & Kjer 2008; Béthoux 2009). As a result previous attempts to incorporate phylogenetic information into the study of diversification in the group have either lacked an explicit basis for the relationships used e.g. (Mayhew 2002; Mayhew 2003) or are subject to methodological issues due to the use of supertrees and non-independent source data, e.g. (Davis et al. 2010; Davis et al. 2011). The recent expansion of molecular markers for hexapod groups (Terry & Whiting 2005; Regier et al. 2008; Wiegmann, Trautwein, et al. 2009; McKenna & Farrell 2010; Ishiwata et al. 2011), as well as enhancements in analytical techniques, have resulted in a growing consensus regarding the deep relationships within the group (reviewed in Section 1.4). As a result of these developments it is now feasible to attempt to infer an explicitly based phylogeny for the clade as a whole in order to provide new insights into the pattern of diversification.

The use of molecular data to define relationships in Hexapoda has the further advantage that it allows the dating of the phylogeny through the use of relaxed molecular clocks (Rutschmann 2006; Drummond et al. 2006). For hexapods this is particularly significant as much of the known fossil record of the group is restricted deposits of exceptional quality (Lagerstätten) (Grimaldi & Engel 2005) and as a result their record is often incomplete (Wills 2001) and subject to geological or sampling noise (Labandeira & Sepkoski Jr 1993; Davis et al. 2010; Nicholson et al. 2014). This effect is most notable during the Paleozoic when insect bearing deposits are rare (Engel & Grimaldi 2004; Béthoux 2009; Garrouste et al. 2012), as well as for lineages with low preservation potential such as Protura, Pthiraptera (lice) and Siphonaptera (fleas) (Grimaldi & Engel 2005; Huang et al. 2012). In resolving the inadequacies of the fossil record a dated phylogeny of the group provides the ability of explore the impact of specific historical links to hexapod diversification, for example Permo-Triassic mass extinction (Labandeira

& Sepkoski Jr 1993; Grimaldi & Engel 2005; Labandeira 2005) and the Cretaceous rise of the flowering plants (angiosperms) (Fiz-Palacios et al. 2011; Clarke et al. 2011).

2.3. Topological reconstruction

2.3.1. Methods

The raw sequence data for the inference of the Hexapod phylogeny used in this study was extracted from Genbank (Benson et al. 2013), and the curated ribosomal database SILVA (Pruesse et al. 2007) and using BLAST (Altschul et al. 1990) searches targeted on *Drosophila* sequences for genes previously used for phylogenetic inference above the family level. Accessions for the target sequences used are shown in tables linked to Appendix 7.1, and were taken directly from previous publications or a published complete mitochondrial genome (ref AJ400907) (Azou & Bregliano 2001). Our initial survey included fourteen genes collectively encompassing the set used in previous super-matrix studies (Kjer et al. 2006; Regier et al. 2008; Wiegmann, Trautwein, et al. 2009; Ishiwata et al. 2011). Following reviewers comments on an early draft publication, reharding high levels of introduced missing data as compromising the underlying inference model (17.1% nucleotide complete dataset of 21,634 bp length, driven primarily by markers available only at the ordinal level, i.e. for approximately 36 of the 880 terminal taxa), this set was reduced to only encompassed the most widely sampled markers, all of which had family level coverage in at least one or two major orders. This reduced set, on which all the analyses presented here are based, incorporated eight molecular markers, including nuclear (CAD, Ef1α, PGD) and mitochondrial (COI, COII) protein coding genes and 16S, 18S and 28S rRNA.

Given that our goal was to investigate patterns at a higher taxonomic level due to the resolution of available richness data and fossil calibrations we took the unusual step of combining sequence data from multiple studies and species in order to maximize coverage across the different studied markers i.e. the terminal taxa used are chimeras with sampling favoring those with highest similarity to the BLAST target. This differs from the more common practice of using sequences for a single taxon (if not a single individual) as an exemplar for the placement of a larger group. While the former approach, which is common in studies using secondary data e.g. (Peters et al. 2011; Fiz-Palacios et al. 2011), has the benefit of maximizing the often sparse coverage of markers across taxonomic groups (Springer et al. 2004; Campbell & Lapointe 2009) it does entail an assertion of

monophyly for the group in question (i.e. that there is no conflict in the signal between different gene partitions (Yoshizawa & Johnson 2010; Leigh et al. 2011; Simon et al. 2012; Caravas & Friedrich 2013)) which as noted in Chapter 1 is not universally true of all hexapod families. In a sense this issue of assumed monophyly is moot, as the downstream analyses of species richness patterns themselves assume that all terminals are monophyletic groups (Chapters 3-5); however it does also introduce a level of conflict among markers that may contribute to phylogenetic instability (Nosenko et al. 2013). In the absence of high resolution phylogenetic data and in the face of current taxonomic understanding, we felt that the benefits of using the most complete available sequences for each marker for each terminal group outweighed any introduced ambiguities in taxonomic placement, particularly given the already sparse nature of available sequence data. Thus we present our work as contingent on current taxonomic assignments within hexapods and expect that future improvements in these hypotheses will enhance the model presented here.

The taxonomy of most hexapod groups was resolved to the family level (tips listed in Appendix 7.1), following Genbank up to August 2013 in order to maintain consistency across genetic partitions. Exceptions to the family level resolution include the Hemiptera suborder Sternorrhyncha (Aphidoidea (10 families), Coccoidea (23 families), Phylloxeroidea (2 families) and Psylloidea (7 families)) and the small parasitic order Strepsiptera (7 families). In the former case this was the result of taxonomic conflict between Genbank and the consulted sources of species richness estimates, while in the case of the latter this represented a deliberate strategy to minimize problems of long branch attraction known to be associated with this taxon (Huelsenbeck 1998; Pohl & Beutel 2005; Niehuis et al. 2012). In total the groups included on the presented tree incorporate a total of 903 of the approximately 1100 recognized extant hexapod families, with the remainder being excluded on account of a lack of suitable sequence information.

The eight markers sampled were individually aligned using MAFFT (local pair distances, max iterations=1000) (Katoh et al. 2002), with the exceptions of 18S and 28S which were aligned using an automated profile alignment based on the structural reference database SILVA (Pruesse et al. 2007). This approach was selected due to the combination of time efficiency and the preservation of conserved structural elements that previous studies have shown to be significant in accurate phylogenetic reconstruction within Hexapoda (Misof et al. 2007; Wang et al. 2013). All partitions were subjected to Gblocks

(Talavera & Castresana 2007) in order to remove regions of poor alignment, and limited overlap among taxa, with the minimum conserved block length set to three for protein coding genes and two for ribosomal sequences. Third codon positions for all protein-coding sequences were also excluded due to the risk of substitution saturation (Ho & Jermiin 2004) following the testing of preliminary datasets using the index proposed by (Xia et al. 2003) implemented in the program DAMBE (Xia & Xie 2001). In total the concatenated sequence had a length of 7021bp and was 50.69% complete at the nucleotide level (aligned sequences linked to in Appendix 7.1).

2.3.1.1. Constraints on Topology

Recent studies attempting to infer broad scale phylogenies for diverse groups have typically relied on a tiered system of phylogenetic inference such that datasets with restricted taxonomic sampling have been used to provide a backbone in order to guide the placement of the rest of the tree e.g. (Wiens et al. 2005; Jetz et al. 2012). For Hexapoda the distribution of sequence information currently available in Genbank (particularly for the reduced set of genes considered here) means that there are restrictions on the availability of data for the most controversial nodes within the tree and as a result a backbone approach fails to adequately control the placement of unstable taxa (Whitfield & Kjer 2008; Trautwein et al. 2012; Yeates et al. 2012). Instead we adopted a constraint-based approach whereby relationships that have been recovered as consistently well supported in previous studies were used to guide the reconstruction of the tree used here. In order to define suitable constraints the available phylogenetic literature for hexapods was reviewed since 2005 (Section 1.4), and used to define suitable constraints in accordance with the following principals:

- A constraint must reflect a recognized and named systematic group that is widely accepted within the literature of the appropriate taxa, defined by reference to appropriate encyclopedic sources e.g. (Grimaldi & Engel 2005; Resh & Cardé 2009).
- Recovery with strong support, defined as bootstrap support of at least 95 under maximum likelihood or parsimony, or alternatively Bayesian posterior probability of 0.99, in all relevant recent molecular phylogenic studies with sufficient resolution to inform the family level analysis conducted here. Where both maximum likelihood and Bayesian trees were available the former were favored

due to the well known tendency for Bayesian analyses to give higher confidence values relative to conventional bootstrap procedures (Alfaro et al. 2003; Douady et al. 2003; Erixon et al. 2003) and in order to maximize methodological comparability with the process of topological inference used here.

Constraints meeting these criteria are listed below with example references that include the most recent and comprehensive taxonomic treatments of the relevant groups. In addition to these we also constrained the monophyly of recognized orders with the following exceptions where paraphyly is known or suspected (Trautwein et al. 2012; Yeates et al. 2012) i.e.: "Blattodea" with respect to Isoptera (Inward et al. 2007), "Pscoptera" w.r.t Pthiraptera (Yoshizawa & Johnson 2010) and "Mecoptera" w.r.t. Siphonaptera (Whiting 2002) and Phasmatodea w.r.t. Embiodea (Zrzay 2008; Friedemann et al. 2012).

Finally, we also used constraints to restrict the movement of representatives of the unstable polyneopteran order Zoraptera which was extremely data-deficient with respect to the studied markers and whose position was strongly influenced by biases induced by the uniquely modified nuclear rRNAs of this taxon (Yoshizawa & Johnson 2005; Yoshizawa 2010; Wang et al. 2013). Traditional classifications placed Zoraptera at the base of Paraneoptera (e.g. (Hennig 1969)); however recent opinion favors an unresolved placement somewhere within Polyneoptera (Terry & Whiting 2005; Yoshizawa & Johnson 2005; Misof et al. 2007; Yoshizawa 2007; Friedrich & Beutel 2008; Ishiwata et al. 2011; Simon et al. 2012; Misof et al. 2014). The constraint used here was based on the most recent study, and the only one to be based on protein-coding markers independent of the problematic rRNAs, which recovered Zoraptera as sister to Dictyoptera (Ishiwata et al. 2011), a position which is also accepted in a recent review of hexapod ordinal relationships (Trautwein et al. 2012). In the absence of constraint Zoraptera was recovered as sister to all other Neoptera, a position that has never been supported in previous work and which has serious consequences for the pattern of diversification due to low richness of this lineage.

Implemented constraints (Figure 15);

- **Ordinal relationships**- reviewed in (Trautwein et al. 2012; Yeates et al. 2012): Insecta, Dicondylia, Pterygota, Neoptera, Holometabola
- **Hymenoptera** (Heraty et al. 2011; Debevec et al. 2012; Wilson et al. 2013): Unicalcarida, Vespina, Aculeata, Chrysidoidea, Anthophila

- **Coleoptera** (Hunt et al. 2007; Lawrence et al. 2011): Hydrophiloidea, Scarabaeoidea, Cucujiformia, Curculionoidea
- **Diptera** (Wiegmann et al. 2011; Caravas & Friedrich 2013; Lambkin et al. 2013): Brachycera, Schizophora
- **Lepidoptera** (Regier et al. 2009; Mutanen et al. 2010; Regier et al. 2013): Glossata, Rhopalocera

2.3.1.2 Tree inference

Tree inference was conducted using the rapid maximum likelihood routine RAxML (Stamatakis et al. 2005; Stamatakis et al. 2008) implemented on the CIPRES webserver (Miller et al. 2010). This procedure was chosen on the basis of rapid implementation and the availability of computational resources, and represents the industry standard in current generation phylogenetic analyses. The matrix in RAxML was partitioned such that codons one and two for nuclear and mitochondrial genes respectively were allotted separate GTR+CAT models, as were each of the three RNA partitions included. Reported bootstrap values are calculated using an MR-based stopping criterion (Stamatakis et al. 2008) resulting in 650 replicated bootstrap samples.

2.3.2. Results

At an ordinal level the recovered topology was broadly consistent with the current consensus regarding hexapod relationships (Section 1.4). Regions of strong support include the increasingly well-established relationships within Holometabola, with Hymenoptera as basal (Bootstrap support- BS = 98) to the super-ordinal groupings Neuropteroidea (Neuropterida, Coleoptera, Strepsiptera; BS 99) and Mecopterida (BS 100), with the later including two traditional groupings; Amphiesmenoptera (Trichoptera and Lepidoptera; BS 100) and Antliophora (Diptera, “Mecoptera” and Siphonaptera; BS 97).

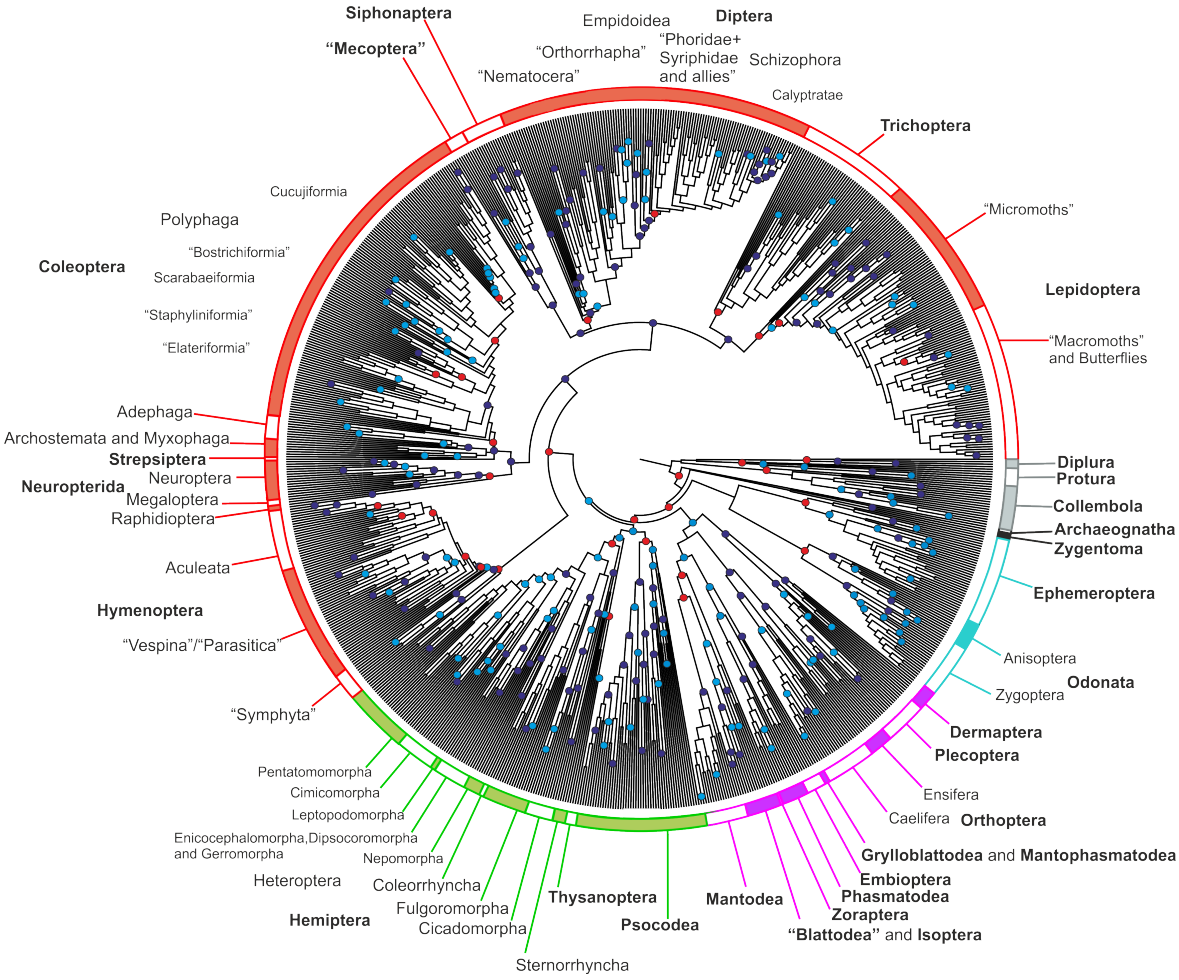


Figure 15 Nodal support on the phylogeny. Nodes marked with circles are either constrained (red) or have high bootstrap support (light blue 50-80%, dark blue: over 80%). Membership of major hexapod clades is denoted by coloration of the ring (Grey: Entognatha, Black: basal insects, Cyan: Palaeoptera, Purple: Polyneoptera, Green: Paraneoptera, Red: Holometabola). Common names for the labeled clades are shown on Figure 18 (Section 3.4.).

With respect to controversial regions of the tree; among the basal winged insects we recover extremely weak support (BS 20) for a monophyletic Palaeoptera (Whitfield & Kjer 2008) consistent with recent morphological (Blanke et al. 2012; Blanke et al. 2013)(Blanke et al. 2013) and genomic findings (Thomas et al. 2013); as well as modest support for the monophyly of Polyneoptera (BS 59) and Paraneoptera (BS 70) and Nonoculata (uniting Diplura and Protura (BS 99)). Within Polyneoptera few conclusions can be drawn regarding ordinal relationships (partially due to the constraint imposed on Zoraptera- see above) although we do find strong evidence for an unnamed grouping of Plecoptera and Dermaptera (BS 96) that is also present on (Kjer et al. 2006; Misof et al.

2007; Ishiwata et al. 2011; Wang et al. 2013), Xenonomia (Grylloblatodea + Mantophasmatodea , BS 93) (Terry & Whiting 2005), as well as weak support for Eukinolabia (Phasmatodea +Embioptera BS 58) (Terry & Whiting 2005; Friedemann et al. 2012). Within Paraneoptera Hemiptera was modestly supported as sister to Thysanoptera, i.e. recovering the clade Condylognatha (BS 68).

As discussed in Section 1.4 many within-ordinal relationships in hexapod groups remain highly contentious and have often been subject to low support and instability in previous phylogenetic surveys. Unsurprisingly, given that we draw exclusively on previously published sequence data, the results presented here show a similar pattern, encompassing a mixture of well-supported and unstable relationships across different parts of the tree (local support summarized on Figure 15). Particularly problematic areas are often associated with regions of the tree that are suspected as having been subject to rapid diversification of large lineages (Whitfield & Kjer 2008) for example Apocrita (Hym.) (Heraty et al. 2011; Sharkey et al. 2012), the "backbone" relationships of Polyphaga (Col.) (Hunt et al. 2007), Schizophora (Dip.) (Wiegmann et al. 2011; Caravas & Friedrich 2013) and Apoditrysia (Lep.) (Mutanen et al. 2010; Cho et al. 2011; Regier et al. 2013), many of which are also areas of ambiguity with respect to the taxonomic limits and monophyly of families e.g. (Marshall 2012). Despite recent progress it seems likely that it will be some years and requiring large amounts of targeted sequencing before these issues are completely addressed and therefore we regard the work presented here as an important first step in defining a testable baseline against which such improvements can be assessed.

To briefly review the areas of phylogeny that do show strong support, beginning at the base of the tree, we recover strong support for Acerentomata within Protura (although with the addition of Berberentomidae) (Dell'Ampio et al. 2011) and find modest evidence to suggest that Sinentomidae may be more derived than in previous studies, with Fujientomidae being recovered as the basal family (BS 60). In Diplura the conventional grouping of the three included taxa is recovered within a monophyletic Japygomorpha (BS 95) (Grimaldi & Engel 2005). Collembola show low support for the family level relationships, driven in part by instability in the placement of Neelidae (Xiong et al. 2008); however we do find strong support for a clade approximating Poduromorpha (BS 93), which is weakly linked with the family Isotomidae (BS 57), traditionally placed in Entomobryomorpha, raising questions reading the monophyly of the latter group. Implemented constraints govern the placement of the primitively wingless insects,

specifically the enforced monophyly of Dicondylia (in the absence of which Zygentoma appears grouped with Odonata, possibly as a result of shared GC compositional bias in 18S rRNA).

As noted previously our results recover Palaeoptera as monophyletic, albeit with trivial support (BS 20). Within the Ephemeroptera, relationships are almost completely unresolved with the only nodes of any note approximating the burrowing mayflies of the Fossoriae (BS 78), although this group is also found to include the non-burrowing Siphlonuridae and Ichthybotidae, both of which are conventionally treated in a more basal position (Ogden et al. 2009). By contrast, within Odonata relationships follow a far more conventional structure, with both the suborders recovered as monophyletic (Epiprocta BS 89, Zygentoma BS 91), and Epiophlebiidae (Anisozygoptera) nested deep within Anisoptera as opposed to being its sister (BS 49) (as in (Bybee et al. 2008) but conflicting with (Dumont et al. 2010)). Strongly supported is a clade which combines members of the superfamily Libelluloidea with Corduliidae and Macromiidae (BS 94), both of the latter being families for which the inference of monophyly has been problematic (Dumont et al. 2010). Within Zygoptera we recover the division between Lestoidea and the remaining taxa (BS 65, although the former is not recovered as monophyletic BS 67 (Dumont et al. 2010)) and otherwise family relationships are poorly supported .

Within the Polyneopteran orders, Dictyoptera follow (Djernæs et al. 2012), showing a monophyletic Mantodae as sister to “Blattodea”/ Isoptera (BS 77), with the later closely linked to the wood roaches of the family Cryptoceridae (BS 72). Within Isoptera the placement of Mastotermittidae is unresolved although the clades of the higher (Rhinotermitidae, Serritermitidae and Termtidae, BS 95) and lower (Hodotermitidae, Termopsidae and Kalotermitidae, BS 61) termites are both recovered. Dermaptera lacks any clear resolution, although the topology of the studied families (which do not include the two parasitic lineages Arixeniina and Hemimerina due to lack of suitable sequence data) is broadly consistent with (Kocarek et al. 2013). Plecoptera are surprisingly well resolved, with moderate support placing Notonemouridae at the base of the order (BS 72), consistent with (Thomas et al. 2000), and support for the monophyly of the southern hemisphere lineages Antarctoperlaria (BS 61) and for the suborder Systellognatha (BS 93) which includes the predatory super-family Perloidea (BS 98). Within Orthoptera all of the suborders and infra-orders are recovered as monophyletic, although as expected in Ensifera there is low support above the family level (Legendre et al. 2010). In Caelifera support

values are also often low, particularly within the important infra-order Acridoidea although the topology is broadly consistent with other recent opinions regarding relationships within the group (Song 2010; Leavitt et al. 2013). Neither Embiodae nor Phasmatodae show strong support for the resolution of their family relationships although the latter is recovered as monophyletic (see constraints Section 2.3.1.1) with low support linking *Timema* (BS 39) to the remaining taxa (Euphasmatodea (BS 100)).

At the base of Paraneoptera relationships within the Psocodea show modest support for the monophyly of Pthiraptera (BS 68) (Yoshizawa & Johnson 2010) which, with the associated families Liposcelididae and Pachytroctidae, are weakly placed as sister to the rest of the order (BS 54) . Within Pthiraptera Amblycera (BS 80) is sister to the rest of the clade (BS 91), with "Ischnocera" forming a paraphyletic series leading to sister grouping of Anoplura (BS 88) and Rhychophthirina (BS 89). In the remaining Psocodea, Trogiomorpha (BS 86) is sister (BS 77) to clade uniting Psocomorpha (BS 99) and Amphientometae (BS 63) which are recovered with modest support (Yoshizawa & Johnson 2010). Thysanopteran relationships are unresolved. Within Hemiptera Sternorrhyncha (BS 96) are basal (BS 62), with very weak support for a monophyletic Auchenorrhyncha (BS 35) and the conventional sister grouping of Coleorrhyncha and Heteroptera (BS 44). Within the latter all of the recognized infra-orders are recovered as monophyletic with the exception of the aquatic bugs Nepomorpha, with both the family Corixidae (water boatmen) and Pleidae (pigmy backswimmers) recovered elsewhere (the former as basal to clade containing Enicocephalomorpha, Dipsocoromorpha (BS 99) and Gerromorpha (BS 89), and the latter as sister to Leptopodomorpha (BS 97)). Geocorisae is recovered with weak support (BS 55) and relationships within the Cimicomorpha (BS 30) are particularly unstable.

The phylogeny of Hymenoptera recovered here corresponds closely to that outlined in Section 1.4.7.1, with modestly strong support for deep relationships among "Symphyta" and very weak support for the majority of clades within Vespina (BS 72, including Orussidae as sister to Stephaniidae (Heraty et al. 2011)). Excluding Platygastroidea, there is some support for the Proctotrupomorph (BS 75) and within this grouping most of the recognized superfamilies are recovered, although with low support. Given that the monophyly of Aculeata was enforced it was not possible to make strong statements regarding its relationships although it should be noted that support values both within and around this important clade are extremely low.

In Neuropterida our results follow those of (Peters et al. 2014) in recovering a monophyletic Megaloptera (BS 100) as sister (BS 82) to Neuroptera (BS 98), thus rendering Raphidoptera (BS 100) basal and supporting a terrestrial ancestral larval state for the clade (Aspöck et al. 2012). Within Neuroptera, Coniopterygidae, Nevrorthidae, Sisyridae and Osmylidae form a poorly resolved basal series, with the remaining taxa robustly recovered as monophyletic (BS 86) including the sub-order Myrmeleontiformia (BS 63).

Within Coleoptera we find very weak support for a novel sub-ordinal topology of (((Archostema, Myxophaga)(BS 34), Adephaga)(BS 23), Polyphaga) conflicting with previous molecular work that has supported a sister grouping of Adephaga and Polyphaga (Hunt et al. 2007)(Maddison et al. 2009). Beyond the constrained nodes described above, there are relatively few regions of strong support within beetles, although the overall shape is rather similar to that proposed by (Hunt et al. 2007), including a monophyletic Geoadephaga (BS 67) nested within "Hydroadephaga" (Maddison et al. 2009), separation of the "basal four" (Hunt et al. 2007) from the remainder of Polyphaga (BS 83), and broad patterns of association within the rump Staphyliniformia, Elateriformia and Bostrichiformia. Cucujiformia is another large and almost completely unresolved group with the only region of strongly supported nodes representing the constrained superfamily Curculionoidea. Low support within Coleoptera is likely to be in part a product of a lack of data to constrain the placement of unstable taxa, as well as the absence of suitable markers, such as slow evolving nuclear genes (Maddison et al. 2009) to resolve deep nodes. The group is thus one of the regions of the tree that would benefit most from further in depth sampling e.g. (Kundrata et al. 2014).

As with other analyses that have included ribosomal markers, e.g. (Whiting 2002; Kjer 2004; Kjer et al. 2006), our tree suports a linkage between Siphonaptera (BS 100) and the mecopteran family Boreidae (BS 80), which together are placed as sister to the remaining Mecoptera (BS 100). The monophyly of the Pistillifera (the "true scorpion flies" excluding Nannochoristidae) is also confirmed with strong support (BS 91) with modest support for its internal relationships.

In Diptera there is very strong support for considering Deuterophlebidae as sister to the rest of the order (BS 93; our analysis excludes Nymphomyiidae (Wiegmann et al. 2011)), as well as for the monophyly of the various "Nematocera" clades and (Brachycera being constrained). Relationships among the basal brachyceran lineages remain poorly

resolved (as in (Wiegmann et al. 2011; Caravas & Friedrich 2013)) and we do not recover the controversial clade "Orthorrhapha". Eremoneura including Empidoidae and Cyclorrhapha is present and well supported (BS 95), and our tree shows the same relationships among the basal Cyclorrhapha as are present in (Wiegmann et al. 2011) (i.e. ((Phoridae, Platypezoidae)(BS 51), (Syrphidae (Pipunculidae, Schizophora)(BS 94))(BS 93))(BS 93);). Within Schizophora there are no well-supported relationships with the exception of Calyptratae (BS 96).

Relationships in both Trichoptera and Lepidoptera suffer generally low support due to the instability of small numbers of data-deficient taxa and the limited resolving power of the available markers. In Trichoptera the only clade with any meaningful support represents part of the sub order Annulipalpia (BS 98 although the large family Hydropsychidae is absent). Lepidopteran relationships are dominated by the implement constraints; however there is at least some support for the nested group Heteroneura (BS 88) and Ditrysia (BS 80) despite poor resolution within the Obtectoma.

2.4. Dating the Tree

2.4.1. Methods

Dating of the topology was conducted using an independent gamma rates clock implemented in MrBayes 3.2 (Ronquist, Teslenko, et al. 2012). The form of clock was chosen based on Bayes factor comparisons conducted on preliminary versions of the data matrix which significantly favored this form of clock over alternative models (Xie et al. 2011). The data were partitioned as above, and optimal models for each partition identified using PHYML, implemented in the TOPALI browser (Milne et al. 2009). Priors on the clock rates were based on those used in a comparable multigene analysis of Hymenoptera (Ronquist, Teslenko, et al. 2012) (overall clock rate; lognormal mean= 6.10, SD 2.458582, independent gamma rates variation; exponential mean= 37.12). Preliminary runs of the Bayesian chain identified a strong tendency for chains to become stuck on distinct local optima resulting in poor convergence and inadequate parameter sampling. In order to resolve this, runs were conducted under low temperature conditions (heating parameter 0.005) with the proposal frequencies of a number of parameters being modified in order to achieve acceptance rates within the recommended range (20-80%) (Ronquist et al. 2009). Two independent sets of four Markov Chain Monte Carlo (MCMC) chains were run for 12 million generations with sampling conducted every 500 generations with a burn-in fraction

of 50% necessary to remove the impact of all the suboptimal peaks obtained during sampling. Convergence on the stable distribution was assessed based on trace outputs analyzed in Tracer v1.5 (Drummond & Rambaut 2007; Rambaut & Drummond 2009) and on the adequate sampling of the majority of the model parameters (average effective sample sizes >200 and potential scale reduction factors of approximately 1) (Ronquist et al. 2009; Ronquist, Teslenko, et al. 2012). Dating was preformed locally on a Mac 3.4Ghz Intel Core i7 processor and required approximately 1100 hours to run to completion.

Calibration was conducted using 86 fossils taken from the recent palaeoentomological literature (Table 2). In defining calibrations we favored specific named fossils, as oppose to commonly used general references e.g. (Benton & Donoghue 2007), following the recommendations of (Parham et al. 2012). All calibrations were implemented as hard minimum bounds on a uniform distribution with a hard maximum based on the upper 95% CI on the basal divergence in Hexapoda in (Rota-Stabelli et al. 2013); 503 Ma. The same maximum was also placed on the age of the root node. Where available minimum radiometric dates for the relevant deposits have been used, otherwise deposits were treated as the age of the termination of the relevant stage in (Gradstein et al. 2012) (Table 2).

2.4.2. Results and Discussion

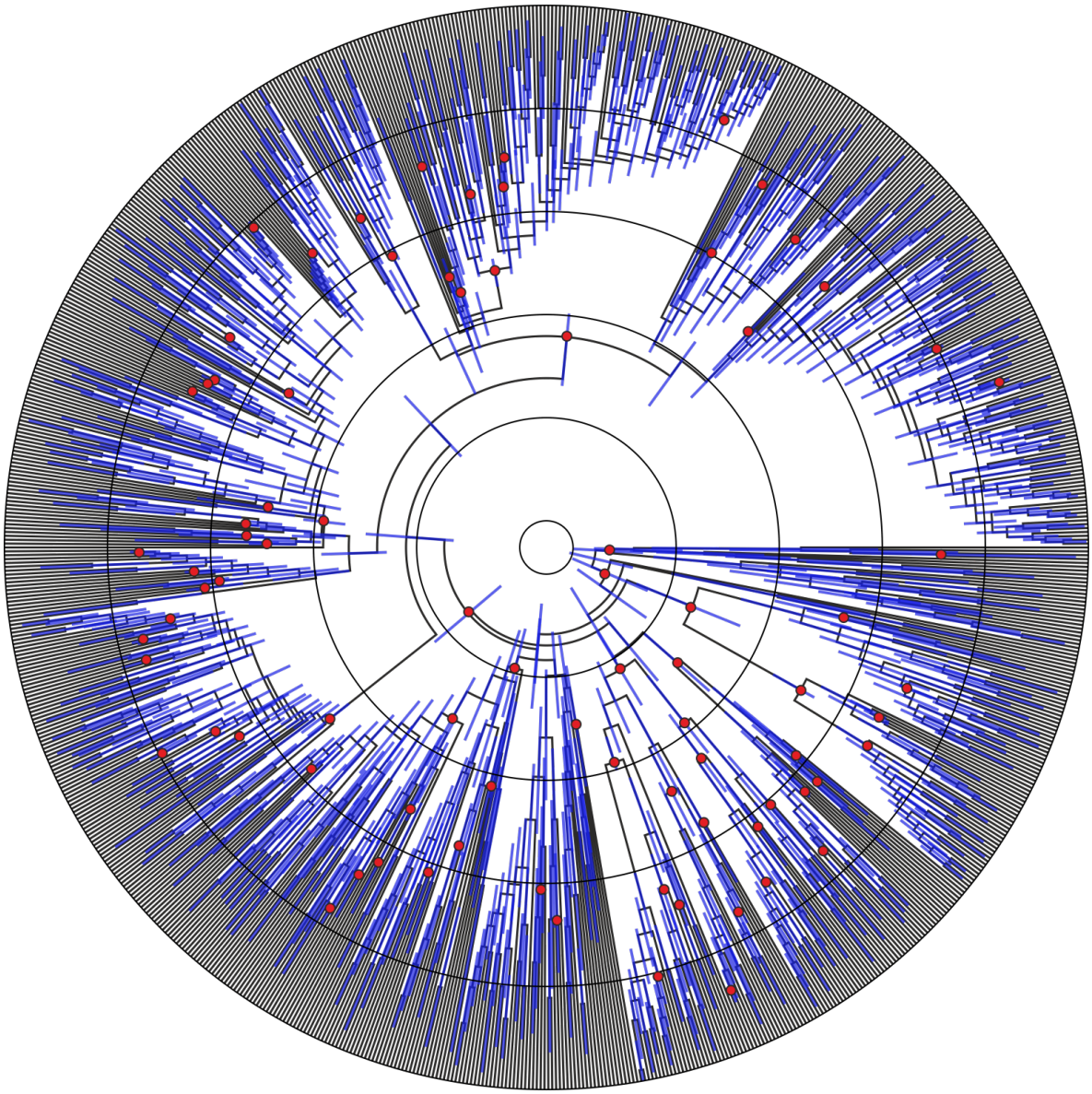


Figure 16 Topology showing 95% confidence intervals on node ages (transparent blue bars). Black rings denote 100Ma intervals from the present. Nodes denoted with red circles are involved in calibration (see Table 2). Tree shown in the same orientation as Figure 15

Comparisons between the node ages generated in this study and those of previous molecular clock analyses and the fossil record for Hexapoda are shown in Table 1. One of the major difficulties in making such comparisons lies in the fact that because of differences in the availability of sequence information used, the coverage of nodes and the estimated tree topology can be highly variable among studies, meaning that identifying sets of comparable nodes across different studies can be challenging (particular where explicit

dates are not given for all the included divergences, e.g. (Rota-Stabelli et al. 2013)). In addition there is substantial variation in calibration between studies in the choice of calibration points; reflecting both differences in the focus of studies (Giribet & Edgecombe 2012; Wheat & Wahlberg 2013), improvements in our understanding of the relevant fossil record (Grimaldi & Engel 2005; Benton & Donoghue 2007; Nel et al. 2013) and changes in the numerical geological timescale (Gradstein et al. 2012). Different studies have also differed in their the implementation of fossils (see Section 1.3.3.) either as; point data (Gaunt & Miles 2002), hard (this study, (Wiegmann, Kim, et al. 2009) and (Rehm et al. 2011)) or soft (Warnock et al. 2011; Rota-Stabelli et al. 2013) bounds, or explicit probability distribution (normal distributions in (Wheat & Wahlberg 2013)).

For comparative purposes I will here focus on three studies (Warnock et al. 2011) (Rehm et al. 2011) and (Wheat & Wahlberg 2013) which show the greatest overall similarity with the work conducted here in terms of both studied nodes and implemented calibrations. As a general rule the dates inferred in this study are typically somewhat older than those identified in these works with the greatest similarity (rarely more than 25 Ma age differences, confidence intervals usually overlapping) shown with (Rehm et al. 2011), with whom we share the same calibration philosophy (broadly distributed hard minimum bounds on node ages with, mostly, arbitrarily high hard maxima). The greatest contrast is seen with (Wheat & Wahlberg 2013), whose estimated ages we often exceed by 50 Ma. We attribute this difference primarily to the use by these authors of explicit, and normally distributed probability distributions (Ho & Phillips 2009).

The appropriate implementation of fossil calibrations for molecular clock studies remains an active area of debate ((Ho & Phillips 2009), Section 1.3.3.) and we recognize that previous work has shown that the use of uniform calibrations as is done here can lead to non-uniform effective priors (Heled & Drummond 2011). However we feel that, due to the incomplete nature of the hexapod fossil record (Wills 2001), attempts to explicitly calibrate node age maxima using probability distributions of fossil absences introduce unknowable levels of error (Nowak et al. 2013), relating primarily to taphonomic bias in the early hexapod record (Grimaldi & Engel 2005). In addition the use by these authors of symmetrical probability distributions for node calibration is problematic (Ho & Phillips 2009), as the result is to treat errors in phylogenetic placement (resulting in nodes being calibrated as earlier then the true age of divergence (Near et al. 2005)) as having equal weight to the uncertainty in maximum age, a position that may downgrade the quality of

hypotheses based on morphological inference. In this context soft bounded distributions, or asymmetric calibrations such as the log normal are more justifiable, and when comparing our results with (Warnock et al. 2011) which used such priors we find greater concordance in inferred node ages, particularly when confidence intervals are taken into account.

As noted in Chapter 1 the fossil record of hexapods is sporadic (Wills 2001), and due a lack of suitable terrestrial deposits (Kenrick et al. 2012), the earliest phases of the diversification of the group remain poorly understood (Giribet & Edgecombe 2012; Nel et al. 2013). Under these circumstances it should come as no surprise that the inferred age of many of the deep nodes within the tree, in both this and previous studies (Table 1), is substantially earlier than the first fossil evidence for the presence of such groups. For example we estimated the age of Hexapoda as a whole, i.e. the divergence of true insects from Entognatha (basal hexapods including springtails) as occurring in the Ordovician (mean estimate 474.4Ma, 95% CI 439.6-502.9Ma), compared with the earliest fossil evidence from the Rynie chert of Scotland dating to around 410Ma (early Devonian (Grimaldi & Engel 2005)(Parry et al. 2011)). Our date estimates suggest that by the time of these earliest fossil deposits, which include a collembolan *Rhyniella praecursor* (Whalley & Jarzembowski 1981)) and a possible winged insect; *Rhyniognatha hirsti*; 410.2 Ma (Engel & Grimaldi 2004), hexapods had already diversified to the point where Paraneoptera had likely separated from Holometabola (426 Ma, 95CI 384-468 Ma), and that by the time of the first diverse faunas (for example the famous Mason Creek fauna of the southern USA; Moscovian stage- 307 Ma (Grimaldi & Engel 2005; Nel et al. 2013)) almost all of the extant orders had diverged and several, notably Hemiptera and Psocodea, were beginning their basal radiations (although the latter may be artifact of elevated rates of substitution in the parasitic Pthiraptera (Cameron et al. 2011; Letsch et al. 2012) which are difficult to constrain given the lack of fossil evidence for the group).

Table 1: Estimated ages on key nodes from this study, the fossil record and other relevant clock analyses. For (Warnock et al. 2011), (Rehm et al. 2011) and (Rota-Stabelli et al. 2013); values are given for the mean of considered calibrations and datasets. Values given in parentheses are the 95% confidence interval associated with the age in question. Values given in braces are the calibration, if any, applied to the node in the relevant study. Different forms of calibration given include: point calibrations (denoted pt) hard minimum/maximum bounds (denoted min/max respectively), soft maxima (denoted s*max), and normal distributions (denoted by the mean (mu) and standard deviation (sd) of the prior distribution). Fossils are given as the radiometric dates of known deposits or termination of the relevant geological stage and were taken from (Nicholson 2012) except where noted. For calibrations used in this analysis see Table 2.

Clade	Estimated Crown Age (Ma)								
	This study	Earliest fossils	(Gaunt & Miles 2002)	(Regier et al. 2004)	(Wiegmann, Trautwein, et al. 2009) /(Wiegmann, Kim, et al. 2009)	(Warnock et al. 2011)	(Rehm et al. 2011)	(Wheat & Wahlberg 2013)	(Rota-Stabelli et al. 2013)
Hexapoda	479 (439-502)	410.2	-	485-488 (467-504)	-	-	485 (447-547) {min= 404}	433 (420-445)	483 (456-504) {max= 543, min= 395}
Insecta	462 (419-498)	410.2	-	402-415 (387-443)	-	-	455 (418-512)	Approximate age = 410 {mu= 425, sd=7)	455 (405-500) {max= 543, min= 383}
Pterygota	442 (397-480)	318	373- 388	368-385 (337-416)	-	-	419 (385–469) {min= 324.8}	384	-
Palaeoptera	373 (323-431)	228	-	-	-	-	388 (343-435)	326 (278-367)	-
Neoptera	431 (393-471)	318	-	-	-	-	397 (364-442)	346 (324-371)	-
Polyneoptera node A (Dermaptera-Dictyoptera)	401 (357-438)	As above		245- 279 (194-328)	-	-	-	-	-

Polyneoptera node B (Orthoptera – Dictyoptera)	387 (347-430)	As above	{pt= 323}	-	-	-	351 (310-389)	279 (218-334)	-
Eumetabola	427 (384-468)	315	-	-	355 {max= 360, min= 280}	476 (439-513) {min= 307.2, s*max=414}	NA	-	-
Paraneoptera	417 (369-457)	315	-	-	-	419 (368-470) {min=283.7, s*max =414}	-	275 (219-325)	-
Holometabola	390 (350-435)	Unclear ; see (Nel et al. 2013)	-	-	350 (336-359)	439 (404-476) {min=307.2 s*max= 414}	372 (340-412)	308 (292-325)	-
Hymenoptera	257 (231-282)	228	-	-	-	-	-	-	-
non hymenopteran Holometabola	362 (324-405)	Unclear; see (Nel et al. 2013);	-	-	300 (287-315)	406 (372-441){min= 307.2 s*max= 414}	353 (322-388)	288 (270-305){mu=300, sd=11)	-
Neuropteroidea	334 (307-371)	299 (Nel et al. 2013)	-	-	286 (274-299)	-		-	-
Neuropterida	300 (259-333)	279	-	-	255 (227-276)	-	-	-	-
Coleoptera	307 (288-329)	237	-	-	-	-	-	-	-
Mecopterida	321 (298-368)	265	338-351	-	282 (264-300)	373 (341-406){min= 238.5, s*max=295.4}	342 (311-373)	267 (249-285)	-
Amphiesmen-optera	302 (279-356)	Unclear	-	-	230 (190-261)	-	-	-	-
Lepidoptera	269 (232-324)	113 (Sohn et al. 2012)	-	-	-	-	-	-	-
Antliophora	317 (291-362)	Unclear/ As below	-	-	256 (234-279)	-	-	-	-
Diptera (Brachycera-Culicomorpha)	290 (269-313)	241	248-283	-	Approx. 240 (secondary calibration used in (Wiegmann et al. 2011))	325 (294-357){min= 238.5, s*max=295.4}	281 (251-295){ min= 238.5, max= 295.4)	227 (210–244){mu=230, sd=11)	-

At the time of the Permo-Triassic extinction event (approximately 252 Ma (Benton & Twitchett 2003; Gradstein et al. 2012)), traditionally seen as a major period of faunal turnover in hexapods (Labandeira & Sepkoski Jr 1993; Labandeira 2005; Nicholson 2012), we find that, consistent with previous work e.g. (Davis et al. 2010; Nicholson et al. 2014)(Nicholson 2012), limited evidence for signal of a hiatus or rebound with respect to the extant fauna (see Chapter 3 for further discussion). For the majority of taxa the family level radiation of groups is estimated as Mesozoic in origin e.g. Hymenoptera in the earliest Triassic (contrasting with the Permian dates given in (Ronquist, Klopfstein, et al. 2012)), although exceptions do occur, notably the late Permian diversification of Coleoptera and Diptera (Chapter 3). Given the coarse taxonomic level used in this study it is difficult to resolve the impact of more recent events, e.g. the Cretaceous diversification of Angiosperms (Fiz-Palacios et al. 2011; Clarke et al. 2011), or the K/T extinction event, as despite these likely playing a major role in hexapod diversification at the species level (See Chapter 3 and Chapter 4) their impacts on hexapod families are regarded as marginal based on fossil surveys (Labandeira 2005; Nicholson 2012).

2.5. Conclusions

This chapter summarizes the dated phylogenetic framework that underpins the comparative work conducted in the remainder of this thesis and illustrates the various data sources and procedures that went into its construction. As noted our goal in this work was to provide a comprehensive framework for the study of hexapod diversity, particularly with respect to times of divergence. Achieving this entailed data-driven compromise to make the best use of patchy available data, with the result being retention of phylogenetic ambiguity in regions of rapidly diverging terminal groups due to limited constraining information and sequence data. For a group as diverse as Hexapoda there will always be a trade-off between the (considerable) benefits of having a single explicit tree to provide the basis for general models vs. some level of ambiguity regarding the placement of some of the constituent lineages (see discussion regarding the widely implemented super-tree of all mammal species (Bininda-Emonds et al. 2007) e.g. (Meredith et al. 2011)). As our focus here was on whole tree processes of diversification and ambiguity applied throughout the studied datasets (e.g. estimated tip richness and ecological state- see subsequent chapters and Section 7), we felt comfortable in accepting the topology estimated here as an approximation for the hexapod phylogeny. We recognize that further data sampling other

improvements in taxonomic understanding are likely to refine the picture of hexapod relations described here (e.g. (Misof et al. 2014)). However based on current understanding as reviewed in Section 1.4 our topology shows no major deviations from the consensus view of hexapod relationships and thus represents a valid first attempt at establishing a comprehensive framework for Hexapoda, all be it one with outstanding regions of uncertainty reflecting rapidly diverging groups. Further discussion of possible improvements in phylogenetic understanding and a horizon scan of near future developments are provided in Section 6.1.

To date our favored topology, we have applied what is, to our knowledge, the most comprehensive set of fossil calibrations yet used with respect to Hexapoda. As discussed in Section 2.4.2 we find that the consistency of our inferred dates with those of previous studies is driven by similarities in the fossil derived distributions implemented (Ho & Phillips 2009; Parham et al. 2012; Nowak et al. 2013). This implies a strong need to justify our choice of calibration scheme. As noted above, our use of multiple widely dispersed hard minimum bounds on node ages is based primarily on a belief that the fossil record of hexapods, and in particular the association of fossil clades within the extant phylogeny, is insufficiently resolved to support explicit probability distributions on node ages. However, due to the large degree of sequence rate variability within the group, there is a need for many localized calibration points. Method selection also included practical considerations such as run-time (always a limitation when working on comparatively large trees) and the availability of suitable software implementations for some of the methods in question. As a general point there is a need to be critical regarding the validity of symmetrical probability functions for node calibration e.g. the normal distribution (Wheat & Wahlberg 2013), as it is unclear how these conform to the expectations of fossil data (Nowak et al. 2013). Given the rapid development of paleoentomology in recent years e.g. (Nicholson 2012; Clapham & Karr 2012; Nel et al. 2013), the calibrations used here represent only a snapshot of current knowledge and are subject to continual updating (e.g. compare with (Misof et al. 2014)). The consequences of this for further studies of hexapod diversification are very much dependent on precisely which nodes are impacted and how these interact within the joint model. Unfortunately computational limitations restricted our ability to make formal cross validation, see (Near et al. 2005; Sanders & Lee 2007; Marshall 2008; Pyron 2010) of such effects although this would provide a potentially valuable addition to subsequent dating studies.

Table 2: Fossil calibrations implemented in dating tree topology. Calibrated nodes are plotted on Figure 15. Where available radiometric date estimates are referenced on the first occurrence of the deposit, alternatively the relevant stage termination is given based on (Gradstein et al. 2012). Recovered age is given as the mean estimate from the post-burnin MCMC samples with associated confidence intervals.

Node	Fossil	Deposit	Age (Ma)	References	Recovered age (Ma, 95% CI's)
Stem Collembola	*Rhyniella praecursor*	Rhynie chert, Dryden Flags Fm., Aberdeenshire, Scotland	410.2 (Parry et al. 2011)	(Whalley & Jarzembowski 1981; Greenslade & Whalley 1996)	464 (432-500)
Stem Japygoidea	*Ferrojapyx vivax*	Crato Fm., Brazil	Aptian -113	(Wilson & Martill 2001)	230 (126-348)
Stem Dicondylia	*Rhyniognatha hirtsi*	Rhynie chert, Dryden Flags Fm., Aberdeenshire, Scotland	410.2	(Engel & Grimaldi 2004)	462 (419-498)
Palaeoptera	*e.g. Eugeropteron lunatum*	Malanzán Fm, Cuesta de la Herradura, Argentina	Serpukhovian-323	(Riek & Kukalova-Peck 1984)	373 (323-432)
Crown Odonata	*Triassothemis mendozensis*	Potrerillos Fm., Cerro Cachueta, Potrerillos Argentina	228 (Spalletti et al. 2008)	(Carpenter 1960; Davis et al. 2011)	243 (228-274)
Odonata Aeshnidae	*Sinacymatophlebia mongolica*	Jiulongshan Fm., Inner Mongolia, China	152 (Liu et al. 2006)	(Huang & Nel 2009a)	155 (152-167)
Odonata Hemiphlebidae	*Mersituria ludmilae*	Doronino Fm., Transbaikalia.	Berriasian-139	(Vasilenko 2005)	160 (139-195)
Ephemeroptera Leptophlebiidae	*Conovirilus poinari*	Lebanese Amber (Collection locality not reported)	Aptian -113	(McCafferty 1997)	230 (174-299)
Ephemeroptera-Baetiscidae	*Protobaetisca bechlyi*	Crato Fm., Brazil	Aptian -113	(Martill et al. 2007)	150 (124-194)
Stem Orthoptera	Unnamed Archaeorthoptera, specimen B13711 Municipal Museum of Ostrava, Czech Republic	Petřkovice Beds, Ostrava Fm., Upper Silesian Basin, Czech Republic	Serpukhovian-318	(Béthoux & Nel 2002; Prokop et al. 2005)	387 (347-430)
Crown Orthoptera	*Raphogla rubra*	Salagou Fm., Saxonian Group, Lodève Basin, France	Kungurian - 272	(Béthoux et al. 2002)	309 (272-368)
Orthoptera Gryllidae	*Araripegryllus orientalis*	Weald Clay Fm., UK	Barremian 126	(Gorochov et al. 2006)	195 (137-258)
Orthoptera-Caelifera	*e.g. Dzhajloutshella sp.*	Madygen Fm., Kyrgyzstan	Carnian- 228	(Gorochov 2005)	271 (232-325)
Orthoptera-	*Eoproscopia martilli*	Crato Fm.,	Aptian -113	(Heads 2008)	138 (113-

Proscopiidae		Brazil			174)
Orthoptera-Tridactylidae	*Cretoxya rasnitsyni*	Lulworth Fm., UK	Berriasian - 139	(Gorochov et al. 2006)	186 (139-236)
Stem Mantodea	*Mesoptilus dolloi*	Upper Coal Measures, Commentry Basin, France	Gzhelian - 299	(Béthoux & Wieland 2009)	307 (299-333)
Blattodea – Ectobiidae	*Piniblattella sharingolensis*	Sharin-Gol Fm., North Mongolia	Valanginian-134	(Vršanský 2005)	155 (134-203)
Stem Isoptera	*Baissatermes lapideus*	Baissa locality, Zaza Fm., Siberia, Russia	Valanginian-134	(Engel et al. 2007; Engel et al. 2009)	174 (134-212)
Mantodea-Chaeteessidae	*Arvenineura insignis*	Menat Puy-de-Dome locality, France	Palaeocene – 56	(Nel & Roy 1996; Grimaldi 2003; Grimaldi & Engel 2005)	96 (68-125)
Stem Embiodea	*Sinembia rossi, Juraembia ningchengensis*	Jiulongshan Fm, Inner Mongolia, China	152	(Huang & Nel 2009b)	260 (211-306)
Embiodea – Australembiidae	*Burmitembia venosa*	Burmese amber	98 (Shi et al. 2012)	(Engel & Grimaldi 2006a)	125 (98-169)*
Phasmatodea-Phyllidae	*Eophyllium messelensis*	Messel Fm., Germany	Middle Eocene- 47.8	(Wedmann et al. 2007)	59 (48-82)
Mantophasmatodea	*Juramantophasma sinica*	Jiulongshan Fm., Inner Mongolia, China	152	(Huang et al. 2008)	220 (152-269)
Stem Plecoptera	*Gulou carpenter*	Tupo Fm., Ningxia Hui Autonomous region, China	Bashkrian - 315	(Béthoux et al. 2011)	*356 (315-400)*
Plecoptera Capniidae	*Dobbertiniopteryx capniomimus*	Upper Lias, Dobbertin, Germany	Toarcian- 174	(Liu et al. 2009)	181 (174-199)
Plecoptera-Perlidae	*Archaeoperla rarissimus*	Yixian Fm., Liaoning Province, China	121.8 (Swisher et al. 1999; Sun et al. 2011)	(Yushuang et al. 2008)	128 (122-146)
Dermaptera- Labiidae	*Kotejalabis haeuseri*	Crato Fm., Brazil	Aptian -113	(Martill et al. 2007)	179 (119-238)
Dermaptera-Pygidicranidae	*Astreptolabis ethirosomatia*	Burmese amber	98	(Engel 2011)	*211 (120-282)*
Stem Thysanoptera	*Westphalothripides oudardi*	Isolated material taken from a slag heap, 'Terril N◦ 7', Avion, Nord, France	Bashkrian - 315	(P. Nel et al. 2012)	405 (356-445)
Pscodea Liposcelididae	*Cretoscelis burmitica*	Burmese amber	98	(Grimaldi & Engel 2006b; Yoshizawa & Lienhard 2010)	351 (291-410)
Pscodea - Sphaeropsocidae	*Sphaeropsocites lebanensis*	Lebanese Amber, Jezzine locaility, Central Lebanon	Aptian -113	(Grimaldi & Engel 2006a)	164 (113-229)
Pscodea - Compscocidae	*Burmacompsocus perreaui*	Burmese amber	98	(Nel & Waller 2007)	193 (112-275)
Stem Heteroptera	*Paraknightia magnifica*	Belmont Conglomerate Member,	Changhsingian - 252	(Grimaldi & Engel 2005)	336 (284-382)

		Croudace Bay Fm., New South Wales, Australia			
Hemiptera-Belostomatidae	Unnamed specimen, Virginia Museum of Natural History 727	Cow branch Fm., Virginia, USA	Carnian 228	(Fraser et al. 1996; Grimaldi & Engel 2005)	239 (228-267)
Hemiptera -Ochteridae	*e.g. Pristinochterus zhangi*	Yixian Fm., Liaoning Province, China	121.8	(Yao et al. 2007; Yao et al. 2011)	179 (124-247)
Hemiptera - Schizopteridae	*Libanohypselosoma popovi*	Amber, Hammana-Mdeirij locality, Boundary of Abeih Fm. and Chouf Sandstone Fm., Central Lebanon	Aptian-113	(Azar & Nel 2010)	160 (113-226)
Hemiptera- Saldidae	*Brevrimatus pulchalifer*	Yixian Fm., Liaoning Province, China	121.8	(Zhang et al. 2011)	173 (122-243)
Hemiptera- Gerridae	*Cretogerris albianus*	Amber from Archingeay-Les Nouillers, Charente-Maritime, France	Albian- 100	(Perrichot et al. 2005)	118 (100-177)
Hemiptera – Coreidae	*Kerjiecoris oopsis*	Huangshanjie Fm., Xinjiang Uygur Autonomous Region, China	Norian -209	(Lin 1992; Yao et al. 2012)	214 (209-226)
Hemiptera Cicadoidea (Tettigarctidae)	*"Liassiocicada" ignotata*	Lilstock Fm., UK	Rhaetian- 201	(Scherbakov 2009)	225 (201-267)
Hemiptera Cixiidae	*"Cixius" petrinus*	Weald Clay Fm., UK	Barremian 126	(Szwedo 2007; Szwedo et al. 2011)	191 (126-262)
Hemiptera-Aphidomorpha	*Leaphis prima (syn. Vosegus triassicus)*	Grès à Voltzia Fm., France	Anisian- 241	(Szwedo & Nel 2011)	286 (241-339)
Stem Holometabola	*Westphalomerope maryvonneae*	Terril n° 5bis, Vicoigne Series, France	Westphalian A, Bashkirian-315	(Nel et al. 2007; Labandeira 2011)	427 (384-468)
Hymenoptera-Xyelidae	*Triassoxyela foveolata, Leioxyela antiqua*	Madygen Fm., Russia	Carnian 228	(Ronquist, Klopfstein, et al. 2012)	257 (231-282)
Hymenoptera-Pelecinidae	*e.g. Archaeopelecinus tebbei*	Jiulongshan Fm., Inner Mongolia, China	152	(Rasnitsyn & Zhang 2004; Shih et al. 2010)	159 (152-175)
Hymenoptera - Heloridae	*Archaeohelorus hoi*	Jiulongshan Fm., Inner Mongolia, China	152	(Wang et al. 2005; Shih et al. 2011)	176 (152-206)
Hymenoptera - Bethylidae	*Lancepyris opertus*	Lebanese amber, Ain Dara locality, Central Lebanon	Aptian-113	(Azevedo & Azar 2012)	124 (113-148)
Hymenoptera-Figitidae	*Jerseucoila*	New Jersey	Turonian- 89.8	(Liu et al.	103 (90-

	plesiosoma	Amber, Raritan Fm., USA		2007)	141)
Hymenoptera – Pompilidae	*Bryopompilus interfector*	Burmese amber	98	(Engel & Grimaldi 2006b)	123 (98-157)
Hymenoptera-Anthophila	*Melittosphex burmensis*	Burmese amber	98	(Poinar Jr 2009; Shi et al. 2012)	*154 (98-180)*
Megaloptera - Sialidae	*Dobbertinia reticulata*	Upper Lias, Dobbertin, Germany	Toarcian- 174	(Wichard & Engel 2006)	193 (174-230)
Neuroptera - Mantispidae	*Liassochrysa stigmatica*	Upper Lias, Dobbertin, Germany	Toarcian- 174	(Wedmann & Makarkin 2007)	184 (174-209)
Neuroptera -Osmylidae	*Allotriosmylus uniramosus*	Jiulongshan Fm., Inner Mongolia, China	152	(Yang et al. 2010)	207 (152-257)
Neuroptera Myrmeleontidae	*Choromyrmeleon aspoeckkorum*	Yixian Fm., Liaoning Province, China	121.8	(Ren & Engel 2008; Makarkin et al. 2012)	*131 (122-151)*
Stem Coleoptera	*e.g. Moravocoleus permianus*	Bačov Fm., Obora, Czech Republic	Sakmarian-290	(Kukalová-Peck & Beutel 2012)	309 (290-331)
Coleoptera -Cupedidae	*Argentinocupes pulcher*	Los Rastros Fm., La Rioja Province, Argentina	Ladinian-237	(Martins-Neto et al. 2006)	255 (237-283)
Coleoptera-Hydroscaphidae	*Hydroscapha jeholensis*	Yixian Fm., Liaoning Province, China	121.8	(Cai et al. 2012)	235 (169-277)
Coleoptera-Carabidae	Unnamed specimen pictured and discussed in (Grimaldi & Engel 2005)	Cow branch Fm. Virginia, USA	Carnian 228	(Grimaldi & Engel 2005)	232 (228-246)
Coleoptera-Staphylinidae	*Leehermania prorova*	Cow branch Fm. Virginia, USA	Carnian 228	(Chatzimanolis et al. 2012)	236 (228-253)
Coleoptera- Lucanidae	*Juraesalus atavus*	Jiulongshan Fm., Inner Mongolia, China	152	(Bai et al. 2012; Nikolajev et al. 2011)	163 (152-188)
Coleoptera-Ochodaeidae	*Mesochodaeus daohugouensis*	Jiulongshan Fm., Inner Mongolia, China	152	(Nikolajev & Ren 2010)	160 (152-186)
Coleoptera Trogossitidae	*Sinopeltis jurassica*	Jiulongshan Fm., Inner Mongolia, China	152	(Yu et al. 2012)	159 (152-175)
Coleoptera Helophoridae	*Helophorus (Mesosperchus) inceptivus*	Shar-teg Fm., Gobi-Altai Province, Mongolia	Tithonian- 145	(Fikáček et al. 2012b; Fikáček et al. 2012a)	151 (145-172)
Coleoptera - Ithyceridae	*Karacar contractus*	Karabastau Fm., Karatau locality Kazakhstan	Oxfordian-157	(Gratshev & Legalov 2011)	161 (157-170)
Coleoptera- Silvanidae	*Pleuroceratos burmiticus*	Burmese amber	98	(Poinar Jr et al. 2008)	105 (98-119)
Coleoptera- Clambidae	*Eoclambus rugidorsum*	Amber, Hammana-Mdeirij	Aptian-113	(Kirejtshuk & Azar 2008)	254 (183-301)

		locality, Boundary of Abeih Fm. and Chouf Sandstone Fm., Central Lebanon			
Mecopterida (i.e. Stem Amphiesmenoptera)	*Cladochorista sp.,*	Chepanikha locality, Russia	Wordian-265	(Aristov & Bashkuev 2008; Minet et al. 2010)	321 (298-368)
Trichoptera – Philopotamidae	*Liadotaulis major*	Upper Lias, Dobbertin, Germany	Toarcian- 174	(Ansorge 2003; Hao & Huang 2012)	198 (174-229)
Trichoptera - Lepidostomatidae	*Eucrunoecia ridicula*	Weald Fm., UK	Barremian-126	(Sukatsheva & Jarzembowski 2001)	142 (126-175)
Trichoptera-*Psychomyiidae*	*Palerasnitsynus ohlhoffi*	Burmese amber	98	(Wichard et al. 2011)	*116 (98-150)*
Lepidoptera-Micropterigidae	Parasabatinca aftimacrai	Amber, Hammana-Mdeirij locality, Boundary of Abeih Fm. and Chouf Sandstone Fm., Central Lebanon	Aptian-113	(Sohn et al. 2012)	242 (160-290)
Lepidoptera-Nepticulidae	*Stigmella* (leaf mine trace)	Dakota Fm., Nebraska, USA	Cenomanian-93.9	(Labandeira et al. 1994; Sohn et al. 2012)	157 (94-201)
Lepidoptera-Gracillariidae	Leaf mine trace attributed to *Phyllocnistis*	Dakota Fm., Nebraska, USA	Cenomanian-93.9	(Labandeira et al. 1994; Sohn et al. 2012)	102 (94-117)
Lepidoptera-Hesperiidae	Undescribed fossil Henrik Madsen Collection, Morsland Historical Museum Mors, Denmark (1 ex: DK 136)	Stolleklint Clay, Fur Fm., Denmark	early Ypresian- 47.8	(Sohn et al. 2012)	59 (48-83)
Siphonaptera+ Boreidae	Unnamed “Giant fleas” Nanjing Institute of Geology and Palaeontology reference numbers 154244a-b, 54245, 154247a, 154249a and 154250a	Jiulongshan Fm., Inner Mongolia, China	152	(Huang et al. 2012; Gao et al. 2012)	206 (158-255)
Mecoptera-Bittacidae	*e.g. Preanabittacus validus*	Jiulongshan Fm., Inner Mongolia, China	152	(Yang et al. 2012)	160 (152-181)
Diptera-Psychodidae	*Triassopsychoda olseni*	Cow branch Fm. Virginia, USA	Carnian 228	(Blagoderov et al. 2007)	247 (228-274)
Diptera- Culicomorpha	*Anisinodus crinitus*	Grès-a-Voltzia Fm., France	Anisian- 241	(Lukashevich et al. 2010)	264 (243-294)
Diptera-Chironomiodae	*Aenne triassica*	Lilstock Fm., UK	Rhaetian- 210	(Blagoderov et al. 2007)	234 (210-260)
Diptera-Perissommatidae	*Palaeoperissomma collessi*	Itat Fm., Kubekovo, Krasnoyarsk Krai, Siberian	Bathonian-166	(Lukashevich et al. 2006)	176 (166-197)

		Federal District, Russia			
Diptera- Brachycera	*Gallia alsatica*	Grès-a-Voltzia Fm., France	Anisian- 241	(Krzemiñski & Krzemiñski 2003)	252 (241-269)
Diptera -Rhagionidae	*e.g. Lithorhagio megalocephalus*	Jiulongshan Fm., Inner Mongolia, China	152	(Zhang & Li 2012)	176 (152-214)
Diptera-Tabanidae	*Eotabanoid lordi*	Durlston Fm., UK	Berriasian - 139	(Zhang 2012)	146 (139-175)
Diptera- Agromyzidae	*Phytomyzites biliapchaensis* (leaf mine trace)	Fort Union Fm., south-eastern Montana, USA	early Paleocene- 61.6	(Winkler et al. 2010)	77 (62-111)

3. Phylogenetic distribution of extant richness suggests metamorphosis is a key innovation in insects

3.1. Abstract

Insects and their six-legged relatives (Hexapoda) comprise more than half of all described species and dominate terrestrial and freshwater ecosystems. Understanding the macroevolutionary processes generating this richness requires a historical perspective, but the fossil record of hexapods is patchy and incomplete. Dated molecular phylogenies provide an alternative perspective on divergence times and have been combined with birth-death models to infer patterns of diversification across a range of taxonomic groups. Here we use a dated phylogeny of hexapod families to identify the broad pattern of macroevolutionary changes responsible for the composition of the extant hexapod fauna. The most prominent increase in diversification identified is associated with the origin of complete metamorphosis, confirming this as a key innovation in promoting insect diversity. Subsequent reductions are recovered for several groups previously identified as having a higher fossil diversity during the Mesozoic. In addition a number of recently derived taxa are found to have radiated following the development of flowering plant (angiosperm) floras during the mid-Cretaceous. These results reveal that the composition of the modern hexapod fauna is a product of a key developmental innovation, combined with multiple and varied evolutionary responses to environmental changes from the mid Cretaceous floral transition onward.

3.2. Introduction

Hexapoda, including the insects and their six-legged relatives, are the most species-rich animal clade in terrestrial ecosystems and collectively comprise over half of all described extant species (Gaston 1991; Grimaldi & Engel 2005). Therefore understanding the origins of this exceptional richness is key to understanding the history of life on land and the assembly of terrestrial ecosystems (Mayhew 2007). In addition to their high overall species richness, insect groups are also remarkable for the degree of disparity in richness existing among the major sub clades. For example the orders Zoraptera (“angel insects”) and Coleoptera (beetles) differ in richness by four orders of magnitude (32 and 350,000

described extant species, respectively (Grimaldi & Engel 2005)). A key part of the discussion on these differences in extant richness relates to the hypothesized effects of potential key innovations that may have acted as drivers for hexapod richness (Mayhew 2007). Such proposed innovations include both major morphological developments including: the origin of the insect body plan, flight, the capacity to fold the wings and the origin of complete metamorphosis (Carpenter 1953; de Queiroz 1998; Yang 2001; Dudley 2002; Mayhew 2002; Grimaldi & Engel 2005; Mayhew 2007; Davis et al. 2010a), and ecological opportunities or innovations, notably the evolution of flowering plants (angiosperms) (Mitter et al. 1988; Farrell 1998; Nyman 2010) and parasitism (Wiegmann et al. 1993).

Attempts to explicitly test these ideas within a phylogenetic framework have either been restricted to particular orders (Hunt et al. 2007; Wiegmann et al. 2011; Heikkilä et al. 2011), thus omitting a wider context, or have ignored variation within orders (Mayhew 2002; Davis et al. 2010a). Here we integrate these disparate approaches by producing a dated hypothesis of phylogenetic relationships across the hexapods that is near-complete at the family level, through the combination of previously published molecular sequence data and a set of literature derived constraints (see Chapter 2). Our goal is therefore not to present a novel estimate of the hexapod phylogeny (see discussion below), but instead to focus on what current taxonomic, phylogenetic and paleontological evidence reveals about broad patterns of diversification within the group, and its relationship with key evolutionary innovations, environmental changes and mass extinctions (Labandeira & Sepkoski Jr 1993; Ross et al. 2000; Labandeira 2005).

3.3. Methods

All analyses of diversification and tree processing were conducted in R v 2.15.1 (R Development Core Team 2011). Estimates of extant species richness for terminal taxa were sourced from previous publications (Durden & Musser 1994; Hunt et al. 2007; Whiting et al. 2008; Kathirithamby 2009; Cryan & Svenson 2010; Vas et al. 2012), recent encyclopedias (Lienhard & Smithers 2002; Grimaldi & Engel 2005; Resh & Cardé 2009; Zhang 2011) and online taxonomic resources (Penny 1997; Noyes 2003; Deitz et al. 2010; Ascher & Pickering 2012; Bourgoin 2012; Deem 2012; DeWalt et al. 2012; Eades 2012; Maehr & Eades 2012; Otte et al. 2012; Pulawski 2012) (tip richness linked in Appendix

7.1). Where taxonomic sources conflicted with the classification given here species assigned to any subgroups were deducted from the more inclusive clade.

The primary algorithm used to infer the topological position of shifts in diversification rate was the stepwise greedy ML algorithm MEDUSA (Alfaro et al. 2009) as implemented in the package TurboMEDUSA (Brown et al. 2012). This algorithm proceeds by estimating optimal parameters for a global birth-death model on a given tree with tip richnesses, and then going through all nodes and identifying the optimal position of a break in the diversification model (henceforth a "shift") that maximally improved the overall AICc score. This process is repeated adding further shifts until some threshold is achieved where there is no further improvement in AICc. The appropriate threshold score was calculated internally to the routine and given as an improvement in AICc of 9.321 units. The resulting optimal model identified 48 shift events with parameter values listed in Table 4 (Figure 18). In order to estimate the impact of particular shift events on the overall richness of hexapods we used simulated birth death models to estimate what the richness of clades would have been had the modeled shifts not occurred. This was done using the crown.limits (Magallon & Sanderson 2001) function in package geiger (Harmon et al. 2008) with parameter values taken from the parental model estimated in MEDUSA (i.e. the model within which the focal shift is nested) and node age based on the consensus tree. The function gives the upper and lower confidence intervals on the richness of a clade of that age, from which a mean estimate was calculated. This was compared with estimates of clade richness which had previously been corrected for the presence of further shifts in diversification rate, by replacing the richness of nodes with subject to further shifts with their mean modeled richness (Figure 19).

The reported shift on Holometabola (Section 3.4) is of particular interest as it is both the first shift recovered by the greedy MEDUSA algorithm, implying that its inclusion makes the greatest overall improvement to AICc scores and it is also significant in ideas relating to key innovations in driving hexapod diversification (Mayhew 2007). In order to explore the importance of this shift and to make comparisons with other potential key innovations we examined the likelihood improvement associated with a range of shifts using the package laser (Rabosky 2007). The functions fitNDR_1rate and fitNDR_2rate were used to find parameter estimates and log likelihood values for the optimal global model and models with a single shift at every possible node respectively. In both cases these functions require that the turnover parameter be specified and this was based on the

value obtained from for the homogenous model (i.e. that with no shifts) estimated in MEDUSA (eps=0.9990712). Parameter estimates and likelihood scores of the resulting models are presented in Table 3 with an emphasis placed on groups associated with potential shifts in diversification rate identified in previous studies (Mayhew 2002; Mayhew 2003; Davis et al. 2010a). The results of likelihood ratio tests are shown for comparison of the Holometabola shift model with these alternatives and the single shift model.

In order to assess the degree to which uncertainty with respect to node age impacted on the pattern of rate shift events, the analysis was repeated across 500 random samples taken from the post-burnin MCMC chain used in dating. Samples were taken so as to be evenly distributed between the two parallel chains and were scaled into time units using the appropriate estimates of the overall clock rate. Note that due to the use of the two-stage phylogenetic inference process described above it was not possible to assess the impact of topological uncertainty on the results of this study. The occurrence of a shift associated with a particular node across the different samples was scored and the proportion of samples in which a node occurs is used as the basis for the colouration of the symbols on Figure 18 and is reported in Table 4. Table 5 lists the top 50 nodes with respect to proportional occurrence and these are depicted on Figure 20. For this study we have elected not to use the alternative diversification model TreePar (Stadler 2011; Jetz et al. 2012; Near et al. 2013) due to limitations on the available computational resources which restricted the capacity to simulate the large numbers of species complete trees needed to calculate appropriate confidence intervals (Stadler 2011a)

3.4. Results

The dated phylogeny used in this study contains 874 higher taxa of Hexapoda (Figure 18). Taxa were variously resolved to a family or superfamily level, such that the presented tree incorporates a total of 903 of the approximately 1100 recognized extant families, with taxonomy following that given by GenBank references up to August 2013 (see Chapter 2 for further discussion). The tree was reconstructed using a combination of eight widely sampled molecular markers and literature-derived constraints on certain widely recognized phylogenetic nodes ((Trautwein et al. 2012; Yeates et al. 2012), see Section 2.3.1). The tree topology was inferred using a partitioned RAxML (maximum likelihood) analysis (Stamatakis et al. 2005; Stamatakis et al. 2008). This topology was

dated using a relaxed molecular clock implemented in Mr Bayes 3.2 (Ronquist et al. 2012) and calibrated using 86 fossil dates taken from the recent palaeoentomological literature (Table 2).

Using our dated tree we estimated the crown divergence of Hexapoda, i.e. the divergence of true insects from Entognatha (basal hexapods including springtails) as occurring in the Ordovician (mean estimate 474.4Ma, 95% CI 439.6-502.9Ma), which is consistent with other recent molecular clock estimates (Rehm et al. 2011; Rota-Stabelli et al. 2013; Wheat & Wahlberg 2013) (Figure 16, Table 1). These estimates greatly exceed the age of the oldest securely placed hexapod fossils including the potential crown winged insect *Rhyniognatha hirsti* from the early Devonian (Engel & Grimaldi 2004). Little is known regarding Devonian insect communities (Grimaldi & Engel 2005), and the nature of terrestrial communities at this early date remains poorly understood (Kenrick et al. 2012). However, our results are in line with recent fossil evidence indicating an early (i.e. prior to the late Carboniferous) origin for major crown lineages, including the stem lineages of several orders of advanced Holometabola (insects that undergo complete metamorphosis) (Labandeira 2011; Nel et al. 2013).

At higher taxonomic levels, lineage-through-time plots (Figure 17) indicate a remarkable stability in divergence rate across all the major hexapod clades, with some suggestion of an elevated diversification rate in Holometabola during the late Permian corresponding to basal divergences within Coleoptera and Diptera (flies) (Grimaldi & Engel 2005; Blagoderov et al. 2007). Despite the conventional division between Paleozoic and post-Paleozoic insect faunas in paleoentomological research (Labandeira 1998; Grimaldi & Engel 2005), our results reveal no evidence for changes in diversification rate around the time of the Permo-Triassic extinction event (P/T) (Figure 17), suggesting that the radiation of extant groups was not strongly impacted by the loss of Paleozoic forms indicated by the fossil record (Labandeira & Sepkoski Jr 1993; Labandeira 2005). A possible exception is an upshift in the diversity of Palaeoptera (dragonflies and mayflies) associated with the origin of crown members of the two orders, both of which undergo major taxonomic turnover during the P/T event (Grimaldi & Engel 2005; Davis et al. 2011) (Figure 17).

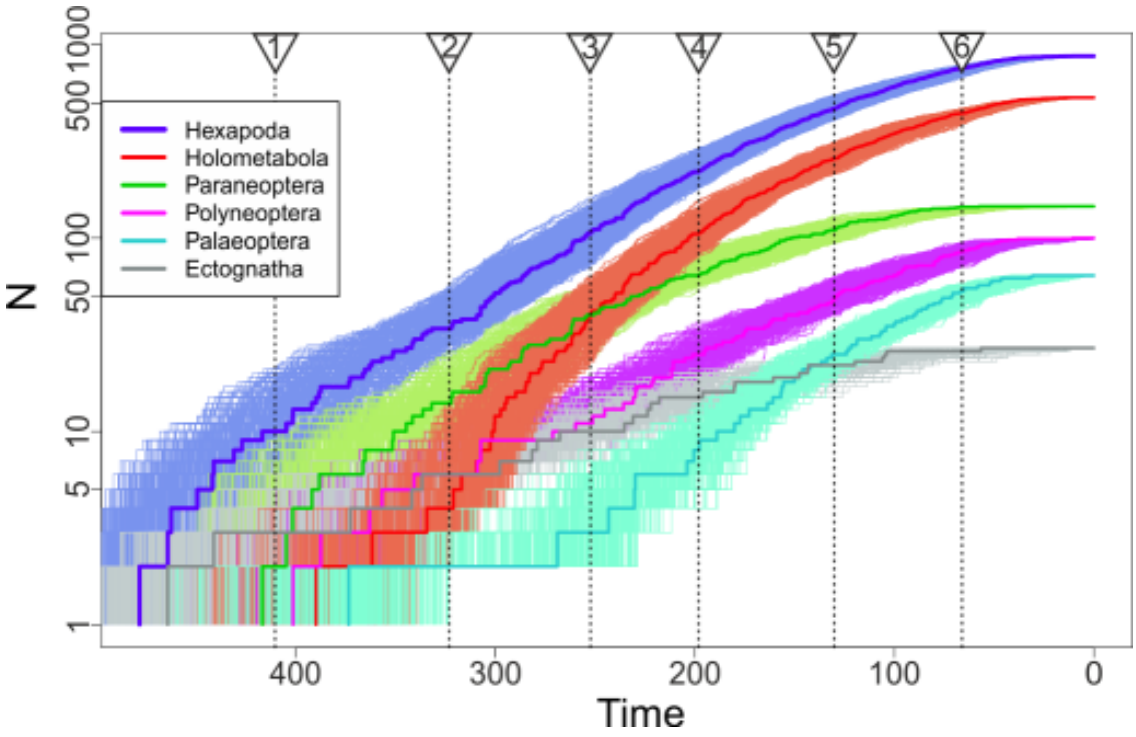


Figure 17 Lineage (y-axis; log scale) through time (x-axis; Ma) plot for the major groups of Hexapoda using the phylogeny in Figure 18 Colors used identify the same clades as the ring in Figure 18. Thick lines are calculated from the mean tree dates (Figure 16). Shaded regions represent 500-scaled samples taken from the MCMC chain used in dating. Major events in the history of the group are denoted using dotted lines: 1. Oldest Hexapod fossil. 2. Oldest member of crown Pterygota (Polyneoptera). 3. Permo-Triassic mass extinction. 4. Origination of crown Angiosperms (Clarke et al 2011). 5. Angiosperms become abundant in fossil record. 6. Cretaceous-Paleocene mass extinction

Despite this apparent stability in the origination of higher taxa, the application of birth-death models (Nee 2006; Alfaro et al. 2009) identifies two major transitions, characterized as shifts in the net diversification rate and turnover in the descendent clades, which together play a major role in defining the overall structure of hexapod diversification. These major shifts correspond to the origins of flight (Pterygota) and of complete metamorphosis (Holometabola) (Figure 18, Table 3 and Table 4). Both in terms of the degree to which its inclusion improves the likelihood of diversification models (Table 3) and in its relative stability with respect to uncertainties in node age estimation (Figure 18, Table 4) the upshift in diversification rate associated with the origin of complete metamorphosis represents the more strongly supported event.

Previous studies proposing a link between complete metamorphosis and elevated diversification rates have been based on evidence in the fossil record (Yang 2001; Nicholson et al. 2014), which for hexapods is highly incomplete (Wills 2001). In contrast, sister group comparisons, using earlier phylogenetic reconstructions (Trautwein et al. 2012; Yeates et al. 2012), failed to recover a diversification shift associated with Holometabola (Mayhew 2002; Davis et al. 2010a). However, likelihood ratio tests indicate that the birth-death models significantly favor this position over alternative proposals including Eumetabola (Holometabola plus its sister group) and Neoptera (insects able to fold their wings; Table 3). Earlier studies (Mayhew 2002; Mayhew 2003) have also provided some evidence supporting the role of flight in promoting hexapod diversification. Although our analysis supports this notion it also shows that the recovery of this shift is sensitive to uncertainties in divergence time estimates within the phylogeny rendering its overall role in hexapod diversification ambiguous (Figure 18).

In addition to these broad patterns, diversification shift models identified a further forty-three clades on the tree potentially associated with shifts in diversification rates (Figure 18, Table 4). These shifts vary in their intensity and robustness with respect to uncertainties in branch length and are distributed across the tree, with the majority occurring within the holometabolan radiation. Among the most robust and phylogenetically inclusive shifts are down-shifts impacting on known or suspected relict groups within the modern fauna. These included holometabolan groups such as Neuropterida (lacewings and their relatives), Mecoptera and Siphonaptera (scorpionflies and fleas (Whiting 2002)) and basal members of Coleoptera (beetles) and Lepidoptera (moths), as well as non-metamorphosing groups such as Ephemeroptera (mayflies) and Psocodea (booklice and parasitic lice) (Yoshizawa & Johnson 2010). The fossil records for a number of these groups indicate a higher family richness during the Mesozoic, suggesting that their current representatives are surviving relics of taxa that were formerly more diverse (Labandeira & Sepkoski Jr 1993; Labandeira 2005), further supporting the results of the diversification shift models.

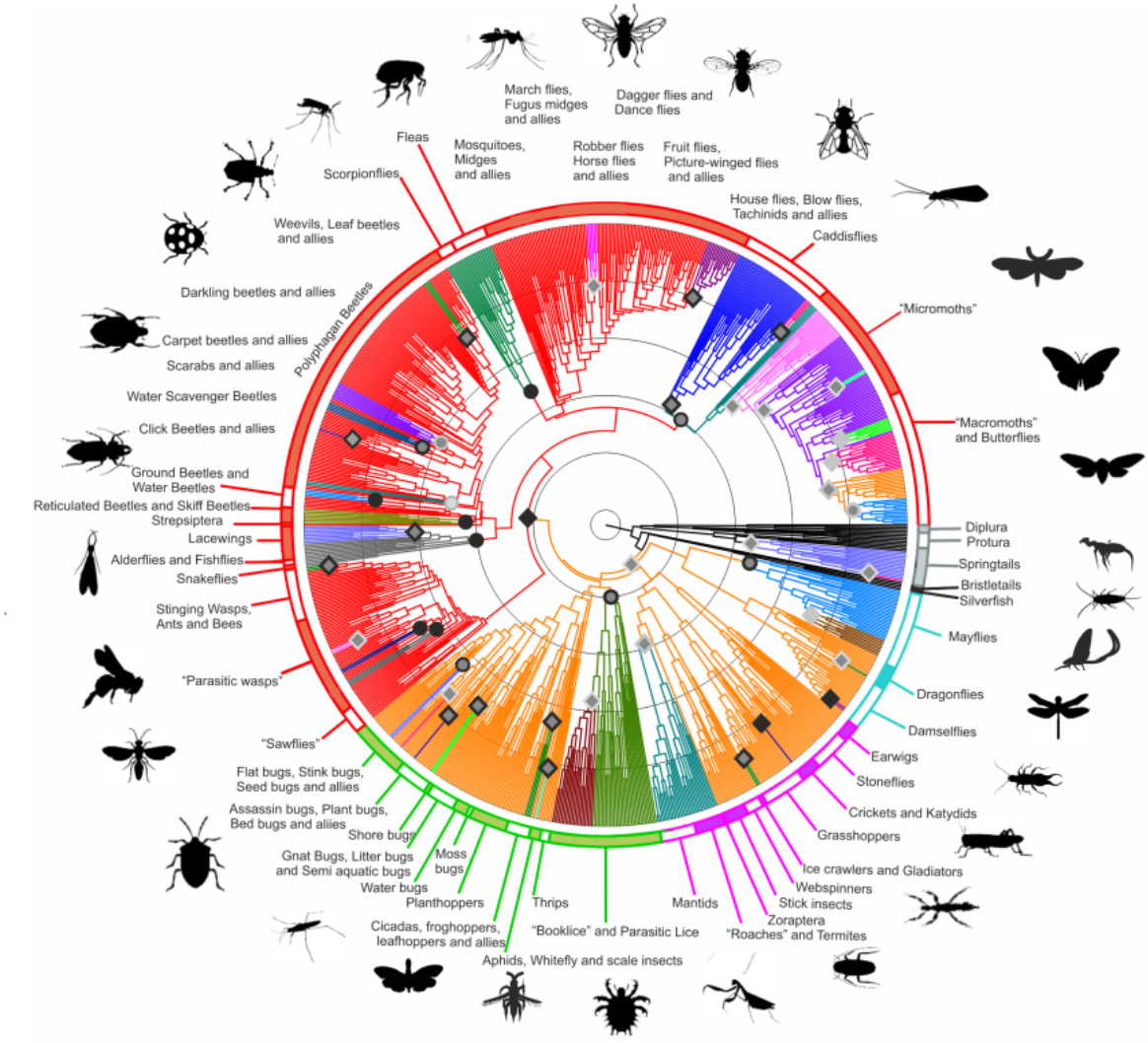


Figure 18 Dated phylogeny of extant hexapod families showing diversification rate shifts. Membership of major clades is denoted by coloration of the ring (Grey: Entognatha, Black: basal insects, Cyan: Palaeoptera, Purple: Polyneoptera, Green: Paraneoptera, Red: Holometabola). Changes in branch coloration denote diversification shifts identified using TurboMEDUSA (Table 4). Symbols at shifts denote a net upshift (diamond) or down shift (circle). Coloration of symbols reflects the robustness of the shift event across dating (Black: shift recovered in >80% of samples, Grey with Black outline: recovery >50%, Grey with Pale outline: recovery >30%, Pale Grey: recovery<30%). Black circles are shown at 100Ma increments from the present.

In contrast with these relict groups, most of the shifts leading to a net increase in taxonomic richness are comparatively recent (Figure 19) and are associated with restricted, but massively diverse lineages many of which are of large ecological significance in recent communities. Among the non-holometabolan insects these include large herbivorous radiations such as the katydids (Tettigoniidae), true and lubber grasshoppers (Acrididae and Romaleidae), aphids (Aphidoidea), leafhoppers and treehoppers (Membracoidea), as well as plant/lace bugs and stink bugs (Miridae/Tingidae and Pentatomidae). Also represented are predatory groups such as assassin bugs (Reduviidae) and certain families of dragonflies and damselflies (Odonata). The pattern of shifts within the Dictyoptera (which includes detritivorous roaches and termites as well as predatory mantids) (Davis et al. 2009) is unstable with respect to branch length, with the majority of samples failing to recover the small proposed shift encompassing the entirety of the group (Figure 20, Table 5). These groups, with the exception of Dictyoptera, radiated during the mid to late Cretaceous, which may imply an association between these radiations and the restructuring of floral and faunal communities during this interval following the radiation of angiosperms (Clarke et al. 2011; Fiz-Palacios et al. 2011).

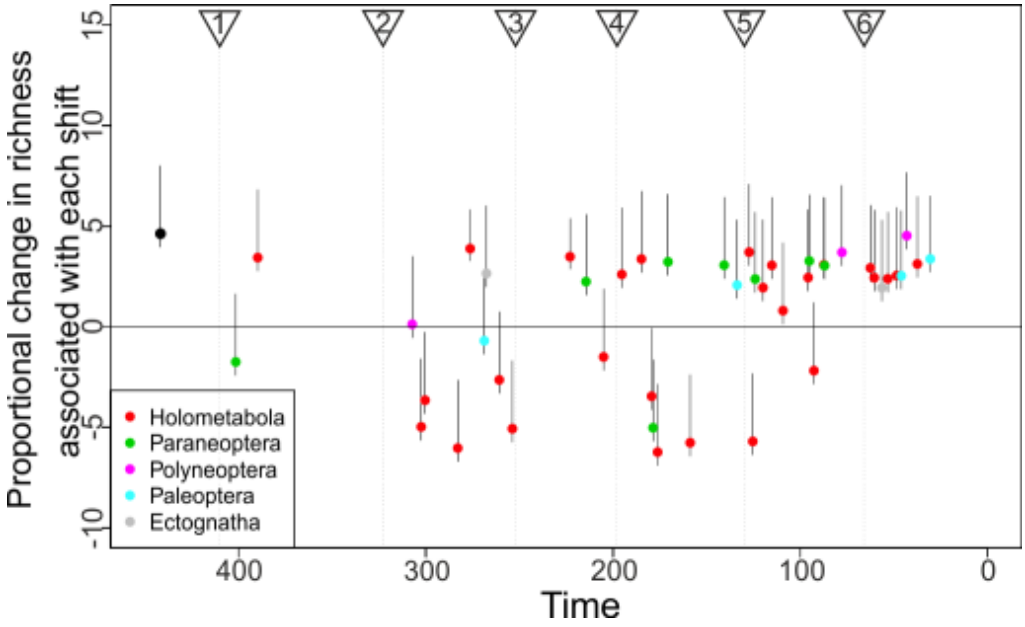


Figure 19 Change in species richness associated with shift events plotted through time. Values plotted are ratio between the observed richness of the clade (after correction for nested shifts) and the mean estimated values of the richness of a clade of the appropriate age under the parental diversification model (see text). Confidence intervals given are based on the change in richness associated with 95% CIs on the estimated outcomes of the stochastic diversification process

Unsurprisingly, several upshifts in diversification within Holometabola also involve groups directly associated with the angiosperm radiation, with notable examples including leaf and longhorn beetles (Chrysomeloidea) (Gómez-Zurita et al. 2007; Hunt et al. 2007) and advanced bees (Apidae and Megachilidae) (Davis et al. 2010b). Our results also strongly support an upshift encompassing the Calyptratae, which includes houseflies and the important parasitoid group Tachinidae (Wiegmann et al. 2011).

The recovered pattern of diversification shifts in Lepidoptera is complex and highly sensitive to uncertainties in branch length estimation, reflecting the difficulties of accurately dating a group for which there is a lack of suitable calibration fossils (Labandeira & Sepkoski Jr 1993; Sohn et al. 2012), and which includes several regions of phylogenetic instability (Cho et al. 2011; Regier et al. 2013). The pattern recovered from the mean estimates of node times indicates a nested model with an overall down-shift associated with the most basal moths followed by a series of up-shifts corresponding to the major clades Glossata (moths with a proboscis) and Ditrysia (moths with partitioned female reproductive tracts). The pattern of shifts within the advanced moths and butterflies is poorly resolved with a number of events showing limited robustness with respect to branch length variation. If shift recovery across multiple samples of the Markov Chain Monte Carlo used in dating is considered ((Ronquist et al. 2012) see Section 2.4.), several of these events are found to be collapsed into a single shift associated with the redefined Obtectomera (Cho et al. 2011; Regier et al. 2013) (Figure 20, Table 5) which also corresponds to the shift associated with second largest improvement in overall model likelihood in single-shift models (Table 3).

Comparable previous work on patterns of diversification within Diptera identified a series of nested shifts within the order that are not recovered in our study (Wiegmann et al. 2011). These differences can be attributed to the placement of radiations within a more inclusive phylogenetic context, i.e. within Holometabola in its entirety, resulting in greater estimated turnover within the group, as well as minor differences in taxonomic sampling and dating between analyses. Contrary to previous views, which have tended to emphasize the role of particular ecologies (notably phytophagy) (Mitter et al. 1988; Farrell 1998) in determining patterns of hexapod richness, our results do not show strong correlation between patterns of diversity and particular life history traits, e.g. upshifted clades show a range of dietary ecologies (Chapter 4). Instead, our results suggest diverse responses

within Mesozoic fauna to the ecological transition and novel opportunities provided by the Cretaceous angiosperm expansion (Grimaldi & Engel 2005).

3.4.1.1. Perspectives on diversification from sampling the MCMC

The majority of studies involved in modeling diversification rates within dated trees have tended to focus exclusively on patterns associated with the mean tree without any consideration of the potential uncertainties involved in estimating node ages. Bayesian methods as implemented here provide a natural way to approximate the confidence intervals on the node ages through the use of samples from the postdated MCMC (see methods). Comparing the pattern of nodes consistently recovered from such samples (the top fifty of which are shown on Figure 20 and listed in Table 5 provides an alternative insight into the processes of diversification active within the group. Overall the pattern of well supported shifts is broadly consistent with that recovered on the mean tree indicating that most of the inferred events, including the shifts associated with Holometabola and Pterygota, are relatively robust with respect to uncertainties in branch length. However there are also a number of differences within certain major clades that change our perspective on diversification patterns.

The majority of changes in the pattern of diversification occur within the megadiverse Holometabolan orders. In Lepidoptera the pattern of shifts alters such that in place of the idiosyncratic shifts associated with butterflies (Rhopalocera) and Gelechioidea a more secure shift is recovered associated with the redefined Obtectomera, which encompasses both these groups as well as macromoths, and is one of the best supported clades in the advanced Lepidoptera (Mutanen et al. 2010; Cho et al. 2011; Regier et al. 2013). This shift also corresponds to that identified as the second best position under the two-rate model in Table 3. Within Coleoptera four novel shifts are highlighted involving three large and recently derived phytophagous groups (Farrell 1998; Hunt et al. 2007): Buprestidae (jewel beetles), Curculionidae (“true” weevils) and Mordellidae (tumbling flower beetles) (the last in association with the also tending to phytophagous Anthicidae (ant-like flower beetles) and the parasitoid Meloidae (Blister beetles)) (Hunt et al. 2007)(Arnett et al. 2010), as well as the large detritivorous family Tenebrionidae (Darkling beetles). Within Diptera the seminal role of Calyptratae, and Tachinidae in particular, in dominating the pattern of diversification is again emphasized. Tachinids are among the most diverse (9626 described species) (Zhang 2011) and youngest (divergence estimated

as 26.27 Ma, CIs 9.96-42.31 Ma) fly families, implying exceptional rates of diversification, which may reflect the group's successful adaptation to a parasitoid lifestyle on a huge variety of arthropod hosts, particularly similarly recently derived Lepidoptera (Marshall 2012).

Outside of Holometabola there are also minor modifications to the apparent pattern of diversification including a potential shift associated with the aquatic bugs (including Nepomorpha, Gerromorpha, Dipsocoromorpha and Enicocephalomorpha), modifications to the apparent pattern within Dictyoptera that highlight three super-rich clades (i.e. Blattidae + Blaberidae (Blattodea), Mantidae (Mantodea), Termitidae (Isoptera)) and bring results more in line with those of previous studies (Davis et al. 2009), and the loss of idiosyncratic up-shifts associated with a subclade of Ephemeroptera and Neanuridae (Collembola). As well as highlighting further candidates for radiations within Hexapoda these results also emphasize the potential dangers of relying on a single set of date estimates when discussing diversification, as the resulting pattern may include shifts shaped by idiosyncrasies of the particular tree chosen and so not be representative of the overall pattern implied by the data.

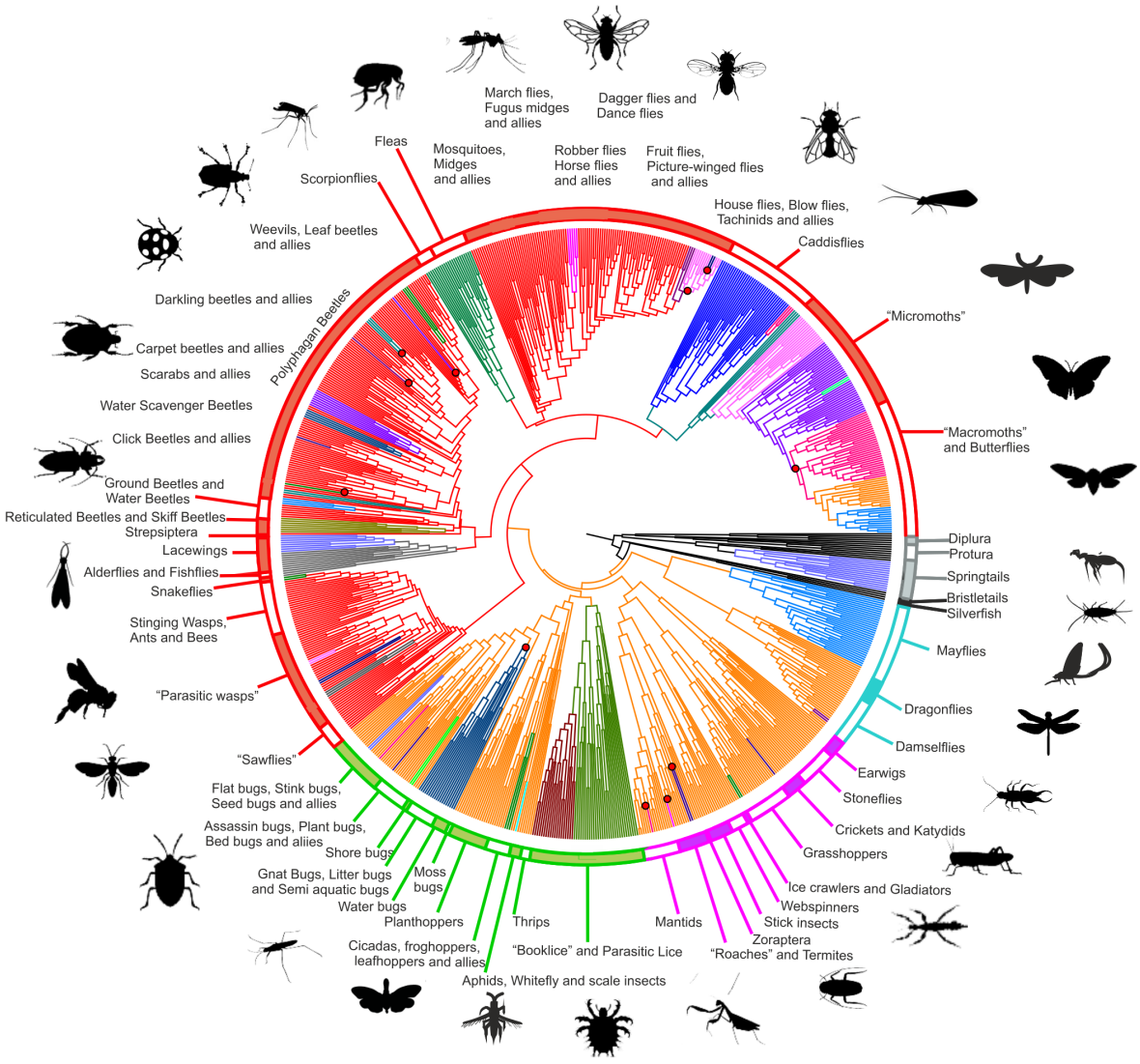


Figure 20: The fifty shifts with the highest rates of recovery in samples from the MCMC chain (Table 5) plotted together on the tree topology. Shifts are denoted as Figure 18 with novel shifts not recovered on the mean tree denoted by red circles

3.5. Discussion

Ultimate explanations of insect diversification can be classified into morphological key innovations, and ecological interactions (Mayhew 2007). Our results highlight the importance of complete metamorphosis as the major key innovation underpinning the pattern of hexapod species richness. The mechanism by which complete metamorphosis promotes diversification is incompletely understood. However, previous workers have suggested that the ecological division of adult and juvenile life stages separated by a pupal stage in Holometabola may play a major role (Carpenter 1953; Yang 2001; Mayhew

2007). The adaptation to novel ecological niches likely played a role in promoting diversity within specific hexapod radiations, such as family-level or lower taxonomic levels, but there is no evidence here to support the idea that a single suite of ecological traits is generally associated with shifts in hexapod diversification. Instead, the patterns observed are consistent with distinct members of the community responding in a wide variety of ways to the ecological changes following the angiosperm radiation and continuing to the present day. However, we did find evidence that the radiation of angiosperms itself triggered a number of upshifts in diversification rate across both non-holometabolan and holometabolan groups, marking the evolution of angiosperms as a key ecological change in the evolutionary history of Hexapoda.

It is important to note that our recognition of these patterns is dependent on the inferred phylogenetic topology, which contains some regions of considerable phylogenetic uncertainty (see Section 2.3.2). However, it is unlikely that the major findings of our analysis – i.e. key roles of complete metamorphosis and angiosperm evolution as well as the failure to recover a distinct suite of ecological traits underlying a species group's phylogenetic richness – will change in the light of future improvements to the topology, dating, and extant species richness of the insect phylogenetic tree, which collectively will combine to further improve our understanding of the origins and diversification of this key component of terrestrial ecosystems.

Table 3 Log likelihood and parameter estimates for models with a single shift in diversification rate. Log likelihood tests performed to compare the optimal model, with shift placed on Holometabola, and various alternative models. Data shown are log likelihoods of the respective models estimated in laser using turnover estimates as estimated in the homogenous model in MEDUSA (i.e. that with no shifts). Net diversification rates estimated for the partition in including the root (Rroot) and the descendants of the focal node (Rclade). Chi squared values, Degrees of Freedom and p values relate to results of likelihood ratio comparisons between denoted models. Comparable parameter values for the Holometabola model are; lnL= -11299.6, R_{root}=0.00359636, R_{clade}=0.0112969.

Taxon 1	*lnL*	Taxon 2	*lnL*	R root	R clade	Chi-Squared	DF	p.value
Holometabola	-11299.6	**Uniform Model**	-11504.6	0.00874	-	409.90	3	< 0.001
Holometabola	-11299.6	**Obtectomera (Second Best Node)**	-11351.0	0.0073964	0.0264802	102.81	1	< 0.001
Holometabola	-11299.6	**Paraneoptera**	-11433.7	0.0096620	0.0039822	268.22	1	< 0.001
Holometabola	-11299.6	**Eumetabola**	-11393.2	0.0032761	0.0098645	187.10	1	< 0.001
Holometabola	-11299.6	**Neoptera**	-11415.3	0.0020526	0.0093177	231.25	1	< 0.001
Holometabola	-11299.6	**Pterygota**	-11474.3	0.0023812	0.0089912	349.32	1	< 0.001
Holometabola	-11299.6	**Insecta**	-11481.9	0.0026486	0.0089487	364.58	1	< 0.001
Holometabola	-11299.6	**Coleoptera**	-11495.2	0.0082814	0.0105405	391.11	1	< 0.001
Holometabola	-11299.6	**Lepidoptera**	-11418.2	0.0075482	0.016791	237.22	1	< 0.001
Holometabola	-11299.6	**Diptera**	-11486.3	0.0082253	0.0120813	373.32	1	< 0.001
Holometabola	-11299.6	**Hymenoptera**	-11497.6	0.0085002	0.0111554	395.96	1	< 0.001

Table 4 Parameter values and shifts in species richness associated with MEDUSA model shifts inferred across the mean topology.

Shift No.	Affected Taxa	Percentage occurrence in samples from the MCMC chain	Log likelihood of model including shift	AIC of cumulative models	Net diversification of shifted clade	Turnover in shifted clade	Age (Myr)	Parent shift	Total richness (Corrected Richness)	Mean richness in the absence of shift (lower- upper 95% CIs)
1	Root	NA	-11504.58	23013.15	0.006306892	0.984480673	478.08	NA	1037967	NA
2	Holometabola	98	-11261.60	22533.20	0.016956804	0.996863114	389.69	42	867118 (912767)	29329 (988- 57669)
3	Macromoths (Hyblaeoidea+ Pyraloidea+ Noctuoidea+ Geometroidea+ Drepanoidea+ Bombycoidea+ Cimeliidae) (Lepidoptera)	44.2	-11160.56	22337.12	0.017787886	0.999846505	127.35	40	89731 (126799)	3107 (105- 6110)
4	Calyptratae + Drosophilidae + Ephydridae + Agromyzidae (Diptera)	61.4	-11108.74	22239.49	0.031770706	0.99907383	95.88	2	22890 (22890)	1988 (67- 3908)
5	Acrididae + Romaleidae (Orthoptera)	70.8	-11062.85	22153.71	0.067112354	0.99470352	43.18	42	6481 (6481)	70 (3-138)
6	Trichoptera	70	-11022.27	22078.54	0.022328614	0.926777663	276.29	30	14193 (12994)	267 (38-495)
7	Archostemata + Myxophaga (Coleoptera)	100	-10987.31	22014.62	0.015592491	6.03E-07	282.92	2	140 (140)	58475 (1969-114981)
8	Psocodea	59.4	-10957.34	21960.68	0.01259899	0.940790255	401.42	42	9234 (6109)	35161 (1184- 69137)
9	Mecoptera + Siphonaptera	84.8	-10928.56	21909.13	0.009646414	0.99145901	260.77	2	2853 (2853)	40013 (1348- 78679)
10	Ephemeroptera	56.6	-10900.43	21858.87	0.021219991	0.785579484	268.74	42	3126 (2230)	4480 (151-8808)
11	Neuropterida	96.2	-10872.32	21808.65	0.017081799	0.865444919	300.35	2	6180 (2039)	78749 (2652- 154847)
12	Chrysomeloidea	75.8	-10847.07	21764.13	0.027013199	0.999342832	115.25	2	62619	2949

	(Coleoptera)								(62619)	(100- 5799)
13	Bombycoidea (in part) + Cimeliidae (Lepidoptera)	45.6	-10824.57	21725.14	0.047987062	0.981655444	92.77	3	4704 (4704)	41772 (1407- 82138)
14	Apataniidae + Goeridae + Limnephilidae (Trichoptera)	57	-10804.30	21690.60	0.015012577	0.99791316	62.64	6	1267 (1267)	68 (3-133)
15	Tettigoniidae (Orthoptera)	95	-10784.19	21656.38	0.103127941	0.552218484	77.82	42	6827 (6827)	169 (6-332)
16	Apidae + Megachilidae (Hymenoptera)	59.2	-10764.50	21623.00	0.007156149	0.99993768	37.56	2	9871 (9871)	435 (15-855)
17	Rhopalocera (with the exception of Papilionidae) (Lepidoptera)	11.8	-10744.78	21589.56	0.029625726	0.998872604	87.46	40	18177 (18177)	839 (29-1650)
18	Membracoidea (Hemiptera)	52.2	-10728.34	21562.68	0.020257069	0.996798633	170.67	42	23492 (23492)	932 (32-1832)
19	Reduviidae (Hemiptera)	78.6	-10712.48	21536.97	0.084227724	0.525345485	95.246	42	6420 (6420)	243 (9-477)
20	Heloridae+ Maamingidae + Mymarommatidae (Hymenoptera)	100	-10696.79	21511.58	0.013650868	6.86E-06	176.23	2	18 (18)	9173 (309- 18036)
21	Amphizoidae + Aspidytidae + Hygrobiidae (Coleoptera)	99.2	-10681.83	21487.67	0.014945337	1.70E-06	125.41	2	12 (12)	3595 (122- 7069)
22	Miridae + Tingidae (Hemiptera)	65.2	-10667.91	21465.83	0.006895962	0.999726953	140.70	42	12000 (12000)	561 (19-1103)
23	Coenagrionidae + Protoneuridae (Odonata)	90.6	-10653.16	21442.32	0.024525083	0.998329638	30.66	42	1344 (1344)	46 (2-89)
24	Aphidoidea (Hemiptera)	62.8	-10637.93	21417.87	0.087179981	0.537060022	87.14	42	4300 (4300)	206 (7-405)
25	Dictyoptera	33.0	-10620.86	21389.71	0.010720132	0.996830072	307.23	42	9253 (9253)	8169 (276- 16063)
26	Asiloidea (in part) (Diptera)	36.0	-10607.69	21369.38	0.015839809	0.999508807	109.28	2	5837 (5837)	2619 (89-5149)
27	Elateridae (Coleoptera)	66.4	-10594.42	21348.85	0.138604856	0.558284995	60.55	2	10000 (10000)	874 (30-1718)

28	Eulophidae + Pteromalidae (Hymenoptera)	49.2	-10580.62	21327.24	5.02E-07	0.999999994	48.58	2	7978 (7978)	624 (21-1227)
29	Gelechioidea + Callidulidae (Lepidoptera)	2.8	-10567.60	21307.21	0.028217669	0.997116756	119.95	40	17121 (17121)	2447 (83-4811)
30	Amphiesmenoptera (Lepidoptera + Trichoptera)	56.4	-10554.44	21286.88	0.016831916	2.76E-06	302.41	2	170209 (562)	81563 (2747- 160379)
31	Glossata (excluding Eriocraniidae) (Lepidoptera)	41,4	-10536.46	21256.91	0.011361995	0.990118042	222.98	30	155829 (3515)	108 (16-201)
32	Ameletidae + Ameletopsidae + Caenidae + Oligoneuriidae + Ephemerellidae + Heptageniidae + Nesameletidae + Rallidentidae (Ephemeroptera)	26.8	-10523.79	21237.57	0.025969093	0.952800982	133.74	10	1025 (1025)	129 (5-253)
33	Neanuridae (Collembola)	32.4	-10511.52	21219.03	0.113979086	0.556941465	56.53	43	1417 (1417)	204 (7-401)
34	Epimetopidae + Georissidae + Helophoridae + Spercheidae (Coleoptera)	72.2	-10499.31	21200.61	0.028544091	2.61E-06	179.43	2	306 (306)	9711 (327- 19095)
35	Libellulidae (Odonata)	31.4	-10487.14	21182.27	0.130729344	0.57268948	46.12	42	970 (970)	77 (3-150)
36	Pentatomidae (Hemiptera)	44.2	-10474.51	21163.02	0.061467249	0.542746523	124.12	42	4500 (4500)	419 (15-823)
37	Psocomorpha (Psocodea)	47.6	-10462.77	21145.55	0.019704411	0.964589998	214.44	8	3496 (3496)	371 (13-728)
38	Myrmeleontoidea + Nemopteridae + Ithonioidea + Chrysopidae + Hemerobiidae (Neuroptera)	69.6	-10451.64	21129.28	0.029858396	0.918567569	195.34	11	4472 (4472)	331 (12-649)
39	Ceratocanthidae + Hybosoridae + Glaresidae + Lucanidae + Pleocomidae + Glaphyridae + Ochodaeidae + Passalidae (Coleoptera)	35.2	-10440.73	21113.46	0.026293101	0.884805968	205.08	2	3402 (3402)	15266 (514- 30017)
40	Ditrysia (Lepidoptera)	38.8	-10430.50	21098.99	0.031681892	0.972384078	184.87	31	153427 (32332)	1113 (38-2188)
41	Pelecinidae + Roproniidae	92.2	-10420.50	21085.01	0.015770387	2.65E-08	158.84	2	21	6706

	(Hymenoptera)								(21)	(226-13185)
42	Pterygota	34.8	-10411.44	21072.88	0.015426853	0.97861985	441.60	1	1028365 (154218)	1506 (51-2961)
43	Collembola (with exceptions of Neelipleona, Katiannidae and Tomoceridae)	40	-10399.43	21054.86	0.011988111	0.992676842	267.57	1	7419 (6206)	438 (15-861)
44	Clambidae + Eucinetidae (Coleoptera)	35.4	-10390.75	21043.51	0.018680745	3.41E-05	253.75	2	223 (223)	35468 (1195-69742)
45	Hyocephalidae + Idiostolidae (Hemiptera)	67.8	-10382.08	21032.16	0.00776074	4.28E-06	178.65	42	7 (7)	1063 (36-2090)
46	Limacodidae + Zygaenidae (Lepidoptera)	36.6	-10373.60	21021.20	0.000317545	0.999987364	53.34	40	2708 (2708)	249 (9-490)

Table 5 The fifty most robustly recovered shifts inferred from 500 samples from the post-burnin Markov Chain Monte Carlo (MCMC). Shifts are plotted on Figure 20. Shifts without equivalents on the mean tree are highlighted in bold.

Percentage occurrence in samples from the MCMC chain	Equivalent Model on Consensus tree	Taxa Impacted
1	20	Heloridae *Maamingidae Mymarommatidae* (Hymenoptera)
1	7	Archostemata + Myxophaga (Coleoptera)
0.992	21	*Amphizoidae* + *Aspidytidae* + *Hygrobiidae* (Coleoptera)
0.98	2	Holometabola
0.962	11	Neuropterida
0.95	15	*Tettigoniidae* (Orthoptera)
0.922	41	*Pelecinidae* + *Roproniidae* (Hymenoptera)
0.906	23	*Coenagrionidae* + *Protoneuridae* (Odonata)
0.848	9	Mecoptera + Siphonaptera
0.824	**-**	**Obtectomera (Lepidoptera)**
0.786	19	*Reduviidae* (Hemiptera)
0.758	12	Chrysomeloidea (Coleoptera)
0.722	34	*Epimetopidae* + *Georissidae* + *Helophoridae* + *Spercheidae* (Coleoptera)
0.708	5	Acrididae + *Romaleidae* (Orthoptera)
0.7	6	Trichoptera
0.696	38	Myrmeleontoidea + Nemopteridae + Ithonioidea + *Chrysopidae* + *Hemerobiidae* (Neuroptera)
0.678	45	*Hyocephalidae* + *Idiostolidae* (Hemiptera)
0.664	27	*Elateridae* (Coleoptera)
0.652	22	*Miridae* + *Tingidae* (Hemiptera)
0.628	24	Aphidoidea (Hemiptera)
0.614	4	Calyptratae + Drosophilidae + Ephydridae + Agromyzidae (Diptera)
0.598	**-**	***Termitidae* (Isoptera)**
0.594	8	Psocodea
0.592	16	*Apidae* + *Megachilidae* (Hymenoptera)
0.57	14	*Apataniidae* + *Goeridae* +*Limnephilidae* (Trichoptera)
0.566	10	Ephemeroptera
0.564	30	Amphiesmenoptera (Lepidoptera + Trichoptera)
0.542	**-**	***Buprestidae* (Coleoptera)**
0.522	18	Membracoidea (Hemiptera)
0.492	28	*Eulophidae* + *Pteromalidae* (Hymenoptera)
0.482	-	*Curculionidae* (Coleoptera)
0.476	37	Psocomorpha (Psocodea)
0.456	13	Bombycoidea (in part) + *Cimeliidae* (Lepidoptera)
0.442	36	*Pentatomidae* (Hemiptera)
0.442	3	Macromoths (Hyblaeoidea+ Pyraloidea+ Noctuoidea+ Geometroidea+ Drepanoidea+Bombycoidea+*Cimeliidae* (Lepidoptera)
0.424	**-**	***Tenebrionidae* (Coleoptera)**
0.414	31	Glossata (excluding Eriocraniidae)(Lepidoptera)
0.4	43	Collembola (with exceptions of Neelipleona, *Katiannidae and Tomoceridae*)
0.388	40	Ditrysia (Lepidoptera)
0.372	**-**	**Calyptrata (Diptera)**
0.368	**-**	***Sarcophagidae* + *Tachinidae* (Diptera)**
0.366	46	*Limacodidae* + *Zygaenidae* (Lepidoptera)
0.36	26	Asiloidea (in part) (Diptera)
0.354	44	*Clambidae* + *Eucinetidae* (Coleoptera)
0.352	39	*Ceratocanthidae* + *Hybosoridae* + *Glaresidae* + *Lucanidae* + *Pleocomidae* + *Glaphyridae* + *Ochodaeidae* + *Passalidae* (Coleoptera)
0.35	**-**	***Anthicidae* + *Meloidae* + *Mordellidae* (Coleoptera)**
0.348	42	Pterygota
0.346	-	Neopomorpha+ Gerromorpha + Enicocephalomorpha+ Dipsocoromorpha (Hemiptera)
0.342	-	*Blattidae* + *Blaberidae* (Blattodea)
0.336	**-**	***Mantidae* (Mantodea)**

4. The Impact of Dietary Ecology on Diversification in Hexapoda

4.1. Abstract

Hexapoda, the insects and their relatives, includes over half of all described species. Because large proportions of this diversity cluster within a small set of phytophagous groups, dietary-substrates have been proposed to shape patterns of richness within the clade through antagonistic co-evolution and zones of ecological opportunity. Here we explore these processes in the context of a recent dated phylogeny of Hexapod families. Our results indicate phylogenetic clustering of specialized ecologies such as phytophagy and parasitism, but reveal no consistent associations between the use of particular dietary substrates and clade richness. We also find no evidence that diets expected to promote antagonistic co-evolution are consistently associated with elevated species richness, nor that sister clades differing in dietary state are associated with greater-than-expected differences in richness. We do, however, identify variation in the age of, and transition rates among, dietary states that are likely to play a role in the observed heterogeneity in richness among dietary classes. Based on these findings we suggest remaining circumspect about the generality of adaptive zones based on broad dietary groupings as an explanation for hexapod richness, and suggest that richness heterogeneity may be better explained by origination and transitions rates, and variation within dietary categories.

4.2. Introduction

A key issue in macroevolution is how ecology affects speciation and extinction to generate differences in species richness among clades (Schluter, 2009). Ecological opportunity is a key potential part of this relationship, and refers to how niche space constrains the richness of clades using these niches (Valentine, 1980; Wellborn and Langerhans, 2015). Zones of ecological opportunity are challenging to visualize, as they exist in a multi-dimensional volume defined by a combination of many ecological traits (Devictor et al., 2010; Futuyma and Moreno, 1988). However, ecological zones may sometimes be approximated by single, simply measured traits, and the distribution of such traits may be studied on phylogenies of radiating taxa (Cantalapiedra et al., 2014; Poisot et

al., 2011). For example, the division of niche space into zones of opportunity would be expected to restrict transitions between such zones, resulting in strong phylogenetic conservatism in correlated traits (Cooper et al., 2010; Poisot et al., 2011). Likewise, because zones of opportunity are expected to differ in their control of net diversification rates and carrying capacity, transitions across different ecological zones should correlate with differences in the inferred diversification process and therefore the species richness of transitioning clades (Maddison et al., 2007; Rabosky, 2009). Understanding the role of ecology in structuring species richness among clades therefore relies on an understanding of diversification rates, the history of ecological evolution within the group, and an appreciation of the limits to zones of ecological opportunity.

Much of the work on diversification and ecology explores the relationship between species richness and host specialization (Poisot et al., 2011; Thompson, 2009). Host specialists, by definition, make use of only part of the resources available in an environment, and therefore zones of opportunity can be occupied by greater numbers of species, potentially resulting in more species-rich clades (Poisot et al., 2011; Vamosi et al., 2014). Furthermore, antagonistic coevolution, and loss of genetic variation, is expected to result in increased specialization among specialized daughter clades leading to potential long lasting impacts on clade diversification (Ehrlich and Raven, 1964). Specialization also imposes macro-evolutionary costs, such as reduced population and range size, which render species more susceptible to extinction, and which may mask or counter the effects of increased diversification rates (Kelley and Farrell, 1998; Nosil, 2002). Whether and how clades overcome this "paradox of parasitism" (Drake, 2003), and how this relates to zones of ecological opportunity, remain major outstanding questions.

A classic system for exploring the relationship between ecology and species richness is the macroevolution of Hexapoda, the six-legged arthropods that include insects and their relatives. Within this clade there is considerable variation among sub-groups in both species richness (Mayhew, 2007), and dietary ecology (Grimaldi and Engel, 2005), presenting an ideal system for studying relationships between these traits. In addition, due to the typical presence of a feeding nymph or larval stage with limited mobility, there is a long tradition in hexapod studies of using dietary substrates, e.g. phytophagy (Mitter et al., 1988; Nyman et al., 2010; Winkler and Mitter, 2008), parasitoidism (Wiegmann et al., 1993), fungivory (Leschen and Buckley, 2007), and generalized diets such as detritivory

and predation, as proxies for zones of ecological opportunity and therefore controls on clade diversification within the group (Mayhew, 2007).

Evidence that use of heterogeneous dietary substrates may promote clade richness in hexapods is based on the widely cited studies of Mitter et al. (1988), Farrell (1998) and Winkler and Mitter (2008). These analyses purported to show that plant feeding (Mitter et al. 1988), and specifically feeding on angiosperms (Farrell 1998; Winkler and Mitter 2008), is correlated with elevated diversity with respect to sister taxa, across insects as a whole and within Coleoptera (beetles). While this view has become standard in discussions of plant feeding and speciation (e.g. (Nyman 2010) and references therein) there remain a number of outstanding issues associated with this interpretation, including questions surrounding selectivity in the choice of sister group contrasts. Mitter et al. (1988) included only 13 comparisons in their analysis, which due to the state of phylogenetic and taxonomic information then available, show an implicit bias towards larger plant feeding groups. The authors acknowledged this ((Mitter et al. 1988): appendix) and justified the exclusion of small phytophagous radiations on the basis of their playing a marginal role in understanding overall patterns of clade diversification, due to their low diversity and that of their probable sister taxa. Attempts to test this assertion within Coleoptera indicated that such small families may in fact play pivotal roles in the clade's diversification, resulting in an analysis which failed to recover any consistent association between phytophagy and species richness (Hunt et al. 2007). In addition, conflicting evidence from parasitic hexapods challenges the generality of heterogeneous diets for promoting clade diversification (Futuyma and Moreno 1988; Wiegmann et al. 1993).

In recent years there has been a steady increase in the phylogenetic information available for Hexapoda e.g. (Misof et al., 2014; Trautwein et al., 2012), and in techniques for assembling such data into increasingly comprehensive frameworks for the group (Chapter 2). As a result it is now possible to extend the methodologies used by (Mitter et al., 1988) and others to consider a more inclusive view of hexapod diversification. The aims of this study are thus: a) to summarize the phylogenetic distribution of diets across higher insect taxa based on a consistent dietary classification (see Appendix 7.2) and evaluate what this implies about the historical patterns of dietary acquisition and loss, b) to demonstrate if there is phylogenetic conservatism in diet across the broad array of hexapod taxa as a prerequisite to a long term macro-evolutionary association between diet and species richness, and c) to investigate the association between net diversification and

dietary ecology, specifically whether the use of particular substrates is correlated with elevated or depressed richness among hexapod clades, and if consistent patterns occur among the set of diets that are expected to promote antagonistic co-evolution.

4.3. Methods

Underlying this study is a dated topology of Hexapoda, including 874 terminal taxa covering 903 of the approximately 1100 extant hexapod families (Chapter 2). Whilst clearly a phylogeny so inclusive will never be error-free, and some regions whilst plausible, are only weakly supported, this topology includes all clades highly supported by previous work at the level of hexapod families (Section 2.3.2), and is broadly consistent with recent opinions regarding the deep structuring of higher taxonomic relationships (Trautwein et al., 2012; Misof et al., 2014). We therefore propose it as the best current working basis for a broad and inclusive comparative study of hexapod diversification. Accompanying this tree are estimates of described species richness for terminal groups taken from recent encyclopedia and related sources (references in Section 3.3).

Dietary ecology for terminal groups was taken from published descriptions (Appendix 7.2) and categorized according to predominant substrate use among subfamilies or comparable groups. The substrates used include fungivory, detritivory, phytophagy (herbivory), predation, parasitoidism, and ecto-parasitism as well as non-feeding and liquid-feeding adults (Appendix 7.2). Diets were coded separately for juveniles and adults, with most non-metamorphosing taxa assumed to maintain the same ecology throughout the lifecycle. Omnivorous taxa or taxa in which subfamilies varied in predominant ecology were coded as mixed states (Appendix 7.2). In order to reflect differences in previous classifications regarding the treatment of marginal diets, such as whether to classify xylophagy and/or pollenivory under phytophagy or detritivory (Mitter et al., 1988; Hunt et al., 2007), and whether to group carnivorous parasites (parasitoids, ecto-parasites and other blood feeding taxa) as a single category (Wiegmann et al., 1993), we developed three distinct coding schemes, details of which are provided in Appendix 7.2. Our favored scheme, emphasizing larval/immature diets, is denoted “Larval Raw”. A scheme that more closely corresponds to the categories used in previous sister-group studies is henceforth “Larval Modified” (in parentheses; Appendix 7.2). Finally, a scheme based on the ecology of adult taxa is henceforth- “Adult”.

To assess the degree to which ecological states demonstrated non-random phylogenetic structure of across terminal groups, e.g. due to clustering or over-dispersion, we used Phylocom (Webb et al., 2008) to calculate two indices; net relatedness index (NRI- measuring total phylogenetic distance between taxa with particular ecologies) and nearest taxon index (NTI- measuring mean distance to nearest neighbor sharing a particular diet) relative to 999 randomized tip permutations. Given that hypothesized zones of ecological opportunity implicitly assume long-term associations of clades with particular dietary substrates (Irwin et al. 2012), the presence of phylogenetic conservatism can be regarded as evidence consistent with such models. Taxa with mixed coding states were treated as contributing to all relevant indices and taxa with for which no ecological information could be obtained (denoted by "?" in Appendix 7.2) as contributing to all studied indices, so as to minimize any biasing effect this lack of data might have on the analyses.

As the basis for subsequent sister-group comparisons (see below) we reconstructed ancestral dietary states under parsimony (using Mesquite (Maddison and Maddison, 2011)), and maximum likelihood (ML) using the hidden rates Markov model rayDISC (R package; corHMM (Beaulieu et al., 2013)). For the ML reconstruction of the "Adult" dataset the rarity of some ecologies, e.g. fungivory, resulted in an overexpression at deep nodes within the phylogeny, (see (Nosil, 2002)). To resolve this we constrained the root state for this reconstruction to detritivory in order to match the parsimony reconstruction. We converted reconstructed probabilities into discrete states using a threshold approach; with all nodes where a single state represented greater than 0.7 of the total probability referred to this state, and the remaining nodes referred to mixed-states encompassing all traits present with a probability of at least 0.05. These values were selected to maximize similarity with previous studies in inclusion of clades showing strong dominance of particular diets, while maintaining ambiguity where ancestral states are uncertain.

Comparisons of richness across sister clades with divergent ecologies were estimated for each novel origination of a trait on our tree following (Mitter et al., 1988). As with these authors our compared richness values subtracted any members of the focal clade belong to terminal taxa lacking the ecology of interest or members of the sister group processing the focal ecology, including within mixed states (henceforth corrected richness). For the purposes of corrected richness, we did not apply subtractions for dietary variation within our terminals due to limits in describing diet and species richness in clades

below the family level for many ecologically diverse groups (see discussion). Contrasts where the ecology of either taxon was unknown were excluded.

Corrected richness values were compared using the sign binomial test (Farrell et al., 1991) and "species diversity contrast" (SDC) methods that incorporate the magnitude of the diversity contrast between groups (Vamosi and Vamosi, 2005). Three SDC statistics were calculated (using Python script "Systers" (Hardy and Cook, 2010)), represented distinct approaches to scaling richness comparisons: raw contrast (Wiegmann et al., 1993), proportional contrast (Barraclough et al., 1995) and log contrast (Barraclough et al., 1996). Following (Vamosi and Vamosi, 2005), sample size dependent statistical tests were applied to these statistics, with very small samples (<6 comparisons) analyzed using a randomization test of matched pairs, and larger sets compared using a Wilcoxon non-parametric test or its normal approximation (for > 20 comparisons). In this study we did not use the well-known (Slowinski and Guyer, 1989) test for diet contrasts following evidence of an elevated type one error rate when multiple comparisons are combined within a single test (de Queiroz, 1998; Vamosi and Vamosi, 2005). Likelihood models of trait-dependent diversification processes, e.g. BiSSE (Maddison et al., 2007), were not used here, as current implementations rely on a species complete transition rate matrix for parameter estimation on trees of higher taxa, resulting in exponential growth in time and memory requirements, that render such methods computationally intractable on the scale of Hexapoda (FitzJohn et al., 2009). We further question whether such approaches, which estimate uniform speciation and extinction rates across whole dietary classes, are appropriate for clades where there is clearly enormous rate heterogeneity across subgroups (Chapter 3).

To assess the hypothesis that use of biochemically heterogeneous states, i.e. those expected to promote antagonistic co-evolution, might collectively act as drivers of species richness in hexapods we combined these states (phytophagy, parasitoidism, ectoparasitism and fungivory) into a single character, which we term "potential for specialization" or PS. Ancestral reconstruction of the PS character was conducted as above and corrected richness contrasts calculated on novel originations vs. "generalized" sister taxa using the SDC methods described above.

We also explored the idea that related clades occupying different ecological zones (i.e. processing different diets) should be associated with larger than expected absolute differences in richness, arising from differences in the control of net diversification and

carrying capacity. To do this we calculated standardized differences in richness at each node in the tree (standardized relative rate difference; stRRD, defined as the absolute contrast in log richness of the two descendant clades) using the trickle down protocol of (Davies et al., 2004) as compensation for phylogenetic nesting. The latter uses single-contrast Slowinski-Guyer (1989) tests to identify nodes associated with significant differences in the species richness of descendant clades and then compensates for such shifts for comparisons of more deeply nested nodes (Davies et al., 2004). Our implementation differs from previous work in that where the richnesses of both descendant clades differ significantly from that of the outgroup, we use the sum of the two descendant richnesses for more deeply nested comparisons, thus avoiding corrections where the direction of richness change is ambiguous (Davies et al., 2004). Following standardisation we calculated the mean stRRD value for nodes where ecology was divergent between descendent clades, including mixed states and compared this with the distribution of means of 1,000 sets of equal length, drawn at random from nodes of the tree.

4.4. Results

Based on our favored, "Raw Larval" classification, just over half of all hexapod species belong to families that contain at least some plant feeding taxa (527,000 species of the 1,038,000 estimated described taxa within the clades present on the discussed tree; (Rainford et al., 2014)). This compares with approximately thirty percent represented each by families including detritivorous and predatory representatives (330,000 and 322,000 species respectively), nineteen percent for fungivory (194,000 species), thirteen percent for parasitoids (136,000 species) and less than one percent ecto-parasites (7700 species). By comparison adult hexapods are dominated by liquid feeding taxa, which comprise the majority of adult Holometabola (411,000 species), with approximately equal proportions of detritivorous (301,000 species), phytophagous (351,000 species) and predatory groups (260,000 species), and minority representation of fungivores (167,000 species), non-feeding groups (116,000 species) and blood-feeders (39,000 species). Note that the percentages given here incorporate terminal taxa with mixed ecologies into each of the relevant dietary categories hence they exceed one hundred percent.

There is significant phylogenetic clustering of fungivory, phytophagy, parasitoidism and ectoparasitism under both the NRI and NTI metrics for the "Raw Larval" and "Larval Modified" coding systems (Table 6). This implies that on average,

taxa with these ecologies tend to be closely related to other taxa with the same diet (Figure 21). This pattern is not observed in detritivory and predation, both of which show non-significant trends towards over-dispersion in both larval datasets with respect to NRI. For adult ecologies, significant clustering is observed for fungivory, blood-feeding and non-feeding diets while predators, detritivores and plant feeders show no significant trend with respect to NRI.

Both parsimony and ML reconstructions identified detritivory as the ancestral and most widespread larval ecology within Hexapoda (Figure 21), although under ML some degree of fungivory is inferred in the early insect radiation, based on the diet of the poorly known basal order Protura (Pass and Szucsich, 2011). The two methods identify broadly similar patterns in resource use across the tree, with most ordinal groups showing strong conservatism with respect to diet, resulting in a pattern dominated by a small number of well characterized radiations, for example that of plant feeding within Lepidoptera and parasitisoidism within Hymenoptera (Grimaldi and Engel, 2005).

Disagreement between reconstruction methods typically reflects taxa showing high degrees of ecological lability, e.g. Coleoptera, where ML identifies fungivory as the ancestral ecology, and multiple originations of detritivory, whereas under parsimony this pattern is reversed. This reflects genuine ambiguity regarding the deep topology of the order, as well as the close relationships of many families associated with both dietary states (e.g. with wood-boring or soil-living lifestyles) (Hunt et al., 2007). Likewise basal members of the fly suborder Brachycera show conflict in the origination of predation, which under ML is recovered as multiple independent origins from a detritivorous ancestor in Asiloidea, Empidoidea and Tabanomorpha, as opposed to a single origin, with a return to detritivory in Stratiomyomorpha and Cyclorrhapha observed under parsimony (Marshall, 2012).

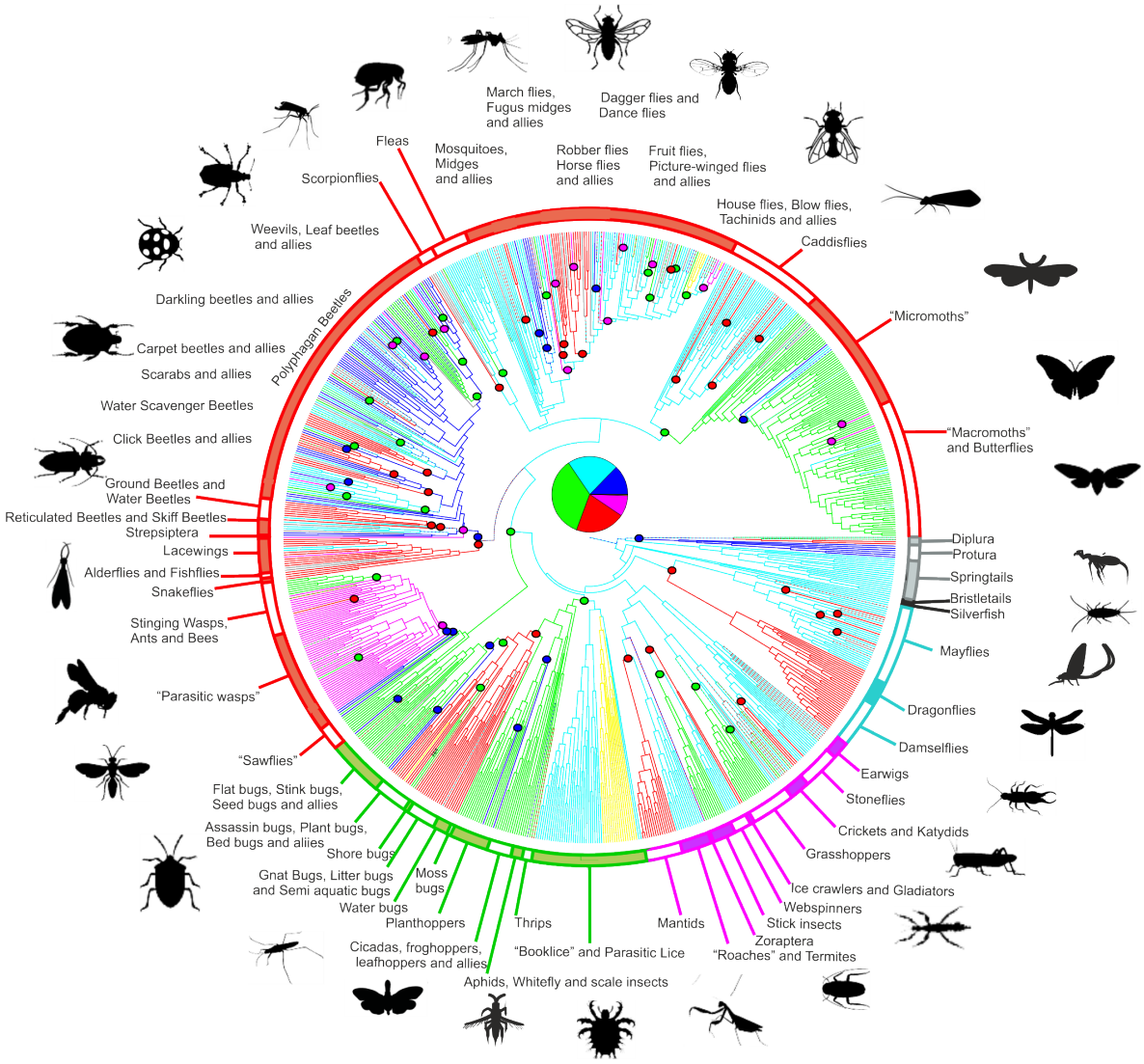


Figure 21 Reconstructed dietary ecologies for the Larval Raw dataset under maximum likelihood. Ecologies are denoted as follows: Dark Blue- Fungivory, Cyan- Detritivory, Green- Phytophagy, Red- Predatory, Magenta- Parasitoids, Yellow- Ectoparasites. Taxa and nodes with mixed states are shown by dashed lines. Taxa with unknown states are shown in Grey. Coloured dots denote the postions of sister group comparisons. The colouration of the outer ring denotes major clades (Grey; Entognatha, Black; basal insects, Cyan; Palaeoptera, Purple; Polyneoptera, Green; Paraneoptera, Red; Holometabola). Internal piechart gives the relative species richness associated with each dietary category, with taxa with mixed ecologies contributing to all relevant states, see Appendix 7.2.

The implemented ML model for the Larval Raw dataset allowed all transition rates to vary independently (all rates variable AICc = 1182.3, vs. equal rates AICc= 1244.2, a significant difference of 61.87 likelihood units). The highest-obtained transition rates occur between fungivory and detritivory (Table 4), reflecting the widespread nature of these diets and frequent transitions within Coleoptera and Diptera (see above). Transitions between detritivory and predation also occur at high rates, being particularly common among freshwater taxa, e.g. caddisflies and mayflies where multiple parallel origins of predatory larvae occur in various families. Ecto-parasitism is identified as a dead end with respect to ecological diversification with no examples of the emergence of other ecologies within a primitively ecto-parasitic group (transition probabilities equal zero). Note that ecto-parasitism in larval hexapods is extremely rare and restricted to four clades, including one ancient origination in Pthiraptera (lice; whose age is highly uncertain due to lack of suitable fossils (Grimaldi and Engel, 2005) and accelerated rates of genomic evolution (Trautwein et al., 2012)), and three further events occurring among young terminal groups with modest extant diversity. Other transitional dead-ends also appear: there are no direct transitions between fungivory and ecto-parasitism or from parasitoidism to fungivory or detritivory. In the later case this result is dependent on the coding of pollenivory based on the results of the Larval Modified dataset (Figure 22, Table 5).

Table 4 Overall Likelihood and Transition rates per million years inferred for the optimal ML model of Larval Raw Data set. Overall LnL: 563.59, AIC: 1187.17, n. taxa: 874. Models are denoted as transition rates from rows to columns.

	Fungivory	Detritivory	Phytophagy	Predators	Parasitoids	Ecto-parasites
Fungivory	NA	0.00187	0.00060	0.00065	0.00022	0.00
Detritivory	0.00016	NA	0.00039	0.00072	0.00016	0.00004
Phytophagy	0.00031	0.00015	NA	0.00003	0.00010	0.00
Predators	0.00007	0.00016	0.00015	NA	0.00009	0.00004
Parasitoids	0.00	0.00	0.00035	0.00016	NA	0.00011
Ecto-parasites	0.00	0.00	0.00	0.00	0.00	NA

Table 5 Overall Likelihood and transition rates per million years inferred for the optimal ML model of Larval Modified dataset. Overall LnL: 561.30, AIC: 1162.61, n. taxa: 874. Models are denoted as transition rates from rows to columns

	Fungivory	Detritivory	Phytophagy	Predators	Parasites (Combined)
Fungivory	NA	0.002054	0.00068	0.00063	0.00019
Detritivory	0.00015	NA	0.00037	0.00074	0.00024
Phytophagy	0.00033	0.00015	NA	0.00003	0.00010
Predators	0.00007	0.00016	0.00021	NA	0.00014
Parasites	0.00	0.00014	0.00009	0.00011	NA

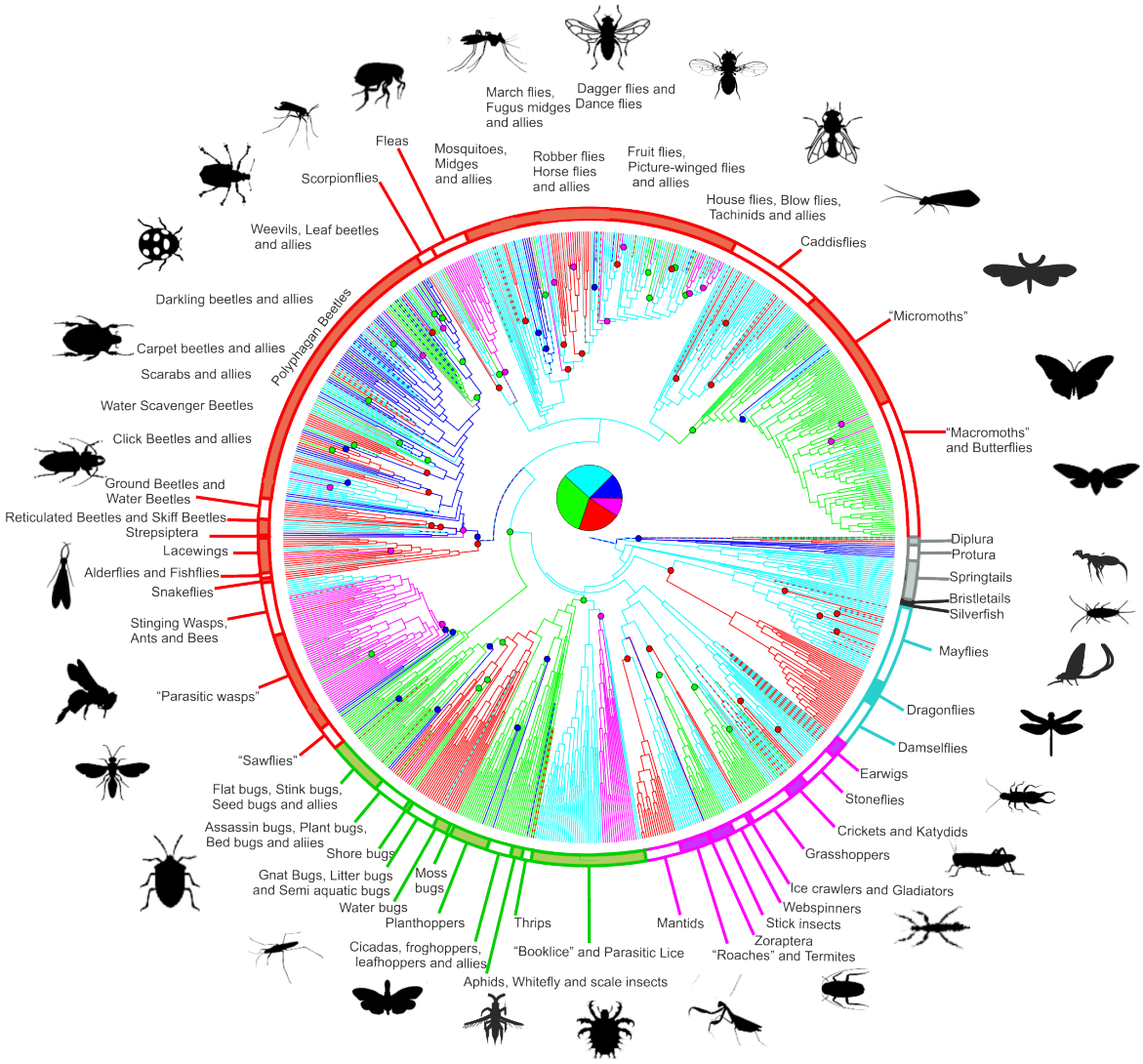


Figure 22 Reconstructed ecologies for the Larval Modified dataset under maximum likelihood. Colours are as Figure 21 with Magenta denoting carnivorous parasites. Internal piechart gives the relative species richness associated with each dietary category with taxa with mixed ecologies contributing to all relevant states, see Appendix 7.2.

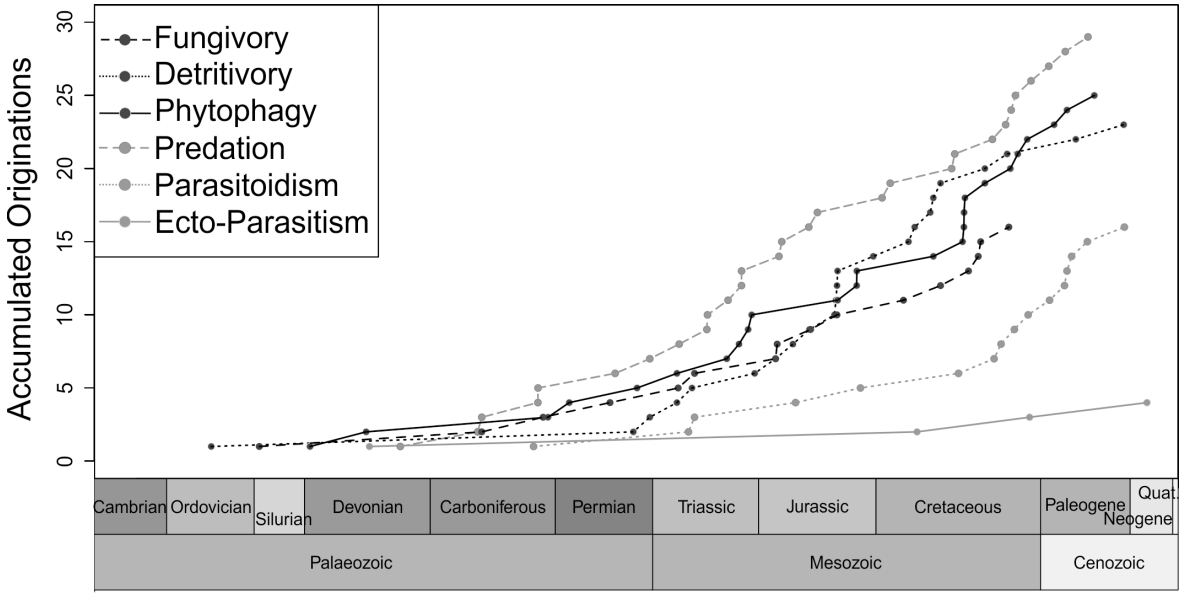


Figure 23 Accumulation Plot of dietary originations through geological time for Larval_Raw ML reconstruction.

The accumulation curve of dietary originations for the ML reconstruction of the Larval Raw dataset demonstrates differences in the relative times and rates of origination between diets that collectively may contribute to their respective differences in richness (Figure 23). These series reveal certain ecologies including fungivory, phytophagy and predation as appearing early in the history of the group and undergoing approximately consistent rates of origination throughout the history of Hexapoda. This contrasts with patterns in detritivory, and parasitoidism both of which are strongly skewed, such that the majority of originations occur within specific time intervals (respectively the Middle and Late Mesozoic). Friedman tests using the number of originations within 50Ma bins reveal significant differences in the origination rate across dietary categories (maxT = 3.068, p-value = 0.02622).

Post-hoc analysis; (Galili, 2010), identifies significance as driven by a large contrast between the predatory and ecto-parasitic dietary classes (p=0.0263) as well as a marginally non-significant contrast between ecto-parasitism and phytophagy (p=0.0574). By comparison the ML reconstructed Larval Modified dataset (where ecto-parasitism and parasitoidism are combined) reveals no comparable differences between binned ecological categories (maxT = 1.6813, p-value = 0.446) suggesting that these differences arise from splitting these two distinct forms of carnivorous parasitism into discrete categories. Likewise there are no significant differences in the binned counts among the ecologies in the Adult dataset (ML reconstruction; maxT = 2.4076, p-value = 0.195).

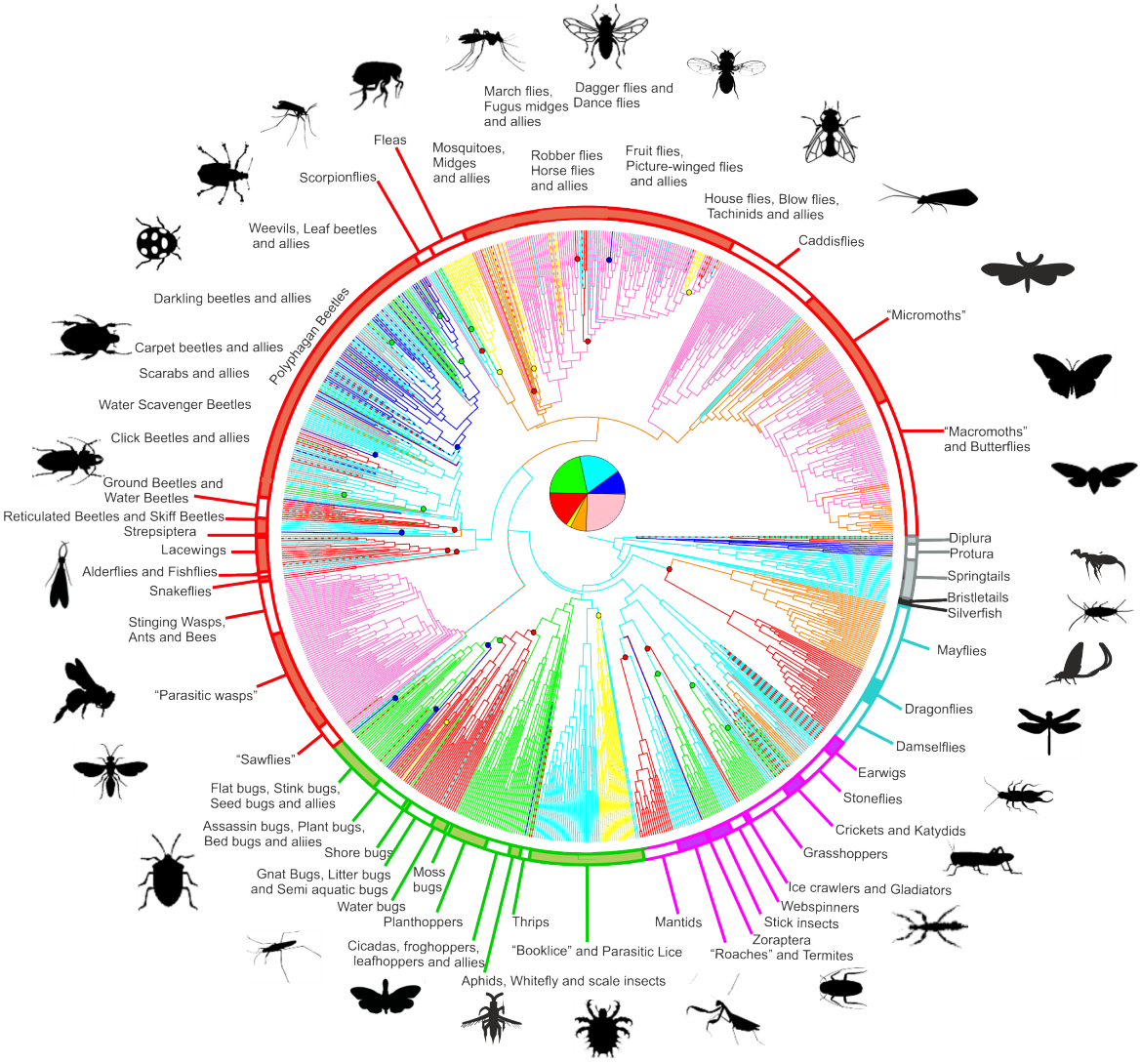


Figure 24: Reconstructed adult dietary ecology under maximum likelihood. Colours are as Figure 1 with the addition of Orange- NonFeeding and Pink- Liquid feeding/Nectivory. Yellow is used to denote both Ectoparasites and adult blood feeding taxa. Coloured dots denote the positions of sister group comparisons. Internal piechart gives the relative species richness associated with each dietary category with taxa with mixed ecologies contributing to all relevant states, see Appendix 7.2.

Compared with the larval phase, information on the extent and significance of adult feeding within many hexapod groups is relatively uncertain, resulting in poorer documentation and fewer records of adult diet. As a result, reconstructions from our adult dataset are subject to extrapolation errors under ML associated with rare ecologies (see above) and there are high degrees of conflict between reconstruction techniques (Figure 24). Major regions of conflict include: the ancestral state of Mecopterida (including Trichoptera, Lepidoptera, Diptera, "Mecoptera" and Siphonaptera) (Trautwein et al., 2012), the relative importance of fungivory in the diets of adult beetles (many of which are polyphagous relative to their larval stage), and ancestral diets within Heteroptera.

Sister group comparisons failed to show any significant effect of any dietary ecology on species richness. The only potential exception was the reconstruction of detritivory under parsimony, which showed a significant trend with respect to Raw contrasts towards increased richness (Table 7), primarily driven by the novel origination of detritivory in Cyclorrhapha (Diptera) (conflicting with ML reconstruction). The analysis of the Larval Modified dataset produced similar results indicating that this lack of previously identified relationship was not simply a manifestation of differences in the coding system between this work and previous studies. Similarly no significant trends in richness association were observed with respect to adult ecology.

The reconstructed history of the PS character state was identical under parsimony and ML methods and corresponded to the major specialized groups previously described (Figure 21, Figure 25). The associated transition matrix for the ML model identifies a marginally significant bias in transition rates towards the evolution of more specialized ecologies (0->1: 0.00072 myr^{-1} vs. 1->0 0.00035 myr^{-1}, AIC= 526.51, vs. an AIC of 528.19 for an equal rates model). Sister group comparisons between PS and non-PS groups failed to recover any evidence for the trait promoting diversity with exactly half of the test comparisons running contrary to the view (24 of 47 instances, Sign Test p value= 0.5106, Wilcoxon Tests: Raw Contrasts; W=504, p= 0.52893, Log Ratio Contrasts; W=545, p=0.84479, Proportional Contrasts; W=524, p=0.67595).

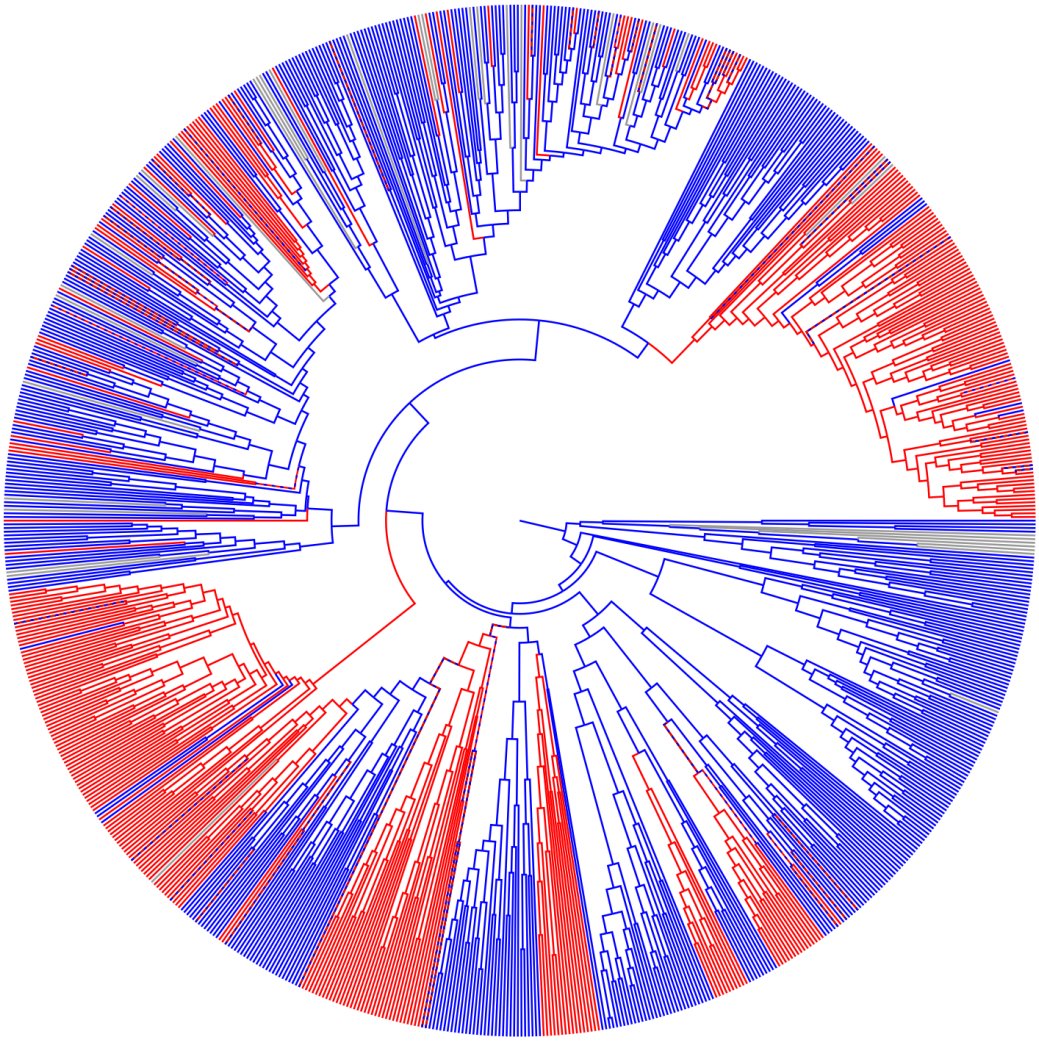


Figure 25; The reconstructed history of the "potential specialization" (PS) character described in the text. Clades with specialized ecologies are denoted in red, generalised ecologies in blue and taxa with unknown states in grey. Mixed states are shown by dashed lines. Tree orientation and clade identities are as Figure 21.

When we compare nodes where ecology diverges we show that for the "Larval Raw" and "Larval Modified" datasets there is no significant trend in terms of greater than expected contrasts at nodes with divergent ecologies. For the "Larval Raw" data, the mean estimate of the stRRD value associated with nodes showing ecological divergence was 2.1245, corresponding to a *p* value of 0.2741, assuming a normal distribution of the log means of the 1000 randomly sampled node sets (mean of random samples 2.061; sd 0.1058; Shapiro-Wilk normality test; W = 0.9994, p-value = 0.0948). For the "Larval Modified" mean stRRD = 2.116, *p* value= 0.2898, (mean of random samples 2.057; sd 0.1072; SW test; W = 0.999, p-value = 0.562).

By contrast in the "Adult" dataset there was a marginally significant trend towards larger contrasts, estimate of stRRD = 2.252, p= 0.0484**,** (Mean of random samples 2.0572; sd 0.1175; SW test; W = 0.998, p-value = 0.296). It is unclear whether this difference between datasets simply reflects the presence of more character states in the Adult dataset (thus generating more divergent nodes) or if it is the influence of a small number of deep dietary shifts, e.g. the transitions to predominantly non-feeding or liquid feeding diets within Holometabola. Based on these results we conclude that one of the basic assertions of a ecological zone model, that groups with divergent ecologies should show greater than average differences in richness (after standardization to render contrasts independent), cannot be demonstrated as holding at the resolution of hexapod families

4.5. Discussion

Our results explore the association between dietary substrate and hexapod richness via adaptive zones of opportunity. Surprisingly, no association was found between particular dietary ecologies and net diversification, in contrast to previous findings relating to phytophagous insects. Also, our specialized and generalized diets do not consistently promote differences in species richness. Nor are nodes involving dietary shifts attributed greater shifts in diversification than random nodes. There is, however, evidence for strong phylogenetic conservatism of specialized diets but not generalized diets. Finally, there are differences in transition rates and origin times between diets, with a tendency towards evolution of diets associated with co-evolution and host specialization. Below we interpret these findings and their consequences for understanding how diet affects hexapod diversification.

Our original motivation was to explore how hexapod diet acts as a proxy for zones of ecological opportunity (Mayhew 2007). The association of diet with zones of ecological opportunity is expected to result in phylogenetic clustering, arising from restricted transitions between diets (Cooper et al. 2010; Poisot et al. 2011). We observed phylogenetic conservatism in the use of biochemically and mechanically heterogeneous resources such as fungivory, phytophagy and parasitoidism. Such clustering is weaker for detritivory and predation (Table 6). Strong phylogenetic conservatism in hexapod diets, e.g. herbivorous insects, is widely acknowledged, e.g. (Ehrlich and Raven 1964; Mitter and Farrell 1991; Futuyma and Agrawal 2009), although comparisons across multiple substrates are rare.

Phylogenetic conservatism in the use of heterogeneous substrates may be generated by a requirement to overcome with host defenses (Mitter and Farrell 1991; Futuyma and Agrawal 2009) and skewed nutrient content (Mattson 1980; Douglas 2009), which may restrict colonization of these resources by novel hexapod lineages. By contrast, intermediate stages such as scavenging, incidental predation of cohabitants, and cannibalism, may serve to lower such barriers for originations of "generalist" diets (Coll and Guershon 2002), allowing their adoption by a wider range of clades (Figure 23), and as a result, reduced phylogenetic clustering across Hexapoda. The higher transition rates from generalized to specialized diets also supports this interpretation, mirroring previous species-level studies within dietary groups (Nosil 2002; Nosil and Mooers 2005). These findings suggest that "specialized" ecologies are consistent with being zones of opportunity, at the family level, in contrast to "generalized" diets. However, this pattern alone is insufficient to define dietary adaptive zones, which are also contingent on the association of different diversification processes with particular diets (Maddison et al. 2007; Rabosky 2009).

Contrary to previous publications, our work finds no evidence that plant-feeding groups are consistently more species rich than their sisters; a view which has been very influential (e.g. (Grimaldi and Engel, 2005; Nyman, 2010) and references therein). We think it unlikely that our lack of evidence reflects a lack of power as our sampling of phytophagous clades was more comprehensive than in previous studies (13 sister comparisons in (Mitter et al., 1988), but 25- Larval Raw or 26 -Larval Modified here). Instead these differences probably arise from earlier selective sampling towards representation of larger and phylogenetically better known plant-feeding groups ((Mitter et al., 1988)- appendix), and heterogeneity in the macro-evolutionary dynamics of phytophagous lineages. Our total number of parasitism comparisons is identical to that of previous studies (15- (Wiegmann et al., 1993)), although the identity of the groups shows minor differences. In agreement with previous work, our analysis fails to show a consistent association of parasitism and species richness, including where invertebrate and vertebrate parasites are treated separately (as in Larval Raw).

Our study does not distinguish between phytophagous clades feeding on angiosperms and gymnosperms (Farrell, 1998; Winkler and Mitter, 2008), because hexapod families often include species feeding on both plant clades. The idea that host clades may produce different diversification dynamics, has been discussed extensively

(Winkler and Mitter, 2008; Nyman, 2010), e.g. the Cretaceous radiation of angiosperms expanding the ecological space available for plant-feeding clades (Labandeira and Eble, 2000; Grimaldi and Engel, 2005). However, testing this assertion is challenging due to uncertainties in the timing of hexapod diversification (Chapter 2), the extended emergence of angiosperms prior to their appearance in the fossil record (Clarke et al., 2011), and because insect radiations may be decoupled from the radiations of their hosts e.g. (McKenna et al., 2009). As such it is beyond the reach of the presented datasets.

Our analysis also fails to identify consistent differences in net diversification between specialist and generalist diets (represented here by our PS character). Likewise, sister taxa that differ in ecology do not experience significantly increased differences in richness, as would be expected if diets represent different adaptive zones (Rabosky, 2009). This opens up alternative perspectives on the controls on hexapod diversification (see below).

Differences in the richness of dietary groups can alternatively be explained by historical factors in the evolution of hexapod diet. For example, the timing of originations, such as bias towards post-Mesozoic originations in parasitoidism (Labandeira and Eble, 2000; Grimaldi and Engel, 2005), may limit the richness of such clades when compared with older (e.g. phytophagous) lineages (Figure 23). Time-lagged evolution of suitable hosts may also limit the richness of ecto-parasites (Whiting et al., 2008), while other diets such as fungivory, detritivory and phytophagy have undergone many parallel and ancient origins which may partly account for differences in their richness (Figure 23) (Grimaldi and Engel, 2005).

The flipside of origination is extinction; however evidence regarding the latter is limited in phylogenetic studies of extant taxa (Labandeira and Eble, 2000). Insect fossils, whilst providing direct evidence of the extinction of higher taxa, provide little evidence regarding the diets of extinct groups (Grimaldi and Engel, 2005). One major, (probably) phytophagous group, the Palaeodictyoptera, went extinct at the Permo-Triassic mass extinction (Labandeira, 2006, 2013), however the implications of this for modern plant-feeding clades remains unclear, and to date no fossil studies have attempted to compare the extinction rates of different dietary categories.

The frequency of transitions between different dietary groups is another historical factor potentially affecting the richness of different diets. Ecto-parasitism originates at a

very low rate compared to other dietary categories, as does parasitoidism (Table 4). Some transition rates to the other dietary categories are much higher; for example from fungivory to detritivory, fungivory to phytophagy, fungivory to predation, and detritivory to predation (Table 4). The high rates associated with transition away from "generalized" ecologies mirrors recent findings with respect to mammalian dietary evolution (Price et al., 2012), although contrary to the latter it seems unlikely that "generalized" categories in hexapods represent unstable intermediates between specialized diets (see discussion of omnivory in (Price et al., 2012)). Ecto-parasitism appears to be an evolutionary dead-end in hexapods (Kelley and Farrell, 1998), resulting in no further transitions to other dietary substrates and likewise, at the studied resolution, there are no transitions from fungivory or phytophagy to ectoparasitism (Table 4). This may be due to extreme differences in nutrient content, the requirements of appropriate nutritional symbionts (Douglas, 2009) and limited opportunity for the establishment of long-term insect-host associations (Lehane, 1991; Balashov, 2006). Thus the data presented here suggest that, even in the absence of consistent differences in net diversification between diets, the historical pattern of originations and transitions between diets goes some way towards explaining the heterogeneity in richness between different dietary categories. On their own however they fail to provide an explanation for the exceptional richness of Hexapoda that the adaptive zones hypothesis potentially provides.

Given these findings, what is the role of dietary adaptive zones in hexapod diversification? One possibility is that the real impact of diet is masked by uncertainties in hexapod taxonomy, phylogeny and ecological description. Discussion of the problems of monophyly and richness estimates can be found in Chapter 2. Phylogenetic uncertainties in hexapod relationships (Trautwein et al., 2012; Misof et al., 2014) could in principle bias the results, particularly if small clades with divergent diets, whose phylogenetic placement is generally less certain, have been systematically wrongly placed next to taxa with greater richness than that of their true sister groups. However, we currently have no reason to suspect such a bias, and thus expect to maintain signal across the implemented tests. Overall, the presence of numerous, and previously neglected, species poor phytophagous taxa give reason to remain circumspect on the generality of the findings of Mitter et al. (1988) and positive and sensible results found elsewhere in this paper suggest substantial signal regarding diet and species richness is present within our dataset.

Unseen ecological variation within families may also bias our results. However, available descriptions severely limit analysis at finer taxonomic scales: there are often no species richness estimates below the family level, or published descriptions may not attribute observed variation in diet to particular sub-taxa, particularly where observations are known for only a few species. Within-family phylogenetic uncertainties are also limiting; for example previous work involved contrasts using subfamilies of Scarabeidae and Coccinellidae (Coleoptera) (Mitter et al., 1988; Hunt et al., 2007), the sister groups of which have been disputed by subsequent phylogenetic work, e.g. (Smith et al., 2006; Magro et al., 2010). It is possible that compiling all sets of ecological contrasts would collectively reveal different patterns to those described here, however this would still leave considerable heterogeneity in diversification within dietary groups to be explained (Mayhew, 2007).

As this study draws extensively on the results of ancestral reconstruction it is important to acknowledge the sensitivity of these techniques to model misspecification, rapidly evolving traits, widespread convergence and transitions via intermediate states which may be lost in extant representatives (Cunningham et al., 1998; Nosil, 2002; Beaulieu et al., 2013). The uncertainties surrounding historical hexapod diets and the timings of dietary transitions (Labandeira, 2006, 2013) limit the extent to which we can test the impact of these limitations. However there tends to be close agreement between the transitions shown here and previous broad historical hypotheses (see Grimaldi & Engel 2005).

Sister clade comparisons of species richness explore only the sum of speciation and extinction impacting on focal taxa. Approaches that attempt to tease apart these processes e.g. BiSSE (Maddison et al., 2007) have become increasingly popular and may play a role in resolving some of the ideas discussed here. Note however that in current implementations these procedures have their own limitations (see methods). Very recently the idea has been proposed of using global inference of diversification processes in combination with tree pruning to describe the subset of diversification rates associated with procession of a particular trait e.g. (Weber and Agrawal, 2014). This is an idea that holds considerable promise for future work, however once again there are issues relating to its implementation for Hexapoda.

Leaving aside the above methodological issues, there remains the possibility that the observed lack of association between diet and diversification reflects real features of

hexapod evolution. This implies that, rather then each diet being linked to a particular diversification process, different clades using the same substrate respond in different ways. In other words, substrate-based classifications, such as that applied here, may be poor approximations for the real zones of ecological opportunity that have shaped hexapod diversification. Instead we should consider how other features of diet or hexapod traits may shape the macroevolution of diversifying clades (Mayhew, 2007). A simple example would be host clade specific diversification, such as between gymnosperms and angiosperms (see discussion above). However, other features, such as differences in spatial context, e.g. between terrestrial and aquatic taxa (Hunt et al., 2007), and the role of ecological co-variates such as body size and dispersal capacity, may modify the effect of diet on clade diversification (Isaac et al., 2005; Phillimore et al., 2006).

A further possibility is that differences among taxa in their ability to transition between ecological zones may have consequences for their relative diversification (Dodd et al., 1999). Evidence for evolvability as a correlate of richness remains limited (Dodd et al., 1999), and is subject to theoretical issues regarding the underlying model of character change (Ricklefs and Renner, 2000; Silvertown et al., 2000). These, as well as data restrictions due to our incomplete sampling of hexapod diets (see above), mean that we do not incorporate such ideas into this study, although we acknowledge the potential for future analysis in exploring these ideas.

One potential source of heterogeneity within diets is the ecological feeding guild (Simberloff and Dayan, 1991), i.e. the manner in which taxa utilize a particular resource. Evidence for the importance of guild-specific processes can be found in community assembly studies (Novotny et al., 2010, 2012), as well as differences in the fossil dynamics of higher taxa (Labandeira, 2006, 2013). However, to date few guilds have been explicitly explored in terms of diversification e.g. leaf-mining; (Connor and Taverner, 1997), galling (Hardy and Cook, 2010). Some others, such as the distinction between idiobiont (which restrict host development from the point of parasitism) and koinobiont (which must deal with active defense by the developing host) parasitoids, have been subject to intensive speculation e.g. (Hawkins, 2005; Santos and Quicke, 2011) and warrant serious consideration in future studies.

Since the work of authors such as (Gould and Calloway, 1980) and (Mitter et al., 1988) the broad emphasis of trait-based diversification studies has predominantly been on testing a-priori hypotheses regarding the association of traits with patterns of

diversification. While acknowledging the power of this approach, there is a need to be rigorous in discussing the relationships between studied proxies (e.g. dietary substrate) and the processes of interest that may have acted to shape clade diversification (e.g. host specialization and zones of opportunity) (de Queiroz, 2002; Vamosi et al., 2014). In our analyses we group a set of ecologies potentially associated with promoting co-evolution and host specialization, under the expectation that these might show common patterns of clade diversification (our PS traits). However, a clear definition of "specialization" that is applicable when comparing different ecologies remains lacking (Devictor et al., 2010; Poisot et al., 2011), rendering comparisons between diets ambiguous (Giller, 1996; Nyman, 2010). Attempts to resolve this issue through metrics of specialisation (e.g. based on the number or phylogenetic diversity of host lineages used by taxa, e.g. (Forister et al., 2015), or on measures of interspecific competition within communities e.g. (Kaplan and Denno, 2007)), have yet to be widely adopted and remain restricted to single dietary classes (Poisot et al., 2012). There is therefore a need to use language rigorously to describe these interactions and their relationship to theoretical models of niche divergence ((Vamosi et al., 2014) and references therein).

To conclude, the work presented here suggests that, while some diets show strong conservatism at the level of hexapod families, and the origination dates and transition rates between different broad diets go some way towards explaining their heterogeneity in species richness, evidence for differential diversification processes operating within these substrate-based categories is lacking. It seems likely that by the restriction of discussion to arbitrary and subjective classifications we are failing to appropriately account for the different processes that may be responsible for shaping clade richness and how these relate to our measured proxy traits (Nyman, 2010). Understanding this linkage will require a combination of detailed ecological study, as well as further investigation into the macro-evolutionary process with a view towards defining appropriate hypotheses to test with comparative methods. Ultimately thereby we may establish a more "insect's eye" view of adaptive landscapes and thus enhance our understanding of the processes that drive diversification within the clade.

Table 3 Clustering analyses of different character states inferred for the different coding systems. Column headings: MPD (MPD.r, MPD.sd)- Mean phylogenetic distance of taxa possessing a particular ecology in the data set, and the mean and standard deviation of the implemented randomizations respectively, NRI- net relatedness index, MNTD (MNTD.r, MNTD.sd)- Mean Nearest Taxonomic distance of dataset and mean and standard deviation of value of the implemented randomizations respectively, NTI- Nearest Taxon index. MPD and MNTD are given in millions of years (the unit of branch length of the underlying phylogeny). NRI and NTI are dimensionless ratios, defined on the difference of the observed and mean expected values divided by the standard deviation of expected values, positive values referring to clustered data (Webb et al. 2008). P-values are calculated based on a two-tailed test.

Coding	Ecology	MPD	MPD.r	MPD.sd	NRI	p	MNTD	MNTD.r	MNTD.sd	NTI	p
Larval Raw	Fungivory	696.96	779.35	11.24	7.33	<0.001	341.46	373.91	15.66	2.073	0.0197
Larval Raw	Detritivory	781.83	781.50	5.45	-0.0597	N.S.	265.96	294.98	6.704	4.329	<0.001
Larval Raw	Phytophagy	730.25	792.10	5.60	11.04	<0.001	250.07	308.86	7.570	7.766	<0.001
Larval Raw	Predators	796.34	792.55	6.50	-0.584	N.S.	277.063	319.32	8.730	4.840	<0.001
Larval Raw	Parasitoids	657.03	789.18	11.20	11.80	<0.001	294.76	378.79	15.90	5.285	<0.001
Larval Raw	Ectoparasites	743.37	783.94	16.82	2.412	0.0124	386.16	438.73	25.53	2.059	0.0206
Larval Mod	Fungivory	698.48	779.32	11.28	7.166	<0.001	344.67	374.66	15.75	1.905	N.S.
Larval Mod	Detritivory	787.41	781.36	5.554	-1.089	N.S.	269.30	296.26	6.844	3.941	<0.001
Larval Mod	Phytophagy	726.20	791.83	5.831	11.256	<0.001	250.05	311.67	7.853	7.846	<0.001
Larval Mod	Predators	796.94	792.79	6.478	-0.6403	N.S.	278.00	320.46	8.816	4.817	<0.001
Larval Mod.	Parasites	704.64	790.12	9.324	9.1671	<0.001	283.24	355.68	12.81	5.657	<0.001
Adult	Fungivory	677.92	779.89	12.64	8.0641	<0.001	347.71	391.23	18.22	2.388	0.0091
Adult	Detritivory	791.08	779.12	7.215	-1.6581	N.S.	313.27	318.95	9.129	0.622	N.S.
Adult	Phytophagy	788.84	782.84	9.091	-0.66	N.S.	320.58	346.67	12.10	2.157	0.0156
Adult	Predators	797.92	786.07	7.640	-1.5506	N.S.	290.80	331.24	10.18	3.974	<0.001
Adult	Blood feeders	731.18	788.42	12.41	4.6106	<0.001	326.60	399.86	18.70	3.918	<0.001
Adult	Non feeding	734.51	792.06	9.852	5.8416	<0.001	276.80	371.21	14.45	6.534	<0.001
Adult	Nectivory	653.43	793.85	5.064	27.731	<0.001	219.49	298.83	6.714	11.82	<0.001

Table 4 Statistical tests of sister group comparisons. Methods of reconstruction are denoted P; parsimony and ML; Maximum Liklihood. Number of contrasts (N. cont.) and number of successes (where the richness of the focal origination was greater than its sister clade ; N succ.) are denoted for interpreting the sign test. SDC methods are given as results of Wilcoxon tests or their normalised equivalent, with the exception of tests denoted by an asterix which are the results of randomised matched pairs.

Coding system	Ecology	Method	N cont.	N succ.	p value	Raw Contrasts		Log Ratio Contrasts		Proportional Contrasts	
						W (S*)	p (two tailed)	W (S*)	p (two tailed)	W (S*)	p (two tailed)
Larval (Raw)	Fungivory	P	15	5	0.3333	45	0.4212	50	0.5995	50	0.560
Larval (Raw)	Fungivory	ML	16	5	0. 3125	42	0.1928	49	0.3484	43	0.211
Larval (Raw)	Detritivory	P	19	4	0.2105	31	0.0082	54	0.1042	49	0.066
Larval (Raw)	Detritivory	ML:	24	13	0.5417	111	0.4202	134	0.9152	120	0.595
Larval (Raw)	Phytophagy	P	25	10	0.400	123	0.2940	127	0.3463	151	0.767
Larval (Raw)	Phytophagy	ML	-	-	-	-	-	-	-	-	-
Larval (Raw)	Predators	P	31	15	0.4839	232.0	0.7613	222	0.6173	222	0.617
Larval (Raw)	Predators	ML	29	13	0.4483	210	0.8797	215	0.9655	214	0.948
Larval (Raw)	Parasitoids	P	15	5	0.3333	45	0.4212	50	0.5995	50	0.600
Larval (Raw)	Parasitoids	ML	-	-	-	-	-	-	-	-	-
Larval (Modified)	Phytophagy	P	25	11	0.440	122	0.2818	127	0.3463	157	0.893
Larval (Modified)	Phytophagy	ML	26	12	0.4615	124	0.1952	144	0.4311	171	0.919
Larval (Modified)	Parasites	P	19	7	0.3684	78	0.5153	81	0.5949	81	0.595
Larval (Modified)	Parasites	ML	-	-	-	-	-	-	-	-	-
Adult	Fungivory	P	18	6	0.3333	52	0.1541	60	0.2837	66	0.417
Adult	Fungivory	ML	10	2	0.2	81492*	0.8320	9.672*	0.1211	1.685*	0.174
Adult	Phytophagy	P	12	7	0.5833	17	0.0923	33	0.6772	151	0.733
Adult	Phytophagy	ML	-	-	-	-	-	-	-	-	-
Adult	Predators	P	13	5	0.385	43	0.8926	43	0.8926	43	0.893
Adult	Predators	ML	12	5	0.417	35	0.7910	36	0.8501	35	0.791
Adult	Blood feeders	P	7	5	0.7143	4684*	0.9688	5.107*	0.2031	1.070*	0.266
Adult	Blood feeders	ML	5	2	0.400	11613*	0.8125	0.261*	1.0	0.068*	0.938

5. Body Size Evolution and Diversity in the Hexapoda

5.1. Abstract

One of the most fundamental attributes of animal communities is the relatively high richness of small-bodied taxa across a wide range of animal clades. Work on the macroevolutionary processes responsible for this has been largely restricted to vertebrate model systems, so relatively little is known from other major taxa. Here we explore the macroevolutionary patterns of body size variation within a phylogeny of hexapod families (insects and their close relatives) and the links between size and diversity in this clade. 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 when compared with other hexapod groups consistent with accelerating rates of size evolution within this clade, while as a whole hexapod body size evolution is found to be dominated by neutral processes including significant phylogenetic conservatism. Overall 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.

5.2. Introduction

One of the most prevalent patterns observed in natural systems is the overrepresentation of small-bodied taxa (Kozłowski & Gawelczyk 2002). The observation of right skew in body size distributions, following transformation to the log scale, has been made for a variety of vertebrate clades (Maurer 1998; Gardezi & Silva 1999; Allen et al. 2006) and provides the basis for a variety of size-selective diversification mechanisms that have been proposed as general models for the macroevolution of animals (Kozłowski & Gawelczyk 2002; Allen et al. 2006). 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 (Orme, Isaac, et al. 2002; Orme, Quicke, et al. 2002). 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 (Hutchinson & MacArthur 1959; Morse et al. 1985), the relative strength of gravitational (i.e. body weight) vs. viscous and inertial forces (Vogel 1994) and, via surface area to volume ratios and the scaling of exchange networks, controls the rates of metabolic processes such as temperature response (Brown et al. 2004) and gas diffusion (Harrison et al. 2010). 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 Chown & Gaston 2010; Davis et al. 2013; Gaston & Chown 2013, and references therein, for review of Hexapoda).

Based on these observations a number of size dependent mechanisms linked to clade diversification have been proposed (reviewed in (Gardezi & Silva 1999; Allen et al. 2006)). These include; hard limits on minimum size which restrict random character change (McKinney 1990), energetic models emphasizing the relative efficiency of small body sizes in the production of offspring (Sokolovska et al. 2000; Brown et al. 2004), and fractal environmental models, exploring the capacity for small-bodied taxa to more finely subdivide a given environmental landscape (Hutchinson & MacArthur 1959). The

relationship of these process to macro-evolutionary diversification remains incompletely understood, for example the relative contributions of size-biased cladogenesis (i.e. small taxa being more prone to speciation) (Maurer 1998), directional bias in size evolution within lineages; e.g. “Copes rule” (Hone & Benton 2005), and size-biased extinction (Monroe & Bokma 2013), in 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 (Harmon et al. 2010).

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 (Mayhew 2007). 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 (Chown & Gaston 2010). 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)(Chown & Gaston 2010).

Evidence to suggest that processes in hexapod size evolution my be distinct from larger vertebrate groups includes taxonomic compilations e.g. (Poulin & Morand 1997), regional faunal data e.g. (Ulrich 2006; Ulrich 2007) and broad-scale continental surveys (Finlay et al. 2006), all of which suggest that compared with the latter hexapods exhibit relatively little right skew in the distribution of log body size (Chown & Gaston 2010; Gaston & Chown 2013). 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. (Katzourakis et al. 2001), with one study identifying the opposite pattern with respect to Anisoptera (dragonflies) (Misof 2002).

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 (Mitter et al. 1988; Mayhew 2007; Rainford et al. 2014). 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 (Yang 2001; Chown & Gaston 2010). This process is taken to extremes in Holometabola, where during metamorphosis there is a fundamental reorganization of the body plan (Grimaldi & Engel 2005), and as a result various authors have suggested divergent processes of size evolution associated with this clade (e.g. accelerated rates of size evolution through time (Chown & Gaston 2010; Nel et al. 2013)). Similarly, as a potential correlate of clade richness ((Mitter et al. 1988; Nyman 2010); see Chapter 4), dietary ecology, may also interact with any association between body size and clade richness, particularly where body size within communities is structured by trophic level (see Novotny & Kindlmann (1996) for discussion).

The recent and growing consensus with regards to hexapod higher taxonomic relationships from molecular markers e.g. (Trautwein et al. 2012; Misof et al. 2014; Rainford et al. 2014) 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 these tools 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) that the apparent lack of skew in body size distributions (on the log scale) for hexapods, identified in regional faunas, persists when the group is considered in a global phylogenetic context, b) that this lack of skew results in a lack of consistent relationships between clade richness and body size within hexapods after accounting for phylogeny and within-taxon variation in size, c) that interactions exist between the body size and other hexapod key innovations (Rainford et al. 2014; Chapter 4), which may alter the combined impact of these traits on clade diversification, and, d) that body size evolution within hexapods can be described by simple mathematical models, and different major sub-clades (Rainford et al. 2014) vary in their fit to these descriptive processes (e.g. due to the presence or absence of complete metamorphosis).

5.3. 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 (Rainford et al. 2014).

The size data for this study is based on family level estimates of minimum and maximum body length collected from global, regional and taxonomic datasets (References given in Appendix 7.3). The use of length as a proxy for size is common in Hexapoda due to difficulties in estimating mass from dried museum specimens (Chown & Gaston 2010; Gaston & Chown 2013). Taxon-specific length to mass conversion factors (e.g. (Wardhaugh 2013)) 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 Appendix 7.3). 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 (Arnett 2000).

Estimates of clade richness follow (Rainford et al. 2014; Section 3.3) with the resolution of taxonomic conflict described in Appendix 7.3. 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 (Appendix 7.3).

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 (Novotny & Kindlmann 1996; Hodkinson & Casson 2000), 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 distributes was estimated using meta-analysis statistics that

assume a sample-size dependent relationship between the estimated sd and the observed range (Hozo et al. 2005). Thus, for very small clades (<15 taxa) sd was calculated using Equation [16] of (Hozo et al. 2005), 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 (Hozo et al. 2005). 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 (Ives et al. 2007) 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 (Trautwein et al. 2012; Misof et al. 2014; Rainford et al. 2014). The normality of the overall mean distributions, both at the level of terminal taxa, and with taxa weighted by their observed species richness (Figure 26) was assessed using an Agostino test (D'Agostino 1970) (implemented in R; package *moments* (Komsta & Novomestky 2012)). The phylogenetic distribution of minimum, maximum and mean body length, as well as the estimates of terminal standard deviation (Figure 27) were plotted using a Brownian motion (BM) ancestral reconstruction (Revell 2013) implemented in the package *phytools* (Revell 2012).

The degree of phylogenetic signal present in the data with respect to mean-of-logs size was assessed using Blomberg's K statistic (Blomberg et al. 2003), and by comparing the observed variance among the phylogenetically independent contrasts (PICs) with 1,000 randomized data replications, applying the correction of (Ives et al. 2007) to account for within-group variance (implemented in the package *phytools*) (Table 8). 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) (Blomberg et al. 2003). Data with no phylognetic 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* (Orme et al. 2012) which is optimized to explore associations of traits values and species richness at the level of higher taxa (Agapow & Isaac 2002; Isaac et al. 2003; Freckleton et al. 2008). Richness contrasts at each node were standardized using two metrics; relative rate difference (RRD); ($\ln(N_1/N_2)$; N_1= richness of descendant clade with larger body size, N_2=the richness of the other descendant clade) (Table 9, Figure 28), and proportion dominance index (PDI); (($N_1/(N_1 + N_2)$)-0.5) (Table 10). 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 (Isaac et al. 2003). 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.

Having identified the relationship between body length and clade richness we then expanded these ideas into a multi-trait analysis (e.g. (Phillimore et al. 2006)) in order to investigate possible interactions between size and other potential key innovations. This was done using the PGLS framework (Rohlf 2001) under the assumption of Brownian covariance structure among tips, implemented using code from the packages *ape* (Paradis et al. 2004) and *nlme* (Pinheiro et al. 2014). As a response variable we modeled clade diversification as the log of the Yule diversification rate potentially associated with each tip (Phillimore et al. 2006). Yule rates were calculated as the log species richness of a clade, corrected as described above, divided by the clade's stem age (Yule 1925; Nee 2006)). Logging the diversification rate in this way limits our interpretation of these analyses, however was essential for limiting the weight given to a few recently diverged and extremely diverse clades, and for maintaining a residual structure within the implemented models consistent with the assumptions of the gls testing framework.

The following traits were implemented as blocks of predictor variables, with different combinations being examined and compared using an AICc framework (Table 11): whether or not a clade displays complete metamorphosis (i.e. whether or not it

belongs to Holometabola (Rainford et al. 2014); implemented as a binary variable), mean of logs body size, dietary substrate- implemented as the presence or absence of particular diets within the known ecology of a clade (taxa within unknown ecologies are coded as present for all categories) , and the interaction of diet and body size. For the coding of diets and interactions fungivory was excluded and treated as part of the overall model intercept. To account for differences in coverage between traits topology was standardized on the pruned tree used above, and the diets of modified groups updated to include all coding states present within any subgroup (Chapter 4, Appendix 7.2). We performed analyses using three different dietary coding schemes described in Chapter 4, reflecting different treatments of various life stages within Hexpaoda: Larval Raw, Larval Modified and Adult (Table 11). The definitions underlying these datasets and the coding of diet are given in Appendix 7.2.

To explore the processes responsible for generating the observed size distribution we used a model testing framework; fitContinuous, in the package *geiger* (Harmon et al. 2008; Pennell et al. 2014). Candidate models fitted were: a simple BM process; the early burst model (EB/ACDC), (Blomberg et al. 2003; Harmon et al. 2010) where rates of evolution through time exponentially increase or decrease; the delta model (Pagel 1999), which scales the phylogeny so as to bias the distribution of rates of trait evolution towards either the root or tips; the SSP model (single stationary peak; modeled as an Ornstein-Uhlenbeck process) (Butler & King 2004), which assumes that trait evolution convergences on a single global optimum value (Table 13). All of these models are all 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 (Pagel 1999), 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 13). All fitted models incorporated estimates of standard error around the mean-of-logs, using the methodology of Ives et al. (2007) (see above for how these are

calculated). Model selection was performed on the basis of AICc values and Akaike weights, see discussion in Harmon et al. (2010).

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 (Rabosky 2014). As implemented here, the analysis fits EB/ACDC models of size evolution to nodes within the tree signifying regime changes among descendent 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 *BAMMtool*s (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.

5.4. Results

Distributions of the observed values of mean-of-logs, log maximum and log minimum body size for terminal taxa are shown in Figure 26. 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 the most diverse clade (more than 75% of all extant hexapods) (Grimaldi & Engel 2005) and accounting 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 (including Orthoptera grasshoppers and crickets) and Phasmatodea (stick-insects)); 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).

This pattern is reinforced on the phylogenetic ancestral reconstruction plots for the group (Figure 27) 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 and lice; 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.

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 (Pagel 1997) 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, Neuropterida (lacewings and relatives), Psocodea and Odonata while

the largest values occur in Coleoptera and advanced Lepidoptera (Figure 27), with the single largest value occurring in the morphologically diverse (4-39 mm) but species poor Lepidoptera family *Aididae* (6 species)

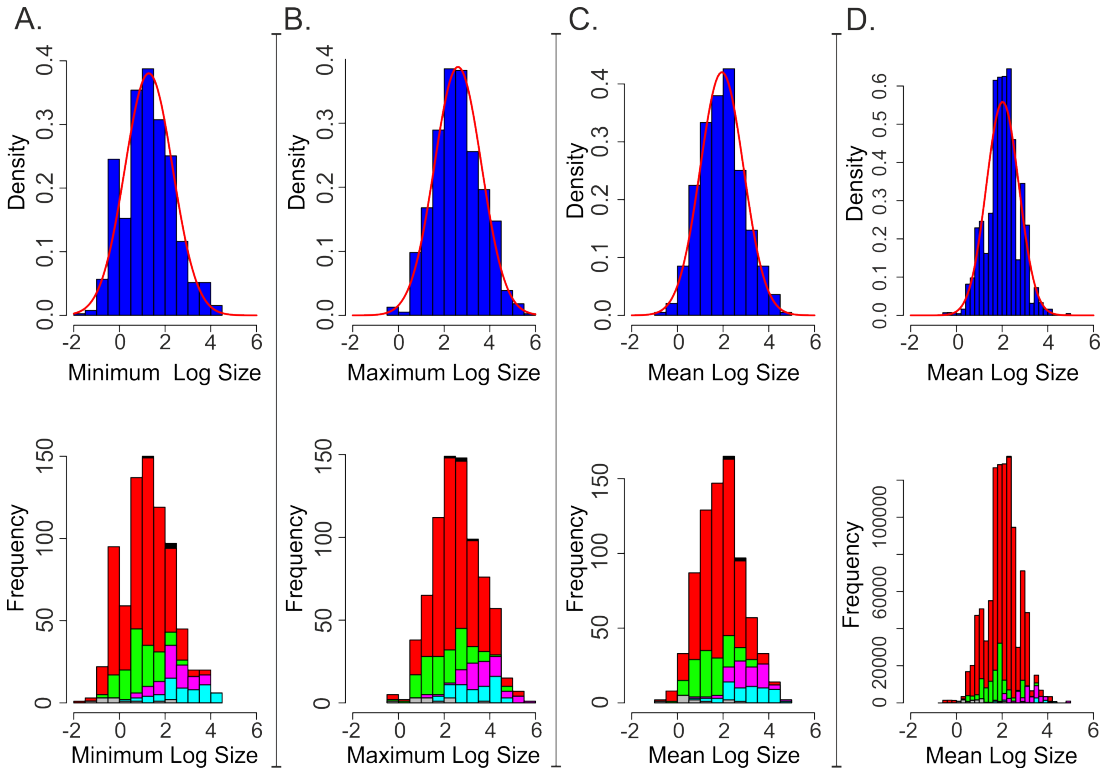


Figure 26 Histograms of A) Minimum log body size (ln(mm), Skewness = 0.3333) B) Maximum log body size (ln(mm), Skewness = 0.07517) C) Calculated mean log body size; for terminal groups used in this analysis (ln(mm), Skewness = 0.2102), D) Mean size with each terminal group represented proportionally to its richness (ln(mm), Skewness =-0.0285) . Curves on upper panels reflect normal distributions with the same mean and standard deviation as the observed data. Colors in lower panels show breakdown of size classes by major taxonomic group; Red - Holometabola, Green - Paraneoptera, Magenta - Polyneoptera, Cyan - Palaeoptera, Black - Basal insects / "Thysanura", Grey - Entognatha.

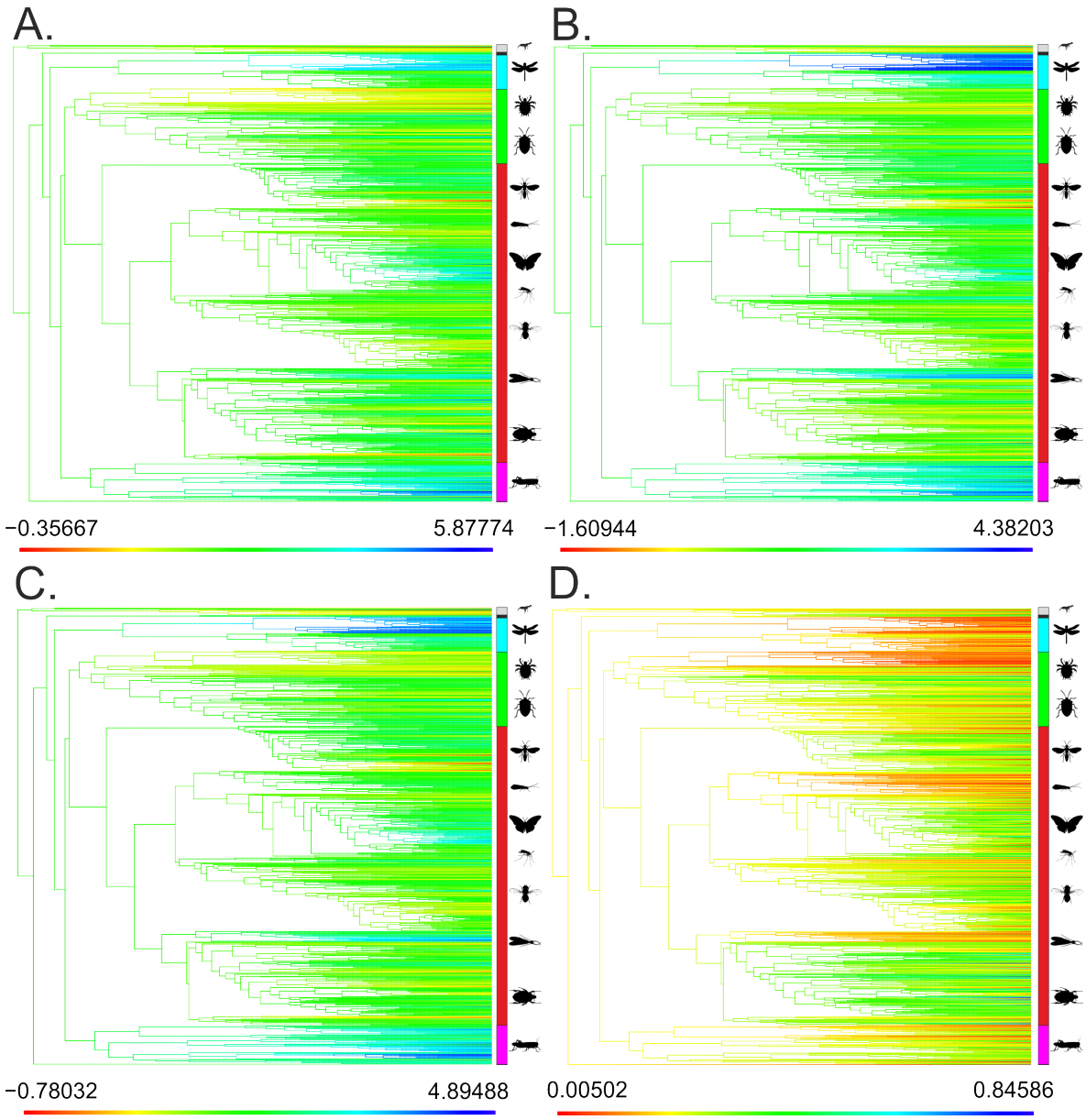


Figure 27 Phylogenetic plot of (log) size traits. A) (log) Maximum body Length; B) (log) Minimum body Length; C) log-mean body length; D) Estimated Standard Deviation. Ancestral reconstruction of internal nodes based on a BM process (ancML) (Revel 2013). Bars denote the minimum and maximum values of observed traits. Coloration on a red to blue scale; minimum and maximum values denoted by the internal bars. Terminal bars denote membership of major clades; colors as Figure 26.

Table 8: Tests of phylogenetic signal in body size within major hexapod clades (with observed standard error). Statistical tests given relative to 1000 tip randomisations.

Taxa	Blomberg's K	Sigma^2 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*

Evidence of phylogenetic signal was recovered in both the full dataset and in all the major sub-clades (Table 8), 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 BM process, consistent with related species resembling one another less than under the expected BM distribution. 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.

The standardized contrasts in body size and RRD across major clades are plotted in Figure 28 with the estimated relationship through the origin calculated on the observed mean-of-log sizes and confidence intervals based on the parametric bootstrap samples as drawn from the estimated terminal distributions for both observed (coloured) and randomized (black) data (parameter values in Table 9). Overall, the data for Hexapoda supports 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 (Table 10) no significant relationships were observed between mean size and richness, rendering further parametric bootstrapping redundant.

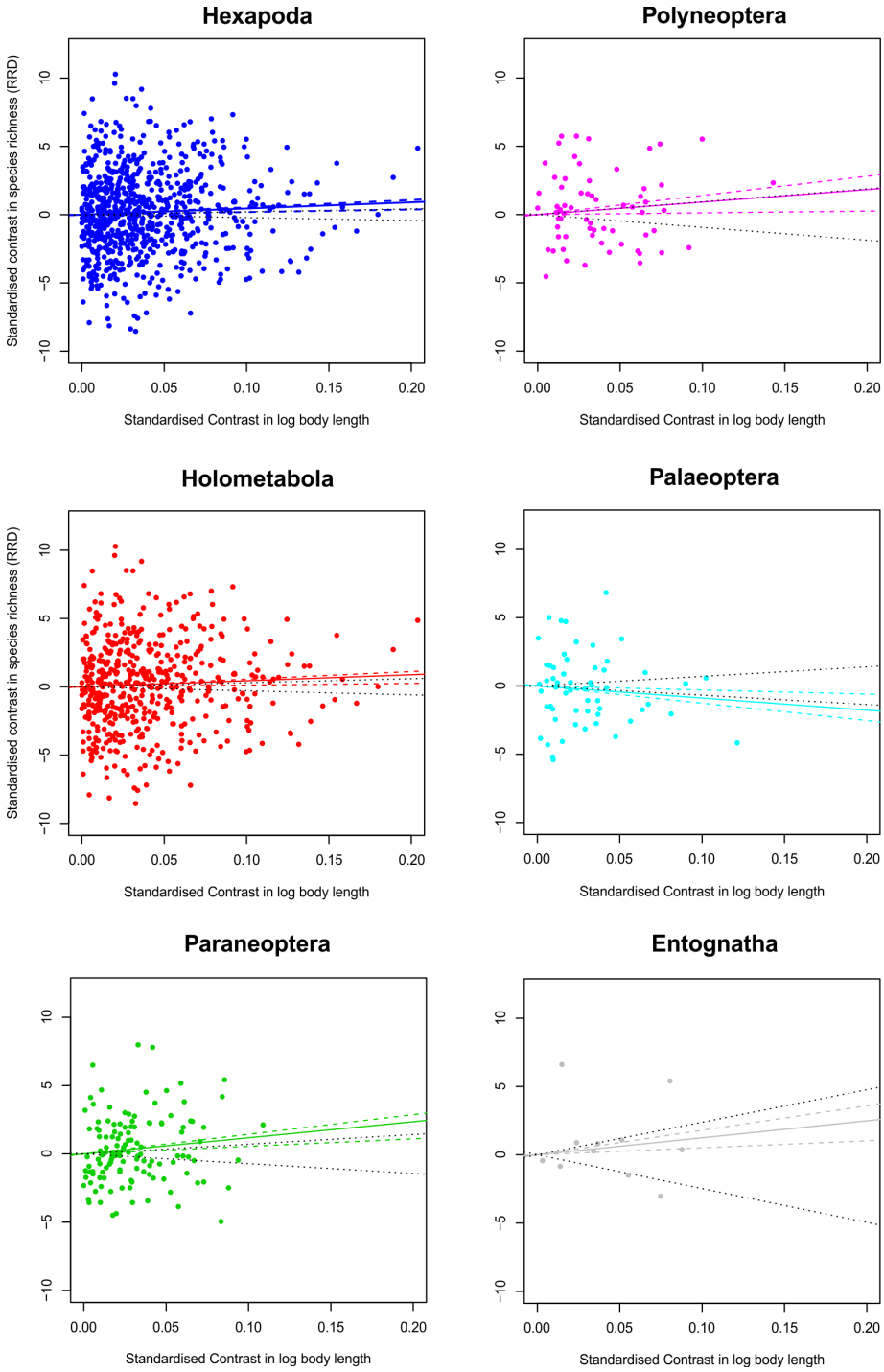


Figure 28 Plots of Standardized contrasts for richness (RDD) and body length (ln(mm)). Solid lines denote the relationship inferred from the mean values in Macrocaic. Dashed colored lines are the 95% CI based on 10,000 parametric bootstraps taking into account the variance present among terminal groups. Dotted black lines denote the equivalent null intervals calculated on the basis of tip randomization. Relevant statistical information for the observed relationships is presented in Table 9.

Table 9: Outputs of Macrocaic analysis: phylogenetically independent node contrasts in RRD and mean log size for major clades (see text). Includes results of parametric bootstrap based on the quartile range of Estimate values from 50000 random draws from the estimated tip distributions vs. the NULL expectation when tip size is randomized. See text for discussion.

Taxa	N.	Estimate	(Adjusted) R^2	SE	t	p	Observed Quartile range (50000 parametric bootstraps)		NULL Quartile range (50000 parametric bootstraps)	
							2.5%	97.5%	2.5%	97.5%
Hexapoda	773	4.538	0.004203	2.219	2.045	0.0412*	1.886	5.383	-2.127	2.106
Holometabola	507	4.415	0.003232	2.715	1.626	0.105	1.246	5.580	-2.944	2.969
Non-Holometabola	265	5.416	0.003874	3.801	1.425	0.155	1.927	7.304	-3.159	3.178
Paraneoptera	126	11.759	0.02523	5.696	2.064	0.0411*	5.495	14.35	-7.172	7.079
Polyneoptera	64	9.135	0.009866	7.139	1.28	0.205	1.256	14.02	-9.385	9.407
Palaeoptera	57	-8.866	-0.000210	8.919	-0.994	0.325	-12.63	-2.987	-6.800	6.986
Entognatha	11	12.43	-0.04417	17.00	0.731	0.481	5.118	17.94	-24.74	23.82

Table 10 Outputs of Macrocaic analysis of relationship between PIC of diversification rate (measured as PDI) and mean log size for major clades. See text for discussion

Taxa	N (Contrasts)	Estimate	(Adj) R^2	SE	t	p
Hexapoda	773	0.4589	0.002572	0.2652	1.73	0.084
Holometabola	507	0.5020	0.003065	0.3138	1.6	0.11
Paraneoptera	126	1.231	0.01155	0.783	1.573	0.118
Polyneoptera	64	0.6955	-0.007437	0.9576	0.726	0.47
Palaeoptera	57	-1.524	0.01207	1.170	-1.303	0.198
Entognatha	11	0.9674	-0.07041	1.8400	0.526	0.611

PGLS modeling of the Yule diversification rates for clades within the tree reveals that, as above, mean-of-logs body size is not alone a significant predictor of this measure of clade diversification, although there is evidence for a marginal effect (Table 11; Log Likelihood ratio (LnL) of Phylogeny + Size model vs. Phylogeny only model; 3.221, p=0.0727, Estimate of Size Parameter = 0.06194). With respect to the role of diet in modulating this relationship, different life history stages and coding philosophies are shown to vary in the implied relationships.

For the majority of hexapod groups, the larval/nymphal stage is the primary feeding phase of the lifecycle and the majority of discussion regarding the roles taxa in natural ecosystems focuses on this developemental stage. For the Larval Raw dataset, the favored model based on AIC comparisons includes the presence of dietary categories (LnL Phylogeny +Diet vs. Phylogeny only; 32.65, p= <.0001*) but precludes any interaction between size and larval diet (LnL Phylogeny +Size +Diet +Interaction vs. Phylogeny + Diet; 7.41516 p =0.2842). By contrast the Larval Modified system, where ecto-parasitism is combined with parasitoidism, does not recover a significant impact of diet on log diversification rates, and instead collapse to the null, phylogeny only model (Table 11; LnL Phylogeny +Diet vs. Phylogeny only; 5.176, p= 0.270).

In contrast with the larval stage, data from Adult insects favors models where (adult) body size and diet interact in their prediction of clade richness (Table 11; LnL Phylogeny +Size +Diet +Interaction vs. Phylogeny + Diet; 16.69, p =0.0195*). The predicted diversification rates for taxa processing different dietary categories vary as a consequence of differences in interaction estimates (Figure 29; values in Table 15). Clades containing phytophagous (mean body length; 2.285 ln(mm), sd=0.900 ln(mm)) and non-feeding taxa (mean = 2.343 ln(mm), sd= 0.7691 ln(mm)) are predicted to have very low diversification rates at small body lengths but to show increased diversification at larger body sizes. However in groups containing ecto-parasitic representatives (mean = 1.780 ln(mm), sd= 0.9113 ln(mm)) this pattern is reversed, such that high rates of diversification are exclusive to small-bodied clades. Predatory lineages (mean = 2.307 ln(mm), sd= 1.061 ln(mm)) are also predicted to diversify most rapidly at small body sizes, although the subsequent drop-off is less severe, such that large bodied predators approximately converge on the overall mean process. Likewise detritivory (mean = 1.774 ln(mm), sd= 0.9124 ln(mm)), nectivory (mean = 1.841 ln(mm), sd= 0.7747 ln(mm)), and taxa for whom ecology is unknown (mean = 2.187 ln(mm), sd= 0.8165 ln(mm)) also approximately follow the overall mean process, although in the latter two cases the inferred diversification rate is always below the global average. The inclusion of holometaboly as a binary variable does not significantly increases the predictive power of any of the models used (Table 11).

Table 11 Model comparisons for PGLS of log (Yule Diversification Rate) assuming BM covariance structure among tips. Parameter blocks are described in text. LnLik is calculated under ML in the package nlme. Highlighted models are the favored explanation for each of the dietary reconstructions.

Diet reconstruction	Model	df	AIC	BIC	lnLik
Larval Raw	Phylogeny +Size +Diet +Interaction+ Holometabola	14	1485.902	1551.024	-728.9512
Larval Raw	Phylogeny +Size +Diet +Interaction	13	1483.937	1544.408	-728.9686
Larval Raw	Phylogeny + Diet+ Holometabola	8	1481.3	1518.513	-732.6501
Larval Raw	**Phylogeny + Diet**	**7**	**1479.352**	**1511.913**	**-732.6761**
Larval Raw	Phylogeny + Size * Holometabola	5	1504.683	1527.94	-747.3413
Larval Raw	Phylogeny + Size	3	1500.779	1514.733	-747.3893
Larval Raw	Phylogeny+ Holometabola	3	1503.922	1517.877	-748.9609
Larval Raw	Phylogeny only	2	1502	1511.303	-749.0001
Larval Modified	Phylogeny +Size +Diet +Interaction+ Holometabola	12	1508.861	1564.68	-742.4303
Larval Modified	Phylogeny +Size +Diet +Interaction	11	1506.947	1558.114	-742.4734
Larval Modified	Phylogeny + Diet+ Holometabola	7	1506.733	1539.294	-746.3667
Larval Modified	Phylogeny + Diet	6	1504.824	1532.733	-746.412
Larval Modified	Phylogeny + Size * Holometabola	5	1504.683	1527.94	-747.3413
Larval Modified	Phylogeny + Size	3	1500.779	1514.733	-747.3893
Larval Modified	Phylogeny+ Holometabola	3	1503.922	1517.877	-748.9609
Larval Modified	**Phylogeny only**	**2**	**1502**	**1511.303**	**-749.0001**
Adult	Phylogeny +Size +Diet +Interaction+ Holo	16	1486.481	1560.906	-727.2403
Adult	**Phylogeny +Size +Diet +Interaction**	**15**	**1484.816**	**1554.589**	**-727.4079**
Adult	Phylogeny + Diet+ Holo	9	1489.174	1531.038	-735.5871
Adult	Phylogeny + Diet	8	1487.511	1524.723	-735.7553
Adult	Phylogeny + Size * Holometabola	5	1504.683	1527.94	-747.3413
Adult	Phylogeny + Size	3	1500.779	1514.733	-747.3893
Adult	Phylogeny+ Holometabola	3	1503.922	1517.877	-748.9609
Adult	Phylogeny only	2	1502	1511.303	-749.0001

Table 12 Parameter estimates for the (favored) Phylogeny + Diet model (Larval Raw dataset)

Parameter	Estimate	Std.Error	t-value	p-value
(Intercept)	-3.845	0.2889	-13.33	0
Detritivory	-0.021	0.0561	-0.377	0.7065
Phytophagy	0.1114	0.0655	1.701	0.0893
Predation	0.0806	0.0645	1.248	0.2124
Parasitoidism	-0.0018	0.0768	-0.023	0.9816
Ecto-parasitism	-0.5745	0.1093	-5.255	>0.001*

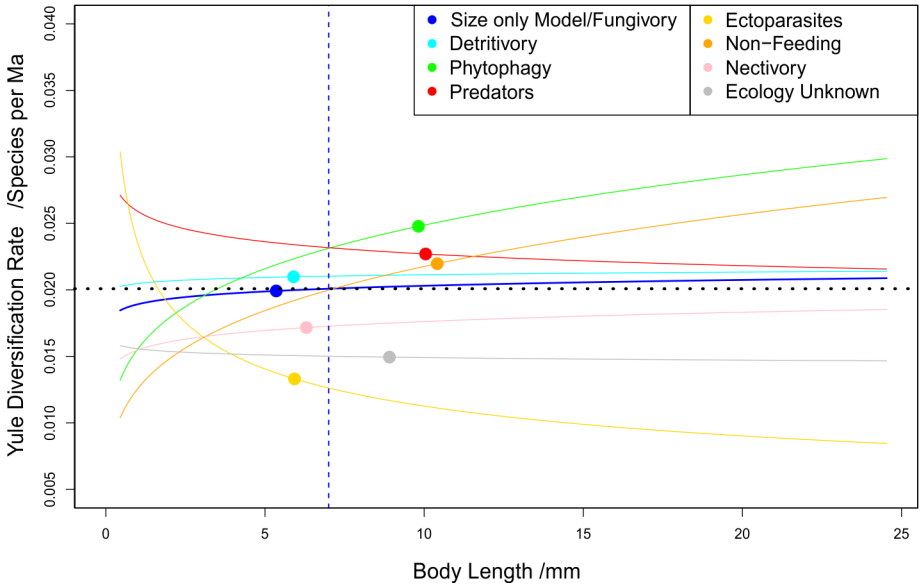


Figure 29 Predicted response curves for the Phylogeny + Size +Diet + Interaction model (Adult dataset) across a 0-25 mm size interval. Data are shown after back transformation from the log values used in the model (parameter estimates for original model provided in Table 15). Colours denoting diet classifications are given in the internal legend. The dashed vertical line gives the overall mean size value for Hexapoda and the horizontal line, the associated intercept diversification rate. Mean body sizes of taxa processing particular diets are denoted by the circles, with the dark blue circle denoting the average size of taxa containing fungivorous lineages (mean = 1.678 ln(mm), sd= 0.8429 ln(mm)). The "ecology unknown" category represents the aggregate process for taxa coded as present for all dietary categories (represented by "?" in Appendix 7.3).

Table 13: Parameter estimates and relative likelihoods of alternative models explaining mean body size for major clades of Hexapoda (including terminal standard error). 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), OU: Ornstein-Uhlenbeck model with central tendency trend (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. See text for further discussion.

Clade	Model	Sigma squared	z0	a/ delta / alpha/ lambda	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	-1e-06*	-779.4	3	1564.7	23.051	0.00001
	delta	0.002196	1.766	1.129	-779.1	3	1564.3	22.627	0.00001
	SSP	0.002666	1.764	0.000591	-778.0	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	-1e-06*	-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	0.001923	-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	-0.000111	-117.0	3	240.3	2.094	0.1382
	delta	0.001559	1.119	0.9031	-117.0	3	240.1	1.9781	0.1465
	SSP	0.001469	1.132	0.00	-117.0	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
Polyneoptera	**BM**	**0.002121**	**2.759**		**-66.26**	**2**	**136.7**	**0.1955**	**0.2922**
	EB	0.002121	2.759	-1e-06*	-66.26	3	138.9	2.3961	0.0972
	delta	0.001389	2.822	2.186	-65.06	3	136.5	**0**	0.3221
	SSP	0.003247	2.812	0.002286	-65.60	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	-0.001169	-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	-40.18	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	-15.71	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

Table 14 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 Table 13.

Clade	Model	Sigma squared	z0	a/delta/alpha	LnLik	k	AICc	Delta AiCc from optimal model	*Akaike weights*
Hymenoptera	**BM**	**0.003168**	**2.091**		**-86.87**	**2**	**177.9**	**0**	**0.4210**
	EB	0.003952	2.105	-0.001230	-86.81	3	179.9	2.043	0.1516
	delta	0.003159	2.090	1.006	-86.87	3	180.1	2.166	0.1425
	SSP	0.003168	2.091	0.000	-86.87	3	180.1	2.167	0.1425
	lambda	0.003168	2.091	1	-86.87	3	180.1	2.167	0.1425
	WN	0.8712	1.784		-104.1	2	212.3	34.39	0.0000
Diptera	BM	0.003120	1.635		-114.8	2	233.7	16.94	0.00014
	EB	0.003121	1.635	-1e-06*	-114.8	3	235.8	19.04	0.00005
	delta	0.001396	1.539	4.392	-106.8	3	219.9	3.117	0.1357
	SSP	0.006735	1.550	0.007896	-106.4	3	219.0	2.195	0.2152
	lambda	**0.001695**	**1.611**	**0.6648**	**-105.3**	**3**	**216.7**	**0**	**0.6449**
	WN	0.3991	1.513		-111.4	2	227.0	10.21	0.0039
Coleoptera	**BM**	**0.002685**	**1.424**		**-153.6**	**2**	**311.3**	**0.5992**	**0.2071**
	EB	0.002686	1.424	-1e-06*	-153.6	3	313.4	2.689	0.0729
	delta	0.002091	1.494	1.656	-152.34	3	310.9	0.1922	0.2538
	SSP	0.003932	1.467	0.002282	-152.3	3	310.7	**0**	0.2794
	lambda	0.00228	1.436	0.8274	-152.7		311.5	0.8054	0.1868
	WN	0.5859	1.625		-162.6	2	329.2	18.47	0.00003
Lepidoptera	**BM**	**0.002756**	**1.368**		**-95.14**	**2**	**194.4**	**1.661**	**0.1996**
	EB	0.002756	1.368	-1e-06*	-95.14	3	196.5	3.778	0.0692
	delta	0.002012	1.488	1.618	-94.46	3	195.2	2.415	0.1369
	SSP	0.003441	1.444	0.001989	-94.47	3	195.2	2.420	0.1365
	lambda	0.002197	1.393	0.88127	-93.26	3	192.7	0	0.4578
	WN	0.5985	2.106		-125.7	2	255.4	62.68	0.0000

Considering the potential processes responsible for generating observed patterns of size evolution, our data suggest that, of our process based models; the majority of hexapod clades favor simple Brownian motion (BM), with the exception of Holometabola, where the favored process is an SSP model with convergence on a single global optimum or elevated diversification at distant tips (Table 13,Table 14). However, when models without an explicit generating process are considered (i.e. lambda and 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 8). Similar patterns are obtained when the major holometabolan orders are examined individually, with Hymenoptera (wasps), Coleoptera (beetles) and Lepidoptera (moths) all favoring BM processes, while Diptera (flies) shows strong evidence for non-phylogenetic signal (and thus favors 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 the BAMM analysis of model heterogeneity further support the idea that the process of size evolution behaves differently in holometabolan and non-holometabolan groups (Figure 30). A single shift in the rate model associated with the origins of Holometabola is recovered with a marginal probability of 0.988, i.e. 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) (Figure 31) 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 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 including (relative frequency 0.17) or excluding (relative frequency 0.18)) the basal family Hydroptilidae.

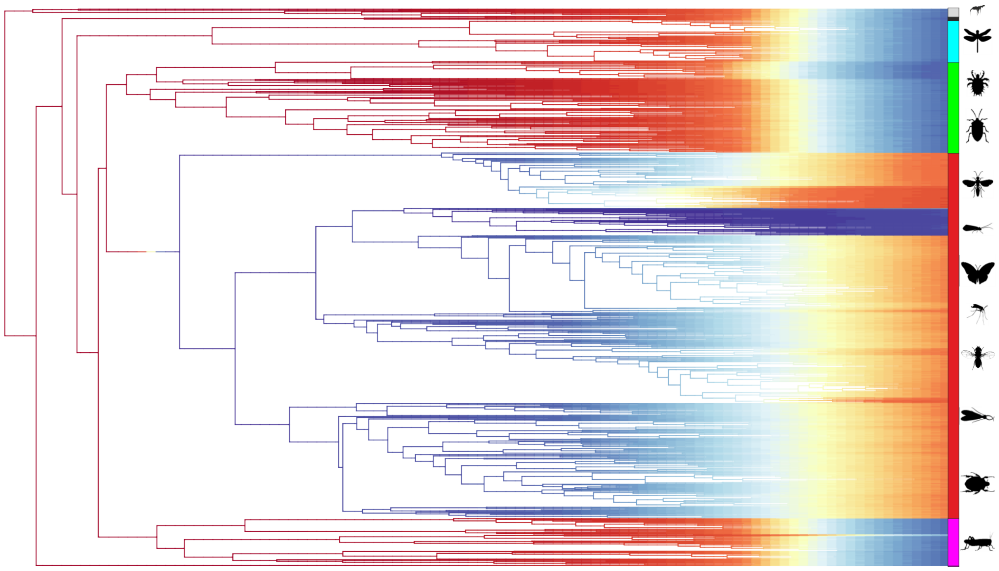


Figure 30 Outputs of Bayesian Analysis of Macroevolutionary Mixtures (BAMM) analysis of log mean body size data. Mean rate of evolution for branches across all post-burnin samples (ln(mm) per million years), denoted by branch coloration (red being high).

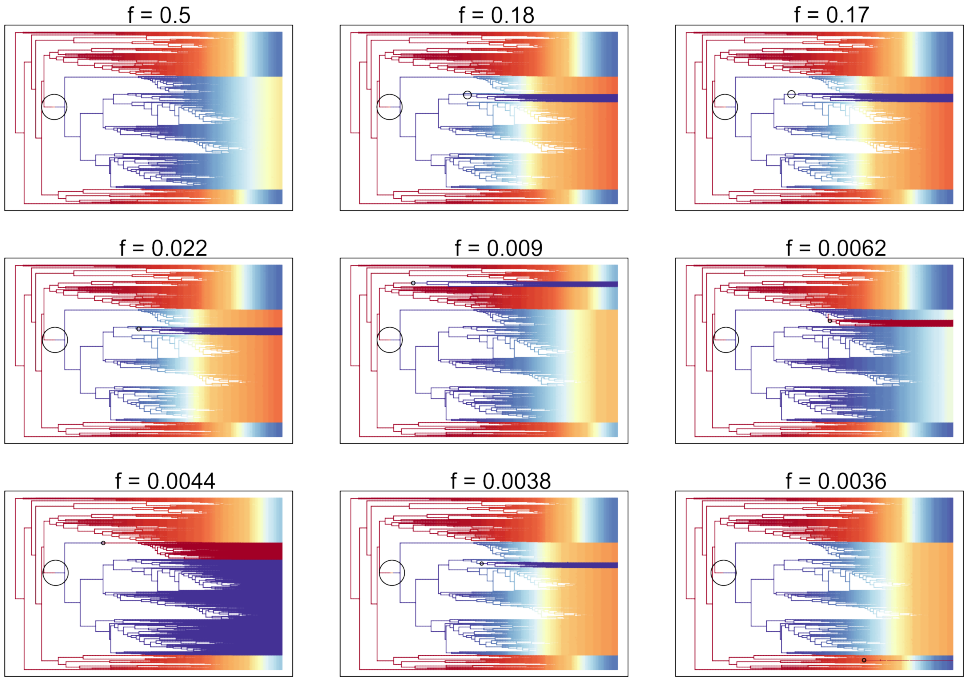


Figure 31 Maximum credible model set from Bayesian Analysis of Macroevolutionary Mixtures (BAMM) corresponding to 95% of the overall model likelihood. Models are listed in order of frequency (f) of obtaining model in the post burnin set corresponding to their inferred probability. Coloration and tree orientation are as Figure 30

5.5. Discussion

The findings of this study corroborate previous work at continental scales e.g. (Finlay et al. 2006; Ulrich 2006; Ulrich 2007) 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 the 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. In addition for adult insects there appears be an interaction between dietary ecology and body size with respect to their effects on diversification; however this interaction does not appear when larval ecologies (which make up the primary feeding stage) are considered independently. 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 (Gardezi & Silva 1999; Kozłowski & Gawelczyk 2002) may be inappropriate when discussing other taxonomic groups (Orme, Isaac, et al. 2002; Orme, Quicke, et al. 2002), and hexapods in particular (Finlay et al. 2006; Chown & Gaston 2010; Gaston & Chown 2013). 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 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 (Vogel 1994; Whitman 2008). Likewise fractal environmental models, which postulate the existence of a higher number of niches at small body sizes (Hutchinson & MacArthur 1959; Morse et al. 1985), 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 therefore subject to local

homogeneity in the composition of their environment across a range of spatial scales (Poulin & Morand 1997; Mouillot et al. 2003; Nyman 2010). 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 (Klingenberg & Spence 1997, see also Blanckenhorn 2000; Sokolovska et al. 2000). 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 (Orme, Quicke, et al. 2002). There is a need to further investigate processes of size evolution across a broader range of invertebrate groups for comparative purposes (e.g. (Nekola et al. 2013)), which when taken together may provide us with new insights into underlying mechanisms controlling the size structuring of natural environments (Woodward et al. 2005).

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 mechanism underlying such constraint is likely to be variable across different lineages. For example the absence of small minimum body sizes within Odonata may be attributed to limitations on the minimum size required for the group's unique flight mechanism (Dudley 2002). In other cases the causes of constraint are much less apparent, 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 proportionally 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 (Rainford et al. 2014). This is particularly striking in that clades previously recovered as downshifted 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 (Ricklefs 2004). Similar ideas have been previously proposed with respect to bird families, (Ricklefs 2004), however formalized testing via multiple regression has been shown to be statistically problematic, due to an inability to distinguish time-dependent and speciation dependent generation of variance (Purvis 2004; Ricklefs 2006). This, in combination with the data abstraction required to treat higher taxonomic groups here (Bokma 2010), and the fact that our approaches to estimate standard deviation are confounded with clade richness (see above; (Hozo et al. 2005)), 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.

PGLS modeling of log Yule diversification rates indicated that, for some coding systems (Larval Raw and Adult), diet is a (marginally) significant predictor of clade diversification. Despite marginal significance in the Larval Raw dataset the predictive power of larval ecology on diversification is inferred to be small. The only coding state to show large loadings is ecto-parasitism (Table 14), a state which is extremely rare in larval hexapods (see Section 4.4), and which is strongly confounded with taxa with unknown ecologies (coded as present in all states, see above). As a result the Larval Modified dataset, where ecto-parasitism is combined with parasitoidism, shows no overall support for a relationship between diet and clade diversity, and analysis favors the null, phylogeny only model (Table 11). Given that the larval/nymphal stage is the primary feeding phase of the hexapod lifecycle this lack of support can be considered consistent with the limited role for diet in structuring clade diversity described in Chapter 4.

By contrast, adult diets are associated with a significant interaction with (adult) body sizes, resulting in different diversification dynamics across the size spectrum for different dietary classes (Figure 29). Given the marginal importance, in terms of overall fitness, of adult feeding in many hexapods (Hill & Pierce 1989), the simplest interpretation of these patterns is to infer optimal size spectrum for taxa utilizing particular resources see discussion in (Lafferty & Kuris 2002)), resulting in differential selection and

diversification of different trophic groups at different sizes (Novotny & Kindlmann 1996). For example, one can envisage that large-bodied ecto-parasites (vertebrate blood-feeders), would be at a competitive disadvantage in their probability of remaining undetected by the host, and in their susceptibility to periods of resource limitation, resulting in few large-bodied and species rich ecto-parasitic clades ((Waage 1979); see discussion in (Poulin & Morand 1997)). On the other hand, large bodied herbivores may gain a mechanical advantage that allows them to overcome structural plant defenses (Andres & Connor 2003; Hanley et al. 2007), and/or may benefit from increased generalism in host-plant use (Wasserman & Mitter 1978; Brändle et al. 2000), which may potentially result in more rapid diversification within large-bodied clades.

Based on this interpretation the expectation is that associations between (adult) body size and diet should be strongest in groups where, relative to the juvenile stage, the adult is long-lived and so makes important contributions to overall lifetime nutrition. This aspect of hexapod ecology is understudied, and not included in the data compiled here, but represents a potentially important macro-evolutionary correlate for incorporation into subsequent analyses of diversification. Both body size and diet are also subject to complex interactions with dispersal ability (Dudley 2002; Jenkins et al. 2007), with the former being potentially responsible for the elevated diversification of large-bodied non-feeding taxa reported here (Figure 29, Table 14). Dispersal ability is complex trait, which has the potential to both accelerate and inhibit clade diversification across the large temporal and spatial scales occupied by hexapod families (Weeks & Claramunt 2014). As with many fundamental hexapod traits, lack of suitable proxies and the sheer scale of the radiation limit the production of multi-trait datasets comparable with those used to explore correlates of clade richness in well-studied vertebrate groups e.g. (Phillimore et al. 2006). As a result a key future priority for hexapod studies is the targeted assembly of disparate ecological descriptions, such that we can make the maximum use of recent improvements in phylogenetic information (Mayhew 2007; Section 6.4.).

In the above models there is no evidence that in inclusion of holometaboly as a categorical trait improves the prediction of log Yule diversification rates, within the limits of our modeling assumptions (Table 11). This implies, no systematic differences between Holometabola and non-holometabolan taxa in terms of their overall diversification that are not accounted for by the other modeled variables (including the covariance associated with the tree structure). Note however, that there are limitations on these models, arising from

the use of (log) pure-birth Yule rate (Yule 1925; Nee 2006) as a proxy for clade diversification, which neglects the role of extinction and turnover in diversification dynamics in terminal clades. Previous birth-death models (Rainford et al. 2014), and fossil evidence (Labandeira 2005; Nicholson et al. 2014), has indicated that, within hexapods, extinction and turnover are key features in understanding clade evolution, and therefore that Yule rates may be poor estimators of the true underlying processes. The use of shift-based frameworks, and method of moments (Magallon & Sanderson 2001) approximations of net diversification in order to incorporate turnover (e.g. MEDUSA (Alfaro et al. 2009)) results in statistical linkage between tip values, which may violate the assumptions of the underlying gls framework, and hence we have not implemented these here. However recent improvements in birth-death models, such the rjMCMC framework in BAMM (Rabosky 2014)) hold promise for incorporating these ideas into future studies of the impact of particular traits on clade diversification (Weber & Agrawal 2014; Setion 6.3).

Turning now to 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 13) lambda model is not in itself a process description, hence why this parameter is most commonly described as a test of phylogenetic signal, e.g. (Pagel 1999). However despite this, 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 (Table 14), 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 (Regier et al. 2013)), 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; mean=1.30mm), Corethrellidae (mean=1.22mm) and Phoridae (mean=1.75mm), which collectively may further skew size distributions across the order (Marshall 2012). 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 (McAlpine et al. 1981; McAlpine et al. 1987; Brown et al. 2009)) as proxies for global size distributions, and phylogenetic uncertainty in relationships, e.g. within Schizophora (Wiegmann et al. 2011; Caravas & Friedrich 2013; Rainford et al. 2014), mean that we should be cautious of over-interpreting these patterns and await better comparative information, preferably incorporating developmental and larval data (Chown & Gaston 2010).

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 (Rainford et al. 2014). Plausible mechanisms a for 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 (Yang 2001; Chown & Gaston 2010), and historical factors relating to the differential extinction of large bodied non-holometabolan groups (Monroe & Bokma 2013; Nicholson et al. 2014). There have been various suggestions, based on the small size of early fossil representatives (Nel et al. 2013), 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 (Hone & Benton 2005). However, the lack of joint systematic framework for extant and fossil taxa has restricted formal testing of this assertion in recent fossil compilations (e.g. (Clapham & Karr 2012)).

Unlike well-studied vertebrate clades, there is currently no universal reference source for comparative data within Hexapoda, with the result that the information used here is derived from a mix of global and regional scale datasets collected at the level of individual clades (Appendix 7.3). This imposes additional assumptions beyond the selection of phylogenetic framework (see discussion of tree in (Rainford et al. 2014)) and the use of described species as proxies for total clade richness (Costello et al. 2012). 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 (Chown & Gaston 2010), 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 (Finlay et al. 2006) and that by combining multiple regional sets we at least attempt to consolidate our size ranges across the known taxonomic range.

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 (Gaston 1991; Blackburn & Gaston 1994; Gaston & Blackburn 1994) or, to put this another way, that the estimates of described species richness for terminal groups are unbiased approximations of their true extant diversity (Costello et al. 2012). The problem of acquiring estimates of "true" species richness based on available, and often incomplete records of described species is one of the most profound challenges facing work on any diverse clade (see discussion in (Costello et al. 2012; Poulin 2014) and references therein). Of the work conducted here, the observed pattern, i.e. a weak and statistically non-significant 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 (Mayhew 2007).

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

(Hozo et al. 2005)), whereas in reality such estimates also encompass other sources of error such as length variation among individual specimens (Gouws et al. 2011) and sexual dimorphism (Cohen et al. 2005) which may contribute to variation observed across lineages. Dealing with within tip variance in trait measurements is perhaps the greatest outstanding challenge in modeling of trait evolution at deep phylogenetic levels (Revell & Reynolds 2012). The methods used here, based on Ives et al. (2007) and Felsenstein (2008), 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" (Slater et al. 2012), 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 (Revell & Reynolds 2012), 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 trait evolution may be modeled across very large taxonomic groups.

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 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 process 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. (Harrison et al. 2010; Clapham & Karr 2012)) and the evolution of vertebrate predators (notably birds) (Dorrington 2012). 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 (Clapham & Karr 2012), the inability of analyses based on extant data to take account of no-longer existing diversity (Finarelli & Goswami 2013) and impacts of phylogenetic non-independence, which are often neglected in fossil analyses of hexapods (Grimaldi & Engel 2005).

The consequences of these findings for the standard size paradigm (e.g. (Kozłowski & Gawelczyk 2002)), with its emphasis on vertebrates, in which size and richness show a strong degree of coupling (Maurer 1998; Gardezi & Silva 1999), are significant in that they attack the universality of these findings to other terrestrial clades (Orme, Quicke, et al. 2002). 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 understating of the processes of size evolution and its consequences for an invertebrate perspective of the natural world.

Table 15 Parameter estimates for the (favored) Phylogeny +Size +Diet +Interaction model (Adult dataset)

Parameter	Value	Std.Error	t-value	p-value
(Intercept)	-3.968	0.3000	-13.22	>0.001*
Mean-of-logs body size	0.0308	0.0528	0.5835	0.5597
Detritivory	0.0792	0.1429	0.5541	0.5796
Phytophagy	-0.1961	0.2118	-0.9253	0.3551
Predators	0.3142	0.1589	1.977	0.0484
Ecto-parasites	0.2178	0.2467	0.8826	0.3777
Non-feeding	-0.4091	0.2233	-1.832	0.0673
Nectivory	-0.2006	0.1507	-1.331	0.1835
Interaction (detritvory)	-0.0171	0.0719	-0.237	0.8126
Interaction (phytophagy)	0.1731	0.0921	1.881	0.0604
Interaction (predators)	-0.0881	0.0772	-1.140	0.2546
Interaction (ecto-parasites)	-0.3507	0.1364	-2.571	0.0103*
Interaction (non-feeding)	0.2076	0.0876	2.371	0.018*
Interaction (nectivory)	0.0254	0.0638	0.3981	0.6907

6. General Conclusions and Future Prospects

Comprising more than half of all described extant species, the diversity of the Hexapoda (insects and their six-legged relatives such as springtails) represents a defining feature of most modern terrestrial communities (May 1988; Gaston 1991b). Within this vast diversity, species richness among the various sub-clades shows striking variation with the extant species richness with different orders varying across four orders of magnitude (Grimaldi & Engel 2005; Mayhew 2007). The overall aims of this thesis were the exploration of diversity dynamics within the Hexapoda, and in particular to test long standing ideas regarding the roles played by key morphological innovations (Chapter 3), ecological determinism and co-evolution (Chapter 4) and body size evolution (Chapter 5) in the structuring of richness within the group (Mayhew 2007). This was done within the context of an explicit dated phylogenetic hypothesis for hexapods generated from publically available molecular sequence data, and incorporating constraints from previous systematic analyses and novel calibration fossils (Chapter 2). In this final section I review the major findings identified here, and attempt to place these within the context of our overall knowledge regarding the hexapod radiation, as well as to explore generalities arising from across the different analyses and look towards how future work may continue to enhance our understanding of the origins and evolution of this vast and vitally important clade.

6.1. Constructing and dating a molecular phylogeny for Hexapoda

Chapter 2 described the process and data used in the construction of the dated phylogenetic framework for Hexapoda used as the basis for all subsequent analyses in this thesis. As noted in Section 2.5, the aim of this work was primarily to make use of the growing topological consensus of molecular and morphological datasets (Beutel et al. 2011; Trautwein et al. 2012; Yeates et al. 2012; Misof et al. 2014) and the publication of novel datasets for markers within major ordinal groups, e.g. (Hunt et al. 2007; Wiegmann et al. 2011; Cho et al. 2011), to provide a comprehensive basis for inferring patterns of diversification within the group. Compared with other major groups for whom broad scale phylogenies have been constructed for the purposes of modeling diversity; e.g. vertebrates

(Alfaro et al. 2009), angiosperms (Smith et al. 2011; Fiz-Palacios et al. 2011), mammals (Yu et al. 2012), birds (Jetz et al. 2012), snakes (Pyron & Burbrink 2012) and ray-fined fish (Near et al. 2013), attempts to resolve Hexapoda must deal with; a) an extreme degree of species richness, limiting any feasible phylogenetic scheme to higher taxonomic levels (thus entailing issues of non-monophyly and conflicting systematic definitions among terminal groups) and b) a lack of universally sequenced markers comparable with, for example, the rbcL marker used in plant surveys (Savolainen et al. 2000; Fiz-Palacios et al. 2011). This, combined with lack of resources for novel-sequencing work to fill gaps in the dataset, resulted the need for some non-standard approaches, outlined in Chapter 2 in order to maximize the information content of available datasets. As a consequence, our analysis entailed a necessary tradeoff between comprehensive reconstructions of all major hexapod groups vs. ambiguity with respect to placements in regions with low phylogenetic signal.

In practice these intrinsic limitations of the dataset means that there is conflict between our recovered results and those of other recent analyses, in particular with respect to more focused studies targeting individual orders, e.g. (Wiegmann et al. 2011; Cho et al. 2011; Regier et al. 2013). The latter tend to have greater consistency in terms of sequence sampling among tips, may be able to make use of more markers, and often have defined strategies for dealing with missing data (e.g. novel sequencing), which may account for their improved coverage. In comparing such studies to that conducted here it must first be recognized that much of the apparent conflict is restricted to regions that receive low support in all analyses, i.e. areas in which the phylogeny remains genuinely ambiguous, such as the backbone of Coleoptera, and diversification within Schizophora (see discussion in Section 1.4, 2.3.2). There are also questions regarding taxonomic overlap, and the monophyly of terminal taxa that could not be addressed here due to restrictions on the resolution of fossil and richness data used in subsequent analyses (Sections 2.4 and 3.2). In terms of hard conflict with well supported nodes, of which there are very few examples in our dataset, we attribute these to greater context provided by placing sequences in the context of the full hexapod radiation (e.g. in terms of alignment) as well as subtle difference in the sequences used, for example by restricting the dataset to a single representative for each family lineage.

One recent study that bears particular comparison with the work presented here is the recently published genome analysis of (Misof et al. 2014), an early output of the “i5k” invertebrate genome sequencing project (Robinson et al. 2011), which represents, for the

first time, a comprehensive effort to use full genome sequencing to resolve many of the outstanding controversies in hexapod relations. This work, the publication of which postdates the analysis presented here, shows a view of hexapod ordinal relationships that is, for the most part, congruent with that shown in Chapter 2, indicating at least some level of validity to the multi-gene public-data approach adopted in this thesis. Areas of conflict include the monophyly of Entognatha; recovered here and disputed by (Misof et al. 2014) with Diplura recovered as sister to insects (>90 bootstrap support- BS), the placement of Zoraptera within Polyneoptera; which we constrained as sister to Dictyoptera following (Ishiwata et al. 2011), while (Misof et al. 2014) weakly recover as sister to Dermaptera (BS <75) (see also (Terry & Whiting 2005)), the exclusion of Siphonaptera from within Mecoptera as opposed to a sister relationship with Boreidae recovered here (see also (Peters et al. 2014)), and, perhaps most strikingly, the failure of these authors to recover a monophyletic Paraneoptera (which these authors term "Acercaria"); with Psocodea instead recovered as sister to Holometabola (as in (Ishiwata et al. 2011))(BS >98).

In terms of divergence times the two studies are also similar although the confidence intervals reported here are larger reflecting differences in calibration philosophy. The dates generated in (Misof et al. 2014) are based on 37 fossil calibration points, which show some overlap with those used here and are implemented as log-normal priors ($\mu = 2$; $d = 0.5$) with an arbitrary root height range of range of 411.5-580 Ma (the authors describe uniform distributions as used here to be "too conservative"). Example nodes for comparison include; the origin of Hexapoda (our analysis: 478 Ma [Confidence Interval; CI; 440-503 Ma], (Misof et al. 2014): 479 Ma [CI; 452-509 Ma]), radiation of true insects (our analysis: 462 [CI 419-498Ma], (Misof et al. 2014): ~441Ma [CI 421-465 Ma]), diversification of Polyneoptera (our analysis 401 Ma [357-438 Ma] , (Misof et al. 2014): ~302 Ma [CI 231- 377 Ma]) and diversification of Holometabola (our analysis: 390 Ma [350-436 Ma], (Misof et al. 2014): 344 Ma [317-372 Ma]), although note the difference in sister group described above).

In assessing the significance of the conflict between these two datasets, it is first important to appreciate the vast differences in scope and scale between these analyses. The focus of this study was to provide a shallow (8 genes, 7kb length of sequence) but comprehensive (874 tips) analysis for the entirety of Hexapoda, as our interest lay in understanding the processes of diversification across the group at a relatively fine taxonomic level. By contrast the (Misof et al. 2014) analysis can be thought of as a very

deep and detailed sampling of sequence information (total sequencing of 2.5 giga-bases, analysis of multiple super matrices of up to 201,000 amino acid sites) for a restricted range of taxa (103 tips) emphasizing ordinal relationships. These differences in emphasis, as well as the vast difference in available resources (notably computation time and genomic expertise), are likely to account for most of the differences between the results. It should also be noted that genomic data, due to its sheer size, contains the potential for severe internal conflict and is susceptible to a number of hard-to-characterize biases, for example in terms of in terms of nucleotide composition and potential paralogy at included loci (Phillips et al. 2004; Jeffroy et al. 2006; Leigh et al. 2011). A great deal of effort has been spent by these authors to deal with some of these issues; however, as they note, there remains a number of unanswered questions within their phylogeny, notably relating to Palaeoptera, the position of Zoraptera and the validity of a paraphyletic Paraneoptera (Misof et al. 2014).

Given these advances in data availability it is important to ask what are the consequences for the validity of our work and its subsequent utility for inferring macroevolutionary hypotheses. As noted, the tree presented in Chapter 2 is the most comprehensive dated tree currently available for Hexapoda and one of very few that attempts to treat multiple orders at a fine (e.g. family) taxonomic level. The value of this approach is that provides consistency when comparing many groups, particularly with respect to divergence times, such as has not been possible in former studies using distributed trees e.g. (Connor & Taverner 1997). The value of such comprehensive trees, even where there are problems with the placements of some of the tips is nicely illustrated by the large, and growing, body of work based on the mammalian super-tree of (Bininda-Emonds et al. 2007), e.g. (Stadler 2011a; Price et al. 2012), despite some concerns over its representation of certain tip groups (Meredith et al. 2011). Given the sheer scale of the hexapod radiation it is unlikely that we will ever attain the species level coverage of the former analysis (which as noted below, requires some rethinking in terms of how we model diversity controlling processes). However having at least a preliminary framework for conducting analyses is likely to be valuable in its own right as a tool for exploring the group.

How long the tree described here remains a useable tool is in many ways dependent on the future of phylogenetic surveying among hexapod clades. The ongoing “i5k” project is ultimately aimed at sequencing 5,000 arthropod genomes, i.e. many more than the

number of tips on the tree presented here, however much of the proposed sampling is driven by concerns other than phylogenetic inference (most notably pest management (Robinson et al. 2011; Consortium 2013)) with the result that currently planned sampling will be less comprehensive in terms of major clades covered than that given here. In any case it seems unlikely that this vast array of data will be fully integrated into a comprehensive dating framework for the group as a whole any time in the near future, and until then our tree (or perhaps a similar one that makes use of the topology of (Misof et al. 2014) as a direct source of topological constraints) should serve as a useful placeholder and to illustrate the potential gains in terms of macro-evolutionary understanding that can potentially be achieved by taking a broad view of the hexapod phylogeny. Given the overall similarity of our tree with that produced by these novel analyses it is unclear what, if any, effect these improvements will have on the conclusions presented here.

6.2. Patterns of diversification rate shifts in Hexapoda

The main motivation behind this work was a reexamination of key innovation hypotheses in the diversification of Hexapoda at a level of phylogenetic resolution beyond that available in previous surveys of the group (Mayhew 2002; Mayhew 2003; Mayhew 2007; Davis et al. 2010) (Reviewed in Section 1.2.). The hope was that by combining enhanced resolution, as well as explicit and direct dating of the node ages (which would not have been possible under previous methodologies (Nee 2006)) we could further our understanding of the radiation and assess the overall significance of proposed key innovations. (Davis et al. 2010a), postulated a major shift in diversification process associated with the development of wing–folding (i.e. at the origin of Neoptera) with a minor effect identified associated with the development of flight, based on a super-tree compilation of available trees to 2007 (Bininda-Emonds et al. 2004; Davis et al. 2010), with diversification modeled using the trickle-down process of (Davies et al. 2004). By contrast our analysis, based on the MEDUSA algorithm strongly indicates the presence of shift associated with Holometabola (i.e. the origin of complete metamorphosis) and provides weak evidence for a shift associated with the origin of flight. Given the numerous differences in terms of topology, dating and taxonomic richness between these two studies such differences in apparent pattern are perhaps unsurprising and we can demonstrate on the basis of likelihood ratio tests that these previous hypotheses represent inferior

descriptions of the assembled data when compared with complete metamorphosis (Section 3.4.).

The improved phylogenetic resolution available to this dataset allows us to further assess the significance of other proposed shifts in diversification that may have played a role in structuring extant patterns of diversity. For example (Davis et al. 2010a) identified a down-shift associated with the clade Neuropterida relative to its sister lineage; however at the time the significance of this even was unclear due to the uncertainties around tree construction. Our recovery of this event, within the context of the Holometabolan radiation, is thus of significance as it underscores the possibility of divergent processes ongoing within the group. What these processes may be is hinted at by combining this apparent event with other similar downshifts associated with unrelated groups such as Mecoptera, Ephemeroptera and basal members of Coleoptera and Lepidoptera. A striking feature of a number of these groups is that their fossil records suggest that, in terms of number of families, these are groups which have declined in richness from a peak in the late Mesozoic (Labandeira & Sepkoski Jr 1993; Nicholson 2012; Nicholson et al. 2014), and there are a number of lineages that appear to have restricted extant distributions relative to recent fossils (e.g. Raphidoptera or Nannochoristidae (Grimaldi & Engel 2005)). Both of these features are compatible with reduced diversification due to extinction (resulting in "relic taxa"), that may be indicative of the failure of these clades to adapt to the massive ecological changes brought about by the angiosperm radiation during the late Cretaceous (Ross et al. 2000; Grimaldi & Engel 2005). It should however, be noted that attributing particular up or downshifts with respect to overall diversification to either speciation or extinction processes is challenging using the methods described here (Nee et al. 1994; Rabosky 2010) and may be contingent on the availability of suitably resolved fossil data (McInnes et al. 2011), although see (Stadler & Bokma 2012) for potential recent developments in this regard.

In addition to the apparent relictual taxa, our analysis also identifies a set of recently derived up-shifted groups many of which are familiar and important components of the recent fauna such as Acrididae, Chrysomeloidea, Apidae and Calyptratae (in particular Tachinidae). We also find evidence, once uncertainties in node ages are taken into account for nested shift dynamics within Lepidoptera, particularly with reference to Obtectomera which may reflect repeated Cretaceous radiations in this predominantly phytophagous lineage (Regier et al. 2013). Interestingly, with the exception of a shift

associated with Calyptratae, our work fails to recover the nested or "episodic" shift pattern previously reported within Diptera (Wiegmann et al. 2011), a difference which we attribute to: differences in the inferred relationships within this group (Caravas & Friedrich 2013), improvements in the dating procedures, and the greater context provided by nesting Diptera within the complete Holometabolan radiation. Once again these differences serve to illustrate the benefits gained by adopting a universal approach to considering hexapod diversification, as oppose to fixating on single clades in the identification of large scale patterns and possible general models involved (Mayhew 2007).

The methods used to model diversification in this study are subject to a number of limitations, some of which have been partially alleviated by work subsequent to this analysis. Firstly there are a number of outstanding questions regarding the use of the MEDUSA algorithm (Alfaro et al. 2009; Brown et al. 2012) that provides the primarily basis for the results described above. These criticisms include the fact that the form of algorithm implemented in this study is "greedy" in the sense that at each time step it is only possible to add further shifts to the joint model, even if the overall likelihood would be improved by the removal of previously inferred events (this feature has subsequently been added to more recent versions (Pennell et al. 2014)). In addition there are also questions regarding the suitability of the underlying method of moments estimator for net diversification rate (Magallon & Sanderson 2001) particularly in the context where rates of diversification might be expect to change through time (Rabosky 2009; Wiens 2011). Currently there are relatively few alternative methodologies that are able to deal with appropriate modeling of terminally incomplete clades. Several of these, including the BiSSE algorithm (Maddison et al. 2007) and its various derivatives e.g. (FitzJohn et al. 2009; FitzJohn 2010; Magnuson-Ford & Otto 2012), rely implicitly on simulation of the species resolved tree as the basis for parameter estimates (Stadler 2011b; Pyron & Burbrink 2013), or highly parameterized and computationally complex transition matrices (FitzJohn et al. 2009) both of which is computationally unfeasible on the scale of Hexapoda due to exponential scaling of time and memeory requirements on the number of species represented by tips.

The recently developed alternative procedure, the BAMM algorithm of (Rabosky 2014), represents the beginnings of a gradual shift within diversification studies into the Bayesian framework (as opposed to the maximum likelihood models described here), that due to its capacity to deal with uncertainty in parameter estimates (represented as

distributions as opposed to single point values) may represent a more natural framework for dealing with the issues surrounding diversity modeling (Bokma 2008; Moore & Donoghue 2009; Silvestro et al. 2011). Preliminary attempts within this project to use this algorithm in its current form resulted in inference of a flat likelihood surface, where the uninformative model prior dominated the position of inferred shifts. At the time of writing there has been insufficient emphasis placed on adapting models of diversification for large unresolved clades which make up the majority of taxonomic diversity, partially due to the availability of species resolved trees for charismatic vertebrate groups and the intrinsic difficulty of scaling current algorithms to deal with large non-model clades. This represents an important future direction for the field and a major area of current development.

While extending available modeling frameworks presents a significant challenge for future studies, an equally great issue arises from the need to acquire accurate proxies for clade richness in groups, such as Hexapoda, where the majority of species remain un-described (Erwin 1982; May 1988; Mayhew 2007). Throughout this thesis we have relied implicitly on estimates of described species richness (which are themselves subject to unknown levels of error relating to the extent of synonymy across groups (Costello et al. 2013)) as proxies for the true richness of different hexapod clades. This assumes, however that the proportion of species described is equal across different clades, which is very unlikely to hold with respect to small bodied, cryptic or parasitic taxa (May 1988; Poulin 2014). Attempts to utilize rates of species description through time to standardize estimates of total diversity have proved to be of mixed utility, with the most well known examples being highly sensitive to the underlying assumptions regarding the rate of increase (Dolphin & Quicke 2001) and some models resulting in statistically undefined estimates of total richness (Bebber et al. 2007; Costello et al. 2012). Alternative approximations such as expert opinion have their own issues of subjectivity and systematic bias ((Appeltans et al. 2012; Poulin 2014) and references therein), although some novel methods, which draw on richness relationships between different levels in the taxonomic hierarchy, hold at least potential promise for enhancing the clade richness estimates used here (Mora et al. 2011).

While acknowledging these potential methodological and data constraints, based on the relative robustness of the findings presented here to the sources of error that we could explicitly examine (e.g. uncertainty in the ages of divergence from the molecular clock study- see section 3.4), it appears likely that the major conclusions of this study; elevated

diversification rate in Holometabola (which contains a large number of under-described lineages), the existence and identity of apparent relictual groups, and the importance of some recently derived ecologically important lineages, will show some level of robustness to improvements in species description, as well as to increased sampling of and redefinition of the terminal lineages used here. Unfortunately the precise results of such data changes cannot be predicted a-priori and will be strongly contingent on precisely which groups are strongly impacted.

6.3. Dietary ecology as a determinate of species richness patterns

Arguably the most influential idea in hexapod diversification is the apparent link between species richness and ecological, and in particular, dietary, zones of adaptive opportunity (Mayhew 2007) (see also (Price et al. 2012)). Of these, by far the most significant is the apparent increase in species richness linked with plant feeding (Mitter et al. 1988; Farrell 1998), as underpinned by both the observation of extreme species richness in certain specialized phytophagous clades (e.g. (Mitter et al. 1988; Farrell 1998; Barraclough et al. 1998; Winkler & Mitter 2008; Nyman 2010)) and a broader theoretical framework linking heterogeneous substrate use with specialization and enhanced clade richness (Ehrlich & Raven 1964; Thompson 2009; Forister et al. 2011; Janz 2011). In Chapter 4, we reexamined this hypothesis, along with related ideas looking at parasitism (Wiegmann et al. 1993), and the role of ecological specialization more generally (Futuyma & Moreno 1988; Forister et al. 2011), in the context of our explicit tree. In addition we also explored patterns of phylogenetic structure and transition rates among dietary types that may have contributed to their respective species richness associations. Our conclusions, that there was no significant association between possession of a particular dietary trait and elevated or depressed richness, nor with combinations of traits that would be expected a-priori to be associated with specialization (due to host parasite co-evolution), are thus novel in that they conflict with established opinion regarding the drivers within Hexapoda. We do however recover evidence that dietary states differ in their degree of phylogenetic conservatism and relative transition rates, specifically evidence of bias transition in favor of more "specialized" dietary states that may reflect key processes or events in the macroevolution of hexapod ecology.

As discussed in Section 4.5, possible sources for heterogeneity within ecological zones include a) guild specific diversification processes (Hardy & Cook 2010; Novotny et al. 2010; Novotny et al. 2012) that are collectively masked by the relatively coarse substrate based descriptors applied here, b) failure of the coding system to adequately denote real zones of ecological opportunity at the clade level (e.g. failure to appropriately consider within tip variation in ecology) (Mitter et al. 1988) or, c) contingent effects of other trait values such as body size or dispersal ability in regulating the effect of a given dietary shift (de Queiroz 2002). Exploring these various possibilities takes us beyond the exploration of single trait axis and into a more integrated perspective on the limits of zones of ecological opportunity, niche space, and the resultant impacts on clade richness (Poisot et al. 2011; Vamosi et al. 2014). Multi-trait models that simultaneously consider multiple impacts on potential diversification are likely to play a major role in unraveling outstanding patterns within hexapods. The application of such modeling frameworks remains limited by the availability of descriptive data and the resolution of underlying phylogenetic frameworks, prompting the need for further synthesis of disparate data sources, possibly in the form of online data-bases (see discussion below).

From a methodological perspective the simple sister-group methods used here have increasingly been superseded, with respect to small (species resolved) clades, by joint trait-diversification rate models, in particular those belonging to the BiSSE family (Maddison et al. 2007) (see discussion above). As noted above, these models were not available for implementation here due to limitations on computational scaling and memory requirements for the study of large unresolved clades. However, as computation improves and modeling frameworks become more able to deal with trees of higher taxa it appears likely that these will be become increasingly able to address issues on these large phylogenetic scales. A possible way forward, involving the fusion of the Bayesian rjMCMC algorithm BAMM with tree pruning, to explore average rates among taxa showing a particular phenotype, is presented in recent work by (Weber & Agrawal 2014) representing the culmination of the current generation of tools for historical rate based analysis, and providing a model for future explorations of state dependent diversification.

As with other macro-evolutionary questions study into the impacts of trait acquisition (e.g. dietary shifts) on diversification would benefit from a more holistic approach that takes into consideration the context and evolutionary mechanisms by which such shifts may occur and the consequences of such mechanisms for the resultant patterns

of diversity (see for example discussion in (Ricklefs 2004b)). Diet, particularly at the broad scale considered here, is a complex trait, evolving over long time scales, and as such is subject to considerable historical effects that may play a role in shaping the resultant patterns of niche space (Grimaldi 1999; Grimaldi & Engel 2005; Pennington et al. 2006; Futuyma & Agrawal 2009). Implicit in the work conducted here and elsewhere is an underlying model where niche space is partitioned into semi-discrete zones of opportunity that, due to differences in rates of diversification and/or carrying capacity, facilitate differences in species richness. Historical factors, such as the relative availability of food resources or the radiation of novel host groups have the potential to modulate the structure and permeability of boundaries between these ecological zones, which may affect their promotion of diversity (Pennington et al. 2006; Futuyma & Agrawal 2009). Assessing such effects may be beyond the capacity of known hexapod and plant fossil records (Grimaldi & Engel 2005), although see (Labandeira 2006; Labandeira 2013) for discussion of recent progress. However growing taxonomic knowledge and in particular greater integration of fossil groups with the phylogeny of extant taxa (thus providing ecological models), may go a long way towards making discussion of these ideas increasingly feasible in future work.

Identifying and categorizing zones of ecological opportunity, and the signal they leave on phylogenies, remains one of the greatest outstanding challenges of macroecology, as intuitive classifications such as that applied here are inherently subjective (Nyman 2010), and may result in missing biologically relevant principals (Bernays & Graham 1988; Singer & Stireman 2005). Recently there have been moves to utilize the growing body of phylogenetic and biochemical data available for insect herbivores and their plant hosts to begin to formalize the concept of adaptive zones within theses groups (Cavender-Bares et al. 2009; Joy & Crespi 2012) although, as yet, there are too few well defined examples to provide a basis for generalities across all hexapod groups (see discussion in (Futuyma & Agrawal 2009)). With respect to other diets, severe knowledge gaps in our understanding of how ecological guilds are partitioned (Giller 1996), and the roles played by other factors such as natural enemies, limit our ability to describe such lifestyles with respect to adaptive zones resulting in a need for further ecological work targeting such systems, as well as efforts to improve our understanding of the phylogenetic context underlying such radiations.

6.4. Body size and diversity

Body size is one of the most significant controls on how organisms interact with local environments (Hutchinson & MacArthur 1959; Morse et al. 1985; Gaston & Blackburn 2000) and has been studied as a correlate of richness in a wide range of vertebrate groups (Maurer 1998; Gardezi & Silva 1999; Kozłowski & Gawelczyk 2002; Albert & Johnson 2012). However to date there have been relatively few attempts to explore such patterns in invertebrates (Orme, Quicke, et al. 2002; Orme, Isaac, et al. 2002) and hexapods in particular (Chown & Gaston 2010; Gaston & Chown 2013). In Chapter 5 we present a series of analyses aimed at exploring patterns of size bias in clade richness within hexapods, and the processes responsible for generating observed patterns of body size evolution on the log scale within the group. This was conducted within the context of the described hexapod phylogeny and coarse-scaled estimates of the observed range of body sizes observed within terminal groups based on log minimum and maximum body length estimates. From these analyses, the details of which are given in Section 5.4 we concluded that, on the log scale, there is no clear evidence of bias in terms of the length distribution of insects and, after taking account both phylogeny and the uncertainty within tip estimates, there is no global association of body length with species richness either for the group as a whole or any of the major sub-clades. We also concluded that the overall process of size evolution within hexapods closely approximates simple Brownian motion on the log scale, with the exception of Holometabola, a clade for which the observed processes are non-Brownian and consistent with accelerating rates of size evolution through time (see Section 5.5. for discussion).

These results, and in particular the possible future development of this work serve to illustrate two distinct issues applicable to the study of macroevolutionary patterns in continuous trait data for higher taxonomic groups. The first is a data-driven issue relating to the availability of information in non-model taxonomic groups and how we can improve the treatment of such information within the modeling process. There are a number of limitations imposed on a dataset concerning a group as diverse as a the Hexapoda, including: geographic bias in the availability of size data (both in terms of the placement of well studied faunas and in collections that could potential be used to model size distributions) (Finlay et al. 2006), descriptive bias relating to possible systematic underreporting of small sized forms (Gaston 1991a; Blackburn & Gaston 1994; Gaston &

Blackburn 1994) and the presence of potential confounding features such as sexual dimorphism that prevent like-for-like comparison of different taxonomic groups (Chown & Gaston 2010; Gaston & Chown 2013). For modestly sized and charismatic groups, such as mammals and birds, such issues are for the most part mitigated through the availability of global comprehensive databases of trait data e.g. PanTHERIA (Jones et al. 2009), that provide the basis for the study of macro-evolutionary patterns e.g. (Slater 2013). For hexapods, due to their extreme taxonomic richness, the problem is several orders of magnitude harder. However, any standardized approach, even if restricted to particular sub-clades (for example the addition of size and ecological meta-data to entries in active online systematic databases e.g. Polyneoptera Species file project (Eades 2012) or Systema Dipterorum (Pape & Evenhius 2013)) would represent a major advance in terms of our capacity to make strong statements regarding the processes involved, and may, through modeling associated uncertainty, prove the basis for resolving some of the deeper data issues (see (Blackburn & Gaston 1994; Gaston & Blackburn 1994) and discussion above on modeling true lineage richness from rates of species description).

The other major trend that this analysis serves to illustrate is the importance of incorporating variance within tips into the modeling of macro-evolutionary processes for the study of large taxonomic groups (Harmon & Losos 2005; Garamszegi & Møller 2010; Silvestro et al. 2015). Due to the primary focus of comparative methods on highly resolved (ideally species level) trees, historically there has been little emphasis placed on considering the degree to which internal variation within tips, (which with respect to insect body size can be a large proportion of the total variance observed across clades (Chapter 5)) have potential to impact on the results of comparative methods (Felsenstein 1985). Only relatively recently have procedures been developed that explicitly deal with this additional source of variance within standard comparative frameworks, and several of these, e.g. (Ives et al. 2007; Felsenstein 2008) are primarily constructed to deal with error arising from measurement variation among samples, particularly where sample sizes vary, although they can potentially be adapted to other situations.

More relevant to the work conducted here (although as yet lacking good implementations) are the development of Brownian procedures that model species means, intraspecific phenotypic variances, and the parameters of the evolutionary process from their joint posterior probability distribution, such as that recently developed by (Revell & Reynolds 2012). This is an exciting addition to available methodological tools; in that it

provides a bi-directional information flow such that modeling parameters and terminal distributions are potentially jointly informative (i.e. in effect it can compensate to some degree for missing data within the terminal lineage by drawing on the joint likelihood of the overall tree). The referenced publication provides a simple example of this approach, which assumes intraspecific variability is the same for all species (their so called reduced model), and does not fully explore the more complex case where variance is itself variable across tip taxa (corresponding to different ranges in the data discussed here) (Revell & Reynolds 2012). However expanding this, particularly if such procedures can be combined with a more diverse array of trait evolution models (the current form fits only the lambda model), and recent innovations to reduce the inference of such models to linear time (Ho & Ane 2014) then this could lead to a major advance in terms of how we explore trait evolution in large and unresolved groups. A nice feature of Bayesian protocols is their inherent hierarchical nature, such that distinct processes can be combined into a single more complex joint model, thus in principle (if less so in practice) it is possible to envisage a future world in which models have been developed that combine this variance based inference of process with the capacity to explore clade heterogeneity (e.g. using of reversible jump Markov chain Mote-Carlo such as BAMM (Rabosky 2014)) in order to provide a comprehensive approach to exploring patterns of trait evolution in the context of large clades.

One other current, but as yet rarely implemented, method that, in principle, is able to deal with unresolved variance within phylogenetic tips is the approximate Bayesian computation (ABC) method, MECCA, developed by (Slater et al. 2012). ABC is a simulation based technique related to Bayesian modeling where parameter values are sampled from an assumed prior distribution (which must be defined a-priori, see (Templeton 2010)) and then the distribution of model outputs compared with that of real data based on a summary statistic (Joyce & Marjoram 2008), in order to generate a parameter distribution approximating the underlying likelihood surface (Marjoram & Tavaré 2006). Thus this approach removes the, often time consuming, step of calculating likelihood at each time step in an MCMC chain, and so is able to implement models for which there is no known explicit likelihood function (Marjoram & Tavaré 2006). This is a relatively novel mathematical tool in diversification modeling, and as currently implemented is computationally taxing (due to the need to generate many simulations, most of which are discarded). However, such approaches have been heavily implemented

in phylo-geographic studies (although not without criticism, see (Templeton 2009; Templeton 2010)) and may, as trait models become increasingly sophisticated, become an important tool in the study of trait evolution.

The work conducted here focuses exclusively on adult body size as an explanatory variable for clade richness within Hexapoda, which despite the data issues discussed above is one of the easiest quantitative ecological traits to characterize in the context of a data deficient clade. There are however ever numerous other continuous variables that would be great interest to study in the context of insect macroevolution. Some of these, such as reproductive rate, life span and home range size have been dealt with in model vertebrate groups e.g. (Phillimore et al. 2006). However, others notably dispersal capacity (briefly examined in the context of the evolution of flightless taxa in (Mitterboeck 2012)), remain poorly characterized for many taxa despite their clear ecological significance, and potential for co-evolution with the dietary traits and size evolution discussed here. Once again I reiterate the need for standardized, and preferably curated reference sources to collate the work of hundreds of years of ecological observation and facilitate comparative work for the study of the controls of richness in this vast and important clade.

6.5. Final thoughts

Above I have outlined the major findings of this thesis and briefly explored the potential impact of recent findings and methodological developments on the conclusions presented here and the potential for further work. To summarize, prior to this work much of the focus of macroevolutionary explanations for hexapod diversity have been restricted to incomplete phylogenetic settings, that have either lacked resolution to deal with questions on interest (e.g. by being restricted to the ordinal level (Mayhew 2002; Mayhew 2003; Davis et al. 2010)), or have lacked an explicit hypothesis of relationships underpinning chosen richness comparisons (e.g. (Mitter et al. 1988; Connor & Taverner 1997)). By constructing a resolved and dated phylogenetic tree, despite outstanding uncertainties, our goal was to establish a new standard in hexapod diversification studies bringing work on the group one step closer to that conducted for well studied vertebrate lineages (e.g. (Price et al. 2012; Jetz et al. 2012)). Analyses in this thesis highlight the significance of metamorphosis as a key innovation, contrasting with previous views focusing predominantly on wing-folding (Davis et al. 2010a); cast doubt on long standing ideas regarding the role of phytophagy in promoting hexapod diversification (Mitter et al.

1988), and raise questions regarding the relationship between body size evolution and diversity within the group (Mayhew 2007).

In terms of forward direction, patterns of diversification in hexapods over long time scales are likely to be driven by an amalgamation of species-level, clade-level and ecosystem level phenomena that may render simple models of diversification inadequate to explain the full complexity of diverse and ancient clades (Benton 2010). In seeking to further our understanding there is a need to integrate our perspectives across various phylogenetic scales, in combination with other sources of information such as the fossil record (Losos 2011), in order to understand the relative roles of different mechanism in the promotion and maintenance of clade richness. Areas of priority for future work include:

- Resolving the outstanding phylogenetic uncertainty among the major extant lineages (see (Misof et al. 2014) and discussion above), and using this knowledge to better contextualize the placement of ambiguous fossil groups (e.g. (Béthoux & Nel 2002; Grimaldi & Engel 2005; Davis et al. 2011)), in order to provide a more complete historical record of hexapod evolution (Nicholson 2012).
- Consideration of the potential impacts of outstanding uncertainty in taxonomic description (estimated clade richness) and the impacts this may have for inferred patterns of diversification.
- Improved data standardization and integration for the exploration of trait data both at a higher taxonomic and species level with an aim to condense hundreds of years of ecological observations into a referenced and reliable format to facilitate future studies on trait mediated diversification within the group.
- Integration of trait based analyses of diversification drivers into a more cohesive multivariate framework (for examples see (Marx & Uhen 2010; Benson & Mannion 2011)) with a particular emphasis on trying to identify the biologically limiting controls that are likely to denote zones of ecological opportunity.
- Expansion of modeling frameworks for dealing with large and terminally unresolved clades in the context of birth-death and related modeling frameworks.

This is an exciting time for the study of hexapod diversity as molecular tools, including the ones described here and the recent genome work of (Misof et al. 2014), are beginning to generate the phylogenetic frameworks needed to support robust analyses comparing different hexapod clades. It is unfortunate therefore that the integration of potential trait correlates of richness such as diet and body size has not kept pace with these improvements, restricting our ability to generate and test sophisticated hypotheses regarding relative diversification rates within the group. In terms of modeling diversity we are again seeing an explosion of novel and potentially powerful methods for exploring species richness, although as noted above, there has as yet been too little emphasis on extending such frameworks in order to be suitable for use on large and (at the species level) unresolved groups that make up the overwhelming majority of life on Earth. The work presented here provides an early phase overview, within a specific hypothesis of hexapod relationships, of some of the potential processes that may be operating within the group. As phylogenies and methods become increasingly sophisticated, and data sources for the various traits on interest continue to improve, the hope is that we can expand our understanding of the potential macro-evolutionary drivers of speciation and extinction, and so further our understanding of the radiation of the most diverse animal clade on earth.

7. Appendix and Supplementary Tables

7.1. Link to supporting information for Chapters 2 and 3

Additional supporting information for Chapters 2 and 3, including Accession Numbers for genes used in tree construction, implemented alignments, and digital copies of the inferred tree available at: [last accessed 26th January 2015]

http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0109085#s

7.2. Link to supporting information for Chapter 4

Supporting files on the Dryad data repository, including diet coding for all terminal groups and associated references available at

http://dx.doi.org/10.5061/dryad.6f75v

7.2.1. Definitions of ecological states

1. **Fungivory-** use of living fungal tissue or symbiotic fungi for external digestion of a detritus substrate (common in wood boring taxa, particularly beetles and primitive Hymenoptera). Includes organisms specialized on slime molds and yeasts. Not spore feeding (detritivory (category 2) unless the organism has some other close association with fungal tissue). Unlike the other categories discussed, the specialization state of fungivory is treated as variable depending on whether the relevant sources note a close relationship between the taxon and a particular subgroup of fungi.
2. **Detritivory**- feeding on decaying substrates and/or associated microbial communities. Includes scavenging and corpse feeding (animals capable of killing live prey are treated under predation (4)), dung feeding, shredding of decaying plant material, microbial film feeders and filter feeders in aquatic settings. All detritivores are treated as generalist in the analysis of specialization.
3. **Phytophagy**- feeding on living plant material including vegetative parts, roots and seeds (the last only if taken in situ with pests of stored grain treated under detritivory (2)). Includes taxa feeding on both vascular plants and lower plants such as mosses and

liverworts. Algal feeders, including wrack, are excluded and treated under detritivory (2). Includes wood boring of the cambium layer of living trees (taxa boring into dead wood treated under detritivory (2)) and larval pollen feeding. The last two are recoded as detritivory (2) for the purposes of the Larval Modified data set. For adults, pollen and nectar feeding are excluded (as detritivory (2) or liquid feeding (9) respectively). Phytophagy is considered a specialized ecology in the specialization analysis.

4. Predation- feeding on multiple other animals that the focal organism has the capacity to kill. Includes taxa which also scavenge invertebrates, but corpse feeding of larger animals and specialized kleptoparasites are treated under detritivory (2). Facultative predation and cannibalism by taxa with predominantly phytophagous or detritivous taxa are excluded, except in the presence of clear adaptions towards hunting behavior. Omnivores are treated as mixed states. Predation is considered generalist ecology in the specialization analysis.

5. **Parasitoidism**- Taxa which, as larvae, complete their development using a single individual arthropod host which dies as a direct result of having been fed upon (Godfray 1994). Also includes similar groups that engage in multiple provisioning e.g. Sphecidae (*sensu lato*) and some *Bombyliidae,* as well as parasitoids of molluscs or other invertebrate groups. Parasitoidism is considered a specialized ecology in the specialization analysis and combined with ecto-parasitism (6) in the Larval-modified data set.

6. **Ecto-parasitism**- Feeding on blood, flesh, or other products on vertebrates in the context of long-term associations. Also includes, for larval surveys, taxa which do not directly feed as larvae but which have blood-feeding adults (e.g. Hippoboscoidea). In the context of adults this category is extended to include micro-predators such as *Tabanidae* and many *Culicomorpha.* With the exception of the later (considered generalist) ecto-parasitism is treated as a specialized ecology in the specialization analysis and combined with parasitoidism (5) in the Larval-modified data set.

7. Non-feeding- Adult organisms within this category have vestigial mouth parts and are incapable of feeding.

8. Liquid feeding and Nectivory- Feeding exclusively or near-exclusively on liquid substrates such as nectar, liquid products of decay, honeydew or host haemolymph, primarily in the context of maintenance of the winged adult. Includes some adult taxa such as Trichoptera and certain Diptera and Hymenoptera for which feeding status is undetermined (see below)

In cases of polymorphic feeding strategies, sexually dimorphic taxa are denoted as the female state (which in most insects in the primary feeding state) with the exception of the parasitoid Strepsiptera, where in most families the larval-form female never leaves its host, and which are coded on the non-feeding male. Species showing caste dimorphisms are coded on the worker stage, although in practice there is no distinction at the level of resolution used in this study. Unless otherwise stated, non-holometabolan insects are assumed to feed on the same substrate throughout the life cycle, resulting in identical coding's for the "larval" and adult ecologies (denoted by – below). Taxa for which there is no ecological information available are denoted by "?".

7.3. Body length data for included terminal groups with references (Chapter 5)

Taxon	**Richness**	**Length Data**		**Raw Data**		**Reference**	**Dietary Substrate**			**Notes**
		Min /mm	Max /mm	Min /mm	Max /mm		**Larval Raw**	**Larval Mod.**	**Adult**	
Archaeognatha	495	10	12			(Arnett 2000)	2	2	2	
Blattodea Blaberidae	1198	2.5	75			(Arnett 2000)/(Hogue 1993)	2	2	2	
Blattodea Blattidae	2381	18	45			(Arnett 2000)	2	2	2	
Blattodea Cryptocercidae	594	24	29			(Arnett 2000)	2	2	2	
Blattodea Ectobiidae	12	8	18			(Arnett 2000)	2	2	2	
Blattodea Corydiidae	247	15	24			(Arnett 2000)	2	2	2	Includes Nocticolidae
Coleoptera Amphizoidae	5	11	16			(Parker 1982)	4	4	4	
Coleoptera Aspidytidae	2	4.8	7			(Beutel & Leschen 2005)	4	4	4	
Coleoptera Carabidae	40000	1	85			(Beutel & Leschen 2005)	4	4	3&4	
Coleoptera Dytiscidae	4015	1	48			(Beutel & Leschen 2005)	4	4	2&4	
Coleoptera Gyrinidae	882	3	15			(Parker 1982)	4	4	4	
Coleoptera Haliplidae	218	2	6			(Parker 1982)	2	2	4	
Coleoptera Hygrobiidae	5	8	10			(Parker 1982)	4	4	4	
Coleoptera Noteridae	250	1	5.8			(Beutel & Leschen 2005)	4	4	4	
Coleoptera	6	3.8	7			(Beutel &	4	4	4	

Trachypachidae				Leschen 2005)			
Coleoptera Cupedidae	31	5	22	(Beutel & Leschen 2005)	1	1	2
Coleoptera Micromalthidae	1	1.5	2.5	(Parker 1982)	1	1	1
Coleoptera Ommatidae	6	6	27	(Beutel & Leschen 2005)	?	?	2
Coleoptera Lepiceridae	1	1.5	2	(Beutel & Leschen 2005)	2	2	2
Coleoptera Hydroscaphidae	22	1	2	(Beutel & Leschen 2005)	?	?	2
Coleoptera Sphaeriusidae	19	0.5	1.2	(Beutel & Leschen 2005)	2	2	2
Coleoptera Torridincolidae	60	1	2.7	(Beutel & Leschen 2005)	2	2	2
Coleoptera Aderidae	900	1	4	(Arnett et al. 2010)	?	?	?
Coleoptera Agyrtidae	70	4	14	(Arnett & Thomas 2000)	2	2	2
Coleoptera Alexiidae	50	1.2	1.7	(F. W. Shockley 2008)	1	1	1
Coleoptera Anobiidae	2084	1	9	(Parker 1982)	2&3	2	2&3
Coleoptera Anthicidae	3000	1.5	15	(Parker 1982)	2	2	2&4
Coleoptera Anthribidae	3900	1	20	(Parker 1982)	1&3	1&3	1&2&3
Coleoptera Polyphaga Artematopodidae	45	2.5	10	(Leschen et al. 2010)	3	3	?
Coleoptera Attelabidae	2500	1	18	(Parker 1982)	3	3	3
Coleoptera Belidae	375	4.5	20	(Parker 1982)	3	3	2
Coleoptera Biphyllidae	200	1.5	8	(Parker 1982)	1	1	1
Coleoptera Boridae	4	8	25	(Arnett et al. 2010)	?	?	?
Coleoptera Bostrichidae	570	1	50	(Parker 1982)	2&3	2	2&3
Coleoptera Bothrideridae	400	1.5	13	(Arnett et al. 2010)	1&5	1&5	4
Coleoptera Brachyceridae	385	1.5	6	(Arnett et al. 2010)	3	3	3
Coleoptera Brentidae	4000	3	80	(Parker 1982)	1&3	1&3	3
Coleoptera Buprestidae	14700	1.5	60	(Parker 1982)	3	2	3
Coleoptera Byrrhidae	430	1.5	10	(Parker 1982)	3	3	3
Coleoptera	24	2.5	8	(Parker	3	3	2&8

Byturidae				1982)			
Coleoptera Callirhipidae	150	9	23	(Arnett et al. 2010)	1	1	?
Coleoptera Cantharidae	5100	1.5	30	(Parker 1982)	4	4	2&4
Coleoptera Cephaloidae	19	4.2	22	(Leschen et al. 2010)	1	1	2
Coleoptera Cerambycidae	30079	2	200	(Parker 1982)	3	2&3	3
Coleoptera Ceratocanthidae	120	2	9	(Arnett et al. 2010)	1	1	1
Coleoptera Cerylonidae	450	1	4	(Parker 1982)	1	1	1
Coleoptera Chelonariidae	250	2.5	10	(Parker 1982)	2	2	?
Coleoptera Chrysomelidae	32500	1	40	(Parker 1982)	3	3	3
Coleoptera Ciidae	650	0.5	7	(Parker 1982)	1	1	1
Coleoptera Clambidae	170	0.7	2	(Parker 1982)	1	1	2
Coleoptera Cleridae	3400	2	25	(Parker 1982)	4	4	2&4
Coleoptera Coccinellidae	6000	1	10	(Parker 1982)	3&4	3&4	3&4
Coleoptera Corylophidae	200	0.7	2.3	(Parker 1982)	1	1	2
Coleoptera Cryptophagidae	600	1	4	(Parker 1982)	1	1	1
Coleoptera Cucujidae	44	2.5	25	(Leschen et al. 2010)	4	4	?
Coleoptera Curculionidae	50615	1	55	(Parker 1982)	1&3	1&3	1&2&3
Coleoptera Dascillidae	80	6	20	(Parker 1982)	2	2	?
Coleoptera Dermestidae	1200	1	12	(Parker 1982)	2	2	2
Coleoptera Derodontidae	30	1.5	4	(Parker 1982)	1	1	1&4
Coleoptera Discolomatidae	400	1.1	8	(Parker 1982)	1	1	1
Coleoptera Drilidae	120	3	10	(Parker 1982)	4	4	4
Coleoptera Dryopidae	300	2	8	(Parker 1982)	2	2	2
Coleoptera Elateridae	10000	1.5	60	(Parker 1982)	3&4	3&4	2&3&4
Coleoptera Elmidae	1500	1	8	(Parker 1982)	2	2	2
Coleoptera Endomychidae	1800	1	18	(Parker 1982)	1	1	1
Coleoptera Epimetopidae	27	1	4	(Beutel & Leschen 2005)	4	4	2
Coleoptera Erotylidae	2500	2.5	25	(Parker 1982)	1	1	1
Coleoptera Eucinetidae	53	0.8	4	(Parker 1982)	1	1	?
Coleoptera Eucnemidae	1500	1.5	40	(Leschen et al. 2010)	1	1	?

Coleoptera Eulichadidae	30	15	25	(Parker 1982)	2	2	?
Coleoptera Georissidae	77	1	3	(Parker 1982)	4	4	2
Coleoptera Geotrupidae	920	5	45	(Arnett et al. 2010)	2	2	2&7
Coleoptera Glaphyridae	204	6	20	(Arnett et al. 2010)	2	2	?
Coleoptera Glaresidae	57	2.5	6	(Arnett et al. 2010)	?	?	1
Coleoptera Helophoridae	183	2	9	(Beutel & Leschen 2005)	4	4	2
Coleoptera Helotidae	107	6	16	(Parker 1982)	2	2	2
Coleoptera Heteroceridae	300	1	8	(Arnett et al. 2010)	2	2	2
Coleoptera Histeridae	4300	0.5	20	Parker et al 82	4	4	2&4
Coleoptera Hybosoridae	572	5	7	(Arnett et al. 2010)	2	2	2
Coleoptera Hydraenidae	1600	1.2	3	(Parker 1982)	2	2	2
Coleoptera Hydrochidae	164	2	4	(Jäch & Balke 2003)	4	4	2
Coleoptera Hydrophilidae	3400	1	40	(Parker 1982)	4	4	2
Coleoptera Ithyceridae	6	12	15	(Parker 1982)	3	3	3
Coleoptera Kateretidae	95	1.3	6	(Hisamatsu 2011)	3	3	3
Coleoptera Laemophloeidae	430	1	5	(Arnett et al. 2010)	1&2	1&2	1&4
Coleoptera Lampyridae	2200	4	30	(Parker 1982)	4	4	4
Coleoptera Languriidae	1000	1.2	25	(Parker 1982)	1&2&3	1&2&3	1&2&3
Coleoptera Latridiidae	1000	1	3	(Shockley et al. 2011)	1	1	1
Coleoptera Leiodidae	3700	1	7	(Parker 1982)	1	1	1&2
Coleoptera Limnichidae	390	0.8	3	(Parker 1982)	2	2	2
Coleoptera Lucanidae	1489	4	80	(Parker 1982)	2	2	2&8
Coleoptera Lutrochidae	11	3	5	(Parker 1982)	2	2	2
Coleoptera Lycidae	4600	3	22	(Parker 1982)	1	1	8
Coleoptera Lymexylidae	70	5	40	(Parker 1982)	1	1	?
Coleoptera Mauroniscidae	26	2	4.5	(Leschen et al. 2010)	?	?	?
Coleoptera Melandryidae	420	1.2	19	(Parker 1982)	1	1	1
Coleoptera Meloidae	3000	5	33	(Parker 1982)	2&5	2&5	3
Coleoptera Melyridae	6000	1	20	(Parker 1982)	2&4	2&4	2&4
Coleoptera	250	1.3	5	(Parker	1&4	1&4	1&4

Monotomidae				1982)			
Coleoptera Mordellidae	1500	2	15	(Parker 1982)	3	3	2
Coleoptera Mycetophagidae	130	0.8	6.5	(Parker 1982)	1	1	1&2
Coleoptera Nemonychidae	70	4	6	(Parker 1982)	3	3	2
Coleoptera Nitidulidae	4500	0.9	14	(Parker 1982)	1&2	1&2	1&2
Coleoptera Nosodendridae	50	2.5	9	(Parker 1982)	1	1	1&4
Coleoptera Ochodaeidae	110	3	10	(Arnett et al. 2010)	?	?	?
Coleoptera Oedemeridae	500	5	20	(Parker 1982)	2	2	2
Coleoptera Omalisidae	8	3	9	(Leschen et al. 2010)	4	4	?
Coleoptera Omethidae	33	3	12	(Arnett et al. 2010)	?	?	?
Coleoptera Orsodacnidae	40	4	15	(Arnett et al. 2010)	3	3	3
Coleoptera Passalidae	800	18	80	(Parker 1982)	2	2	2
Coleoptera Passandridae	109	3	35	(Leschen et al. 2010)	5	5	?
Coleoptera Perimylopidae	19	6	10	(Parker 1982)	2	2	2
Coleoptera Phalacridae	640	1.2	4.5	(Parker 1982)	1	1	1
Coleoptera Phengodidae	250	3	65	(Parker 1982)	4	4	4
Coleoptera Phloeostichidae	14	2.4	15	(Leschen et al. 2010)	1	1	1
Coleoptera Phloiophilidae	1	2	3	(Parker 1982)	1	1	1
Coleoptera Pleocomidae	50	15	45	(Parker 1982)	3	3	7
Coleoptera Prionoceridae	160	5.5	20	(Leschen et al. 2010)	2&4	2&4	2
Coleoptera Propalticidae	30	1.2	1.8	(Parker 1982)	1	1	1
Coleoptera Prostomidae	30	5	10	(Arnett et al. 2010)	1	1	1
Coleoptera Protocucujidae	7	3.5	5.8	(Leschen et al. 2010)	?	?	?
Coleoptera Psephenidae	290	2	7	(Parker 1982)	2	2	2
Coleoptera Ptiliidae	650	0.3	2	(Parker 1982)	1	1	1
Coleoptera Ptilodactylidae	500	2	16	(Parker 1982)	2	2	2
Coleoptera Ptinidae	500	1	5	(Parker 1982)	2&3	2&3	2
Coleoptera Pyrochroidae	167	7	18	(Parker 1982)	1&2	1&2	2
Coleoptera Pythidae	23	3	20	(Parker 1982)	2&3	2	2&4
Coleoptera Rhipiceridae	70	10	25	(Parker 1982)	5	5	7
Coleoptera	400	2	38	(Parker	5	5	?

Ripiphoridae				1982)				
Coleoptera Salpingidae	300	1.5	12	(Parker 1982)	2&4	2&4	2	
Coleoptera Scarabaeidae	27000	1	160	(Parker 1982)	2&3	2&3	2&3&7	
Coleoptera Scirtidae	800	1.5	12	(Parker 1982)	2	2	2	
Coleoptera Scraptiidae	500	1.3	15	(Parker 1982)	1&2	1&2	?	
Coleoptera Scydmaenidae	4586	0.5	7	(Parker 1982)	4	4	4	
Coleoptera Silphidae	200	7	45	(Parker 1982)	2	2	2&4	
Coleoptera Silvanidae	500	2	15	(Arnett et al. 2010)	1&2	1&2	1&2	
Coleoptera Spercheidae	19	3	7	(Darilmaz & Kiyak 2011)	2	2	2	
Coleoptera Sphaeritidae	5	4	6	(Parker 1982)	4	4	?	
Coleoptera Sphindidae	59	1.5	3.5	(Leschen et al. 2010)	1	1	1	
Coleoptera Staphylinidae	56000	0.5	50	(Parker 1982)	1&2&4	1&2&4	1&2&4	
Coleoptera Synchroidae	8	7	13	(Parker 1982)	1	1	1	
Coleoptera Tenebrionidae	20000	1	50	(Parker 1982)	1&2	1&2	1&2	
Coleoptera Tetratomidae	150	2.8	15	(Parker 1982)	1	1	1	
Coleoptera Throscidae	150	1.2	6	(Leschen et al. 2010)	1	1	2	
Coleoptera Trictenotomidae	13	32	80	(Leschen et al. 2010)	1	1	?	
Coleoptera Trogossitidae	600	1	50	(Parker 1982)	2&4	2&4	1&2&4	
Coleoptera Zopheridae	1700	2	40	(Parker 1982)	1&2&4	1&2&4	1&2&4	
Collembola Entomobryidae	2189	1	10	(Arnett 2000)	2	2	2	Includes Paronellidae -
Collembola Hypogastruridae	682	0.8	3	(Arnett 2000)	2	2	2	
Collembola Isotomidae	1346	0.7	6	(Arnett 2000)	2	2	2	
Collembola Neanuridae	1546	2	3.5	(Arnett 2000)	2	2	2	includes Brachystomellidae
Collembola Neelidae	33	0.3	0.7	(Arnett 2000)	2	2	2	
Collembola Onychiuridae	913	0.5	3	(Arnett 2000)	2	2	2	includes Odontellidae + Tullbergiidae
Collembola Poduridae	1	1.3	2	(Arnett 2000)/(Hopkin 1997)	2	2	2	
Collembola Sminthuridae	742	0.4	2.7	(Arnett 2000)	2	2	2	includes Bourletiellidae, Dicyrtomida

										e and Oncopoduridae
Collembola Tomoceridae	354	6	10			(Arnett 2000)	2	2	2	includes Katiannidae
Dermaptera Anisolabididae	38	9	13			(Arnett 2000)	2&4	2&4	2&4	
Dermaptera Apachyidae	15	11	25			(Boeseman 1954)	2&4	2&4	2&4	
Dermaptera Chelisochidae	95	16	20			(Arnett 2000)	2&4	2&4	2&4	
Dermaptera Forficulidae	485	10	18			(Arnett 2000)	2&3&4	2&3&4	2&3&4	
Dermaptera Labiduridae	64	18	80			(Arnett 2000)/(Berenbaum 2007)	2&4	2&4	2&4	
Dermaptera Labiidae	495	4	7			(Arnett 2000)	2&4	2&4	2&4	
Dermaptera Pygidicranidae	181	9	45			(Parker 1982)	2&4	2&4	2&4	
Diplura Campodeidae	448	8	10			(Arnett 2000)	2&3&4	2&3&4	2&3&4	
Diplura Japygoidea	590	8	50			(Arnett 2000)	2&3&4	2&3&4	2&3&4	
Diptera Acartophthalmidae	6	2.5	3			(McAlpine et al. 1987)	?	?	8	
Diptera Acroceridae	400	2	21			(Brown et al. 2009)	5	5	8	
Diptera Agromyzidae	3017	0.9	6.5			(Brown et al. 2009)	3	3	8	
Diptera Anisopodidae	196	2	18			(McAlpine et al. 1981)	2	2	8	
Diptera Anthomyiidae	1941	2	12			(McAlpine et al. 1987)	2&3&5	2&3&5	4&8	
Diptera Anthomyzidae	100	1.1	3.4			(Brown et al. 2009)	2	2	8	
Diptera Apioceridae	143	7.5	35			(McAlpine et al. 1981)	4	4	8	
Diptera Apsilocephalidae	7	4.5	5.5			(Nagatomi et al. 1991)	?	?	8	
Diptera Asilidae	7531	3	60			(Brown et al. 2009)	4	4	4	
Diptera Asteiidae	138	1	5			(Brown et al. 2009)	2	2	8	
Diptera Atelestidae	22	1.5	4			(Wiegmann 1989)/(Capinera 2008)	?	?	2	
Diptera Athericidae	133	7	10			(Brown et al. 2009)	4	4	6&8	
Diptera Aulacigastridae	19	1.5	4			(Brown et al. 2009)	2	2	8	
Diptera Australimyzidae	9	1.3	2.6			(Brake & Mathis 2007)	?	?	8	
Diptera Austroleptidae	8	3.1	5.3			(Nagatomi & Nagatomi 1987)	?	?	8	
Diptera Axymyiidae	8	4	7	5	8	(Schneeberg et al. 2013)	2	2	8	Data given as wing length
Diptera	1382	2	15			(Brown et	2	2	8	Includes

Bibionidae				al. 2009)			Pleciidae
Diptera Blephariceridae	331	3	13	(McAlpine et al. 1981)	2	2	4
Diptera Bombyliidae	5382	4	40	(Arnett 2000)	5	5	2&8
Diptera_Braulidae	7	1	1.7	(McAlpine et al. 1987)	2	2	8
Diptera Calliphoridae	1525	4	16	(McAlpine et al. 1987)	2&5	2&5	4&8
Diptera Canthyloscelidae	14	2	3.5	Manual of Neoarctic diptera	2	2	8
Diptera Carnidae	92	1	3	(Brown et al. 2009)	2	2	2&8
Diptera Cecidomyiidae	6296	1	8	(McAlpine et al. 1981)	3	3	2&8
Diptera Ceratopogonidae	5902	1	6	(McAlpine et al. 1981)	2&4	2&4	4&6
Diptera Chaoboridae	89	1.4	10	(McAlpine et al. 1981)	4	4	8
Diptera Chironomidae	7290	1	13	(Brown et al. 2009)	2&4	2&4	8
Diptera Chloropidae	2885	1	7	(Karpa 2001)	2&3&4	2&3&4	2&8
Diptera Chyromyidae	139	0.5	4.5	(Brown et al. 2009)/ (McAlpine et al. 1987)	2	2	8
Diptera Clusiidae	363	1.8	7.5	(McAlpine et al. 1987)	4	4	8
Diptera Coelopidae	35	3	16	(Brown et al. 2009)	2	2	8
Diptera Conopidae	831	2.5	30	(Brown et al. 2009)	5	5	8
Diptera Corethrellidae	111	0.6	2.5	(Brown et al. 2009)	4	4	6
Diptera Culicidae	3725	3	9	(McAlpine et al. 1981)	2	2	6
Diptera Deuterophlebiidae	14	2	4	(Arnett 2000)	2	2	7
Diptera Diadocidiidae	39	3	10	(Bechev & Chandler 2011)	1	1	8
Diptera Diopsidae	194	4	12	(McAlpine et al. 1987)	2	2	8
Diptera Dixidae	197	4.5	7	(Arnett 2000)	2	2	7
Diptera Dolichopodidae	7358	0.8	9	(McAlpine et al. 1981)	4	4	4
Diptera Drosophilidae	4017	1	7	(Brown et al. 2009)	1&2	1&2	2&8
Diptera Dryomyzidae	30	4	18	(Mathis & Sueyoshi 2011)	2	2	8
Diptera Empididae	3142	2	12	(Capinera 2008)	4	4	2&4
Diptera Ephydridae	1994	0.6	11	(Brown et al. 2009)	2&4	2&4	4&8
Diptera Fanniidae	359	3.5	7.5	(Brown et al. 2009)	2	2	8
Diptera Fergusoninidae	29	2	3	(Nelson et al. 2011)	3	3	8

Diptera Glossinidae	25	6	14	(Wall & Shearer 2008)	6	5	6
Diptera Helcomyzidae	12	3	16	(Mathis 2011a)	2	2	8
Diptera Helosciomyzidae	23	5	11	(Barnes 1981)	4	4	8
Diptera Hesperinidae	10	4.7	12	(Papp 2010)	2	2	8
Diptera Heterocheilidae	2	4.2	6.5	(Mathis 2011b)	2	2	8
Diptera Hilarimorphidae	36	1.8	7.2	(McAlpine et al. 1981)	?	?	?
Diptera Hippoboscidae	271	1.5	12	(McAlpine et al. 1987)	6	5	6
Diptera Hybotidae	2005	1	9	(Capinera 2008)	4	4	4
Diptera Keroplatidae	993	2.8	8.8	(Brown et al. 2009)	1&4	1&4	7
Diptera Lauxaniidae	1900	2	11	(Brown et al. 2009)	1&2	1&2	1
Diptera Lonchaeidae	504	3	6	(Brown et al. 2009)	2	2	8
Diptera Lonchopteridae	65	2	4	(Brown et al. 2009)	2	2	2&8
Diptera Lygistorrhinidae	44	3	5	(Brown et al. 2009)	?	?	8
Diptera Marginidae	3	1.5	2	(McAlpine 1991)	?	?	8
Diptera Micropezidae	583	5	17	(Brown et al. 2009)	2&3	2&3	8
Diptera Milichiidae	288	1	7	(Brown et al. 2009)	2	2	2&8
Diptera Muscidae	5218	2	20	(Brown et al. 2009)	2&4	2&4	4&6&8
Diptera Mycetophilidae	4525	2.2	13.3	(McAlpine et al. 1981)	1	1	8
Diptera Mydidae	498	9	60	(McAlpine et al. 1981)	4	4	8
Diptera Mythicomyiidae	350	0.8	3	(Brown et al. 2009)	4	4	8
Diptera Nemestrinidae	300	4	16	(Brown et al. 2009)	5	5	8
Diptera Neurochaetidae	22	1.5	4.1	(McAlpine 1993)	2	2	?
Diptera Nycteribiidae	274	1.5	5.5	(Brown et al. 2009)	6	5	6
Diptera Odiniidae	65	2.5	6	(Brown et al. 2009)	2	2	8
Diptera Oestridae	176	8	25	(McAlpine et al. 1987)	6	5	7
Diptera Opomyzidae	61	2	4.4	(McAlpine et al. 1987)	3	3	8
Diptera Pachyneuridae	8	5	6	(Arnett 2000)	1	1	8
Diptera Pallopteridae	71	3	5	(McAlpine et al. 1987)	3&4	3&4	2&8
Diptera Pelecorhynchidae	49	4	18	(McAlpine et al. 1981)	4	4	8
Diptera Periscelididae	91	2.5	5	(Brown et al. 2009)	2	2	8
Diptera	9	1	2	(Colless	1	1	8

Perissommatidae						1969)				
Diptera Phoridae	4200	0.5	6			(Brown et al. 2009)	1&2&3&4&5	1&2&3&4&5	2&8	
Diptera Piophilidae	83	3	8			(Brown et al. 2009)	2	2	8	
Diptera Pipunculidae	1428	2	11.5			(Brown et al. 2009)	5	5	8	
Diptera Platypezidae	277	1.4	10			(Brown et al. 2009)	1	1	8	
Diptera Platystomatidae	1164	2.5	20			(Brown et al. 2009)	2&3	2&3	8	
Diptera Psilidae	322	3	12			(Brown et al. 2009)	3	3	2&8	
Diptera Psychodidae	3026	1	5			(Brown et al. 2009)	2	2	6&7	
Diptera Ptychopteridae	156	7	14			(Brown et al. 2009)	2	2	8	
Diptera Pyrgotidae	351	5	30			(Brown et al. 2009)	5	5	8	
Diptera Rhagionidae	756	4	12			(Brown et al. 2009)	4	4	6&7	
Diptera Rhinophoridae	174	3.5	8			(Brown et al. 2009)	5	5	8	
Diptera Richardiidae	178	3	15			(Brown et al. 2009)	2	2	8	
Diptera Sarcophagidae	3094	5	25			(Brown et al. 2009)	2&4&5	2&4&5	8	
Diptera Scathophagidae	419	3	13			(McAlpine et al. 1987)	2&3&4	2&3&4	4	
Diptera Scatopsidae	407	0.6	4.1			(McAlpine et al. 1981)	2	2	8	
Diptera Scenopinidae	420	1	8.5			(Oosterbroek 1998)	4	4	8	
Diptera Sciaridae	2455	1	11			(McAlpine et al. 1981)	2	2	8	
Diptera Sciomyzidae	618	2	13			(Brown et al. 2009)	4&5	4&5	2&8	
Diptera Sepsidae	345	2	7			(Brown et al. 2009)	2	2	8	
Diptera Simuliidae	2121	1	5.5			(McAlpine et al. 1981)	2	2	6	
Diptera Somatiidae	7	3.5	5			(Brown et al. 2009)	?	?	2&8	
Diptera Sphaeroceridae	1571	0.7	6			(Brown et al. 2009)	2	2	8	
Diptera Stratiomyidae	2690	2	28			(Brown et al. 2009)	2	2	8	
Diptera Streblidae	237	0.7	5.5			(Brown et al. 2009)	6	5	6	
Diptera Strongylophthalmyiidae	45	2	6			(Palaczyk et al. 2013)	2	2	8	
Diptera Synneuridae	3	2	3.5			(McAlpine et al. 1981)	2	2	8	
Diptera Syrphidae	6107	4	25			(Brown et al. 2009)	2&4	2&4	2&8	
Diptera Tabanidae	4434	6	30			(McAlpine et al. 1981)	4	4	6	
Diptera Tachinidae	9626	3	25			(Brown et al. 2009)	5	5	8	
Diptera Tanyderidae	55	11	23	20	42	(Arnett 2000)	?	?	?	Data given as wingspan

Diptera Tephritidae	4716	2	35			(Brown et al. 2009)	3	3	8	
Diptera Thaumaleidae	183	2	4.5			(Arnett 2000)	2	2	7&8	
Diptera Therevidae	1143	2.5	15			(McAlpine et al. 1981)	4	4	8	
Diptera Tipulidae	1577 0	6	60			(McAlpine et al. 1981)	2&3& 4	2&3& 4	7&8	
Diptera Trichoceridae	183	3	9			(Brown et al. 2009)	2	2	8	
Diptera Ulidiidae	678	2	14			(Brown et al. 2009)	2	2	8	
Diptera Vermileonidae	61	7	12			(Brown et al. 2009)	4	4	8	
Diptera Xenasteiidae	13	1.2	2			(Evenhius 2011)	2	2	8	
Diptera Xylomyidae	138	5	15			(McAlpine et al. 1981)	2	2	8	
Diptera Xylophagidae	145	2	25			(McAlpine et al. 1981)	4	4	8	
Embioptera	337	4	22			(Arnett 2000)	2	2	2	
Ephemeroptera Ameletidae	56	7	21			(Zloty & Pritchard 1997)	2	2	7	
Ephemeroptera Ameletopsidae	6	15.5	22			(Mercado & Elliot 2005)	4	4	7	
Ephemeroptera Ametropodidae	3	13	15	13	15	(Edmunds et al. 1976)	2	2	7	Data as forewing length
Ephemeroptera Baetidae	860	3	10			(Arnett 2000)	2&4	2&4	7	
Ephemeroptera Baetiscidae	12	8	16	8	16	(Edmunds et al. 1976)	2	2	7	Data as forewing length
Ephemeroptera Behningiidae	7	12	18			(Parker 1982)	4	4	7	
Ephemeroptera Caenidae	211	2	6			(Arnett 2000)	2	2	7	
Ephemeroptera Coloburiscidae	6	13	18			(Marsh 2004)	2	2	7	
Ephemeroptera Dipteromimidae	2	13	23.5			(Tojo & Matsukawa 2003)	?	?	7	
Ephemeroptera Ephemerellidae	91	5	12			(Arnett 2000)	2	2	7	
Ephemeroptera Ephemeridae	160	10	32			(Parker 1982)	2&4	2&4	7	
Ephemeroptera Euthyplociidae	19	11	16			(Gillies 1980)	2	2	7	
Ephemeroptera Heptageniidae	529	4	14			(Arnett 2000)	2&4	2&4	7	
Ephemeroptera Ichthybotidae	2	19	22			(Phillips 1930)	2	2	7	
Ephemeroptera Isonychiidae	30	9	16			(Arnett 2000)	2	2	7	
Ephemeroptera Leptohyphidae	157	2	10			(Dominguez et al. 2006)	2	2	7	
Ephemeroptera Leptophlebiidae	623	4	12	4	14	(Edmunds et al. 1976)	2	2	7	Data as forewing length
Ephemeroptera Metretopodidae	13	9	16	9	16	(Edmunds et al. 1976)	2&4	2&4	7	Data as forewing length

Ephemeroptera_ Neoephemeridae	7	6	13			(Bae & McCafferty 1998)	2	2	7	
Ephemeroptera Nesameletidae	11	10.5	16.5			(Hitchings & Staniczek 2003)	2	2	7	
Ephemeroptera Oligoneuriidae	54	6	10	6	10	(Edmunds et al. 1976)	2	2	7	Data as forewing length
Ephemeroptera Oniscigastridae	8	10	11			(Heckman 2002)	2	2	7	
Ephemeroptera Palingeniidae	32	15	35			(Parker 1982)	2	2	7	
Ephemeroptera Polymitarcyidae	84	12	35			(Parker 1982)	2	2	7	
Ephemeroptera Potamanthidae	23	8	25			(Parker 1982)	2	2	7	
Ephemeroptera Prosopistomatidae	19	1.5	4.5			(Pearson & Penridge 1979)	4	4	7	
Ephemeroptera Rallidentidae	1	10.5	12			(Penniket 1966)	2	2	7	
Ephemeroptera Siphlaenigmatidae	1	8	9			(Penniket 1962)	2	2	7	
Ephemeroptera Siphlonuridae	49	9	13			(Arnett 2000)	2&4	2&4	7	
Ephemeroptera Tricorythidae	34	4	6.5			(Edmunds et al. 1976)	2	2	7	Data as forewing length
Grylloblattidae	27	10	30			(Arnett 2000)	4	4	4	
Hemiptera Acanthosomatidae	200	6	18			(Schuh & Slater 1995)	3	3	3	
Hemiptera Achilidae	503	3	13			(Capinera 2008)	1	1	3	
Hemiptera Achilixiidae	24	4	8			(Capinera 2008)	3	3	3	
Hemiptera Aetalionidae	42	3	30			(Deitz et al. 2010)	3	3	3	
Hemiptera Aleyrodoidea	1560	1	4			(Capinera 2008)	3	3	3	
Hemiptera Alydidae	250	8	20			(Schuh & Slater 1995)	3	3	3	
Hemiptera Anthocoridae	600	1.4	4.5			(Schuh & Slater 1995)	4	4	4	
Hemiptera Aphelocheiridae	400	3.5	11.5			(Schuh & Slater 1995)	4	4	4	
Hemiptera Aphidoidea	4375	1	8			(Capinera 2008)	3	3	3	Includes Phylloxeroidea
Hemiptera Aradidae	2000	3	11			(Schuh & Slater 1995)	1	1	1	
Hemiptera Belostomatidae	150	9	110			(Schuh & Slater 1995)	4	4	4	
Hemiptera Berytidae	100	2.5	11			(Schuh & Slater 1995)	3&4	3&4	3&4	
Hemiptera Caliscelidae	202	1	5			(Capinera 2008)	3	3	3	
Hemiptera Canopidae	8	5	7			(Schuh & Slater 1995)	1	1	1	

Hemiptera Cercopidae	2410	5	20	(Arnett 2000)	3	3	3	Includes Aphrophoridae, Clastopteridae, Machaerotidae
Hemiptera Cicadellidae	20000	1.7	28	(Evans 1966)	3	3	3	
Hemiptera Cicadidae	1300	10	100	(Capinera 2008)	3	3	3	
Hemiptera Cimicidae	100	2	12	(Schuh & Slater 1995)	6	5	6	
Hemiptera Cixiidae	2223	3	13	(Capinera 2008)	3	3	3	
Hemiptera Coccoidea	8000	0.6	35	(Arnett 2000)	3	3	3	
Hemiptera Colobathristidae	90	6	20	(Schuh & Slater 1995)	3	3	3	
Hemiptera Coreidae	1900	7	45	(Schuh & Slater 1995)	3	3	3	
Hemiptera Corixidae	600	2.5	15	(Schuh & Slater 1995)	2&4	2&4	2&4	
Hemiptera Cydnidae	617	2	20	(Schuh & Slater 1995)	3	3	3	
Hemiptera Delphacidae	2029	2	10	(Capinera 2008)	3	3	3	
Hemiptera Derbidae	1700	4	11	(Capinera 2008)/ (Arnett 2000)	1	1	3	
Hemiptera Dictyopharidae	731	3	33	(Capinera 2008)	3	3	3	
Hemiptera Dinidoridae	90	9	27	(Schuh & Slater 1995)	3	3	3	
Hemiptera Dipsocoridae	30	0.8	3	(Schuh & Slater 1995)	4	4	4	
Hemiptera Enicocephalidae	400	2	15	(Schuh & Slater 1995)	4	4	4	
Hemiptera Eurybrachyidae	189	7	29	(Capinera 2008)	3	3	3	
Hemiptera Flatidae	1446	4	32	(Capinera 2008)	3	3	3	
Hemiptera Fulgoridae	687	4	100	(Capinera 2008)	3	3	3	
Hemiptera Gelastocoridae	100	7	15	(Schuh & Slater 1995)	4	4	4	
Hemiptera Gerridae	620	1.6	36	(Schuh & Slater 1995)	4	4	4	
Hemiptera Hebridae	150	1.3	3.7	(Schuh & Slater 1995)	4	4	4	
Hemiptera Hermatobatidae	8	2.5	4	(Schuh & Slater 1995)	4	4	4	
Hemiptera Hydrometridae	110	2.7	22	(Schuh & Slater 1995)	4	4	4	
Hemiptera Hyocephalidae	3	8	15	Resh and Carde 2009	3	3	3	
Hemiptera Idiostolidae	4	5	7	(Schuh & Slater 1995)	?	?	?	
Hemiptera Issidae	924	2	19	(Capinera 2008)	3	3	3	

Hemiptera Joppeicidae	1	2.5	3	(Schuh & Slater 1995)	4	4	4	Modified to avoid zero variance
Hemiptera Largidae	120	7	55	(Arnett 2000)/ (Schuh & Slater 1995)	3	3	3	
Hemiptera Leptopodidae	40	1.8	7	(Schuh & Slater 1995)	4	4	4	
Hemiptera Lestoniidae	2	3.5	5.6	(Schuh & Slater 1995)	3	3	3	
Hemiptera Lophopidae	138	5	15	(Capinera 2008)	3	3	3	
Hemiptera Lyctocoridae	27	2	6	(Schuh & Slater 1995)	4	4	4	
Hemiptera Lygaeidae	4400	1.2	12	(Schuh & Slater 1995)	3&4	3&4	3&4	
Hemiptera Macroveliidae	3	2.5	5.6	(Schuh & Slater 1995)	4	4	4	
Hemiptera Malcidae	20	3	4	(Schuh & Slater 1995)	3	3	3	
Hemiptera Meenoplidae	158	3	7	(Capinera 2008)	3	3	3	
Hemiptera Membracidae	3450	2	24	(Deitz et al. 2010)	3	3	3	
Hemiptera Mesoveliidae	35	1.2	4.2	(Schuh & Slater 1995)	4	4	4	
Hemiptera Microphysidae	30	1.5	3	(Schuh & Slater 1995)	4	4	4	
Hemiptera Miridae	10000	2	15	(Schuh & Slater 1995)	3&4	3&4	3&4	
Hemiptera Nabidae	400	7	11	(Arnett 2000)	4	4	4	
Hemiptera Naucoridae	500	5	20	(Schuh & Slater 1995)	4	4	4	
Hemiptera Nepidae	225	15	45	(Schuh & Slater 1995)	4	4	4	
Hemiptera Nogodinidae	286	4	17	(Capinera 2008)	3	3	3	
Hemiptera Notonectidae	350	5	15	(Schuh & Slater 1995)	4	4	4	
Hemiptera Ochteridae	50	4.5	9	(Schuh & Slater 1995)	4	4	4	
Hemiptera Paraphrynoveliidae	2	1.7	2.4	(Schuh & Slater 1995)	4	4	4	
Hemiptera Peloridiidae	12	2	5	(Resh & Cardé 2009)	3	3	3	
Hemiptera Pentatomidae	4500	4	20	(Schuh & Slater 1995)	3&4	3&4	3&4	
Hemiptera Phloeidae	3	20	30	(Schuh & Slater 1995)	3	3	3	
Hemiptera Piesmatidae	40	2.5	5	(Schuh & Slater 1995)	3	3	3	
Hemiptera Plataspidae	500	2	20	(Schuh & Slater 1995)	3	3	3	
Hemiptera Pleidae	40	1.5	3	(Schuh & Slater 1995)	4	4	4	
Hemiptera Plokiophilidae	6	1.2	3	(Schuh & Slater 1995)	4	4	4	
Hemiptera	2500	1	8	(Capinera	3	3	3	

Psylloidea				2008)				
Hemiptera Pyrrhocoridae	225	8	30	(Schuh & Slater 1995)	3	3	3	
Hemiptera Reduviidae	6700	7	40	(Schuh & Slater 1995)	4	4	4	Includes Phymatidae
Hemiptera Rhopalidae	200	4	15	(Schuh & Slater 1995)	3	3	3	
Hemiptera Ricaniidae	417	4	12	(Capinera 2008)	3	3	3	
Hemiptera Saldidae	265	2.3	7.4	(Schuh & Slater 1995)	4	4	4	
Hemiptera Schizopteridae	120	0.8	2	(Schuh & Slater 1995)	4	4	4	
Hemiptera Scutelleridae	500	5	20	(Schuh & Slater 1995)	3	3	3	
Hemiptera Stenocephalidae	30	8	15	(Schuh & Slater 1995)	3	3	3	
Hemiptera Termitaphididae	9	2	3	(Schuh & Slater 1995)	1	1	1	
Hemiptera Tessaratomidae	250	15	40	(Foottit & Adler 2009)	3	3	3	
Hemiptera Tettigometridae	73	3	11	(Capinera 2008)	3	3	3	
Hemiptera Thaumastocoridae	19	2	4.6	(Schuh & Slater 1995)	3	3	3	
Hemiptera Tingidae	2000	2	8	(Schuh & Slater 1995)	3	3	3	
Hemiptera Tropiduchidae	575	5	13	(Capinera 2008)	3	3	3	
Hemiptera Veliidae	720	1	10	(Schuh & Slater 1995)	4	4	4	
Hemiptera Velocipedidae	31	10	15	(Schuh & Slater 1995)	4	4	4	
Hymenoptera Agaonidae	757	1	3	(Parker 1982)	3&5	3&5	8	
Hymenoptera Ampulicidae	200	5	15	(Arnett 2000)	5	5	8	
Hymenoptera Anaxyelidae	1	7.5	8	(Parker 1982)	1	1	8	Modified to avoid zero variance
Hymenoptera Andrenidae	2938	4	22	(Arnett 2000)	3	2	8	
Hymenoptera Aphelinidae	1168	0.35	2.5	(Parker 1982)	5	5	8	
Hymenoptera Apidae	5751	3.5	27	(Arnett 2000)	3	2	8	
Hymenoptera Argidae	800	4	15	(Parker 1982)	3	3	8	
Hymenoptera Aulacidae	200	1	20	(Parker 1982)	5	5	8	
Hymenoptera Bethylidae	2000	1	20	(Parker 1982)	5	5	8	
Hymenoptera Blasticotomidae	10	6	10	(Parker 1982)	3	3	8	
Hymenoptera Braconidae	20000	2	15	(Parker 1982)	5	5	8	
Hymenoptera Bradynobaenidae	200	3	20	(Parker 1982)	5	5	8	
Hymenoptera Cephidae	80	5	25	(Parker 1982)	3	3	8	

Hymenoptera Ceraphronidae	350	0.5	5	(Parker 1982)	5	5	8	
Hymenoptera Chalcididae	1464	2	12	(Parker 1982)	5	5	8	
Hymenoptera Chrysididae	3000	2.5	20	(Parker 1982)	5	5	8	
Hymenoptera Cimbicidae	130	18	25	(Arnett 2000)	3	3	8	
Hymenoptera Colletidae	2545	3.5	20	(Arnett 2000)	3	2	8	
Hymenoptera Crabronidae	8774	6	20	(Arnett 2000)	5	5	8	
Hymenoptera Cynipidae	1000	1	8	(Parker 1982)	3	3	8	
Hymenoptera Diapriidae	2300	3	15	(Parker 1982)	5	5	8	
Hymenoptera Diprionidae	90	5	12	(Parker 1982)	3	3	8	
Hymenoptera Encyrtidae	3735	0.5	5	(Parker 1982)	5	5	8	
Hymenoptera Eucharitidae	423	3	10	(Parker 1982)	5	5	8	
Hymenoptera Eulophidae	4472	1	5	(Parker 1982)	5	5	8	
Hymenoptera Eupelmidae	907	1	8	(Parker 1982)	5	5	8	
Hymenoptera Eurytomidae	1424	3	5	(Parker 1982)	3&5	3&5	8	
Hymenoptera Evaniidae	500	2	15	(Parker 1982)	5	5	8	
Hymenoptera Figitidae	1500	1.5	5	(Parker 1982)	5	5	8	
Hymenoptera Formicidae	10000	1	33	(Arnett 2000)/(Lenhart et al. 2013)	4	4	8	Sizes given based on workers
Hymenoptera Gasteruptiidae	420	13	40	(Arnett 2000)	5	5	8	
Hymenoptera Halictidae	4338	4	10	(Arnett 2000)	3	2	8	
Hymenoptera Heloridae	7	4	7	(Parker 1982)	5	5	8	
Hymenoptera Ibaliidae	50	8	25	(Parker 1982)	5	5	8	
Hymenoptera Ichneumonidae	22000	3	40	(Parker 1982)	5	5	8	
Hymenoptera Liopteridae	50	4	15	(Parker 1982)	5	5	8	
Hymenoptera Maamingidae	2	1	2	(Early et al. 2001)	5	5	8	
Hymenoptera Megachilidae	4120	7	39	(Arnett 2000)/(Messer 1984)	3	2	8	
Hymenoptera Megalodontesidae	40	5	20	(Parker 1982)	3	3	8	
Hymenoptera Megalyridae	50	4	20	(Parker 1982)	5	5	8	
Hymenoptera Megaspilidae	450	1	5	(Parker 1982)	5	5	8	
Hymenoptera Melittidae	191	7	12	(Arnett 2000)	3	2	8	
Hymenoptera Monomachidae	20	7	22	(Parker 1982)	5	5	8	

Hymenoptera Mutillidae	5000	3	30	(Parker 1982)	5	5	8
Hymenoptera Mymaridae	1424	0.2	2	(Parker 1982)	5	5	8
Hymenoptera Mymarommatidae	9	0.3	0.8	(Gibson et al. 2007)	5	5	8
Hymenoptera Orussidae	75	5	20	(Parker 1982)	5	5	8
Hymenoptera Pamphiliidae	250	8	15	(Arnett 2000)	3	3	8
Hymenoptera Pelecinidae	3	30	60	(Parker 1982)	5	5	8
Hymenoptera Pergidae	500	7	10	(Arnett 2000)	3	3	8
Hymenoptera Perilampidae	277	1.5	7	(Parker 1982)	5	5	8
Hymenoptera Platygastridae	1100	0.5	5	(Parker 1982)	5	5	8
Hymenoptera Plumariidae	20	3	10	(Parker 1982)	5	5	8
Hymenoptera Pompilidae	4000	3	60	(Parker 1982)	5	5	8
Hymenoptera Proctotrupidae	310	6	8	(Parker 1982)	5	5	8
Hymenoptera Pteromalidae	3506	1	4	(Parker 1982)	5	5	8
Hymenoptera Roproniidae	18	8	10	(Parker 1982)	5	5	8
Hymenoptera Rotoitidae	2	0.7	0.9	(Bouček & Noyes 1987)	5	5	8
Hymenoptera Sapygidae	80	6	22	(Parker 1982)	5	5	8
Hymenoptera Scelionidae	3000	0.5	15	(Parker 1982)	5	5	8
Hymenoptera Scolebythidae	3	7	10	(Cambra & Oliveira 2003)	5	5	8
Hymenoptera Scoliidae	300	8	60	(Parker 1982)	5	5	8
Hymenoptera Siricidae	95	20	40	(Parker 1982)	5	5	8
Hymenoptera Sierolomorphidae	10	3.5	6	(Parker 1982)	1	1	8
Hymenoptera Sphecidae	724	18	55	(Arnett 2000)	5	5	8
Hymenoptera Stenotritidae	21	14	20.5	(Houston 1983)	3	2	8
Hymenoptera Stephanidae	200	4	40	(Parker 1982)	5	5	8
Hymenoptera Tenthredinidae	4000	3	20	(Parker 1982)	3	3	3&4&8
Hymenoptera Tetracampidae	50	0.5	2	(Doganler 2003)	5	5	8
Hymenoptera Tiphiidae	1500	4	30	(Parker 1982)	5	5	8
Hymenoptera Torymidae	986	1	15	(Parker 1982)	3&5	3&5	8
Hymenoptera Trichogrammatidae	839	0.5	1	(Parker 1982)	5	5	8
Hymenoptera Trigonalidae	100	8	17	(Parker 1982)	5	5	8

Hymenoptera Vanhorniidae	5	3	10			(Arnett 2000)	5	5	8	
Hymenoptera Vespidae	4000	8	25			(Arnett 2000)	3&4	2&4	8	
Hymenoptera Xiphydriidae	100	7	25			(Parker 1982)	1	1	8	
Hymenoptera Xyelidae	50	5	15			(Parker 1982)	3	3	2&3	
Isoptera	2658	4	20			(Robinson 2005)	2	2	2	Sizes given based on winged forms
Lepidoptera Acanthopterocte tidae	5	4	6	11	16	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Acrolophidae	300	4	28	9	60	(Capinera 2008)	2	2	7	Data given as wingspan
Lepidoptera Adelidae	294	2	11	4	28	(Arnett 2000)	3	3	8	Data given as wingspan
Lepidoptera Agathiphagidae	2	3	5	9	14	(Capinera 2008)	3	3	?	Data given as wingspan
Lepidoptera Agonoxenidae	4	2	6	6	15	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Aididae	6	4	39	10	90	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Alucitidae	216	2	10	7	28	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Amphisbatidae	21	7	8	17	19	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Andesianidae	3	10	23	27	61	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Anomoeotidae	40	5	8	22	31	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Anthelidae	94	9	68	22	166	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Apatelodidae	145	10	37	20	74	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Arctiidae	6000	3	44	8	115	(Capinera 2008)	2&3	2&3	7	Data given as wingspan
Lepidoptera Argyresthiidae	157	2	5	6	15	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Arrhenophanida e	26	5	29	12	69	(Capinera 2008)	1	1	7	Data given as wingspan
Lepidoptera Autostichidae	585	4	7	10	20	(Capinera 2008)	2	2	8	Data given as wingspan
Lepidoptera Batrachedridae	99	2	9	7	28	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Blastobasidae	377	1	11	5	35	(Capinera 2008)	2	2	8	Data given as wingspan
Lepidoptera Bombycidae	185	10	33	19	64	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Brachodidae	137	3	17	8	42	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Brahmaeidae	44	18	66	50	180	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Bucculatricidae	297	2	5	5	16	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Callidulidae	49	8	14	22	38	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Carposinidae	283	3	14	10	40	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Carthaeidae	1	28	37	75	100	(Capinera 2008)	3	3	8	Data given as wingspan

Lepidoptera Castniidae	113	10	82	24	190	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Choreutidae	406	3	10	7	24	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Cimeliidae	6	8	11	22	28	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Coleophoridae	1386	2	7	5	24	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Copromorphidae	43	3	11	12	37	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Cosmopterigidae	1792	2	11	6	32	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Cossidae	971	5	136	9	240	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Crinopterygidae	1	2	2.5	3	3.5	(Kristensen et al. 2007)	3	3	8	Data given as forewing length
Lepidoptera Cyclotornidae	5	3	10	10	30	(Capinera 2008)	5	5	7	Data given as wingspan
Lepidoptera Dalceridae	80	4	18	11	50	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Douglasiidae	29	2	5	6	15	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Drepanidae	660	6	22	18	66	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Dudgeoneidae	57	11	29	28	72	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Elachistidae	3197	2	9	5	23	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Endromidae	56	11	27	29	74	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Epicopeiidae	20	13	38	36	126	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Epipyropidae	32	1	13	4	35	(Capinera 2008)	5	5	7	Data given as wingspan
Lepidoptera Eriocottidae	80	2	21	5	50	(Capinera 2008)	2	2	7	Data given as wingspan
Lepidoptera Eriocraniidae	28	2	5	6	13.5	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Eupterotidae	339	8	47	23	140	(Capinera 2008)	3	3	7&8	Data given as wingspan
Lepidoptera Gelechiidae	4700	1	12	4	35	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Geometridae	23002	3	42	8	120	(Capinera 2008)	3	3	7&8	Data given as wingspan
Lepidoptera Glyphidoceridae	49	5	7	13	19	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Glyphipterigidae	535	2	14	5	35	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Gracillariidae	1864	2	10	4	25	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Hedylidae	36	15	27	35	65	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Heliozelidae	123	1	3	3	9	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Hepialidae	604	8	104	20	250	(Capinera 2008)	1&3	1&3	7	Data given as wingspan
Lepidoptera Hesperiidae	4113	7	37	16	82	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Heterobathmiidae	3	3	4	10	11	(Capinera 2008)	3	3	2	Data given as wingspan
Lepidoptera	10	3	9	9	29	(Capinera	3	3	8	Data given

Heterogynidae						2008)				as wingspan
Lepidoptera Himantopteridae	40	5	12	16	42	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Hyblaeidae	18	11	22	25	49	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Immidae	245	5	14	14	42	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Incurvariidae	50	2	6	7	18	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Lacturidae	120	4	22	11	65	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Lasiocampidae	1952	10	92	19	172	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Lecithoceridae	1200	2	10	5	30	(Capinera 2008)	2	2	8	Data given as wingspan
Lepidoptera Lemoniidae	21	9	28	20	65	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Limacodidae	1672	4	35	9	80	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Lycaenidae	5201	2	33	6	92	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Lymantriidae	2500	7	58	16	135	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Lyonetiidae	220	2	5	4	12	(Capinera 2008)	3	3	8	Data given as wingspan, includes Bedelliidae
Lepidoptera Megalopygidae	232	5	44	10	90	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Micropterigidae	154	1	3	5	12	(Capinera 2008)	2&3	2&3	2	Data given as wingspan
Lepidoptera Mimallonidae	194	10	28	22	60	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Mnesarchaeidae	7	2	4	5	10	(Capinera 2008)	2&3	2&3	8	Data given as wingspan
Lepidoptera Momphidae	115	2	5	8	18	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Neopseustidae	14	3	6	14	27	(Capinera 2008)	?	?	8	Data given as wingspan
Lepidoptera Nepticulidae	806	1	2	2.5	8	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Noctuidae	30579	3	154	7	360	(Arnett 2000)/(Ohl & Thiele 2007)	2&3	2&3	8	Data given as wingspan
Lepidoptera Notodontidae	3800	9	57	20	124	(Capinera 2008)	3	3	7&8	Data given as wingspan
Lepidoptera Nymphalidae	6131	6	56	20	180	(Arnett 2000)/(Hogue 1993)	3	3	8	Data given as wingspan
Lepidoptera Oecophoridae	3304	2	30	5	80	(Capinera 2008)	2&3	2&3	8	Data given as wingspan
Lepidoptera Opostegidae	192	1	6	3	16	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Palaephatidae	57	3	11	8	36	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Papilionidae	566	10	83	35	285	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Pieridae	1164	7	31	23	100	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Plutellidae	150	2	17	7	55	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Prodoxidae	98	2	11	5	33	(Capinera 2008)	3	3	8	Data given as wingspan

Lepidoptera Prototheoridae	12	2	15	6	40	(Capinera 2008)	?	?	7	Data given as wingspan
Lepidoptera Psychidae	1324	2	37	4	60	(Arnett 2000)	2&3	2&3	7	Data given as wingspan
Lepidoptera Pterolonchidae	8	11	12	24	27	(Arnett 2000)	3	3	8	Data given as wingspan
Lepidoptera Pterophoridae	1318	2	15	6	40	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Pyralidae	5921	4	67	5	75	(Resh & Cardé 2009)	2&3	2&3	7&8	Data given as Forewing length, includes Crambidae
Lepidoptera Riodinidae	1532	8	13	20	35	(Arnett 2000)	3	3	8	Data given as wingspan
Lepidoptera Roeslerstammiidae	53	4	8	11	22	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Saturniidae	2349	12	117	30	300	(Capinera 2008)	3	3	7	Data given as wingspan
Lepidoptera Sematuridae	40	19	44	42	100	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Sesiidae	1397	5	34	5	28	(Resh & Cardé 2009)	3	3	7&8	Data given as Forewing length
Lepidoptera Somabrachyidae	8	7	8	18	22	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Sphingidae	1461	11	94	23	200	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Thyrididae	940	4	42	9	90	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Tineidae	2093	2	21	5	54	(Capinera 2008)	1&2	1&2	7	Data given as wingspan
Lepidoptera Tineodidae	19	7	16	15	34	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Tischeriidae	110	2	3	6	11	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Tortricidae	10387	3	24	7	60	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Uraniidae	686	10	50	31	160	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Urodidae	66	4	14	10	37	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Xyloryctidae	524	5	29	12	75	(Pohl et al. 2010)/(Zborowski & Edwards 2007)	3	3	8	Data given as wingspan
Lepidoptera Yponomeutidae	363	2	11	3.2	15	(Resh & Cardé 2009)	3	3	8	Data given as Forewing length
Lepidoptera Ypsolophidae	163	3	6	9	17	(Capinera 2008)	3	3	8	Data given as wingspan
Lepidoptera Zygaenidae	1036	4	40	5	50	(Resh & Cardé 2009)	3	3	8	Data given as Forewing length
Mantodea	2163	10	170			(Prete 1999)	4	4	4	
Mantophasmatodea	16	10	30			(Buder & Klass 2013)	4	4	4	
Mecoptera Apteropanorpidae	1	5.5	11			(Palmer & Siebke 2008)	?	?	2	
Mecoptera Bittacidae	214	14	34			(Parker 1982)	2	2	4	

Mecoptera Boreidae	38	2	7.5			(Parker 1982)	3	3	2&3	
Mecoptera Choristidae	12	11	14	13	17	(Riek 1973)	?	?	2	Data as forewing length
Mecoptera Meropeidae	2	10	12			(Arnett 2000)	?	?	2	
Mecoptera Nannochoristidae	9	5	9	6	12	(Byers 1989)	4	4	8	Data given as wingspan
Mecoptera Panorpidae	480	9	25			(Arnett 2000)	2	2	2	
Mecoptera Panorpodidae	19	7	17			(Byers 1990)	?	?	3	
Megaloptera Corydalidae	200	20	80			(Arnett 2000)	4	4	8	
Megaloptera Sialidae	70	13	18			(Arnett 2000)	4	4	2	
Neuroptera Ascalaphidae	430	40	80			(Arnett 2000)	4	4	4	
Neuroptera Berothidae	115	6	15	6	15	(Resh & Cardé 2009)	4	4	4	Data given as Forewing length
Neuroptera Chrysopidae	1200	10	25			(Arnett 2000)	4	4	4	
Neuroptera Coniopterygidae	450	2	3			(Arnett 2000)	4	4	4	
Neuroptera Hemerobiidae	550	6	12			(Arnett 2000)	4	4	4	
Neuroptera Ithonidae	53	21	40			(Arnett 2000)	2	2	?	
Neuroptera Mantispidae	400	20	35			(Arnett 2000)	5	5	4	
Neuroptera Myrmeleontidae	2100	40	80			(Arnett 2000)	4	4	2&4	
Neuroptera Nemopteridae	100	15	35	15	35	(Resh & Cardé 2009)	4	4	2	Data given as Forewing length
Neuroptera Nevrorthidae	12	6	10	6	10	(Resh & Cardé 2009)	?	?	?	Data given as Forewing length
Neuroptera Nymphidae	35	18	40	18	40	(Resh & Cardé 2009)	4	4	?	Data given as Forewing length
Neuroptera Osmylidae	160	15	30	15	30	(Resh & Cardé 2009)	?	?	?	Data given as Forewing length
Neuroptera Polystoechotidae	4	35	75			(Arnett 2000)	2	2	4	
Neuroptera Psychopsidae	26	10	35	10	35	(Resh & Cardé 2009)	4	4	?	Data given as Forewing length
Neuroptera Sisyridae	50	6	8			(Arnett 2000)	4	4	2&4	
Odonata Aeshnidae	428	50	100			(Garrison et al. 2006)	4	4	4	Data given as wingspan
Odonata Austropetaliidae	11	57	86			(Garrison et al. 2006)	4	4	4	
Odonata Chlorogomphidae	45	60	78			(Wilson undated)	4	4	4	
Odonata Cordulegastrida	51	55	88			(Garrison et al. 2006)	4	4	4	

e										
Odonata Corduliidae	285	28	68			(Garrison et al. 2006)	4	4	4	As subfamily of Libellulidae, Includes Synthemistidae
Odonata Gomphidae	945	25.5	90			(Garrison et al. 2006)	4	4	4	
Odonata Libellulidae	970	17	63			(Garrison et al. 2006)	4	4	4	
Odonata Macromiidae	123	56	91			(Garrison et al. 2006)	4	4	4	As subfamily of Libellulidae
Odonata Neopetaliidae	1	57	58			(Garrison et al. 2006)	4	4	4	
Odonata Petaluridae	11	54	88			(Garrison et al. 2006)	4	4	4	
Odonata Epiophlebiidae	2	48	60			(Fleck et al. 2013)	4	4	4	
Odonata Calopterygidae	172	45	60			(Esquivel 1997)	4	4	4	
Odonata Chlorocyphidae	143	26	30			(Serrano-Meneses et al. 2008)	4	4	4	
Odonata Chorismagrionidae	1	38	40			(Morton. 1914)	4	4	4	
Odonata Coenagrionidae	1104	16	60			(Silsby 2001)/(Howarth & Mull 1992)	4	4	4	
Odonata Diphlebiidae	9	45	55			(Stewart 1980)	4	4	4	
Odonata Euphaeidae	68	26	38			(Hayashi 1990)	4	4	4	
Odonata Hemiphlebiidae	1	23	25			(Rivera 2014)	4	4	4	
Odonata Isostictidae	45	15	40			(Watson 1974)	4	4	4	
Odonata Lestidae	150	40	75			(Esquivel 1997)	4	4	4	
Odonata Megapodagrionidae	285	40	75			(Esquivel 1997)	4	4	4	
Odonata Perilestidae	19	50	55			(Esquivel 1997)	4	4	4	
Odonata Platycnemididae	222	40	50			(Silsby 2001)	4	4	4	
Odonata Platystictidae	189	40	50			(Esquivel 1997)	4	4	4	
Odonata Polythoridae	58	30	40			(Esquivel 1997)	4	4	4	
Odonata Protoneuridae	240	30	35			(Esquivel 1997)	4	4	4	
Odonata Pseudostigmatidae	19	80	120			(Esquivel 1997)	4	4	4	
Odonata Synlestidae	33	35	60	50	85	(Picker et al. 2004)	4	4	4	Data given as wingspan
Orthoptera Acrididae	6016	9	120			(Parker 1982)	3	3	3	
Orthoptera Cylindrachetida	16	35	75			(Günther 1992)/(Bail	2	2	2	

e				ey 2007)				
Orthoptera Eumastacoidae	645	10	45	(Arnett 2000)	3	3	3	Includes Euschmidtiidae, Episactidae, Chorotypidae and Thericleidae
Orthoptera Lentulidae	35	12	25	(Parker 1982)	3	3	3	
Orthoptera Pamphagidae	448	30	90	(Parker 1982)	3	3	3	
Orthoptera Pneumoridae	17	11.5	100	(Parker 1982)	3	3	3	
Orthoptera Proscopiidae	214	25	165	(Parker 1982)	3	3	3	
Orthoptera Pyrgomorphidae	455	10	90	(Parker 1982)	3	3	3	
Orthoptera Romaleidae	465	18	80	(Arnett 2000)	3	3	3	
Orthoptera Tanaoceridae	3	10.3	25	(Parker 1982)	3	3	3	
Orthoptera Tetrigidae	1246	6	16	(Arnett 2000)	2	2	2	
Orthoptera Tridactylidae	201	4	15	(Naskrecki 2001)	2	2	2	Includes Rhipipterygidae
Orthoptera Trigonopterygidae	16	29	40	(Ng et al. 2011)	3	3	3	
Orthoptera Xyronotidae	4	17	30	(Parker 1982)	3	3	3	
Orthoptera Anostostomatidae	206	20	80	(Pratt et al. 2008)	2&4	2&4	2&3&4	
Orthoptera Gryllacrididae	675	7	50	(Arnett 2000)	3&4	3&4	3&4	
Orthoptera Gryllidae	4664	4	50	(Otte 2007)	2&3&4	2&3&4	2&3&4	
Orthoptera Gryllotalpidae	100	20	40	(Arnett 2000)	3&4	3&4	3&4	
Orthoptera Myrmecophilidae	8	2	4	(Arnett 2000)	2	2	2	
Orthoptera Prophalangopsidae	71	17	30	(Walker 2013)	3	3	3	
Orthoptera Rhaphidophoridae	497	10	30	(Richards 1968)/(Richards 1959)	2&4	2&4	2&4	
Orthoptera Stenopelmatidae	28	30	50	(Arnett 2000)	2&4	2&4	2&4	
Orthoptera Tettigoniidae	6827	5	90	(Rentz 2010)	3&4	3&4	3&4	
Phasmatodea Agathemeridae	8	40	70	(Zompro 2004)	3	3	3	
Phasmatodea Aschiphasmatidae	96	20	60	(Ng et al. 2011)	3	3	3	
Phasmatodea Bacillidae	54	40	110	(Scali et al. 2012) /(Picker et al. 2004)	3	3	3	
Phasmatodea	1210	17.5	140	(Zompro	3	3	3	

Diapheromeridae				1999)/ (Brock & Hasenpusch 2009)				
Phasmatodea Heteropterygidae	103	20	150	(Ng et al. 2011)	3	3	3	
Phasmatodea Phasmatidae	991	50	357	(Ng et al. 2011)/ (Hennemann & Conle 2008)	3	3	3	
Phasmatodea Phylliidae	51	24	90	(Zompro 2001)/(Ng et al. 2011)	3	3	3	
Phasmatodea Pseudophasmatoidea	406	17.5	250	(Zompro 1998)/(Picker et al. 2004)	3	3	3	Includes Heteronemiidae
Phasmatodea Timematidae	21	12	25	(Arnett 2000)	3	3	3	
Phthiraptera Boopidae	55	1.3	3.14	(Parker 1982)	6	5	6	
Phthiraptera Gyropidae	93	0.8	1	(Parker 1982)	6	5	6	
Phthiraptera Haematomyzidae	3	1.9	3	(Parker 1982)	6	5	6	
Phthiraptera Heptapsogasteridae	130	0.81	4.44	(Parker 1982)	6	5	6	
Phthiraptera Laemobothriidae	20	6.5	11	(Parker 1982)	6	5	6	
Phthiraptera Menoponidae	1039	1.1	6	(Parker 1982)	6	5	6	
Phthiraptera Philopteridae	2698	1.12	9.72	(Parker 1982)	6	5	6	
Phthiraptera Ricinidae	109	1.6	5.5	(Parker 1982)	6	5	6	
Phthiraptera Trichodectidae	362	0.92	2.73	(Parker 1982)	6	5	6	
Phthiraptera Anoplura	446	0.5	5	(Arnett 2000)	6	5	6	includes, Echinophthiriidae, Hoplopleuridae, Linognathidae, Pedicinidae, Pediculidae, Pthiridae and Polyplacidae
Plecoptera Austroperlidae	15	10	35	(Parker 1982)	2	2	?	
Plecoptera Capniidae	287	3	25	(Parker 1982)	2	2	2	
Plecoptera Chloroperlidae	187	6	40	(Parker 1982)	4	4	2	
Plecoptera Diamphipnoidae	6	25	45	(Parker 1982)	2	2	?	
Plecoptera Eustheniidae	23	15	35	(Parker 1982)	4	4	2	
Plecoptera	270	5	25	(Michaelis	2	2	2	

Gripopterygidae				et al. 2011)			
Plecoptera Leuctridae	360	6	13	(Arnett 2000)	2	2	2
Plecoptera Nemouridae	674	6	15	(Arnett 2000)	2	2	2
Plecoptera Notonemouridae	118	5	8	(Picker et al. 2004)	2	2	2
Plecoptera Peltoperlidae	69	34	49	(Arnett 2000)	2	2	7
Plecoptera Perlidae	965	10	50	(Parker 1982)	4	4	7
Plecoptera Perlodidae	310	8	50	(Parker 1982)	4	4	2&7
Plecoptera Pteronarcyidae	12	38	63	(Arnett 2000)	2	2	7
Plecoptera Scopuridae	8	16	25	(Jin & Bae 2005)	2	2	?
Plecoptera Taeniopterygidae	103	10	25	(Arnett 2000)	2	2	?
Protura	712	0.6	2.5	(Arnett 2000)	1	1	1
Psocoptera Amphientomidae	100	2.3	5	(New & Lienhard 2007)	2	2	2
Psocoptera Amphipsocidae	180	2.8	5.5	(New & Lienhard 2007)	2	2	2
Psocoptera Archipsocidae	81	1.2	1.8	(New & Lienhard 2007)	2	2	2
Psocoptera Caeciliusidae	566	2.5	4.5	(New & Lienhard 2007)	2	2	2
Psocoptera Calopsocidae	34	4.3	7	(New & Lienhard 2007)	2	2	2
Psocoptera Ectopsocidae	177	2	2.5	(New & Lienhard 2007)	2	2	2
Psocoptera Elipsocidae	129	2	2.6	(New & Lienhard 2007)	2	2	2
Psocoptera Epipsocidae	138	2.5	5.7	(New & Lienhard 2007)	2	2	2
Psocoptera Hemipsocidae	24	2.5	2.8	(New & Lienhard 2007)	2	2	2
Psocoptera Lachesillidae	271	1.8	2.2	(New & Lienhard 2007)	2	2	2
Psocoptera Lepidopsocidae	206	2	2.5	(New & Lienhard 2007)	2	2	2
Psocoptera Liposcelididae	181	1	1.5	(New & Lienhard 2007)	2	2	2
Psocoptera Mesopsocidae	75	3.8	4.2	(New & Lienhard 2007)	2	2	2
Psocoptera Myopsocidae	159	3	5	(Arnett 2000)	2	2	2
Psocoptera	87	1.4	1.8	(New &	2	2	2

Pachytroctidae						Lienhard 2007)				
Psocoptera Peripsocidae	235	2	4			(New & Lienhard 2007)	2	2	2	
Psocoptera Philotarsidae	111	2.2	3.8			(New & Lienhard 2007)	2	2	2	
Psocoptera Prionoglarididae	7	3	3.4			(New & Lienhard 2007)	2	2	2	
Psocoptera Pseudocaeciliidae	899	2.5	8			(New & Lienhard 2007)	2	2	2	
Psocoptera Psilopsocidae	300	1.9	3.2			(New & Lienhard 2007)	2	2	2	
Psocoptera Psocidae	7	3.2	5.4			(New & Lienhard 2007)	2	2	2	
Psocoptera Psoquillidae	27	1.1	2			(New & Lienhard 2007)	2	2	2	
Psocoptera Psyllipsocidae	26	1.3	2			(New & Lienhard 2007)	2	2	2	
Psocoptera Stenopsocidae	95	3.5	4.5			(New & Lienhard 2007)	2	2	2	
Psocoptera Trichopsocidae	11	2	2.5			(Arnett 2000)	2	2	2	
Psocoptera Troctopsocidae	22	1.4	4.1			(New & Lienhard 2007)	2	2	2	
Psocoptera Trogiidae	52	1.6	2.5			(New & Lienhard 2007)	2	2	2	
Raphidioptera	225	5	20	5	20	(Resh & Cardé 2009)	4	4	4	Data as Forewing length
Siphonaptera	2078	1	10			(Whiting et al. 2008)	2	2	6	
Strepsiptera	590	1	7.5			(Parker 1982)	5	5	7	
Thysanoptera Aeolothripidae	201	1.4	2.6			(Treherne 1919)	3&4	3&4	3&4	
Thysanoptera Heterothripidae	76	0.6	1.5			(Retana-Salazar 2009)	3	3	3	
Thysanoptera Phlaeothripidae	3532	2	14			(Lewis 1973)	1&3	1&3	1&3	
Thysanoptera Thripidae	2066	1	3			(Arnett 2000)	3&4	3&4	3&4	
Trichoptera Anomalopsychidae	27	4	8			(Holzenthal & Flint Jr 1995)	2	2	8	
Trichoptera Apataniidae	203	4	13	8	15	(Ivanov & Menshutkina 1996)	2	2	8	Data as Forewing length
Trichoptera Atriplectididae	6	7	10	20	28	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Beraeidae	57	4	5			(Arnett 2000)	2	2	8	
Trichoptera	111	6	11			(Arnett	2	2	8	

Brachycentridae						2000)				
Trichoptera Calamoceratidae	182	6	10	15	26	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Calocidae	23	2	10	5	25	(Arnett 2000) /(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Chathamiidae	5	6	9	15	22	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Conoesucidae	43	4	10	10	25	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Dipseudopsidae	114	4	14	4	16	(Olah & Johanson 2010)	2	2	8	Data as forewing length
Trichoptera Ecnomidae	469	2	7	6	18	(Neboiss 1986)	2&4	2&4	8	Data as wingspan
Trichoptera Glossosomatidae	682	3	10	8	12	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Goeridae	184	4	10	5	12	(Parker 1998)/(Greenhalgh & Ovenden 2004)	2	2	8	Data as forewing length
Trichoptera Helicophidae	44	3	6	8	15	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Helicopsychidae	269	4	6	10	16	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Hydrobiosidae	407	4	13	10	35	(Neboiss 1986)	4	4	8	Data as wingspan
Trichoptera Hydropsychidae	1808	3	21	8	56	(Neboiss 1986) /(Picker et al. 2004)	2&4	2&4	8	Data as wingspan
Trichoptera Hydroptilidae	2124	1.5	4	4	12	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Kokiriidae	15	5	9	14	24	(Neboiss 1986)	4	4	8	Data as wingspan
Trichoptera Lepidostomatidae	471	8	10			(Arnett 2000)	2	2	8	
Trichoptera Leptoceridae	2020	4	15	10	40	(Neboiss 1986)	2&4	2&4	8	Data as wingspan
Trichoptera Limnephilidae	880	7	23	25	40	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Limnocentropodidae	15	10	12	27	33	(Wiggins 1956)	4	4	?	Synonym Kitagamiidae, Data as wingspan
Trichoptera Molannidae	41	10	17			(Arnett 2000)	2&4	2&4	8	
Trichoptera Odontoceridae	154	5	14	14	-	(Neboiss 1986)/(Arnett 2000)	2&4	2&4	8	Minimum as wingspan/ body length
Trichoptera Oeconesidae	18	12	16	30	40	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Philopotamidae	1168	6	9	12	20	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Philorheithridae	30	6	13	16	35	(Neboiss 1986)	4	4	8	Data as wingspan
Trichoptera Phryganeidae	84	12	28	18	43	(Wiggins 1998)	2&4	2&4	8	Data as forewing length

Trichoptera Pisuliidae	19	6	19	-	40	(Morse 1974)/ (Picker et al. 2004)	2	2	8	Maximum as wingspan
Trichoptera Polycentropodidae	806	5	10	8	25	(Neboiss 1986)	4	4	8	Data as wingspan
Trichoptera Psychomyiidae	522	4	6			(Arnett 2000)	2	2	8	
Trichoptera Rhyacophilidae	774	8	13			(Arnett 2000)	4	4	8	
Trichoptera Sericostomatidae	107	8	14	20	35	(Picker et al. 2004)	2	2	8	Data as wingspan
Trichoptera Stenopsychidae	94	6	12	18	35	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Tasimiidae	9	4	6	12	18	(Neboiss 1986)	2	2	8	Data as wingspan
Trichoptera Uenoidae	31	7	9	8	10	(Houghton 2012)	2	2	8	Data given as forewing length
Trichoptera Xiphocentronidae	172	3	4	3	4	(Munoz-Quesada & Holzenthal 1997)	2	2	8	Data given as forewing length
Zoraptera	35	2	3			(Parker 1982)	1&4	1&4	1&4	
Zygentoma Lepidotrichidae	1	12	14			(Arnett 2000)/(Resh & Cardé 2009)	2	2	2	
Zygentoma Lepismatidae	200	8	20			(Arnett 2000)	2	2	2	
Zygentoma Nicoletiidae	30	4	29			(Arnett 2000)/(Espinasa et al. 2013)	2	2	2	

8. References

Adams, D.C. et al., 2009. Are rates of species diversification correlated with rates of morphological evolution? *Proceedings of the Royal Society B: Biological Sciences*, 276(1668), pp.2729–2738.

Agapow, P.-M. & Isaac, N.J.B., 2002. MacroCAIC: Revealing Correlates of Species Richness by Comparative Analysis. *Diversity and Distributions*, 8(1), pp.41–43.

Albert, J.S. & Johnson, D.M., 2012. Diversity and Evolution of Body Size in Fishes. *Evolutionary Biology*, 39(3), pp.324–340.

Alfaro, M.E. et al., 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences*, 106, pp.13410–13414.

Alfaro, M.E. & Holder, M.T., 2006. The Posterior and the Prior in Bayesian Phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, 37, pp.19–42.

Alfaro, M.E., Zoller, S. & Lutzoni, F., 2003. Bayes or Bootstrap? A Simulation Study Comparing the Performance of Bayesian Markov Chain Monte Carlo Sampling and Bootstrapping in Assessing Phylogenetic Confidence. *Molecular Biology and Evolution*, 20(2), pp.255 –266.

Allen, A.P. & Gillooly, J.F., 2006. Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecology Letters*, 9(8), pp.947–954.

Allen, C. et al., 2006. Patterns in body size distributions: sifting among alternative hypotheses. *Ecology Letters*, 9, pp.630–643.

Alroy, J. et al., 2008. Phanerozoic Trends in the Global Diversity of Marine Invertebrates. *Science*, 321(5885), pp.97–100.

Altschul, S.F. et al., 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), pp.403–410.

Andres, M.R. & Connor, E.F., 2003. The community-wide and guild-specific effects of pubescence on the folivorous insects of manzanitas Arctostaphylos spp. *Ecological Entomology*, 28(4), pp.383–396.

Andrew, D.R., 2011. A new view of insect–crustacean relationships II. Inferences from expressed sequence tags and comparisons with neural cladistics. *Arthropod Structure & Development*, 40(3), pp.289–302.

Ansorge, J., 2003. Upper Liassic Amphiesmenopterans (Trichoptera + Lepidoptera) from Germany – a review. *Acta zoologica cracoviensia,*, 46, pp.285–290.

Appeltans, W. et al., 2012. The Magnitude of Global Marine Species Diversity. *Current Biology*, 22(23), pp.2189–2202.

Arillo, A. & Engel, M.S., 2006. Rock crawlers in the Baltic amber (Notoptera :Mantophasmatodea). *American Museum Novitates*, 3539.

Aristov, D. & Bashkuev, A., 2008. New insects (Insecta: Mecoptera, Grylloblattida) from the Middle Permian Chepanikha locality, Udmurtia. *Paleontological Journal*, 42(2), pp.159–165.

Arnett, R.H. et al., 2010. *American Beetles, Volume II: Polyphaga: Scarabaeoidea through Curculionoidea*, CRC Press.

Arnett, R.H., 2000. *American Insects: A Handbook of the Insects of America north of Mexico*, Boca Raton, Fla.: CRC Press.

Arnett, R.H. & Thomas, M.C. eds., 2000. *American Beetles, Volume I: Archostemata, Myxophaga, Adephaga, Polyphaga: Staphyliniformia*, CRC Press.

Ascher, J.S. & Pickering, J., 2012. Discover life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila. Available at: http://www.discoverlife.org/mp/20q?guide=Apoidea_species.
Aspöck, U., Haring, E. & Aspöck, H., 2012. The phylogeny of the Neuropterida: long lasting and current controversies and challenges (Insecta: Endopterygota). *Arthropod Systematics and Phylogeny*, 70(2), pp.119–129.
Azar, D. & Nel, A., 2010. The earliest fossil schizopterid bug (Insecta: Heteroptera) in the Lower Cretaceous amber of Lebanon. *Annales de la Société Entomologique de France*, 46(1-2), pp.193–197.
Azevedo, C.O. & Azar, D., 2012. A new fossil subfamily of Bethylidae (Hymenoptera) from the Early Cretaceous Lebanese amber and its phylogenetic position. *Zoologia (Curitiba)*, 29(3), pp.210–218.
Azou, Y. & Bregliano, J.-C., 2001. I–R system of hybrid dysgenesis in Drosophila melanogaster: analysis of the mitochondrial DNA in reactive strains exhibiting different potentials for I factor transposition. *Heredity*, 86(1), pp.110–116.
Bae, Y. & McCafferty, W.P., 1998. Phylogenetic Systematics and Biogeography of the Neoephemeridae (Ephemeroptera: Pannota). *Aquatic Insects*, 20(1), pp.35–68.
Bailey, P.T., 2007. *Pests of Field Crops and Pastures: Identification and Control*, Csiro Publishing.
Bai, M., Ren, D. & Yang, X., 2012. Prosinodendron krelli from the Yixian Formation, China: A missing link among Lucanidae, Diphyllostomatidae and Passalidae (Coleoptera: Scarabaeoidea). *Cretaceous Research*, 34(0), pp.334–339.
Balashov, Y.S., 2006. Types of parasitism of acarines and insects on terrestrial vertebrates. *Entomological Review*, 86(8), pp.957–971.
Baldauf, S.L., 2003. Phylogeny for the faint of heart: a tutorial. *Trends in Genetics*, 19(6), pp.345–351.
Barber-James, H.M. et al., 2007. Global diversity of mayflies (Ephemeroptera, Insecta) in freshwater. *Hydrobiologia*, 595(1), pp.339–350.
Barnes, J.K., 1981. Revision of the Helosciomyzidae (Diptera). *Journal of the Royal Society of New Zealand*, 11(1), pp.45–72.
Barraclough, T.G., Barclay, M.V.. & Vogler, A.P., 1998. Species richness: Does flower power explain beetle-mania? *Current Biology*, 8(23), pp.R843–R845.
Barraclough, T.G., Harvey, P.H. & Nee, S., 1996. Rate of rbcL Gene Sequence Evolution and Species Diversification in Flowering Plants (Angiosperms). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263(1370), pp.589–591.
Barraclough, T.G., Harvey, P.H. & Nee, S., 1995. Sexual Selection and Taxonomic Diversity in Passerine Birds. *Proceedings of the Royal Society B: Biological Sciences*, 259(1355), pp.211–215.
Baum, E., Dressler, C. & Beutel, R.G., 2007. Head structures of Karoophasma sp. (Hexapoda, Mantophasmatodea) with phylogenetic implications. *Journal of Zoological Systematics and Evolutionary Research*, 45(2), pp.104–119.
Beaulieu, J., O'Meara, B.C. & Donoghue, M.J., 2013. Identifying Hidden Rate Changes in the Evolution of a Binary Morphological Character: The Evolution of Plant Habit in Campanulid Angiosperms. *Systematic Biology*, 62(5), pp.725–737.
Bebber, D.P. et al., 2007. Predicting unknown species numbers using discovery curves. *Proceedings of the Royal Society B: Biological Sciences*, 274(1618), pp.1651–1658.
Bechev, D. & Chandler, P., 2011. Catalogue of the Bolitophilidae and Diadocidiidae of the World (Insecta: Diptera). *Zootaxa*, 2741, pp.38–58.
Benson, D.A. et al., 2013. GenBank. *Nucleic Acids Research*, 41(Database issue), pp.D36–42.

Benson, R.B.J. & Mannion, P.D., 2011. Multi-variate models are essential for understanding vertebrate diversification in deep time. *Biology Letters*, p.rsbl20110460.
Benton, M.J., 2010. The origins of modern biodiversity on land. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1558), pp.3667–3679.
Benton, M.J. & Donoghue, P.C.J., 2007. Paleontological Evidence to Date the Tree of Life. *Molecular Biology and Evolution*, 24(1), pp.26 –53.
Benton, M.J. & Emerson, B.C., 2007. How Did Life Become so Diverse? the Dynamics of Diversification According to the Fossil Record and Molecular Phylogenetics. *Palaeontology*, 50(1), pp.23–40.
Benton, M.J. & Twitchett, R.J., 2003. How to kill (almost) all life: the end-Permian extinction event. *Trends in Ecology & Evolution*, 18(7), pp.358–365.
Berenbaum, M., 2007. Lend me your earwigs. *American Entomologist*, 53(4), pp.196–197.
Bergsten, J., 2005. A review of long-branch attraction. *Cladistics*, 21(2), pp.163–193.
Bernays, E. & Graham, M., 1988. On the Evolution of Host Specificity in Phytophagous Arthropods. *Ecology*, 69(4), p.886.
Bernhart, S. et al., 2008. RNAalifold: improved consensus structure prediction for RNA alignments. *BMC Bioinformatics*, 9, p.474.
Bertone, M.A., Courtney, G.W. & Wiegmann, B.M., 2008. Phylogenetics and temporal diversification of the earliest true flies (Insecta: Diptera) based on multiple nuclear genes. *Systematic Entomology*, 33(4), pp.668–687.
Béthoux, O., 2007. Archaeorthoptera wing venation nomenclature: A reply to Gorokhov. *Paleontological Journal*, 41(3), pp.338–340.
Béthoux, O. et al., 2011. At last, a Pennsylvanian stem-stonefly (Plecoptera) discovered. *BMC Evolutionary Biology*, 11(1), p.248.
Béthoux, O., 2009a. Gaps and nodes between fossil and extant insects. *Systematic Entomology*, 34(4), pp.599–609.
Béthoux, O. et al., 2002. Raphogla rubra gen. n. sp. n. the oldest representative of the clade of modern Ensifera (Orthoptera: Tettigoniidea, Gryllidea). *European Journal of Entomology*, 99, pp.111–116.
Béthoux, O., 2009b. The earliest beetle identified. *Journal of Paleontology*, 83(6), pp.931–937.
Béthoux, O. & Nel, A., 2002a. Venation pattern and revision of Orthoptera sensu nov. and sister groups. *Zootaxa*, 96, pp.1–88.
Béthoux, O. & Nel, A., 2002b. Venation pattern and revision of Orthoptera sensu nov. and sister groups. Phylogeny of Palaeozoic and Mesozoic Orthoptera sensu nov. *Zootaxa*, 96, pp.1–88.
Béthoux, O. & Wieland, F., 2009. Evidence for Carboniferous origin of the order Mantodea (Insecta: Dictyoptera) gained from forewing morphology. *Zoological Journal of the Linnean Society*, 156(1), pp.79–113.
Beutel, R.G. et al., 2011. Morphological and molecular evidence converge upon a robust phylogeny of the megadiverse Holometabola. *Cladistics*, 27(4), pp.341–355.
Beutel, R.G. et al., 2013. On the phylogeny and evolution of Mesozoic and extant lineages of Adephaga (Coleoptera, Insecta). *Cladistics*, 29(2), pp.147–165.
Beutel, R.G. & Gorb, S.N., 2006. A revised interpretation of the evolution of attachment structures in Hexapoda (Arthropoda), with special emphasis on Mantophasmatodea. *Arthropod Systematics and Phylogeny*, 64(1), pp.3–25.
Beutel, R.G. & Gorb, S.N., 2008. Evolutionary scenarios for unusual attachment devices of Phasmatodea and Mantophasmatodea (Insecta). *Systematic Entomology*, 33(3), pp.501–510.

Beutel, R.G. & Gorb, S.N., 2001. Ultrastructure of attachment specializations of hexapods (Arthropoda): evolutionary patterns inferred from a revised ordinal phylogeny. *Journal of Zoological Systematics and Evolutionary Research*, 39(4), pp.177–207.
Beutel, R.G. & Haas, F., 2000. Phylogenetic Relationships of the Suborders of Coleoptera (Insecta). *Cladistics*, 16(1), pp.103–141.
Beutel, R.G. & Leschen, R., 2005. *Coleoptera. Beetles. Volume 1: Morphology and Systematics (Archostemata, Adephaga, Myxophaga, Polyphaga partim)*, Walter de Gruyter.
Beutel, R.G. & Weide, D., 2005. Cephalic anatomy of Zorotypus hubbardi (Hexapoda: Zoraptera): new evidence for a relationship with Acercaria. *Zoomorphology*, 124, pp.121–136.
Bininda-Emonds, O.R.. et al., 2004. Garbage in, Garbage out; Data issues in supertree construction. In O. R. . Bininda-Emonds, ed. *Phylogenetic supertrees, combining information to reveal the tree of life*. Dordrecht, The Neatherlands: Kluwer Academic Publishers.
Bininda-Emonds, O.R.P. et al., 2007. The delayed rise of present-day mammals. *Nature*, 446(7135), pp.507–512.
Blackburn, T.M. & Gaston, K.J., 1994a. Animal Body Size Distributions Change as More Species are Described. *Proceedings of the Royal Society B: Biological Sciences*, 257(1350), pp.293–297.
Blackburn, T.M. & Gaston, K.J., 1994b. The Distribution of Body Sizes of the World's Bird Species. *Oikos*, 70(1), pp.127–130.
Blagoderov, V., Grimaldi, D.A. & Fraser, N., 2007. How time flies for flies: Diverse Diptera from the Triassic of Virginia and early radiation of the order. *American Museum Novitates*, 3572, pp.1–39.
Blanckenhorn, W., 2000. The evolution of body size: what keeps organisms small? *The Quarterly Review of Biology*, 75(4), pp.385–407.
Blanke, A., Greve, C., Mokso, R., et al., 2013. An updated phylogeny of Anisoptera including formal convergence analysis of morphological characters. *Systematic Entomology*, 38(3), pp.474–490.
Blanke, A. et al., 2012. Revival of Palaeoptera—head characters support a monophyletic origin of Odonata and Ephemeroptera (Insecta). *Cladistics*, 28(6), pp.560–581.
Blanke, A., Greve, C., Wipfler, B., et al., 2013. The Identification of Concerted Convergence in Insect Heads Corroborates Palaeoptera. *Systematic Biology*, 62(2), pp.250–263.
Blomberg, S.P. et al., 2012. Independent Contrasts and PGLS Regression Estimators Are Equivalent. *Systematic Biology*, p.syr118.
Blomberg, S.P., Garland, T. & Ives, A.R., 2003. Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits Are More Labile. *Evolution*, 57(4), pp.717–745.
Boeseman, M., 1954. *The Dermaptera in the museums at Leiden and Amsterdam*, Brill Academic Publishing.
Bokma, F., 2008. Bayesian Estimation of Speciation and Extinction Probabilities from (in)complete Phylogenies. *Evolution*, 62(9), pp.2441–2445.
Bokma, F., 2010. Time, Species, and Separating Their Effects on Trait Variance in Clades. *Systematic Biology*, 59(5), pp.602–607.
Bokma, F., Baek, S.K. & Minnhagen, P., 2014. 50 Years of Inordinate Fondness. *Systematic Biology*, 63(2), pp.251–256.
Bouček, Z. & Noyes, J.S., 1987. Rotoitidae, a curious new family of Chalcidoidea (Hymenoptera) from New Zealand. *Systematic Entomology*, 12(4), pp.407–412.

Bourgoin, T., 2012. FLOW (Fulgoromorpha Lists on The Web): a world knowledge base dedicated to Fulgoromorpha. Version 8 [October 2012]. Available at: http://hemiptera-databases.org/flow/.
Boussau, B. et al., 2014. Strepsiptera, Phylogenomics and the Long Branch Attraction Problem. *PLoS ONE*, 9(10), p.e107709.
Bradler, S., 2009. *Die Phylogenie der Stab- und Gespenstschrecken (Insecta: Phasmatodea)*, Universitätsverl.
Bradler, S. & Buckley, T.R., 2011. Stick insect on unsafe ground: does a fossil from the early Eocene of France really link Mesozoic taxa with the extant crown group of Phasmatodea? *Systematic Entomology*, 36(2), pp.218–222.
Bradler, S., Robertson, J.A. & Whiting, M.F., 2014. A molecular phylogeny of Phasmatodea with emphasis on Necrosciinae, the most species-rich subfamily of stick insects. *Systematic Entomology*, 39(2), pp.205–222.
Brake, I. & Mathis, W.N., 2007. Revision of the genus Australimyza Harrison (Diptera: Australimyzidae). *Systematic Entomology*, 32(2), pp.252–275.
Brändle, M., Stadler, J. & Brandl, R., 2000. Body size and host range in European Heteroptera. *Ecography*, 23(1), pp.139–147.
Brock, P.D. & Hasenpusch, J.W., 2009. *The Complete Field Guide to Stick and Leaf Insects of Australia*, Csiro Publishing.
Bromham, L. & Penny, D., 2003. The modern molecular clock. *Nature Reviews Genetics*, 4(3), pp.216–224.
Brown, B.V. et al., 2009. *Manual of Central American Diptera*, Ottawa: NRC Research Press.
Brown, J.H. et al., 2004. Towards a Metabolic Theory of Ecology. *Ecology*, 85(7), pp.1771–1789.
Brown, J.M. & Lemmon, A.R., 2007. The Importance of Data Partitioning and the Utility of Bayes Factors in Bayesian Phylogenetics. *Systematic Biology*, 56(4), pp.643 –655.
Brown, J.W. et al., 2012. MEDUSA: Modeling Evolutionary Diversification Using Stepwise AIC,. Available at: http://www.webpages.uidaho.edu/~lukeh/software/software.html [Accessed February 24, 2013].
Buckley, T.R., Attanayake, D. & Bradler, S., 2009. Extreme Convergence in Stick Insect Evolution: Phylogenetic Placement of the Lord Howe Island Tree Lobster. *Proceedings of the Royal Society B: Biological Sciences*, 276(1659), pp.1055–1062.
Buckman, R.S., Mound, L.A. & Whiting, M.F., 2013. Phylogeny of thrips (Insecta: Thysanoptera) based on five molecular loci. *Systematic Entomology*, 38(1), pp.123–133.
Buder, G. & Klass, K.-D., 2013. The morphology of tarsal processes in Mantophasmatodea. *Deutsche Entomologische Zeitschrift*, 60(1), pp.5–23.
Burleigh, J.G., Hilu, K.W. & Soltis, D.E., 2009. Inferring phylogenies with incomplete data sets: a 5-gene, 567-taxon analysis of angiosperms. *BMC Evolutionary Biology*, 9(1), p.61.
Burnham, K.P. & Anderson, D.R., 2002. *Model Selection and Multimodel Inference - A Practical Information-Theoretic Approach,* 2nd ed., New York: Springer-Verlag. Available at: http://www.springer.com/statistics/statistical+theory+and+methods/book/978-0-387-95364-9 [Accessed November 21, 2014].
Butler, M.A. & King, A.A., 2004. Phylogenetic Comparative Analysis: A Modeling Approach for Adaptive Evolution. *The American Naturalist*, 164(6), pp.683–695.
Bybee, S.M. et al., 2008. Molecules, morphology and fossils: a comprehensive approach to odonate phylogeny and the evolution of the odonate wing. *Cladistics*, 24(4), pp.477–514.

Byers, G.W., 1990. Brachypanorpa sacajawea n. sp. (Mecoptera: Panorpodidae) from the Rocky Mountains. *Journal of the Kansas Entomological Society*, 63(2), pp.211–217.
Byers, G.W., 1989. The Nannochoristidae of South America (Mecoptera). *University of Kansas Science Bulletin*, 54, pp.25–34.
Cai, C., Short, A.E.Z. & Huang, D.-Y., 2012. The first Skiff Beetle (Coleoptera: Myxophaga: Hydroscaphidae) from Early Cretaceous Jehol Biota. *Journal of Paleontology*, 86(1), pp.116–119.
Cambra, R.A. & Oliveira, A.., 2003. First Central American record of Clystopsenella longiventris (Hymenoptera: Scolebythidae) with comments on the variation of the species. *Entomotropica*, 18(2), pp.147–148.
Cameron, S.L. et al., 2009. A mitochondrial genome phylogeny of the Neuropterida (lace-wings, alderflies and snakeflies) and their relationship to the other holometabolous insect orders. *Zoologica Scripta*, 38(6), pp.575–590.
Cameron, S.L., 2014. Insect Mitochondrial Genomics: Implications for Evolution and Phylogeny. *Annual Review of Entomology*, 59(1), pp.95–117.
Cameron, S.L. et al., 2011. Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). *BMC Genomics*, 12(1), p.394.
Cameron, S.L., Barker, S.C. & Whiting, M.F., 2006. Mitochondrial genomics and the new insect order Mantophasmatodea. *Molecular Phylogenetics and Evolution*, 38(1), pp.274–279.
Campbell, B.C. et al., 1995. Paraphyly of Homoptera and Auchenorrhyncha inferred from 18S rDNA nucleotide sequences. *Systematic Entomology*, 20(3), pp.175–194.
Campbell, V. & Lapointe, F.-J., 2009. The Use and Validity of Composite Taxa in Phylogenetic Analysis. *Systematic Biology*, 58(6), pp.560–572.
Cantalapiedra, J.L. et al., 2014. Dietary innovations spurred the diversification of ruminants during the Caenozoic. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1776), p.20132746.
Capinera, J.L., 2008. *Encyclopedia of Entomology*, Springer.
Carapelli, A. et al., 2006. A review of molecular data for the phylogeny of basal hexapods. *Pedobiologia*, 50(2), pp.191–204.
Caravas, J. & Friedrich, M., 2013. Shaking the Diptera tree of life: performance analysis of nuclear and mitochondrial sequence data partitions. *Systematic Entomology*, 38(1), pp.93–103.
Carpenter, F.M., 1960. A Triassic Odonate From Argentina. *Psyche: A Journal of Entomology*, 67(3), pp.71–75.
Carpenter, F.M., 1953. The Geological History and Evolution of Insects. *American Scientist*, 41(2), pp.256–270.
Castro, L.R. & Dowton, M., 2005. The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of Perga condei (Hymenoptera: Symphyta: Pergidae). *Molecular Phylogenetics and Evolution*, 34(3), pp.469–479.
Cavender-Bares, J. et al., 2009. The merging of community ecology and phylogenetic biology. *Ecology Letters*, 12(7), pp.693–715.
Chan, K.M.A. & Moore, B.R., 2002. Whole-Tree Methods for Detecting Differential Diversification Rates. *Systematic Biology*, 51(6), pp.855 –865.
Chatzimanolis, S. et al., 2012. Leehermania prorova, the earliest staphyliniform beetle, from the Late Triassic of Virginia (Coleoptera: Staphylinidae). *American Museum Novitates*, 3761, pp.1–28.
Chenna, R. et al., 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Research*, 31(13), pp.3497–3500.

Cho, S. et al., 2011. Can Deliberately Incomplete Gene Sample Augmentation Improve a Phylogeny Estimate for the Advanced Moths and Butterflies (Hexapoda: Lepidoptera)? *Systematic Biology*, 60(6), pp.782–796.
Chown, S.L. & Gaston, K.J., 2010. Body size variation in insects: a macroecological perspective. *Biological Reviews*, 85(1), pp.139–169.
Christiansen, K. & Nascimbene, P., 2006. Collembola (Arthropoda, Hexapoda) from the mid Cretaceous of Myanmar (Burma). *Cretaceous Research*, 27(3), pp.318–363.
Clack, J.A., 2012. *Gaining Ground: The Origin and Evolution of Tetrapods*, Indiana University Press.
Clapham, M.E. & Karr, J.A., 2012. Environmental and biotic controls on the evolutionary history of insect body size. *Proceedings of the National Academy of Sciences*, 109(27), pp.10927–10930.
Clarke, J.T., Warnock, R.C.M. & Donoghue, P.C.J., 2011. Establishing a time-scale for plant evolution. *New Phytologist*, 192(1), pp.266–301.
Cohen, J.E. et al., 2005. Body sizes of hosts and parasitoids in individual feeding relationships. *Proceedings of the National Academy of Sciences*, 102(3), pp.684–689.
Colless, D., 1969. The genus Perissomma (Diptera : Perissommatidae) with new species from Australia and Chile. *Australian Journal of Zoology*, 17(4), pp.719–728.
Coll, M. & Guershon, M., 2002. Omnivory in Terrestrail Arthropods: Mixing Plant and Prey Diets. *Annual Review of Entomology*, 47(1), pp.267–297.
Comandi, S. et al., 2009. The complete mitochondrial genome of Atelura formicaria (Hexapoda: Zygentoma) and the phylogenetic relationships of basal insects. *Gene*, 439(1–2), pp.25–34.
Condamine, F.L., Rolland, J. & Morlon, H., 2013. Macroevolutionary perspectives to environmental change. *Ecology Letters*, 16, pp.72–85.
Connor, E.F. & Taverner, M.P., 1997. The Evolution and Adaptive Significance of the Leaf-Mining Habit. *Oikos*, 79(1), pp.6–25.
Consortium, i5K, 2013. The i5K Initiative: Advancing Arthropod Genomics for Knowledge, Human Health, Agriculture, and the Environment. *Journal of Heredity*, 104(5), pp.595–600.
Cooper, N., Jetz, W. & Freckleton, R.P., 2010. Phylogenetic comparative approaches for studying niche conservatism. *Journal of Evolutionary Biology*, 23(12), pp.2529–2539.
Costello, M.J., May, R.M. & Stork, N.E., 2013. Can We Name Earth's Species Before They Go Extinct? *Science*, 339(6118), pp.413–416.
Costello, M.J., Wilson, S. & Houlding, B., 2012. Predicting Total Global Species Richness Using Rates of Species Description and Estimates of Taxonomic Effort. *Systematic Biology*, 61(5), pp.871–883.
Coyne, J.A. & Orr, H.A., 2004. *Speciation*, Sinauer Associates, Incorporated Publishers.
Cryan, J.R., 2005. Molecular phylogeny of Cicadomorpha (Insecta: Hemiptera: Cicadoidea, Cercopoidea and Membracoidea): adding evidence to the controversy. *Systematic Entomology*, 30(4), pp.563–574.
Cryan, J.R. & Svenson, G.J., 2010. Family-level relationships of the spittlebugs and froghoppers (Hemiptera: Cicadomorpha: Cercopoidea). *Systematic Entomology*, 35(3), pp.393–415.
Cryan, J.R. & Urban, J.M., 2012. Higher-level phylogeny of the insect order Hemiptera: is Auchenorrhyncha really paraphyletic? *Systematic Entomology*, 37(1), pp.7–21.
Cunningham, C.W., Omland, K.E. & Oakley, T.H., 1998. Reconstructing ancestral character states: a critical reappraisal. *Trends in Ecology & Evolution*, 13(9), pp.361–366.
D'Agostino, R.B., 1970. Transformation to Normality of the Null Distribution of g1. *Biometrika*, 57(3), pp.679–681.

Damgaard, J. et al., 2008. Phylogeny of the Heelwalkers (Insecta: Mantophasmatodea) based on mtDNA sequences, with evidence for additional taxa in South Africa. *Molecular Phylogenetics and Evolution*, 47(2), pp.443–462.
Darilmaz, M.C. & Kiyak, S., 2011. A study of the family Spercheidae (Coleoptera) from Turkey. *Turkish Journal of Zoology*, 35(3), pp.441–444.
Davies, T.J., Barraclough, T.G., et al., 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences*, 101(7), pp.1904 –1909.
Davies, T.J., Savolainen, V., et al., 2004. Environmental energy and evolutionary rates in flowering plants. *Proceedings of the Royal Society B: Biological Sciences*, 271(1553), pp.2195–2200.
Davis, R.B. et al., 2013. Degree of Specialization Is Related to Body Size in Herbivorous Insects: A Phylogenetic Confirmation. *Evolution*, 67(2), pp.583–589.
Davis, R.B. et al., 2011. Fossil gaps inferred from phylogenies alter the apparent nature of diversification in dragonflies and their relatives. *BMC Evolutionary Biology*, 11(1), p.252.
Davis, R.B., Baldauf, S.L. & Mayhew, P.J., 2009. Eusociality and the success of the termites: insights from a supertree of dictyopteran families. *Journal of Evolutionary Biology*, 22(8), pp.1750–1761.
Davis, R.B., Baldauf, S.L. & Mayhew, P.J., 2010a. Many hexapod groups originated earlier and withstood extinction events better than previously realized: inferences from supertrees. *Proceedings of the Royal Society B: Biological Sciences*, 277(1687), pp.1597 – 1606.
Davis, R.B., Baldauf, S.L. & Mayhew, P.J., 2010b. The origins of species richness in the Hymenoptera: insights from a family-level supertree. *BMC Evolutionary Biology*, 10(1), p.109.
Debevec, A.H., Cardinal, S. & Danforth, B.N., 2012. Identifying the sister group to the bees: a molecular phylogeny of Aculeata with an emphasis on the superfamily Apoidea. *Zoologica Scripta*, 41(5), pp.527–535.
Deem, L.S., 2012. Dermaptera Species File. Version 1.0/4.1. [October 2012]. Available at: http://Dermaptera.SpeciesFile.org.
Deitz, L. et al., 2010. Treehoppers; Aetalionidae, Melizoderidae, and Membracidae (Hemiptera)[October 2012]. Available at: http://treehoppers.insectmuseum.org/public/site/treehoppers/home.
Dell'Ampio, E., Szucsich, N.U. & Pass, G., 2011. Protura and molecular phylogenetics: status quo of a young love. *Soil Organisms*, 83(3), pp.347–358.
Devictor, V. et al., 2010. Defining and measuring ecological specialization. *Journal of Applied Ecology*, 47(1), pp.15–25.
DeWalt, R.E., Neu-Becker, U. & Stueber, G., 2012. Plecoptera Species File Online. Version 1.0/4.1. [October 2012]. Available at: http://Plecoptera.SpeciesFile.org.
D'Haese, C.A., 2003. Morphological appraisal of Collembola phylogeny with special emphasis on Poduromorpha and a test of the aquatic origin hypothesis. *Zoologica Scripta*, 32(6), pp.563–586.
Dijkstra, K.-D.B. et al., 2014. Redefining the damselfly families: a comprehensive molecular phylogeny of Zygoptera (Odonata). *Systematic Entomology*, 39(1), pp.68–96.
Djernæs, M. et al., 2012. Phylogeny of cockroaches (Insecta, Dictyoptera, Blattodea), with placement of aberrant taxa and exploration of out-group sampling. *Systematic Entomology*, 37(1), pp.65–83.
Dodd, M.E., Silvertown, J. & Chase, M.W., 1999. Phylogenetic Analysis of Trait Evolution and Species Diversity Variation among Angiosperm Families. *Evolution*, 53(3), pp.732–744.

Doganler, M., 2003. Notes on the species of Tetracampidae with descriptions of some new species from Turkey. *Entomofuna*, 24(27), pp.381–396.
Dolphin, K. & Quicke, D.L.J., 2001. Estimating the global species richness of an incompletely described taxon: an example using parasitoid wasps (Hymenoptera: Braconidae). *Biological Journal of the Linnean Society*, 73(3), pp.279–286.
Dominguez, E. et al., 2006. *Ephemeroptera de América Del Sur*, Pensoft Publishers.
Donoghue, P.C.J. & Benton, M.J., 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends in Ecology & Evolution*, 22(8), pp.424–431.
Dorrington, G.E., 2012. On flying insect size and Phanerozoic atmospheric oxygen. *Proceedings of the National Academy of Sciences*, 109(50), pp.E3393–E3393.
Douady, C.J. et al., 2003. Comparison of Bayesian and Maximum Likelihood Bootstrap Measures of Phylogenetic Reliability. *Molecular Biology and Evolution*, 20(2), pp.248 – 254.
Douglas, A.E., 2009. The microbial dimension in insect nutritional ecology. *Functional Ecology*, 23(1), pp.38–47.
Douzery, E.J.P. et al., 2004. The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? *Proceedings of the National Academy of Sciences*, 101(43), pp.15386–15391.
Drake, J.M., 2003. The paradox of the parasites: implications for biological invasion. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(Suppl 2), pp.S133–S135.
Dressler, C. & Beutel, R.G., 2010. The morphology and evolution of the adult head of Adephaga (Insecta: Coleoptera). *Arthropod Systematics and Phylogeny*, 68(2), pp.239–287.
Drummond, A.J. et al., 2006. Relaxed Phylogenetics and Dating with Confidence. *PLoS Biology*, 4(5), p.e88.
Drummond, A.J. & Suchard, M.A., 2010. Bayesian random local clocks, or one rate to rule them all. *BMC Biology*, 8(1), p.114.
Drummond, A. & Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7, p.214.
Dudley, R., 2002. *The biomechanics of insect flight: form, function, evolution*, Princeton University Press.
Dumont, H.J., Vierstraete, A. & Vanfleteren, J.R., 2010. A molecular phylogeny of the Odonata (Insecta). *Systematic Entomology*, 35(1), pp.6–18.
Durden, L. & Musser, G., 1994. The sucking lice (insecta Anoplura) of the world: a Taxonomic checklist with records of mammalian hosts and geographical distributions. *Bulletin American Museam of Natural History*, 218, pp.1–90.
Eades, D.C., 2012. Polyneoptera Species File Online. Version 1.0/4.1. [October 2012]. Available at: http://Polyneoptera.SpeciesFile.org.
Early, J.W. et al., 2001. Maamingidae, a new family of proctotrupoid wasp (Insecta : Hymenoptera) from New Zealand. *Invertebrate Systematics*, 15(3), pp.341–352.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), pp.1792 –1797.
Edgar, R.C. & Batzoglou, S., 2006. Multiple sequence alignment. *Current Opinion in Structural Biology*, 16(3), pp.368–373.
Edmunds, G.F., Jensen, S.L. & Berner, L., 1976. *The Mayflies of North and Central America*, University of Minnesota Press.
Ehrlich, P.R. & Raven, P.H., 1964. Butterflies and Plants: A Study in Coevolution. *Evolution*, 18(4), pp.586–608.

Engel, M.S., 2006. A Note on the Relic Silverfish Tricholepidion gertschi (Zygentoma). *Transactions of the Kansas Academy of Science (1903-)*, 109(3/4), pp.236–238.
Engel, M.S., 2011. New earwigs in mid-Cretaceous amber from Myanmar (Dermaptera, Neodermaptera). *ZooKeys*, (130), pp.137–152.
Engel, M.S., Davis, S.R. & Prokop, J., 2013. Insect Wings: The Evolutionary Development of Nature's First Flyers. In A. Minelli, G. Boxshall, & F. Giuseppe, eds. *Arthropod Biology and Evolution*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 269–298.
Engel, M.S. & Grimaldi, D.A., 2004. New light shed on the oldest insect. *Nature*, 427(6975), pp.627–630.
Engel, M.S. & Grimaldi, D.A., 2006a. The earliest webspinners (Insecta: Embiodea). *American Museum Novitates*, 3514, pp.1–15.
Engel, M.S. & Grimaldi, D.A., 2006b. The first Cretaceous Spider Wasp (Hymenoptera: Pompilidae). *Journal of the Kansas Entomological Society*, 79(4), pp.359–368.
Engel, M.S., Grimaldi, D.A. & Krishna, K., 2007. Primitive termites from the Early Cretaceous of Asia (Isoptera). *Stuttgarter Beiträge zur Naturkunde Serie B (Geologie und Paläontologie)*, 371, pp.1–32.
Engel, M.S., Grimaldi, D.A. & Krishna, K., 2009. Termites (Isoptera): Their phylogeny, classification, and rise to ecological dominance. *American Museum Novitates*, 3650, pp.1–27.
Erixon, P. et al., 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Systematic Biology*, 52(5), pp.665–673.
Ertas, B. et al., 2009. Hemocyanin Suggests a Close Relationship of Remipedia and Hexapoda. *Molecular Biology and Evolution*, 26(12), pp.2711–2718.
Erwin, T.L., 1982. Tropical forests: their richness in Coleoptera and other arthropod species. *The Coleopterists Bulletin*, 36(1), pp.74–75.
Espinasa, L., Botelho, M. & Socci, Ki., 2013. A new species of genus Squamigera (Insecta: Zygentoma: Nicoletiidae) from the Mayan ruins of Palenque, Chiapas, Mexico. *Journal of Entomology and Nematology*, 52(2), pp.24–28.
Esquivel, C., 1997. The Families of Insects Costa Rica- Odonta. In A. Solís, ed. *Las familias de insectos de Costa Rica*. Available at: http://www.inbio.ac.cr/papers/insectoscr/Odonata.html.
Etienne, R.S. et al., 2011. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proceedings of the Royal Society B: Biological Sciences*, 279(1732), pp.1300–1309.
Evans, J.W., 1966. The leafhoppers and froghoppers of Australia and New Zealand (Homoptera: Cicadelloidea and Cercopoidea). *Australian Museum Memoir*, 12, pp.1–347.
Evenhius, N.L., 2011. family Xenasteiidae. In *Evenhuis, N.L. (ed.), Catalog of the Diptera of the Australasian and Oceanian Regions. (online version)*. Available at: http://hbs.bishopmuseum.org/aocat/hybotidae.html.
Farrell, B.D., 1998. "Inordinate fondness" explained: Why are there so many beetles? *Science*, 281(5376), pp.555 –559.
Farrell, B.D., Dussourd, D.E. & Mitter, C., 1991. Escalation of Plant Defense: Do Latex and Resin Canals Spur Plant Diversification? *The American Naturalist*, 138(4), pp.881–900.
Felsenstein, J., 2012. A Comparative Method for Both Discrete and Continuous Characters Using the Threshold Model. *The American Naturalist*, 179(2), pp.145–156.
Felsenstein, J., 2008. Comparative Methods with Sampling Error and Within-Species Variation: Contrasts Revisited and Revised. *The American Naturalist*, 171(6), pp.713–725.
Felsenstein, J., 1985. Phylogenies and the Comparative Method. *The American Naturalist*, 125(1), pp.1–15.

Fikáček, M. et al., 2012a. Phylogeny and the fossil record of the Helophoridae reveal Jurassic origin of extant hydrophiloid lineages (Coleoptera: Polyphaga). *Systematic Entomology*, 37(3), pp.420–447.
Fikáček, M. et al., 2012b. Revision of Mesozoic fossils of the helophorid lineage of the superfamily Hydrophiloidea (Coleoptera: Polyphaga). *Acta Entomologica Musei Nationalis Pragae*, 52(1), pp.89–127.
Finarelli, J.A., 2007. Mechanisms behind Active Trends in Body Size Evolution of the Canidae (Carnivora: Mammalia). *The American Naturalist*, 170(6), pp.876–885.
Finarelli, J.A. & Goswami, A., 2013. Potential Pitfalls of Reconstructing Deep Time Evolutionary History with Only Extant Data, a Case Study Using the Canidae (Mammalia, Carnivora). *Evolution*, 67(12), pp.3678–3685.
Finlay, B.J. et al., 2006. Self-similar patterns of nature: insect diversity at local to global scales. *Proceedings of the Royal Society B: Biological Sciences*, 273(1596), pp.1935–1941.
FitzJohn, R.G., 2010. Quantitative Traits and Diversification. *Systematic Biology*, 59(6), pp.619–633.
FitzJohn, R.G., Maddison, W.P. & Otto, S.P., 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology*, 58(6), pp.595 –611.
Fiz-Palacios, O. et al., 2011. Diversification of land plants: insights from a family-level phylogenetic analysis. *BMC Evolutionary Biology*, 11, p.341.
Fleck, G. et al., 2008. A phylogeny of anisopterous dragonflies (Insecta, Odonata) using mtRNA genes and mixed nucleotide/doublet models. *Journal of Zoological Systematics and Evolutionary Research*, 46(4), pp.310–322.
Fleck, G. et al., 2013. Epiophlebia sinensis Li & Nel 2011 in Li et al. (2012) (Odonata) newly recorded in North Korea. *International Dragonfly Fund Report*, 61, pp.1–4.
Foote, M., 1988. Survivorship Analysis of Cambrian and Ordovician Trilobites. *Paleobiology*, 14(3), pp.258–271.
Foottit, R.G. & Adler, P.H., 2009. *Insect Biodiversity: Science and Society*, John Wiley & Sons.
Forister, M.L. et al., 2011. Revisiting the evolution of ecological specialization, with emphasis on insect–plant interactions. *Ecology*, 93(5), pp.981–991.
Forister, M.L. et al., 2015. The global distribution of diet breadth in insect herbivores. *Proceedings of the National Academy of Sciences*, 112(2), pp.442–447.
Fraser, N. et al., 1996. A Triassic Lagerstatte from eastern North America. *Nature*, 380, pp.615–619.
Fraulob, M. et al., 2012. The larval abdomen of the enigmatic Nannochoristidae (Mecoptera, Insecta). *Arthropod Structure & Development*, 41(2), pp.187–198.
Freckleton, R.P. & Harvey, P.H., 2006. Detecting Non-Brownian Trait Evolution in Adaptive Radiations. *PLoS Biol*, 4(11), p.e373.
Freckleton, R.P., Phillimore, A.B. & Pagel, M., 2008. Relating Traits to Diversification: A Simple Test. *The American Naturalist*, 172(1), pp.102–115.
Friedemann, K. et al., 2012. On the head morphology of Phyllium and the phylogenetic relationships of Phasmatodea (Insecta). *Acta Zoologica*, 93(2), pp.184–199.
Friedrich, F. et al., 2014. Insect morphology in the age of phylogenomics: innovative techniques and its future role in systematics. *Entomological Science*, 17(1), pp.1–24.
Friedrich, F. & Beutel, R.G., 2010a. Goodbye Halteria? The thoracic morphology of Endopterygota (Insecta) and its phylogenetic implications. *Cladistics*, 26(6), pp.579–612.
Friedrich, F. & Beutel, R.G., 2010b. The thoracic morphology of Nannochorista (Nannochoristidae) and its implications for the phylogeny of Mecoptera and Antliophora. *Journal of Zoological Systematics and Evolutionary Research*, 48(1), pp.50–74.

Friedrich, F. & Beutel, R.G., 2008. The thorax of Zorotypus (Hexapoda, Zoraptera) and a new nomenclature for the musculature of Neoptera. *Arthropod Structure & Development*, 37(1), pp.29–54.
Friedrich, F., Farrell, B.D. & Beutel, R.G., 2009. The thoracic morphology of Archostemata and the relationships of the extant suborders of Coleoptera (Hexapoda). *Cladistics*, 25(1), pp.1–37.
Fritz, S.A. et al., 2013. Diversity in time and space: wanted dead and alive. *Trends in Ecology & Evolution*, 28(9), pp.509–516.
Futuyma, D.J. & Agrawal, A.A., 2009. Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences*, 106(43), pp.18054 –18061.
Futuyma, D.J. & Moreno, G., 1988. The Evolution of Ecological Specialization. *Annual Review of Ecology and Systematics*, 19, pp.207–233.
Galili, T., 2010. Post hoc analysis for Friedman's Test (R code). *R-statistics blog*. Available at: http://www.r-statistics.com/tag/friedman-test/ [Accessed March 4, 2015].
Gandolfo, M.A., Nixon, K.C. & Crepet, W.L., 2008. Selection of Fossils for Calibration of Molecular Dating Models. *Annals of the Missouri Botanical Garden*, 95(1), pp.34–42.
Gao, T. et al., 2012. Mid-Mesozoic flea-like ectoparasites of feathered or haired vertebrates. *Current Biology*, 22(8), pp.732–735.
Gao, T. et al., 2013. New Transitional Fleas from China Highlighting Diversity of Early Cretaceous Ectoparasitic Insects. *Current Biology*, 23(13), pp.1261–1266.
Gao, Y., Bu, Y. & Luan, Y.-X., 2008. Phylogenetic Relationships of Basal Hexapods Reconstructed from Nearly Complete 18S and 28S rRNA Gene Sequences. *Zoological Science*, 25, pp.1139–1145.
Garamszegi, L.Z. & Møller, A.P., 2010. Effects of sample size and intraspecific variation in phylogenetic comparative studies: a meta-analytic review. *Biological Reviews*, 85(4), pp.797–805.
Gardezi, T. & Silva, J. da, 1999. Diversity in Relation to Body Size in Mammals: A Comparative Study. *The American Naturalist*, 153(1), pp.110–123.
Garland, T., Bennett, A.F. & Rezende, E.L., 2005. Phylogenetic approaches in comparative physiology. *Journal of Experimental Biology*, 208(16), pp.3015–3035.
Garland, T., Harvey, P.H. & Ives, A.R., 1992. Procedures for the Analysis of Comparative Data Using Phylogenetically Independent Contrasts. *Systematic Biology*, 41(1), pp.18–32.
Garrison, R.W., Ellenrieder, N. von & Louton, J.A., 2006. *Dragonfly Genera of the New World: An Illustrated and Annotated Key to the Anisoptera*, JHU Press.
Garrouste, R. et al., 2012. A complete insect from the Late Devonian period. *Nature*, 488(7409), pp.82–85.
Gaston, K.J., 1991a. Body size and probability of description: the beetle fauna of Britain. *Ecological Entomology*, 16(4), pp.505–508.
Gaston, K.J., 1991b. The magnitude of global insect species richness. *Conservation Biology*, 5(3), pp.283–296.
Gaston, K.J. & Blackburn, T., 2000. *Pattern and Process in Macroecology* First Edition., Oxford, OX ; Malden, MA, USA: Wiley-Blackwell.
Gaston, K.J. & Blackburn, T.M., 1994. Are Newly Described Bird Species Small-Bodied? *Biodiversity Letters*, 2(1), pp.16–20.
Gaston, K.J. & Chown, S.L., 2013. Macroecological Patterns in Insect Body Size. In F. Smith & S. K. Lyons, eds. *Animal Body Size: Linking Pattern and Process across Space Time and Taxonomic group*. University of Chicago Press, pp. 13–61.

Gaunt, M.W. & Miles, M.A., 2002. An Insect Molecular Clock Dates the Origin of the Insects and Accords with Paleontological and Biogeographic Landmarks. *Molecular Biology and Evolution*, 19(5), pp.748 –761.
Gibson, G. a. P., Read, J. & Huber, J.T., 2007. Diversity, classification and higher relationships of Mymarommatoidea (Hymenoptera). Available at: http://www.cfs.nrcan.gc.ca/publications?id=34849&lang=en_CA [Accessed April 4, 2014].
Giller, P.S., 1996. The diversity of soil communities, the "poor man"s tropical rainforest'. *Biodiversity & Conservation*, 5(2), pp.135–168.
Gillespie, J.J. et al., 2005. Assessing the odd secondary structural properties of nuclear small subunit ribosomal RNA sequences (18S) of the twisted-wing parasites (Insecta: Strepsiptera). *Insect Molecular Biology*, 14(6), pp.625–643.
Gillies, M.T., 1980. The african Euthyplociidae (Ephemeroptera), (Exeuthyplociinae subfam. n.). *Aquatic Insects*, 2(4), pp.217–224.
Giribet, G. et al., 2004. Is Ellipura monophyletic? A combined analysis of basal hexapod relationships with emphasis on the origin of insects. *Organisms Diversity & Evolution*, 4(4), pp.319–340.
Giribet, G. & Edgecombe, G.D., 2012. Reevaluating the Arthropod Tree of Life. *Annual Review of Entomology*, 57(1), pp.167–186.
Giribet, G. & Wheeler, W.C., 1999. On Gaps. *Molecular Phylogenetics and Evolution*, 13(1), pp.132–143.
Gittleman, J.L. & Purvis, A., 1998. Body size and species–richness in carnivores and primates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1391), pp.113–119.
Godfray, H.C.J., 1994. *Parasitoids: Behavioral and Evolutionary Ecology*, Princeton University Press.
Goldberg, E.E., Lancaster, L.T. & Ree, R.H., 2011. Phylogenetic Inference of Reciprocal Effects between Geographic Range Evolution and Diversification. *Systematic Biology*, 60(4), pp.451–465.
Gómez-Zurita, J. et al., 2007. Recalibrated Tree of Leaf Beetles (Chrysomelidae) Indicates Independent Diversification of Angiosperms and Their Insect Herbivores. *PLoS ONE*, 2(4), p.e360.
Gorochov, A.V., 2005. Review of Triassic Orthoptera with descriptions of new and little known taxa: Part 2. *Paleontological Journal,*, 39, pp.272–279.
Gorochov, A.V., Jarzembowski, E.A. & Coram, R.A., 2006. Grasshoppers and crickets (Insecta: Orthoptera) from the Lower Cretaceous of southern England. *Cretaceous Research*, 27(5), pp.641–662.
Gould, S.J. & Calloway, C.B., 1980. Clams and Brachiopods-Ships that Pass in the Night. *Paleobiology*, 6(4), pp.383–396.
Gouws, E.J., Gaston, K.J. & Chown, S.L., 2011. Intraspecific Body Size Frequency Distributions of Insects. *PLoS ONE*, 6(3), p.e16606.
Gradstein, F.M., Ogg, G. & Schmitz, M., 2012. *The Geologic Time Scale 2012*, Elsevier.
Gratshev, V. & Legalov, A., 2011. New Mesozoic Ithyceridae beetles (Coleoptera). *Paleontological Journal*, 45(1), pp.77–82.
Graur, D. & Martin, W., 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics*, 20(2), pp.80–86.
Greenhalgh, M. & Ovenden, M.G. and D., 2004. *Malcolm Greenhalghs Flyfishers Handbook*, Coch Y Bonddu Books.

Greenslade, P. & Whalley, P.E.S., 1996. The systematic position of Rhyniella Praecursor Hirst and Maulik (Collembola). The earliest known Hexapod. In R. Dallai, ed. *2nd International Seminar on Apterygota*. University of Siena.
Grimaldi, D.A., 2010. 400 million years on six legs: On the origin and early evolution of Hexapoda. *Arthropod Structure & Development*, 39(2–3), pp.191–203.
Grimaldi, D.A., 2003. A revision of Cretaceous mantises and their relationships, including new taxa (Insecta, Dictyoptera, Mantodea). *American Museum Novitates*, 3412, pp.1–47.
Grimaldi, D.A., 1999. The Co-Radiations of Pollinating Insects and Angiosperms in the Cretaceous. *Annals of the Missouri Botanical Garden*, 86(2), p.373.
Grimaldi, D.A. & Engel, M.S., 2005. *Evolution of the Insects*, Cambridge University Press.
Grimaldi, D.A. & Engel, M.S., 2006a. Extralimital fossils of the "Gondwanan" family Sphaeropsocidae (Insecta, Psocodea). *American Museum Novitates*, 3523, pp.1–18.
Grimaldi, D.A. & Engel, M.S., 2006b. Fossil Liposcelididae and the lice ages (Insecta: Psocodea). *Proceedings of the Royal Society B: Biological Sciences*, 273(1586), pp.625–633.
Günther, K.K., 1992. Revision der Familie Cylindrachetidae Giglio-Tos, 1914 (Orthoptera, Tridactyloidea). *Deutsche Entomologische Zeitschrift*, 39(4-5), pp.233–291.
Haas, F. & Kukalová-Peck, J., 2001. Dermaptera hindwing structure and folding: New evidence for familial, ordinal and superordinal relationships within Neoptera (Insecta). *European Journal of Entomology*, 98, pp.445–509.
Hallinan, N., 2012. The generalized time variable reconstructed birth–death process. *Journal of Theoretical Biology*, 300, pp.265–276.
Hamilton, A.J. et al., 2010. Quantifying Uncertainty in Estimation of Tropical Arthropod Species Richness. *The American Naturalist*, 176(1), pp.90–95.
Hanley, M.E. et al., 2007. Plant structural traits and their role in anti-herbivore defence. *Perspectives in Plant Ecology, Evolution and Systematics*, 8(4), pp.157–178.
Hao, W. & Huang, D.-Y., 2012. A New Species of Liadotaulius (Insecta: Trichoptera) from the Middle Jurassic of Daohugou, Inner Mongolia. *Acta Geologica Sinica - English Edition*, 86(2), pp.320–324.
Hardy, N. & Cook, L., 2010. Gall-induction in insects: evolutionary dead-end or speciation driver? *BMC Evolutionary Biology*, 10(1), p.257.
Haring, E. & Aspöck, U., 2004. Phylogeny of the Neuropterida: a first molecular approach. *Systematic Entomology*, 29(3), pp.415–430.
Harmon, L.J. et al., 2010. Early Bursts of Body Size and Shape Evolution Are Rare in Comparative Data. *Evolution*, 64(8), pp.2385–2396.
Harmon, L.J. et al., 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics*, 24(1), pp.129–131.
Harmon, L.J. & Losos, J.B., 2005. The Effect of Intraspecific Sample Size on Type I and Type Ii Error Rates in Comparative Studies. *Evolution*, 59(12), pp.2705–2710.
Harrison, J.F., Kaiser, A. & VandenBrooks, J.M., 2010. Atmospheric oxygen level and the evolution of insect body size. *Proceedings of the Royal Society B: Biological Sciences*, 277(1690), pp.1937–1946.
Hartmann, S. & Vision, T.J., 2008. Using ESTs for phylogenomics: Can one accurately infer a phylogenetic tree from a gappy alignment? *BMC Evolutionary Biology*, 8(1), p.95.
Harzsch, S., 2006. Neurophylogeny: Architecture of the nervous system and a fresh view on arthropod phyologeny. *Integrative and Comparative Biology*, 46(2), pp.162–194.
Hawkins, B.A., 2005. *Pattern and Process in Host-Parasitoid Interactions*, Cambridge University Press.
Hayashi, F., 1990. Convergence of insular dwarfism in damselflies (Euphaea) and dobsonflies (Protohermes). *Freshwater Biology*, 23(2), pp.219–231.

Heads, S.W., 2008. The First Fossil Proscopiidae (Insecta Orthoptera Eumastacoidea) with comments on the historical biogeography and evolution of the family. *Palaeontology*, 51(2), pp.499–507.
Heckman, C.W., 2002. *Encyclopedia of South American Aquatic Insects: Ephemeroptera: Illustrated Keys to Known Families, Genera, and Species in South America.*, Springer.
Heikkilä, M. et al., 2011. Cretaceous origin and repeated tertiary diversification of the redefined butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 279, pp.1093–1099.
Heled, J. & Drummond, A.J., 2011. Calibrated tree priors for relaxed phylogenetics and divergence time estimation. *Systematic Biology*, 61(1), pp.138–149.
Hennemann, F.H. & Conle, O.V., 2008. Revision of Oriental Phasmatodea: The tribe Pharnaciini Günther, 1953, including the description of the world's longest insect, and a survey of the family Phasmatidae Gray, 1835 with keys to the subfamilies and tribes (Phasmatodea: "Anareolatae": Phasmatidae). *Zootaxa*, 1906, pp.1–316.
Hennig, W., 1969. *Stammesgeschichte der Insekten*, Kramer, Frankfurt am Main.
Heraty, J. et al., 2011. Evolution of the hymenopteran megaradiation. *Molecular Phylogenetics and Evolution*, 60(1), pp.73–88.
Hill, C.J. & Pierce, N.E., 1989. The effect of adult diet on the biology of butterflies. *Oecologia*, 81(2), pp.249–257.
Hirst, S. & Maulik, S., 1926. On Some Arthropod Remains from the Rhynie Chert (Old Red Sandstone). *Geological Magazine*, 63(02), pp.69–71.
Hisamatsu, S., 2011. A review of the Japanese Kateretidae fauna (Coleoptera: Cucujoidea). *Acta Entomologica Musei Nationalis Pragae*, 51(2), pp.551–585.
Hitchings, T.R. & Staniczek, A.H., 2003. Nesameletidae (Insecta: Ephemeroptera). *Fauna of New Zealand*, 46(0). Available at: http://www.biotaxa.org/fnz/article/view/1818 [Accessed April 3, 2014].
Hodkinson, I.D. & Casson, D.S., 2000. Patterns within patterns: abundance-size relationships within the Hemiptera of tropical rain forest or why phylogeny matters. *Oikos*, 88(3), pp.509–514.
Hogue, C.L., 1993. *Latin American Insects and Entomology*, University of California Press.
Holder, M. & Lewis, P.O., 2003. Phylogeny estimation: traditional and Bayesian approaches. *Nature Reviews Genetics*, 4(4), pp.275–284.
Ho, L.S.T. & Ane, C., 2014. A Linear-Time Algorithm for Gaussian and Non-Gaussian Trait Evolution Models. *Systematic Biology*, 63(3), pp.397–408.
Holzenthal, R.W. et al., 2007. Order Trichoptera Kirby 1813 (Insecta), Caddisflies. *Zootaxa*, 1668, pp.639–698.
Holzenthal, R.W. & Flint Jr, O., 1995. Studies of Neotropical Caddisflies, LI : Systematics of th e Neotropical Caddisfly Genus Contulma (Trichoptera : Anomalopsychidae).
Hone, D.W.E. & Benton, M.J., 2005. The evolution of large size: how does Cope's Rule work? *Trends in Ecology & Evolution*, 20(1), pp.4–6.
Hopkin, S.P., 1997. *Biology of the Springtails : (Insecta: Collembola)*, Oxford University Press.
Ho, S.Y.W., 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology*, 38(4), pp.409–414.
Ho, S.Y.W. & Jermiin, L.S., 2004. Tracing the Decay of the Historical Signal in Biological Sequence Data. *Systematic Biology*, 53(4), pp.623–637.
Ho, S.Y.W. & Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, 58(3), pp.367 –380.

Houghton, D.C., 2012. Biological diversity of the Minnesota caddisflies (Insecta, Trichoptera). *ZooKeys*, (189), pp.1–389.
Houston, T.F., 1983. A Revision of the Bee Genus Ctenocolletes (Hymenoptera: Stenotritidae). *Records of the Western Australian Museam*, 10(3), pp.269–306.
Hovmöller, R., Pape, T. & Källersjö, M., 2002. The Palaeoptera Problem: Basal Pterygote Phylogeny Inferred from 18S and 28S rDNA Sequences. *Cladistics*, 18(3), pp.313–323.
Howarth, F.G. & Mull, W.P., 1992. *Hawaiian Insects and Their Kin*, University of Hawaii Press.
Hozo, S.P., Djulbegovic, B. & Hozo, I., 2005. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology*, 5(1), p.13.
Hua, J. et al., 2008. Comparative and phylogenomic studies on the mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera). *BMC Genomics*, 9(1), pp.1–15.
Hua, J. et al., 2009. Phylogenetic analysis of the true water bugs (Insecta: Hemiptera: Heteroptera: Nepomorpha): evidence from mitochondrial genomes. *BMC Evolutionary Biology*, 9(1), p.134.
Huang, D.-Y. et al., 2012. Diverse transitional giant fleas from the Mesozoic era of China. *Nature*, 483, pp.210–204.
Huang, D.-Y., Nel, A., Zompro, O., et al., 2008. Mantophasmatodea now in the Jurassic. *Naturwissenschaften*, 95(10), pp.947–952.
Huang, D.-Y. & Nel, A., 2009a. First Chinese Cymatophlebiidae from the Middle Jurassic of Inner Mongolia (Odonata: Anisoptera: Aeshnoptera). *Palaeodiversity*, 2, pp.199–204.
Huang, D.-Y. & Nel, A., 2009b. Oldest webspinners from the Middle Jurassic of Inner Mongolia, China (Insecta: Embiodea). *Zoological Journal of the Linnean Society*, 156(4), pp.889–895.
Huang, D.-Y., Nel, A. & Minet, J., 2010. A New Family of Moths from the Middle Jurassic (Insecta: Lepidoptera). *Acta Geologica Sinica - English Edition*, 84(4), pp.874–885.
Huang, D.-Y., Nel, A. & Petrulevičius, J.F., 2008. New evolutionary evidence of Grylloblattida from the Middle Jurassic of Inner Mongolia, north-east China (Insecta, Polyneoptera). *Zoological Journal of the Linnean Society*, 152(1), pp.17–24.
Huelsenbeck, J.P. et al., 2001. Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science*, 294(5550), pp.2310 –2314.
Huelsenbeck, J.P., 1998. Systematic bias in phylogenetic analysis: Is the Strepsiptera problem solved? *Systematic Biology*, 47(3), pp.519–537.
Huelsenbeck, J.P., Larget, B. & Swofford, D., 2000. A compound poisson process for relaxing the molecular clock. *Genetics*, 154(4), pp.1879–1892.
Hughes, J. et al., 2006. Dense Taxonomic EST Sampling and Its Applications for Molecular Systematics of the Coleoptera (Beetles). *Molecular Biology and Evolution*, 23(2), pp.268 –278.
Hunt, T. et al., 2007. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science*, 318(5858), pp.1913 –1916.
Hutchinson, G.E. & MacArthur, R.H., 1959. A Theoretical Ecological Model of Size Distributions Among Species of Animals. *The American Naturalist*, 93(869), pp.117–125.
Hwang, W.S. & Weirauch, C., 2012. Evolutionary History of Assassin Bugs (Insecta: Hemiptera: Reduviidae): Insights from Divergence Dating and Ancestral State Reconstruction. *PLoS ONE*, 7(9), p.e45523.
Hynes, H.B.N., 1942. A Study Of The Feeding Of Adult Stoneflies (Plecoptera). *Proceedings of the Royal Entomological Society of London. Series A, General Entomology*, 17(7-9), pp.81–82.

Ikeda, H., Nishikawa, M. & Sota, T., 2012. Loss of flight promotes beetle diversification. *Nature Communications*, 3(648). Available at: http://www.nature.com/ncomms/journal/v3/n1/full/ncomms1659.html [Accessed August 28, 2013].
Ingram, T. & Mahler, D.L., 2013. SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike Information Criterion. *Methods in Ecology and Evolution*, 4(5), pp.416–425.
Inward, D., Beccaloni, G. & Eggleton, P., 2007. Death of an order: a comprehensive molecular phylogenetic study confirms that termites are eusocial cockroaches. *Biology Letters*, 3(3), pp.331 –335.
Irwin, N.R. et al., 2012. Complex patterns of host switching in New World arenaviruses. *Molecular Ecology*, 21(16), pp.4137–4150.
Isaac, N.J.B. et al., 2005. Correlates of Species Richness in Mammals: Body Size, Life History, and Ecology. *The American Naturalist*, 165(5), pp.600–607.
Isaac, N.J.B. et al., 2003. Phylogenetically Nested Comparisons for Testing Correlates of Species Richness: A Simulation Study of Continuous Variables. *Evolution*, 57(1), pp.18–26.
Ishikawa, S.A., Inagaki, Y. & Hashimoto, T., 2012. RY-Coding and Non-Homogeneous Models Can Ameliorate the Maximum-Likelihood Inferences From Nucleotide Sequence Data with Parallel Compositional Heterogeneity. *Evolutionary Bioinformatics Online*, 8, pp.357–371.
Ishiwata, K. et al., 2011. Phylogenetic relationships among insect orders based on three nuclear protein-coding gene sequences. *Molecular Phylogenetics and Evolution*, 58(2), pp.169–180.
Ivanov, V.D. & Menshutkina, T.., 1996. Endemic Caddisflies of Lake Baikal (Trichoptera Apataniidae). *Braueria*, 23, pp.13–28.
Ives, A.R., Midford, P.E. & Garland, T., 2007. Within-Species Variation and Measurement Error in Phylogenetic Comparative Methods. *Systematic Biology*, 56(2), pp.252–270.
Jäch, M.A. & Balke, M., 2003. Key to the adults of Chinease water beetle families. In *Water Beetles of China*. Wien, pp. 21–36.
Janz, N., 2011. Ehrlich and Raven Revisited: Mechanisms Underlying Codiversification of Plants and Enemies. *Annual Review of Ecology, Evolution, and Systematics*, 42(1), pp.71–89.
Jardine, B., 2010. Evolution & Phylogeny of Some Major Aquatic Radiations of Coleoptera. *Aquatic Entomology*.
Jarvis, K.J., Haas, F. & Whiting, M.F., 2005. Phylogeny of earwigs (Insecta: Dermaptera) based on molecular and morphological evidence: reconsidering the classification of Dermaptera. *Systematic Entomology*, 30(3), pp.442–453.
Jeffroy, O. et al., 2006. Phylogenomics: the beginning of incongruence? *Trends in Genetics*, 22(4), pp.225–231.
Jenkins, D.G. et al., 2007. Does size matter for dispersal distance? *Global Ecology and Biogeography*, 16(4), pp.415–425.
Jetz, W. et al., 2012. The global diversity of birds in space and time. *Nature*, 491(7424), pp.444–448.
Jin, Y. & Bae, Y., 2005. The wingless stonefly family Scopuridae (Plecoptera) in Korea. *Aquatic Insects*, 27(1), pp.21–34.
Johnson, B.R. et al., 2013. Phylogenomics Resolves Evolutionary Relationships among Ants, Bees, and Wasps. *Current Biology*, 23(20), pp.2058–2062.

Johnson, E.W. et al., 1994. Non-marine arthropod traces from the subaerial Ordovician Borrowdale Volcanic Group, English Lake District. *Geological Magazine*, 131(03), pp.395–406.
Johnson, K.P., Yoshizawa, K. & Smith, V.S., 2004. Multiple origins of parasitism in lice. *Proceedings of the Royal Society B: Biological Sciences*, 271(1550), pp.1771–1776.
Jones, K.E. et al., 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, 90(9), pp.2648–2648.
Joyce, P. & Marjoram, P., 2008. Approximately Sufficient Statistics and Bayesian Computation. *Statistical Applications in Genetics and Molecular Biology*, 7(1). Available at: http://www.degruyter.com/view/j/sagmb.2008.7.1/sagmb.2008.7.1.1389/sagmb.2008.7.1.1389.xml [Accessed October 28, 2014].
Joy, J.B. & Crespi, B.J., 2012. Island phytophagy: explaining the remarkable diversity of plant-feeding insects. *Proceedings of the Royal Society B: Biological Sciences*, 279(1741), pp.3250–3255.
Jung, S. & Lee, S., 2012. Correlated evolution and Bayesian divergence time estimates of the Cimicoidea (Heteroptera: Cimicomorpha) reveal the evolutionary history. *Systematic Entomology*, 37(1), pp.22–31.
Kaila, L., Mutanen, M. & Nyman, T., 2011. Phylogeny of the mega-diverse Gelechioidea (Lepidoptera): Adaptations and determinants of success. *Molecular Phylogenetics and Evolution*, 61(3), pp.801–809.
Kaplan, I. & Denno, R.F., 2007. Interspecific interactions in phytophagous insects revisited: a quantitative assessment of competition theory. *Ecology Letters*, 10(10), pp.977–994.
Karpa, A., 2001. Revision of the Chloropidae of the collection of B.A. Gimmerthal and a check list of Latvian Chloropidae (Diptera). *Latvijas entomologs*, 38, pp.44–49.
Kathirithamby, J., 2009. Host-Parasitoid Associations in Strepsiptera. *Annual Review of Entomology*, 54(1), pp.227–249.
Katoh, K. et al., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), pp.3059 –3066.
Katzourakis, A. et al., 2001. Macroevolution of hoverflies (Diptera: Syrphidae): the effect of using higher-level taxa in studies of biodiversity, and correlates of species richness. *Journal of Evolutionary Biology*, 14(2), pp.219–227.
Kawahara, A.Y. & Breinholt, J.W., 2014. Phylogenomics provides strong evidence for relationships of butterflies and moths. *Proceedings of the Royal Society B: Biological Sciences*, 281(1788), p.20140970.
Kelley, S.T. & Farrell, B.D., 1998. Is Specialization a Dead End? The Phylogeny of Host Use in Dendroctonus Bark Beetles (Scolytidae). *Evolution*, 52(6), pp.1731–1743.
Kenrick, P. et al., 2012. A timeline for terrestrialization: consequences for the carbon cycle in the Palaeozoic. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1588), pp.519–536.
Kirejtshuk, A.G., 2009. A new genus and species of Sphaeriusidae (Coleoptera, Myxophaga) from Lower Cretaceous Burmese amber. *Denisia*, 26(86), pp.99–102.
Kirejtshuk, A.G. & Azar, D., 2008. New taxa of beetles (Insecta, Coleoptera) from Lebanese amber with evolutionary and systematic comments. *Alavesia*, 2, pp.15–46.
Kjer, K.M., 2004. Aligned 18S and insect phylogeny. *Systematic Biology*, 53(3), pp.506–514.
Kjer, K.M. et al., 2006. A Molecular Phylogeny of Hexapoda. *Arthropod Systematics and Phylogeny*, 64(1), pp.35–44.

Kjer, K.M., Gillespie, J.J. & Ober, K.A., 2007. Opinions on Multiple Sequence Alignment, and an Empirical Comparison of Repeatability and Accuracy between POY and Structural Alignment. *Systematic Biology*, 56(1), pp.133 –146.
Klass, K.-D. et al., 2002. Mantophasmatodea: A New Insect Order with Extant Members in the Afrotropics. *Science*, 296(5572), pp.1456–1459.
Klingenberg, C.P. & Spence, J., 1997. On the role of body size for life-history evolution. *Ecological Entomology*, 22(1), pp.55–68.
Kocarek, P., John, V. & Hulva, P., 2013. When the Body Hides the Ancestry: Phylogeny of Morphologically Modified Epizoic Earwigs Based on Molecular Evidence. *PLoS ONE*, 8(6), p.e66900.
Koeth, M. et al., 2012. The thoracic skeleto-muscular system of Mengenilla (Strepsiptera: Mengenillidae) and its phylogenetic implications. *Arthropod Structure & Development*, 41(4), pp.323–335.
Kômoto, N. et al., 2011. Exploring the molecular phylogeny of phasmids with whole mitochondrial genome sequences. *Molecular Phylogenetics and Evolution*, 58(1), pp.43–52.
Kômoto, N. et al., 2012. Novel gene rearrangements in the mitochondrial genome of a webspinner, Aposthonia japonica (Insecta: Embioptera). *Genome*, 55(3), pp.222–233.
Komsta, L. & Novomestky, F., 2012. moments: Moments, cumulants, skewness, kurtosis and related tests. R package version 0.13. Available at: http://CRAN.R-project.org/package=moments.
Kozłowski, J. & Gawelczyk, A.T., 2002. Why are species' body size distributions usually skewed to the right? *Functional Ecology*, 16(4), pp.419–432.
Kristensen, N.P., 1999. *Lepidoptera, Moths and Butterflies, Volume 1: Evolution Systematics and Biogeography*, Walter de Gruyter.
Kristensen, N.P., Scoble, M. & Karsholt, O., 2007. Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfly diversity. *Zootaxa*, 1668, pp.699–747.
Krzemiñski, W. & Krzemiñski, E., 2003. Triassic Diptera: descriptions, revisions and phylogenetic relations. *Acta zoologica cracoviensia,*, 46, pp.153–184.
Ksepka, D.T. et al., 2011. Synthesizing and databasing fossil calibrations: divergence dating and beyond. *Biology Letters*, 7(6), pp.801–803.
Kühl, G. & Rust, J., 2009. Devonohexapodus bocksbergensis is a synonym of Wingertshellicus backesi (Euarthropoda) – no evidence for marine hexapods living in the Devonian Hunsrück Sea. *Organisms Diversity & Evolution*, 9(3), pp.215–231.
Kukalova-Peck, J., 1987. New Carboniferous Diplura, Monura and Thysanura, the hexapods gorund plan and the role of thoracic side lobes in the origin of wings (insecta). *Canadian Journal of Zoology*, 65, pp.2327–2345.
Kukalová-Peck, J. & Beutel, R.G., 2012. Is the Carboniferous Adiphlebia lacoana really the "oldest beetle"? Critical reassessment and description of a new Permian beetle family. *European Journal of Entomology*, 109, pp.633–645.
Kundrata, R., Bocakova, M. & Bocak, L., 2014. The comprehensive phylogeny of the superfamily Elateroidea (Coleoptera: Elateriformia). *Molecular Phylogenetics and Evolution*, 76, pp.162–171.
Labandeira, C.C., 2013. Deep-time patterns of tissue consumption by terrestrial arthropod herbivores. *Naturwissenschaften*, 100(4), pp.355–364.
Labandeira, C.C., 1998. Early History of Arthropod and Vascular Plant Associations. *Annual Review of Earth and Planetary Sciences*, 26(1), pp.329–377.

Labandeira, C.C., 2011. Evidence for an Earliest Late Carboniferous Divergence Time and the Early Larval Ecology and Diversification of Major Holometabola Lineages. *Entomologica Americana*, 117(1 & 2), pp.9–21.
Labandeira, C.C. et al., 1994. Ninety-seven million years of angiosperm-insect association: Paleobiological insights into the meaning of coevolution. *Proceedings of the National Academy of Sciences*, 91, pp.12278–12282.
Labandeira, C.C., 2005. The fossil record of insect extinction: new approaches and future directions. *American Entomologist*, 51(1), pp.14–29.
Labandeira, C.C., 2006. The Four Phases of Plant-Arthropod Associations in Deep Time. *Geologica Acta*, 4, pp.409–438.
Labandeira, C.C. & Eble, G.J., 2000. The Fossil Record of Insect Diversity and Disparity. In J. Anderson et al., eds. *Gondwana Alive: Biodiversity and the Evolving Biosphere*. Witwatersrand University Press. Available at: http://www.santafe.edu/media/workingpapers/00-08-044.pdf.
Labandeira, C.C. & Sepkoski Jr, J.J., 1993. Insect diversity in the fossil record. *Science*, 261(5119), pp.310–315.
Lafferty, K.D. & Kuris, A.M., 2002. Trophic strategies, animal diversity and body size. *Trends in Ecology & Evolution*, 17(11), pp.507–513.
Lambkin, C.L. et al., 2013. The phylogenetic relationships among infraorders and superfamilies of Diptera based on morphological evidence. *Systematic Entomology*, 38(1), pp.164–179.
Larkin, M.A. et al., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), pp.2947 –2948.
Lawrence, J.F. et al., 2011. Phylogeny of the Coleoptera Based on Morphological Characters of Adults and Larvae. *Annales Zoologici*, 61(1), pp.1–217.
Leavitt, J.R. et al., 2013. Searching for the optimal data partitioning strategy in mitochondrial phylogenomics: A phylogeny of Acridoidea (Insecta: Orthoptera: Caelifera) as a case study. *Molecular Phylogenetics and Evolution*, 67(2), pp.494–508.
Lee, C., Grasso, C. & Sharlow, M.F., 2002. Multiple sequence alignment using partial order graphs. *Bioinformatics*, 18(3), pp.452 –464.
Lee, M., 2001. Unalignable sequences and molecular evolution. *Trends in Ecology & Evolution*, 16(12), pp.681–685.
Legendre, F. et al., 2010. One hundred years of instability in ensiferan relationships. *Systematic Entomology*, 35(3), pp.475–488.
Lehane, M.J., 1991. The evolution of the blood-sucking habit. In *Biology of Blood-Sucking Insects*. Springer Netherlands, pp. 6–13. Available at: http://link.springer.com/chapter/10.1007/978-94-011-7953-9_2 [Accessed March 2, 2015].
Leigh, J.W. et al., 2011. Evaluating Phylogenetic Congruence in the Post-Genomic Era. *Genome Biology and Evolution*, 3, pp.571–587.
Lemmon, A.R. et al., 2009. The Effect of Ambiguous Data on Phylogenetic Estimates Obtained by Maximum Likelihood and Bayesian Inference. *Systematic Biology*, 58(1), pp.130 –145.
Lemmon, A.R. & Moriarty, E.C., 2004. The Importance of Proper Model Assumption in Bayesian Phylogenetics. *Systematic Biology*, 53(2), pp.265–277.
Lenhart, P., Dash, S.T. & Mackay, W.P., 2013. A revision of the giant Amazonian ants of the genus Dinoponera (Hymenoptera, Formicidae). *Journal of Hymenoptera Research*, 31(0), pp.119–164.
Lepage, T. et al., 2007. A General Comparison of Relaxed Molecular Clock Models. *Molecular Biology and Evolution*, 24(12), pp.2669–2680.

Leschen, R.A.B., Beutel, R.G. & Lawrence, J.F., 2010. *Coleoptera. Beetles. Volume 2: Morphology and Systematics (Elateroidea, Bostrichiformia, Cucujiformia partim)*, Walter de Gruyter.
Leschen, R.A.B. & Buckley, T.R., 2007. Multistate Characters and Diet Shifts: Evolution of Erotylidae (Coleoptera). *Systematic Biology*, 56(1), pp.97–112.
Letsch, H.O. et al., 2012. Insect phylogenomics: results, problems and the impact of matrix composition. *Proceedings of the Royal Society B: Biological Sciences*, 279(1741), pp.3282–3290.
Letsch, H.O. et al., 2010. The impact of rRNA secondary structure consideration in alignment and tree reconstruction: simulated data and a case study on the phylogeny of hexapods. *Molecular Biology and Evolution*, 27(11), pp.2507–2521.
Lewis, T., 1973. *Thrips: Their Biology, Ecology and Economic Importance*, Academic Press.
Li, D., Qin, J.-C. & Zhou, C.-F., 2014. The phylogeny of Ephemeroptera in Pterygota revealed by the mitochondrial genome of Siphluriscus chinensis (Hexapoda: Insecta). *Gene*, 545(1), pp.132–140.
Lienhard, C. & Smithers, C.N., 2002. *Psocoptera (Insecta): World Catalogue and Bibliography*, Muséum d'histoire naturelle.
Li, H.-M. et al., 2005. A preliminary phylogeny of the Pentatomomorpha (Hemiptera: Heteroptera) based on nuclear 18S rDNA and mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 37(2), pp.313–326.
Li, M. et al., 2012. Higher Level Phylogeny and the First Divergence Time Estimation of Heteroptera (Insecta: Hemiptera) Based on Multiple Genes. *PLoS ONE*, 7(2), p.e32152.
Lin, Q., 1992. Late Triassic insect fauna from Toksun Xinjiang. *Acta Palaeont Sinica*, 31, pp.313–335.
Liu, Y. et al., 2006. U-Pb zircon age for the Daohugou Biota at Ningcheng of Inner Mongolia and comments on related issues. *Chinese Science Bulletin*, 51(21), pp.2634–2644.
Liu, Y., Sinitshenkova, N.D. & Ren, D., 2009. A revision of the Jurassic stonefly Genera Dobbertiniopteryx Ansorge and Karanemoura Sinitshenkova (Insecta: Plecoptera), with the description of new species from the Daohugou locality, China. *Paleontological Journal*, 43(2), pp.183–190.
Liu, Z., Engel, M.S. & Grimaldi, D.A., 2007. Phylogeny and geological history of the cynipoid wasps (Hymenoptera: Cynipoidea). *American Museum Novitates*, 3583, pp.1–48.
Li, W.-H. & Tanimura, M., 1987. The molecular clock runs more slowly in man than in apes and monkeys. *Nature*, 326(6108), pp.93–96.
Lo, N. et al., 2000. Evidence from multiple gene sequences indicates that termites evolved from wood-feeding cockroaches. *Current Biology*, 10(13), pp.801–804.
Longhorn, S.J., Pohl, H.W. & Vogler, A.P., 2010. Ribosomal protein genes of holometabolan insects reject the Halteria, instead revealing a close affinity of Strepsiptera with Coleoptera. *Molecular Phylogenetics and Evolution*, 55(3), pp.846–859.
Losos, J.B., 2010. Adaptive Radiation, Ecological Opportunity, and Evolutionary Determinism. *The American Naturalist*, 175(6), pp.623–639.
Losos, J.B., 2011. Seeing the Forest for the Trees: The Limitations of Phylogenies in Comparative Biology. *The American Naturalist*, 177(6), pp.709–727.
Luan, Y. et al., 2005. The Phylogenetic Positions of Three Basal-Hexapod Groups (Protura, Diplura, and Collembola) Based on Ribosomal RNA Gene Sequences. *Molecular Biology and Evolution*, 22(7), pp.1579 –1592.

Lukashevich, E.D. et al., 2010. The oldest occurrence of immature Diptera (Insecta), Middle Triassic, France. *Annales de la Société Entomologique de France*, 46(1-2), pp.4–22.
Lukashevich, E.D., Huang, D.-Y. & Lin, Q.-B., 2006. Rare families of lower Diptera (Hennigmatidae, Blephariceridae, Perissommatidae) from the Jurassic of China. *Studia dipterologica*, 13, pp.127–143.
Maddison, D.R. et al., 2009. Monophyly of terrestrial adephagan beetles as indicated by three nuclear genes (Coleoptera: Carabidae and Trachypachidae). *Zoologica scripta*, 38(1), pp.43–62.
Maddison, W.P. & Maddison, D.R., 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75. Available at: http://mesquiteproject.org.
Maddison, W.P., Midford, P.E. & Otto, S.P., 2007. Estimating a Binary Character's Effect on Speciation and Extinction. *Systematic Biology*, 56(5), pp.701–710.
Maehr, M.D. & Eades, D., 2012. Embioptera Species File Online. Version 1.0/4.1. [October 2012]. Available at: http://Embioptera.SpeciesFile.org.
Magallon, S. & Sanderson, M.J., 2001. Absolute Diversification Rates in Angiosperm Clades. *Evolution*, 55(9), pp.1762–1780.
Magnuson-Ford, K. & Otto, S.P., 2012. Linking the Investigations of Character Evolution and Species Diversification. *The American Naturalist*, 180(2), pp.225–245.
Magro, A. et al., 2010. Phylogeny of ladybirds (Coleoptera: Coccinellidae): Are the subfamilies monophyletic? *Molecular Phylogenetics and Evolution*, 54(3), pp.833–848.
Makarkin, V.N. et al., 2012. A comparative overview of the neuropteran assemblage of the Lower Cretaceous Yixian Formation (China), with description of a new genus of Psychopsidae (Insecta: Neuroptera). *Cretaceous Research*, 35, pp.57–68.
Mallatt, J. & Giribet, G., 2006. Further use of nearly complete 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. *Molecular Phylogenetics and Evolution*, 40(3), pp.772–794.
Malm, T., Johanson, K.A. & Wahlberg, N., 2013. The evolutionary history of Trichoptera (Insecta): A case of successful adaptation to life in freshwater. *Systematic Entomology*, 38(3), pp.459–473.
Mardis, E.R., 2013. Next-Generation Sequencing Platforms. *Annual Review of Analytical Chemistry*, 6(1), pp.287–303.
Marinho, M.A.T. et al., 2012. Molecular phylogenetics of Oestroidea (Diptera: Calyptratae) with emphasis on Calliphoridae: Insights into the inter-familial relationships and additional evidence for paraphyly among blowflies. *Molecular Phylogenetics and Evolution*, 65(3), pp.840–854.
Marjoram, P. & Tavaré, S., 2006. Modern computational approaches for analysing molecular genetic variation data. *Nature Reviews Genetics*, 7(10), pp.759–770.
Marshall, C.R., 2008. A Simple Method for Bracketing Absolute Divergence Times on Molecular Phylogenies Using Multiple Fossil Calibration Points. *The American Naturalist*, 171(6), pp.726–742.
Marshall, S., 2012. *Flies: The Natural History and Diversity of Diptera*, Firefly books.
Marsh, N., 2004. *Trout Stream Insects of New Zealand*, Stackpole Books.
Martill, D.M., Bechly, G. & Loveridge, R.F., 2007. *The Crato fossil beds of Brazil: Window into an ancient world*, Cambridge University Press.
Martins, E.P. & Hansen, T.F., 1997. Phylogenies and the Comparative Method: A General Approach to Incorporating Phylogenetic Information into the Analysis of Interspecific Data. *The American Naturalist*, 149(4), pp.646–667.

Martins, E.P. & Hansen, T.F., 1996. The statistical analysis of interspecific data: a review and evaluation of phylogenetic comparative methods. In E. P. Martins, ed. *Phylogenies and Comparative Method in Animal Behavior*. Oxford University Press.
Martins-Neto, R.G., Gallego, O. & Mancuso, A., 2006. The Triassic insect fauna from Argentina. Coleoptera from the Los Rastros Formation (Bermejo Basin), La Rioja Province. *Ameghiniana*, 43(3), pp.591–609.
Marx, F.G. & Uhen, M.D., 2010. Climate, Critters, and Cetaceans: Cenozoic Drivers of the Evolution of Modern Whales. *Science*, 327(5968), pp.993–996.
Mathis, W.N., 2011a. World Catalog and Conspectus on the Family Helcomyzidae (Diptera: Schizophora). *Myia*, 12, pp.267–280.
Mathis, W.N., 2011b. World Catalog and Conspectus on the Family Heterocheilidae (Diptera: Schizophora). *Myia*, 12, pp.281–289.
Mathis, W.N. & Sueyoshi, M., 2011. World Catalog and Conspectus on the Family Dryomyzidae. *Myia*, 12, pp.207–233.
Mattson, W.J., 1980. Herbivory in Relation to Plant Nitrogen Content. *Annual Review of Ecology and Systematics*, 11(1), pp.119–161.
Maurer, B. a, 1998. The evolution of body size in birds. I. Evidence for non-random diversification. *Evolutionary Ecology*, 12(8), pp.925–934.
Mayhew, P.J., 2003. A tale of two analyses: estimating the consequences of shifts in hexapod diversification. *Biological Journal of the Linnean Society*, 80(1), pp.23–36.
Mayhew, P.J., 2006. *Discovering Evolutionary Ecology: Bringing Together Ecology and Evolution*, Oxford: Oxford University Press.
Mayhew, P.J., 2002. Shifts in Hexapod Diversification and what Haldane could have said. *Proceedings of the Royal Society B: Biological Sciences*, 269(1494), pp.969–974.
Mayhew, P.J., 2007. Why are there so many insect species? Perspectives from fossils and phylogenies. *Biological Reviews*, 82(3), pp.425–454.
May, R.M., 1988. How Many Species Are There on Earth? *Science*, 241(4872), pp.1441–1449.
May, R.M., 2000. The dimensions of life on Earth. In P. H. Raven & T. Williams, eds. *Nature and human society: the quest for a sustainable world*. Washington DC: National Academy Press.
McAlpine, D.K., 1991. Marginidae a new afrotropical family of diptera schizophora ? opomyzoidea. *Annals Of The Natal Museum*, 32, pp.167–178.
McAlpine, D.K., 1993. Review of the Upside-down Flies (Diptera: Neurochaetidae) of Madagascar and Africa, and Evolution of Neurochaetid Host Plant Associations. *Records of the Australian Museum*, 45, pp.221–239.
McAlpine, J.F. et al. eds., 1981. *Manual of Nearctic Diptera. Volume 1.*, Research Branch, Agriculture Canada.
McAlpine, J.F. et al. eds., 1987. *Manual of Nearctic Diptera. Volume 2.*, Research Branch, Agriculture Canada.
McCafferty, W.P., 1997. Discovery and analysis of the oldest mayflies (Insecta, Ephemeroptera) known from amber. *Bulletin Societe D Histoire Naturelle de Toulouse*, 133, pp.77–82.
McInnes, L., Orme, C.D.L. & Purvis, A., 2011. Detecting shifts in diversity limits from molecular phylogenies: what can we know? *Proceedings of the Royal Society B: Biological Sciences*, 278(1722), pp.3294–3302.
McKenna, D.D. et al., 2009. Temporal lags and overlap in the diversification of weevils and flowering plants. *Proceedings of the National Academy of Sciences*, 106(17), pp.7083–7088.

McKenna, D.D. & Farrell, B.D., 2010. 9-Genes reinforce the phylogeny of Holometabola and yield alternate views on the phylogenetic placement of Strepsiptera. *PLoS ONE*, 5(7), p.e11887.
McKinney, M.L., 1990. Trends in body size evolution. In K. J. McNamara, ed. *Evolutionary trends*. Tucson: University of Arizona Press, pp. 75–118.
McMahon, D.P., Hayward, A. & Kathirithamby, J., 2011. The First Molecular Phylogeny of Strepsiptera (Insecta) Reveals an Early Burst of Molecular Evolution Correlated with the Transition to Endoparasitism. *PLoS ONE*, 6(6), p.e21206.
McPeek, M.A. & Brown, J.M., 2007. Clade Age and Not Diversification Rate Explains Species Richness among Animal Taxa. *The American Naturalist*, 169(4), pp.E97–E106.
Mercado, M. & Elliot, S., 2005. Taxonomic revision of the genus Chiloporter Lestage (Ameletopsidae: Ephemeroptera) with notes on its biology and distribution. *Studies on Neotropical Fauna and Environment*, 40(3), pp.229–236.
Meredith, R.W. et al., 2011. Impacts of the Cretaceous Terrestrial revolution and KPg extinction on mammal diversification. *Science*, 334(6055), pp.521–524.
Messer, A.C., 1984. Chalicodoma pluto: The World's Largest Bee Rediscovered Living Communally in Termite Nests (Hymenoptera: Megachilidae). *Journal of the Kansas Entomological Society*, 57(1), pp.165–168.
Meusemann, K. et al., 2010. A Phylogenomic Approach to Resolve the Arthropod Tree of Life. *Molecular Biology and Evolution*, 27(11), pp.2451–2464.
Michaelis, F.B., Yule, C. & Calder, A., 2011. Family Gripopterygidae. In *Australian Faunal Directory*. Canberra, Australia: CSIRO Entomology. Available at: http://archive.is/Ckwa#selection-249.0-249.22.
Miller, K.B. et al., 2012. The phylogeny and classification of Embioptera (Insecta). *Systematic Entomology*, 37(3), pp.550–570.
Miller, M.A., Pfeiffer, W. & Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop*. New Orleans: GCE, pp. 1–8.
Milne, I. et al., 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, 25(1), pp.126–127.
Minet, J., 1990. Tentative reconstruction of the ditrysian phylogeny (Lepidoptera: Glossata). *Insect Systematics & Evolution*, 22(1), pp.69–95.
Minet, J., Huang, D.-Y. & Nel, A., 2010. Early Mecopterida and the systematic position of the Microptysmatidae (Insecta: Endopterygota). *Annales de la Société Entomologique de France*, 46(1-2), pp.262–270.
Misof, B., 2002. Diversity of Anisoptera (Odonata): Inferring speciation processes from patterns of morphological diversity. *Zoology*, 105(4), pp.355–365.
Misof, B. et al., 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210), pp.763–767.
Misof, B. et al., 2007. Towards an 18S phylogeny of hexapods: Accounting for group-specific character covariance in optimized mixed nucleotide/doublet models. *Zoology*, 110(5), pp.409–429.
Mitterboeck, T.F., 2012. Consequences of Insect Flight Loss for Molecular Evolutionary Rates and Diversification, Msc Thesis, University of Guelph. Available at: https://atrium.lib.uoguelph.ca/xmlui/handle/10214/3679.
Mitter, C., Farrell, B. & Wiegmann, B., 1988. The Phylogenetic Study of Adaptive Zones: Has Phytophagy Promoted Insect Diversification? *The American Naturalist*, 132(1), pp.107–128.

Mitter, C.W. & Farrell, B.D., 1991. Macroevolutionary Aspects of Insect-Plant Relationships. In E. Bernays, ed. *Insect-Plant Interactions*. Boca Raton, Fla.: CRC Press, pp. 35–78.
Monroe, M.J. & Bokma, F., 2013. Mass extinctions do not explain skew in interspecific body size distributions. *Journal of Zoological Systematics and Evolutionary Research*, 51(1), pp.13–18.
Mooers, A.O. & Heard, S.B., 1997. Inferring Evolutionary Process from Phylogenetic Tree Shape. *The Quarterly Review of Biology*, 72(1), pp.31–54.
Moore, B.R., Chan, K.M.A. & Donoghue, M.J., 2004. Detecting diversification rate variation in supertrees. In O. R. P. Bininda-Emonds, ed. *Phylogenetic supertrees, combining information to reveal the tree of life*. Dordrecht, The Neatherlands: Kluwer Academic Publishers, pp. 488–533.
Moore, B.R. & Donoghue, M.J., 2009. A Bayesian approach for evaluating the impact of historical events on rates of diversification. *Proceedings of the National Academy of Sciences*, 106(11), pp.4307–4312.
Mora, C. et al., 2011. How Many Species Are There on Earth and in the Ocean? *PLoS Biol*, 9(8), p.e1001127.
Morlon, H., 2014. Phylogenetic approaches for studying diversification. *Ecology Letters*, 17(4), pp.508–525.
Morlon, H., Parsons, T.L. & Plotkin, J.B., 2011. Reconciling molecular phylogenies with the fossil record. *Proceedings of the National Academy of Sciences*, 108(39), pp.16327–16332.
Morse, D.R. et al., 1985. Fractal dimension of vegetation and the distribution of arthropod body lengths. *Nature*, 314(6013), pp.731–733.
Morse, J.C., 1974. New Caddisflies (Trichoptera) from Southern Africa. *Journal of the Kansas Entomological Society*, 47(3), pp.328–344.
Morton., K.J., 1914. IV. A remarkable new genus and new species of Odonata, of the legion Podagrion, Selys, from North Queensland. *Transactions of the Royal Entomological Society of London*, 62(1), pp.169–172.
Mouillot, D., George-Nascimento, M. & Poulin, R., 2003. How parasites divide resources: a test of the niche apportionment hypothesis. *Journal of Animal Ecology*, 72(5), pp.757–764.
Mound, L.A., 2010. World Thysanoptera. Available at: http://anic.ento.csiro.au/thrips/index.html [Accessed May 14, 2013].
Mound, L.A. & Morris, D., 2007. The insect Order Thysanoptera: Classification versus Systematics. *Zootaxa*, 1668, pp.395–411.
Mulcahy, D.G. et al., 2012. Estimating divergence dates and evaluating dating methods using phylogenomic and mitochondrial data in squamate reptiles. *Molecular Phylogenetics and Evolution*, 65(3), pp.974–991.
Munoz-Quesada, F. & Holzenthal, R.W., 1997. A new species of Xiphoncentron (Antillotrichia) from Costa Rica with semiterrestrial immatrure stages (Trichoptera Xiphoncentronidae). *Proceedings of the 8th Internationa Symposium on Trichoptera*, pp.355–363.
Munro, J.B. et al., 2011. A Molecular Phylogeny of the Chalcidoidea (Hymenoptera). *PLoS ONE*, 6(11), p.e27023.
Murrell, A. & Barker, S., 2005. Multiple origins of parasitism in lice: phylogenetic analysis of SSU rDNA indicates that the Phthiraptera and Psocoptera are not monophyletic. *Parasitology Research*, 97(4), pp.274–280.

Mutanen, M., Wahlberg, N. & Kaila, L., 2010. Comprehensive Gene and Taxon Coverage Elucidates Radiation Patterns in Moths and Butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 277(1695), pp.2839–2848.
Nagatomi, A. et al., 1991. Apsilocephalidae, a new family of orthorrhaphous Brachycera (Insecta, Diptera). *Zoological Science (Tokyo)*, 8(3), pp.579–591.
Nagatomi, A. & Nagatomi, H., 1987. The Genus Austroleptis from South Chile and Patagonia (Diptera, Rhagionidae). *Memoirs of the Kagoshima University Research Center for the South Pacific*, 8(2), pp.139–156.
Nardi, F. et al., 2003. Hexapod Origins: Monophyletic or Paraphyletic? *Science*, 299(5614), pp.1887 –1889.
Naskrecki, P., 2001. Grasshoppers and their realtives. In *Encylopedia of Biodiversity*. pp. 247–264. Available at: http://enviro.doe.gov.my/lib/digital/1385475746-3-s2.0-B0122268652001401-main.pdf.
Near, T.J. et al., 2013. Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. *Proceedings of the National Academy of Sciences*, 110(31), pp.12738–12743.
Near, T.J., Meylan, P.A. & Shaffer, H.B., 2005. Assessing Concordance of Fossil Calibration Points in Molecular Clock Studies: An Example Using Turtles. *The American Naturalist*, 165(2), pp.137–146.
Neboiss, A., 1986. *Atlas of Trichoptera of the SW Pacific-Australian region*, W. Junk.
Needleman, S.B. & Wunsch, C.D., 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48(3), pp.443–453.
Nee, S., 2006. Birth-death models in macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, 37, pp.1–17.
Nee, S., Holmes, E.C., et al., 1994. Extinction Rates can be estimated from Molecular Phylogenies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 344(1307), pp.77–82.
Nee, S., 2000. Inferring speciation rates from phylogenies. *Evolution*, 55(4), pp.661–668.
Nee, S., May, R.M. & Harvey, P.H., 1994. The Reconstructed Evolutionary Process. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 344(1309), pp.305–311.
Nekola, J. et al., 2013. Latitudinal Variation of Body Size in Land Snail Populations and Communities. In F. Smith & S. K. Lyons, eds. *Animal Body Size: Linking Pattern and Process across Space Time and Taxonomic group*. University of Chicago Press.
Nel, A. et al., 2010. An early winged crown group stick insect from the Early Eocene amber of France (Insecta, Phasmatodea). *Systematic Entomology*, 35, pp.340–346.
Nel, A., Aria, C., et al., 2012. Evolution and palaeosynecology of the Mesozoic earwigs (Insecta: Dermaptera). *Cretaceous Research*, 33(1), pp.189–195.
Nel, A. et al., 2007. The earliest holometabolous insect from the Carboniferous: a “crucial” innovation with delayed success (Insecta Protomeropina Protomeropidae). *Annales de la Société Entomologique de France*, 43(3), pp.349–355.
Nel, A. et al., 2013. The earliest known holometabolous insects. *Nature*, 503(7475), pp.257–261.
Nel, A., Prokop, J., et al., 2012. Traits and evolution of wing venation pattern in paraneopteran insects. *Journal of Morphology*, 273(5), pp.480–506.
Nel, A. & Delfosse, E., 2011. A new Chinese Mesozoic stick insect. *Acta Palaeontologica Polonica*, 56(2), pp.429–432.

Nel, A. & Roy, R., 1996. Revision of fossil “mantid” and “ephemerid” species described by Piton from the Paleocene of Menat (France) Mantodea Chaeteessidae; Ensifera Tettigonioidea. *European Journal of Entomology*, 93, pp.223–234.
Nel, A. & Waller, A., 2007. The first fossil Compsocidae from Cretaceous Burmese amber (Insecta, Psocoptera, Troctomorpha). *Cretaceous Research*, 28(6), pp.1039–1041.
Nel, P., Azar, D., et al., 2012. From Carboniferous to Recent: wing venation enlightens evolution of thysanopteran lineage. *Journal of Systematic Palaeontology*, 10(2), pp.385–399.
Nel, P. et al., 2010. Modern thrips families Thripidae and Phlaeothripidae in Early Cretaceous amber (Insecta: Thysanoptera). *Annales de la Société Entomologique de France*, 46(1-2), pp.154–163.
Nelson, L.A., Scheffer, S.J. & Yeates, D.K., 2011. Two new species of sympatric Fergusonina Malloch flies (Diptera: Fergusoninidae) from bud galls on high-elevation snow gums (Eucalyptus pauciflora Sieb. ex Spreng. complex) in the Australian Alps. *Australian Journal of Entomology*, 50(4), pp.356–364.
New, T. & Lienhard, C., 2007. *The Psocoptera of Tropical South East Asia*, Brill Academic Publishing.
Ng, P.K.L., Corlett, R. & Tan, H.T.W., 2011. *Singapore Biodiversity: An Encyclopedia of the Natural Environment and Sustainable Development*, Editions Didier Millet.
Nicholson, D.B., 2012. Fossil perspectives on the evolution of Insect Diversity. PhD thesis, Department of Biology, University of York, UK. Available at: http://etheses.whiterose.ac.uk/3713/.
Nicholson, D.B., Ross, A.J. & Mayhew, P.J., 2014. Fossil evidence for key innovations in the evolution of insect diversity. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1793), p.20141823.
Niehuis, O. et al., 2012. Genomic and Morphological Evidence Converge to Resolve the Enigma of Strepsiptera. *Current Biology*, 22(14), pp.1309–1313.
Nikolajev, G.V. et al., 2011. Stag Beetles from the Mesozoic of Inner Mongolia, China (Scrabaeoidea: Lucanidae). *Acta Palaeont Sinica*, 50, pp.41–47.
Nikolajev, G.V. & Ren, D., 2010. The oldest fossil Ochodaeidae (Coleoptera: Scarabaeoidea) from the Middle Jurassic of China. *Zootaxa*, 2553, pp.65–68.
Nosenko, T. et al., 2013. Deep metazoan phylogeny: When different genes tell different stories. *Molecular Phylogenetics and Evolution*, 67(1), pp.223–233.
Nosil, P., 2002. Transition Rates Between Specialization and Generalization in Phytophagous Insects. *Evolution*, 56(8), pp.1701–1706.
Nosil, P. & Mooers, A.Ø., 2005. Testing Hypotheses About Ecological Specialization Using Phylogenetic Trees. *Evolution*, 59(10), pp.2256–2263.
Novotny, V. et al., 2010. Guild-specific patterns of species richness and host specialization in plant–herbivore food webs from a tropical forest. *Journal of Animal Ecology*, 79(6), pp.1193–1203.
Novotny, V. et al., 2012. Insects on Plants: Explaining the Paradox of Low Diversity within Specialist Herbivore Guilds. *The American Naturalist*, 179(3), pp.351–362.
Novotny, V. & Kindlmann, P., 1996. Distribution of Body Sizes in Arthropod Taxa and Communities. *Oikos*, 75(1), pp.75–82.
Nowak, M.D. et al., 2013. A Simple Method for Estimating Informative Node Age Priors for the Fossil Calibration of Molecular Divergence Time Analyses. *PLoS ONE*, 8(6), p.e66245.
Noyes, J.S., 2003. Universal Chalcidoidea Database. Online electronic publication. Available at: http://www.nhm.ac.uk/chalcidoids.

Nyman, T. et al., 2010. How common is ecological speciation in plant-feeding insects? A "Higher" Nematinae perspective. *BMC Evolutionary Biology*, 10(1), p.266.
Nyman, T., 2010. To speciate, or not to speciate? Resource heterogeneity, the subjectivity of similarity, and the macroevolutionary consequences of niche-width shifts in plant-feeding insects. *Biological Reviews*, 85(2), pp.393–411.
Ogden, T. h et al., 2009. Towards a new paradigm in mayfly phylogeny (Ephemeroptera): combined analysis of morphological and molecular data. *Systematic Entomology*, 34(4), pp.616–634.
Ogden, T.H. & Rosenberg, M.S., 2007. Alignment and Topological Accuracy of the Direct Optimization Approach Via POY and Traditional Phylogenetics Via ClustalW + PAUP*. *Systematic Biology*, 56(2), pp.182–193.
Ogden, T.H. & Whiting, M.F., 2003. The problem with "the Paleoptera Problem:" sense and sensitivity. *Cladistics*, 19(5), pp.432–442.
Ohl, M. & Thiele, K., 2007. Estimating body size in apoid wasps: the significance of linear variables in a morphologically diverse taxon (Hymenoptera, Apoidea). *Zoosystematics and Evolution*, 83(2), pp.110–124.
Olah, J. & Johanson, K.A., 2010. Contributions to the systematics of the genera Dipseudopsis , Hyalopsyche and Pseudoneureclipsis (Trichoptera: Dipseudopsidae), with descriptions of 19 new species from the Oriental Region. *Zootaxa*, 2658, pp.1–37.
O'Meara, B.C., 2012. Evolutionary Inferences from Phylogenies: A Review of Methods. *Annual Review of Ecology, Evolution, and Systematics*, 43(1), pp.267–285.
Oosterbroek, P., 1998. *The Families of Diptera of the Malay Archipelago*, Brill Academic Publishing.
Orme, C.D.L., Quicke, D.L.J., et al., 2002. Body size does not predict species richness among the metazoan phyla. *Journal of Evolutionary Biology*, 15(2), pp.235–247.
Orme, C.D.L. et al., 2012. caper: Comparative Analyses of Phylogenetics and Evolution in R. Available at: http://CRAN.R-project.org/package=caper.
Orme, C.D.L., Isaac, N.J.B. & Purvis, A., 2002. Are most species small? Not within species–level phylogenies. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1497), pp.1279–1287.
Otte, D., 2007. *Australian Crickets (Orthoptera: Gryllidae)*, Academy of Natural Sciences.
Otte, D., Spearman, L. & Stiewe, M., 2012. Mantodea Species File Online. Version 1.0/4.1. [October 2012]. Available at: http://Mantodea.SpeciesFile.org.
Pagel, M., 1994. Detecting Correlated Evolution on Phylogenies: A General Method for the Comparative Analysis of Discrete Characters. *Proceedings of the Royal Society B: Biological Sciences*, 255(1342), pp.37–45.
Pagel, M., 1997. Inferring evolutionary processes from phylogenies. *Zoologica Scripta*, 26(4), pp.331–348.
Pagel, M., 1999. Inferring the historical patterns of biological evolution. *Nature*, 401(6756), pp.877–884.
Palaczyk, A., Klasa, A. & Slowinska-Krysiak, I., 2013. The family Strongylophthalmyiidae of Poland with catalog of European species (Insecta: Diptera). *Genus*, 24(3-4), pp.425–438.
Palmer, C.M. & Siebke, K., 2008. Cold hardiness of Apteropanorpa tasmanica Carpenter (Mecoptera: Apteropanorpidae). *Journal of Insect Physiology*, 54(7), pp.1148–1156.
Pape, T. & Evenhius, N.L. eds., 2013. Systema Dipterorum, Version 1.5. Available at: http://www.diptera.org/ [Accessed November 4, 2014].
Papp, L., 2010. A study on Hesperinus (Walker) with description of a new species (Diptera: Hesperinidae). *Acta Zoologica Academiae Scientiarum Hungaricae*, 56(4), pp.347–370.

Paradis, E., 2012. Shift in Diversification in Sister-Clade Comparisons: A More Powerful Test. *Evolution*, 66(1), pp.288–295.
Paradis, E., Claude, J. & Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, 20(2), pp.289–290.
Parham, J.F. et al., 2012. Best practices for justifying fossil calibrations. *Systematic Biology*, 61(2), pp.346–359.
Parker, C., 1998. A review of Goerita (Trichoptera: Goeridae), with description of a new species. *Insecta Mundi*, 377.
Parker, S.P., 1982. *Synopsis and Classification of Living Organisms*, Mc-Graw Hill Book Company.
Parry, S.F. et al., 2011. A high-precision U–Pb age constraint on the Rhynie Chert Konservat-Lagerstatte: time scale and other implications. *Journal of Geological Society London*, 168, pp.863–872.
Pass, G. & Szucsich, N.U., 2011. 100 years of research on the Protura: many secrets still retained. *Soil Organisms*, 83(3), pp.309–334.
Pearson, R.G. & Penridge, L.K., 1979. First records of Prosopistoma sedlaceki in Australia (Ephemeroptera Prosopistomatidae). *Journal of the Australian Entomological Society*, 18, p.362.
Pennell, M.W. et al., 2014. geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, 30(15), pp.2216–2218.
Penniket, J.G., 1962. Notes on New Zealand Ephemeroptera III. A new family genus and species. *Records of the Canterbury Museum*, 7(5), pp.389–398.
Penniket, J.G., 1966. Notes on New Zealand Ephemeroptera IV. A new Siphlonurid subfamily; Rallidentinae. *Records of the Canterbury Museum*, 8(2), pp.163–175.
Pennington, R.T., Richardson, J.E. & Lavin, M., 2006. Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. *New Phytologist*, 172(4), pp.605–616.
Penny, N., 1997. World Checklist of Extant Mecoptera Species. Available at: http://research.calacademy.org/redirect?url=http://researcharchive.calacademy.org/research/entomology/Entomology_Resources/mecoptera/index.htm.
Perrichot, V., Nel, A. & Neraudeau, D., 2005. Gerromorphan bugs in Early Cretaceous French amber (Insecta: Heteroptera): first representatives of Gerridae and their phylogenetic and palaeoecological implications. *Cretaceous Research*, 26(5), pp.793–800.
Peters, R.S. et al., 2014. The evolutionary history of holometabolous insects inferred from transcriptome-based phylogeny and comprehensive morphological data. *BMC Evolutionary Biology*, 14(1), p.52.
Peters, R.S. et al., 2011. The taming of an impossible child: a standardized all-in approach to the phylogeny of Hymenoptera using public database sequences. *BMC Biology*, 9(55).
Philippe, H. et al., 2004. Phylogenomics of Eukaryotes: Impact of Missing Data on Large Alignments. *Molecular Biology and Evolution*, 21(9), pp.1740–1752.
Philippe, H. et al., 2011. Resolving Difficult Phylogenetic Questions: Why More Sequences Are Not Enough. *PLoS Biol*, 9(3), p.e1000602.
Philippe, H. & Forterre, P., 1999. The Rooting of the Universal Tree of Life Is Not Reliable. *Journal of Molecular Evolution*, 49(4), pp.509–523.
Phillimore, A.B. et al., 2006. Ecology Predicts Large-Scale Patterns of Phylogenetic Diversification in Birds. *The American Naturalist*, 168(2), pp.220–229.
Phillips, A., Janies, D. & Wheeler, W., 2000. Multiple Sequence Alignment in Phylogenetic Analysis. *Molecular Phylogenetics and Evolution*, 16(3), pp.317–330.
Phillips, J.S., 1930. A revision of New Zealand Ephemeroptera Part 1. *Transactions and Proceedings of the Royal Society of New Zealand*, 61, pp.271–334.

Phillips, M.J., Delsuc, F. & Penny, D., 2004. Genome-Scale Phylogeny and the Detection of Systematic Biases. *Molecular Biology and Evolution*, 21(7), pp.1455 –1458.
Picker, M., Griffiths, C. & Weaving, A., 2004. *Field Guide to Insects of South Africa*, Struik Nature.
Pickett, K.M. & Randle, C.P., 2005. Strange bayes indeed: uniform topological priors imply non-uniform clade priors. *Molecular Phylogenetics and Evolution*, 34(1), pp.203–211.
Pilgrim, E.M., Von Dohlen, C.D. & Pitts, J.P., 2008. Molecular phylogenetics of Vespoidea indicate paraphyly of the superfamily and novel relationships of its component families and subfamilies. *Zoologica Scripta*, 37(5), pp.539–560.
Pinheiro, J. et al., 2014. *nlme: Linear and Nonlinear Mixed Effects Models*, Available at: http://cran.r-project.org/web/packages/nlme/index.html [Accessed July 2, 2014].
Plazzi, F., Ricci, A. & Passamonti, M., 2011. The mitochondrial genome of Bacillus stick insects (Phasmatodea) and the phylogeny of orthopteroid insects. *Molecular Phylogenetics and Evolution*, 58(2), pp.304–316.
Pohl, G. et al., 2010. *Annotated List of the Lepidoptera of Alberta, Canada*, PenSoft Publishers LTD.
Pohl, H. & Beutel, R.G., 2005. The phylogeny of Strepsiptera (Hexapoda). *Cladistics*, 21(4), pp.328–374.
Poinar Jr, G., 2009. Mellittosphex (Hymenoptera:Melittosphecidae), a Primitive Bee not a a Wasp. *Palaeontology*, 52(2), pp.483–484.
Poinar Jr, G., Kirejtshuk, A.G. & Buckley, R., 2008. Pleuroceratos burmiticus, n. gen., n. sp.(Coleoptera: Silvanidae) from early Cretaceous Burmese amber. *Proceedings of the Entomological Society of Washington*, 110(1), pp.250–257.
Poisot, T. et al., 2012. A comparative study of ecological specialization estimators. *Methods in Ecology and Evolution*, 3(3), pp.537–544.
Poisot, T. et al., 2011. A conceptual framework for the evolution of ecological specialisation. *Ecology Letters*, 14(9), pp.841–851.
Posada, D., 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, 25(7), pp.1253 –1256.
Poulin, R., 2014. Parasite biodiversity revisited: frontiers and constraints. *International Journal for Parasitology*, 44(9), pp.581–589.
Poulin, R. & Morand, S., 1997. Parasite body size distributions: Interpreting patterns of skewness. *International Journal for Parasitology*, 27(8), pp.959–964.
Powell, J., Mitter, C. & Farrell, B.D., 1998. Evolution of larval food preference in Lepidoptera. In N. P. Kristensen, ed. *Moths and Butterflies. Volume 1. Evolution, Systematics, and Biogeography. Handbook of Zoology. Volume IV. Arthropoda: Insecta*. Berlin: Walter deGruyter.
Pratt, R.C., Morgan-Richards, M. & Trewick, S.A., 2008. Diversification of New Zealand weta (Orthoptera: Ensifera: Anostostomatidae) and their relationships in Australasia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1508), pp.3427–3437.
Prete, F.R., 1999. *The Praying Mantids*, JHU Press.
Price, S.A. et al., 2012. Tempo of trophic evolution and its impact on mammalian diversification. *Proceedings of the National Academy of Sciences*, 109(18), pp.7008–7012.
Prokop, J., Nel, A. & Hoch, I., 2005. Discovery of the oldest known Pterygota in the Lower Carboniferous of the Upper Silesian Basin in the Czech Republic (Insecta: Archaeorthoptera). *Geobios*, 38(3), pp.383–387.

Pruesse, E. et al., 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, 35(21), pp.7188 –7196.
Pulawski, W., 2012. Catalog of Sphecidae sensu lato. Available at: http://research.calacademy.org/ent/catalog_sphecidae.
Purvis, A., 2004. Evolution: How do characters evolve? *Nature*, 432(7014). Available at: http://www.nature.com/nature/journal/v432/n7014/full/nature03092.html [Accessed June 19, 2015].
Pybus, O.G. & Harvey, P.H., 2000. Testing macro–evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1459), pp.2267 –2272.
Pyron, R.A., 2010. A Likelihood Method for Assessing Molecular Divergence Time Estimates and the Placement of Fossil Calibrations. *Systematic Biology*, 59(2), pp.185–194.
Pyron, R.A. & Burbrink, F.T., 2012. Extinction, Ecological Opportunity, and the Origins of Global Snake Diversity. *Evolution*, 66(1), pp.163–178.
Pyron, R.A. & Burbrink, F.T., 2013. Phylogenetic estimates of speciation and extinction rates for testing ecological and evolutionary hypotheses. *Trends in Ecology & Evolution*, 28(12), pp.729–736.
De Queiroz, A., 2002. Contingent predictability in evolution: key traits and diversification. *Systematic biology*, 51(6), pp.917–929.
De Queiroz, A., 1998. Interpreting Sister-Group Tests of Key Innovation Hypotheses. *Systematic Biology*, 47(4), pp.710–718.
Quental, T.B. & Marshall, C.R., 2010. Diversity dynamics: molecular phylogenies need the fossil record. *Trends in Ecology & Evolution*, 25(8), pp.434–441.
Rabosky, D.L., 2014. Automatic Detection of Key Innovations, Rate Shifts, and Diversity-Dependence on Phylogenetic Trees. *PLoS ONE*, 9(2), p.e89543.
Rabosky, D.L., 2009a. Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecology Letters*, 12(8), pp.735–743.
Rabosky, D.L., 2009b. Ecological Limits on Clade Diversification in Higher Taxa. *The American Naturalist*, 173(5), pp.662–674.
Rabosky, D.L. et al., 2007. Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proceedings of the Royal Society B: Biological Sciences*, 274(1628), pp.2915 –2923.
Rabosky, D.L., 2010. Extinction Rates Should Not Be Estimated from Molecular Phylogenies. *Evolution*, 64(6), pp.1816–1824.
Rabosky, D.L., 2007. LASER: A Maximum Likelihood Toolkit for Detecting Temporal Shifts in Diversification Rates From Molecular Phylogenies. *Evolutionary Bioinformatics Online*, 2, pp.247–250.
Rabosky, D.L., Slater, G.J. & Alfaro, M.E., 2012. Clade Age and Species Richness Are Decoupled Across the Eukaryotic Tree of Life. *PLoS Biol*, 10(8), p.e1001381.
Rainford, J.L. et al., 2014. Phylogenetic Distribution of Extant Richness Suggests Metamorphosis Is a Key Innovation Driving Diversification in Insects. *PLoS ONE*, 9(10), p.e109085.
Rambaut, A. & Drummond, A.J., 2009. Tracer [computer program]. Available at: http://tree.bio.ed.ac.uk/software/tracer/.
Randle, C.P., Mort, M.E. & Crawford, D.J., 2005. Bayesian Inference of Phylogenetics Revisited: Developments and Concerns. *Taxon*, 54(1), pp.9–15.

Rasnitsyn, A., 2010. Molecular phylogenetics, morphological cladistics, and fossil record. *Entomological Review*, 90(3), pp.263–298.
Rasnitsyn, A.P., 1988. An outline of evolution of hymenopterous insects (order Vespida). *Oriental Insects*, 22, pp.115–145.
Rasnitsyn, A.P. & Quicke, D.L.J., 2002. *History of the Insects*, AA Dordrecht, the Neatherlands: Springer Netherlands.
Rasnitsyn, A.P. & Zhang, H., 2004. Composition and age of the Daohugou hymenopteran (Insecta, Hymenoptera - Vespida) assemblage from Inner Mongolia, China. *Palaeontology*, 47(6), pp.1507–1517.
Rasnitsyn, A.P. & Zhang, H., 2010. Early Evoltuion of Apocrita (Insecta, Hymenoptera) as indicated by New Findings in the Middle Jurassic of Daohugou, Northeast China. *Acta Zoologica Sinica*, 84(4), pp.834–873.
R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. Available at: http://www.R-project.org.
Regier, J.C. et al., 2013. A Large-Scale, Higher-Level, Molecular Phylogenetic Study of the Insect Order Lepidoptera (Moths and Butterflies). *PLoS ONE*, 8(3), p.e58568.
Regier, J.C. et al., 2010. Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature*, 463(7284), pp.1079–1083.
Regier, J.C. et al., 2008. Resolving Arthropod phylogeny: exploring phylogenetic signal within 41 Kb of protein-coding nuclear gene sequence. *Systematic Biology*, 57(6), pp.920–938.
Regier, J.C. et al., 2009. Toward reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia): an initial molecular study. *BMC Evolutionary Biology*, 9(1), p.280.
Regier, J.C., Shultz, J.W. & Kambic, R.E., 2004. Phylogeny of Basal Hexapod Lineages and Estimates of Divergence Times. *Annals of the Entomological Society of America*, 97, pp.411–419.
Rehm, P. et al., 2011. Dating the arthropod tree based on large-scale transcriptome data. *Molecular Phylogenetics and Evolution*, 61(3), pp.880–887.
Ren, D. & Engel, M.S., 2008. A second antlion from the Mesozoic of northeastern China (Neuroptera: Myrmeleontidae). *Alavasia*, 2, pp.183–186.
Rentz, D., 2010. *A Guide to the Katydids of Australia*, Csiro Publishing.
Resh, V.H. & Cardé, R.T., 2009. *Encyclopedia of Insects, Second Edition*, Elsevier.
Retana-Salazar, A., 2009. Species of Heterothrips Hood (Terebrantia: Heterothripidae) from Central America. *Ceiba*, 50(1), pp.10–17.
Von Reumont, B.M. et al., 2012. Pancrustacean Phylogeny in the Light of New Phylogenomic Data: Support for Remipedia as the Possible Sister Group of Hexapoda. *Molecular Biology and Evolution*, 29(3), pp.1031–1045.
Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3(2), pp.217–223.
Revell, L.J., 2013. Two new graphical methods for mapping trait evolution on phylogenies. *Methods in Ecology and Evolution*, 4(8), pp.754–759.
Revell, L.J. & Reynolds, R.G., 2012. A New Bayesian Method for Fitting Evolutionary Models to Comparative Data with Intraspecific Variation. *Evolution*, 66(9), pp.2697–2707.
Ribera, I., Hogan, J.E. & Vogler, A.P., 2002. Phylogeny of Hydradephagan Water Beetles Inferred from 18S rRNA Sequences. *Molecular Phylogenetics and Evolution*, 23(1), pp.43–62.
Richards, A., 1959. Revision of the Rhaphidophoridae (Orthopera) of New Zealand Part IV: The Raphidophoridae of Thames Gold Mines. *Transactions of the Royal Society of New Zealand*, 87, pp.27–33.

Richards, A., 1968. The Raphidophoridae (Orthopera) or Australia: Part 6 Two New Species from Northern Tasmania. *Pacific Insects*, 10(1), pp.167–176.
Ricklefs, R.E., 2004a. A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, 7(1), pp.1–15.
Ricklefs, R.E., 2004b. Cladogenesis and morphological diversification in passerine birds. *Nature*, 430(6997), pp.338–341.
Ricklefs, R.E., 2007. Estimating diversification rates from phylogenetic information. *Trends in Ecology & Evolution*, 22(11), pp.601–610.
Ricklefs, R.E., 2006. Time, Species, and the Generation of Trait Variance in Clades. *Systematic Biology*, 55(1), pp.151–159.
Ricklefs, R.E. & Renner, S.S., 2000. Evolutionary Flexibility and Flowering Plant Familial Diversity: A Comment on Dodd, Silvertown, and Chase. *Evolution*, 54(3), pp.1061–1065.
Riek, E.F., 1973. A Revision of Australian Scorpion Flies of the Family Choristidae (mecoptera). *Australian Journal of Entomology*, 12(2), pp.103–112.
Riek, E. & Kukalova-Peck, J., 1984. A new interpretation of dragonfly wing venation based upon Early Upper Carboniferous fossils from Argentina (Insecta: Odonatoidea) and basic character states in pterygote wings. *Canadian Journal of Zoology*, 62, pp.1150–1166.
Rivera, A.C., 2014. Behaviour and ecology of Hemiphlebia mirabilis (Odonata: Hemiphlebiidae). Available at: http://natureglenelg.org.au/wp-content/uploads/2014/02/Behaviour-and-ecology-of-Hemiphlebia-mirabilis-by-Adolfo-Cordero-Rivera-2014.pdf.
Robinson, G.E. et al., 2011. Creating a Buzz About Insect Genomes. *Science*, 331(6023), pp.1386–1386.
Robinson, W.H., 2005. *Urban Insects and Arachnids: A Handbook of Urban Entomology*, Cambridge University Press.
Rohlf, F.J., 2001. Comparative methods for the analysis of continuous variables: Geometric interpretations. *Evolution*, 55(11), pp.2143–2160.
Ronquist, F., Klopfstein, S., et al., 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology*, 61(6), pp.973–999.
Ronquist, F., Teslenko, M., et al., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3), pp.539–542.
Ronquist, F., van der Mark, P. & Huelsenbeck, J.P., 2009. Bayesian phylogenetic analysis using MrBayes. In P. Lemey, M. Salemi, & A. M. Vandamme, eds. *The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing*. Cambridge University Press.
Ross, A.J., Jarzembowski, E.A. & Brooks, S.J., 2000. The Cretaceous and Cenozoic record of insects (Hexapoda) with regard to global change. In S. J. Culver & P. Rawson, eds. *Biotic Response to Global Change: The Last 145 Million Years*. Cambridge: Cambridge University Press, pp. 288– 302.
Rota-Stabelli, O., Daley, A.C. & Pisani, D., 2013. Molecular timetrees reveal a Cambrian colonization of land and a new scenario for Ecdysozoan evolution. *Current Biology*, 23(5), pp.392–398.
Roure, B., Baurain, D. & Philippe, H., 2012. Impact of Missing Data on Phylogenies Inferred from Empirical Phylogenomic Data Sets. *Molecular Biology and Evolution*, 30(1), pp.197–214.
Rutschmann, F., 2006. Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. *Diversity and Distributions*, 12(1), pp.35–48.
Sanders, K.L. & Lee, M.S.Y., 2010. Arthropod molecular divergence times and the Cambrian origin of pentastomids. *Systematics and Biodiversity*, 8(1), pp.63–74.

Sanders, K.L. & Lee, M.S.Y., 2007. Evaluating molecular clock calibrations using Bayesian analyses with soft and hard bounds. *Biology Letters*, 3(3), pp.275 –279.
Sanderson, M.J. & Driskell, A.C., 2003. The challenge of constructing large phylogenetic trees. *Trends in Plant Science*, 8(8), pp.374–379.
Sanderson, M.J., McMahon, M.M. & Steel, M., 2010. Phylogenomics with incomplete taxon coverage: the limits to inference. *BMC Evolutionary Biology*, 10(1), p.155.
Sanderson, M.J. & Shaffer, H.B., 2002. Troubleshooting Molecular Phylogenetic Analyses. *Annual Review of Ecology and Systematics*, 33, pp.49–72.
Santos, A.M.C. & Quicke, D.L.J., 2011. Large-scale diversity patterns of parasitoid insects. *Entomological Science*, 14(4), pp.371–382.
Sauquet, H. et al., 2012. Testing the Impact of Calibration on Molecular Divergence Times Using a Fossil-Rich Group: The Case of Nothofagus (Fagales). *Systematic Biology*, 61(2), pp.289–313.
Savard, J. et al., 2006. Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects. *Genome Research*, 16(11), pp.1334 – 1338.
Savolainen, V. et al., 2000. Phylogeny of the Eudicots: A Nearly Complete Familial Analysis Based on rbcL Gene Sequences. *Kew Bulletin*, 55(2), pp.257–309.
Scali, V., MIlani, L. & Passmonti, M., 2012. Revision of the stick insect genus Leptynia : description of new taxa, speciation mechanism and phylogeography. *Contributions to Zoology*, 81(1), pp.25–42.
Scherbakov, D.E., 2009. Review of fossil and extant genera of the cicada family Tettigarctidae. *Russiand Entomological Journal*, 17(4), pp.343–348.
Schluter, D., 2009. Evidence for Ecological Speciation and Its Alternative. *Science*, 323(5915), pp.737–741.
Schneeberg, K., Krause, K. & Beutel, R.G., 2013. The adult head of Axymyia furcata (Insecta: Diptera: Axymyiidae). *Arthropod Systematics & Phylogeny*, 71(2), pp.91–102.
Schuh, R.T. & Slater, J.A., 1995. *True Bugs of the World (Hemiptera:Heteroptera): Classification and Natural History*, Cornell University Press.
Serrano-Meneses, M.A. et al., 2008. Sexual selection, sexual size dimorphism and Rensch's rule in Odonata. *Journal of Evolutionary Biology*, 21(5), pp.1259–1273.
Sharkey, M.J. et al., 2012. Phylogenetic relationships among superfamilies of Hymenoptera. *Cladistics*, 28(1), pp.80–112.
Shi, G. et al., 2012. Age constraint on Burmese amber based on U–Pb dating of zircons. *Cretaceous Research*, 37, pp.155–163.
Shih, C. et al., 2010. Morphology, phylogeny, evolution, and dispersal of pelecinid wasps (Hymenoptera: Pelecinidae) Over 165 million years. *Annals of the Entomological Society of America*, 103(6), pp.875–885.
Shih, C., Feng, H. & Ren, D., 2011. New fossil Heloridae and Mesoserphidae Wasps (Insecta, Hymenoptera, Proctotrupoidea) from the Middle Jurassic of China. *Annals of the Entomological Society of America*, 104(6), pp.1334–1348.
Shmakov, A., 2009. The oldest members of the families Aeolothripidae and Thripidae (Insecta: Thysanoptera) from the Lower Cretaceous of Transbaikalia. *Paleontological Journal*, 43(4), pp.428–432.
Shockley, F.W., 2008. Alexiidae Imhoff 1856. Sphaerosoma Samouelle 1819. Available at: http://tolweb.org/Sphaerosoma/65850/2008.06.2.
Shockley, F.W., Hartley, C.S. & Lord, N., 2011. Latridiidae. Minute brown scavenger beetles. Available at: http://tolweb.org/Latridiidae/9172/2011.03.23.
Silsby, J., 2001. *Dragonflies of the World*, Csiro Publishing.

Silvertown, J. et al., 2000. “Flexibility” as a Trait and Methodological Issues in Species Diversity Variation Among Angiosperm Families. *Evolution*, 54(3), pp.1066–1068.
Silvestro, D. et al., 2015. Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution*, 6(3), pp.340–346.
Silvestro, D., Schnitzler, J. & Zizka, G., 2011. A Bayesian framework to estimate diversification rates and their variation through time and space. *BMC Evolutionary Biology*, 11(1), p.311.
Simberloff, D. & Dayan, T., 1991. The Guild Concept and the Structure of Ecological Communities. *Annual Review of Ecology and Systematics*, 22(1), pp.115–143.
Simmons, M.P., Mueller, K.F. & Norton, A.P., 2010. Alignment of, and phylogenetic inference from, random sequences: The susceptibility of alternative alignment methods to creating artifactual resolution and support. *Molecular Phylogenetics and Evolution*, 57(3), pp.1004–1016.
Simmons, M.P., Pickett, K.M. & Miya, M., 2004. How Meaningful Are Bayesian Support Values? *Molecular Biology and Evolution*, 21(1), pp.188–199.
Simon, S. et al., 2009. A Phylogenomic Approach to Resolve the Basal Pterygote Divergence. *Molecular Biology and Evolution*, 26(12), pp.2719 –2730.
Simon, S. et al., 2012. Insect Phylogenomics: Exploring the Source of Incongruence Using New Transcriptomic Data. *Genome Biology and Evolution*, 4(12), pp.1295–1309.
Simpson, G.G., 1967. *The Major Features of Evolution*, Touchstone.
Singer, M.S. & Stireman, J.O., 2005. The tri-trophic niche concept and adaptive radiation of phytophagous insects. *Ecology Letters*, 8(12), pp.1247–1255.
Slater, G.J. et al., 2012. Fitting Models of Continuous Trait Evolution to Incompletely Sampled Comparative Data Using Approximate Bayesian Computation. *Evolution*, 66(3), pp.752–762.
Slater, G.J., 2013. Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution*, 4(8), pp.734–744.
Slowinski, J.B. & Guyer, C., 1989. Testing the Stochasticity of patterns of organismal diversity: an improved null model. *American Naturalist*, 134(6), pp.907–921.
Smith, A.B.T., Hawks, D.C. & Heraty, J.M., 2006. An Overview of the Classification and Evolution of the Major Scarab Beetle Clades (Coleoptera: Scarabaeoidea) Based on Preliminary Molecular Analyses. *The Coleopterists Bulletin*, 60(sp5), pp.35–46.
Smith, S.A. et al., 2011. Understanding angiosperm diversification using small and large phylogenetic trees. *American Journal of Botany*, 98(3), pp.404–414.
Sohn, J.-C. et al., 2012. An annotated catalog of fossil and subfossil Lepidoptera (Insecta: Holometabola) of the world. *Zootaxa*, 3286, pp.1–132.
Sokolovska, N., Rowe, L. & Johansson, F., 2000. Fitness and body size in mature odonates. *Ecological Entomology*, 25(2), pp.239–248.
Song, H., 2010. Grasshopper systematics: past, present and future. *Journal of Orthoptera Research*, 19(1), pp.57–68.
Song, N. & Liang, A.-P., 2013. A Preliminary Molecular Phylogeny of Planthoppers (Hemiptera: Fulgoroidea) Based on Nuclear and Mitochondrial DNA Sequences. *PLoS ONE*, 8(3), p.e58400.
Song, N., Liang, A.-P. & Bu, C.-P., 2012. A Molecular Phylogeny of Hemiptera Inferred from Mitochondrial Genome Sequences. *PLoS ONE*, 7(11), p.e48778.
Soubrier, J. et al., 2012. The Influence of Rate Heterogeneity among Sites on the Time Dependence of Molecular Rates. *Molecular Biology and Evolution*, 29(11), pp.3345–3358.

Spalletti, L.A., Fanning, C.M. & Rapela, C.W., 2008. Dating the Triassic continental rift in the southern Andes: the Potrerillos Formation, Cuyo Basin, Argentina. *Geologica Acta*, 6(3), pp.267–283.
Springer, M.S. et al., 2004. The use of composite taxa in supermatrices. *Molecular Phylogenetics and Evolution*, 30(3), pp.883–884.
Stadler, T., 2011a. Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences*, 108(15), pp.6187–6192.
Stadler, T., 2011b. Simulating Trees with a Fixed Number of Extant Species. *Systematic Biology*, 60(5), pp.676–684.
Stadler, T. & Bokma, F., 2012. Estimating speciation and extinction rates for phylogenies of higher taxa. *Systematic Biology*, 62(2), pp.220–230.
Stamatakis, A., Hoover, P. & Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology*, 57(5), pp.758–771.
Stamatakis, A., Ludwig, T. & Meier, H., 2005. RAxML-III: A fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics*, 21(4), pp.456–463.
Stanley, S.M., 1998. *Macroevolution, pattern and process*, Johns Hopkins University Press.
Stewart, W., 1980. The Australian genus Diphlebia Selys (Odonata : Amphipterygidae). I. Taxonomic revision of the adults. *Australian Journal of Zoology Supplementary Series*, 28(75), pp.1–57.
Stireman, J.O., 2005. The evolution of generalization? Parasitoid flies and the perils of inferring host range evolution from phylogenies. *Journal of Evolutionary Biology*, 18(2), pp.325–336.
Sukatsheva, I.D. & Jarzembowski, E.A., 2001. Fossil caddisflies (Insecta: Trichoptera) from the Early Cretaceous of southern England II. *Cretaceous Research*, 22(6), pp.685–694.
Sullivan, J. & Joyce, P., 2005. Model Selection in Phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, 36, pp.445–466.
Sun, G. et al., 2011. A eudicot from the Early Cretaceous of China. *Nature*, 471(7340), pp.625–628.
Suzuki, Y., Glazko, G.V. & Nei, M., 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proceedings of the National Academy of Sciences*, 99(25), pp.16138 –16143.
Svenson, G.J. & Whiting, M.F., 2004. Phylogeny of Mantodea based on molecular data: evolution of a charismatic predator. *Systematic Entomology*, 29(3), pp.359–370.
Svenson, G.J. & Whiting, M.F., 2009. Reconstructing the origins of praying mantises (Dictyoptera, Mantodea): the roles of Gondwanan vicariance and morphological convergence. *Cladistics*, 25(5), pp.468–514.
Swisher, C.C. et al., 1999. Cretaceous age for the feathered dinosaurs of Liaoning, China. *Nature*, 400(6739), pp.58–61.
Swofford, D. et al., 1996. Chapter 11: Phylogenetic inference. In D. M. Hillis, C. Moritz, & B. Mable, eds. *Molecular Systematics*. Sunderland: Sinauer Associates, Incorporated Publishers, pp. 407–514.
Szumik, C., Edgerly, J.S. & Hayashi, C.Y., 2008. Phylogeny of embiopterans (Insecta). *Cladistics*, 24(6), pp.993–1005.
Szwedo, J., 2007. Glisachaemus jonasdamzeni gen. et sp. nov. of Cixiidae from the Eocene Baltic amber (Hemiptera: Fulgoromorpha). *Alavasia*, 1, pp.109–116.
Szwedo, J. & Nel, A., 2011. The oldest aphid insect from the Middle Triassic of the Vosges, France. *Acta Palaeontologica Polonica*, 56(4), pp.757–766.

Szwedo, J., Wang, B. & Zhang, H., 2011. An extraordinary Early Jurassic planthopper from Hunan (China) representing a new family Qiyangiricaniidae fam. nov. (Hemiptera: Fulgoromorpha: Fulgoroidea). *Acta Geologica Sinica - English Edition*, 85(4), pp.739–748.
Talavera, G. & Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, 56(4), pp.564 –577.
Talavera, G. & Vila, R., 2011. What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta class phylogeny. *BMC Evolutionary Biology*, 11(1), p.315.
Templeton, A.R., 2010. Coherent and incoherent inference in phylogeography and human evolution. *Proceedings of the National Academy of Sciences*, 107(14), pp.6376–6381.
Templeton, A.R., 2009. Statistical Hypothesis Testing in Intraspecific Phylogeography: NCPA versus ABC. *Molecular ecology*, 18(2), pp.319–331.
Terry, M.D., 2003. *Phylogeny of the Polyneopterous Insects With Emphasis on Plecoptera: Molecular and Morpological Evidence*, PhD thesis Department of Integrative Biology, Brigham Young University. Available at: http://scholarsarchive.byu.edu/etd/1134/.
Terry, M.D. & Whiting, M.F., 2005. Mantophasmatodea and phylogeny of the lower neopterous insects. *Cladistics*, 21(3), pp.240–257.
Thomas, G.H. & Freckleton, R.P., 2012. MOTMOT: models of trait macroevolution on trees. *Methods in Ecology and Evolution*, 3(1), pp.145–151.
Thomas, J.A. et al., 2013. Relaxed phylogenetics and the Palaeoptera problem: Resolving deep ancestral splits in the insect phylogeny. *Systematic Biology*, 62(2), pp.285–297.
Thomas, M.A. et al., 2000. Molecular phylogenetic analysis of evolutionary trends in stonefly wing structure and locomotor behavior. *Proceedings of the National Academy of Sciences*, 97(24), pp.13178–13183.
Thompson, J.N., 2009. The Coevolving Web of Life. *The American Naturalist*, 173(2), pp.125–140.
Thorne, J.L. & Kishino, H., 2002. Divergence Time and Evolutionary Rate Estimation with Multilocus Data. *Systematic Biology*, 51(5), pp.689 –702.
Timmermans, M., Roelofs, D. & van Straalen, N., 2008. Revealing pancrustacean relationships: Phylogenetic analysis of ribosomal protein genes places Collembola (springtails) in a monophyletic Hexapoda and reinforces the discrepancy between mitochondrial and nuclear DNA markers. *BMC Evolutionary Biology*, 8, p.83.
Tojo, K. & Matsukawa, K., 2003. A Description of the Second species of the Family Dipteromimidae (Insecta, Ephemeroptera), and Genetic Relationship of Two Dipteromimid Mayflies Inferred from Mitochondrial 16S rRNA Gene Sequences. *Zoological Science*, 20(10), pp.1249–1259.
Trautwein, M.D. et al., 2012. Advances in insect phylogeny at the dawn of the postgenomic era. *Annual Review of Entomology*, 57(1), pp.449–468.
Trautwein, M.D., Wiegmann, B.M. & Yeates, D.K., 2010. A multigene phylogeny of the fly superfamily Asiloidea (Insecta): Taxon sampling and additional genes reveal the sister-group to all higher flies (Cyclorrhapha). *Molecular Phylogenetics and Evolution*, 56(3), pp.918–930.
Treherne, R.C., 1919. Notes on the Aeolothripidae. *Journal of the Entomological Society of British Columbia*, 12(0), pp.27–33.
Uchifune, T. & Machida, R., 2005. Embryonic development of Galloisiana yuasai Asahina, with special reference to external morphology (Insecta: Grylloblattodea). *Journal of Morphology*, 266(2), pp.182–207.

Ulrich, W., 2007. Body weight distributions of central European Coleoptera. *European Journal of Entomology*, 104, pp.769–776.
Ulrich, W., 2006. Body weight distributions of European Hymenoptera. *Oikos*, 114(3), pp.518–528.
Valentine, J.W., 1980. Determinants of Diversity in Higher Taxonomic Categories. *Paleobiology*, 6(4), pp.444–450.
Vamosi, J.C., Armbruster, W.S. & Renner, S.S., 2014. Evolutionary ecology of specialization: insights from phylogenetic analysis. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1795), p.20142004.
Vamosi, S. & Vamosi, J., 2005. Endless tests: guidelines for analyzing non-nested sister group comparisons. *Evolutionary Ecology Research*, (7), pp.567–579.
Vasilenko, D.V., 2005. New Damselflies (Odonata: Synlestidae, Hemiphlebiidae) from the Mesozoic Transbaikalian Locality of Chernovskie Kopi. *Paleontological Journal*, 39(3), p.280.
Vas, Z., Csorba, G. & Rózsa, L., 2012. Evolutionary co-variation of host and parasite diversity—the first test of Eichler's rule using parasitic lice (Insecta: Phthiraptera). *Parasitol Research*, 111, pp.393–401.
Vilhelmsen, L., Mikó, I. & Krogmann, L., 2010. Beyond the wasp-waist: structural diversity and phylogenetic significance of the mesosoma in apocritan wasps (Insecta: Hymenoptera). *Zoological Journal of the Linnean Society*, 159(1), pp.22–194.
Vogel, S., 1994. *Life in Moving Fluids: The Physical Biology of Flow*, Princeton University Press.
Vršanský, P., 2005. Lower Cretaceous cockroaches and mantids (Insecta: Blattaria, Mantodea) from the Sharin-Gol in Mongolia. *Entomological Problems,*, 35(2), pp.163–167.
Vršanský, P., 2008. New Blattarians and a Review of Dictyopteran Assemblages from the Lower Cretaceous of Mongolia. *Acta Palaeontologica Polonica*, 53(1), pp.129–136.
Waage, J.K., 1979. The evolution of insect/vertebrate associations. *Biological Journal of the Linnean Society*, 12(3), pp.187–224.
Walker, T., 2013. Family Prophalangopsidae (hump-winged grigs) in North America north of Mexico. In *Encyclopedia of Life*. Available at: http://eol.org/data_objects/25214510.
Wall, R. & Shearer, D., 2008. *Veterinary Ectoparasites: Biology, Pathology and Control*, John Wiley & Sons.
Wang, X. et al., 2005. Stratigraphy and age of the Daohugou Bed in Ningcheng, Inner Mongolia. *Chinese Science Bulletin*, 50(20), pp.2369–2376.
Wang, Y. et al., 2013. A Unique Box in 28S rRNA Is Shared by the Enigmatic Insect Order Zoraptera and Dictyoptera. *PLoS ONE*, 8(1), p.e53679.
Wan, X. et al., 2012. Complete mitochondrial genome of the free-living earwig, Challia fletcheri (Dermaptera: Pygidicranidae) and phylogeny of Polyneoptera. *PLoS ONE*, 7(8), p.e42056.
Wardhaugh, C.W., 2013. Estimation of biomass from body length and width for tropical rainforest canopy invertebrates: Invertebrate biomass measures. *Australian Journal of Entomology*, 52(4), pp.291–298.
Ware, J.L. et al., 2008. Relationships among the major lineages of Dictyoptera: the effect of outgroup selection on dictyopteran tree topology. *Systematic Entomology*, 33(3), pp.429–450.
Warnock, R.C.M., Yang, Z. & Donoghue, P.C.J., 2011. Exploring uncertainty in the calibration of the molecular clock. *Biology Letters*, 8(1), pp.156–159.
Wasserman, S.S. & Mitter, C., 1978. The relationship of body size to breadth of diet in some Lepidoptera. *Ecological Entomology*, 3(2), pp.155–160.

Watson, J.A.L., 1974. The distributions of the Australian dragonflies (Odonata). *Journal of the Australian Entomological Society*, 13, pp.137–149.
Webb, C.O., Ackerly, D.D. & Kembel, S.W., 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, 24(18), pp.2098–2100.
Weber, M.G. & Agrawal, A.A., 2014. Defense mutualisms enhance plant diversification. *Proceedings of the National Academy of Sciences*, 111(46), pp.16442–16447.
Wedmann, S., Bradler, S. & Rust, J., 2007. The first fossil leaf insect: 47 million years of specialized cryptic morphology and behavior. *Proceedings of the National Academy of Sciences*, 104(2), pp.565–569.
Wedmann, S. & Makarkin, V.N., 2007. A new genus of Mantispidae (Insecta: Neuroptera) from the Eocene of Germany, with a review of the fossil record and palaeobiogeography of the family. *Zoological Journal of the Linnean Society*, 149(4), pp.701–716.
Weeks, B.C. & Claramunt, S., 2014. Dispersal has inhibited avian diversification in Australasian archipelagoes. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1791), p.20141257.
Wellborn, G.A. & Langerhans, R.B., 2015. Ecological opportunity and the adaptive diversification of lineages. *Ecology and Evolution*, 5(1), pp.176–195.
Whalley, P. & Jarzembowski, E.A., 1981. A new assessment of Rhyniella, the earliest known insect, from the Devonian of Rhynie, Scotland. *Nature*, 291(5813), pp.317–317.
Wheat, C. & Wahlberg, N., 2013. Phylogenomic insights into the Cambrian Explosion, the colonization of land and the evolution of flight in Arthropoda. *Systematic Biology*, 62(1), pp.93–109.
Wheeler, Q.D., 2008. Undisciplined thinking: morphology and Hennig's unfinished revolution. *Systematic Entomology*, 33(1), pp.2–7.
Wheeler, W. et al., 2006. *Dynamic Homology and Phylogenetic Systematics: A Unified Approach Using POY*, American Museum of Natural History.
Wheeler, W., 1995. Sequence Alignment, Parameter Sensitivity, and the Phylogenetic Analysis of Molecular Data. *Systematic Biology*, 44(3), pp.321–331.
Wheeler, W. et al., 2001. The Phylogeny of the Extant Hexapod Orders. *Cladistics*, 17(2), pp.113–169.
Whitfield, J.B. & Kjer, K.M., 2008. Ancient Rapid Radiations of Insects: Challenges for Phylogenetic Analysis. *Annual Review of Entomology*, 53, pp.449–472.
Whiting, M.F. et al., 2008. A molecular phylogeny of fleas (Insecta: Siphonaptera): origins and host associations. *Cladistics*, 24(5), pp.677–707.
Whiting, M.F., 2002. Mecoptera is paraphyletic: multiple genes and phylogeny of Mecoptera and Siphonaptera. *Zoologica Scripta*, 31(1), pp.93–104.
Whiting, M.F., 1998. Phylogenetic position of the Strepsiptera: Review of molecular and morphological evidence. *International Journal of Insect Morphology and Embryology*, 27(1), pp.53–60.
Whiting, M.F. et al., 1997. The Strepsiptera Problem: Phylogeny of the Holometabolous Insect Orders Inferred from 18S and 28S Ribosomal DNA Sequences and Morphology. *Systematic Biology*, 46(1), pp.1–68.
Whiting, M.F. & Wheeler, W.C., 1994. Insect homeotic transformation. *Nature*, 368(6473), p.696.
Whitman, D.W., 2008. The Significance of Body Size in the Orthoptera: A Review. *Journal of Orthoptera Research*, 17(2), pp.117–134.
Wichard, W. & Engel, M.S., 2006. A new alderfly in Baltic amber (Megaloptera, Sialidae). *American Museum Novitates*, 3513, pp.1–9.

Wichard, W., Ross, E. & Ross, A.J., 2011. Palerasnitsynus gen. n. (Trichoptera, Psychomyiidae) from Burmese amber. *ZooKeys*, (130), pp.323–330.
Wiegmann, B.M., 1989. *A Phylogenetic Revision of the Family Atelestidae (diptera: Empidoidea) and Its Implications for the Origin of the Cyclorrhaphous Diptera*, University of Maryland, College Park.
Wiegmann, B.M. et al., 2011. Episodic radiations in the fly tree of life. *Proceedings of the National Academy of Sciences*, 108, pp.5690–5695.
Wiegmann, B.M. et al., 2000. Nuclear Genes Resolve Mesozoic-Aged Divergences in the Insect Order Lepidoptera. *Molecular Phylogenetics and Evolution*, 15(2), pp.242–259.
Wiegmann, B.M., Trautwein, M.D., et al., 2009. Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC Biology*, 7, pp.34–34.
Wiegmann, B.M., Kim, J.-W. & Trautwein, M.D., 2009. Holometabolous insects (Holometabola). In S. B. Hedges & S. Kumar, eds. *The Timetree of Life*. Oxford University Press, pp. 260–263.
Wiegmann, B.M., Mitter, C. & Farrell, B., 1993. Diversification of Carnivorous Parasitic Insects: Extraordinary Radiation or Specialized Dead End? *The American Naturalist*, 142(5), pp.737–754.
Wiegmann, B.M., Regier, J.C. & Mitter, C., 2002. Combined molecular and morphological evidence on the phylogeny of the earliest lepidopteran lineages. *Zoologica Scripta*, 31(1), pp.67–81.
Wiens, J.J., 2005. Can Incomplete Taxa Rescue Phylogenetic Analyses from Long-Branch Attraction? *Systematic Biology*, 54(5), pp.731–742.
Wiens, J.J. et al., 2005. Hylid frog phylogeny and sampling strategies for speciose clades. *Systematic Biology*, 54(5), pp.778–807.
Wiens, J.J., 2003a. Incomplete taxa, incomplete characters, and phylogenetic accuracy: is there a missing data problem? *Journal of Vertebrate Paleontology*, 23(2), pp.297–310.
Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. *Journal of Biomedical Informatics*, 39(1), pp.34–42.
Wiens, J.J., 2003b. Missing data, incomplete taxa, and phylogenetic accuracy. *Systematic Biology*, 52(4), pp.528 –538.
Wiens, J.J., 2011. The causes of species richness patterns across space, time, and clades and the role of "ecological limits." *The Quarterly Review of Biology*, 86(2), pp.75–96.
Wiens, J.J. & Morrill, M.C., 2011. Missing Data in Phylogenetic Analysis: Reconciling Results from Simulations and Empirical Data. *Systematic Biology*, 60(5), pp.719–731.
Wiggins, G.B., 2004. *Caddisflies: the underwater architects*, University of Toronto Press.
Wiggins, G.B., 1998. *The caddisfly family Phryganeidae (Trichoptera)*, University of Toronto Press.
Wiggins, G.B., 1956. *The Kitagamiidae : a family of caddisflies new to North America (Trichoptera)*, Toronto : Royal Ontario Museum. Available at: http://archive.org/details/kitagamiidaefami00wigg [Accessed April 8, 2014].
Wills, M.A., 2001. How good is the fossil record of arthropods? An assessment using the stratigraphic congruence of cladograms. *Geological Journal*, 36(3-4), pp.187–210.
Wilson, E.O., 2010. *The Creation: An Appeal to Save Life on Earth*, W. W. Norton & Company.
Wilson, H.M. & Martill, D.M., 2001. A new Japygid Dipluran from the Lower Cretaceous of Brazil. *Palaeontology*, 44(5), pp.1025–1031.
Wilson, J.S. et al., 2013. Family-Level Divergences in the Stinging Wasps (Hymenoptera: Aculeata), with Correlations to Angiosperm Diversification. *Evolutionary Biology*, 40(1), pp.101–107.

Wilson, K.D.., undated. Dragonfly groups of conservation interest confined to the Oriental Region. Available at: http://www. asia-dragonfly.net/Ar ticles/ OrientalOdonateBiodiversity.pdf.
Winkler, I.S. et al., 2010. Distinguishing Agromyzidae (Diptera) leaf mines in the fossil record: New taxa from the Paleogene of North America and Germany and their evolutionary implications. *Journal of Paleontology*, 84(5), pp.935–954.
Winkler, I.S. & Mitter, C., 2008. The phylogenetic dimension of insect plant interactions: a review of recent evidence. In K. J. Tilmon, ed. *Specialization, Speciation, and Radiation: The Evolutionary Biology of Herbivorous Insects*. University of California Press, pp. 240–263.
Winterton, S.L., Hardy, N.B. & Wiegmann, B.M., 2010. On wings of lace: phylogeny and Bayesian divergence time estimates of Neuropterida (Insecta) based on morphological and molecular data. *Systematic Entomology*, 35(3), pp.349–378.
Wipfler, B. et al., 2011. On the head morphology of Grylloblattodea (Insecta) and the systematic position of the order, with a new nomenclature for the head muscles of Dicondylia. *Systematic Entomology*, 36(2), pp.241–266.
Woodward, G. et al., 2005. Body size in ecological networks. *Trends in Ecology & Evolution*, 20(7), pp.402–409.
Xia, X. et al., 2003. An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution*, 26(1), pp.1–7.
Xia, X. & Xie, Z., 2001. DAMBE: Software Package for Data Analysis in Molecular Biology and Evolution. *Journal of Heredity*, 92(4), pp.371–373.
Xie, Q. et al., 2008. 18S rRNA hyper-elongation and the phylogeny of Euhemiptera (Insecta: Hemiptera). *Molecular Phylogenetics and Evolution*, 47(2), pp.463–471.
Xie, Q. et al., 2009. Phylogenetic comparison of local length plasticity of the small subunit of nuclear rDNAs among all Hexapoda orders and the impact of hyper-length-variation on alignment. *Molecular Phylogenetics and Evolution*, 50(2), pp.310–316.
Xie, W. et al., 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology*, 60(2), pp.150 –160.
Xiong, Y. et al., 2008. Molecular phylogeny of Collembola inferred from ribosomal RNA genes. *Molecular phylogenetics and evolution*, 49(3), pp.728–735.
Yang, A.S., 2001. Modularity, evolvability, and adaptive radiations: a comparison of the hemi- and holometabolous insects. *Evolution & Development*, 3(2), pp.59–72.
Yang, Q., Makarkin, V.N. & Ren, D., 2010. Remarkable New Genus of Gumillinae (Neuroptera: Osmylidae) From the Jurassic of China. *Annals of the Entomological Society of America*, 103(6), pp.855–859.
Yang, X.-G. et al., 2012. New Middle Jurassic hangingflies (Insecta: Mecoptera) from Inner Mongolia, China. *Alcheringa: An Australasian Journal of Palaeontology*, 36(2), pp.195–201.
Yang, Z., 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution*, 24(8), pp.1586–1591.
Yang, Z. & Rannala, B., 2006. Bayesian Estimation of Species Divergence Times Under a Molecular Clock Using Multiple Fossil Calibrations with Soft Bounds. *Molecular Biology and Evolution*, 23(1), pp.212 –226.
Yao, Y. et al., 2011. New fossil Ochteridae (Hemiptera: Heteroptera: Ochteroidea) from the Upper Mesozoic of north-eastern China, with phylogeny of the family. *Systematic Entomology*, 36(3), pp.589–600.
Yao, Y. et al., 2012. Phylogeny of the infraorder Pentatomomorpha based on fossil and extant morphology, with description of a new fossil family from China. *PLoS ONE*, 7(5), p.e37289.

Yao, Y., Cai, W. & Ren, D., 2007. Pristinochterus gen. n. (Hemiptera: Ochteridae) from the Upper Mesozoic of northeastern China. *European Journal of Entomology*, 104(4), p.827.
Yeates, D.K., Cameron, S.L. & Trautwein, M., 2012. A view from the edge of the forest: recent progress in understanding the relationships of the insect orders. *Australian Journal of Entomology*, 51(2), pp.79–87.
Yoshizawa, K., 2010. Direct optimization overly optimizes data. *Systematic Entomology*, 35(2), pp.199–206.
Yoshizawa, K., 2011. Monophyletic Polyneoptera recovered by wing base structure. *Systematic Entomology*, 36(3), pp.377–394.
Yoshizawa, K., 2007. The Zoraptera problem: evidence for Zoraptera + Embiodea from the wing base. *Systematic Entomology*, 32(2), pp.197–204.
Yoshizawa, K. & Johnson, K.P., 2005. Aligned 18S for Zoraptera (Insecta): Phylogenetic position and molecular evolution. *Molecular Phylogenetics and Evolution*, 37(2), pp.572–580.
Yoshizawa, K. & Johnson, K.P., 2010. How stable is the "Polyphyly of Lice" hypothesis (Insecta: Psocodea)?: A comparison of phylogenetic signal in multiple genes. *Molecular Phylogenetics and Evolution*, 55(3), pp.939–951.
Yoshizawa, K. & Johnson, K.P., 2006. Morphology of male genitalia in lice and their relatives and phylogenetic implications. *Systematic Entomology*, 31(2), pp.350–361.
Yoshizawa, K. & Johnson, K.P., 2003. Phylogenetic position of Phthiraptera (Insecta: Paraneoptera) and elevated rate of evolution in mitochondrial 12S and 16S rDNA. *Molecular Phylogenetics and Evolution*, 29(1), pp.102–114.
Yoshizawa, K. & Johnson, K.P., 2014. Phylogeny of the suborder Psocomorpha: congruence and incongruence between morphology and molecular data (Insecta: Psocodea: "Psocoptera"). *Zoological Journal of the Linnean Society*, 171(4), pp.716–731.
Yoshizawa, K. & Lienhard, C., 2010. In search of the sister group of the true lice: A systematic review of booklice and their relatives, with an updated checklist of Liposcelididae (Insecta: Psocodea). *Arthropd Systematics and Phylogeny*, 68(2), pp.181–195.
Yoshizawa, K. & Saigusa, T., 2001. Phylogenetic analysis of paraneopteran orders (Insecta: Neoptera) based on forewing base structure, with comments on monophyly of Auchenorrhyncha (Hemiptera). *Systematic Entomology*, 26(1), pp.1–13.
Yule, G.U., 1925. A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 213(402-410), pp.21–87.
Yushuang, L. et al., 2008. Three New Stoneflies (Insecta: Plecoptera) from the Yixian Formation of Liaoning, China. *Acta Geologica Sinica - English Edition*, 82(2), pp.249–256.
Yu, W. et al., 2012. A Comparative Study of Mammalian Diversification Pattern. *International Journal of Biological Sciences*, 8(4), pp.486–497.
Yu, Y. et al., 2012. The first fossil Bark-Gnawing Beetle from the Middle Jurassic of Inner Mongolia, China (Coleoptera: Trogossitidae). *Annales Zoologici*, 62(2), pp.245–252.
Zborowski, P. & Edwards, T., 2007. *A Guide to Australian Moths*, Csiro Publishing.
Zdobnov, E.M. & Bork, P., 2007. Quantification of insect genome divergence. *Trends in Genetics*, 23(1), pp.16–20.
Zhang, J., 2012. New horseflies and water snipe-flies (Diptera: Tabanidae and Athericidae) from the Lower Cretaceous of China. *Cretaceous Research*, 36, pp.1–5.
Zhang, J. & Li, H., 2012. New taxa of snipe flies (Diptera: Brachycera: Rhagionidae) in the Daohugou Biota, China. *Paleontological Journal*, 46(2), pp.157–163.

Zhang, W., Yao, Y. & Ren, D., 2011. New shore bug (Hemiptera, Heteroptera, Saldidae) from the Early Cretaceous of China with phylogenetic analyses. *ZooKeys*, (130), pp.185–198.
Zhang, Z., 2011. *Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness*, Magnolia Press.
Zhao, J. et al., 2010. Transitional fossil earwigs - a missing link in Dermaptera evolution. *BMC Evolutionary Biology*, 10(1), p.344.
Zheng, Y. et al., 2011. Exploring Patterns and Extent of Bias in Estimating Divergence Time from Mitochondrial DNA Sequence Data in a Particular Lineage: A Case Study of Salamanders (Order Caudata). *Molecular Biology and Evolution*, 28(9), pp.2521–2535.
Zimmermann, D. et al., 2011. The function and phylogenetic implications of the tentorium in adult Neuroptera (Insecta). *Arthropod Structure & Development*, 40(6), pp.571–582.
Zloty, J. & Pritchard, G., 1997. Larvae and adults of Ameletus Mayflies (Ephemeroptera Ameletidae) from Alberta. *The Canadian Entomologist*, 129, pp.251–289.
Zompro, O., 1999. Microphasma, a new stick insect genus from Sri Lanka. (Phasmatodea: Pachymorphinae). *Entomological Journal*, 109(3), pp.124–127.
Zompro, O., 1998. New phasmids from Venezuela and Ecuador Phasmatodea Neue Phasmiden aus Venezuela und Ecuador Phasmatodea. *Entomologische Zeitschrift*, 10811, pp.456–459.
Zompro, O., 2001. Philippine phasmids from the collection of the Staatliches Museums für Tierkunde, Dresden (Insecta: Phasmatodea). *Reichenbachia, v.34, 49-56 (2001)*.
Zompro, O., 2004. *Revision of the Genera of the Areolatae, Including the Status of Timema and Agathemera (Insecta, Phasmatodea)*, Goecke & Evers.
Zrzay, J., 2008. Four chapters about the monophyly of insect "orders": A review of recent phylogenetic contributions. *Acta Entomologica Musei Nationalis Pragae*, 48(2), pp.217–232.
Zuckerkandl, E. & Pauling, L., 1962. Molecular disease, evolution and genetic heterogeneity. In *Kasha M. and Pullman B. (Eds.) Horizons in Biochemistry*. New York: Academic Press.
Zwick, P., 2000. Phylogenetic System and Zoogeography of the Plecoptera. *Annual Review of Entomology*, 45(1), pp.709–746.