

# The First Molecular Phylogeny of Strepsiptera (Insecta) Reveals an Early Burst of Molecular Evolution Correlated with the Transition to Endoparasitism

Dino P. McMahon\*, Alexander Hayward, Jeyaraney Kathirithamby\*

Department of Zoology, University of Oxford, Oxford, United Kingdom

## Abstract

A comprehensive model of evolution requires an understanding of the relationship between selection at the molecular and phenotypic level. We investigate this in Strepsiptera, an order of endoparasitic insects whose evolutionary biology is poorly studied. We present the first molecular phylogeny of Strepsiptera, and use this as a framework to investigate the association between parasitism and molecular evolution. We find evidence of a significant burst in the rate of molecular evolution in the early history of Strepsiptera. The evolution of morphological traits linked to parasitism is significantly correlated with the pattern in molecular rate. The correlated burst in genotypic-phenotypic evolution precedes the main phase of strepsipteran diversification, which is characterised by the return to a low and even molecular rate, and a period of relative morphological stability. These findings suggest that the transition to endoparasitism led to relaxation of selective constraint in the strepsipteran genome. Our results indicate that a parasitic lifestyle can affect the rate of molecular evolution, although other causal life-history traits correlated with parasitism may also play an important role.

**Citation:** McMahon DP, 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): e21206. doi:10.1371/journal.pone.0021206

**Editor:** Hector Escriva, Laboratoire Arago, France

**Received:** February 1, 2011; **Accepted:** May 23, 2011; **Published:** June 28, 2011

**Copyright:** © 2011 McMahon et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by The Leverhulme Trust (F/08 502/G). DPM was supported by an Elizabeth Hannah Jenkinson Fund for laboratory funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: dino.mcmahon@googlemail.com (DPM); jeyaraney.kathirithamby@zoo.ox.ac.uk (JK)

## Introduction

A central focus in evolutionary research is the interaction between molecular evolution and selection at the level of the phenotype, the interface of which unifies aspects of evolutionary research often examined independently [1]. Such an approach offers insight into the factors shaping the rate of molecular evolution, and into the link between genome evolution and species divergence [2]. Strepsiptera is an order of insect parasitoids which display a variety of unusual genetic and phenotypic features [3–8]. Targeting groups with complex biologies such as Strepsiptera is useful for testing the validity and generality of ecological and evolutionary theory [9,10]. However, insufficient molecular data have prevented the study of a number of interesting questions, such as the relationship between genotypic and phenotypic evolution. Strepsiptera display characteristics that are close to the parasite/parasitoid boundary [4,11]. Female morphology is highly derived (eyes, antennae, mouthparts, legs, wings and reproductive characters are lost) and accompanied by an endoparasitic lifestyle that is host-dependent throughout the lifecycle (except for the family Mengerillidae). In contrast the male is free-flying as an adult and possesses typical insect characteristics. Strepsiptera infect a broad range of hosts, and are recorded from at least 34 families of insects distributed across 7 orders [3,4,12]. As with many other parasitic taxa, relatively little research has examined the evolution of host-usage in the group and its effect on speciation. However, recognition of the contribution of parasitic taxa to total animal diversity [13] has

emphasised the need to understand the basis of parasite diversification and host usage [14].

The relationship between genotype and phenotype can be examined in a variety of ways. Positive selection in candidate gene phylogenies has been paired with extant phenotypic traits on terminal or internal branches of a phylogeny [15,16]. Alternatively, a null model of evolution can be compared against models that specify positive selection [17]. These methods specifically target associations in genes responsible for particular phenotypic adaptations (i.e. those under positive selection). Another approach focuses instead on the differing molecular evolutionary rates between species, and in identifying potential life-history traits that influence molecular evolution. Understanding is limited by the availability of suitable methodology, since the field has emerged recently in response to increased DNA sequence data [2]. Investigations have searched for meaningful associations between the rate of molecular evolution and key phenotypic or other extrinsic factors [18,19], or between phenotypic factors and significant shifts in the pattern of lineage diversification [20]. Results from such studies must be carefully interpreted [21,22], due to errors associated with phylogeny and rate estimation, or ancestral state reconstruction [1,2,23].

Here we investigate the nature and underlying cause of a common feature of higher-level insect phylogenetic analyses: the long-branch separating Strepsiptera from other insect groups [24–30]. We explore the link between the strepsipteran phenotype's evolution and: i) variation in the rate of molecular evolution; ii) the pattern of lineage diversification. We reconstruct the first robust

molecular phylogeny of Strepsiptera, and use this as a framework to investigate the history of morphological and host-use evolution. Key characteristics include the loss of compound eyes, antennae, legs, wings and reproductive structures in the female, and modifications to the legs and tarsi, and loss of mandibles in some males. The questions we address include the point at which strepsipteran traits evolved, if they emerged more than once and how they are associated with variation in molecular evolution. We explore the hypothesis that the parasitic lifestyle exerts an effect on the rate of molecular evolution [31], an assertion that few studies have so far been able to support [32,33]. To do so, we examine molecular rate variation in Strepsiptera, and establish a model of evolutionary history that encompasses both molecular and phenotypic evolution.

## Results

### Strepsiptera phylogeny

We used 41 strepsipteran taxa, across 16 genera, and data from four genes: the mitochondrial genes cytochrome c oxidase I (*cox1*), NADH dehydrogenase I (*nad1*), and *small subunit ribosomal RNA* (*16S rRNA*), and the nuclear gene *small subunit ribosomal RNA* (*18S rRNA*) to generate a final alignment of 3930 nucleotides. This consisted of 967 bp 803 bp, and 2160 bp for the *18S rRNA*, *16S rRNA*, and *cox1+nad1* partitions respectively.

Across all phylogenetic analyses, we identified monophyletic groupings for extant Strepsiptera, Stylopodia and Stylopiformia, corroborating the findings from previous studies [3,4,12,34] using molecular data for the first time. In the concatenated Bayesian Inference (BI) analyses, all nodes (bar one) at and above the family level receive 100 posterior support (Figure 1A). Maximum Likelihood (ML) bootstrap values are  $\geq 75/100$  with *Lychnocolax* (except one: node within grey oval); and  $\geq 80/100$  without *Lychnocolax*. We find Myrmecolacidae as the sister-clade to all remaining families within Stylopiformia, and a sister-group relationship between Stylopidae (which parasitize bees) and Xenidae (which parasitize crabronid, sphecids, eumenid and vespids wasps), and between Halictophagidae and Elenchidae that predominantly parasitize Auchenorrhyncha (a “true bug” group, containing amongst others the cicadas, leafhoppers, treehoppers, planthoppers and spittlebugs). The genus *Lychnocolax* has no host records, and was historically placed within Myrmecolacidae [3]. Here, we find evidence that it is an older taxon, as the sister-group to Stylopidae, Xenidae, Elenchidae and Halictophagidae. Its position in the analyses is supported by ribosomal nucleotide composition data (Figure S1), but this is only moderately supported in the concatenated phylogenetic analyses (Figure 1A, grey oval). Removal of *Lychnocolax* led to increased ML bootstrap support in a descendent node (Figure 1A; green oval). Genera are all returned as monophyletic except *Halictophagus*, which occurred as a poorly resolved polyphyletic grouping with *Tridactylophagus* and *Callipharixenos*. The latter species is placed within Halictophagidae, arguing against the separate family-status hypothesized for this lineage.

The dated phylogeny based on the MIT1+2 dataset (see methods) is given in Figure 2, with 95% credibility interval (CI) bars positioned over relevant nodes. The tree reflects the topology produced using the entire (concatenated) dataset presented in Figure 1 (based on the nuclear *18S rRNA*+ mitochondrial *16S rRNA*+ *cox1/nad1* partitions), although there is minor incongruence within Myrmecolacidae, Xenidae, Halictophagidae, Elenchidae. Differences between MIT1+2 and MIT123 (the mitochondrial dataset including 3<sup>rd</sup> codon positions) on date estimation was minor, with a marginal increase in 95% CIs using MIT1+2 (Table

S1). Imprecise CIs concentrate in regions less well informed by available fossil prior information.

### Molecular evolution

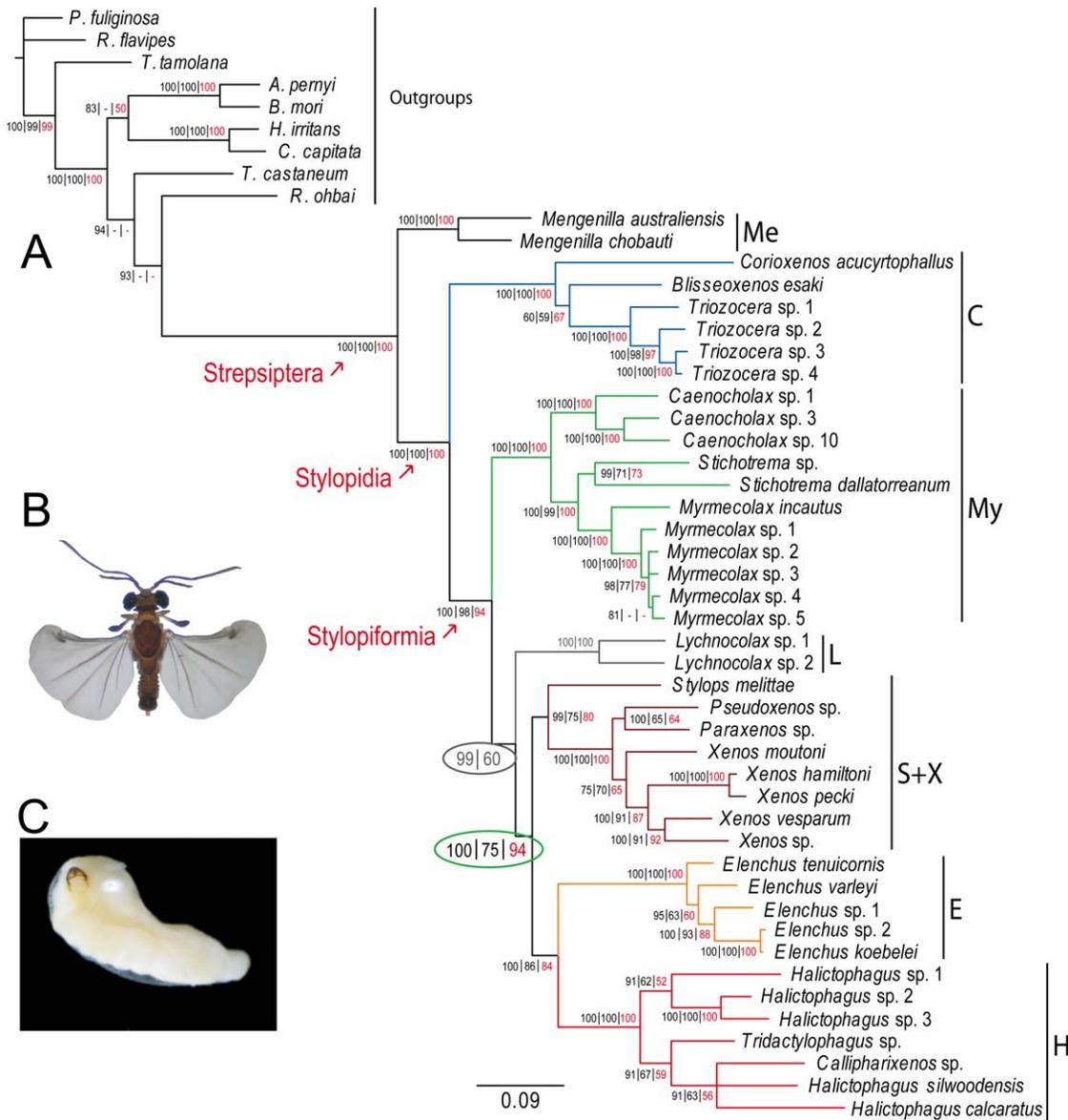
Firstly, we investigated the history of molecular rate across the MIT1+2 and MIT123 phylogenies. We found molecular rate estimates ranging from 1–1.5% pairwise sequence divergence per million years (Table S1) for analyses including 3<sup>rd</sup> codon positions. These are lower than the commonly cited value of 2.3% [35] and more in-keeping with 1.5% [36] for insect mitochondrial DNA and other rates reported for Strepsiptera [10]. However, within this overall pattern, the dated trees and relative rate analyses revealed significant variation in molecular rate, notably at the time to most recent common ancestor (tMRCA) of Strepsiptera and Stylopodia (Figure 2). Both nodes are associated with clades with high relative rates of molecular evolution. The rates at descendent strepsipteran nodes are lower, at around 1% pairwise sequence divergence per million years. The pattern of molecular evolution in the nuclear *18S rRNA* gene is in good overall agreement with the mitochondrial *cox1+nad1* gene (MIT1+2, MIT123) datasets (Table S2, Figure S2).

We then compared the evolution of molecular rate with an investigation into the history of diversification rate. The relative cladogenesis test (RCT) indicated a significant shift in Stylopodia and Stylopiformia (Figure 2). But statistics from the topological method in Symmetree were not significant, with upper and lower bound confidence intervals (CI) (at .025 and .975 percentiles) of 0.079–0.168 and 0.042–0.095 in the  $M_R$  and  $M_\Sigma$  tests respectively. Inclusion of 560 missing taxa produced p-value CIs for whole-tree test statistics ( $M_R$ ;  $I_C$ ;  $M_{IT}^*$ ;  $M_{IT}$ ;  $M_\Sigma^*$ ;  $M_\Sigma$ ;  $B_I$ ) between 0.001 - 0.000. This discrepancy could stem from over-representation of Mengenillidae and Corioxenidae, and under-representation of Stylopidae, Halictophagidae and Myrmecolacidae in the taxon set. In both analyses, individual nodes were not associated with rate shifts, with p-values of 0.121 and 0.209 and 0.107 and 0.185 (with missing taxa) for Stylopodia and Stylopiformia respectively. Furthermore, the branching pattern from the maximum clade credibility (MCC) tree did not depart significantly from a constant-rate/null speciation model: 0.999 ( $b = 0.5$ ,  $d = 0.5$ ,  $m = 560$ ); 0.998 ( $b = 0.5$ ,  $d = 0.0$ ,  $m = 560$ ); 0.836 ( $b = 0.5$ ,  $d = 0.5$ ,  $m = 60$ ); 0.701 ( $b = 0.5$ ,  $d = 0.0$ ,  $m = 60$ ). These results indicate that a burst of molecular rate evolution characterised the early evolution of Strepsiptera, but this did not coincide with a significant shift in lineage diversification.

### Reconstruction of the strepsipteran phenotype

Having established a basic framework for genotypic evolution, we directed attention towards understanding evolution of the strepsipteran phenotype. The morphological character reconstructions used in BI approach are summarized over the MIT1+2 phylogeny in Figure 2, details of the character state reconstructions that were recovered at each node can be found in Table S3. These corroborate the reconstruction of morphological evolution from a previous phylogeny using parsimony [34]. Both closely mirror the pattern of molecular rate depicted in Figure 2.

The long-branch leading to Strepsiptera is linked with phenotypic modifications relating to extreme sexual dimorphism, obligate endoparasitism in the larval stages, and entomophagy (consumption of insects as food). Stylopodia is associated with the evolution of the endoparasitic female (and the continuation of endoparasitism through pupation for males). In males, this node is linked to the reduction or loss of spiracles in the adult and larvae respectively, and the loss of pupal claws. Stylopiformia is associated with modifications to the tarsi, reduction of tarsal



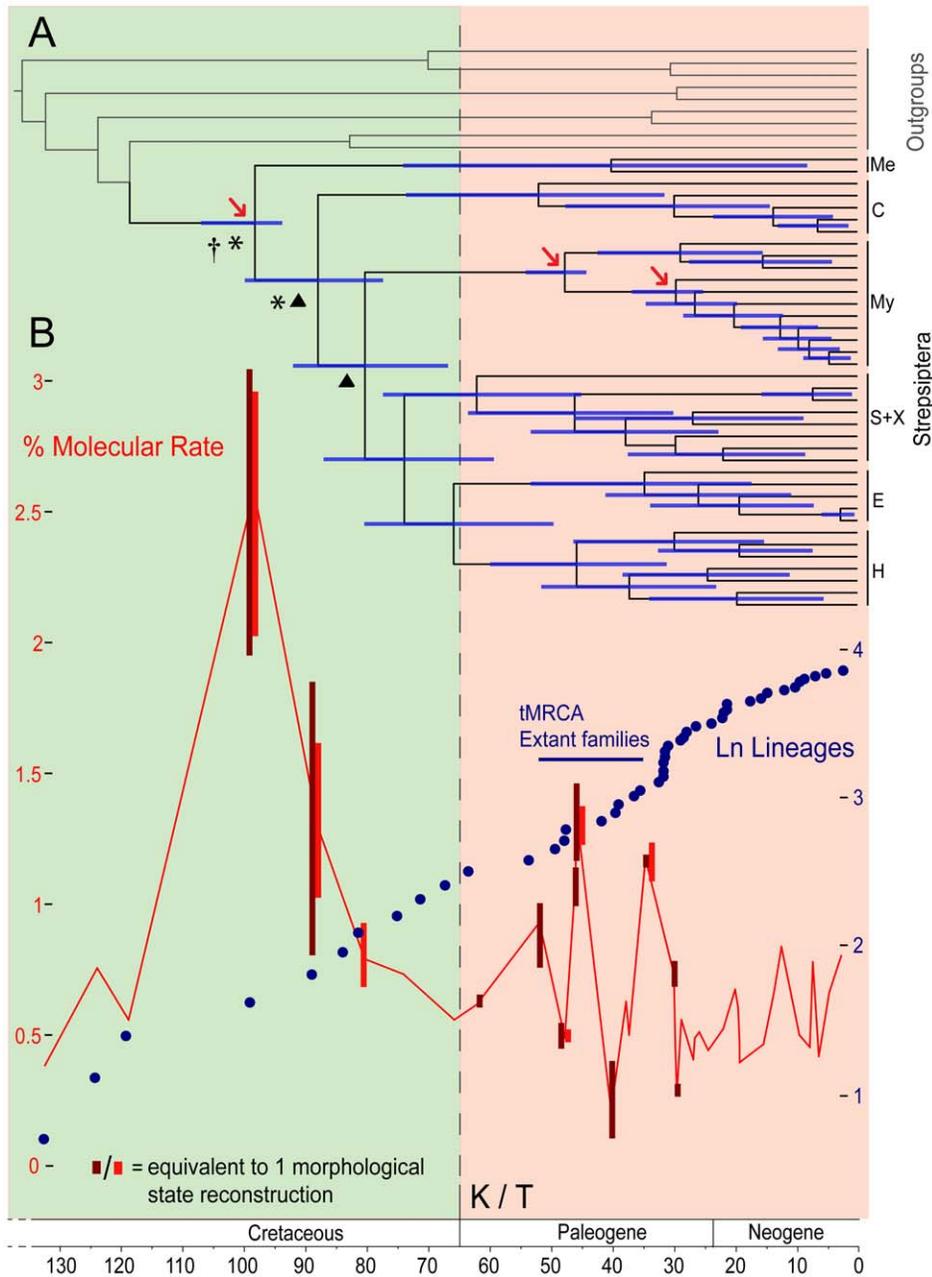
**Figure 1. Molecular phylogeny of Strepsiptera.** (A) Branch lengths from the BI 50% majority rule tree, with support values from BI and ML analyses appearing next to nodes. Support values (%) = BI posterior support | 1000 ML parametric bootstraps with *Lychnocolax* | 500 ML parametric bootstraps without *Lychnocolax*. Grey oval: support values = BI posterior support | 1000 ML parametric bootstraps including *Lychnocolax*. Green oval: increased ML support following removal of *Lychnocolax*. Me = Mengenillidae; C = Corioxenidae; My = Myrmecolacidae; L = *Lychnocolax*; S+X = Stylopidae+Xenidae; E = Elenchidae; H = Halictophagidae. (B) Male *Caenocholax fenyesi* sensu lato (C) Female *Caenocholax fenyesi* sensu lato [4]. doi:10.1371/journal.pone.0021206.g001

number and loss of larval legs in males, and the evolution of the cephalothorax in females.

The history of strepsipteran host-use is summarized in Figure 3. Parasitization of aculeate hymenopterans is predicted to have originated in Stylopiformia, or possibly earlier in the ancestor to Stylopida, where a secondary switch would be implicated in Corioxenidae to Heteroptera (a “true bug” group, containing amongst others the assassin bugs, bed bugs, seed bugs and shield bugs). In both models, a subsequent switch to Auchenorrhyncha in the ancestor of Elenchidae and Halictophagidae is strongly supported. Outside of Myrmecolacidae, host switching between infraordinal host groups occurs only in Halictophagidae. The ancestral host of Strepsiptera remains unresolved given currently available data.

### Comparison of molecular and phenotypic rates of evolution

We undertook a number of analyses to test the statistical validity of the association between molecular and phenotypic rates of evolution. A linear model indicated that molecular rate is positively associated with morphological branch length variation (T-statistic = 7.360, p-value = 2.29E-07). Molecular rate contributed the majority of variation in branch length (Adjusted R-squared = 0.698). However, non-linearity of error and heterogeneity of variance undermined the assumptions of a parametric statistical approach. The concentration of molecular and morphological rate evolution in the node leading to Strepsiptera represents a significant component of the skew in the distribution (Figure 4). We therefore re-examined the correlation by using a



**Figure 2. History of divergence and rate of molecular evolution in Strepsiptera.** (A) BI phylogeny using the MIT1+2 dataset calibrated against time. Node age 95% credibility intervals are indicated over nodes. †Increased relative rate of *18S rRNA*. \*Increased relative rate of MIT1+2. ▲Significant Relative Cladogenesis (RCT) statistics. Arrows indicate fossil calibrated nodes. Clade abbreviations follow Figure 1. (B) History of molecular rate using MIT1+2 scaled to the tree in panel A with number of ancestral character reconstructions at corresponding nodes (dark red bars=total non-homoplastic state changes in morphology under parsimony [34], red bars=total morphological reconstructions using Bayesian ancestral reconstruction). Blue: Log number of lineages at corresponding distance from root. tMRCA=time to Most Recent Common Ancestor; K/T=Cretaceous/Tertiary boundary. doi:10.1371/journal.pone.0021206.g002

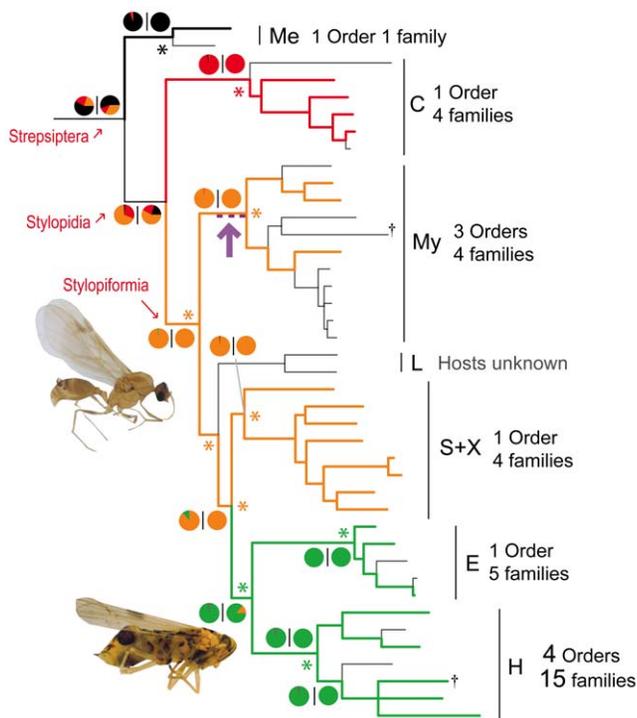
spearman test ( $\rho = 0.542$ ,  $S = 1054.628$ ,  $p\text{-value} = 0.003$ ) (Figure 4A).

We also took an alternative approach by making a whole-tree comparison of morphological versus molecular branch lengths (instead of molecular rate). We generated a null distribution of K-scores (mean = 0.579, S.D = 0.000134, min = 0.5780, max = 0.5794), and compared our K-test statistic of 0.553 against this distribution (see methods). The hypothesis that the observed K-score was due to random processes could be rejected ( $p$ -

value  $\ll 0.001$ ) (Figure 4B). Overall, these results indicate a correlated pattern of molecular and morphological evolution in Strepsiptera.

## Discussion

In this study, we develop a framework for understanding strepsipteran molecular evolutionary history, link it with existing knowledge of strepsipteran morphological evolution, and establish



**Figure 3. Bayesian reconstruction of male host-usage according to infra-ordinal grouping.** Known host records are given next to clades. Unshaded lines = unknown records/equivocal reconstructions. Black = Lepismatidae; Red = Heteroptera; Orange = Hymenoptera; Green = Auchenorrhyncha; Purple = possible origin of heteronomy. \*Significant node reconstructions using BFs. Pie charts = posterior probability | ML support. BFs and support charts not shown below family. †Probable parthenogens. Clade abbreviations follow Figure 1B. Images: *Pheidole* sp. (Hymenoptera) with male cephalotheca (top). *Sogatella furcifera* (Homoptera) with *Elenchus japonicas* male puparium. Photographs © J. Kathirithamby. doi:10.1371/journal.pone.0021206.g003

a foundation for further research into the evolution and ecology of this unusual host-parasite system. We retrieve high support for the monophyly of Strepsiptera, Stylopodia and Styloipiformia, and for interrelationships between the extant families. There remain areas of uncertainty, in particular the equivocal position of *Lychnocolax* and the polyphyly of *Halictophagus*, which require taxonomic revision through re-analysis of morphology and inclusion of additional taxa and alternative DNA markers.

### Molecular rate, diversification rate, and phenotypic evolution

We detected a significant shift in molecular rate in the early history of Strepsiptera. Instead of remaining uniformly high, the rate returned to a low and even rate across the phylogeny. Minor peaks in molecular rate are also linked with the diversification of lineages post K/T, in particular the tMRCAs of extant families Xenidae, Halictophagidae, and Elenchidae between 50–30 MYA. Interestingly, the K/T boundary was not closely linked with shifts in either molecular rate or diversification rate. A similar pattern has also been observed in other terrestrial animal groups, including mammals [37], squamates and passerine birds [38]. This trend reflects contemporaneous changes in strepsipteran morphology, which also changed significantly during the group's early history, followed by a period of stability in more recent history (<70 MYA), notwithstanding minor modifications to

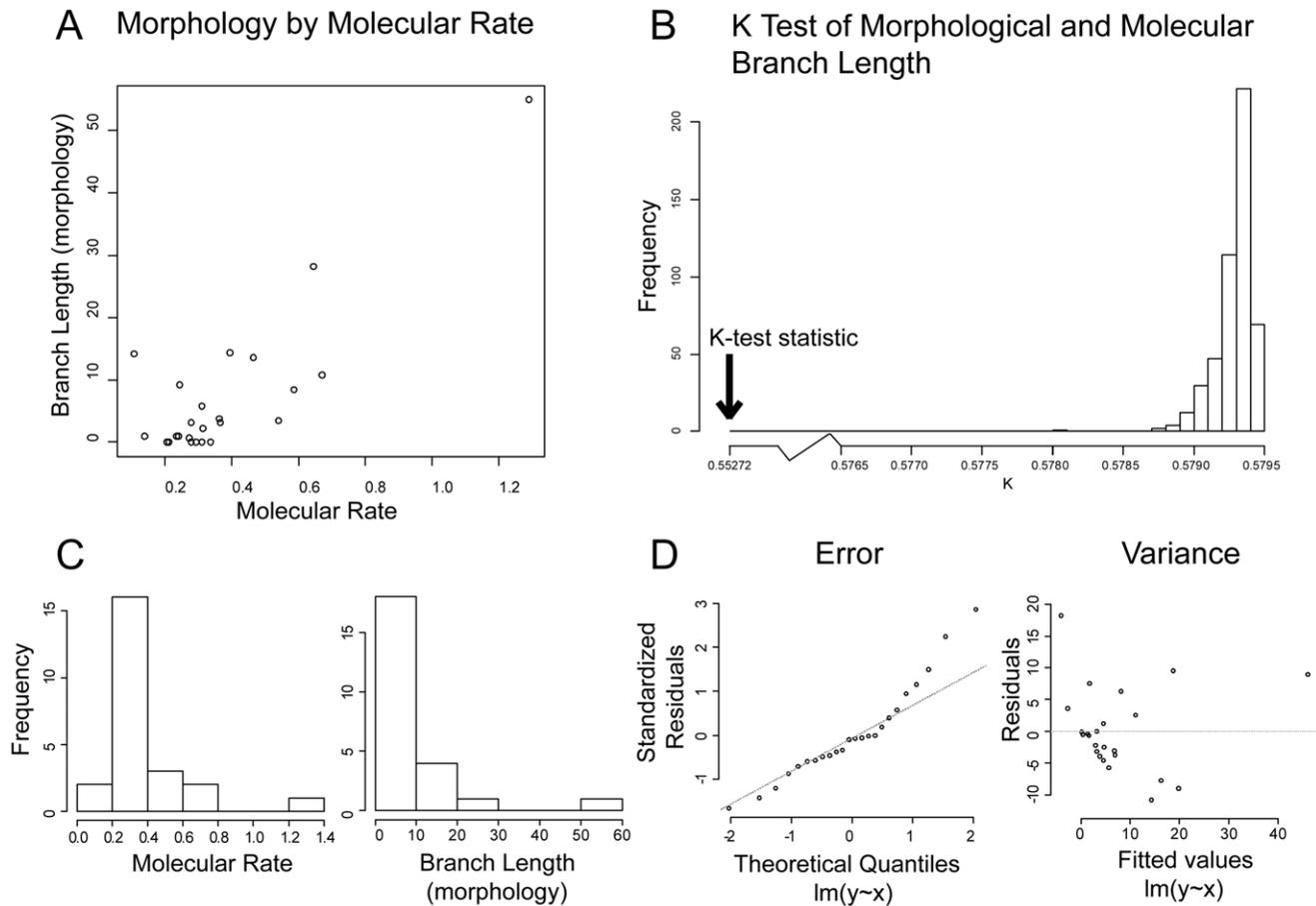
morphology linked with the tMRCAs of several extant families (Figure 2). Relationships within Styloipiformia are inconsistent with the only other (morphology-based) phylogenetic analysis of Strepsiptera. This important result may be due to the low number of non-homoplasious morphological state changes at intermediate depths of the prior study (Figure 29 in [34]). In the current study, inter-node distances are short compared with surrounding branches in the equivalent region of the tree (Figure 1). This is consistent with the individual node comparison of molecular rate and morphological branch length, which show a positive correlation.

The evolution of strepsipteran structural morphology involved a range of adaptations associated with increasingly specialised parasitism. Extant Mengerillidae represent a transitional condition, in that females reproduce and release progeny whilst outside of the host, with the faculty to leave the free-living pupa to lead a motile lifestyle [39]. We hypothesize that complete female endoparasitism in Stylopodia led to strong sexual selection on free-living adult males, which to copulate successfully must engage with highly modified female structures (the cephalothorax) protruding from living hosts. This may have led to the evolution of hairy adhesive tarsal pads; needed to adhere to diverse host substrates during insemination of the endoparasitic female [40].

The diversification rate in Strepsiptera increased after the initial burst in molecular rate, but a significant individual node shift was only detected in the RCT statistic, at the origin of Stylopodia/Styloipiformia. Lack of evidence for an increase in diversification across other methods suggests this result should be interpreted with caution. We refrain from discussing in depth which phenotypic traits (if any) might be causally linked to the main phase of strepsipteran diversification due to difficulties associated with identifying trait(s) that are responsible for speciation [41,42]. One might hypothesize that after endoparasitism became an obligate component of all aspects of female life-history (in the ancestor of Stylopodia), the evolution of a more effective method of host immune evasion may have enhanced the ability of Strepsiptera to successfully infect novel hosts, thereby opening opportunities for speciation. During infection, Strepsiptera are contained within a host-derived epithelial membrane, which is thought to conceal the endoparasite from the host's immune system [5]. However, its point of origin remains unknown.

### Did parasitism cause the burst in molecular rate?

We discovered a correlated burst in the rate of molecular and morphological evolution, which coincided with significant increases in the relative rate of molecular evolution, and abrupt shifts in the evolution of rRNA structures (Figure S1). We showed in an overall comparison of branch length that the observed similarity between molecular and morphological trees (K-score) was not due to random processes. The traits that evolved during the early history of Strepsiptera were broadly adaptations relating to the evolution to endoparasitism. These results are consistent with the hypothesis that parasitism may be an important cause of molecular evolutionary rate variation [31–33]. A plausible scenario could have involved deleterious mutations in free-living species becoming neutral/nearly-neutral in progressively host-dependent endoparasites. An adaptive interpretation could be that parasitism indirectly led to the selection of increased variation (through recombination or mutation), due to increased red-queen pressures between the host and parasite. But such a hypothesis does not explain why a subsequent decrease in substitution rate is observed in descendent nodes within Strepsiptera. Under the first (non adaptive) model, once most sites had been exposed to novel evolutionary forces, the substitution rate returned to a background



**Figure 4. Graphical summary of molecular rate and morphological branch length variation.** (A) Individual node comparison of molecular rate (pairwise sequence divergence / million years) versus morphological branch length (steps required under parsimony). (B) K-tree null distribution, with test-statistic indicated by arrow. (C) Distribution of molecular rate (left) and morphological branch length (right) variation. (D) Graphical summary of linear model assumptions: non-linearity of error (left) and heterogeneity of variance (right). doi:10.1371/journal.pone.0021206.g004

level. Our findings are more in-keeping with relaxation of selective constraint as a dominating force in the early evolution of Strepsiptera.

However, the precise relationship between molecular rate and parasitism cannot be conclusively resolved in the current framework. Confounding correlates of endoparasitism may prove causally more relevant [2]. For example, in studies of mammalian molecular evolution, the increase in availability of DNA sequence data has questioned initial hypotheses positing a simple correlation with body size. Later studies used more reliable rate estimates and better methods to demonstrate that rate in the nuclear genome covaried with generation time and fecundity (and body size) but that variation in the rate of the mitochondrial genome was explained by longevity [43] (although in our study, rates between nuclear and mitochondrial genomes are similar). An alternative explanation in Strepsiptera could be that endoparasitism enabled females to increase individual fecundity by being able to concentrate more resources on one aspect of life-history: reproduction. Higher mutation rate could have subsequently stemmed from the associated increase in germline replications per generation. Endoparasitism may have also been correlated with increased generation time and shorter lifespan, where pressures to reproduce prior to host-death or clearance are considered to be critical components of parasite evolution [44]. These factors could be causally important in explaining the evolution of molecular rate

in Strepsiptera, but determining which requires a more detailed understanding of life-history, ecology and the fine-scale interaction between Strepsiptera and host. Uncovering the mechanism of immune evasion could represent a particularly important target for future research.

### The Strepsiptera long-branch

This study indicates that elevated molecular evolutionary rate was an important contributing cause of the strepsipteran long-branch. However, missing data in the form of undiscovered extinct (or extant) transitional lineages and imprecision over the nearest extant sister-lineage are also relevant to improving understanding of the causes of molecular rate variation in Strepsiptera. A number of recent studies consolidate the view that Strepsiptera are closely related to Coleoptera [27–30] but a precise hypothesis has still not been reached. Increased knowledge of strepsipteran life-history and ecology, in combination with a more detailed understanding of strepsipteran sister-relationships, will lead to better estimations of divergence, allowing for more informative date priors to be incorporated into a relaxed phylogenetic approach [45,46]. Together, these will help to develop a more accurate picture of the forces responsible for variation in the rate of genome evolution in Strepsiptera. Revisiting hypotheses, like a possible association with rhipiphorine beetles [47] may help to identify potential candidate taxa that interrupt the branch. Alongside approaches

that implement more sensitive phylogenetic methodology and larger data sets [48], new data may offer greater understanding of strepsipteran origins. However, this study suggests that the strepsipteran long-branch may never be easy to “break up”.

## Conclusions

In this report, we present the first molecular phylogeny of Strepsiptera. Estimates of morphological branch length, alongside reconstruction of the strepsipteran phenotype reveal a correlation between morphological traits linked to endoparasitism and rate of molecular evolution. The main phase of diversification (Stylopodia, Stylopiformia) is associated with a return to a low and even rate of molecular evolution, and a period of relative morphological stability. This pattern supports the hypothesis that the transition to parasitism from a free-living insect ancestor can affect molecular rate. Greater precision over the nearest extant strepsipteran sister group will lead to better estimations of both divergence and molecular rate. Improved understanding of strepsipteran biology will in future permit the causes of molecular rate variation in Strepsiptera to be examined in greater detail. Together, these results establish an important foundation for further research into the evolution and ecology of a highly unusual host-parasite system.

## Materials and Methods

### Taxon and DNA sampling

Individuals were included from 41 strepsipteran taxa, across 16 genera (50% coverage). Bohartillidae and Bahiixenidae, which are rare and represented by few specimens, were not included [49]. Three hemi- and six holometabolous outgroup species were selected from nucleotide data in Genbank. Due to the nature of mitochondrial gene evolution in Hymenoptera [50,51] and the possibility of long-branch attraction between Strepsiptera and Hymenoptera (Hayward et al. in preparation; Figure 4 in [27]), the latter were not included. Specimens were preserved at 4°C in 95% ethanol, and protocols employed for sequence generation follow [8]. The mitochondrial genes cytochrome c oxidase I (*cox1*), NADH dehydrogenase I (*nad1*), and *small subunit ribosomal RNA* (*16S rRNA*), and nuclear gene *small subunit ribosomal RNA* (*18S rRNA*) were chosen to represent independent and variable evolutionary rates (Genbank accession JN082786–JN082922). Chromatograms were inspected manually using FinchTV ([www.geospiza.com](http://www.geospiza.com)), and *cox1* and *nad1* fragments were aligned by eye in BioEdit [52], using translated nucleotides to guide the management of indels. *16S* and *18S rRNA* fragments were aligned manually using the comparative structural method [6,53,54] and mfold [55], but these do not correspond strictly to the category definitions *sensu* Gillespie [54].

A final alignment consisting of 3930 nucleotides was used in subsequent analyses, consisting of 967 bp 803 bp, and 2160 bp for the *18S rRNA*, *16S rRNA*, and *cox1+nad1* partitions respectively, each with 339, 433, and 433 parsimony-informative positions. This approach was compared against an automated alignment strategy using the default settings in MUSCLE [56] and Gblocks [57], but retaining columns with a gap at greater than 50% of taxa. The resultant alignment contained 21% fewer characters (3048 nucleotides) of 642 bp, 774 bp and 1632 bp in the *18S rRNA*, *16S rRNA* and *cox1+nad1* partitions, each with 229, 546, and 341 parsimony informative positions respectively. All analyses in this study are based on the structurally-informed “manual” alignment as trees based on the automated approach produced trees with limited support and equivocal topologies (data not shown). For the estimation of molecular rates, divergence estimates and date-informed branch lengths, the mitochondrial *cox1* and *nad1* genes were combined into a single data partition and

analysed separately with/without the 3<sup>rd</sup> codon position (datasets MIT1+2 and MIT123 respectively). Specimen information (including accession numbers, primer information and 18S/16S rRNA template alignments) appears in Table S4.

### Evolutionary model selection and phylogenetic analysis

For Bayesian analyses (BI), the most appropriate models of evolution were selected by comparing harmonic means across separate gene partitions in MrBayes v3.1.2 [58,59], and then calculating Bayes Factor (BF) values. For maximum likelihood (ML), the Akaike Information Criterion (AIC) approach in MrModelTest v2 [60] and ProtTest v2.4 [61–63] was employed to select the most suitable models for RAxML v7.0.3 [64]. For all nucleotide partitions, the GTR+Γ+I model was preferred by BF and AIC with the following harmonic means: −7209.42 (*18S rRNA*); −9235.59 (*16S rRNA*); −14796.14 (MIT1+2); −30436.50 (MIT123) and log-likelihoods: −7160.8037; −9189.0898; −14758.0146; −31345.5293. For partition MIT123, the number of transitions and transversions estimated under the F84 model were plotted against genetic distance for each codon position using DAMBE v5.0.8 [65] (Figure S3, Table S5). A test of substitution saturation [66] and quartet likelihood mapping (TREE-PUZZLE v5.2) [67,68] indicated high percentages of noise versus signal (20.6% and 32.7% in *cox1* and *nad1* respectively) in synonymous 3<sup>rd</sup> codon position, and little correspondence between 3<sup>rd</sup> codon position transition frequencies and genetic distance. Consequently, the mixed amino acid model facilitated by MrBayes was selected for use in all concatenated BI analyses. The amino acid substitution model favoured by the posterior density was MtRev [69] +Γ+I (+F). ProtTest found highest support for MtArt [70] and LG [71], but these are unavailable to MrBayes v3.1.2 and RAxML v7.0.3.

After models were selected, concatenated BI analyses consisted of two independent (MC)<sup>3</sup> algorithms running for 2 million generations, each with four chains (3 hot, 1 cold), sampling one tree in 200, burn-in cutoffs were inspected manually for each parameter file in Tracer v1.4 [72]: the first 40000 steps were discarded. Inspection of the standard deviation of split frequencies confirmed that runs had converged (0.0059). All parameters except topology were unlinked between partitions. Data were summarized over a majority rule consensus tree (50% cutoff). 1000 ML nonparametric bootstrap pseudoreplicates were estimated in RAxML v7.0.3 [73]. 500 ML bootstrap pseudoreplicates without *Lychocolax* were also estimated. Trees were imported into FigTree v1.2.3 for editing [74].

### Divergence time estimation

MIT1+2 was employed in a Bayesian relaxed clock framework in BEAST v1.4.8 [75] using the GTR+Γ+I model. *Lychocolax* taxa were removed prior to analyses and Holometabola was constrained as monophyletic. Likelihood ratio tests using least and most complex evolutionary models in PAML v4 [76], with/without the 3<sup>rd</sup> codon position were overdispersed with respect to a molecular clock (2ΔlnL = 1398.37, 1539.43, 836.45, 948.41; *df* = 7, *P* < 0.001). Significant rate-heterogeneity was accommodated by employing the relaxed-clock MCMC with an uncorrelated lognormal model (UCLN) [77], calibrated using three strepsipteran fossils. The implementation of fossil priors is described in Text S1.

MCMC analyses ran for 10 million iterations, sampling every 1000<sup>th</sup> step. The effect of A+T-rich 3<sup>rd</sup> codon positions was investigated using the MIT123 dataset, in which the two partitions (1+2)(3) were unlinked. Analyses were repeated using the *18S rRNA* dataset. The effect of model choice was assessed by comparing

GTR+ $\Gamma$ +I with the SRD framework [78]. Molecular rate estimates were calculated as % pairwise sequence divergences per million years: equal to twice the per lineage rate. Dates were specified as millions of years before present, the Yule process was employed as the tree prior. Parameter files were inspected manually to ensure chain stability across parameters, and to select an appropriate burn-in. Tree files were summarized on a maximum clade credibility (MCC) tree.

### Molecular rate, diversification rate, and tree shape

Relative rates of molecular evolution were examined via cross comparison of families and outgroups in RRTree v1.1.11 [79]. Whole-tree and single node methods were employed to test for departures in diversification rate [39]. The temporal rate cladogenesis test (RCT) statistic was calculated [80,81]: nodes showing a “trickle-down” effect [39] were excluded. Whole tree simulations of rate-constant/rate-variable variants of the birth-death model in the Laser R-package [82] were conducted. The simulation ran for 1000 trees, comparing the best *constant* speciation model versus best *variable* speciation model ( $\Delta$ AICr) using the MIT1+2 tree. Outgroup taxa were pruned, and the birth rate ( $b$ ), death rate ( $d$ ) and unsampled taxa ( $m$ ) were varied.  $m$  represents unsampled Strepsiptera species diversity. SymmeTree v1.1 [83] was used as an independent topological method. To investigate the impact of missing taxa, 560 tips were assigned to known groupings as soft polytomies. Whole-tree and single-node statistics were calculated using 100000 Bayesian simulations.

### Evolutionary trait reconstruction

For reconstructions of morphology, data were imported into BayesTraits v1.0 [84]. Ancestral states were enforced using the ‘fossil’ prior. Harmonic means were compared for fossilized states, and accepted or rejected using BF values. 2 million MCMC iterations were conducted using the final consensus branching pattern and repeated if harmonic means did not stabilize. A reverse-jump hyperprior with exponential distribution 0–30 was set. ‘ratedev’ was optimized so that proposals were accepted 20–40% of the time. For host-use, major infraordinal divisions were treated as states. Aculeata (ants) were placed as the primary (ancestral) host for Myrmecolacidae: males of Myrmecolacidae parasitize only ants [85] and evidence for a female myrmecolacid in a fossilized ant host [86] indicated this was appropriate. A maximum likelihood approach, using the symmetrical method (Mesquite v2.71 [87]) was implemented to offer an independent measure of support.

### Comparison of molecular and phenotypic rates of evolution

State changes corresponding to non-homoplasious steps from a morphological phylogeny of Strepsiptera [34] were mapped to shared nodes of the MIT1+2 MCC tree (excluding branches leading to missing taxa, conflicting nodes, and clades represented by one taxon – representing 6, 6 and 9 steps respectively). A smaller list of discrete adult and secondary larval characters was used in a Bayesian reconstruction approach to ensure this pattern was repeatable across methods.

As a formal comparison of the relationship between molecular and phenotypic evolution, morphological branch lengths were estimated in Phylip [88], using the genus-level matrix of adult and secondary larval characters from [34], updated to include the current set of taxa. The topology was constrained to follow the MIT1+2 dated MCC tree. For individual nodes, we tested the correlation between % pairwise sequence divergence and

morphological branch length using standard statistics in R v2.9.2. As an independent whole-tree method, the K-score was calculated using Ktreedist v1.0 [89] and compared against a null K-distribution (500 simulated trees; following [18]). In this approach, molecular branch length (instead of pairwise sequence divergence) was assessed against morphological branch length.

### Supporting Information

**Figure S1 rRNA variable and core domain structural attributes mapped onto the Strepsiptera phylogeny.** (A) *18S* variable (bar) and core (filled circle) A+T% content. (B) *16S* variable (bar) and core (filled circle) A+T% content. (C) Variable domain size (nucleotide length) for the *18S* (red) and *16S* (black) genes. Outgroups grey and highlighted. Clade abbreviations and colour scheme follow Figure 1. Note the shifts in variable domain bp length, in both the *18S* (length increase) and *16S* (length decrease) genes at the node leading to Strepsiptera in (C). (TIF)

**Figure S2 Divergence time and molecular rate patterns using the nuclear *18S* rRNA dataset.** Red: % molecular rate mapped for each node at corresponding distances from root. Blue:  $L_n$  number of cumulative lineages at corresponding distances from root. This corroborates the analysis using the mitochondrial partition (Figure 2), confirming that the observed pattern is consistent across genomic compartments. (TIF)

**Figure S3 Exploration of data quality across the mitochondrial genes.** Transitions and transversions estimated under the F84 model were plotted against genetic distance for each codon position: Green = 1sts, Blue = 2nds, Orange = 3rds. Signal versus noise was graphically visualized using quartet likelihood mapping. (TIF)

**Table S1 Summary of Strepsiptera divergence times.** Summary of divergence time estimates for the major nodes in the Strepsiptera phylogeny using the combined mitochondrial coding gene (*cox1+nad1*) partition. \*Pairwise sequence divergences per million years. Clade abbreviations follow figure 1. †Node ages defined by exponential priors. (DOC)

**Table S2 RRTTest comparative analysis across strepsipteran clades.** Bold = P-value with significant rate comparison (bonferroni corrected). \*Marginally non-significant after bonferroni adjustment in the mitochondrial (A) and *18S* rRNA partition (B). Clade abbreviations follow figure 1. (DOC)

**Table S3 List of characters and corresponding states recovered in the reconstruction of strepsipteran morphological traits.** The position in the phylogeny of significant character reconstructions appears in brackets next to the corresponding state, followed by the BF range supporting that reconstruction. Some characters may be considered dependent, if single genotypic events can be demonstrated to produce pleiotropic effects. Possible examples include male/female larval spiracles, and male/female larval legs. \* equivocal BFs (0.2–3.8). This might disguise a potentially apomorphic loss of tarsomeres in the Elenchidae+Halictophagidae ancestor. (DOC)

**Table S4 Specimen, primer information and rRNA template alignments.** Genbank accession and specimen source information; list of primers used in this study (\*Primers designed

for short-fragment PCR) and 18 rRNA and 16S rRNA template secondary structural alignments.  
(DOC)

**Table S5 Test of saturation by mitochondrial gene and codon position.** \*Statistics indicating little saturation. †Statistics with substantial saturation (bold). ‡Statistics indicating useless/very poor sequence for phylogenetics (bold). Ts = symmetrical T-statistic. Tns = non-symmetrical T-statistic.  
(DOC)

**Text S1**  
(DOC)

## References

- O'Conner TD, Mundy NI (2009) Genotype-phenotype associations: substitution models to detect evolutionary associations between phenotypic variables and genotypic evolutionary rate. *Bioinformatics* 25: i94–i100.
- Lanfear R, Welch JJ, Bromham L (2010) Watching the clock: studying variation in rates of molecular evolution between species. *Trends Ecol Evol* 25: 495–503.
- Kathirithamby J (1989) Review of the order Strepsiptera. *Syst Entomol* 14: 41–62.
- Kathirithamby J (2009) Host-parasitoid associations in Strepsiptera. *Annu Rev Entomol* 54: 227–249.
- Kathirithamby J, Ross LD, Johnston JS (2003) Masquerading as self? endoparasitic Strepsiptera (Insecta) enclose themselves in host-derived epithelial bag. *P Natl Acad Sci USA* 100: 7655–7659.
- Gillespie JJ, McKenna CH, Yoder MJ, Gutell RR, Johnston JS, 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 Mol Biol* 14: 625–643.
- Pohl H, Beutel RG (2007) The evolution of Strepsiptera. *Zoology* 111: 318–338.
- McMahon DP, Hayward A, Kathirithamby J (2009) The mitochondrial genome of the 'twisted-wing parasite' *Mengenilla australiensis*: a comparative study. *BMC Genomics* 10: 603.
- Hunter MS, Woolley JB (2001) Evolution and behavioural ecology of heteronomous aphelinid parasitoids. *Annu Rev Entomol* 46: 251–90.
- Hayward A, McMahon DP, Kathirithamby J (2011) Cryptic diversity and host specificity in a parasitoid where the sexes utilize hosts from separate orders. *Mol Ecol* 20: 1508–1528.
- Eggleton P, Belshaw R (1992) Insect parasitoids: an evolutionary overview. *Philos Trans R Soc London Ser B* 337: 1–20.
- Kinzelbach RK (1978) Strepsiptera. *Die Tierwelt Deutschlands* 65: 166.
- Poulin R, Morand S (2004) Parasite biodiversity. Washington DC: Smithsonian Institution Press.
- Poulin R, Keeney DB (2008) Host specificity under molecular and experimental scrutiny. *Trends Parasitol* 24: 24–28.
- Dorus S, Evans PD, Wyckoff GJ, Choi SS, Lahn BT (2004) Rate of molecular evolution of the seminal protein gene SEMG2 correlates with levels of female promiscuity. *Nat Genet* 36: 1326–1329.
- Nadeau NJ, Burke T, Mundy NI (2007) Evolution of an avian pigmentation gene correlates with a measure of sexual selection. *Proc R Soc B Biol Sci* 255: 37–45.
- Ramm SA, Oliver PL, Ponting CP, Stockley P, Emes RD (2008) Sexual selection and the adaptive evolution of mammalian ejaculate proteins. *Mol Biol Evol* 25: 207–219.
- Ahrens D, Ribera I (2009) Inferring speciation modes in a clade of Iberian chafers from rates of morphological evolution in different character systems. *BMC Evol Biol* 9: 234.
- Kim H, Lee W, Lee S (2010) Morphometric relationship, phylogenetic correlation, and character evolution in the species-rich genus *Aphis* (Hemiptera: Aphididae). *PLoS ONE* 5: e11608.
- Hardman M, Hardman LM (2008) The relative importance of body size and paleoclimatic change as explanatory variables influencing lineage diversification rate: an evolutionary analysis of bullhead catfishes (Siluriformes: Ictaluridae). *Syst Biol* 57: 116–130.
- Omland KE (1997) Correlated rates of molecular and morphological evolution. *Evolution* 51: 1381–1393.
- Bromham L, Woolfit M, Lee MSY, Rambaut A (2002) Testing the relationship between morphological and molecular rates of change along phylogenies. *Evolution* 56: 1921–1930.
- Ekman S, Andersen HL, Wedin M (2007) The limitations of ancestral state reconstruction and the evolution of the Ascus in the Lecanorales (lichenized Ascomycota). *Syst Biol* 57: 141–156.
- Whiting MF, Carpenter JC, Wheeler QD, Wheeler WC (1997) The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Syst Biol* 46: 1–68.
- Huelsbeck JP (1998) Systematic bias in phylogenetic analysis: is the Strepsiptera problem solved? *Syst Biol* 47: 519–537.
- Huelsbeck JP (2001) A Bayesian perspective of the Strepsiptera problem. *Tidsskr Ent* 144: 165–178.
- Wiegmann BM, Trautwein MD, Kim JW, Cassel BK, Bertone MA, et al. (2009) Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC Biology* 7: 34.
- Longhorn SJ, Pohl H, Vogler AP (2010) Ribosomal protein genes of holometabolous insects reject the Halteria, instead revealing a close affinity of Strepsiptera with Coleoptera. *Mol Phylogenet Evol* 55: 846–859.
- McKenna DD, Farrell BD (2010) 9-genes reinforce the phylogeny of Holometabola and yield alternative views on the phylogenetic placement of Strepsiptera. *PLoS ONE* 5: e11887.
- Ishiwata K, Sasaki G, Owaga J, Miyata T, Su ZH (2010) Phylogenetic relationships among insect orders based on three nuclear protein-coding gene sequences. *Mol Phylogenet Evol* (doi:10.1016/j.ympev.2010.11.001).
- Bromham L (2009) Why do species vary in their rate of molecular evolution? *Biol Lett* 5: 401–404.
- Downton M, Austin AD (1995) Increased genetic diversity in mitochondrial genes is correlated with the evolution of parasitism in the Hymenoptera. *J Mol Evol* 41: 958–965.
- Duff RJ, Nickrent DL (1997) Characterization of mitochondrial small-subunit ribosomal RNAs from holoparasitic plants. *J Mol Evol* 45: 631–639.
- Pohl H, Beutel RG (2005) The phylogeny of Strepsiptera (Hexapoda). *Cladistics* 21: 328–374.
- Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *P Natl Acad Sci USA* 91: 6491–6495.
- Farrell BD (2001) Evolutionary assembly of the milkweed fauna: cytochrome oxidase I and the age of *Tetraopes* beetles. *Mol Phylogenet Evol* 18: 46–478.
- Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, et al. (2007) The delayed rise of present-day mammals. *Nature* 446: 507–512.
- Ricklefs RE, Losos JB, Townsend TM (2007) Evolutionary diversification of clades of squamate reptiles. *J Evol Biol* 20: 1751–1762.
- Vestri F (1943) Studi sugli 'Strepsiptera' Insecta. III. Descrizione e biologia di 6 specie italiane di *Mengenilla*. *Boll Lab Zool Gen Agric Portici* 32: 197–282.
- Pohl H, Beutel RG (2004) Fine structure of adhesive devices of Strepsiptera (Insecta). *Arthropod Struct Dev* 33: 31–43.
- Moore BR, Donoghue MJ (2007) Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovation. *Am Nat* 170: S28–S55.
- Venditti C, Meade A, Pagel M (2009) Phylogenies reveal new interpretation of speciation and the red queen. *Nature* 463: 349–352.
- Welch JJ, Bininda-Emonds ORP, Bromham L (2008) Correlates of substitution rate variation in mammalian protein-coding sequences. *BMC Evol Biol* 8: 53.
- Schmid-Hempel P (2008) Parasite immune evasion: a momentous molecular war. *Trends Ecol Evol* 23: 318–26.
- Welch JJ, Bromham L (2005) Molecular dating when rates vary. *Trends Ecol Evol* 20: 320–327.
- Ho SYW (2009) An examination of phylogenetic models of substitution rate variation among lineages. *Biol Lett* 5: 421–424.
- Crowson RA (1960) The phylogeny of Coleoptera. *Ann Rev Entomol* 5: 111–134.
- Lartillot N, Philippe H (2004) A Bayesian mixture model for across-site heterogeneities in the amino acid replacement process. *Mol Biol Evol* 21: 1095–1099.
- Bravo F, Pohl H, Silvo-Neto A, Beutel RG (2009) Bahiixenidae, a "living fossil" and a new family of Strepsiptera (Hexapoda) discovered in Brazil. *Cladistics* 25: 1–10.
- Castro LR, Downton M (2005) The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of *Perga condei* (Hymenoptera: Symphyta: Pergidae). *Mol Phylogenet Evol* 34: 469–479.
- Castro LR, Downton M (2006) Mitochondrial genomes in the Hymenoptera and their utility as phylogenetic markers. *Syst Entomol* 32: 60–69.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41: 95–98.

## Acknowledgments

We thank Peter W. H. Holland for use of laboratory facilities; St Hugh's College, Oxford; the National Museum of Natural History; the Smithsonian Institution (Washington DC) and the Smithsonian Tropical Research Institute (Panama) for support to JK. We thank three anonymous reviewers and Jordi Paps for helpful suggestions. We are indebted to many collectors for providing invaluable material (listed in Text S1).

## Author Contributions

Conceived and designed the experiments: DPM. Performed the experiments: DPM. Analyzed the data: DPM AH JK. Contributed reagents/materials/analysis tools: DPM AH JK. Wrote the paper: DPM AH JK.

53. Kjer KM, Baldrige GD, Fallon AM (1994) Mosquito large subunit ribosomal RNA: simultaneous alignment of primary and secondary structure. *Biochim Biophys Acta* 1217: 147–155.
54. Gillespie JJ (2004) Characterizing regions of ambiguous alignment caused by the expansion and contraction of hairpin-stem loops in ribosomal RNA molecules. *Mol Phylogenet Evol* 33: 936–943.
55. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31: 3406–3415.
56. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792–1797.
57. Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent ambiguously aligned blocks from protein sequence alignments. *Syst Biol* 56: 564–577.
58. Huelsenbeck J, Ronquist F (2001) MrBayes: Bayesian inference of phylogeny. *Bioinformatics* 17: 754–755.
59. Ronquist F, Huelsenbeck JP (2003) MrBayes 3, Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
60. Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
61. Drummond A, Strimmer K (2001) PAL: An object-oriented programming library for molecular evolution and phylogenetics. *Bioinformatics* 17: 662–663.
62. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696–704.
63. Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21: 2104–2105.
64. Stomatakis A (2006) RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.
65. Xia X, Xie Z (2001) DAMBE: data analysis in molecular biology and evolution. *J Heredity* 92: 371–373.
66. Xia X, Xie Z, Salemi M, Chen L, Wang Y (2003) An index of substitution saturation and its application. *Mol Phylogenet Evol* 26: 1–7.
67. Strimmer K, von Haeseler A (1997) Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proc Natl Acad Sci USA* 94: 6815–6819.
68. Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18: 502–504.
69. Adachi J, Hasegawa M (1996) Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J Mol Evol* 42: 459–468.
70. Abascal F, Posada D, Zardoya R (2007) MtArt: A new model of amino acid replacement for Arthropoda. *Mol Biol Evol* 24: 1–5.
71. Le SQ, Gascuel O (2008) LG: an improved, general amino-acid replacement matrix. *Mol Biol Evol* 25: 1307–20.
72. Rambaut A, Drummond (2007) AJ Tracer v1.4, Available from <http://beast.bio.ed.ac.uk/Tracer>.
73. Stomatakis A (2006) RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.
74. Rambaut A, Drummond AJ (2007) FigTree v1.0, Available from <http://tree.bio.ed.ac.uk/FigTree>.
75. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7: 214.
76. Yang Z (2007) PAML 4, a program package for phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24: 1586–1591.
77. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4: e88.
78. Shapiro B, Rambaut A, Drummond AJ (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol Biol Evol* 23: 7–9.
79. Robinson-Rechavi M, Huchon D (2000) RRTree: Relative-rate tests between groups of sequences on a phylogenetic tree. *Bioinformatics* 16: 296–297.
80. Nee S, Barraclough TG, Harvey PH (1996) Temporal changes in biodiversity: detecting patterns and identifying causes. In: Gaston KJ, ed. *Biodiversity: a biology of numbers and differences*. Oxford: Blackwell Science, pp 230–252.
81. Rambaut A, Harvey PH, Nee S (1997) End-Epi: an application for inferring phylogenies and population dynamic processes from molecular sequences. *Comput Appl Biosci* 13: 303–306.
82. Rabosky DL (2006) Likelihood methods for inferring temporal shifts in diversification rates. *Evolution* 60: 1152–1164.
83. Chan KM, Moore BR (2005) SYMMETREE: whole-tree analysis of differential diversification rates. *Bioinformatics* 21: 1709–1710.
84. Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* 53: 673–684.
85. Kathirithamby J, Hayward A, McMahon DP, Ferreira RS, Andreazze R, et al. (2010) Conspecifics of a heterotrophic heteronomous species of Strepsiptera (Insecta) are matched by molecular characterization. *Syst Entomol* 35: 234–242.
86. Pohl H, Kinzelbach KJ (2001) First record of a female stylopod (Strepsiptera: Myrecolacidae) parasite of prionomyrmecine ant (Hymenoptera: Formicidae) in Baltic amber. *Insect Syst Evol* 32: 143–146.
87. Maddison WP, Maddison DR (2006) Mesquite, a modular system for evolutionary analysis, Version 1.12. Available at: <http://mesquiteproject.org>.
88. Felsenstein J (1989) PHYLIP-phylogeny inference package (version 3.2). *Cladistics* 5: 164–166.
89. Soria-Carrasco V, Talavera G, Idea J, Castresana J (2007) The K tree score: quantification of differences in the relative branch length and topology of phylogenetic trees. *Bioinformatics* 23: 2954–2956.