Revisiting the Insect Mitochondrial Molecular Clock: The Mid-Aegean Trench Calibration

Anna Papadopoulou,*,1,2 Ioannis Anastasiou,3 and Alfried P. Vogler1,2

1Department of Entomology, Natural History Museum, London, United Kingdom
2Division of Biology, Imperial College London, Silwood Park Campus, Ascot, United Kingdom
3Department of Biology, University of Athens, Panepistimioupolis, Athens, Greece

*Corresponding author: E-mail: a.papadopoulou@nhm.ac.uk.
Associate editor: Barbara Holland

Abstract

Phylogenetic trees in insects are frequently dated by applying a “standard” mitochondrial DNA (mtDNA) clock estimated at 2.3% My–1, but despite its wide use reliable calibration points have been lacking. Here, we used a well-established biogeographic barrier, the mid-Aegean trench separating the western and eastern Aegean archipelago, to estimate substitution rates in tenebrionid beetles. Cytochrome oxidase I (cox1) for six codistributed genera across 28 islands (444 individuals) on both sides of the mid-Aegean trench revealed 60 independently coalescing entities delimited with a mixed Yule-coalescent model. One representative per entity was used for phylogenetic analysis of mitochondrial (cox1, 16S rRNA) and nuclear (Mbp20, 28S rRNA) genes. Six nodes marked geographically congruent east–west splits whose separation was largely contemporaneous and likely to reflect the formation of the mid-Aegean trench at 9–12 Mya. Based on these “known” dates, a divergence rate of 3.54% My–1 for the cox1 gene (2.69% when combined with the 16S rRNA gene) was obtained under the preferred partitioning scheme and substitution model selected using Bayes factors. An extensive survey suggests that discrepancies in mtDNA substitution rates in the entomological literature can be attributed to the use of different substitution models, the use of different mitochondrial gene regions, mixing of intraspecific with interspecific data, and not accounting for variance in coalescent times or postseparation gene flow. Different treatments of these factors in the literature confound estimates of mtDNA substitution rates in opposing directions and obscure lineage-specific differences in rates when comparing data from various sources.

Key words: molecular clock calibration, Bayesian relaxed clock, insects, mitochondrial DNA rates, Tenebrionidae, mid-Aegean trench.

Introduction

Evolutionary scenarios frequently rely on estimates of node ages in time-calibrated phylogenetic trees. Where fossils are unavailable, as is the case for many groups of insects, a standardized rate of molecular change may be applied to obtain such estimates. For insect mitochondrial DNA (mtDNA), the most widely quoted rate of molecular evolution is Brower’s (1994) calibration based on a set of seven studies that provided age estimates of lineage splits ranging from 300 to 3,250,000 years ago. Regression on uncorrected pairwise distances against inferred number of years since lineage divergence revealed a surprisingly strong linear relationship with \( y = 2.34 \times 10^{-6}x \) and \( r^2 = 0.996 \), for a substitution rate of 0.0115 per site per My equal to 2.3% divergence. Although Brower’s (1994) rate appears to be satisfactory to estimate ages in many groups, its uniformity in the seven original studies is surprising given the mixed use of protein-coding and ribosomal markers, whose rates are expected to differ greatly. In addition, mutation rates in recently diverged haplotypes at or near the population level have now been found to be much higher than phylogenetic rates (Ho et al. 2005, 2007), possibly due to the delayed effect of purifying selection. This “time dependency” of the clock may affect a period of up to \( \sim 1 \) Ma, well within the time window from which most of Brower’s (1994) calibration points were drawn.

Subsequent estimates of insect mtDNA substitution rates based on biogeographic vicariance, island ages, fossils, or other independent evidence (supplementary table S1, Supplementary Material online), reported both decreased (Sperling et al. 1997; Pruser and Mossakowski 1998; Andersen et al. 2000) or elevated (Fleischer et al. 1998; Luchetti et al. 2005; Shapiro et al. 2006) rates, which is generally attributed to lineage-specific or gene-specific effects (e.g., a lower rate of 1.5% divergence in cytochrome oxidase I [cox1]; Farrell 2001). However, the effects of the methodology used for rate or age estimation have not been fully appreciated. Recently developed “relaxed-clock” methods, which allow substitution rates to vary among branches either in an autocorrelated or uncorrelated manner (Sanderson 2002, 2003; Thorne and Kishino 2002; Yang 2004; Drummond et al. 2006), hold great promise for a more accurate calibration of the mtDNA clock and its rate variation among insect lineages. Yet, accurate estimations of substitution rates require careful choice of the model of sequence evolution used to correct for multiple hits (Yang 1996; Arbogast et al. 2002), an issue that has received little attention in the entomological literature for its effect on
clock estimates. Moreover, the preponderance of biogeographic and paleoecological data used to calibrate the clock requires well-documented geological or paleoclimatic evidence and a demonstration that the assumed barriers constitute a true obstacle to dispersal for the focal group. Finally, several factors may confound the correlation between geological time and lineage divergence, such as ancestral polymorphism (Edwards and Beerli 2000) or postseparation gene flow that delay lineage sorting along an assumed biogeographic boundary (Carstens and Knowles 2007).

Improved estimates of the mtDNA clock in insects therefore require further study of systems that are less affected by frequently encountered sources of error. The mid-Aegean trench in the eastern Mediterranean is highly suited for this kind of analysis. This geological formation originated in the Upper Miocene (12–9 Mya) and led to the initial split of the united landmass of Agaïs (Creutzburg 1963; Dermitzakis and Papanikolaou 1981), separating the western and eastern Aegean archipelago (fig. 1). These areas have remained subdivided by the surrounding ocean ever since, except for the Messinian desiccation of the Mediterranean basin from 5.96 to 5.33 Mya (Krijgsman et al. 1999) when a land bridge may have existed via a deep rift valley.

The Aegean islands harbor species-rich assemblages of darkling beetles (Coleoptera: Tenebrionidae) highly suited for biogeographic clock calibrations. Many species are flightless with limited dispersal capabilities, in particular, the “geophilic” lineages associated with ecologically stable soil types. Genetic variation in a few of these species has already been shown to be highly structured geographically and is deeply subdivided along the mid-Aegean trench (Papadopoulou, Anastasiou, et al. 2009). Codistributed lineages of tenebrionids therefore can be used to calibrate rates simultaneously in multiple taxon pairs that were subjected to a single geological event, thus reducing stochastic effects of lineage sorting and variation in substitution rates (Edwards and Beerli 2000; Hickerson et al. 2003). A comprehensive survey of genetic variation of six flightless, geophilic tenebrionid genera with wide distribution throughout the Aegean established several temporally congruent lineage splits that were attributable to the formation of the mid-Aegean trench. This provides an independent calibration of the mtDNA clock in a major lineage of insects and permits an assessment of the sensitivity of inferred substitution rates to model selection, alternative partitioning schemes, the use of relaxed versus strict clock methods, and gene-specific differences in rates between protein-coding and rRNA genes. Comparisons with literature data show that discrepancies in rates found by different studies can be attributed largely to methodological issues, although substitutions rates in the Aegean tenebrionids may still be higher than most existing estimates for insects.

**Material and Methods**

**Taxon Sampling and DNA Sequencing**

A total of 444 specimens from six flightless, geophilic tenebrionid genera were sampled from 11 to 25 islands each, and 51–122 individuals were sequenced per genus (table 1, supplementary table S2, Supplementary Material online). Total genomic DNA was extracted from thorax or leg tissue using the Promega 96-well plate kit. An 826–829 bp fragment of the third end of the cox1 gene was amplified using primers C1-J-2183 (Jerry) and TL2-N-3014 (Pat) (Simon et al. 1994) or JerryTen and PatTen (Papadopoulou, Anastasiou, et al. 2009). A 513–524 bp fragment of the single-copy nuclear mucosal protein 20 (Mp20) locus, including 469 bp of coding region and one intron, was amplified using the primer pair Mp205’ and Mp203’ (Pons et al. 2004) for Pilinellinae or Mp20Trib5’ and Mp20Trib3’ for Tenebrioninae (Papadopoulou, Anastasiou, et al. 2009). A 433–437 bp portion of the mitochondrial 16S rRNA gene (rrnL) was amplified using LR-N-13398 (16Sar) (Simon et al. 1994) and LR-J-12961 (16Sb2) (Cognato and Vogler 2001), and a 646–655 bp fragment of the nuclear 28S rRNA gene was amplified using 28SFF and 28SDD (Monaghan et al. 2007). Amplification products were purified using Millipore Multiscreen 96-well plates (Millipore, Billerica, MA) and sequence in both directions using the BigDye technology and an ABI PRISM 3700 DNA Analyzer (Applied Biosystems). Sequence chromatograms were assembled and edited using the Sequencher 4.6 software (Gene Codes Corporation, Ann Harbor, MI). cox1 and Mp20 sequences for the genera Dailognatha, Eutagenia, and Zophosis are from Papadopoulou, Anastasiou, et al. (2009). New sequences have been submitted to the EMBL Nucleotide Sequence Database (supplementary tables S2–S3, Supplementary Material online).

**Species Delimitation, Alignment, and Phylogenetic Analysis**

To minimize the effect of increased mutation rates at the intraspecific level (Ho et al. 2005, 2007), sequence variation was divided into within- and between-species groups using the generalized mixed Yule-coalescent (GMYC) model (Pons et al. 2006; Fontaneto et al. 2007). This procedure identifies a threshold value for the shift in branching rate from coalescent lineage branching to interspecific diversification on an ultrametric tree and delimits “independently evolving” mtDNA clusters. The analysis was carried out using the R package SPLIT (Species Limits by Threshold Statistics) available at http://r-forge.r-project.org/projects/splits/ with the “single-threshold” option. A clock-constrained tree required for the analysis was built separately for the full cox1 data sets of each genus, after removal of identical haplotypes. Tree searches were performed with Bayesian analysis in MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003), applying separate models for two partitions (1st and 2nd codon positions together vs. 3rd codon position) as selected by Akaïke information criterion (AIC) in MrModeltest 2.2 (Nylander 2004), with two parallel runs of 5 million generations each and using one cold and two incrementally heated Markov chains (λ = 0.1) and sampling every 1,000 steps. Trace plots were visually inspected, and convergence diagnostics (standard deviation [SD] of split frequencies, effective sample size), as implemented in
MrBayes and Tracer 1.4.1 (Drummond and Rambaut 2007), were checked to ensure that the Markov chain had reached stationarity. After discarding the first 2.5 million generations as burn-in, trees were summarized using an “all-compatible” consensus. Each consensus tree was converted to ultrametric using penalized likelihood as implemented in r8s 1.7 (Sanderson 2003) with the optimal smoothing parameter selected by cross-validation of values between 0.01 and 1,000 following the procedure described in the r8s manual.

For the phylogenetic analysis at the species level, a single exemplar representing each GMYC group in the cox1 analysis was selected for sequencing of three additional markers given above. A range of different alignment strategies was assessed for the length-variable *rrnL*, 28S, and the intron region of Mp20. We compared five gap penalty combinations (gap opening penalty: 5-6.66-15-20 vs. gap extension penalty: 6.66) in ClustalW (Thompson et al. 1994) and three different iterative refinement methods (E-INS-i, G-INS-i, and L-INS-i) in MAFFT 6.240 (Katoh et al. 2005; Katoh and Toh 2008). Each resulting alignment was assessed for congruence with the unambiguously aligned protein-coding regions of cox1 and Mp20 using the

![Figure 1](image_url)

**Figure 1.** (a) Relatively dated four-gene tree (combined data set of *cox1*, *rrnL*, 28S, and Mp20) of six tenebrionid genera generated using an uncorrelated lognormal clock in BEAST. Eight east–west nodes (EW1–EW8) are compared for temporal congruence (using a relative scale 0–100). Node heights correspond to mean values across 9,000 post burn-in trees, whereas gray bars indicate ± 1 SD for the eight focal nodes (see table 2 for values). (b) Map of the central Aegean region showing the sampled islands and mainland regions on the east and on the west of the mid-Aegean trench, dashed lines represent the presumed position of the trench.
Table 1. Sampled Taxa and Results of the GMYC Model for Species Delimitation.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Tribe</th>
<th>spa</th>
<th>islb</th>
<th>Sequences</th>
<th>Entities</th>
<th>Clusters</th>
<th>logLGMYC</th>
<th>logLnull</th>
<th>T (E-W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D Dialognatha</td>
<td>Tentyriini</td>
<td>2</td>
<td>25</td>
<td>122</td>
<td>23</td>
<td>18-20</td>
<td>160.9728</td>
<td>174.6146</td>
<td>22 (9-13)</td>
</tr>
<tr>
<td></td>
<td>Dendarus</td>
<td>7</td>
<td>12</td>
<td>51</td>
<td>10</td>
<td>9-13</td>
<td>5.6441</td>
<td>11.1804*</td>
<td>9 (3-6)</td>
</tr>
<tr>
<td></td>
<td>Eutenia</td>
<td>1</td>
<td>18</td>
<td>63</td>
<td>12</td>
<td>10-12</td>
<td>28.4019</td>
<td>34.4749**</td>
<td>12 (2-10)</td>
</tr>
<tr>
<td></td>
<td>Pimelia</td>
<td>1</td>
<td>11</td>
<td>56</td>
<td>12</td>
<td>11-12</td>
<td>13.1996</td>
<td>18.2333*</td>
<td>10 (5-5)</td>
</tr>
<tr>
<td></td>
<td>Tentyriini</td>
<td>1</td>
<td>18</td>
<td>59</td>
<td>2</td>
<td>2-4</td>
<td>8.4808</td>
<td>18.3449*</td>
<td>3 (1-2)</td>
</tr>
<tr>
<td></td>
<td>Zophosis</td>
<td>1</td>
<td>25</td>
<td>93</td>
<td>8</td>
<td>3-16</td>
<td>102.1494</td>
<td>114.2980***</td>
<td>4 (2-2)</td>
</tr>
</tbody>
</table>

a Number of morphologically described species.
b Number of islands or mainland regions sampled per taxon.
c Number of cox1 sequences used to apply the GMYC model.
d Total number of independent entities identified by the GMYC model including singletons (range of entities within 2 logL of the model).
e Number of entities with more than one individual.
f The likelihood of the null model.
g The likelihood of the GMYC model, likelihood ratio test ***P < 0.001, **P < 0.01, *P < 0.05.
h Number of terminals per genus in the six-genera phylogenetic tree sampled from the east or the west of the mid-Aegean trench.
i Genera that were analyzed in Papadopoulou, Anastasiou, et al. (2009).

Incongruence Length Difference (ILD) test (Farris et al. 1994). Tree lengths were calculated using parsimony searches in PAUP* 4.0b10 (Swofford 2002) with 1,000 random addition sequence replicates and gaps treated as “5th state.” We separately selected the alignment strategy for each locus that gave the lowest score for the ratio ILD/length of the combined-analysis tree (Wheeler and Hayashi 1998) with respect to the protein-coding regions and then tested the overall congruence of the resulting data set (partitioned as cox1—rrnL—28S—Mp20 exon—Mp20 intron) using the partition homogeneity test in PAUP* with 100 replicates.

Parsimony analysis of the combined data set was performed in PAUP* with 1,000 random addition sequence replicates and gaps treated either as 5th state or as “missing data,” and nonparametric bootstrap was conducted with 1,000 pseudoreplicates. We used PhyML 3.0 (Guindon and Gascuel 2003) to perform unpartitioned maximum likelihood (ML) analysis under a general time reversible (GTR)+Γ+I substitution model and calculate bootstrap support values with 100 replicates. We also carried out partitioned ML searches with RAxML 7.0.4 (Stamatakis 2006) under seven different partitioning schemes (supplementary table S4, Supplementary Material online). The GTR MIX model was employed, so the initial tree searches were conducted with the GTRCAT approximation but the final tree topology was evaluated under a separate GTR+Γ+I model for each partition. In total, 1,000 bootstrap replicates were performed for each partitioning scheme using the rapid RAxML bootstrapping algorithm (Stamatakis et al. 2008).

We tested the null hypothesis that all pairs of east/west clades are reciprocally monophyletic and sister to each other, comparing parsimony and RAxML searches under that topological constraint with those of unconstrained searches and assessing the significance of the observed differences with the Shimodaira–Hasegawa test (Shimodaira and Hasegawa 1999) as implemented in PAUP*. We applied a single GTR+Γ+I model (as PAUP* does not permit applying a separate substitution model to each partition) with parameters estimated by PhyML, and the significance of the test was evaluated using RELL sampling with 10,000 replicates.

Relative Node Ages and Clock Calibrations

Node ages and substitution rates were estimated using an uncorrelated lognormal relaxed clock in BEAST 1.4.8 (Drummond et al. 2006; Drummond and Rambaut 2007). In all analyses, the among-genera relationships and the monophyly of the eastern and western clades were constrained according to the results of the topology tests conducted on the four-gene data set. This was considered necessary because otherwise, if a small proportion of the sampled trees from the posterior distribution did not include all eastern and western clades as reciprocally monophyletic, the 9–12 Mya calibration would have been partially assigned to incorrect nodes. Relative node ages were estimated by fixing the root node to an arbitrary value (normal prior distribution with a mean of 100 and SD of 1). Two independent runs of 50 million generations (sampling every 5,000th generation) were performed for each analysis, using a Yule tree prior and the default options for all other prior and operator settings. The convergence and mixing of each Markov chain Monte Carlo chain was assessed by inspection of the trace plots and the effective sample sizes using Tracer 1.4.1 (Drummond and Rambaut 2007). Samples from both independent runs were then pooled after removing a 10% burn-in using LogCombiner 1.4.8. The means and standard errors (SEs) of the node heights were summarized using Tracer, and SDs were calculated by multiplying the SEs by the square root of the effective sample size in each case.

Estimates of substitution rates were performed under an uncorrelated lognormal relaxed clock in BEAST as described before, but instead of fixing the root node, a normal prior distribution was applied on the ages of the selected calibration points, with a mean of 10.5 My and SD = 1.5 (0.05 quantile: 8.033, 0.95 quantile: 12.97), to reflect the geological age of the Aegean trench at 9–12 Mya. We estimated rates for the combined cox1 and rrnl data set under an Hasegawa-Kishino-Yano (HKY) model, a GTR model, a GTR+Γ+I, and four partitioning schemes (P1: cox1 vs. rrnl, P2: 3rd codons vs. all other sites, P3: 1st and 2nd positions vs. 3rd positions vs. rrnl, P4: 1st vs. 2nd vs. 3rd vs. rrnl). We applied a separate substitution
model to each partition as selected by the AIC implemented in MrModeltest 2.2 (Nylander 2004) (supplementary table S5, Supplementary Material online). Moreover, we used the same priors on the node ages and models selected by the AIC to estimate rates for each of the five gene regions separately (cox1, rnl, Mp20 exon, Mp20 intron, 28S), with cox1 treated either as a single partition or divided by codon position.

A strict clock was applied for comparison using MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003) under a uniform prior on branch lengths. All searches were conducted with 5 million generations and two parallel runs using one cold and two incrementally heated Markov chains (\(\lambda = 0.1\)) and sampling every 1,000 steps, and the first 2.5 million generations were discarded as burn-in. Mean node heights and 95% higher posterior limits across all post burn-in trees were calculated using TreeAnnotator 1.4.8, and these numbers were higher than the posterior means.

Testing for Contemporaneous Divergence

The eight east/west pairs were tested for temporal congruence using a relative dating approach (Loader et al. 2007), by comparing their relative ages under a relaxed-clock model based on all four genes or only on the mtDNA data

Mean node heights and 95% higher posterior limits across all post burn-in trees were calculated using TreeAnnotator 1.4.8, and these numbers were higher than the posterior means.

Alignment and Phylogenetic Analysis

Alignment parameters for each length-variable gene region were selected to reduce incongruence with the protein-coding regions (supplementary table S6, Supplementary Material online). The resulting concatenated matrix was highly congruent among the five partitions (\(P = 0.98\); partition homogeneity test). Parsimony and ML analyses supported the monophyly of each genus, whereas the topology of defining the monophyly of the GMYC model. One individual per cluster and selected singletons were chosen for sequencing of the rnl, Mp20, and 28S markers. Confidence intervals for the number of clusters were calculated (table 1), but we selected exemplars to match the ML solution, except in the case of Tentyria, where we used the suboptimal solution of three GMYC clusters (A1, A2, and B in supplementary fig. S3, Supplementary Material online). After failure to sequence two of the chosen individuals, the final four-gene data set comprised 60 terminals, including 22 eastern and 38 western lineages (fig. 1 and supplementary table S3, Supplementary Material online).

MEGA4 (Tamura et al. 2007) was used to calculate average pairwise uncorrected and Kimura 2-parameter distances between each east/west pair, and distances were converted to maximum and minimum divergences per My, corresponding to 9 and 12 Mya, respectively.

Results

Phylogenetic Analysis of cox1 and mtDNA Cluster Delimitation

The cox1 Bayesian trees of Dendarus, Pimelia, and Tentyria (supplementary figs. S1–S3, Supplementary Material online) and Dialognatha, Eutagenia, and Zophosis (Papadopoulou, Anastasiou, et al. 2009) revealed strong phylogenetic clustering and geographical structure. The GMYC model had a significantly better fit to the data than the null model of uniform coalescent branching for all lineages (\(P < 0.05\)) and identified between 2 (Tentyria) and 23 (Dialognatha) GMYC entities (table 1 and supplementary figs. S1–S3, Supplementary Material online; Papadopoulou, Anastasiou, et al. 2009). The number of GMYC groups greatly exceeded that of Linnaean names, as most of these highly subdivided lineages are currently described as a single species (table 1). Each of the GMYC clusters was geographically restricted to a single island or a group of adjacent islands either on the eastern or western side of the mid-Aegean trench. A notable exception was the eastern GMYC cluster in the genus Tentyria found on the volcanic island of Santorini located to the west of the mid-Aegean trench (supplementary fig. S3, Supplementary Material online). This may be attributed to a recent recolonization event, after the last catastrophic eruption of the volcano 3,500 years ago, a pattern that has been suggested for many other taxa (Thornton 2007). In total, 67 entities (55 clusters and 12 singletons) were recognized at the point of the highest likelihood of the GMYC model. One individual per cluster and selected singletons were chosen for sequencing of the rnl, Mp20, and 28S markers. Confidence intervals for the number of clusters were calculated (table 1), but we selected exemplars to match the ML solution, except in the case of Tentyria, where we used the suboptimal solution of three GMYC clusters (A1, A2, and B in supplementary fig. S3, Supplementary Material online). After failure to sequence two of the chosen individuals, the final four-gene data set comprised 60 terminals, including 22 eastern and 38 western lineages (fig. 1 and supplementary table S3, Supplementary Material online).

Alignment and Phylogenetic Analysis

Alignment parameters for each length-variable gene region were selected to reduce incongruence with the protein-coding regions (supplementary table S6, Supplementary Material online). The resulting concatenated matrix was highly congruent among the five partitions (\(P = 0.98\); partition homogeneity test). Parsimony and ML analyses supported the monophyly of each genus, whereas the topology (Dendarus, ((Dialognatha, Tentyria), Zophosis), (Eutagenia, Pimelia)) was favored by all searches except for parsimony when gaps were treated as 5th state (supplementary table S4, Supplementary Material online). Within each genus, most of the eastern (E) and western (W) lineages were reciprocally monophyletic, with generally high bootstrap support (fig. 1; EW1–8), except for clades E6 and W6 (Dailognatha) that were unresolved with respect to each other (supplementary table S7, Supplementary Material online). Tree scores from searches constrained for monophyly of all eight east/west pairs and their reciprocal monophyly were very similar to those from unconstrained analyses (supplementary table S8, Supplementary Material online), and when scores were slightly lower these differences were not significant in the Shimodaira–Hasegawa test under parsimony (all \(P >> 0.05\)) nor ML (\(P > 0.05\); supplementary table S9, Supplementary Material online). These results provide justification to conduct mtDNA clock calibration on trees constrained for the monophyly of these nodes.
set. Mean values of node ages were within ± 1 SD of each other for six east/west pairs (table 2), supporting a hypothesis of contemporaneous divergence. The two remaining east/west pairs were either marginally (Dailognatha, EW6) or significantly (Zophosis, EW3) more recent (table 2). In light of the geological history of the region, the six contemporaneous nodes were attributed to the formation of the mid-Aegean trench at 9–12 Mya, whereas the EW3 node was assumed to reflect the younger east/west subdivision after the end of the Messinian desiccation at 5.96 Mya. The EW6 split may also be associated to the initial formation of the mid-Aegean trench but affected by postseparation dispersal (e.g., an eastern sample was found on Amorgos island immediately to the west of the trench), and the E6/W6 separation was poorly supported in the phylogenetic analysis. Therefore, only the six temporally congruent east/west nodes were employed as calibration points based on the age of the mid-Aegean trench.

**Calibration of Substitution Rates Based on the Mid-Aegean Trench Geological Age**

Substitution rates for the combined cox1 + rrnL data set on trees calibrated at 9–12 Mya for the six nodes revealed great differences depending on the substitution model used for correction. The estimated divergences per My ranged from 1% when using uncorrected distances and approximately 1.2% when using a HKY or a GTR model without accounting for rate heterogeneity among sites. When applying a GTR+\(\Gamma\)+I model, this divergence estimate increased to 2.23% (strict clock in MrBayes) or 2.39% (relaxed clock in BEAST) without partitioning and to 2.69% when using the P3 or P4 partitioning scheme in BEAST (fig. 2). The differences in estimated rates among partitioning schemes were consistent between MrBayes and BEAST analyses (fig. 2) but were in all cases higher by 0.1–0.2% with the latter. Bayes factor comparisons favored the P3 partitioning scheme in both MrBayes and BEAST analyses when applying the ln(Bayes Factor)/\(\Delta p\) ≥ 10 criterion or the P4 partitioning scheme applying the Kass and Raftery (1995) criterion of 2lnBF ≥ 10 (table 3), in either case supporting a rate of 2.7% (relaxed clock; BEAST) or 2.5% (strict clock; MrBayes). Applying a strict clock in BEAST under the preferred P3 partitioning scheme gave an estimated rate of 2.6% and Bayes factors comparisons between the strict and the uncorrelated lognormal relaxed clock favored the relaxed clock (lnBF = 18.32). Our estimates for cox1 on its own were higher than the average of the two mitochondrial genes (3.36% unpartitioned or 3.54% for the preferred partitioning scheme), whereas they were lower for rrnL (1.06%). For the nuclear genes, we estimated mean divergence rates of 3.68% My⁻¹ for the intron region of Mp20, 0.66% My⁻¹ for the exon, and 0.12% My⁻¹ for 28S (table 4).

The use of an uncorrelated relaxed-clock approach, as implemented in BEAST, also permits comparisons of the clocklikeness in different gene regions (Drummond et al. 2006) and measurement of the degree of rate autocorrelation among lineages (Ho 2009). Our results indicate that cox1 evolves in a more clock-like manner than rrnL (lower ucld.mean and coefficient of variation; table 3), which is in agreement with findings of other studies (Gaunt and Miles 2002). Both mitochondrial regions do not deviate greatly from the strict clock, whereas all nuclear regions (including the presumably neutrally evolving intron) showed much greater variation in rates among branches, as the SD of branch rates and the coefficient of variation are both greater than 1 (table 4). In terms of autocorrelation of rates between neighboring branches, we only found a significantly positive covariance in the case of the Mp20 exon (table 4).

**Applying the “Standard” Insect Mitochondrial Clock**

When the mean of the branch rates on the combined cox1 and rrnL data was fixed at 0.0115, that is, 2.3%...
Table 3. Bayes Factor Comparisons for Selection of Substitution Model and Partitioning Scheme. a) MrBayes Analyses Under a Strict Clock b) BEAST Analyses Under an Uncorrelated Lognormal Relaxed Clock.

<table>
<thead>
<tr>
<th></th>
<th>HME</th>
<th>HKY</th>
<th>GTR</th>
<th>GTR + Γ^-1</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) MrBayes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HKY (4)</td>
<td>-18486.14</td>
<td>-</td>
<td>68.31</td>
<td>504.49</td>
<td>198.40</td>
<td>217.10</td>
<td>138.83</td>
<td>104.09</td>
</tr>
<tr>
<td>GTR (8)</td>
<td>-18212.91</td>
<td>273.23</td>
<td>-</td>
<td>1376.85</td>
<td>238.43</td>
<td>262.88</td>
<td>150.58</td>
<td>108.30</td>
</tr>
<tr>
<td>GTR + Γ^-1 (10)</td>
<td>-15459.21</td>
<td>302.69</td>
<td>2753.70</td>
<td>-</td>
<td>31.44</td>
<td>60.34</td>
<td>39.11</td>
<td>29.01</td>
</tr>
<tr>
<td>P1 (21)</td>
<td>-15113.36</td>
<td>3372.78</td>
<td>3099.55</td>
<td>345.85</td>
<td>-</td>
<td>n/a</td>
<td>46.77</td>
<td>27.74</td>
</tr>
<tr>
<td>P2 (21)</td>
<td>-14795.46</td>
<td>3690.68</td>
<td>3417.45</td>
<td>663.75</td>
<td>317.90</td>
<td>-</td>
<td>17.87</td>
<td>12.61</td>
</tr>
<tr>
<td>P3 (32)</td>
<td>-14598.88</td>
<td>3887.26</td>
<td>3614.03</td>
<td>860.33</td>
<td>514.48</td>
<td>196.58</td>
<td>-</td>
<td>6.81</td>
</tr>
<tr>
<td>P4 (42)</td>
<td>-14530.75</td>
<td>3955.39</td>
<td>3682.16</td>
<td>928.46</td>
<td>582.61</td>
<td>264.71</td>
<td>68.13</td>
<td>-</td>
</tr>
<tr>
<td>b) BEAST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HKY (4)</td>
<td>-18356.50</td>
<td>-</td>
<td>66.49</td>
<td>499.19</td>
<td>192.18</td>
<td>210.09</td>
<td>134.67</td>
<td>101.25</td>
</tr>
<tr>
<td>GTR (8)</td>
<td>-18090.54</td>
<td>265.97</td>
<td>-</td>
<td>1364.60</td>
<td>230.85</td>
<td>254.28</td>
<td>146.04</td>
<td>105.34</td>
</tr>
<tr>
<td>GTR + Γ^-1 (10)</td>
<td>-15361.34</td>
<td>2995.16</td>
<td>2729.19</td>
<td>-</td>
<td>24.72</td>
<td>52.40</td>
<td>35.26</td>
<td>26.64</td>
</tr>
<tr>
<td>P1 (21)</td>
<td>-15089.46</td>
<td>3267.05</td>
<td>3001.08</td>
<td>271.89</td>
<td>-</td>
<td>n/a</td>
<td>45.80</td>
<td>27.65</td>
</tr>
<tr>
<td>P2 (21)</td>
<td>-14784.94</td>
<td>3571.56</td>
<td>3305.59</td>
<td>576.40</td>
<td>304.51</td>
<td>-</td>
<td>18.12</td>
<td>13.15</td>
</tr>
<tr>
<td>P3 (32)</td>
<td>-14585.67</td>
<td>3770.84</td>
<td>3504.87</td>
<td>775.68</td>
<td>503.79</td>
<td>199.28</td>
<td>-</td>
<td>7.68</td>
</tr>
<tr>
<td>P4 (42)</td>
<td>-14508.90</td>
<td>3847.61</td>
<td>3581.64</td>
<td>852.45</td>
<td>580.56</td>
<td>276.05</td>
<td>76.77</td>
<td>-</td>
</tr>
</tbody>
</table>

Note.—n/a, not applicable.

- The harmonic mean of sampled likelihoods as estimated by MrBayes or Tracer.
- Numbers in brackets: total number of free parameters required for each model or partitioning scheme.
- Above the diagonal: ln(Bayes Factor)/Δp (where Δp: difference in total number of free parameters between two models).
- Below the diagonal: ln(Bayes Factor).

 divergence My^-1 (Brower 1994), analyses under five partitioning schemes and using an uncorrelated lognormal relaxed clock in BEAST, resulted in estimated mean ages ± 1 SD for the six contemporaneous east/west nodes that were compatible with the age of the mid-Aegean trench (table 5). The node EW3 (genus Zophosis) was estimated to be much more recent (4.78–4.91 Mya), an age compatible with post-Messinian divergence, whereas EW6 ranged between 7.26 and 7.65 Mya. When the 2.3% standard clock was applied to the individual mitochondrial genes, ages of the six contemporaneous east/west nodes were much higher for cox1 (12.3–17.5 Mya) and much lower for rrnL (3–4.9 Mya) (table 5). The effect of data partitioning was investigated, both for the combined cox1 and rrnL data set (P0–P4) and cox1 on its own. Comparisons of the resulting ages suggested that a greater number of partitions resulted in slightly higher estimates (table 5). For example, compared with the unpartitioned data the P4 partitioning caused an increase in estimated node ages of 0.15–1.7 My (i.e., an increase by 2.8–14.4%), with the lowest percentage corresponding to the most recent node (EW3) and the highest percentage to the oldest node (EW7).

Comparing Existing Calibrations from the Literature
An extensive literature search starting from publications citing Brower (1994) found 30 other studies (supplementary table S1, Supplementary Material online) that estimated substitution rates for insect mtDNA based on biogeographic, paleoclimatic, fossil, or other independent evidence. These data were compiled for comparisons, separately for studies that did not account for rate heterogeneity among sites (uncorrected distances or using simple substitution models) and those that used models

Table 4. Estimated Rates Per Gene Region Based on Six East–West Calibration Points and Using a Lognormal Uncorrelated Relaxed Clock in BEAST, Mean Values ± 1SD.

<table>
<thead>
<tr>
<th>Gene Region</th>
<th>Mean Rate</th>
<th>ucl.mean</th>
<th>ucl.sd</th>
<th>Coefficient Variance</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>cox1</td>
<td>0.0168 ± 0.0018</td>
<td>0.0169 ± 0.0019</td>
<td>0.2571 ± 0.0674</td>
<td>0.2609 ± 0.0702</td>
<td>-0.0120 ± 0.0903</td>
</tr>
<tr>
<td>cox1 (2)</td>
<td>0.0177 ± 0.0019</td>
<td>0.0178 ± 0.0019</td>
<td>0.2973 ± 0.0644</td>
<td>0.3031 ± 0.0686</td>
<td>-0.0134 ± 0.0898</td>
</tr>
<tr>
<td>16S</td>
<td>0.0054 ± 0.0009</td>
<td>0.0049 ± 0.0008</td>
<td>0.5106 ± 0.1238</td>
<td>0.5418 ± 0.1468</td>
<td>0.0001 ± 0.0895</td>
</tr>
<tr>
<td>mtDNA</td>
<td>0.0120 ± 0.0012</td>
<td>0.0119 ± 0.0011</td>
<td>0.1863 ± 0.0619</td>
<td>0.1878 ± 0.0632</td>
<td>-0.0176 ± 0.0914</td>
</tr>
<tr>
<td>mtDNA (P3)</td>
<td>0.0133 ± 0.0013</td>
<td>0.0131 ± 0.0013</td>
<td>0.2602 ± 0.0554</td>
<td>0.264 ± 0.0578</td>
<td>-0.0141 ± 0.0913</td>
</tr>
<tr>
<td>Mpl20 intron</td>
<td>0.0184 ± 0.0152</td>
<td>0.0496 ± 0.2628</td>
<td>2.5717 ± 0.5399</td>
<td>4.4511 ± 1.1499</td>
<td>0.0781 ± 0.1136</td>
</tr>
<tr>
<td>Mpl20 exon</td>
<td>0.0033 ± 0.0006</td>
<td>0.0024 ± 0.0008</td>
<td>1.6353 ± 0.2375</td>
<td>2.3847 ± 0.5021</td>
<td>0.2118 ± 0.1454</td>
</tr>
<tr>
<td>28S</td>
<td>0.0006 ± 0.0003</td>
<td>0.8273 ± 6.2150</td>
<td>3.1057 ± 1.0567</td>
<td>5.4504 ± 1.8708</td>
<td>0.0140 ± 0.0538</td>
</tr>
<tr>
<td>nDNA</td>
<td>0.0017 ± 0.0003</td>
<td>0.0012 ± 0.0003</td>
<td>1.4434 ± 0.2044</td>
<td>2.0133 ± 0.4148</td>
<td>0.2058 ± 0.1387</td>
</tr>
</tbody>
</table>

- Number of substitutions per site divided by tree length.
- Mean of branch rates.
- The SD of the branch rates.
- Coefficient of variation.
- Covariance between parent and child branch rates.
- Unpartitioned.
- Preferred partitioning scheme.
Papadopoulou et al. · doi:10.1093/molbev/msq051

...distribution; Rate(1 - 1666 times relies on the model of sequence evolution used to

...accurate estimation of substitution rates and divergence

...The Importance of Model Selection

...accurate rates obtained in these studies were plotted against

...% divergence per My

...average calibration age (Mya)

...Study a) Studies that did not account for rate heterogeneity among sites; Rate(t) = \( \frac{5.408e^{-1.794t} + 1.5231}{1.794} \), \( R^2 = 0.60 \), (b) Studies that accounted for rate heterogeneity among sites using a gamma distribution; Rate(t) = \( \frac{17.256e^{-1.157t} + 2.0968}{1.157} \), \( R^2 = 0.97 \).

...incorporating gamma-distributed rate variation. The estimated rates obtained in these studies were plotted against the average calibration age (fig. 3), and an exponential curve with three parameters was fitted under the least-squares criterion (Rate(t) = \( \mu e^{-2t} + k \)). Following Ho et al. (2005), the constant k represents the “long-term” substitution rate, which was \( k = 1.52 \) (% divergence per My) for uncorrected distances and \( k = 2.10 \) for the studies that had accounted for among-site rate variation. The latter showed a better fit to the exponential curve (\( R^2 = 0.97 \), fig. 3a) than the uncorrected rates (\( R^2 = 0.60 \), fig. 3b). In both cases, linear regression would result in a much poorer fit (\( R^2 = 0.14 \) for fig. 3a, \( R^2 = 0.19 \) for fig. 3b).

...Discussion

...The Importance of Model Selection

...Accurate estimation of substitution rates and divergence times relies on the model of sequence evolution used to correct for multiple hits (Yang 1996; Arbogast et al. 2002). However, Brower’s (1994) widely accepted estimate was based on uncorrected pairwise distances. The application of this “standard rate” using uncorrected distances has been common practice in the entomological literature (e.g., Bernhard et al. 2005; Baker et al. 2008; Canfield et al. 2008) but may lead to incorrect conclusions, particularly when estimating older divergences. In the case of Aegean tenebrionids, if we had applied a 2.3% divergence rate without correcting for multiple hits, estimated ages of the six east/west nodes would be 3.6–5.4 Mya and support a scenario of post-Messinian diversification for these lineages.

...Conversely, using the geological age of the mid-Aegean trench to estimate the substitution rate of the coxl + rnl data set without using a gamma correction resulted in rates as low as 1% or 1.2% (fig. 2). However, when accounting for rate heterogeneity among sites, either assuming a strict or a relaxed clock, the estimates were much higher (2.23% with MrBayes, 2.39% with BEAST), demonstrating that rates are greatly underestimated if they are derived from uncorrected pairwise distances (Yang 1996; Arbogast et al. 2002). This was also evident from the compilation of literature values of studies using uncorrected versus model-based methods (including a gamma distribution to account for rate heterogeneity among sites) to estimate substitution rates, which showed that the values plateau at 1.5% versus 2.1%, respectively (fig. 3).

...The choice of sequence evolution models requires careful consideration. Recent advances in model-based phylogenetics have suggested that it is often preferable to apply partition-specific rate models instead of using a single gamma distribution to describe the heterogeneity of the substitution process across multiple-gene regions and codon positions (Sullivan and Joyce 2005). Partitioned models have been shown to improve likelihood scores considerably (Castoe et al. 2004; Brandley et al. 2005), while having a great effect on branch lengths (Marshall et al. 2006) and divergence time estimates (Yang and Yoder 2003). In this study, we found the partitioning to cause an increase in inferred rates or ages by up to 12–14% (fig. 2 and table 5), with the deeper nodes affected proportionally more than the most recent nodes and a parallel increase in marginal likelihood as assessed by the harmonic mean estimator (table 4). This is in agreement with other recent studies showing an effect of the partitioning scheme on estimated divergence times (Poux et al. 2008; Torres-Carvajal and de Queiroz 2009), particularly in cases where a single calibration point is used or when an externally estimated standard molecular clock is applied (Marshall et al. 2006; Papadopoulou, Jones, et al. 2009). However, the different criteria for selecting the preferred partitioning scheme using Bayes factors (favoring the P3 or the most parameter-rich partitioning scheme P4) resulted in very similar substitution rates (2.68% vs. 2.69%), indicating that overparameterization has little impact on the estimates.
Table 5. Estimated Mean Ages ± 1 SD for Each of the East–West Nodes When Applying the Standard 2.3% Divergence My⁻¹ Rate on the mtDNA Data set Either Unpartitioned (P0) or Under Four Alternative Partitioning Schemes (P1–P4) or on cox1 and rrnL Separately.

<table>
<thead>
<tr>
<th>mtDNA P0</th>
<th>P1b</th>
<th>P2c</th>
<th>P3d</th>
<th>P4e</th>
<th>cox1 (1)⁺</th>
<th>cox1 (2)⁺</th>
<th>cox1 (3)⁺</th>
<th>rrnL</th>
</tr>
</thead>
<tbody>
<tr>
<td>EW1</td>
<td>10.9 ± 1.37</td>
<td>10.97 ± 1.47</td>
<td>11.77 ± 1.5</td>
<td>11.8 ± 1.58</td>
<td>11.79 ± 1.58</td>
<td>14.81 ± 2.25</td>
<td>17 ± 2.7</td>
<td>16.81 ± 2.73</td>
</tr>
<tr>
<td>EW2</td>
<td>8.67 ± 0.87</td>
<td>9.61 ± 1.07</td>
<td>9.47 ± 1.01</td>
<td>9.76 ± 1.07</td>
<td>9.87 ± 1.08</td>
<td>12.32 ± 1.46</td>
<td>12.9 ± 1.61</td>
<td>13.09 ± 1.66</td>
</tr>
<tr>
<td>EW3</td>
<td>4.78 ± 0.78</td>
<td>4.85 ± 0.83</td>
<td>4.85 ± 0.85</td>
<td>4.92 ± 0.87</td>
<td>4.91 ± 0.86</td>
<td>7.09 ± 1.3</td>
<td>7.45 ± 1.53</td>
<td>7.46 ± 1.55</td>
</tr>
<tr>
<td>EW4</td>
<td>9.67 ± 1.65</td>
<td>9.87 ± 1.72</td>
<td>9.81 ± 1.78</td>
<td>10.03 ± 1.84</td>
<td>10.01 ± 1.82</td>
<td>13.38 ± 2.33</td>
<td>13.98 ± 2.69</td>
<td>13.96 ± 2.77</td>
</tr>
<tr>
<td>EW5</td>
<td>11.64 ± 1.35</td>
<td>12.34 ± 1.51</td>
<td>12.38 ± 1.5</td>
<td>12.47 ± 1.54</td>
<td>12.42 ± 1.54</td>
<td>16.41 ± 2.03</td>
<td>16.88 ± 2.13</td>
<td>16.68 ± 2.15</td>
</tr>
<tr>
<td>EW6</td>
<td>7.26 ± 0.8</td>
<td>7.38 ± 0.87</td>
<td>7.42 ± 0.85</td>
<td>7.65 ± 0.91</td>
<td>7.63 ± 0.9</td>
<td>9.46 ± 1.24</td>
<td>10.18 ± 1.41</td>
<td>10.17 ± 1.43</td>
</tr>
<tr>
<td>EW7</td>
<td>12.22 ± 1.96</td>
<td>13.68 ± 2.21</td>
<td>12.78 ± 2.19</td>
<td>13.64 ± 2.38</td>
<td>13.98 ± 2.41</td>
<td>17.55 ± 3.07</td>
<td>17.59 ± 3.4</td>
<td>18.25 ± 3.56</td>
</tr>
<tr>
<td>EW8</td>
<td>10.73 ± 1.33</td>
<td>12.04 ± 1.59</td>
<td>11.52 ± 1.52</td>
<td>11.9 ± 1.59</td>
<td>12.15 ± 1.62</td>
<td>15.57 ± 2.16</td>
<td>15.53 ± 2.35</td>
<td>15.83 ± 2.38</td>
</tr>
</tbody>
</table>

Time Dependency of Molecular Rate Estimates: Phylogenetic Versus Coalescent Dating

Brower’s (1994) estimate was based on relatively recent calibration points (300 years–3.25 My) and assumed a linear relationship of sequence divergence with time over this period. This ignores the much higher apparent rate of divergence in recently split lineages (Ho et al. 2005; Ho and Larson 2006) which is possibly due to the fact that (slightly) deleterious mutations persist as transient polymorphisms in large populations, before they are lost at phylogenetic scales (Penny 2005). Even if the exact timescale affected by these higher rates and the exponential distribution describing this effect has been debated (Emerson 2007; Ho et al. 2007), the J-shaped curve of inferred rates in the literature survey fits an exponential function (fig. 3) and is consistent with an increased intraspecific rate. However, there is uncertainty with these data because a recent calibration point is needed when using population-level data to estimate mutation rates (Ho et al. 2008), which is rarely available. Consequently, the upturn in this curve is largely due to a small number of attempts to calibrate an intraspecific rate for insect taxa (Clarke et al. 2001; Gratton et al. 2008), which resulted in very high mtDNA divergence estimates (5.7% and 19.2% My⁻¹, respectively). More extensive studies on mammals (including human populations) and birds where demographic data are available confirm this exponential relationship, in broad agreement with the extrapolations of our figure 3. Yet, studies in insects continue to apply the standard phylogenetic rate of 2.3% at the population level where intraspecific calibration points are not available (Pfeiler et al. 2007; Anducho-Reyes et al. 2008; Avtjis et al. 2008; Leschen et al. 2008; McLean et al. 2008; Meng et al. 2008) and therefore risk to overestimate the inferred evolutionary ages.

Likewise, given these great differences between genealogical and phylogenetic rates, it appears critical not to mix intraspecific and interspecific sequence data when estimating lineage ages. Here, we applied an explicit procedure for removing intraspecific variation using the GMYC model, which separates independently coalescing entities from interspecific divergences based on the branching rates that are described by different equations in population-level and phylogenetic data. This mixed model was strongly favored over the null model of a uniform branching process, indicating that indeed such coalescent groups exist in this data set and can be identified by the threshold value that defines the ML point for the transition in branching rates. The GMYC model therefore was used here to reduce the phylogenetic tree to those portions that are expected to exhibit the long-term substitution rate, thereby eliminating fast-evolving population level variation that might have resulted in an overall higher rate estimate.

Caveats of Biogeographic Calibrations

The dense sampling across a putative biogeographic barrier also addresses the problem of a priori species delimitation and selection of target groups suitable for clock calibrations. Clock calibrations using biogeographical data, including those used by Brower (1994), are usually assessed based on population sampling around a “known” boundary, rather than detected in a broad survey of populations, and the divergence on either side of a given barrier is automatically ascribed to the geological events that led to the formation of the barrier. However, lineage divergence and barrier formation (and its timing) may be uncorrelated, even when multiple codistributed taxon pairs are found to be subdivided across a given barrier (Heads 2005). For example, the final rise of the Isthmus of Panama at 2.7–3.5 Mya has been widely employed for molecular clock calibrations in marine taxa, but their ancestors may have diverged well before the closure of the Isthmus (Knowlton and Weigt 1998; Marko 2002) resulting in overestimated substitution rates. Here, instead of assuming a priori that all east/west splits can be attributed to a single biogeographic event, we sampled populations comprehensively across a wide geographic area without prior assumption of a biogeographic boundary and, once a major genetic break was found consistent with known geological features (the mid-Aegean trench), only those taxa showing broad temporal congruence across this border were retained for the calibration, using a relative dating approach (Loader et al. 2007).
Another important consideration when using biogeographic separations to calibrate molecular clocks is the fact that gene divergence generally predates population divergence due to ancestral polymorphism (Edwards and Beerli 2000), which may greatly affect the divergence time estimates when looking at recently separated populations. However, the amount of error will be proportionally smaller when divergence time is long relatively to the effective population size of the ancestral population (Edwards and Beerli 2000; Hurt et al. 2009). In the example of the Isthmus of Panama, recent coalescent-based reanalyses of multiple “geminates” species data sets found that a large proportion of the observed variation in genetic distances among taxa can be explained without invoking scenarios of nonsimultaneous divergence if ancestral polymorphism and among-taxon differences in demographic history are taken into account (Hickerson et al. 2006; Hurt et al. 2009). Nonetheless, applying a similar approach on the Aegean tenebrioid data set was not considered appropriate as this data set concerns more ancient divergences, whereas each of the eastern and western clades are composed of multiple independently coalescing groups confined to sets of adjacent islands. Lineage sorting therefore might affect the relationships at the level of sets of coalescence groups within each of the eastern and western clades but is unlikely to go as far back as the east/west splits themselves. Hence, the assumptions of current coalescence-based approaches for time calibrations that require an estimate of the effective population size to be taken into account (as implemented, e.g., in the msBayes software; Hickerson et al. 2007) are not fulfilled. Moreover, our calibration was performed on all nodes simultaneously, using an uncorrelated relaxed-clock method, which permitted us to estimate a mean rate in the face of stochastic variation in coalescent times and rate heterogeneity among lineages.

We also intended to minimize the confounding effects of postseparation gene flow by selecting focal taxa that are predicted to exhibit very low dispersal propensity given their morphological traits and their habitat association (Papadopoulou, Anastasiou, et al. 2009). For example, the geophilic lineages used here coexist on the same islands with “psammophilic” tenebrioid confined to ephemeral sandy habitats (beaches, sand dunes) and winged lineages. In these groups, dispersal becomes a prerequisite for lineage persistence and consequently these taxa show much weaker geographic structure in the same island system, (Papadopoulou, Anastasiou, et al. 2009), so they are not suitable for clock calibrations based on vicariance patterns. However, despite the careful choice of study taxa, our results illustrate the difficulty of correlating a geologically dated event with the timing of the lineage split. Ocean barriers that separated eastern and western lineages may have broken down during the Messinian drying events, and therefore, the six older contemporaneous east/west splits might be attributable to a more recent event along the same biogeographic boundary. Under this scenario, the original geological separation that lasted at least 3 My (9 to 5.96 Mya) would have left no signature on the diversification of these lineages, that is, all lineages would have acquired wide distribution in the Messinian before again being subjected to vicariant separation, which seems unlikely. This scenario would also leave the younger EW3 node unexplained which now can be associated with the post-Messinian split. Attributing the six older divergence events to the original formation of the trench therefore is considered the most parsimonious explanation for the observed patterns.

**Calibrating the Insect Molecular Clock**

The final estimate for the mean mtDNA divergence rate in the Aegean tenebrioid of 2.39% My⁻¹ under a relaxed clock and applying a GTR+Γ+I or 2.69% under the preferred partitioning scheme (fig. 2) was remarkably similar to Brower’s (1994) estimate and might give credibility to this widely applied rate. However, the current analysis shows that we should guard ourselves against accepting this number as a “universal” clock rate. The way these two figures were arrived at could not be more different, as they were obtained from very recent (mostly < 1 My) versus older (10.5 ± 1.5 My) calibration ages, uncorrected versus corrected substitution rates, and separate versus simultaneous estimates from multiple calibrations. In addition, Brower’s (1994) estimations were based on a diversity of data sets from cox1, cox2, rrnS, rrnL, and restriction sites that are not easily comparable. Our estimates for Aegean tenebrioid are a composite of cox1 and slowly evolving rrnL genes (for a rate of 3.54% and 1.06%, respectively), and therefore the close similarity to the standard rate reflects an average of two quite divergent estimates. Finally, Brower’s (1994) estimate refers to a path across the tree (i.e., % divergence between terminal taxa per time unit), whereas BEAST estimates a mean rate along the root-to-tip axis (expected number of substitutions per site per time unit). Under certain conditions, these two ways to assess rates can produce different results, particularly if there is not rate constancy across the tree. Therefore, it would be preferable to express our rate estimates as substitutions/site/My rather than % divergence/My, although we maintained the latter here (and converted all numbers accordingly) for consistency with the existing literature.

Even if Brower’s (1994) calibration was based on a range of different genes, as cox1 is widely used in insect phylogenetics and in many studies is the only marker available, the 2.3% rate has been often applied to cox1-only data sets. Furthermore, since Farrell (2001) and Quek et al. (2004) both reported a rate of 1.5% for cox1, several studies (Dick et al. 2004; Kandul et al. 2004; Zhang et al. 2005; Steiner et al. 2006; Bell et al. 2007; Aoki et al. 2008; Canfield et al. 2008; Kiyoshi 2008; Leschen et al. 2008; Wirata et al. 2008; Kawakita and Kato 2009) employed this lower rate for cox1 data sets. Here, we found the substitution rate of cox1 to be more than twice as fast as the estimates of Farrell (2001) and Quek et al. (2004). Their low estimates can easily be explained as an artifact of using uncorrected pairwise distances, whereas their calibration points were comparatively ancient (up to 20 Mya). Higher rates for
cox1 have been estimated in various insect lineages by Pons and Vogler (2005) (3.34%; Coleoptera: Cucindelidae), Shapiro et al. (2006) (3-4%; Orthoptera: Tettigoniidae), Kiyoshi and Sota (2006) (3.1%; Odonata: Gomphidae), Percy et al. (2004) (2.35–3.15%; Hemiptera: Psyllidae), and Nazari and Sperling (2007) (2.3–3.1%; Lepidoptera: Papilionidae). Nonetheless, cox1 appears to be among the most slowly evolving protein-coding mitochondrial genes (Crozier RH and Crozier YC 1993), which is corroborated by recent studies (e.g., Pons and Vogler 2005: Cytochrome b (cob) 4.22% vs. cox1 3.34%). Therefore, our estimates of the mtDNA clock would likely be higher using other protein-coding genes. In contrast, the mitochondrial rRNA genes generally diverge more slowly (Trewick and Wallis 2001), in agreement with our estimate for rnr (1.08%) which is still higher than the rates proposed by Gómez-Zurita et al. (2000) (0.45%, Chrysomelidae), Pons and Vogler (2005) (0.76%, Cicindelidae) and it falls into the range reported by Percy et al. (2004) (0.95–1.9%, Psyllidae). Hence, there are consistent, large differences in rates among mitochondrial genes and therefore standard rates to date phylogenies in the absence of other independent evidence need to be based on gene-specific calibrations.

Regarding the nuclear genes, the mean rate obtained for the intron region of Mp20 (3.68%) is close to the numbers reported for synonymous substitutions in Drosophila (3% by Rowan and Hunt 1991; 3.08% by Li 1997) and for a "numt" pseudogene in Cicindela (3.33%; Pons and Vogler 2005). Moreover, it has been proposed that the average rate of neutral single-copy nuclear DNA in insects is similar to the mtDNA rate (Caccone et al. 1988; Sharp and Li 1989), which is consistent with our results, although our estimate for the intron rate remains preliminary, as the SD is large due to the small size of the intron and the stochastic nature of the substitution process. However, the fact that this rate is similar to those expected from independent studies provides confidence in the calibration based on the formation of the mid-Aegean trench and therefore also supports the mtDNA rates established here.

Overall, the substitution rates found here for the Aegean tenebrionids appear slightly higher than most previous estimates for insects, which could be partially due to the particular mode of diversification of these island lineages. High speciation rates have been associated with elevated rates of molecular evolution (Barraclough and Savolainen 2001; Webster et al. 2003), whereas lineages with small effective populations sizes are also predicted to have increased substitution rates due to the greater rate of fixation of nearly neutral alleles (Ohta 1987; Woolfit 2009). However, the hypothesis that island radiations can speed up the molecular clock has been tested explicitly in a range of different data sets and has found no support (Bromham and Woolfit 2004).

Conclusion

MtDNA substitution rates are frequently the only source of information to date historical scenarios. The desire for a standard rate of change is therefore not surprising but

the uncritical use of these rates to calibrate phylogenetic trees is fraught with errors. Here, we established the methodological issues that impact rate estimates, with the correction for rate heterogeneity among sites and the choice of genes having the greatest effect. However, despite great differences in reported rate estimates, the discrepancies can be largely explained by differences in methodology (fig. 3) without invoking large differences in the actual mtDNA substitution rate. More data sets of this kind that reduce stochastic effects by using multiple independent calibration points are needed to assess the subtle changes in rates among lineages with greater precision. This will establish whether or not the comparatively high rates found in the Aegean tenebrionids are unique to this group or reflect generally an underestimate of the standard rate for cox1.

Supplementary Material

Supplementary tables S1–S9 and figures S1–S3 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We are grateful to Bekir Keskin (Ege University, Izmir) for help with fieldwork and identification of specimens and F. Spagopoulou, M. Stalimerou, and S. Terzopoulou (University of Athens) for field assistance. We also thank four anonymous referees and the Associate Editor for helpful comments on the manuscript. This work was funded by a Greek State Scholarship Foundation studentship to A.P.V. and Natural Environment Research Council grant NE/C510908/1 to A.P.V.

References


Steiner FM, Schlick-Steiner BC, Konrad H, Linksvayer TA, Quack SP, Christian E, Stauffer C, Buschinger A. 2006. Phylogeny and...