Estimation of phylogenetic divergence times in Panagrolaimidae and other nematodes using relaxed molecular clocks calibrated with insect and crustacean fossils

Lorraine M. McGill [1,*], David A. Fitzpatrick [1], Davide Pisani [1,2] and Ann M. Burnell [1]

[1] *Department of Biology, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland*
[2] *School of Biological Sciences and School of Earth Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK*

Summary – This study presents the use of relaxed molecular clock methods to infer the dates of divergence between *Panagrolaimus* species. Autocorrelated relaxed tree methods, combined with well characterised fossil calibration dates, yield estimates of nematode divergence dates in accordance with the palaeontological age of fossil ascarid eggs and with the previously estimated date of 18 Ma (range 11.6 to 29.9 Ma) for the divergence of the *Caenorhabditis* lineage. Our data indicate that *Panagrolaimus davidi* from Antarctica separated *ca* 21.98 Ma from its currently known, most closely related strain. Thus, *P. davidi* may have existed in Antarctica prior to the Last Glacial Maximum, although this seems unlikely as it shares physiological and life history traits with closely related nematodes from temperate climates. These traits may have facilitated colonisation of Antarctica by *P. davidi* after the quaternary glaciation, analogous to the colonisation of Surtsey Island, Iceland, by *P. superbus* after its volcanic formation. This study demonstrates that autocorrelated relaxed tree methods combined with well characterised fossil calibration dates may be used as a method to estimate the divergence dates within nematodes in order to gain insight into their evolutionary history.

Keywords – 18S rDNA SSU, 28S rDNA, Antarctica, Bayesian methods, CIR process, Crustacea, divergence dates, evolutionary history, Insecta, Nematoda evolution, palaeoendemism, *Panagrolaimus*, *Panagrolaimus davidi*, *Panagrolaimus superbus*, Phylobayes, phylogeny, relaxed tree methods, Surtsey Island.

The family Panagrolaimidae comprises predominantly free-living nematodes that have evolved to survive in a wide range of substrates and locations. These free-living bacteriophage nematodes have been associated with soil, leaf-litter, rotting fruit, rotting wood and other fermenting substrates (Lazarova *et al*., 2004; Barrière & Félix, 2006; Stock & Nadler, 2006; Fonderie *et al*., 2009; McGill *et al*., 2015). They have been isolated from diverse habitats such as terrestrial deep subsurface water (Borgonie *et al*., 2011), deserts (Zhi *et al*., 2008; Darby *et al*., 2010) and polar regions (Boström, 1988; Wharton & Brown, 1989). Many of these locations have unfavourable growth conditions; however, several members of the Panagrolaimidae have adapted to survive in these extreme environments (Shannon *et al*., 2005; McGill *et al*., 2015).

Two forms of adaptive response to unfavourable environmental conditions have been described: capacity adaptations, and resistance adaptations (Wharton *et al*., 2002). Capacity adaptations enable extremophile organisms to grow and reproduce under conditions that would be lethal to most mesophiles, while resistance adaptations allow organisms to survive environmental stress by entering into a dormant state until favourable conditions return. Capacity and resistance adaptations have both been described among members of the Panagrolaimidae. *Turbatrix aceti* can tolerate extreme pH environments, maintaining activity from pH 1.6-11.0 and growing between pH 3.5-9.0 (Nicholas, 1984), and *Halicephalobus mephisto* isolated at a depth of 1.3 km from hypoxic subsurface fracture water is capable of growing at 41°C (Borgonie *et al*., 2011).

* Corresponding author, e-mail: lorrainemcgill@gmail.com

Potent resistance adaptation in the form of anhydrobiosis and cryobiosis occurs in several *Panagrolaimus* species (Aroian *et al.*, 1993; Wharton & Barclay, 1993; Wharton & Ferns, 1995; Shannon *et al.*, 2005; McGill *et al.*, 2015). Anhydrobiosis and cryobiosis refer to the reversible ametabolic state that organisms utilise to survive conditions of extreme desiccation (Crowe *et al.*, 1992) and freezing temperatures, respectively (Clegg, 2001). These forms of resistance adaptation are suited to extreme environments, and may aid in dispersal in the case of desiccated anhydrobiotes (Nkem *et al.*, 2006). *Panagrolaimus davidi*, isolated from Ross Island, Antarctica, is the best characterised example of a nematode that is anhydrobiotic but can also survive freezing when fully hydrated (Wharton & Ferns, 1995). Phylogenetically, *P. davidi* is contained within a clade of parthenogenetic, anhydrobiotic and cryotolerant nematodes (McGill *et al.*, 2015). Although *P. davidi* was isolated in Antarctica, its most closely related strain identified to date is *Panagrolaimus* sp. PS1579 isolated in San Marino, California (Shannon *et al.*, 2005; Lewis *et al.*, 2009). Outside of this clade there are several other anhydrobiotic and cryobiotic *Panagrolaimus* species and strains from diverse geographical regions (McGill *et al.*, 2015).

Panagrolaimus superbus was isolated from a gull's nest on Surtsey Island, Iceland (Boström, 1988). Surtsey Island was formed from 1963-1967 by volcanic eruptions (Bladursson & Ingadóttir, 2007). Therefore, *P. superbus* cannot have evolved on Surtsey Island; it was transported there following the volcanic origin of the island. A substantial number of Antarctic micro-invertebrates, including nematodes, show a high degree of endemism (Wharton & Ferns, 1995; Andrássy, 1998; Maslen & Convey, 2006; Convey & Stevens, 2007; Convey *et al.*, 2008; Pugh & Convey, 2008). The close relationship between *P. davidi* and other parthenogenetic, anhydrobiotic and cryotolerant nematodes from temperate regions suggests the possibility that *P. davidi*, like *P. superbus*, may not be endemic to Antarctica (Lewis *et al.*, 2009; McGill *et al.*, 2015) but may have been transported there more recently as an anhydrobiotic propagule. Information on divergence dates among the Panagrolaimidae would greatly increase our understanding of the evolution and dispersal of *P. davidi*.

Estimates of nematode divergence have been hampered both by the paucity of nematode fossils (Poinar & Boucot, 2006) to calibrate a molecular clock and the substantial heterogeneity of nucleotide substitution rates in different nematode lineages (Blaxter *et al.*, 1998). Despite these problems, molecular clock methods have been used to estimate divergence times amongst nematodes. Various strategies have been utilised: *i*) a strict molecular clock, with the molecular clock rate for globin and cytochrome being extrapolated from metazoan phylogenies, predominantly chordate and arthropod (Vanfleteren *et al.*, 1994); *ii*) the use of a single calibration point for the time of divergence of nematodes from arthropods (Coughlan & Wolfe, 2002; Stein *et al.*, 2003); and *iii*) from a taxonomically local clock inferred from internal calibration dates from the neutral mutation rate of selected genes in living populations of *Caenorhabditis elegans* (Cutter, 2008). The method employed has also been used to approximate the divergence date between *P. davidi* and its closest known relatives (Lewis *et al.*, 2009).

Strict molecular clock methods generally assign a single substitution rate to the entire tree (Zuckerkandl & Pauling, 1962). However, the realisation that heterogeneity of substitution rates between lineages and rates is common (Welch & Bromham, 2005) has led to a methodological shift towards relaxed clock methods that does not assume a constant evolutionary rate across the phylogeny (Drummond *et al.*, 2006; Lepage *et al.*, 2007; Lartillot *et al.*, 2009). Since nematodes display great heterogeneity of nucleotide substitution rates between lineages, in this study we employed relaxed molecular clock methods to infer divergence dates within the Panagrolaimidae. We used five palaeontological minimum and maximum date estimates for arthropod lineage splitting events (Benton *et al.*, 2009; Rota-Stabelli *et al.*, 2013) to calibrate the clock. Three relaxed molecular clock methods were tested: the autocorrelated CIR (Lepage *et al.*, 2007) and LogNormal (Thorne *et al.*, 1998) models, and the uncorrelated gamma multipliers (Ugam) model (Drummond *et al.*, 2006). The results obtained by these relaxed molecular clock methods were compared with those obtained using the strict molecular clock method. The results of our analyses show that relaxed molecular clock models, when combined with well characterised insect and crustacean fossil calibration dates, give nematode divergence dates that agree with the palaeontological age of fossil ascarid eggs (Poinar & Boucot, 2006) and the date for the divergence of *C. elegans* and *C. briggsae* as estimated by Cutter (2008) using internal calibration dates derived from neutral substitution rates in living populations of *C. elegans*. These correlations suggest that our estimates for lineage splitting within the Panagrolaimidae are reliable and informative. As far as we are aware, this is the first report on the use of relaxed molecular clock methodology

combined with fossil-based calibration dates to estimate nematode divergence times.

Materials and methods

SOURCES AND CULTURING OF NEMATODES

The sources and geographic origins of *Panagrolaimus* isolates used in this study are listed in Table 1. The nematodes were cultured at 20°C in the dark on nematode growth medium (NGM) plates supplemented with streptomycin sulphate (30 μg ml^{-1}) and containing a lawn of streptomycin-resistant *Escherichia coli* strain HB101 obtained from the Caenorhabditis Genetics Center (CGC) (http://www.cgc.cbs.umn.edu). Nematodes were harvested from the NGM plates using sterile distilled water as described by McGill *et al.* (2015).

DNA EXTRACTION AND rDNA SEQUENCING

Nematode DNA was extracted using a modified version of the DNeasy Blood and Tissue Extraction kit (Qiagen). A 100-200 μl packed nematode pellet was ground in 200 μl of nematode lysis buffer (20 mM Tris-HCl pH 7.5, 50 mM EDTA, 200 mM NaCl, 0.5% (w/v) SDS) under liquid nitrogen. The remainder of the extraction protocol followed the manufacturer's instructions for animal tissues, including RNase A digestion.

Primers were designed for the rDNA 18S small subunit (SSU) based on alignment of existing panagrolaimid and closely related nematode sequences available in GenBank (http://www.ncbi.nlm.nih.gov/genbank/). The 18S rDNA SSU genes of the *Panagrolaimus* species (*P. superbus*, *P. rigidus* sp. AF36, *Panagrolaimus* sp. PS1579 and *Panagrolaimus* sp. AS01) were amplified in two overlapping regions as follows: the first half was amplified using 18S_StartF (5′-TAAACACGAAACCGCGTA-3′) and 18S_InternalR (5′-ATCTGATCGCCTTCGATCCT-3′) primers. The second half was amplified with 18S_InternalF (5′-GTGAAATTCGTGGACCCTTG-3′) and 18S_EndR (5′-TACGGCCACCTTGTTACGAC-3′) primers. PCR amplicons were cloned into the pJet1.2/blunt vector (Thermo Fisher Scientific) and transformed into *E. coli* TOP10 cells (Invitrogen). Plasmids were purified using the Qiaprep Spin Miniprep kit (Qiagen) and sequenced by LGC Genomics (Berlin, Germany). Inserts from a minimum of two plasmids were sequenced in forward and reverse direction and assembled using the CAP3 assembly program (Huang & Madan, 1999). These new sequences were deposited in GenBank and are listed in Table 2.

PHYLOGENY RECONSTRUCTION

The GenBank accession numbers of nematode, arthropod and annelid 18S rDNA SSU and 28S rDNA D3 expansion region of the large ribosomal subunit (LSU) sequences used in this analysis are presented in Table 2. The sequences were aligned using MUSCLE alignment software (Edgar, 2004). The 18S rDNA SSU alignment was optimised by structural alignment with RNAsalsa (Stocsits *et al.*, 2009), using the *Saccharomyces cerevisiae* rDNA SSU structure as a reference. The 1520 character 18S rDNA SSU and 28S rDNA D3 expansions region of LSU sequences were trimmed, concatenated (Fig. S1 in the Supplementary material), and used to infer a Bayesian phylogeny with Phylobayes (version 3.3f) (Lartillot *et al.*, 2009). The annelid and kinorhynch species, *Eisenia foetida* and *Pycnophyes kielensis*, were used as outgroups.

Table 1. Source of the *Panagrolaimus* isolates used in this study.

Species and strain	Location	Habitat	Source	Reproduction
Panagrolaimus davidi	Ross Island, McMurdo Sound, Antarctica	Moss and algae	Prof. David Wharton	Parthenogenetic
Panagrolaimus rigidus sp. AF36	Fayette County, PA, USA	Soil	CGC (isolated by Prof. Andras Fodor)	Male/female (amphimictic)
Panagrolaimus superbus	Surtsey Island, Iceland	Gull's nest in a lava cavity	Prof. Bjorn Sohlenius	Male/female (amphimictic)
Panagrolaimus sp. AS01	Leixlip, Co. Kildare, Ireland	A roof gutter	Dr Adam Shannon	Male/female (amphimictic)
Panagrolaimus sp. PS1579	Huntington Botanical Gardens, San Marino, CA, USA	Soil	CGC (isolated by Prof. M.-A. Félix)	Parthenogenetic

Table 2. GenBank accession numbers for the rDNA sequences used in the molecular clock analyses.

Organism	GenBank accession number	
	28S rDNA D3 region	18S rDNA
Panagrolaimus superbus	AY878376	KC522707*
Panagrolaimus rigidus sp. AF36	AY878379	KC522706*
Panagrolaimus davidi	AY878385	AJ567385
Panagrolaimus sp. PS1579	AY878383	KC522714*
Panagrolaimus sp. AS01	FJ717472	KC522708*
Halicephalobus mephisto	GU811759	GQ918144
Halicephalobus gingivalis	DQ145637	AF202156
Caenorhabditis elegans	X03680	X03680
Caenorhabditis briggsae	AY604481	FJ380929
Nippostrongylus brasiliensis	AM039748	AJ920356
Trichostrongylus colubriformis	AM039743	AJ920350
Ascaris suum	FJ418792	U94367
Pseudoterranova decipiens	AY821763	U94380
Dorylaimus stagnalis	AY592994	AY284777
Trichodorus primitivus	AM180729	AJ439517
Nasonia vitripennis	GQ374784	GQ410677
Apis mellifera	AY703551	AB126807
Chironomus tentans	X99212	X99212
Aedes albopictus	L22060	X57172
Musca domestica	AJ551427	DQ133074
Drosophila melanogaster	M21017, M29800	M21017, M29800
Daphnia magna	AF532883	AM490278
Artemia salina	AF169697	X01723
Pycnophyes kielensis	AY863411	PKU67997
Eisenia fetida	X79872	AF212166

* Sequences generated in this study.

Each phylogenetic reconstruction for each nucleotide substitution test was repeated ten times over several random splits and the log likelihood scores averaged. The substitution rate variation was modelled using a Gamma distribution with four discrete categories. Two independent Markov chains were run in parallel and the chains compared for convergence using the tracecomp and bpcomp program with default parameters. When the chains had converged sufficiently (maxdiff < 0.3 and all effective sizes larger than 50) the chains were stopped, one-fifth of the total number of trees was removed as 'burnin' and the majority-rule posterior consensus tree constructed with the readpb program in Phylobayes.

Cross-validation tests implemented within Phylobayes were used to compare and select the best fitting nucleotide substitution model (Lartillot *et al.*, 2009). The cross-validation program within Phylobayes splits the dataset into two (unequal) parts: the learning set and the test set. The parameters of the model are trained on the learning set and these parameter values are then used to compute log likelihood scores of how well the set test is predicted by the model. When comparing two models, a positive value indicates that the test model fits the data better than the reference model (model tested against).

Saturation (*i.e.*, multiple recurrent substitutions at a given site in the sequence alignment) results from homoplasies and it can create convergences between unrelated taxa and inaccurate phylogenetic signal. Therefore, it is important that the molecular clock model anticipates the sequence saturation. In Phylobayes this is achieved using the posterior predictive model checking program ppred using the –sat option. The value of a given summary statistic observed on the true dataset is compared with the null distribution of the summary statistic under the model in a method similar to parametric bootstrapping. Each substitution model was tested for how well the model could anticipate sequence saturation using the Phylobayes ppred program. Where the posterior predictive value returned by

Table 3. The maximum and minimum date estimates used to calibrate the molecular clock (data from Benton *et al.*, 2009), together with their corresponding node numbers in Figure 2.

Node	Minimum age (Ma)	Maximum age (Ma)
Nematode and Arthropod (Node 23)	520.5	581
Crustaceans and Insects (Node 20)	510	543
Hymenoptera and Diptera (Node 18)	238.5	307.2
Within Diptera (Node 16)	238.5	295.4
Within Hymenoptera (Node 19)	152	243

Phylobayes ppred is similar to the observed homoplasy the model accurately predicts the level of saturation.

CLOCK ANALYSES

Divergence times were calculated using Phylobayes (version 3.3f) (Lartillot *et al.*, 2009) using the concatenated 18S/28S rDNA sequence alignment, the optimal tree inferred from the best fitting substitution model (Categories (CAT)-General Time Reversible (GTR)), calibration dates (Table 3), and the outgroup species identities. Two autocorrelated molecular clock models were tested: the CIR process of Lepage *et al.* (2007) and the Log-Normal (Thorne *et al.*, 1998). These were compared with the uncorrelated Ugam molecular clock model (Drummond *et al.*, 2006) and a strict molecular clock. The best-fitting relaxed clock model was selected based on the results of the cross-validation tests and its capacity to anticipate sequence saturation as described above. For non-calibrated nodes, we used a birth-date prior on divergence times and a gamma distributed root prior of 550 Ma (Precambrian, Upper Edicaran) with an SD of 50. We calibrated the clock at five nodes and treated all calibrations as soft-bounds to allow for the possibility that true divergence times may lie outside the specified calibration bounds (Yang & Rannala, 2006). A value of 2.5% of the probability to lie outside each calibration interval was allocated (see sensitivity assays). A Markov chain was run for a minimum of 10 000 cycles and the posterior mean tree was obtained using the readdiv program implemented with Phylobayes, discarding one-third of the trees as 'burnin'. Five replicate chains were run for each model and priors tested, and ten replicates were used in the final analysis to infer the final molecular clock.

SENSITIVITY ASSAYS

Bayesian interference as implemented by Phylobayes creates a posterior probability tree using a model of evolution based on prior probability distributions (*i.e.*, to allow for the probability that the true divergence that may lie outside the specified calibration bounds) for each parameter. The following tests were performed to assess the effect of varying the priors on the estimated divergence dates: *i*) relaxing the soft-bound by allowing either 2.5, 10, 20 or 50% of probability mass to be allocated outside the min-max calibration interval; *ii*) employing a less fit nucleotide substitution model; and *iii*) altering the SD of the root age. For each test, five replicate chains were ran and only the prior under investigation was changed. Each chain was run for a minimum of 10 000 cycles and the posterior mean chromatogram was obtained using the readdiv program implemented with Phylobayes, discarding one-third of the trees as 'burnin'.

Results

PHYLOGENETIC RELATIONSHIP BETWEEN THE NEMATODES AND ARTHROPODS

A Bayesian tree was constructed for the *Panagrolaimus* and other test nematode species along with the arthropod species that were used as palaeontological calibrators (Fig. 1). The CAT-GTR (Fig. 1) and CAT (Fig. S2 in the Supplementary material) substitution models gave similar topologies and fitted the data better than the GTR model (Fig. S3 in the Supplementary material) when compared by cross-validation tests (Table 4) and the estimation of saturation level (Table 5). CAT-GTR was selected as the substitution model for subsequent analysis as it performed better than the CAT model in the cross-validation and saturation tests. The tree topology inferred by the CAT-GTR model was used as the fixed topology for molecular dating under a variety of clock models. This tree topology has high Bayesian posterior probability support and is in accordance with the current hypothesis of phylogenetic relationship for these taxa. The *Panagrolaimus* nematodes form a single clade. *Halicephalobus mephisto* and *H. gingivalis*, also members of the family Panagrolaimidae, form a sister clade to the *Panagrolaimus* species. All other nematodes are grouped outside the Panagrolaimidae, with Dorylaimia and Enoplia nematodes the most basal in the clade. The arthropods form a single clade. The Diptera (*Musca domestica*, *Drosophilia melanogaster*, *Chironomus tentans* and *Aedes albopictus*) form a clade, with the Hymenoptera (*Nasonia vitripennis* and *Apis*
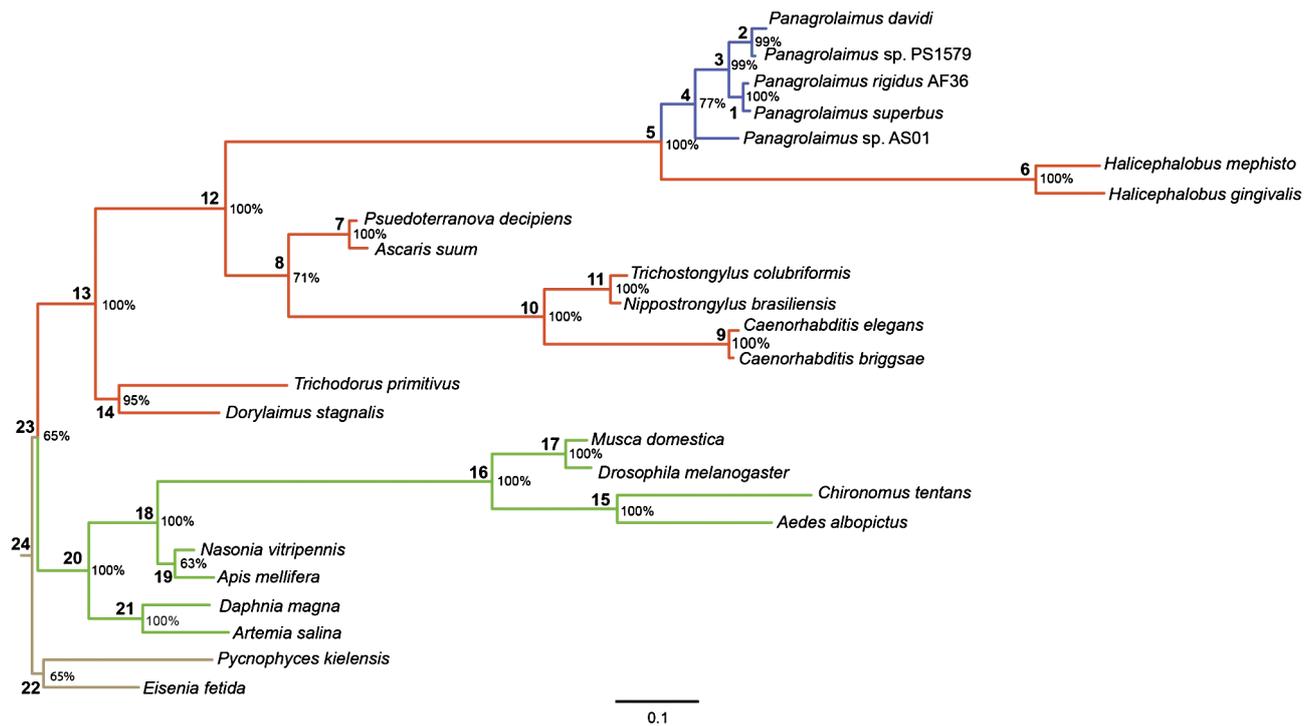
Fig. 1. Hypothesis of phylogenetic relationships among nematodes and arthropod species used for molecular clock analyses. Bayesian inference tree derived from concatenated sequences from the rDNA SSU and 28S rDNA D3 expansion regions of the LSU under the CAT-GTR nucleotide substitution model. The sequences were aligned using MUSCLE (Edgar, 2004) and the 18S rDNA alignment was optimised with RNAsalsa (Stocsits *et al.*, 2009). Branch supports and the node numbers are shown in bold. *Panagrolaimus* sp. are collared, other nematodes are indicated in red, arthropod calibrator species are indicated in green and outgroups are in brown.

Table 4. Cross-validation of substitution models ± SD.

Reference*	GTR-CAT	CAT	GTR
GTR-CAT	–	−2.636 ± 6.47421	−7.197 ± 5.77984
CAT	2.636 ± 6.47421	–	−4.561 ± 9.11368
GTR	7.197 ± 5.77984	4.561 ± 9.11368	–

* A positive value for the test model against the reference models indicates that the test model fits the data better.

Table 5. Substitution model prediction of sequence saturation ± SD.

Model	Observed homoplasy	Posterior predictive
GTR-CAT	3.15361 ± 0.251034	3.12606 ± 0.260003
CAT	3.06436 ± 0.24575	3.05596 ± 0.262086
GTR	1.8788 ± 0.0708481	1.8116 ± 0.095836

mellifera) forming a separate group. The crustaceans (*Daphnia magna* and *Artemia salina*) form a distinct clade constituting the sister group of the insects.

MOLECULAR CLOCK ANALYSIS

Having selected CAT-GTR as the best-fitting nucleotide substitution model (Tables 4, 5) we compared divergence dates under four different molecular clock models: the autocorrelated LogNormal and CIR models, the uncorrelated gamma model, and the strict molecular clock model. Cross-validation tests of the molecular clock models tested under a CAT-GTR fixed tree topology found that there was no clear best-fit model between CIR, LogNormal and Ugam (Table 6). The strict molecular clock was found to fit the data considerably less than the CIR, Log-

Table 6. Cross-validation of molecular clock models ± SD.

Reference*	CIR	LN	Ugam	Strict
CIR	–	1.242 ± 2.0816	2.582 ± 3.21491	−70.878 ± 11.417
LN	−1.242 ± 2.0816	–	1.34 ± 1.48416	−72.12 ± 12.852
Ugam	−2.582 ± 3.21491	−1.34 ± 1.48416	–	−73.46 ± 13.0805
Strict	70.878 ± 11.417	72.12 ± 12.852	73.46 ± 13.0805	–

* A positive value for the test model against the reference models indicates that the test model fits the data better.

Table 7. Molecular clock model prediction of sequence saturation ± SD.

Model	Observed homoplasy	Posterior predictive
CIR + GTR-CAT	2.67699 ± 0.175591	2.66249 ± 0.187497
CIR + GTR	1.848 ± 0.0709461	1.78675 ± 0.0871034
CIR + CAT	2.48757 ± 0.150912	2.48319 ± 0.160678
LN + GTR-CAT	2.93967 ± 0.212665	2.92201 ± 0.221946
Ugam + GTR-CAT	3.78092 ± 1.01431	3.77079 ± 1.04156
Strict + GTR-CAT	2.80579 ± 0.195787	2.76971 ± 0.202151

Normal and Ugam models (Table 6). Each model was also tested for its ability to predict sequence saturation when run under a CAT-GTR fixed topology (Table 7). All models predicted similar levels of sequence saturation to those observed in the sequence data. The uncorrelated Ugam model predicted higher saturation that the other models, but, with a comparatively high SD for these values, it was not considered to be significantly different to the other models with low SD values.

The optimal dates at each node for the CIR, LogNormal and Ugam model were compared and plotted (Fig. 2A, B). The optimal dates between each model for each node were highly correlated (Table S1 in the Supplementary material) with a highly significant P-value of <0.0001. The dates between the CIR and LogNormal autocorrelated models were similar (Spearman $r = 0.9757$, $R^2 = 0.9682$) (Fig. 2A). The Ugam and CIR results were also similar (Spearman $r = 0.9704$, $R^2 = 0.9075$), with the largest difference between the models at node 14 (Dorylaimia and Enoplia) 21 (crustaceans) and 22 (Annelida and Kinorhyncha) (Fig. 2B). For nodes 14, 21 and 22, the Ugam model predicts a wide range of minimum and maximum divergence dates (Fig. S4B in the Supplementary material). The CIR optimal date lies within the dates of the Ugam for nodes 14, 21 and 22, but with a lower optimal divergence date. The optimal divergence dates when compared to the strict molecular clock (Fig. 2C) were correlated (Spearman $r = 0.947$,

$R^2 = 0.9030$); however, when comparing the maximum and minimum divergence dates of each model the strict model differs (Fig. S4C). The strict model has a narrow range between its minimum and maximum dates. For many of the nodes the maximum dates (Table S1) are lower than the optimal dates of the other models, so the strict model may underestimate the divergence dates.

Since the strict model was predicted to fit to the data least by cross-validation it was not considered further. The Ugam model was also discarded because the range between the minimum and maximum divergence dates that it generated was wide with differences of up to 457 Ma (Fig. S4B). The correlated CIR and LogNormal autocorrelated models gave similar optimal, minimum and maximum dates and from cross-validation and saturation tests either model would be suitable for further analysis (Fig. S4A). The CIR model was selected for the final molecular clock (Fig. 3; Table 8) as globally the range between the minimum and maximum dates was narrower with this model than with the LogNormal model.

The estimated molecular divergence times for four of the five calibration nodes obtained using the autocorrelated CIR model were similar to the fossil calibration dates (Table 3), with the optimal divergence dates lying within the range of the minimum and maximum dates proposed by Benton *et al*. (2009). The only date that differed was the estimated date of 321.2 Ma (306.5-342.6 Ma) for the divergence between the Hymenoptera and the Diptera. This is older than the date used for calibrating the tree (238.5-307.2 Ma). However, recent phylogenies present evidence that the divergence of the Hymenoptera from the other holometabolous insects (including the Diptera) is much older than 238.5-307.2 Ma. Misof *et al*. (2014) estimated the divergence date between the Hymenoptera and the Diptera at 344.7 Ma (317.8-372.4 Ma) – the optimal divergence date that we obtained for the divergence between the Hymenoptera and the Diptera under the CIR model lies within this range.
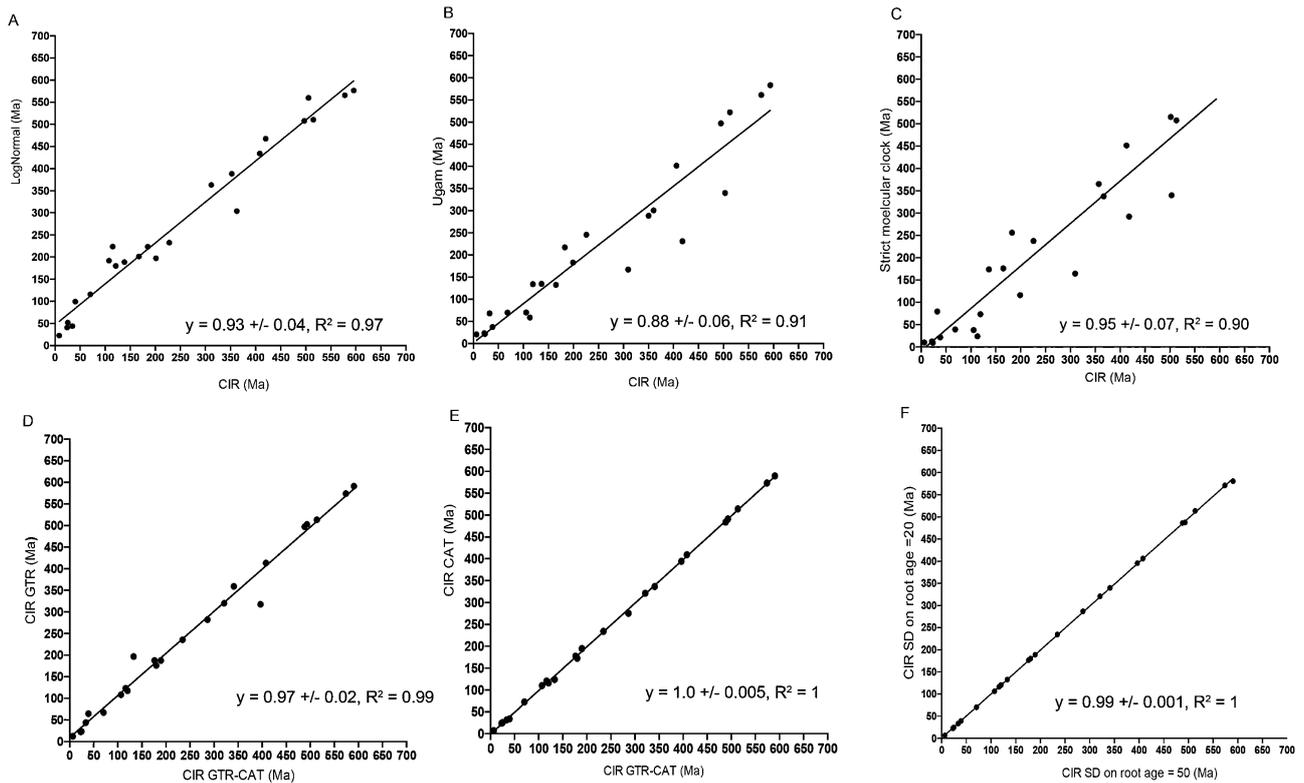
Fig. 2. Sensitivity experiments showing the effect of different models and parameters. Divergence dates were assessed after varying different priors, and dates plotted (*y*-axis) against the dates obtained under the best-fitting model and priors (CIR + CAT-GTR models, with default soft-bound and using calibrator dates) to investigate the properties (slope and R^2) of the regression line interpolating the two estimates. A: CIR model compared to the LogNormal model; B: CIR model compared to the Ugam model; C: CIR model compared to the strict molecular clock; D: CIR model with the CAT-GTR nucleotide substitution model compared to the CIR model with the GTR nucleotide substitution model; E: CIR model with the CAT-GTR nucleotide substitution model compared to the CIR model with the CAT nucleotide substitution model; F: CIR with a root-age SD = 50 Ma compared to CIR with a root-age SD = 20. Each dot represents a node as indicated in Figure 1.

Sᴇɴsɪᴛɪᴠɪᴛʏ ᴀssᴀʏs

Relaxing the soft-bounds

A soft-bound prior probability distribution may be applied to the calibration dates to allow for the probability that the true divergence times may lie outside the specified calibration bounds (Yang & Rannala, 2006). The effect of varying the soft-bound prior on the divergence dates for each node was tested for 10, 20 and 50% (Table S2 in the Supplementary material). Changing the soft-bound prior does not affect the majority of the nodes but the divergence dates for nodes 15, 16, 17 become younger and the divergence dates for nodes 18 and 21 become older as the soft-bounds are relaxed. Altering the soft-bound has the largest impact on nodes that have been calibrated rather than nodes near the root or within the nematodes.

As we do not have any calibration points within the nematode section of the tree, the strictest setting of a default value of 2.5% was maintained as the soft-bound prior in the final molecular clock (Fig. 3) and for all subsequent analyses.

Employing a less fit nucleotide substitution model

Cross-validation and saturation tests found the CAT-GTR nucleotide substitution model to fit the data best. Saturation test showed that the CIR model combined with either CAT-GTR or CAT nucleotide substitution models predicted similar levels of saturation and the predicted levels of saturation were similar to those observed in the sequence data (Table 7). The CIR model, when combined with the GTR nucleotide substitution model, predicted
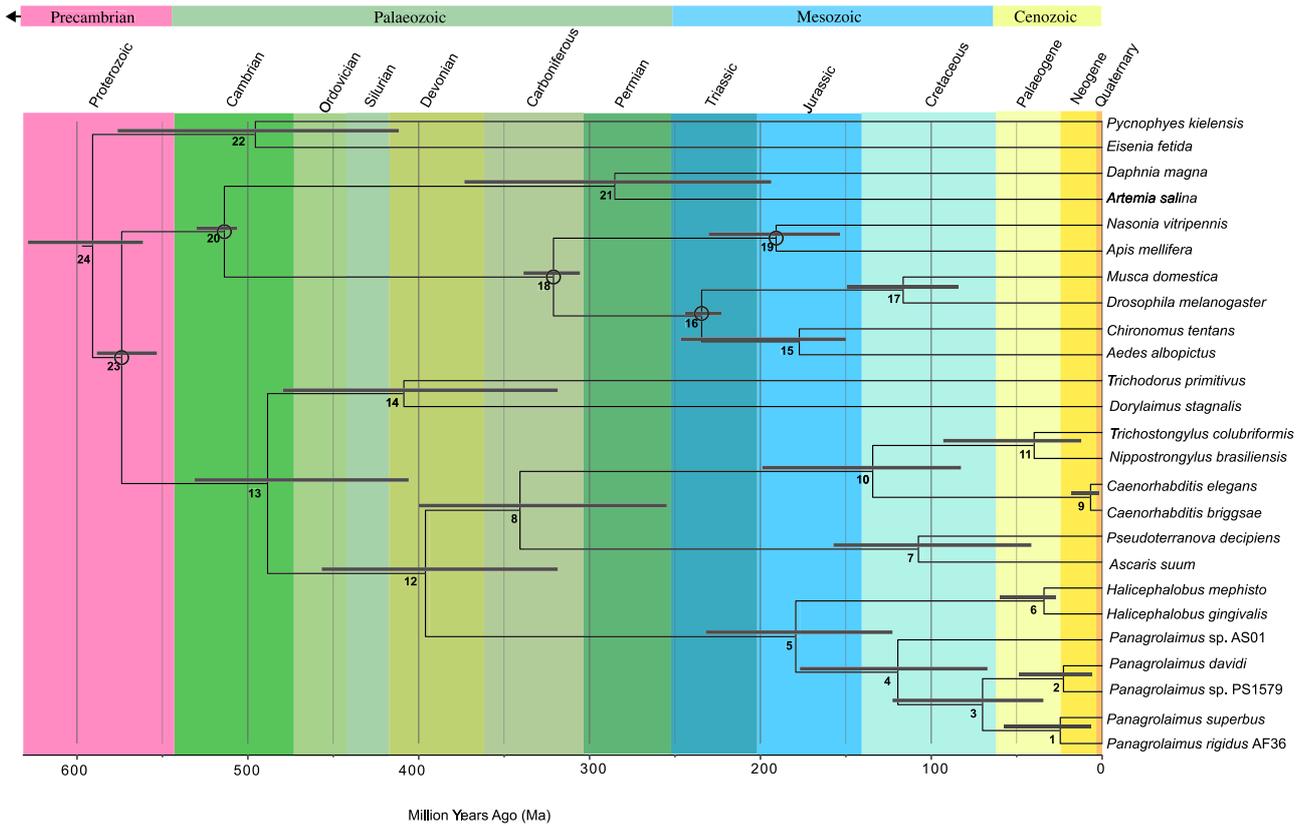
Fig. 3. Chronogram obtained for selected nematode taxa using the autocorrelated CIR model (Lepage *et al.*, 2007) applied to rDNA SSU and 28S rDNA D3 expansion regions of the LSU. The 95% credibility intervals for the age expansion estimates (grey bars) are shown at each node. Fossil calibration dates for nodes 16, 18, 19, 20 and 23 (Table 3) were obtained from Benton *et al.* (2009) and are indicated by open circles.

less saturation than observed and detected less saturation than all other model combinations.

There was no effect on the optimal, minimum and maximum estimated divergences when comparing either nucleotide substitution model with the CIR model for the majority of nodes (Fig. 2D, E; Table S3 in the Supplementary material). Under the CIR + GTR, node 10 (*C. elegans*/*C. briggsae* and *T. colubriformis*/*N. brasiliensis*) and node 12 (Panagrolaimidae and other nematodes) estimated older divergence dates than the CIR + CAT-GTR or CIR + CAT combinations and, as nematode divergence dates are the subject of this study and the cross-validation and saturation data found GTR to be the least fit model, the GTR substitution model was not used in subsequent analyses. When combined with CIR, the CAT-GTR and CAT models give near identical divergence dates. Since the CAT-GTR model was found to fit better following saturation and cross-validation tests, it was used for all

subsequent analysis in combination with the CIR molecular clock model.

Altering the SD on the root age prior

Lartillot *et al.* (2009) strongly advise the explicit declaration of a prior value for the root when calculating molecular divergence times. We specified a prior root age of 550 Ma (Precambrian, Upper Edicaran (Cohen *et al.*, 2013; updated), and investigated the effect of changing the SD of the root age prior probability value from 50-20 Ma. Five replicate chains were ran and the mean divergence times were calculated and compared to a SD of the root age prior of 50 Ma (Table S4 in the Supplementary material). Changing the SD did not affect the divergence time for any node ($R^2 = 0.999$, *P*-value $= 0.0001$) (Fig. 2F). A prior age of 550 Ma with an SD of 50 was used in the final CIR CAT-GTR analysis (Fig. 3).

Table 8. The optimal, minimum and maximum divergence dates estimated for each node by the CIR molecular clock model + CAT-GTR nucleotide substitution model.

	CIR optimal	CIR min	CIR max
Node 1	24.76 ± 0.13	7.28 ± 0.06	56.65 ± 0.68
Node 2	22.94 ± 0.1	7.81 ± 0.1	49.96 ± 0.2
Node 3	70.51 ± 0.33	31.97 ± 0.21	122.17 ± 0.94
Node 4	120.37 ± 0.38	68.72 ± 0.26	178.64 ± 0.73
Node 5	179.99 ± 0.33	123.52 ± 0.91	241.31 ± 0.47
Node 6	33.52 ± 0.08	17 ± 0.09	62.94 ± 0.23
Node 7	106.93 ± 0.26	50.58 ± 0.32	163.49 ± 0.57
Node 8	341.16 ± 0.44	269.22 ± 1.17	400.77 ± 0.65
Node 9	6.62 ± 0.03	2.18 ± 0.01	17.52 ± 0.24
Node 10	133.09 ± 0.4	78.34 ± 0.54	198.48 ± 0.95
Node 11	39.34 ± 0.22	12.51 ± 0.11	85.11 ± 0.84
Node 12	396.66 ± 0.48	322.37 ± 0.92	456.92 ± 0.72
Node 13	488.25 ± 0.38	430.82 ± 1.2	530.55 ± 0.43
Node 14	407.92 ± 0.28	319.79 ± 0.82	473.66 ± 0.42
Node 15	176.4 ± 0.22	147.25 ± 0.54	201.2 ± 0.33
Node 16	234.54 ± 0.02	221.56 ± 0.11	242.4 ± 0.09
Node 17	116.35 ± 0.22	77.91 ± 0.71	148.98 ± 0.41
Node 18	321.16 ± 0.07	306.45 ± 0.05	342.57 ± 0.33
Node 19	189.65 ± 0.19	153.47 ± 0.2	232.72 ± 0.36
Node 20	513.67 ± 0.02	507.98 ± 0.02	525.62 ± 0.12
Node 21	286.34 ± 0.61	191.57 ± 0.83	369.6 ± 0.46
Node 22	493.13 ± 0.59	411.15 ± 1.48	577.9 ± 1.11
Node 23	573.95 ± 0.04	555.4 ± 0.12	584.03 ± 0.05
Node 24	590.58 ± 0.22	564.84 ± 0.13	627.85 ± 0.55

Discussion

Nematoda are soft-bodied and as most are only millimetres in length they fossilise poorly. There are two main sources of nematode fossils: in amber or in coprolites of subfossils. Since the nematode fossil record is poor and palaeontological evidence of divergence is lacking for the vast majority of nematodes, including the Panagrolaimidae, other methods to calculate divergence times are required. Our data demonstrate that phylogenetic relaxed molecular clock methods, combined with well-established fossil dates of arthropods and crustaceans, provide an accurate method to estimate nematode divergence dates and gain insight into their evolutionary history.

SELECTION OF THE MOST APPROPRIATE NUCLEOTIDE SUBSTITUTION MODEL AND MOLECULAR CLOCK MODEL FOR THE NEMATODE AND ARTHROPOD LINEAGES IN THIS DATASET

The rate of nucleotide substitution during evolutionary time varies between nucleotide character states at a given

site and can also vary across the nucleotide sites in a DNA sequence. For example, sites that are more constrained by natural selection show fewer substitutions than sites that are less constrained. Failure to take into account this rate heterogeneity among sites can lead to biased estimations of branch lengths, with corresponding impacts on estimates of evolutionary timescales (Jia *et al.*, 2014). In cross-validation and nucleotide saturation tests the CAT-GTR substitution model was found to fit best the observed substitution rates across the nucleotide sites in our rDNA alignment. Under the CAT (Categories) model, each site in the alignment is potentially heterogeneous with respect to the substitution process, with the total number of classes being a free variable that reflects the substitutional complexity of the underlying data set (Lartillot & Philippe, 2004). The general time-reversible (GTR) model includes parameters that allow unequal frequencies for the four nucleotides and a distinct rate for each of the six possible pairwise nucleotide substitutions (Jia *et al.*, 2014). Combining the GTR model for nucleotide substitution at a given nucleotide site with the CAT across sites substitution rate model has frequently been found to fit both nucleotide and amino acid alignment data better than either the GTR or CAT model alone (Lartillot *et al.*, 2009).

The divergence time estimates obtained using four molecular clock evolutionary models were compared in a series of cross-validation and sensitivity tests to determine which molecular clock model had best described our nematode and arthropod rDNA alignment. Three relaxed molecular clock methods were tested; *i)* the autocorrelated CIR (Lepage *et al.*, 2007) and LogNormal (Thorne *et al.*, 1998) models; *ii)* the uncorrelated gamma multipliers (Ugam) model (Drummond *et al.*, 2006); and *iii)* the strict molecular clock molecular clock model. A strict molecular clock model considers that nucleotide substitutions occur at a constant rate through time among all lineages in a phylogeny. Relaxed molecular clock models allow the substitution rate to vary among lineages, and autocorrelated relaxed clock models assume that neighbouring branches in a phylogeny share similar rates of evolution and that the rate of evolutionary change along a branch depends on the time duration of the branch (Lepage *et al.*, 2007). In this way, the substitution rate under the relaxed clock model can be regarded as a trait that evolves through time, perhaps in correlation with life-history characteristics (Ho & Duchene, 2014). Cross-validation assays indicated that for our dataset the CIR (Lepage *et al.*, 2007) and LogNormal (Thorne *et al.*, 1998) autocorrelated molecular clock models outperformed the uncorre-

lated gamma multipliers model (Drummond *et al*., 2006) and the strict molecular clock model. The divergence date estimates obtained under the CIR model were selected for the final molecular time tree (Fig. 3) because the ranges between the minimum and maximum divergence dates at individual nodes were narrower with this model than with the LogNormal model. Sensitivity assays showed that the divergence dates obtained using the CIR clock model remained robust when tested for changes to the soft-bound probabilities of the calibration dates when using an inferior nucleotide substitution model or when the standard deviation of the root age was changed. Having tested a representative selection of the available molecular models and approaches for estimating variations in nucleotide substitution rates among lineages in phylogenetic trees, we consider that the estimated divergence dates we present for the nematode lineages in this study are robust and reliable.

SUPPORT FOR MOLECULAR DATE ESTIMATES FROM FOSSIL ASCARID EGGS

Validation of our molecular clock divergence estimates are found in nematode eggs identified as members of the family Ascarididae that were isolated in dinosaur coprolites from the Bernissart Iguanodon beds in Belgium (Poinar & Boucot, 2006). This bone bed is dated between the upper Barremian and lowermost Aptian age of the Cretaceous period 124-127.24 Ma (Schnyder *et al*., 2009). According to our molecular clock analyses, the estimated divergence time for *Ascaris* (Ascaridida: Ascarididae) and *Pseudoterranova* (Ascaridida: Anisakidae) under the CIR model is (50.58-163.49 Ma), with an optimal date of 106.9 Ma. This is in good agreement with the data for the Benissart nematode fossils, considering that palaeontological dates always post-date the cladogenetic events identified by molecular data.

DIVERGENCE TIME ESTIMATES FOR *C. ELEGANS* AND *C. BRIGGSAE*

The divergence time of *C. elegans* and *C. briggsae* is of much scientific interest. Previously, the divergence time of *C. elegans* and *C. briggsae* was estimated to be 80-110 Ma based on the divergence date between nematodes and arthropods of *ca* 800-1000 Ma used to calibrate the clock (Stein *et al*., 2003). However, it is now clear that the 800-1000 Ma split between arthropods and nematodes is inaccurate (Erwin *et al*., 2011; Rota-Stabelli *et al*., 2013). The split between arthropods and

nematodes occurred more recently at 520-581 Ma (Benton *et al*., 2009). Cutter (2008) obtained a divergence date of 18 Ma (range 11.6-29.9 Ma) for these two *Caenorhabditis* species. This result was obtained analytically, assuming: *i*) a fixed mutation rate, estimated from *C. elegans* data; *ii*) that each year corresponds to six generations in *C. elegans* and *C. briggsae*; and *iii*) that equal rates of substitution had occurred in the seven *Caenorhabditis* species considered. Our results agree with those of Cutter (2008) in supporting a recent divergence time between *C. elegans* and *C. briggsae* (6.6 Ma, range 2.2-17.5 Ma). A large number of inferences have been drawn with reference to rates of genomic evolution in Nematoda and in animals more broadly (*e.g*., Lynch, 2007). Many were based on divergence estimates of *ca* 100 Ma between *C. elegans* and *C. briggsae*. Based on our results and the results of Cutter (2008) it is clear that many of these conclusions need to be reassessed using a more recent divergence date estimate.

MOLECULAR DATE ESTIMATES LINEAGE SPLITTING EVENTS IN PANAGROLAIMIDAE

Within the order Rhabditida our estimates place the divergence of Panagrolaimidae within the Devonian at 406.18 Ma (322.37-456.92 Ma). The estimated divergence time of *Panagrolaimus* and *Halicephalobus* (both members of the superfamily Panagrolaimidae) was in the Jurassic at 178 Ma (123.5-241.3). By the Cretaceous an anhydrobiotic phenotype had evolved in the ancestor of the lineage leading to the *davidi*, *superbus* and *Panagrolaimus* sp. AS01 clades. *Halicephalobus mephisto* was originally recovered from a borehole of palaeowater, ^{14}C dating estimating the age of this borehole palaeowater at 4500-6000 BP (Borgonie *et al*., 2011). Our molecular clock analyses estimates the date of divergence between *H. mephisto* and *H. gingivalis* at 33.5 Ma (17.0-62.9 Ma), indicating that these two lineages are considerably older than the date of the origin of the borehole palaeowater where *H. mephisto* was originally isolated.

Our data predict that the optimal mean divergence times for the Antarctic *P. davidi* and the Californian *Panagrolaimus* sp. PS1579 is 21.98 Ma (Aquitanian, Miocene). Prior to this lineage-splitting event, the common ancestor of the members of the *davidi* clade already possessed a robust freezing-tolerant phenotype along with the capacity to inhibit the growth and recrystallisation of ice (McGill *et al*., 2015). Lewis *et al*. (2009) inferred divergence estimates of 14 000-140 000 years between *P. davidi* and *Panagrolaimus* sp. PS1579, a date that is sig-

nificantly younger than the divergence dates estimated in our analyses. The divergence dates estimated by Lewis *et al*. (2009) are based on extrapolation from the rate of silent mutation in *C. elegans* and an assumption of either ten or one nematode generations per year in these *Panagrolaimus* species. It is difficult to estimate the number of generations per year for a genus that can survive for more than 8 years in an anhydrobiotic state (Aroian *et al*., 1993). If *P. davidi* and/or its sister species have experienced cumulative periods of anhydrobiosis during their evolutionary history, their divergence time is likely to be older than that proposed by Lewis *et al*. (2009).

IS *P. DAVIDI* PALAEOENDEMIC TO THE ANTARCTIC?

In the early Eocene Epoch (*ca* 55 Ma) Australia began to separate from Antarctica and drift northwards, allowing the Antarctic Circumpolar Current to develop and resulting in the thermal isolation of the Antarctic continent (Kennett, 1977). During the Oligocene (38-22 Ma) glacial conditions became established throughout Antarctica. High cooling rates occurred in Antarctica during the middle Miocene (*ca* 14 Ma), a step often referred to as the Miocene climate transition (Flower & Kennett, 1993). Our data indicate that *P. davidi* separated from *Panagrolaimus* sp. PS1579 22.9 Ma (7.8-50 Ma) – a period when the Antarctic cryosphere was expanding rapidly. A similar divergence date was estimated for the anhydrobiotic and freezing-tolerant nematodes *P. superbus* (Surtsey, Iceland) and *P. rigidus* (Pennsylvania, USA).

The reproductive and ecophysiological characters required for survival in Antarctic environments are generally consistent with the predictions of adversity (A) or stress (S) selection (Greenslade, 1983; Convey, 2000, 2009). Such A/S selected traits, *viz*., long life span, parthenogenesis, low reproductive rates, low temperature thresholds for activity, and tolerance of extreme cold during low energy-cost dormancy (Greenslade, 1983), are clearly seen in the Antarctic nematode *Scottnema lindsayae*. *Scottnema lindsayae* is the sole member of a monotypic genus that is anhydrobiotic (Treonis *et al*., 2000) and freezing-tolerant (Wharton & Raymond, 2015), and is abundant and widespread in coastal regions and some nunataks in continental Antarctica. In the laboratory its optimal growth temperature is 10°C (Caldwell, 1981), its reproductive cycle is 218 days and its fecundity declines if cultured at a higher temperature (Overhoff *et al*., 1993). By contrast, *P. davidi* possesses *r*-selected traits: high fecundity, rapid growth rates, and a short life cycle (Brown *et al*., 2004; Stocsits *et al*., 2009) – traits that are typical of many temperate nematodes with high colonising ability (Bongers, 1990). The optimal growth temperature of *P. davidi* is *ca* 25°C and its life cycle at this temperature is *ca* 10 days. At a lower temperature of 10°C its life cycle increases to *ca* 36 days, with a greatly reduced fecundity (Brown *et al*., 2004; Stocsits *et al*., 2009).

While many Antarctic invertebrates (including *Scottnema*) are likely to be palaeoendemic to Antarctica (Convey & Stevens, 2007; Pugh & Convey, 2008), the great majority of Antarctica biota (particularly arthropods and bryophytes) are most probably recent colonists that became established in Antarctica after the end of the quaternary glaciation *ca* 10 000-15 000 years ago (Convey, 2009). The time since the last Pleistocene glaciation is relatively short in evolutionary terms for new capacity and resistance adaptations to arise in Antarctic invertebrates, particularly allowing for their extended life cycles. Thus, Convey (2009) hypothesises that Antarctic colonisers need to have prior possession of the behavioural, ecophysiological and biochemical phenotypes necessary for their survival in Antarctica. Our data indicate that *P. davidi* separated from *Panagrolaimus* sp. PS1579 *ca* 22.9 Ma, and thus *P. davidi* may have existed in Antarctica prior to the Last Glacial Maximum. However, we feel that this is unlikely and hypothesise that *P. davidi* may have been transported in an anhydrobiotic state to Antarctica at a more recent date. We believe this because the physiological and life history traits of *P. davidi* show no evidence of an evolved response to polar conditions that differs from closely related nematodes from temperate climates. All members of the *P. davidi* clade are anhydrobiotic, freezing-tolerant, parthenogenetic, and have *r*-selected reproduction. This combination of reproductive and A/S selected resistance phenotypes is likely to have contributed to the wide geographic dispersal of the members of the clade and facilitated the colonisation of Antarctica by *P. davidi* after the end of the quaternary glaciation in a manner analogous to the recent colonisation of Surtsey Island by *P. superbus*.

Conclusion

The lack of nematode fossils has made the estimation of divergence dates of nematodes difficult. This study demonstrates that relaxed molecular clock methods, in combination with known insect and crustacean fossil divergence dates, can provide accurate and reliable divergence date estimates for nematodes. Using molecular phylogenetic methods, we have estimated the divergence

dates for anhydrobiotic and freezing tolerant lineages of *Panagrolaimus* nematodes, thereby providing an insight into their evolutionary history. On the basis of these divergence dates, combined with physiological traits, we hypothesise that the nematode *P. davidi* is a recent coloniser of the Antarctic. This analysis was completed using *ca* 1500 characters from two ribosomal gene sequences. With the ongoing generation of large-scale nematode sequence datasets, relaxed molecular methods are likely to be more widely used to investigate and robustly confirm nematode divergence dates.

Acknowledgement

References

Andrássy, I. (1998). Nematodes of the sixth continent. *Journal of Nematode Morphology and Systematics* 1, 107-186.

Aroian, R.V., Carta, L., Kaloshian, I. & Sternberg, P.W. (1993). A free-living *Panagrolaimus* sp. from Armenia can survive in anhydrobiosis for 8.7 years. *Journal of Nematology* 25, 500-502.

Barrière, A. & Félix, M. (2006). Isolation of *C. elegans* and realted nematodes. In: The *C. elegans* Research Community (Eds). *WormBook*. Pasadena, CA, USA, pp. 1-9.

Benton, M.J., Donoghue, P.C.J. & Asher, R.J. (2009). Calibrating and constraining molecular clocks. In: Hedges, S.B. & Kumar, S. (Eds). *The time tree of life*. Oxford, UK, Oxford University Press, pp. 35-86.

Bladursson, S. & Ingadóttir, A. (2007). *Nomination of Surtsey for the UNESCO world heritage list*. Reykjavík, Iceland, Icelandic Institute of Natural History.

Blaxter, M.L., De Ley, P., Garey, J.R., Liu, L.X., Scheldeman, P., Vierstraete, A., Vanfleteren, J.R., Mackey, L.Y., Dorris, M., Frisse, L.M. *et al.* (1998). A molecular evolutionary framework for the phylum Nematoda. *Nature* 392, 71-75. DOI: 10.1038/32160

Bongers, T. (1990). The maturity index – an ecological measure of environmental disturbance based on nematode species composition. *Oecologia* 83, 14-19. DOI: 10.1007/Bf00324627

Borgonie, G., Garcia-Moyano, A., Litthauer, D., Bert, W., Bester, A., van Heerden, E., Moller, C., Erasmus, M. & Onstott, T.C. (2011). Nematoda from the terrestrial deep subsurface of South Africa. *Nature* 474, 79-82. DOI: 10.1038/nature09974

Boström, S. (1988). Descriptions and morphological variability of three populations of *Panagrolaimus* Fuchs, 1930 (Nema-

toda, Panagrolaimidae). *Nematologica* 34, 144-155. DOI: 10.1163/002825988X00233

Brown, I.M., Wharton, D.A. & Millar, R.B. (2004). The influence of temperature on the life history of the Antarctic nematode *Panagrolaimus davidi*. *Nematology* 6, 883-890. DOI: 10.1163/1568541044038641

Caldwell, J.R. (1981). The Signy Island terrestrial reference sites: XIII. Population dynamics of the nematode fauna. *British Antarctic Survey Bulletin* 54, 33-46.

Clegg, J.S. (2001). Cryptobiosis – a peculiar state of biological organization. *Comparative Biochemistry and Physiology – Part B* 128, 613-624. DOI: 10.1016/S1096-4959(01)00300-1

Cohen, K.M., Finney, S., Gibbard, P.L. & Fan, J.-X. (2013; updated). International chronostratigraphic chart. Available online at http://www.stratigraphy.org/ICSchart/ChronostratChart2017-02.pdf.

Convey, P. (2000). How does cold constrain life cycles of terrestrial plants and animals? *CryoLetters* 21, 73-82.

Convey, P. (2009). Comparative studies of Antarctic arthropod and bryophyte lifecycles: do they have a common strategy? In: Battaglia, B., Valencia, J. & Walton, D.W.H. (Eds). *Antarctic communities: species, structure and survival*. Cambridge, UK, Cambridge University Press, pp. 321-327.

Convey, P. & Stevens, M.I. (2007). Antarctic biodiversity. *Science* 317, 1877-1878. DOI: 10.1126/Science.1147261

Convey, P., Gibson, J.A.E., Hillenbrand, C.D., Hodgson, D.A., Pugh, P.J.A., Smellie, J.L. & Stevens, M.I. (2008). Antarctic terrestrial life – challenging the history of the frozen continent? *Biological Reviews* 83, 103-117. DOI: 10.1111/J.1469-185x.2008.00034.X

Coughlan, A. & Wolfe, K.H. (2002). Fourfold faster rate of genome rearrangement in nematodes than in *Drosophilia*. *Genome Research* 12, 857-867. DOI: 10.1101/gr.172702

Crowe, J.H., Hoekstra, F.A. & Crowe, L.M. (1992). Anhydrobiosis. *Annual Review of Physiology* 54, 579-599. DOI: 10.1146/annurev.physiol.54.1.579

Cutter, A.D. (2008). Divergence times in *Caenorhabditis* and *Drosophila* inferred from direct estimates of the neutral mutation rate. *Molecular Biology and Evolution* 25, 778-786. DOI: 10.1093/Molbev/Msn024

Darby, B.J., Neher, D.A. & Belnap, J. (2010). Impact of biological soil crusts and desert plants on soil microfaunal community composition. *Plant and Soil* 328, 421-431. DOI: 10.1007/s11104-009-0122-y

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLOS Biology* 4, 699-710. DOI: 10.1371/Journal.Pbio.0040088

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32, 1792-1797. DOI: 10.1093/Nar/Gkh340

Erwin, D.H., Laflamme, M., Tweedt, S.M., Sperling, E.A., Pisani, D. & Peterson, K.J. (2011). The Cambrian conundrum: early divergence and later ecological success in the

early history of animals. *Science* 334, 1091-1097. DOI: 10. 1126/Science.1206375

Flower, B.P. & Kennett, J.P. (1993). Middle Miocene ocean-climate transition – high-resolution oxygen and carbon isotopic records from deep-sea drilling project site 588a, southwest Pacific. *Paleoceanography* 8, 811-843. DOI: 10.1029/93pa02196

Fonderie, P., Willems, M., Bert, W., Houthoofd, W., Steel, H., Claeys, M. & Borgonie, G. (2009). Intestine ultrastructure of the facultative parasite *Halicephalobus gingivalis* (Nematoda: Panagrolaimidae). *Nematology* 11, 859-868. DOI: 10.1163/156854109X428025

Greenslade, P.J.M. (1983). Adversity selection and the habitat templet. *American Naturalist* 122, 352-365. DOI: 10.1086/284140

Ho, S.Y. & Duchene, S. (2014). Molecular-clock methods for estimating evolutionary rates and timescales. *Molecular Ecology* 23, 5947-5965. DOI: 10.1111/mec.12953

Huang, X.Q. & Madan, A. (1999). Cap3: a DNA sequence assembly program. *Genome Research* 9, 868-877. DOI: 10.1101/Gr.9.9.868

Jia, F., Lo, N. & Ho, S.Y. (2014). The impact of modelling rate heterogeneity among sites on phylogenetic estimates of intraspecific evolutionary rates and timescales. *PLOS ONE* 9, e95722. DOI: 10.1371/journal.pone.0095722

Kennett, J.P. (1977). Cenozoic evolution of Antarctic glaciation, the circum-Antarctic Ocean, and their impact on global paleocenography. *Journal of Geophysical Research-Oceans and Atmospheres* 82, 3843-3860. DOI: 10.1029/JC082i027p03843

Lartillot, N. & Philippe, H. (2004). A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Molecular Biology and Evolution* 21, 1095-1109. DOI: 10.1093/Molbev/Msh112

Lartillot, N., Lepage, T. & Blanquart, S. (2009). PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25, 2286-2288. DOI: 10.1093/Bioinformatics/Btp368

Lazarova, S.S., de Goede, R.G.M., Peneva, V.K. & Bongers, T. (2004). Spatial patterns of variation in the composition and structure of nematode communities in relation to different microhabitats: a case study of *Quercus dalechampii* Ten. forest. *Soil Biology & Biochemistry* 36, 701-712. DOI: 10.1016/j.soilbio.2004.01.005

Lepage, T., Bryant, D., Philippe, H. & Lartillot, N. (2007). A general comparison of relaxed molecular clock models. *Molecular Biology and Evolution* 24, 2669-2680. DOI: 10.1093/Molbev/Msm193

Lewis, S.C., Dyal, L.A., Hilburn, C.F., Weitz, S., Liau, W.S., LaMunyon, C.W. & Denver, D.R. (2009). Molecular evolution in *Panagrolaimus* nematodes: origins of parthenogenesis, hermaphroditism and the Antarctic species *P. davidi. BMC Evolutionary Biology* 9, 15. DOI: 10.1186/1471-2148-9-15

Lynch, M. (2007). *The origins of genome architecture*. Sunderland, MA, USA, Sinauer Associates.

Maslen, N.R. & Convey, P. (2006). Nematode diversity and distribution in the southern maritime Antarctic – clues to history? *Soil Biology & Biochemistry* 38, 3141-3151. DOI: 10.1016/J.Soilbio.2005.12.007

McGill, L.M., Shannon, A.J., Pisani, D., Felix, M.A., Ramlov, H., Dix, I., Wharton, D.A. & Burnell, A.M. (2015). Anhydrobiosis and freezing-tolerance: adaptations that facilitate the establishment of *Panagrolaimus* nematodes in polar habitats. *PLOS ONE* 10, e0116084. DOI: 10.1371/Journal.Pone.0116084

Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G. *et al.* (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346, 763-767. DOI: 10.1126/science.1257570

Nicholas, W.L. (1984). *The biology of free-living nematodes*. Oxford, UK, Oxford University Press.

Nkem, J.N., Wall, D.H., Virginia, R.A., Barrett, J.E., Broos, E.J., Porazinska, D.L. & Adams, B.J. (2006). Wind dispersal of soil invertebrates in the McMurdo Dry Valleys, Antarctica. *Polar Biology* 29, 346-352. DOI: 10.1007/S00300-005-0061-X

Overhoff, A., Freckman, D.W. & Virginia, R.A. (1993). Life-cycle of the microbivorous Antarctic dry valley nematode *Scottnema lindsayae* (Timm 1971). *Polar Biology* 13, 151-156.

Poinar Jr, G.O. & Boucot, A.J. (2006). Evidence of intestinal parasites of dinosaurs. *Parasitology* 133, 245-249. DOI: 10.1017/S0031182006000138

Pugh, P.J.A. & Convey, P. (2008). Surviving out in the cold: Antarctic endemic invertebrates and their refugia. *Journal of Biogeography* 35, 2176-2186. DOI: 10.1111/J.1365-2699.2008.01953.X

Rota-Stabelli, O., Daley, A.C. & Pisani, D. (2013). Molecular timetrees reveal a Cambrian colonization of land and a new scenario for ecdysozoan evolution. *Current Biology* 23, 392-398. DOI: 10.1016/j.cub.2013.01.026

Schnyder, J., Dejax, J., Keppens, E., Tu, T.T.N., Spagna, P., Boulila, S., Galbrun, B., Riboulleau, A., Tshibangu, J.P. & Yans, J. (2009). An Early Cretaceous lacustrine record: organic matter and organic carbon isotopes at Bernissart (Mons Basin, Belgium). *Palaeogeography Palaeoclimatology Palaeoecology* 281, 79-91. DOI: 10.1016/J.Palaeo.2009.07.014

Shannon, A.J., Browne, J.A., Boyd, J., Fitzpatrick, D.A. & Burnell, A.M. (2005). The anhydrobiotic potential and molecular phylogenetics of species and strains of *Panagrolaimus* (Nematoda, Panagrolaimidae). *Journal of Experimental Biology* 208, 2433-2445. DOI: 10.1242/jeb.01629

Stein, L.D., Bao, Z.R., Blasiar, D., Blumenthal, T., Brent, M.R., Chen, N.S., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A. *et al.* (2003). The genome sequence of *Caenorhabditis brig-*

gsae: a platform for comparative genomics. *PLOS Biology* 1, 166-192. DOI: 10.1371/Journal.Pbio.0000045

Stock, S.P. & Nadler, S.A. (2006). Morphological and molecular characterisation of *Panagrellus* spp. (Cephalobina: Panagrolaimidae): taxonomic status and phylogenetic relationships. *Nematology* 8, 921-938. DOI: 10.1163/156854106779799277

Stocsits, R.R., Letsch, H., Hertel, J., Misof, B. & Stadler, P.F. (2009). Accurate and efficient reconstruction of deep phylogenies from structured RNAs. *Nucleic Acids Research* 37, 6184-6193. DOI: 10.1093/Nar/Gkp600

Thorne, J.L., Kishino, H. & Painter, I.S. (1998). Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15, 1647-1657.

Treonis, A.M., Wall, D.H. & Virginia, R.A. (2000). The use of anhydrobiosis by soil nematodes in the Antarctic Dry Valleys. *Functional Ecology* 14, 460-467. DOI: 10.1046/J.1365-2435.2000.00442.X

Vanfleteren, J.R., Van de Peer, Y., Blaxter, M.L., Tweedie, S.A., Trotman, C., Lu, L., Van Hauwaert, M.-L. & Moens, L. (1994). Molecular genealogy of some nematode taxa as based on cytochrome c and globin amino acid sequences. *Molecular Phylogenetics and Evolution* 3, 92-101. DOI: 10.1006/mpev.1994.1012

Welch, J.J. & Bromham, L. (2005). Molecular dating when rates vary. *Trends in Ecology & Evolution* 20, 320-327. DOI: 10.1016/j.tree.2005.02.007

Wharton, D.A. & Barclay, S. (1993). Anhydrobiosis in the free-living antarctic nematode *Panagrolaimus davidi* (Nematoda: Rhabditida). *Fundamental and Applied Nematology* 16, 17-22.

Wharton, D.A. & Brown, I.M. (1989). A survey of terrestrial nematodes from the McMurdo Sound region, Antarctica. *New Zealand Journal of Zoology* 16, 467-470. DOI: 10.1080/03014223.1989.10422914

Wharton, D.A. & Ferns, D.J. (1995). Survival of intracellular freezing by the Antarctic nematode *Panagrolaimus davidi*. *Journal of Experimental Biology* 198, 1381-1387.

Wharton, D.A. & Raymond, M.R. (2015). Cold tolerance of the Antarctic nematodes *Plectus murrayi* and *Scottnema lindsayae*. *Journal of Comparative Physiology B – Biochemical Systemic and Environmental Physiology* 185, 281-289. DOI: 10.1007/s00360-014-0884-2

Wharton, D.A., Goodall, G. & Marshall, C.J. (2002). Freezing rate affects the survival of a short-term freezing stress in *Panagrolaimus davidi*, an Antarctic nematode that survives intracellular freezing. *CryoLetters* 23, 5-10.

Yang, Z.H. & Rannala, B. (2006). Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution* 23, 212-226. DOI: 10.1093/molbev/msj024

Zhi, D.J., Li, H.Y. & Nan, W.B. (2008). Nematode communities in the artificially vegetated belt with or without irrigation in the Tengger Desert, China. *European Journal of Soil Biology* 44, 238-246. DOI: 10.1016/j.ejsobi.2007.09.006

Zuckerkandl, E. & Pauling, L.B. (1962). Molecular disease, evolution and genic heterogeneity. In: Kasha, M. & Pullman, B. (Eds). *Horizons in biochemistry: Albert Szent-Györgyi dedicatory volume*. New York, NY, USA, Academic Press, pp. 97-166.

Supplementary material

Fig. S1. Concatenated and trimmed 18S rDNA SSU and 28S rDNA D3 expansions region of LSU sequences generated using MUSCLE alignment software.
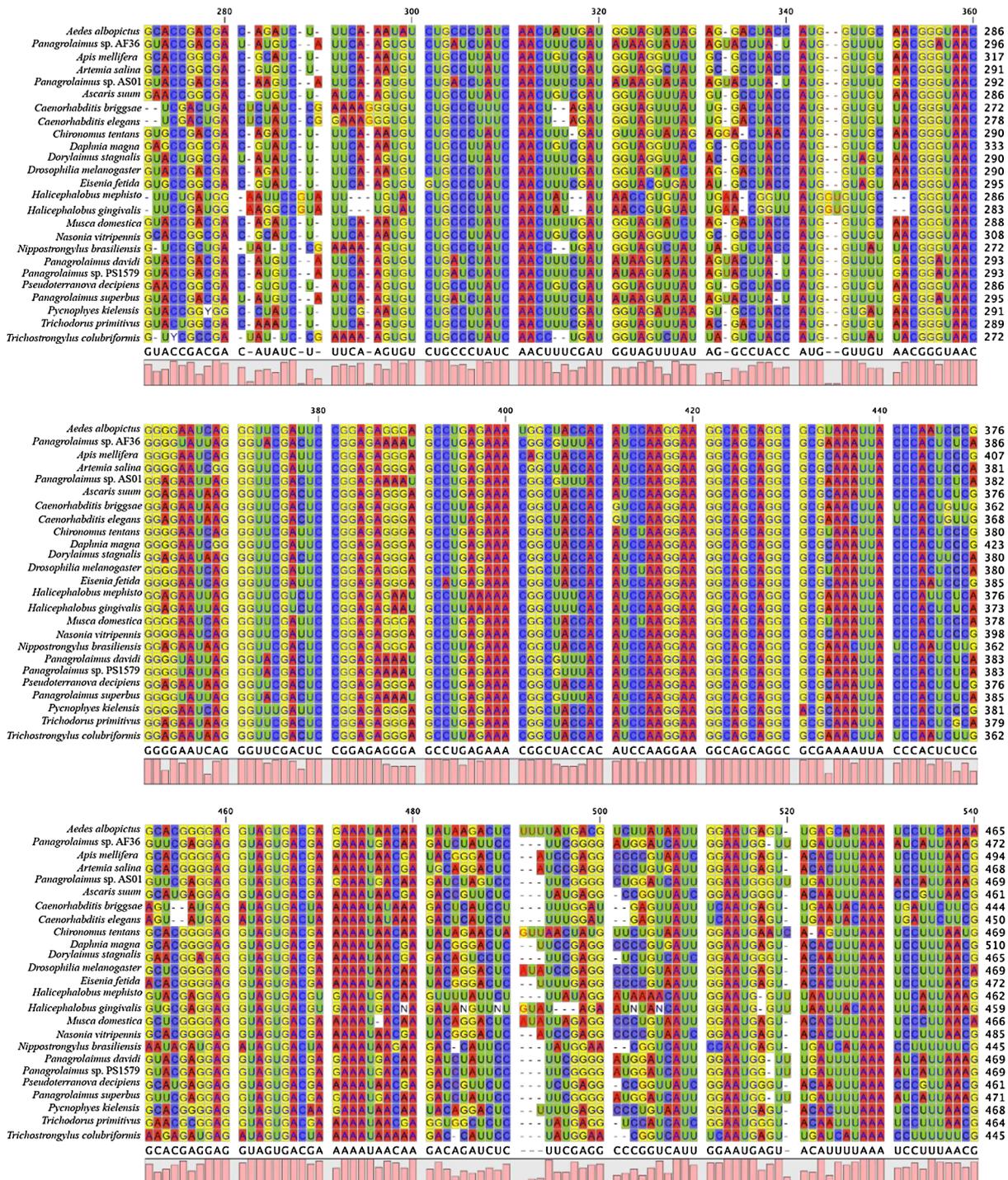
Fig. S1. (Continued.)

Fig. S1. (Continued.)
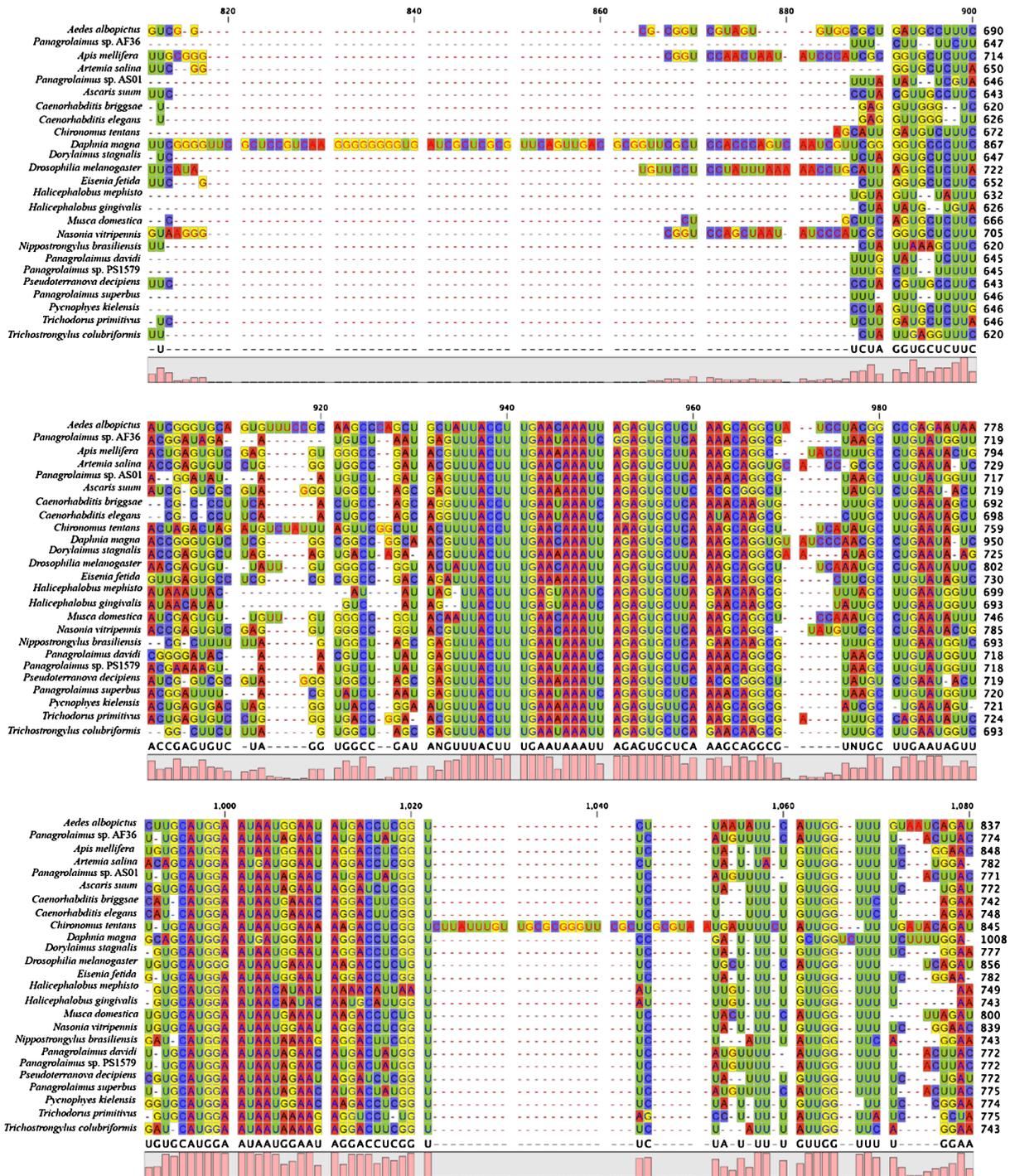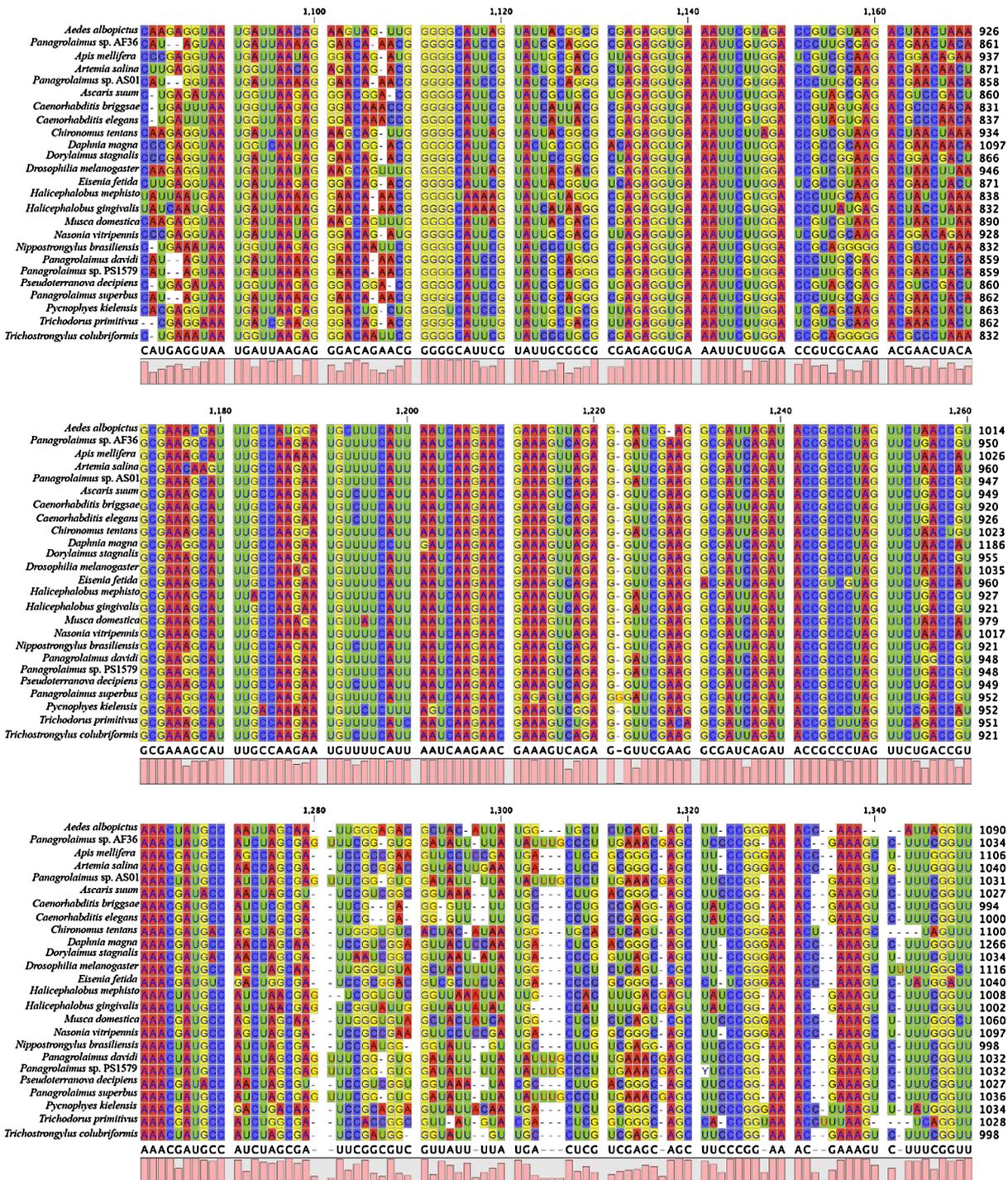
Fig. S1. (Continued.)
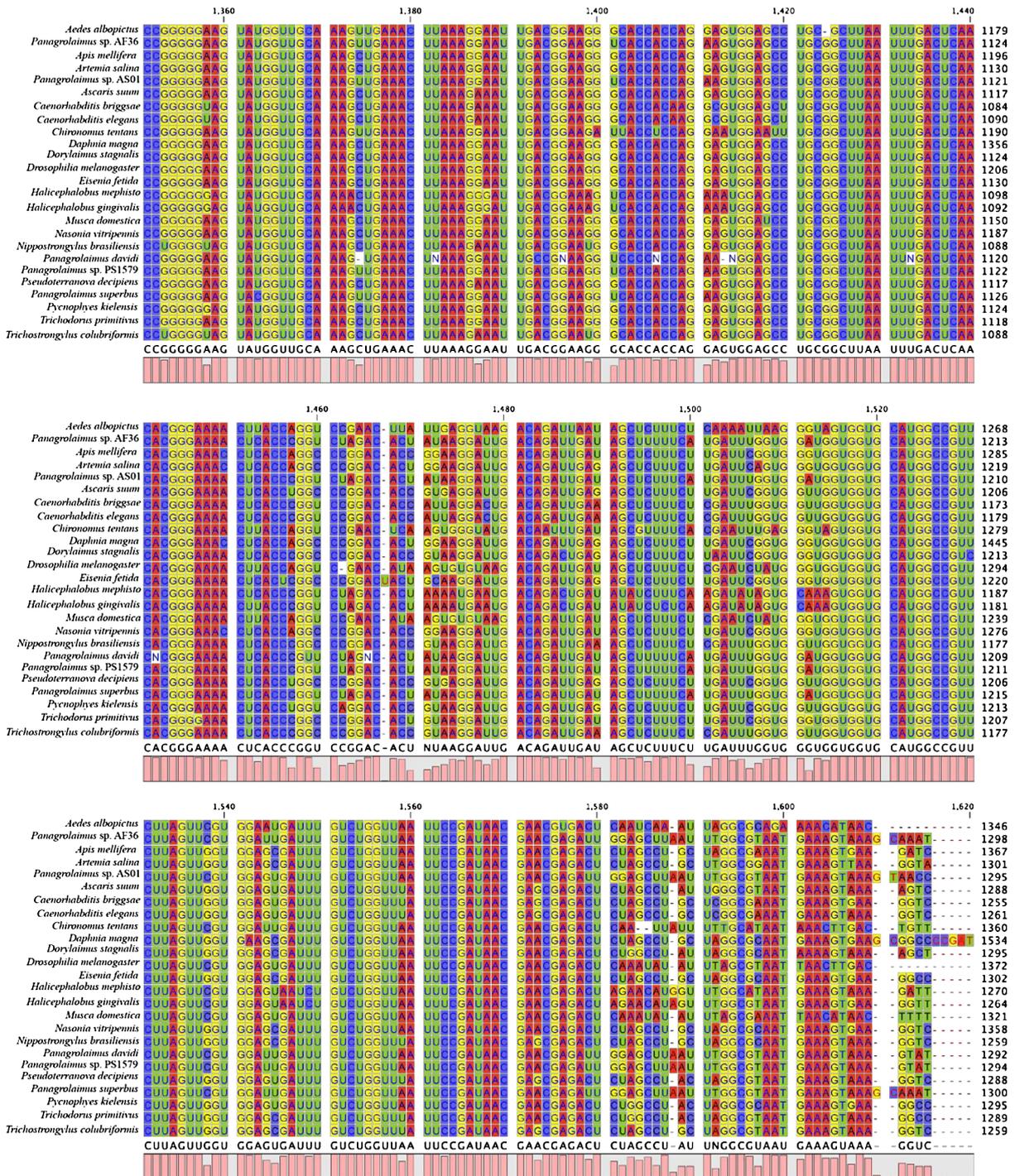
Fig. S1. (Continued.)

Fig. S1. (Continued.)
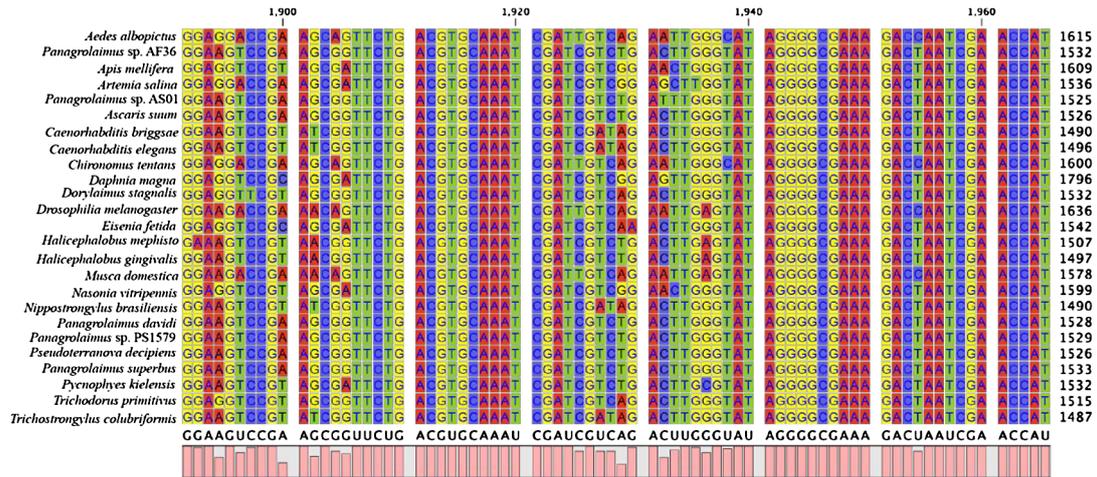
Fig. S1. (Continued.)

Fig. S1. (Continued.)
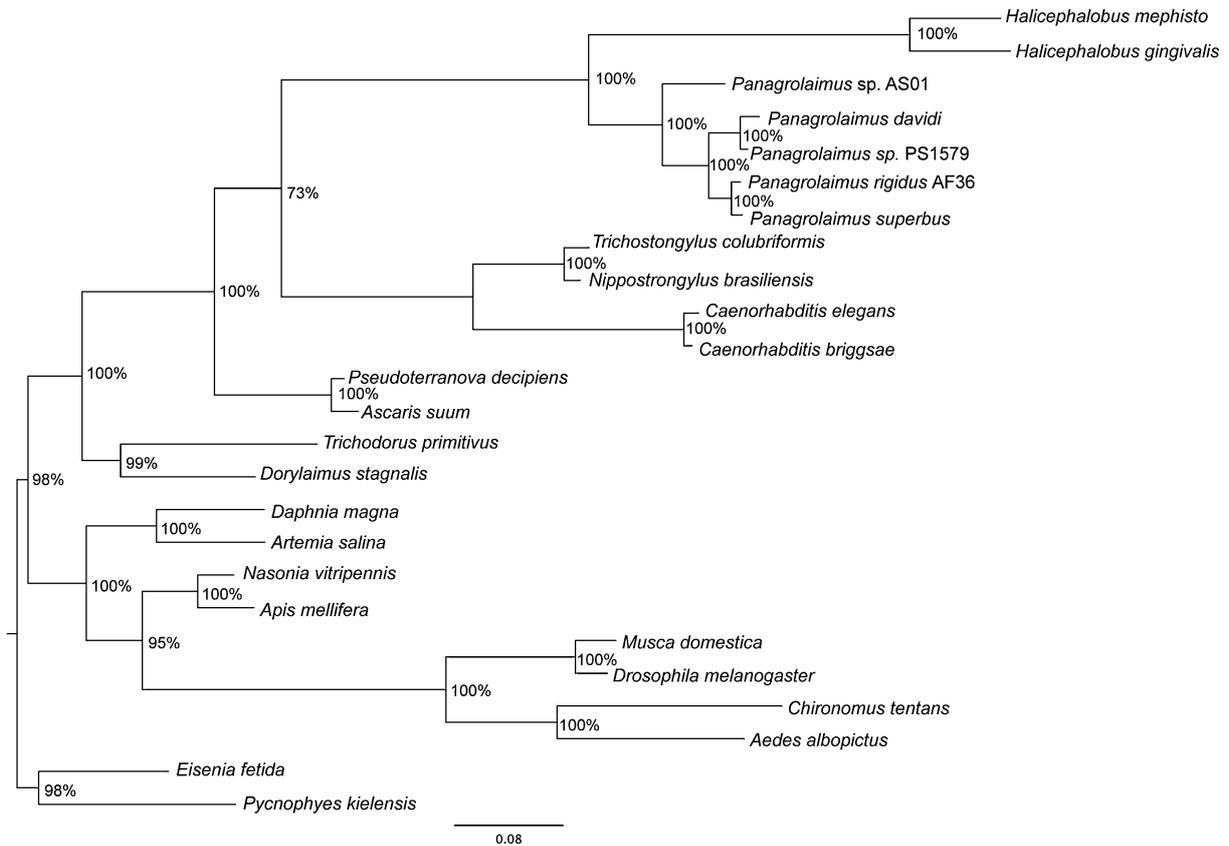
Fig. S2. Hypothesis of phylogenetic relationships among nematodes and arthropod species used for molecular clock analyses under the CAT nucleotide substitution model. Bayesian inference tree derived from concatenated sequences from the rDNA SSU and 28S rDNA D3 expansion regions of the LSU under the CAT nucleotide substitution model. The sequences were aligned using MUSCLE (Edgar, 2004) and the 18S rDNA alignment was optimized with RNAsalsa (Stocsits *et al.*, 2009). Branch supports are shown.
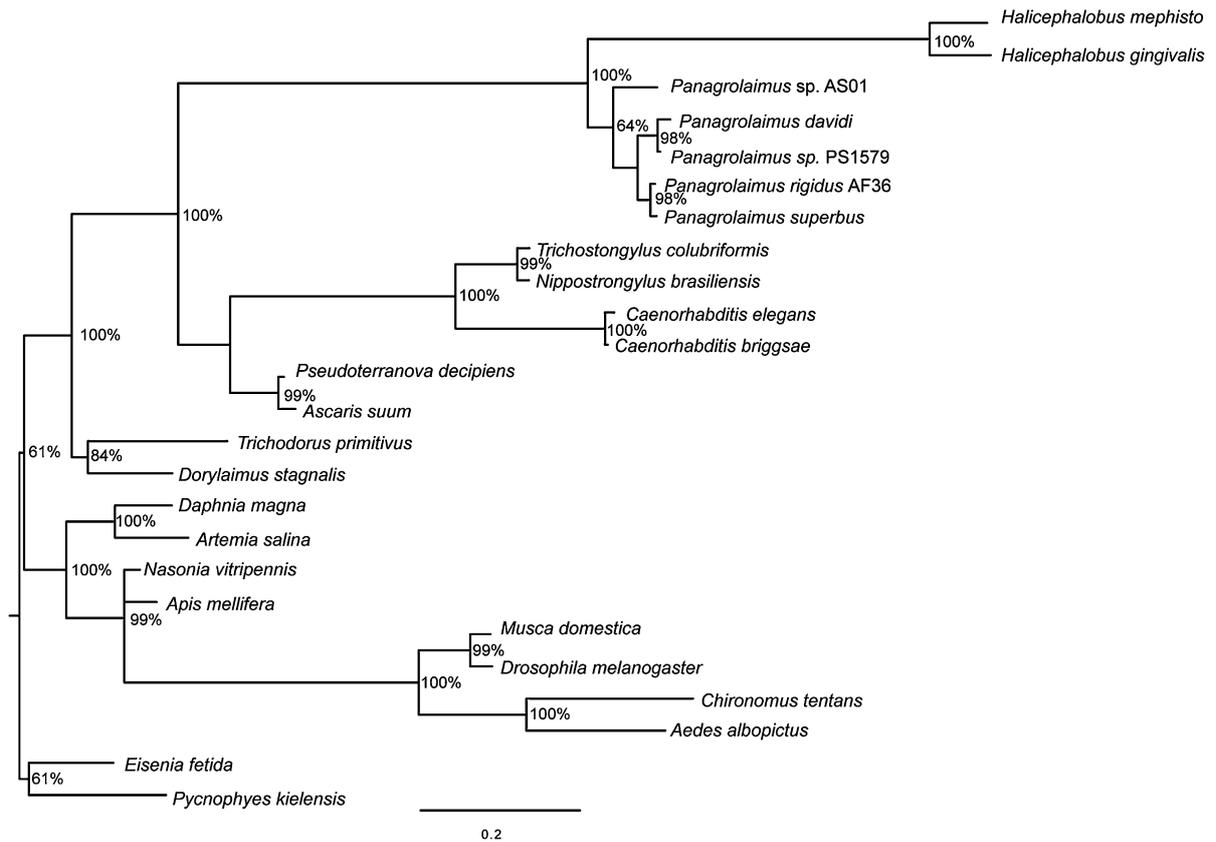
Fig. S3. Hypothesis of phylogenetic relationships among nematodes and arthropod species used for molecular clock analyses under the GTR nucleotide substitution model. Bayesian inference tree derived from concatenated sequences from the rDNA SSU and 28S rDNA D3 expansion regions of the LSU under the GTR nucleotide substitution model. The sequences were aligned using MUSCLE (Edgar, 2004) and the 18S rDNA alignment was optimized with RNAsalsa (Stocsits *et al*., 2009). Branch supports are shown.
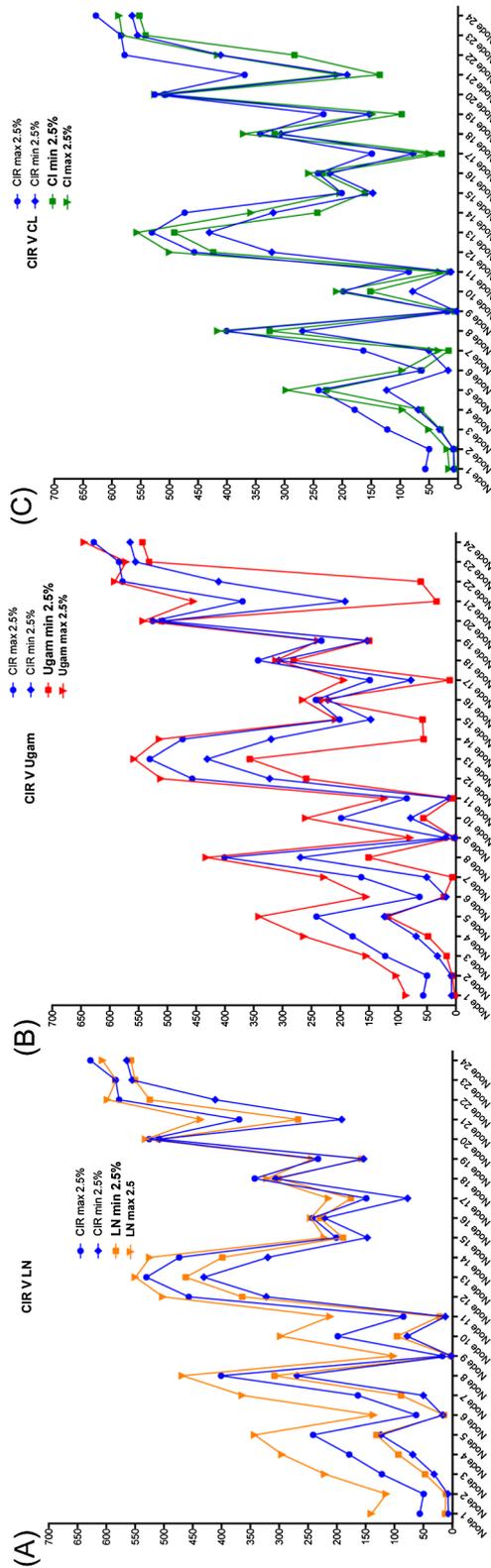
Fig. S4. Comparison of the estimated minimum and minimum node ages (95% credibility intervals) for each model. For these analyses the CAT-GTR model of nucleotide substitution evolution was used and a prior root age of 550 Ma with a SD of 50 and the default soft-bound value was set. The estimated dates under the LogNormal (A), Ugam (B) and strict molecular clock (C) models were compared to the CIR model.

Table S1. Minimum, maximum and optimal divergence dates (± SE) under the CIR, LogNormal, Ugam and Strict molecular clock models. Nodes are as indicated in Figure 1.

|	CIR optimal	CIR min	CIR max	LN optimal	LN min	LN max	Ugam optimal	Ugam min	Ugam max	Strict optimal	Strict min	Strict max
Node 1	24.76 ± 0.13	7.28 ± 0.06	56.65 ± 0.68	57.39 ± 0.4	13 ± 0.1	140.13 ± 0.78	21.72 ± 0.1	2.41 ± 0.01	86.08 ± 1.37	9.71 ± 0.02	5.55 ± 0.03	15.06 ± 0.07
Node 2	22.94 ± 0.1	7.81 ± 0.1	49.96 ± 0.2	46.29 ± 0.24	11.64 ± 0.04	114.3 ± 0.7	23.31 ± 0.09	6.56 ± 3.25	102.87 ± 16.86	12.37 ± 0.03	7.73 ± 0.05	18.18 ± 0.08
Node 3	70.51 ± 0.33	31.97 ± 0.21	122.17 ± 0.94	121.21 ± 0.51	47.56 ± 0.27	220.81 ± 0.6	69.93 ± 0.2	16.09 ± 3.19	154.52 ± 16.36	39.38 ± 0.09	30.24 ± 0.17	49.96 ± 0.14
Node 4	120.37 ± 0.38	68.72 ± 0.26	178.64 ± 0.73	185.44 ± 0.47	93.52 ± 1.53	294.85 ± 1.08	133.89 ± 0.54	48.28 ± 0.28	262.94 ± 1.79	78.96 ± 0.12	63.69 ± 0.28	95.78 ± 0.22
Node 5	179.99 ± 0.33	123.52 ± 0.91	241.31 ± 0.47	228.84 ± 0.54	131.47 ± 1.23	342.5 ± 1.16	217.07 ± 0.48	117.38 ± 0.59	340.43 ± 0.7	261.53 ± 0.25	228.07 ± 0.63	296.81 ± 0.81
Node 6	33.52 ± 0.08	17 ± 0.09	62.94 ± 0.23	49.45 ± 0.29	14.56 ± 0.26	136.13 ± 2.01	68.03 ± 0.19	20.87 ± 0.11	154.6 ± 1.03	79.5 ± 0.12	64.14 ± 0.33	96.18 ± 0.19
Node 7	106.93 ± 0.26	50.58 ± 0.32	163.49 ± 0.57	229.2 ± 0.4	89.1 ± 1.19	363.71 ± 0.61	58.78 ± 0.23	5.89 ± 0.09	227.98 ± 1.48	23.92 ± 0.05	16.19 ± 0.09	33.33 ± 0.12
Node 8	341.16 ± 0.44	269.22 ± 1.17	400.77 ± 0.65	393.95 ± 0.41	308.34 ± 2.76	467.72 ± 0.69	288.62 ± 0.43	151.02 ± 0.44	432.05 ± 0.94	370.92 ± 0.32	326.02 ± 0.9	415.83 ± 0.54
Node 9	6.62 ± 0.03	2.18 ± 0.01	17.52 ± 0.24	28.11 ± 0.1	2.58 ± 0.04	101.53 ± 0.82	20.44 ± 0.08	2.62 ± 0.01	78.98 ± 0.49	10.19 ± 0.02	5.94 ± 0.04	15.67 ± 0.05
Node 10	133.09 ± 0.4	78.34 ± 0.54	198.48 ± 0.95	194.26 ± 0.51	95.87 ± 1.16	296.93 ± 1.72	134.73 ± 0.19	56.3 ± 0.28	259.56 ± 1.34	179.22 ± 0.22	151.35 ± 0.18	209.38 ± 0.45
Node 11	39.34 ± 0.22	12.51 ± 0.11	85.11 ± 0.84	104.96 ± 0.42	22.42 ± 0.23	211.37 ± 0.74	37.13 ± 0.18	5.63 ± 0.04	123.09 ± 1.47	21.73 ± 0.05	14.56 ± 0.12	30.5 ± 0.09
Node 12	396.66 ± 0.48	322.37 ± 0.92	456.92 ± 0.72	439.75 ± 0.37	364.24 ± 0.9	501.79 ± 1.03	401.66 ± 0.36	259.23 ± 0.98	511.07 ± 0.52	461.89 ± 0.24	423.96 ± 0.63	499.61 ± 0.69
Node 13	488.25 ± 0.38	430.82 ± 1.2	530.55 ± 0.43	513.41 ± 0.2	462.28 ± 0.67	549.54 ± 0.39	497.17 ± 0.29	357.08 ± 2.58	557.47 ± 0.25	525.92 ± 0.29	491.6 ± 1	556.02 ± 0.64
Node 14	407.92 ± 0.28	319.79 ± 0.82	473.66 ± 0.42	472.92 ± 0.36	398.73 ± 1.57	524.27 ± 0.51	230.91 ± 0.26	55.79 ± 0.58	513.3 ± 0.57	297.57 ± 0.59	243.33 ± 1.27	357.99 ± 1.28
Node 15	176.4 ± 0.22	147.25 ± 0.54	201.2 ± 0.33	206.91 ± 0.1	189.32 ± 0.33	221.37 ± 0.21	132.6 ± 0.19	57.69 ± 0.31	207.57 ± 0.69	181.57 ± 0.1	160.65 ± 0.25	203.5 ± 0.1
Node 16	234.54 ± 0.02	221.56 ± 0.11	242.4 ± 0.09	237.93 ± 0.01	229.58 ± 0.07	245.4 ± 0.07	245.67 ± 0.03	235.63 ± 0.05	264.89 ± 0.09	242.91 ± 0.05	234.79 ± 0.06	257.77 ± 0.3
Node 17	116.35 ± 0.22	77.91 ± 0.71	148.98 ± 0.41	197.4 ± 0.07	175.73 ± 0.47	214.29 ± 0.14	70.05 ± 0.47	10.68 ± 0.12	193.48 ± 1	38.05 ± 0.07	28.44 ± 0.15	49.02 ± 0.07
Node 18	321.16 ± 0.07	306.45 ± 0.05	342.57 ± 0.33	309.34 ± 0.02	301.09 ± 0.05	321.19 ± 0.09	300.61 ± 0.04	281.43 ± 0.17	311.21 ± 0.09	342.91 ± 0.17	317.5 ± 0.35	370.81 ± 0.34
Node 19	189.65 ± 0.19	153.47 ± 0.2	232.72 ± 0.36	202.71 ± 0.15	157.44 ± 0.14	243 ± 0.08	182.85 ± 0.57	149.96 ± 0.04	238.39 ± 0.24	121.73 ± 0.21	97.41 ± 0.42	147.62 ± 0.3
Node 20	513.67 ± 0.02	507.98 ± 0.02	525.62 ± 0.12	515.92 ± 0.03	508.59 ± 0.02	531.63 ± 0.1	522.38 ± 0.02	509.45 ± 0.02	541.85 ± 0.06	513.33 ± 0.03	507.86 ± 0.05	524.83 ± 0.21
Node 21	286.34 ± 0.61	191.57 ± 0.83	369.6 ± 0.46	368.57 ± 0.4	267.5 ± 2.08	436.05 ± 0.44	167.1 ± 0.37	33.52 ± 0.21	454.58 ± 1.42	169.78 ± 0.21	135.25 ± 0.4	207.82 ± 0.36
Node 22	493.13 ± 0.59	411.15 ± 1.48	577.9 ± 1.11	565.21 ± 0.3	524.75 ± 0.79	597.86 ± 0.19	340.09 ± 1.09	60.72 ± 0.26	591.32 ± 0.75	345.21 ± 0.73	283.31 ± 0.73	415.76 ± 1.81
Node 23	573.95 ± 0.04	555.4 ± 0.12	584.03 ± 0.05	571.26 ± 0.04	550.41 ± 0.1	583.26 ± 0.07	561.55 ± 0.09	531.87 ± 0.17	571.74 ± 9.91	566.31 ± 0.15	541.57 ± 0.69	582.05 ± 0.03
Node 24	590.58 ± 0.22	564.84 ± 0.13	627.85 ± 0.55	581.89 ± 0.1	557.32 ± 0.15	606.55 ± 0.24	583.61 ± 0.2	543.38 ± 0.27	644.19 ± 0.93	569.42 ± 0.15	552.42 ± 8.78	587.54 ± 0.17

Table S2. Optimal (± SEM) divergence dates under the CIR, LogNormal, Ugam and strict molecular clock models with a soft-bounds of 2.5%, 10%, 20% and 50%. Nodes are as indicated in Figure 1.

	CIR 2.5% SB		CIR 10% SB			CIR 20% SB			CIR 50% SB			LogNormal 2.5% SB			LogNormal 10%			LogNormal 20% SB			LogNormal 50% SB		
	Optimal	SEM	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N
Node 1	24.75967	0.1321312	23.51493	0.0944436	10	20.72206	0.1045801	10	16.17815	0.1220712	10	57.38704	0.4046616	10	46.90549	0.4706227	10	34.14125	0.1608546	10	17.75457	0.0865853	10
Node 2	22.9384	0.0971284	21.98127	0.066423	10	19.6972	0.0854785	10	15.80917	0.1001257	10	46.28854	0.2366066	10	38.8354	0.2678433	10	29.46513	0.1147407	10	16.74204	0.0869885	10
Node 3	70.5143	0.330877	68.45871	0.1910048	10	62.1961	0.2360142	10	50.9398	0.3087822	10	121.2087	0.5086319	10	106.7912	0.5185544	10	86.48285	0.3278216	10	52.68539	0.2247195	10
Node 4	120.3744	0.3786204	119.1469	0.1830376	10	110.1139	0.3437658	10	92.37581	0.4758354	10	185.4433	0.4731953	10	167.5522	0.4296809	10	142.6046	0.3586656	10	92.63531	0.2994213	10
Node 5	179.9936	0.3265133	182.8135	0.2334708	10	174.896	0.2238799	10	157.2556	0.3773119	10	228.8368	0.5394631	10	212.0317	0.6000964	10	189.4045	0.4255522	10	137.9408	0.3303336	10
Node 6	33.52287	0.0789786	32.60791	0.1450706	10	29.72862	0.1471837	10	25.53014	0.08434056	10	49.45164	0.2941594	10	40.84907	0.5077615	10	31.15005	0.2161757	10	19.65452	0.1189064	10
Node 7	106.9274	0.2551931	113.1151	0.2038767	10	113.3817	0.2290754	10	97.0325	0.2825607	10	229.2045	0.4022935	10	208.7683	0.5864612	10	181.384	0.4400688	10	116.8332	0.4391538	10
Node 8	341.1577	0.4446785	350.4314	0.330834	10	347.7697	0.331509	10	325.3239	0.5685447	10	393.9453	0.4072427	10	384.6502	0.382747	10	375.394	0.486519	10	325.8013	0.3431739	10
Node 9	6.61604	0.0345326	6.37322	0.0257674	10	5.7312	0.0196297	10	4.4717	0.02640986	10	28.1067	0.0961446	10	20.00601	0.1775791	10	11.80406	0.1308604	10	4.80075	0.0146467	10
Node 10	133.0904	0.395656	136.2507	0.2880948	10	130.0417	0.2741987	10	109.4711	0.4261563	10	194.2564	0.5100425	10	177.1693	0.4935155	10	153.7493	0.4844393	10	105.8253	0.2014708	10
Node 11	39.34217	0.2230935	38.36933	0.1354717	10	33.43812	0.1410621	10	24.67117	0.2488082	10	104.9581	0.4197157	10	86.58816	0.4567213	10	62.64246	0.2567295	10	30.20839	0.1504386	10
Node 12	396.6571	0.4815949	406.1814	0.3037213	10	408.0108	0.2962329	10	390.9991	0.5168335	10	439.7458	0.3672037	10	436.0875	3.286722	10	427.169	0.4980598	10	391.341	0.5279142	10
Node 13	488.2538	0.3844444	495.0168	0.2846422	10	498.2892	0.4310772	10	489.9158	0.4596291	10	513.4115	0.203396	10	510.1763	0.3051708	10	507.785	0.1985136	10	494.1315	0.1317465	10
Node 14	407.9158	0.2814049	418.1537	0.3913164	10	422.9371	0.3898842	10	410.8454	0.5441319	10	472.9173	0.3636469	10	467.1256	0.3244395	10	461.6519	0.249023	10	441.9973	0.38774	10
Node 15	176.396	0.2240007	165.2205	0.1796325	10	132.0073	0.2318886	10	80.08774	0.2550828	10	206.9064	0.0962463	10	194.0586	0.1362754	10	161.1596	0.2232609	10	81.05422	0.2767664	10
Node 16	234.5385	0.016603	225.8239	0.0809676	10	192.1818	0.3014079	10	127.9145	0.3577609	10	237.9297	0.0126364	10	231.0797	0.0262771	10	208.8978	0.1826286	10	129.2393	0.2601351	10
Node 17	116.3491	0.223011	105.5807	0.2032389	10	76.50692	0.2305956	10	38.89162	0.123895	10	197.4001	0.0675944	10	180.848	0.1457299	10	139.874	0.2334542	10	52.41564	0.1107135	10
Node 18	321.1582	0.0696994	360.4944	0.1852291	10	386.9641	0.4342252	10	379.1789	0.4453482	10	309.3393	0.0201462	10	327.3092	0.0569564	10	361.1472	0.2150786	10	368.3022	0.2499139	10
Node 19	189.6458	0.1853402	199.1442	0.2268142	10	213.4866	0.2324767	10	214.1302	0.3133514	10	202.709	0.1488398	10	205.5071	0.1776241	10	217.4886	0.2131669	10	251.8445	0.4974794	10
Node 20	513.6747	0.0180868	513.0989	0.0236153	10	510.5628	0.0429646	10	491.8071	0.1548471	10	515.9247	0.0269245	10	515.0612	0.0306981	10	513.0558	0.02563	10	497.8339	0.0710739	10
Node 21	286.3367	0.6138791	309.6031	0.4827273	10	325.7171	0.4077928	10	312.8536	0.2916259	10	368.5732	0.3967083	10	376.0866	0.5437716	10	388.695	0.3452219	10	366.1755	0.2182483	10
Node 22	493.13	0.5911129	503.3911	0.306663	10	513.2887	0.9597232	10	507.7074	0.9410246	10	565.2134	0.2968729	10	567.2006	0.2967394	10	568.6583	0.263332	10	576.8464	0.224151	10
Node 23	573.9534	0.0379933	575.7788	0.0355846	10	580.1783	0.0657855	10	587.0161	0.1950709	10	571.264	0.0427972	10	572.6506	0.0444993	10	575.8951	0.0835817	10	581.0582	0.1251338	10
Node 24	590.5829	0.2179079	593.6435	0.1020852	10	599.3414	0.2629502	10	606.2703	0.2000059	10	581.8855	0.0970129	10	584.3661	0.1180181	10	588.3766	0.0992278	10	596.1808	0.1031652	10

Table S2. (Continued.)

	Ugam 2.5% SB			Ugam 10% SB			Ugam 20% SB			Ugam 0.5			Strict clock 2.5% SB			Strict clock 10% SB			Strict clock 20% SB			Strict clock 50% SB		
	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N	Optimal	SEM	N
Node 1	21.71916	0.0984816	10	21.96739	0.0924649	10	21.19177	0.0740548	10	18.05099	0.0903613	10	9.71245	0.0249979	10	9.851171	0.036335	10	9.99245	0.0226694	10	9.73973	0.0293453	10
Node 2	23.31285	0.0865824	10	23.36755	0.1377799	10	22.85626	0.1067508	10	19.43122	0.0680276	10	12.37494	0.0269275	10	12.55602	0.0349552	10	12.69341	0.0297924	10	12.42017	0.0276546	10
Node 3	69.9336	0.197452	10	70.27481	0.1497604	10	68.95229	0.173731	10	59.16384	0.1297308	10	39.3794	0.089316	10	39.91124	0.0705106	10	40.20962	0.0819722	10	39.49829	0.062777	10
Node 4	133.8914	0.538487	10	134.3497	0.3950786	10	132.9121	0.3744653	10	120.6385	0.4261728	10	78.95575	0.1244208	10	79.14547	0.1398367	10	79.31441	0.1660382	10	78.32713	0.1128573	10
Node 5	217.071	0.4755656	10	216.9524	0.2530373	10	217.8235	0.4451243	10	219.8443	0.2364236	10	261.5279	0.2545951	10	264.9261	0.3098789	10	268.3219	0.2695636	10	267.9682	0.2457079	10
Node 6	68.0255	0.1935029	10	68.11148	0.1502461	10	68.01244	0.2557423	10	64.30667	0.165275	10	79.49665	0.1247948	10	79.74827	0.1187414	10	79.72112	0.1383889	10	79.24672	0.1102159	10
Node 7	58.7785	0.2290793	10	58.15228	0.2386354	10	56.48364	0.3385531	10	46.98497	0.2688842	10	23.91691	0.0458551	10	24.24843	0.0459431	10	24.60632	0.0598957	10	24.08572	0.0432174	10
Node 8	288.6174	0.4254768	10	289.8787	0.4177839	10	295.14	0.2756509	10	310.9599	0.1438377	10	370.9247	0.3164521	10	377.0679	0.3214285	10	383.2519	0.2830559	10	384.0379	0.515889	10
Node 9	20.4414	0.084181	10	20.48436	0.0816094	10	20.11364	0.0741528	10	17.64057	0.0454209	10	10.18522	0.0238222	10	10.33453	0.0234929	10	10.46632	0.0240432	10	10.24882	0.0321171	10
Node 10	134.7288	0.1857278	10	134.3615	0.1964139	10	135.0382	0.2694962	10	135.2214	0.2605728	10	179.222	0.2163857	10	182.04	0.2287885	10	184.5699	0.1877803	10	183.526	0.1774762	10
Node 11	37.13399	0.1766024	10	36.98261	0.0977096	10	36.6682	0.0885848	10	32.85327	0.1191655	10	21.73241	0.0534991	10	22.08543	0.0387139	10	22.39812	0.0472457	10	21.96412	0.0385043	10
Node 12	401.6624	0.3590557	10	402.9314	0.4285959	10	407.2379	0.4482883	10	419.0501	0.242783	10	461.8867	0.2380582	10	467.2955	0.2978247	10	474.646	0.3703938	10	477.8402	0.169492	10
Node 13	497.168	0.291916	10	498.6058	0.2183122	10	500.8215	0.2129093	10	506.5464	0.17769	10	525.9208	0.2941027	10	531.045	0.188214	10	537.3752	0.4294658	10	541.0597	0.3066175	10
Node 14	230.908	0.2635597	10	233.8181	0.1246097	10	240.4493	0.4406007	10	269.9951	0.8159245	10	297.5721	0.5913624	10	301.3379	0.3430332	10	304.5659	0.6521948	10	302.4501	0.320749	10
Node 15	132.6039	0.1927061	10	131.5381	0.3501824	10	129.5229	0.1810467	10	123.6796	0.312257	10	181.5686	0.0987625	10	193.1029	0.1992732	10	204.0221	0.2547747	10	203.5509	0.216776	10
Node 16	245.6737	0.0285419	10	244.2537	0.0332276	10	240.5441	0.0469767	10	229.8431	0.1255689	10	242.9146	0.0512683	10	255.1765	0.2295352	10	269.2206	0.3184294	10	269.0388	0.1737491	10
Node 17	70.05267	0.4683768	10	69.59967	0.2380229	10	68.31575	0.2252646	10	58.12749	0.2547435	10	38.05266	0.0657147	10	39.11897	0.0697992	10	39.8543	0.064879	10	39.05317	0.0529587	10
Node 18	300.6087	0.036334	10	302.9766	0.0300638	10	311.0469	0.0503845	10	357.7168	0.2133763	10	342.9087	0.16559	10	405.9578	0.3809994	10	464.4753	0.5748757	10	475.6492	0.3413088	10
Node 19	182.8463	0.5745046	10	180.2377	0.0683616	10	174.6365	0.0865895	10	127.5313	0.2222288	10	121.7269	0.2079039	10	93.00298	0.0287196	10	82.318	0.0940543	10	69.41048	0.1129977	10
Node 20	522.3752	0.0156127	10	521.962	0.0160392	10	520.5636	0.0329077	10	510.3498	0.0856047	10	513.332	0.0268769	10	512.0308	0.1553287	10	507.3796	0.0810384	10	491.6036	0.1906701	10
Node 21	167.0952	0.3749342	10	167.5302	0.4457262	10	168.7915	0.5248513	10	167.4749	0.5160646	10	169.7753	0.2079041	10	171.839	0.2836632	10	174.4163	0.351848	10	170.8884	0.2901567	10
Node 22	340.0919	1.089712	10	342.0514	0.8211871	10	351.4377	1.411147	10	387.7429	0.8972987	10	345.2148	0.7279733	10	353.3231	0.5617071	10	358.4642	0.747986	10	355.3023	0.6627004	10
Node 23	561.5519	0.0900824	10	562.6782	0.3298116	10	564.3087	0.056663	10	568.7781	0.1005505	10	566.3132	0.1469217	10	571.0396	0.1138266	10	577.2527	0.1660698	10	581.3114	0.1519695	10
Node 24	583.6077	0.2006502	10	584.6611	0.1629618	10	586.1503	0.1232267	10	589.0835	0.1541675	10	569.4232	0.1459314	10	574.326	0.1087797	10	580.5598	0.1461204	10	584.6406	0.1437806	10

Table S3. Minimum, maximum and optimal divergence dates ± SE of the CIR model with the GTR-CAT, GTR and CAT nucleotide substitution models. Nodes are as indicated in Figure 1.

	CIR GTR-CAT			CIR GTR			CIR CAT		
	Optimal	Min	Max	Optimal	Min	Max	Optimal	Min	Max
Node 1	24.63 ± 0.18	7.28 ± 0.06	56.65 ± 0.68	23.64 ± 0.07	7.5 ± 0.07	53.74 ± 0.21	25.38 ± 0.21	7.54 ± 0.03	58.57 ± 0.19
Node 2	22.85 ± 0.14	7.81 ± 0.1	49.96 ± 0.2	22.42 ± 0.09	8.09 ± 0.07	48.63 ± 0.26	23.87 ± 0.07	7.87 ± 0.28	51.53 ± 0.16
Node 3	70.24 ± 0.46	31.97 ± 0.21	122.17 ± 0.94	67.15 ± 0.14	31.56 ± 0.16	116.28 ± 0.31	72.54 ± 0.25	32.84 ± 0.14	124.82 ± 0.28
Node 4	120.04 ± 0.5	68.72 ± 0.26	178.64 ± 0.73	117.58 ± 0.1	67.01 ± 0.31	174.87 ± 0.37	116.16 ± 0.35	65.33 ± 0.28	174.13 ± 0.67
Node 5	180 ± 0.55	123.52 ± 0.91	241.31 ± 0.47	176.14 ± 0.28	117.21 ± 0.5	236.73 ± 0.31	172.56 ± 0.41	117.84 ± 0.36	233.65 ± 0.42
Node 6	33.66 ± 0.11	17 ± 0.09	62.94 ± 0.23	44.09 ± 0.14	22.29 ± 0.11	79.75 ± 0.43	31.81 ± 0.13	16.16 ± 0.1	59.43 ± 0.35
Node 7	106.98 ± 0.44	50.58 ± 0.32	163.49 ± 0.57	108.53 ± 0.15	50.77 ± 0.42	169.35 ± 0.56	110.13 ± 0.09	53.49 ± 0.1	167.58 ± 0.14
Node 8	340.51 ± 0.73	269.22 ± 1.17	400.77 ± 0.65	359.15 ± 0.3	309.78 ± 28.32	423.33 ± 0.2	336.56 ± 0.3	262.43 ± 0.33	409.97 ± 11.49
Node 9	6.62 ± 0.05	2.18 ± 0.01	17.52 ± 0.24	12.26 ± 0.08	4.07 ± 0.04	31.34 ± 0.39	7.06 ± 0.02	5.65 ± 3.37	19.14 ± 0.08
Node 10	132.41 ± 0.66	78.34 ± 0.54	198.48 ± 0.95	196.84 ± 0.21	130.95 ± 0.49	262.04 ± 2.06	124.08 ± 0.3	71.36 ± 0.27	190.56 ± 0.36
Node 11	38.96 ± 0.36	12.51 ± 0.11	85.11 ± 0.84	64.55 ± 0.05	24.39 ± 0.19	120.16 ± 0.32	33.99 ± 0.14	10.67 ± 0.07	77.69 ± 0.36
Node 12	396.42 ± 0.63	322.37 ± 0.92	456.92 ± 0.72	317.86 ± 0.44	241.95 ± 0.67	383.85 ± 0.22	394.39 ± 0.37	317.7 ± 0.72	456.47 ± 0.31
Node 13	487.72 ± 0.6	430.82 ± 1.2	530.55 ± 0.43	497.3 ± 0.29	437.69 ± 0.6	537.8 ± 0.27	483.73 ± 0.27	421.38 ± 0.54	529.95 ± 0.26
Node 14	407.87 ± 0.28	319.79 ± 0.82	473.66 ± 0.42	413.26 ± 0.62	311.49 ± 1.63	480.68 ± 0.49	409.51 ± 0.41	321.17 ± 0.87	475.52 ± 0.2
Node 15	176.57 ± 0.35	147.25 ± 0.54	201.2 ± 0.33	187.71 ± 0.23	161.86 ± 0.25	208.07 ± 0.09	177.52 ± 0.13	147.91 ± 0.15	202.19 ± 0.13
Node 16	234.52 ± 0.02	221.56 ± 0.11	242.4 ± 0.09	235.48 ± 0.04	223.82 ± 0.09	243 ± 0.05	234.16 ± 0.01	220.53 ± 0.06	242.22 ± 0.02
Node 17	116.34 ± 0.3	77.91 ± 0.71	148.98 ± 0.41	123.39 ± 0.25	86.23 ± 0.11	154.06 ± 0.34	120.73 ± 0.11	81.61 ± 0.21	152.66 ± 0.07
Node 18	321.15 ± 0.09	306.45 ± 0.05	342.57 ± 0.33	320.29 ± 0.1	306.37 ± 0.03	340.77 ± 0.17	321.1 ± 0.01	306.41 ± 0.01	342.46 ± 0.03
Node 19	189.75 ± 0.35	153.47 ± 0.2	232.72 ± 0.36	187.6 ± 0.22	152.56 ± 0.04	229.73 ± 0.39	194.87 ± 0.03	155.5 ± 0.04	237.6 ± 0.07
Node 20	513.67 ± 0.04	507.98 ± 0.02	525.62 ± 0.12	513.37 ± 0.02	507.95 ± 0.04	524.68 ± 0.1	513.97 ± 0.01	508.08 ± 0	526.55 ± 0.02
Node 21	286.31 ± 0.73	191.57 ± 0.83	369.6 ± 0.46	282.17 ± 0.55	193.94 ± 0.37	364.01 ± 0.55	275.33 ± 0.19	184.66 ± 0.11	360.9 ± 0.24
Node 22	492.68 ± 1.12	411.15 ± 1.48	577.9 ± 1.11	502.67 ± 0.34	422.1 ± 0.45	582.83 ± 0.72	490.95 ± 0.45	408.5 ± 0.35	575.99 ± 1.12
Node 23	574 ± 0.04	555.4 ± 0.12	584.03 ± 0.05	574.1 ± 0.06	555.93 ± 0.16	584.17 ± 0.09	573.29 ± 0.02	553.38 ± 0.07	583.9 ± 0.01
Node 24	590.63 ± 0.14	564.84 ± 0.13	627.85 ± 0.55	591.21 ± 0.2	565.18 ± 0.21	628.62 ± 0.29	589.74 ± 0.16	562.86 ± 0.11	613.77 ± 12.71

Phylogenetic divergence times using relaxed molecular clocks

Vol. 19(8), 2017

S14

Downloaded from Brill.com09/18/2019 03:16:11PM
via Maynooth University

Table S4. Optimal divergence dates ± SEM of the CIR model with a soft-bound of 2.5% and a SD of 50 and 20. Nodes are as indicated in Figure 1.

	CIR 2.5% SD 50			CIR 2.5% SD 20		
	Optimal	SEM	*N*	Optimal	SEM	*N*
Node 1	24.75967	0.1321312	10	24.72956	0.1550461	5
Node 2	22.9384	0.0971284	10	22.85796	0.0898685	5
Node 3	70.5143	0.330877	10	70.25706	0.3607412	5
Node 4	120.3744	0.3786204	10	120.3562	0.4740049	5
Node 5	179.9936	0.3265133	10	179.9369	0.4681844	5
Node 6	33.52287	0.0789786	10	33.59406	0.1303464	5
Node 7	106.9274	0.2551931	10	106.2985	0.1627454	5
Node 8	341.1577	0.4446785	10	339.7329	0.4945394	5
Node 9	6.61604	0.0345326	10	6.66732	0.0298782	5
Node 10	133.0904	0.395656	10	132.7643	0.4853723	5
Node 11	39.34217	0.2230935	10	39.43874	0.170724	5
Node 12	396.6571	0.4815949	10	395.3095	0.4893206	5
Node 13	488.2538	0.3844444	10	486.2714	0.4770738	5
Node 14	407.9158	0.2814049	10	406.1426	0.2275182	5
Node 15	176.396	0.2240007	10	176.9531	0.1891135	5
Node 16	234.5385	0.016603	10	234.5495	0.0260538	5
Node 17	116.3491	0.223011	10	116.6393	0.3407698	5
Node 18	321.1582	0.0696994	10	320.8926	0.1477149	5
Node 19	189.6458	0.1853402	10	188.9605	0.1164389	5
Node 20	513.6747	0.0180868	10	513.5178	0.0239012	5
Node 21	286.3367	0.6138791	10	286.7066	0.6535592	5
Node 22	493.13	0.5911129	10	487.2673	0.7695808	5
Node 23	573.9534	0.0379933	10	571.1469	0.0449154	5
Node 24	590.5829	0.2179079	10	580.62	0.0879979	5