JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12068

SHORT COMMUNICATION

You are where you live: parasitic nematode mitochondrial genome size is associated with the thermal environment generated by hosts

M. LAGISZ*†, R. POULIN* & S. NAKAGAWA*†

*Department of Zoology, University of Otago, Dunedin, New Zealand
†National Research Centre for Growth & Development, University of Otago, Dunedin, New Zealand

Keywords:

body size;
longevity;
metabolism;
mtDNA;
temperature.

Abstract

There exists remarkable interspecific variation in mitochondrial sequence evolution rates and in mitochondrial genome sizes. A number of hypotheses based on the forces of mutation and selection have been proposed to explain this variation. Among such hypotheses, we test three: 1) the 'longevity-dependent selection', 2) the 'functional constraints' and 3) the 'race for replication' hypotheses, using published mtDNA genomic sequences of 47 Nematoda species. We did not find any relationship between body size (used as a proxy for longevity) and genome size or the substitution rate of protein sequences, providing little evidence for the first hypothesis. Parasitic species from different thermal habitats, as determined by their definitive host type (ectothermal vs. endothermal), did not differ in their rates of protein evolution. Therefore, little support was obtained for the second hypothesis. However, we revealed that mitogenomes of parasites of endotherms were significantly smaller than those of parasites of ectotherms, supporting the race for replication hypothesis. As mitochondrial genomes of endothermal animals are usually more compact than those of ectothermal animals, intriguingly, nematode parasites of endotherms and ectotherms exhibit similar patterns of mtDNA length variation to their hosts.

Introduction

Mitochondria form a small but essential component of every animal cell by generating the majority of its energy. The animal mitochondrial genome (mitogenome) is typically quite compact (16–17 Kb) and contains around 40 genes, including 13 genes coding elements of the respiratory chain (Boore, 1999). The rapid rate of changes observed at the mitochondrial DNA (mtDNA) sequence level has made the mtDNA a popular marker for systematics and population genetics (Galtier *et al.*, 2009). The evolution of mtDNA was initially assumed to be neutral and constant in time and across lineages. Mounting evidence indicates, however, that such a 'molecular clock' is not universal (Bromham & Penny, 2003); multiple studies have now demonstrated variation in the evolutionary rates of mtDNA (e.g. Rottenberg, 2007; Nabholz *et al.*, 2009). Also, the physical size of mitochondrial genomes varies greatly in some animal taxa (Rand, 1993). Two main lines of reasoning attempt to explain variation in the rates of evolution and mitogenome sizes: one relates to longevity and the other to metabolic rate and the thermal environment.

The first major hypothesis concerning the variation in the evolutionary rate of mtDNA, the 'longevity-dependent selection' hypothesis, is linked to the role of mitochondria in the ageing process. The association between accumulation of mtDNA mutations and ageing in many species leads to the idea that selection acting to reduce the mutation rate may increase the longevity of a species (Samuels *et al.*, 2004; Nabholz *et al.*, 2008; Galtier *et al.*, 2009). As a consequence, we might expect

Correspondence: Malgorzata Lagisz, Department of Zoology, University of Otago, 340 Great King Street, PO Box 56, Dunedin 9054, New Zealand. Tel.: +64 3 479 5046; fax: +64 3 479 7584; e-mail: losialagisz@yahoo.com

low substitution rates in mtDNA of long-lived animals, as observed among mammals (Welch *et al.*, 2008). Certain properties of the genomes can make mutations more likely – for example, indels can act as 'mutators', especially in noncoding regions (Tian *et al.*, 2008). In accordance with this reasoning, the copy numbers of repeats present in noncoding mtDNA are negatively associated with longevity in mammals (Samuels *et al.*, 2004; Khaidakov *et al.*, 2006). At the whole-genome level, if selection against repeats and indels is operating, more compact mitogenomes are likely to be a property of longer lived animals (Monaghan & Metcalfe, 2000; Griffith *et al.*, 2003). A clear prediction from the longevity-selection hypothesis is, thus, that longer life is linked with lower substitution rates and smaller mitogenome sizes.

The second line of thought relates to metabolic rate and the thermal environment. The original metabolic rate hypothesis argues that higher metabolic rate increases DNA mutation rate via generation of reactive oxygen species (Martin & Palumbi, 1993). Endothermal animals not only have higher metabolic rates than ectotherms, but their body temperature is less affected by exogenous factors, which leads us to the 'functional constraints' hypothesis. The variation in the thermal environment may restrict physiologically acceptable amino acid substitutions, reducing the rate of protein sequence evolution (Rand, 1994; Thomas & Beckenbach, 1989). Thus, the 'functional constraints' hypothesis predicts that amino acid sequences will be more conserved in ectotherms than in endotherms. Indeed, higher protein evolutionary rates have been observed in birds and mammals than in amphibians and fishes (Adachi *et al.*, 1993). However, there is still very limited evidence on how the rate of molecular evolution is affected by environmental temperatures among different ectotherms (Bromham & Penny, 2003; Estabrook *et al.*, 2007; McGaughran & Holland, 2010).

Temperature and metabolic rate are also expected to influence the physical size of mitochondrial genomes, as implied by the 'race for replication' hypothesis. This hypothesis predicts that increased metabolic demands in species with faster metabolism or living at higher temperatures will select for higher replication efficiency offered by smaller mtDNA molecules (Rand, 1994; Selosse *et al.*, 2001). A study by Rand (1993) found that endotherms tend to have smaller and less variable mitogenomes than ectotherms, although empirical work on this particular hypothesis seems to be extremely limited.

The phylum Nematoda offers a unique opportunity to test the aforementioned hypotheses. There are currently more than 40 complete mitochondrial genomic sequences available in this taxonomic group. These sequences belong to species that are mostly endoparasitic and their definitive hosts include both endotherms and ectotherms; also few free-living or plant-parasitic species have mitogenome sequences available. Therefore, while all nematodes are themselves ectothermic, their predominant thermal habitats can be drastically different both in terms of average values and stability. Furthermore, Nematoda species can differ dramatically in their body size (Blaxter, 2011) and lifespan (Gems, 2000), which are found to be closely linked in other taxa (Speakman, 2005). Given that larger parasites are more likely to live in larger hosts (which are usually more long-lived themselves and provide more nutrients allowing more time and energy for growth), it seems reasonable to assume that adult body size might reflect nematode lifespan (Poulin, 2007). Also, because definitive hosts often provide an environment for most of the individual's growth and for reproduction, adaptation to the definitive host type could potentially shape parasite's mitogenome.

In this article, using mitochondrial genomes of 47 Nematoda species and a Bayesian phylogenetic mixed model, we test two specific questions: 1) whether the variation observed in the mtDNA length and evolutionary rate of mitochondrial proteins is related to species body size (used as a proxy for longevity)? and 2) whether this variation can be linked to thermal habitat as determined by endothermal vs. ectothermal definitive hosts of parasitic species?

Materials and methods

We downloaded mtDNA genomic sequences for all Nematoda species ($n = 47$ and two other Pseudocoelomata species used as an out-group) available online from GenBank before April 2011. Lengths of the total mitochondrial genome, coding and noncoding sequences were obtained from the GenBank database. We ran RepeatMasker Open (v.3.3.0; Smit *et al.*, 1996–2010), at default settings, on the nucleotide sequences of whole genomes to find lengths of sequences occupied by repetitive elements of various types. From the GeneBank files, we extracted translated protein-coding sequences (CDS) with the online tool ReadSeq (v.2.1; Gilbert, 2003). We used amino acid sequences in our analyses of substitution rates because the third codon position experienced substitution saturation.

We aligned amino acid sequences for each gene separately using ClustalW (Thompson *et al.*, 1994) and manually checked end edited gaps, where necessary, in BioEdit (Hall, 1999). Then, we concatenated alignments for all genes. We observed that ND3 is triplicated in *Romanomermis culicivorax* (all copies being identical) and duplicated in *R. iyengari* (differing by one synonymous substitution). *Hexamermis agrotis* has 3 identical copies of ATP6. Subsequently, only one copy of each duplicated/triplicated gene was included in alignments. *Trichinella spiralis* has one additional gene – ATP8, which was not used in the analyses.

To improve reliability, from the concatenated amino acid alignment we extracted conserved (ungapped) blocks of sequence by using Gblocks server with the less stringent settings (Castresana, 2000; Talavera & Castresana, 2007). This resulted in a protein alignment of 3106 aa (81% of the original aligned 3835 aa positions), which was used for the tree construction and phylogenetic analyses.

Aligned and concatenated protein-coding sequences were used to estimate substitution rates for each species with the software BEAST (v.1.7; Drummond *et al.*, 2012). We tested our sequence dataset for the best-fit substitution model with ProtTest (v.2.4; Abascal *et al.*, 2005) and Aminosan (Tanabe, 2011). Both tests showed that mtArt, mtZoa, and mtRev are the most likely substitution models. We estimated a rooted, ultrametric, time-calibrated phylogeny with BEAST, applying uncorrelated lognormal relaxed-clocks, Yule model as a tree prior and Gamma site heterogeneity model with four site categories and MtRev substitution model (MtArt and mtZoa are not available in BEAST). For the tree dating, estimates of divergence times (Enoplea/Chromadorea – 451 MYA, and Spirurida/Rhabditidae – 250 MYA) were taken after Douzery *et al.* (2004), and used in normal prior distribution as mean values with standard deviation of 50 MY. Two independent Markov chains were run in BEAST for 10 million generations, sampling every 1000th generation. We summarized the chains using Tracer (v.1.4.1; Rambaut & Drummond, 2007) and visually inspected the trace plots (all showed good mixing and convergence). The effective sample sizes (ESS) for the runs were above 100 for most parameters reported in Tracer. We converted the posterior tree distributions into a maximum clade credibility (MCC) tree using TreeAnotator (v.1.4.8; from the BEAST package). Estimated amino acid substitution rates were extracted manually from the terminal branches for each Nematoda species for further analyses.

From the GenBank files, we also retrieved the order and orientation of the protein-coding genes. This information was coded numerically using a custom script written in R (v.2.12.0; R Development Core Team, 2011). The numbers of rearrangements that occurred on each tree edge were recovered with the MGR software (Bourque & Pevzner, 2002), which can work with circular genomes and data constrained on an existing tree topology. From the resulting tree that had branches labelled with the number of rearrangements, we calculated rearrangement distances from the last common ancestor to the tree tips, that is, for each of the nematode species in our dataset.

We collected the following information from the primary literature for the Nematoda species under analysis: thermal environment/definitive host type (coded as: endothermal host, ectothermal host, free-living/plant parasite), average adult body size expressed as volume (mm^3). For body volume calculations, length and body width in mm, for both males and females, were collected from the primary literature for each Nematoda species. An average value for each species was calculated if multiple measurements were obtained from one or a couple of studies. As values for different species range over several orders of magnitude, the average values are perfectly suitable for an interspecific analysis. The body length and width were converted to body volumes, assuming a cylindrical shape for nematodes, using the formula: body volume = length * pi * $(width/2)^2$. Female and male body sizes were highly correlated and therefore only values for females were used in analyses. Female body volume was log-transformed to satisfy normal distribution requirements of linear regression. We note that the free-living/plant-parasitic species were included in the analyses of the effects of thermal habitat type, although we had no clear prediction regarding the expected effects for this group because of its life-style and habitat-related heterogeneity being larger than in the other two groups. Unfortunately, the small number of species with sequenced mitogenomes falling into this group did not allow finer data subdivision.

To investigate whether mitochondrial genome sizes and substitution rates are associated with nematode species body size and thermal environment determined by host type in parasitic species, we used Bayesian Phylogenetic Generalized Linear Mixed-effects Model (PGLMM) implemented in the R statistical package MCMCglmm (v.2.11; Hadfield, 2010; Hadfield & Nakagawa, 2010). This statistical approach accounts for the lack of independence among the data points owing to the evolutionary history of the species, assuming a Brownian motion model of evolution. Phylogeny and species identities were entered into models as random factors, thermal environment type and log-transformed female body volume as fixed factors. The first model was constructed for the log-transformed genome size as a response variable, and the second model for the log-transformed protein substitution rate. We used the inverse Wishart prior ($V = 0.002$, $\upsilon = 1$), equivalent to the inverse Gamma prior. For both models, we ran one chain with 4 000 000 iterations with a burn-in of 3 000 000 and a thinning interval of 1000, resulting in proper mixing and uncorrelated posterior distributions made of 1000 samples per chain per model. Analogous models with interactions between predictors are described in the Supporting Information.

A topological species tree for phylogeneticaly controlled analyses was constructed based on the consensus Nematoda phylogeny from the NCBI taxonomic database (http://www.ncbi.nlm.nih.gov/Taxonomy/). Polytomies were randomly resolved resulting in a binomial ultrametric tree with branch lengths standardized to 1. We used an NCBI taxonomic tree, rather than a tree estimated from the mitochondrial sequences, to

avoid having exactly the same data source for the estimation of the rates of sequence evolution and the phylogeny in the PGLMM model. The NCBI taxonomic trees are usually built using multiple traits, although we acknowledge that on the fine-scale level our NCBI phylogeny is likely to be based on the fragments of mitochondrial sequences. We repeated the PGLMM analyses using a Bayesian tree and a Maximum-Likelihood tree, to confirm that our results are robust (details of the tree construction methods are provided in the Supporting Information and the details of the PGLMM analyses are the same as described above).

We calculated phylogenetic heritability, H^2 *sensu* Lynch (Lynch, 1991), as a measure of the degree of phylogenetic dependence (i.e. phylogenetic signal) of mitogenome length and aa substitution rate. H^2 is equivalent to Pagel's λ (Pagel, 1999). H^2 values close to zero indicate that the character is totally independent from the phylogeny, whereas H^2 values close to one suggest that related species tend to resemble each other more than they resemble species drawn at random from the tree. We interpret parameter estimates for the fixed effects as statistically significant when 95% credible intervals (CI) do not span across zero. As parameter estimates for random effects would never go below zero, we focus our interpretations on effect sizes (Nakagawa & Cuthill, 2007).

Results

We observed considerable variation in mitogenome length (range: 12 625–26 194 bp), gene order (3–25 rearrangement events relative to the last common ancestor), amounts of noncoding sequence (212–12 218 bp) and repeat sequence (88–10 857 bp) among the 47 Nematoda species included in this study. As the above variables were intercorrelated (Fig. S1 in Supporting Information), only mitogenome length was used in the analyses of the effects of thermal habitat and body size. Based on our data set we inferred the following ancestral gene order for Nematoda: ND6, CYTB, ND4L, ND4, ND5, COX2, ND1, -ND2, ATP6, COX3, ND3, COX1. The mean rate of protein evolution estimated using a relaxed molecular clock with a log-normal distribution in BEAST was 0.00283 substitutions per site per million year per lineage. The 95% confidence interval for the rate value ranged from 0.00207 to 0.00364 subs/s/my/l, with a median of 0.00279. We inferred phylogenetic relationships among the mitogenomes of 47 Nematoda species using four different approaches: Bayesian, NCBI taxonomic topology, Maximum Likelihood and phylogenetic network (trees are shown in Fig. 1, Fig. S2, Fig. S3 and Fig. S4, respectively). All methods gave very similar clustering of species and were also congruent with other phylogenetic studies for this taxon (De Ley & Blaxter, 2002; Park *et al.*, 2007). The estimated age of the root of the mitochondrial tree had a mean at 756.6 MYA (SEM = 12.5, median = 745.4, 95% HPD lower = 540.4, 95% HPD lower upper = 1000.2).

Figure 1 visually summarizes our data on thermal habitat type, body size, protein sequence evolution rate and mitogenome lengths in a phylogenetic context. Phylogeneticaly controlled analyses revealed that the parasites of ectothermal hosts have significantly larger mitochondrial genomes than the parasites of endothermal hosts (PGLMM: $b_{[\text{diff. between host ectotherm and host endotherm}]} = 0.289$, 95% CI = 0.184–0.373; Fig. 2a, Table S1). Data for the free-living and plant-parasitic species were pooled together due to small sample sizes within each of these species types ($n = 3$ and 2, respectively). Mean lengths of mitochondrial DNA sequences for this group were significantly smaller than genome sizes of parasites of ectotherms, but not significantly different from the parasites of endotherms (PGLMM: $b_{[\text{diff. between host ectotherm and free-living/plant parasite}]} = 0.207$, 95% CI = 0.093–0.367 and PGLMM: $b_{[\text{diff. between host endotherm and free-living/plant parasite}]} = 0.077$, 95% CI = −0.082 to 0.183), but noticeably more variable (Fig. 2a). The species body size overall accounted for little variation in the mitochondrial DNA length (PGLMM: $b_{[\log(\text{body size})]} = 0.001$, 95% CI = −0.034 to 0.047). In the model with thermal habitat (or host type) and body size as fixed effects, there was a very weak phylogenetic signal in the mitochondrial genome size among species (PGLMM: $H^2 = 0.060$, 95% CI = 0.017–0.517). In other words, when accounting for the thermal habitat and body size, related species can have dissimilar mitogenome sizes.

In contrast to mtDNA length, there were little differences in protein substitution rates among thermal habitats (Fig. 2b) and also no significant association between substitution rates and nematode body size (Table S2). We detected a weak phylogenetic signal in this trait (PGLMM: $H^2 = 0.012$, 95% credible interval, CI = 0.03–0.507). In other words, related species have dissimilar substitution rates.

Discussion

In the current study, we set out to elucidate the mechanisms shaping variation in mitochondrial genome sizes and rates of protein evolution among Nematoda species. We did not find any support either for the longevity-selection hypothesis or for the functional constraints hypothesis. However, we did find that mitogenomes are smaller in parasites of endotherms in comparison to parasites of ectotherms, supporting the race for replication hypothesis. Below, we elaborate on our results in relation to the three hypotheses in turn.

On the basis of the longevity-selection hypothesis, we predicted that longer life would be associated with lower protein substitution rates and smaller mitogenome sizes. Our data did not support these predictions, as neither substitution rate, nor genome size, was

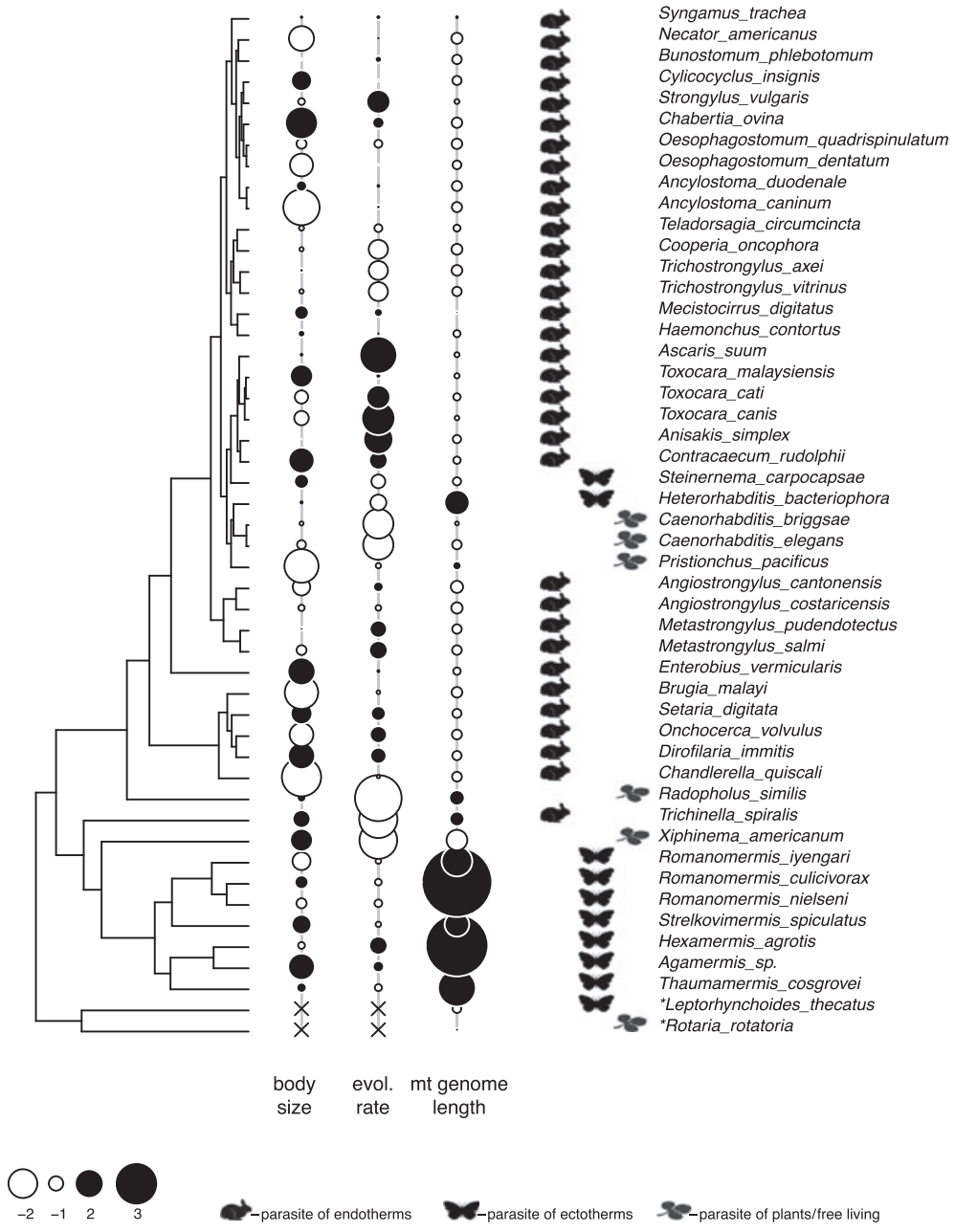

Fig. 1 Bayesian phylogenetic tree showing relationships among 47 Nematoda species. The circles represent log-transformed, centred and standardized to a common scale variables used in our analyses: female body size (mm^3), amino acid substitution rate (subs/site/MY/lineage) and mitochondrial genome length (bp). Positive and negative values are represented using black and white circles, respectively, with a size proportional to the absolute value (i.e. big black circles stand for the largest original values and big white ones for the smallest original values, small dots are intermediate values). Stars indicate the two other Pseudocoelomata species that were used as an outgroup for reconstructing the phylogeny with the BEAST software. Icons represent thermal habitat/host type (see the legend).

related to nematode body size. We used body size as a proxy for longevity because, due to obscure lifestyles, there is very little information available on the life spans of Nematoda (Gems, 2000). Although the size-longevity relationship holds for many animals (Speakman, 2005), it is possible that body size may not accurately reflect nematode longevity due to multiple parasite life-history and host-derived factors (Poulin, 2007). Temperature might be one of the confounding factors here – in ectotherms, colder temperatures not

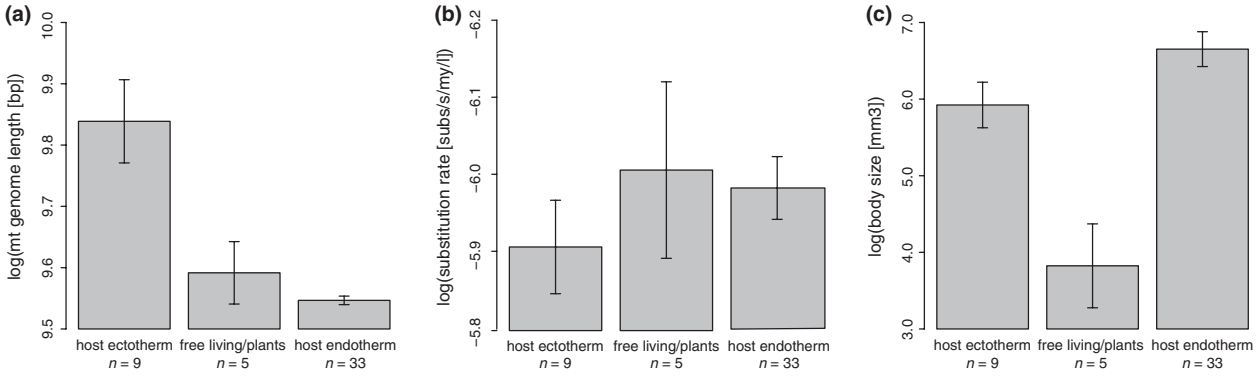

Fig. 2 Bar plots for: (a) the log-transformed mitochondrial genome sizes, (b) the log-transformed amino acid substitution rates of mitochondrial protein-coding genes, and (c) the log-transformed female body volumes of 47 Nematoda species grouped by thermal habitat/definitive host type. The values shown are mean ± SE. The data in this Figure are not controlled for phylogenetic relationships among the species.

only elicit lower metabolic rates but also extend life span relative to the body size (Speakman, 2005).

The functional constraints hypothesis predicts that amino acid sequences will be faster evolving in endotherms than in ectotherms (Rand, 1994; Thomas & Beckenbach, 1989). We did not observe significant differences in evolutionary rates at the protein sequence level between nematodes parasitizing ectotherms and endotherms. The phylogenetic relationship among species does not explain the variation in substitution rates, suggesting that some other environmental and intrinsic factors play a role in shaping this trait.

Smaller mtDNA molecules, according to the race for replication hypothesis, allow for faster generation of new mitochondria, cell proliferation (Blank *et al.*, 2008) and growth of the organism. Selection for smaller mitogenomes could be beneficial if there are no other limitations to growth – energetic, spatial or temporal. Nematode parasites of birds and mammals tend to be larger than those parasitizing arthropods and free-living species (Fig. 2c; Kirchner *et al.*, 1980), suggesting that the growth of this group is least constrained (Calow, 1983) and, therefore, they can benefit most from small mitogenomes.

We found that parasitic Nematoda species spending a significant proportion of their life inside endothermic definitive hosts have more compact mitochondrial genomes than those living inside ectothermal hosts (Figs 1 and 2a). Intriguingly, the architecture of the mitogenomes of the studied parasites reflects that of their hosts. Birds and mammals have compact and stable mitogenomes (Rand, 1993), and, we show, so do the nematode species which inhabit endotherms for most of their active lives. On the other hand, studies on ectothermal invertebrates (e.g. Placozoa (Signorovitch *et al.*, 2007), Bivalvia (Gjetvaj *et al.*, 1992), Insecta (Boyce *et al.*, 1989)) revealed that mtDNA in these taxa can vary greatly in lengths, reaching up to 41 Kb. The Nematoda species included in our study that have ectotherms as definitive hosts, usually parasitize arthropods and their mitogenomes also show substantial variation in size (13–26 Kb).

In addition, the observed variation in the sizes of mitogenomes was not related to the gene content in the studied Nematoda species, which is in agreement with observations on a wide variety of species (Burger *et al.*, 2003). The differences in size of mitochondrial genomes are usually attributed to the variation in the size of intergenic regions, which often harbour duplications and repeat sequences. Likewise, in our data, we observed that mitogenome length is positively correlated with the amount of noncoding DNA, and with the amount of repeat sequence (Fig. S1). A similar pattern was found in other invertebrate taxa (e.g. Boyce *et al.*, 1989; Gjetvaj *et al.*, 1992). Interestingly, we noticed that shorter mitogenomes also have less gene order rearrangements, relative to the ancestral state, which suggests that more compact genomes are more structurally stable.

The results of our study might be limited in some respects, for instance, by other factors potentially involved and not accounted for in our study: 1) the multi-stage lifecycle of parasitic nematodes with different stages often exposed to contrasting conditions, 2) population sizes, 3) mode of reproduction, 4) coevolution with endosymbiotic bacteria, 5) coevolution with nuclear genes, and 6) varying availability of oxygen and varying risks of oxidative damage. The last issue warrants further attention, because different Nematoda species are either 'forced' or 'chose' to rely on anaerobic respiration thus avoiding reactive oxygen species formation (Oliveira & Oliveira, 2002) and mutations (Ragu *et al.*, 2007). Further, our study also has a very limited coverage of soil-associated Nematoda species

(free-living/plant-parasitic), which is bound to affect the conclusions we can draw for this group. The sample size in the soil-associated group of species is small ($n = 5$) and its composition is ecologically heterogeneous, comprising three free-living species and two plant parasites. The free-living species in our study represent very different lifestyles: one lives on beetles (*P. pacificus*) and thus is exposed to similar thermal environment as species parasitizing terrestrial arthropods. The two other species (*C. elegans* and *C. briggsae*) are commonly found in compost heaps, which are thermally more similar to the stable conditions encountered by parasites of endotherms. The two plant parasitizing species (*X. americanum* and *R. similis*) live in plant roots, and are exposed to local climatic conditions which might differ significantly between geographical locations. The very limited conclusion drawn from our study is that mitogenome sizes in this group are very variable and should be analysed in the context of detailed species and location-specific habitat information. Therefore, we should treat all our results for this group as at least provisional until more data becomes available for analysis.

To conclude, the three hypotheses we tested in this study (longevity-selection, functional constraints and race for replication) were initially proposed for and tested on free-living endothermal taxa (birds and mammals). Our results present the first attempt to look at the evolutionary processes shaping mitochondrial genome architecture in ectothermal endoparasites. We discovered that the thermal environment generated by the definitive hosts is associated with mitogenome size of the Nematoda parasites. Therefore, we have shown how considering interactions between species could add a new layer to the emerging picture of mitochondrial genome evolution. Our results should stimulate further studies in invertebrate and vertebrate taxa and parasitic and nonparasitic species alike.

Acknowledgments

SN and ML were supported by the National Research Centre for Growth and Development, New Zealand. All authors declare no conflict of interest.

References

Abascal, F., Zardoya, R. & Posada, D. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**: 2104–2105.

Adachi, J., Cao, Y. & Hasegawa, M. 1993. Tempo and mode of mitochondrial DNA evolution in vertebrates at the amino acid sequence level: rapid evolution in warm-blooded vertebrates. *J. Mol. Evol.* **36**: 270–281.

Blank, H.M., Li, C., Mueller, J.E., Bogomolnaya, L.M., Bryk, M. & Polymenis, M. 2008. An increase in mitochondrial DNA promotes nuclear DNA replication in yeast. *PLoS Genet.* **4**: E1000047.

Blaxter, M. 2011. Nematodes: the worm and its relatives. *PLoS Biol.* **9**: e1001050.

Boore, J.L. 1999. Animal mitochondrial genomes. *Nucl. Acids Res.* **27**: 1767–1780.

Bourque, G. & Pevzner, P. 2002. Genome-scale evolution: reconstructing gene orders in the ancestral species. *Genome Res.* **12**: 26–36.

Boyce, T.M., Zwick, M.E. & Aquadro, C.F. 1989. Mitochondrial DNA in the bark weevils: size, structure and heteroplasmy. *Genetics* **123**: 825–836.

Bromham, L. & Penny, D. 2003. The modern molecular clock. *Nat. Rev. Genet.* **4**: 216–224.

Burger, G., Gray, M.W. & Lang, B.F. 2003. Mitochondrial genomes: anything goes. *Trends Genet.* **19**: 709–716.

Calow, P. 1983. Pattern and paradox in parasite reproduction. *Parasitology* **86**: 197–207.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**: 540–552.

De Ley, P. & Blaxter, M.L. 2002. Systematic position and phylogeny. In: *The Biology of Nematodes* (D.L. Lee, ed.), pp. 1–30. Taylor and Francis, London.

Douzery, E.J.P., Snell, E.A., Bapteste, E., Delsuc, F. & Philippe, H. 2004. The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proc. Natl. Acad. Sci. USA* **101**: 15386–15391.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**: 1969–1973.

Estabrook, G.F., Smith, G.R. & Dowling, T.E. 2007. Body mass and temperature influence rates of mitochondrial DNA evolution in North American cyprinid fish. *Evolution* **61**: 1176–1187.

Galtier, N., Jobson, R.W., Nabholz, B., Glémin, S. & Blier, P.U. 2009. Mitochondrial whims: metabolic rate, longevity and the rate of molecular evolution. *Biol. Lett.* **5**: 413–416.

Gems, D. 2000. Longevity and ageing in parasitic and free-living nematodes. *Biogerontology* **1**: 289–307.

Gilbert, D. 2003. Sequence file format conversion with command-line Readseq. *Curr. Prot. Bioinf.* Appendix 1E.

Gjetvaj, B., Cook, D.I. & Zouros, E. 1992. Repeated sequences and large-scale size variation of mitochondrial DNA: a common feature among scallops (Bivalvia: Pectinidae). *Mol. Biol. Evol.* **9**: 106–124.

Griffith, O.L., Moodie, G.E. & Civetta, A. 2003. Genome size and longevity in fish. *Exp. Gerontol.* **38**: 333–337.

Hadfield, J.D. 2010. MCMC methods for multi–response generalised linear mixed models: the MCMCglmm R Package. *J. Stat. Softw.* **33**: 1–22.

Hadfield, J.D. & Nakagawa, S. 2010. General quantitative genetic methods for comparative biology: phylogenies, taxonomies, and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* **23**: 494–508.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* **41**: 95–98.

Khaidakov, M., Siegel, E.R. & Shmookler Reis, R.J. 2006. Direct repeats in mitochondrial DNA and mammalian lifespan. *Mech. Ageing Dev.* **127**: 808–812.

Kirchner, T.B., Anderson, R.V. & Ingham, R.E. 1980. Natural selection and the distribution of nematode sizes. *Ecology* **61**: 232–237.

Lynch, M. 1991. Methods for the analysis of comparative data in evolutionary biology. *Evolution* **45**: 1065–1080.

Martin, A.P. & Palumbi, S.R. 1993. Body size, metabolic rate, generation time, and the molecular clock. *Proc. Natl Acad. Sci. USA* **90**: 4087–4091.

McGaughran, A. & Holland, B.R. 2010. Testing the effect of metabolic rate on DNA variability at the intra-specific level. *PLoS ONE* **5**: e9686.

Monaghan, P. & Metcalfe, N.B. 2000. Genome size and longevity. *Trends Genet.* **16**: 331–332.

Nabholz, B., Glémin, S. & Galtier, N. 2008. Strong variations of mitochondrial mutation rate across mammals – the longevity hypothesis. *Mol. Biol. Evol.* **251**: 120–130.

Nabholz, B., Glémin, S. & Galtier, N. 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evol. Biol.* **9**: 54.

Nakagawa, S. & Cuthill, I. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev.* **82**: 591–605.

Oliveira, P.L. & Oliveira, M.F. 2002. Vampires, Pasteur and reactive oxygen species. Is the switch from aerobic to anaerobic metabolism a preventive antioxidant defence in blood-feeding parasites? *FEBS Lett.* **525**: 3–6.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. *Nature* **401**: 877–884.

Park, J.K., Sultana, T., Lee, S.H., Kang, S., Kim, H.K., Min, G.S. *et al.* 2007. Monophyly of clade III nematodes is not supported by phylogenetic analysis of complete mitochondrial genome sequences. *BMC Genomics* **12**: 392.

Poulin, R. 2007. *Evolutionary Ecology of Parasites*, 2nd edn. Princeton University Press, Princeton, NJ.

R Development Core Team. 2011. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. (http://www.R-project.org).

Ragu, S., Faye, G., Iraqui, I., Masurel-Heneman, A., Kolodner, R.D. & Huang, M.E. 2007. Oxygen metabolism and reactive oxygen species cause chromosomal rearrangements and cell death. *Proc. Natl Acad. Sci. USA* **104**: 9747–9752.

Rambaut, A. & Drummond, A.J. 2007. Tracer v1.4. (http://beast.bio.ed.ac.uk/Tracer).

Rand, D.M. 1993. Endotherms, ectotherms, and mitochondrial genome-size variation. *J. Mol. Evol.* **37**: 281–295.

Rand, D.M. 1994. Thermal habit, metabolic rate and the evolution of mitochondrial DNA. *Trends Ecol. Evol.* **9**: 125–131.

Rottenberg, H. 2007. Exceptional longevity in songbirds is associated with high rates of evolution of cytochrome b, suggesting selection for reduced generation of free radicals. *J. Exp. Biol.* **210**: 2170–2180.

Samuels, D.C., Schon, E.A. & Chinnery, P.F. 2004. Two direct repeats cause most human mtDNA deletions. *Trends Genet.* **20**: 393–398.

Selosse, M., Albert, B. & Godelle, B. 2001. Reducing the genome size of organelles favours gene transfer to the nucleus. *Trends Ecol. Evol.* **16**: 135–141.

Signorovitch, A.Y., Buss, L.W. & Dellaporta, S.L. 2007. Comparative genomics of large mitochondria in Placozoans. *PLoS Genet.* **3**: E13.

Smit, A.F.A., Hubley, R. & Green, P. 1996–2010. RepeatMasker Open-3.0. (http://www.repeatmasker.org).

Speakman, J.R. 2005. Body size, energy metabolism and life span. *J. Exp. Biol.* **208**: 1717–1730.

Talavera, G. & Castresana, J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**: 564–577.

Tanabe, A.S. 2011. Kakusan4 and Aminosan: two programs for comparing nonpartitioned, proportional, and separate models for combined molecular phylogenetic analyses of multilocus sequence data. *Mol. Ecol. Res.* **11**: 914–921.

Thomas, W.K. & Beckenbach, A.T. 1989. Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. *J. Mol. Evol.* **29**: 233–245.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.

Tian, D., Wang, Q., Zhang, P., Araki, H., Yang, S., Kreitman, M. *et al.* 2008. Single-nucleotide mutation rate increases close to insertions/deletions in eukaryotes. *Nature* **455**: 105–108.

Welch, J.J., Bininda-Emonds, O.R. & Bromham, L. 2008. Correlates of substitution rate variation in mammalian protein-coding sequences. *BMC Evol. Biol.* **8**: 53.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Details of methods and additional results

Data deposited at Dryad: doi: 10.5061/dryad.kf490

Received 6 July 2012; revised 5 November 2012; accepted 5 November 2012