

Mitochondrial DNA of *Clathrina clathrus* (Calcarea, Calcinea): Six Linear Chromosomes, Fragmented rRNAs, tRNA Editing, and a Novel Genetic Code

Dennis V. Lavrov,^{*,1} Walker Pett,¹ Oliver Voigt,² Gert Wörheide,^{2,3} Lise Forget,⁴ B. Franz Lang,⁴ and Ehsan Kayal^{†,1}

¹Department of Ecology, Evolution, and Organismal Biology, Iowa State University

²Department of Earth and Environmental Sciences & GeoBio-Center^{LMU}, Ludwig-Maximilians-Universität München, München, Germany

³Bayerische Staatssammlung für Paläontologie und Geologie, München, Germany

⁴Centre Robert Cedergren, Département de Biochimie, Université de Montréal, Québec, Canada

[†]Present address: Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC

*Corresponding author: E-mail: dlavrov@iastate.edu.

Associate editor: Todd Oakley

Abstract

Sponges (phylum Porifera) are a large and ancient group of morphologically simple but ecologically important aquatic animals. Although their body plan and lifestyle are relatively uniform, sponges show extensive molecular and genetic diversity. In particular, mitochondrial genomes from three of the four previously studied classes of Porifera (Demospongiae, Hexactinellida, and Homoscleromorpha) have distinct gene contents, genome organizations, and evolutionary rates. Here, we report the mitochondrial genome of *Clathrina clathrus* (Calcinea, Clathrinidae), a representative of the fourth poriferan class, the Calcarea, which proves to be the most unusual. *Clathrina clathrus* mitochondrial DNA (mtDNA) consists of six linear chromosomes 7.6–9.4 kb in size and encodes at least 37 genes: 13 protein codings, 2 ribosomal RNAs (rRNAs), and 24 transfer RNAs (tRNAs). Protein genes include *atp9*, which has now been found in all major sponge lineages, but no *atp8*. Our analyses further reveal the presence of a novel genetic code that involves unique reassignments of the UAG codons from termination to tyrosine and of the CGN codons from arginine to glycine. *Clathrina clathrus* mitochondrial rRNAs are encoded in three (srRNA) and ≥ 6 (lrRNA) fragments distributed out of order and on several chromosomes. The encoded tRNAs contain multiple mismatches in the aminoacyl acceptor stems that are repaired posttranscriptionally by 3'-end RNA editing. Although our analysis does not resolve the phylogenetic position of calcareous sponges, likely due to their high rates of mitochondrial sequence evolution, it confirms mtDNA as a promising marker for population studies in this group. The combination of unusual mitochondrial features in *C. clathrus* redefines the extremes of mtDNA evolution in animals and further argues against the idea of a "typical animal mtDNA."

Key words: Porifera, mitochondrial DNA, genome evolution.

Introduction

Mitochondria, the double-membrane organelles found in most eukaryotic cells and best known for their role in energy production, contain their own genome (mitochondrial DNA [mtDNA]), which is maintained, expressed, and inherited separately from the nuclear genome (reviewed in Gibor and Granick 1964; Birky 2001). Although, the genetic function of mtDNA is well conserved, its size, structure, and gene expression mechanisms vary greatly among and within major lineages of eukaryotes (Lang et al. 1999; Burger, Gray, et al. 2003). By contrast, mtDNA of animals is generally considered remarkably uniform and is typically described as a small and economically organized circular molecule with a conserved gene set, modified genetic code, stable gene order,

and high rates of sequence evolution. However, such portrayal of animal mtDNA is based on early studies limited primarily to bilaterian animals (e.g., Clary and Wolstenholme 1985) and is now out of date. A more expanded sampling of bilaterian animals and, especially, recent work on nonbilaterian animals (phyla Cnidaria, Ctenophora, Placozoa, and Porifera) revealed substantial mitochondrial genomic diversity within these groups (reviewed in Lavrov 2011). This diversity is particularly pronounced in sponges, which should be expected given their ancient origin (Love et al. 2009) and early diversification (Pisera 2006).

Sponges (phylum Porifera) are sessile aquatic animals that play a major role in marine and many freshwater benthic communities (Van Soest et al. 2012). They are subdivided into

four classes: Calcarea (calcareous sponges), Demospongiae (demosponges), Hexactinellida (glass sponges), and Homoscleromorpha (homoscleromorphs) (Hooper et al. 2002; Gazave et al. 2011). The first two complete mitochondrial genomes of sponges were determined only in 2005 (Lavrov et al. 2005). Since then, approximately 30 other sponge mt-genomes have been published (Erpenbeck et al. 2007, 2009; Haen et al. 2007; Wang and Lavrov 2007, 2008; Belinky et al. 2008; Lukic-Bilela et al. 2008; Rosengarten et al. 2008; Gazave et al. 2010; Lavrov 2010, 2012; Ereskovsky et al. 2011; Pleše et al. 2012; Sperling et al. 2012). This body of work revealed distinct evolutionary trajectories in mitochondrial genome evolution among and within major lineages of sponges. Thus, mitochondrial genomes of demosponges and homoscleromorphs are characterized by the retention of several features characteristic of many nonmetazoan eukaryotes, including a minimally modified genetic code, presence of several extra genes, more bacteria-like structures of encoded transfer and ribosomal RNAs (rRNAs), low rate of sequence evolution, and the presence of multiple noncoding regions (Lavrov et al. 2005; Wang and Lavrov 2007, 2008). By contrast, mitochondrial genomes of glass sponges have evolved several features in parallel with their counterparts in bilaterian animals and share with them a similar mtDNA organization with a single large noncoding region, a similar nucleotide composition of the coding strand, and, most surprisingly, an identical reassignment of the AGR codons from arginine to serine (Haen et al. 2007; Rosengarten et al. 2008). Within Demospongiae, loss of transfer RNA (tRNA) genes has occurred in several lineages (Wang and Lavrov 2008; Erpenbeck et al. 2009), and a transfer of *atp9* to the nucleus has been reported (Erpenbeck et al. 2007). Within Homoscleromorpha two different mitochondrial organizations have been found: all genomes from the family Oscarellidae are characterized by the presence of *tatC*, a gene for subunit C of the twin arginine translocase, otherwise not known in animals, as well 27 tRNA genes, whereas those from the family Plakinidae lack all but five tRNA genes present in Oscarellidae (Gazave et al. 2010). Remarkably, to date there are no complete and only a few partial mt sequences from Calcarea (Voigt, Eichmann, et al. 2012), the fourth class of sponges. Moreover, several mitochondrial sequences in GenBank assigned to this group appear to be contaminations of demosponge origin (Borchiellini et al. 2004).

Calcareous sponges (class Calcarea) are the only sponges with skeletons made of calcium carbonate (CaCO_3) spicules. The majority of approximately 680 described species are small, colorless, and inconspicuous organisms found mostly in cryptic marine habitats (e.g., beneath rocks and inside cavities) in shallow waters (Wörheide and Hooper 2003); however, a few can reach up to 50 cm in size and some have been collected from bathyal and abyssal zones (Van Soest et al. 2012). Calcareous sponges are present in seas at all latitudes and display a wide variety of organizations, including the famous morphocline in the organization of the aquiferous system, which can be of an asconoid, syconoid, solenoid, syllaebid, and leuconoid type (Voigt, Wülfing, et al.

2012). Based on morphological and cytological data, calcareous sponges have been placed as a sister group to either Demospongiae (e.g., Cellularia [Reiswig et al. 1983]) or Demospongiae + Hexactinellida (e.g., Silicea [Gray 1867]). By contrast, most molecular phylogenetic studies using rRNA (Cavalier-Smith et al. 1996; Zrzavy et al. 1998; Borchiellini et al. 2001; Medina et al. 2001; Manuel et al. 2003) and nuclear housekeeping (Sperling et al. 2009) genes reconstructed Porifera as a paraphyletic group, with calcareous sponges being more closely related to Eumetazoa than to other sponges. Recent phylogenomic analyses based on expressed sequence tags (ESTs) data reversed this trend and supported the monophyly of sponges (Philippe et al. 2009, 2011; Pick et al. 2010) but placed Calcarea as the sister group to Homoscleromorpha, a newly established class of sponges (Gazave et al. 2011) (reviewed in Wörheide et al. 2012). Within Calcarea, two subclasses Calcinea and Calcaronea are recognized and well supported by the analyses of cytological, developmental, and molecular characters (see Manuel 2006, for a review). However, the phylogenetic relationships within these subclasses remain highly contentious (Manuel et al. 2003; Dohrmann et al. 2006; Voigt, Wülfing, et al. 2012).

Here, we report the mitochondrial genome sequence of the calcareous sponge *Clathrina clathrus* (Calcinea, Clathrinidae) and four mitochondrial complementary DNA (cDNA) sequences of *Leucetta chagosensis* (Calcinea, Leucettidae). These highly unusual sequences provide a first glimpse of mitochondrial genome evolution in Calcarea and broaden our view on the existing diversity of mitochondrial genomes in sponges and animals in general.

Results

Multipartite and Linear Architecture of the *C. clathrus* Mitochondrial Genome

The mitochondrial genome of *C. clathrus* collected at the Station Marine d'Endoume, Marseille (France) was assembled from paired end Illumina DNA sequencing reads in six contigs 5,738; 6,382; 7,920; 8,013; 8,413; and 9,199 bp in size, which were extended by polymerase chain reaction (PCR)/Sanger sequencing using primers designed for the ends of the longer contigs to 7,162+; 8,857; 7,920; 8,866; 9,269; and 9,370 bp, respectively (fig. 1). The sequences of all contigs were identical at one end and similar at the other. The contigs also displayed a comparable internal organization with genes compactly arrayed in approximately one-half of each of them and another half containing exclusively noncoding sequence. The total size of the assembly was > 50 kb, making *C. clathrus* mtDNA the largest animal mitochondrial genome reported to date.

Three disjoint mtDNA regions were also PCR-amplified and sequenced from another specimen of *C. clathrus* collected off Isola del Giglio (Italy) (fig. 1A) and used to design probes for Southern hybridization (SH) analysis. SH analysis with probes for *ml_5*, *cob*, and *cox1* revealed a peculiar double-band hybridization pattern, with the estimated sizes of the lower bands at 7, 9, and ~10 kb, and



pattern can be interpreted either as that of partially dimerized linear DNA molecules, or of circular plasmids with fast-migrating supercoiled circles and slowly migrating open circles. However, because no Illumina sequences

have been found that would link the two terminal regions and no PCR reaction across the ends was successful, we interpreted *C. clathrus* mitochondrial chromosomes as linear.

High Inter- and Intraspecific Variation in Coding Sequences

Thirteen protein genes (*atp6*, *atp9*, *cob*, *cox1–3*, *nad1–6*, and *nad4L*) were identified in the mitochondrial genome of *C. clathrus* (France). These genes were localized to five different chromosomes in both transcriptional directions (fig. 1). Seven of them (*atp6*, *cob*, *cox1*, *nad1–4*) were also present in PCR-amplified fragments from *C. clathrus* (Italy). In addition, four mitochondrial protein genes (*cob*, *cox1*, *cox3*, and *nad1*) were identified in an EST library of *L. chagosensis* and PCR-amplified from another specimen of the same species. Analyses of calcareous mitochondrial coding sequences revealed substantial variation at both the intra- and interspecific levels. The two specimens of *C. clathrus* had overall 7.9% sequence divergence in coding sequences (from 4% for *cob* to 17.9% for *nad2*) (table 1). Similarly, the two specimens of *L. chagosensis* displayed on average 4.3% sequence divergence at the nucleotide level. Finally, coding sequences in *C. clathrus* and *L. chagosensis* differed at more than 50% of sites, comparable with the differences between *C. clathrus* and the homoscleromorph *Oscarella carmela* (table 1). These large sequence divergences in both intra- and interspecific comparisons reflect high rates of sequence evolution in calcareous mtDNA as proposed earlier (Lavrov et al. 2006; Voigt, Eichmann, et al. 2012), which also manifest themselves as exceptionally long branches in phylogenetic analyses (discussed later). At the same time, the ratios of synonymous to nonsynonymous substitutions in mitochondrial coding sequences were consistently <1 (table 1), indicating the

presence of purifying selection and hence the functionality of these genes.

A Novel Mitochondrial Genetic Code in Calcarean Sponges

Conceptual translation of *C. clathrus* mitochondrial coding regions revealed multiple internal UAG “termination” codons (63 in total), suggesting a modified genetic code. Comparative sequence analysis using GenDecoder (Abascal et al. 2006) showed that 79% of *C. clathrus* UAG codons present at highly conserved positions (entropy <1.0 and gaps <20%) corresponded to tyrosine codons in other animals. In addition, this analysis uncovered an unexpected pattern of occurrence of CGN codons, preferentially used at position specifying glycine in other animals (>85% codons at highly conserved sites) and not found at highly conserved arginine positions. Finally, *C. clathrus* UGA codons were mostly located at tryptophan position in the alignments (~88% codons at highly conserved sites), consistent with the use of this codon to specify tryptophan in animal mitochondria. We also analyzed patterns of codon usage in four mitochondrial genes from the *L. chagosensis* cDNA library (Philippe et al. 2009) with similar results: 72% of UAG codons and >80% of CGN codons at highly conserved sites corresponded to tyrosine and glycine, respectively. Because unusual codon usage was observed in both gDNA and cDNA sequences, we ruled out mRNA editing as an explanation for it. Furthermore, the fact that identical changes were found in two phylogenetically distant species of calcarean sponges (Rossi et al. 2011) indicates that they may be characteristic for the whole subclass. Together, these results suggested a novel mitochondrial genetic code in calcarean sponges with the reassignment of CGN codons from arginine to glycine, and the UAG codon from termination to tyrosine. The latter reassignment is supported

Table 1. Mitochondrial Coding Sequences in *Clathrina clathrus*.

Gene	Size ^a (codons)	Rates of Evolution ^b			Sequence Identity ^c (%)			Initiation Codon	Stop Codon
		dN	dS	dN/dS	CC/CC1	CC/LC	CC/OC		
<i>atp6</i>	254	0.046	0.217	0.213	91.6		39.3	ATG	TAA
<i>atp9</i>	78	—	—	—			61.0	ACA ^d	TAA
<i>Cob</i>	372	0.004	0.123	0.031	96.0	53.0	46.1	ATT	TAA
<i>cox1</i>	474	0.006	0.187	0.032	95.2	58.3	50.2	AAA ^d	TAA
<i>cox2</i>	216	—	—	—			43.5	ATG	TAG
<i>cox3</i>	269	—	—	—		44.3	43.3	CCT ^d	TAA
<i>Nad1</i>	312	0.019	0.140	0.136	95.1	46.6	45.9	ATC	TAA
<i>Nad2</i>	462	0.143	0.468	0.306	82.1		33.6	ATG	TAA
<i>Nad3</i>	169	0.047	0.326	0.146	89.3		29.7	ATG	TAA
<i>Nad4</i>	477	0.010	0.201	0.051	94.3		40.4	ATG	TAG
<i>nad4L</i>	89	—	—	—	—		39.7	GGG ^d	TAA
<i>Nad5</i>	530	—	—	—	—		38.6	ATG	TAA
<i>Nad6</i>	198	—	—	—	—		34.6	ATG	AAT ^e

^aIncluding stop codon.
^bIn comparison between two *C. clathrus* specimens.
^cPercent sequence identity between the following specimens: CC, *C. clathrus* (France); CC1, *C. clathrus* (Italy); LC, *Leucetta chagosensis* (Australia); and OC, *Oscarella carmela*.
^dNo conventional start codon has been found for the reading frame.
^eNo conventional termination codon has been found for the reading frame.

by the presence of an encoded tRNA with a CUA anticodon in *C. clathrus* mtDNA as described in the following section.

Clathrina clathrus mt-tRNAs Include an Unusual tRNA_{CUA} and Show Evidence of Template-Dependent Editing

We identified 23 sequences in *C. clathrus* mtDNA that could be folded into characteristic tRNA secondary structure with a capacity to decode all codon families except for alanine GCN and tyrosine UAY. The set of potential tRNAs included an unusual tRNA with an CUA anticodon, inferred to be responsible for the translation of the reassigned UAG and possibly also UAY codons as tyrosine. The predicted structure of this

tRNA supports its identification as tRNA^{Tyr}_{CUA}, as it belongs to so-called type II tRNAs—distinguished by a long variable arm—which is restricted to three families: tRNAs^{Tyr}, tRNAs^{Leu}, and tRNAs^{Ser} (Normanly et al. 1992). However, *C. clathrus* tRNA^{Tyr}_{CUA} has an unusual structure as it contains an extra nucleotide between the aminoacyl acceptor stem and the D-arm, which can form a Watson–Crick pair with nucleotide 26 (fig. 2B). In addition, the set of encoded tRNAs includes two with the CAU anticodon, inferred to code for tRNA^{Met}_{CAU} and tRNA^{Ile}_{CAU} (supplementary fig. S3, Supplementary Material online) as in demosponges, homoscleromorphs, placozoans, and many nonmetazoan eukaryotes (Lavrov 2011).

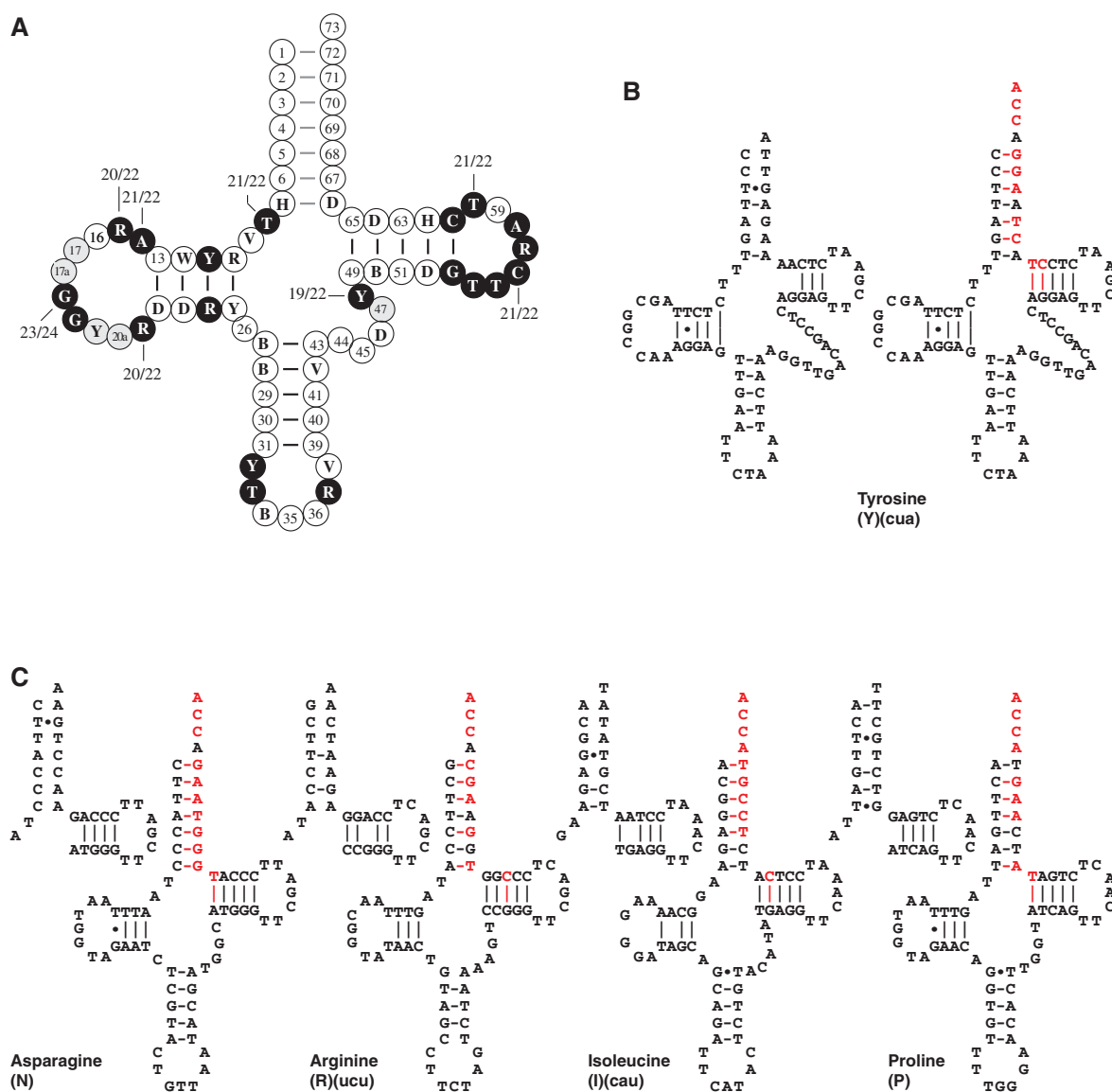


FIG. 2. Sequence conservation and editing in mitochondrial tRNAs of *Clathrina clathrus*. (A) Sequence conservation in mitochondrial tRNA genes of *C. clathrus* (except trnY(cua)). Numbering of nucleotides is based on the convention used for yeast tRNA phenylalanine (Robertus et al. 1974). Open circles with numbers: nucleotides are present in all tRNAs; open circles with letters: nucleotide combinations present in all tRNAs; filled black circle: nucleotides or nucleotide combinations that are described as invariant or semi-invariant in prokaryotic and eukaryotic nuclear-encoded tRNAs with frequencies shown by accompanying numbers when not 100% conserved and filled gray circles: nucleotides present in some but not all tRNAs. (B) The mtDNA-encoded and the posttranscriptionally edited sequence of the unusual tRNA^{Tyr}_{CUA}. Edited nucleotides are marked by color. (C) Additional examples of RNA editing in *C. clathrus* mt-tRNAs. mtDNA encoded (unedited) amino-acyl and TΨC stems are shown next to each tRNA.

All encoded tRNAs displayed well-conserved sequences of the anticodon, DHU (D-) and TΨC (T-) loops (fig. 2A), but had multiple mismatches in the aminoacyl acceptor stems, suggesting tRNA editing. The occurrence of editing has been experimentally confirmed for five tRNAs (tRNA^{Arg}_{UCU}, tRNA^{Asn}_{GUU}, tRNA^{Ile}_{CAU}, tRNA^{Pro}_{UGG}, and tRNA^{Tyr}_{CUA}) using the reverse transcriptase-PCR (RT-PCR) approach described in Price and Gray (1999) (fig. 2B and C). The edited tRNAs formed perfect Watson–Crick complementary paired T-arms and aminoacyl acceptor arms with the CCA sequence added at the 3'-end (Chen et al. 1992). The editing involved all four nucleotides and thus appears to be template dependent (but see Schürer et al. 2001).

An additional interesting feature of the set of *C. clathrus* mt-tRNAs is the presence of an Y11–R24 pair in all of them. Although Y11–R24 is a conserved feature of most tRNAs, some have a characteristic R–Y pair at these positions. tRNAs with the R11–Y24 pair include prokaryotic and organellar initiator tRNA methionine (Marck and Grosjean 2002), animal mt-tRNA tryptophan (Wolstenholme 1992), and mt-tRNA proline in glass sponges, demosponges, and homoscleromorphs (Lavrov 2007). The exclusive occurrence of the Y11–R24 in all *C. clathrus* mt-tRNAs suggests concomitant changes at these positions in methionine, tryptophan, and proline tRNAs in this species.

SSU- and LSU-rRNA Are Encoded by Fragmented Genes Distributed Across Several Chromosomes

We used Infernal 1.0.2 (Nawrocki et al. 2009) and BLAST (Altschul et al. 1990) to identify genes for the large and small subunit rRNA (LSU-rRNA and SSU-rRNA, respectively) in the mitochondrial sequences of *C. clathrus*. Both rRNA genes were discontinuous and located on several chromosomes (fig. 1), but when pieced together, encoded well-conserved rRNA structures (fig. 3). The gene for SSU-rRNA (*rns*) was split into three fragments located on chromosomes 1 and 2 (fig. 3A). The gene for LSU-rRNA (*rnl*) was split into at least six fragments located on three different chromosomes. The 5' half of LSU-rRNA (helices 1–60) was encoded by three genomic regions on chromosome 2 in a nonsequential order (fig. 1). Because the 5' half is a less conserved part of the molecule, additional unidentified regions encoding this part may be present. The better conserved 3'-end of the molecule is also split into three fragments: two encoded on chromosome 1 and separated by *trnP(ugg)* and *trnK(uuu)*, and one encoded on chromosome 6 (figs. 1 and 3B). We used RT-PCR to confirm that the transcribed fragments are not spliced together (amplifications were observed within but not between individual fragments) and to confirm the expression of *trnP(ugg)* located between *rnl_4* and *rnl_5*, as described in the previous section. Interestingly, despite their unusual organization, the rate of sequence evolution in *C. clathrus* mitochondrial rRNAs appeared to be similar to those in other animals, in contrast to protein genes, where it was highly accelerated (supplementary fig. S4, Supplementary Material online).

Noncoding Regions and Terminal Repeats

Each mitochondrial chromosome of *C. clathrus* contained a large terminal noncoding region (hereafter referred to as noncoding region 1 or *nc1*) that comprised about half of its length (3,593–5,210 bp). In addition, a second large noncoding region (noncoding region 2 or *nc2*) was present at the other end of the chromosomes and varied from 909 to 1,671 bp in size. The sequences of the *nc1* showed 62–79% sequence identity among chromosomes (fig. 4A), with the distal part being more conserved (fig. 4C). In particular, the terminal sequence of approximately 1,000 bp in the *nc1* was identical among all chromosomes. This terminal sequence contained five direct repeats and had a potential to form multiple closely spaced hairpin structures (fig. 4E).

The sequence of *nc2* was very similar among the chromosomes 2, 4, and 5 (97–98% sequence identity), and, to a lesser extent, the chromosome 6 (~90% identity with the other three chromosomes) (fig. 4B and D). By contrast, the sequence of *nc2* of the chromosome 1 showed on average only ~73% overall identity and that of the chromosome 3—only 61% overall identity with the corresponding regions on the other chromosomes (fig. 4B and D). However, on all chromosomes, there was a strong drop in the number of guanines in the terminal part of *nc2* (this terminal part was preceded by a dinucleotide [TC] repeat in four of the chromosomes). Very few sequences ≥10 bp in length were shared between the two terminal noncoding regions in any of the chromosomes (supplementary fig. S5, Supplementary Material online).

Phylogenetic Analysis of Mitochondrial Protein Sequences Does Not Resolve the Phylogenetic Position of Calcareous Sponges

We used maximum likelihood (ML) and Bayesian (BI) methods to explore the phylogenetic position of calcareous sponges based on mitochondrial sequence data (fig. 5 and supplementary fig. S6, Supplementary Material online). Representatives of all phyla of nonbilaterian animals except Ctenophora as well as of major phyla of bilaterian animals and several outgroups were included in the analysis. The two available mitochondrial genomes of ctenophores (Pett et al. 2011; Kohn et al. 2012) were not included in the analysis because they have been shown to form extremely long branches on phylogenetic trees that are likely to distort the placement of other long branches due to the long branch attraction (LBA) artifact (Felsenstein 1978).

Most of the reconstructed phylogenetic relationships corresponded closely to the current view on animal phylogeny, including the relationships among bilaterian animals and those within major lineages of nonbilaterian animals. However, several unconventional relationships have been recovered, consistent with the earlier phylogenetic studies based on mtDNA data (Lavrov et al. 2005; Haen et al. 2007; Lavrov 2011). The phylogenetic position of calcareous sponges was unstable in our analyses, and was influenced by the type of the analysis, the model of sequence evolution used, and the inclusion/exclusion of the partial sequence from

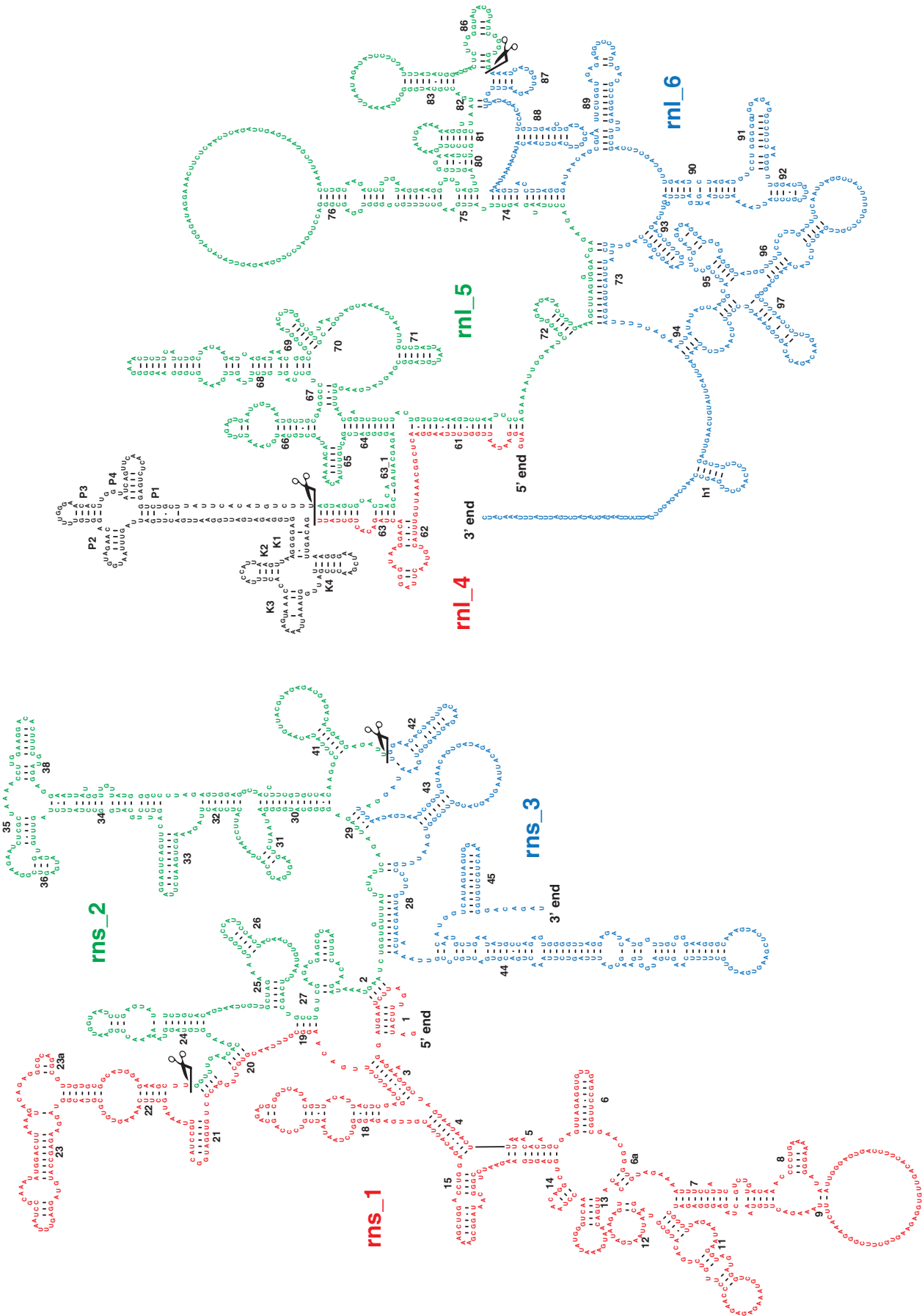


Fig. 3. Inferred secondary structure of mitochondrial rRNAs in *Clathrina clathrus*. (A) Small subunit rRNA. (B) 3' part of the large subunit rRNA. Evolutionary cuts in rRNA structures are shown by scissor symbols. Fragments of rRNA genes are labeled as in figure 1. Helices in rRNA are numbered as in Brimacombe (1995); those in lrrNA as in Leffers et al. (1987). rRNA genes intercalated between fragment 1 and 2 of lrrNA are folded into cloverleaf structures with arms labeled as K1–K4 for *trnK*(*uuu*) and P1–P4 for *trnP*(*ugc*).

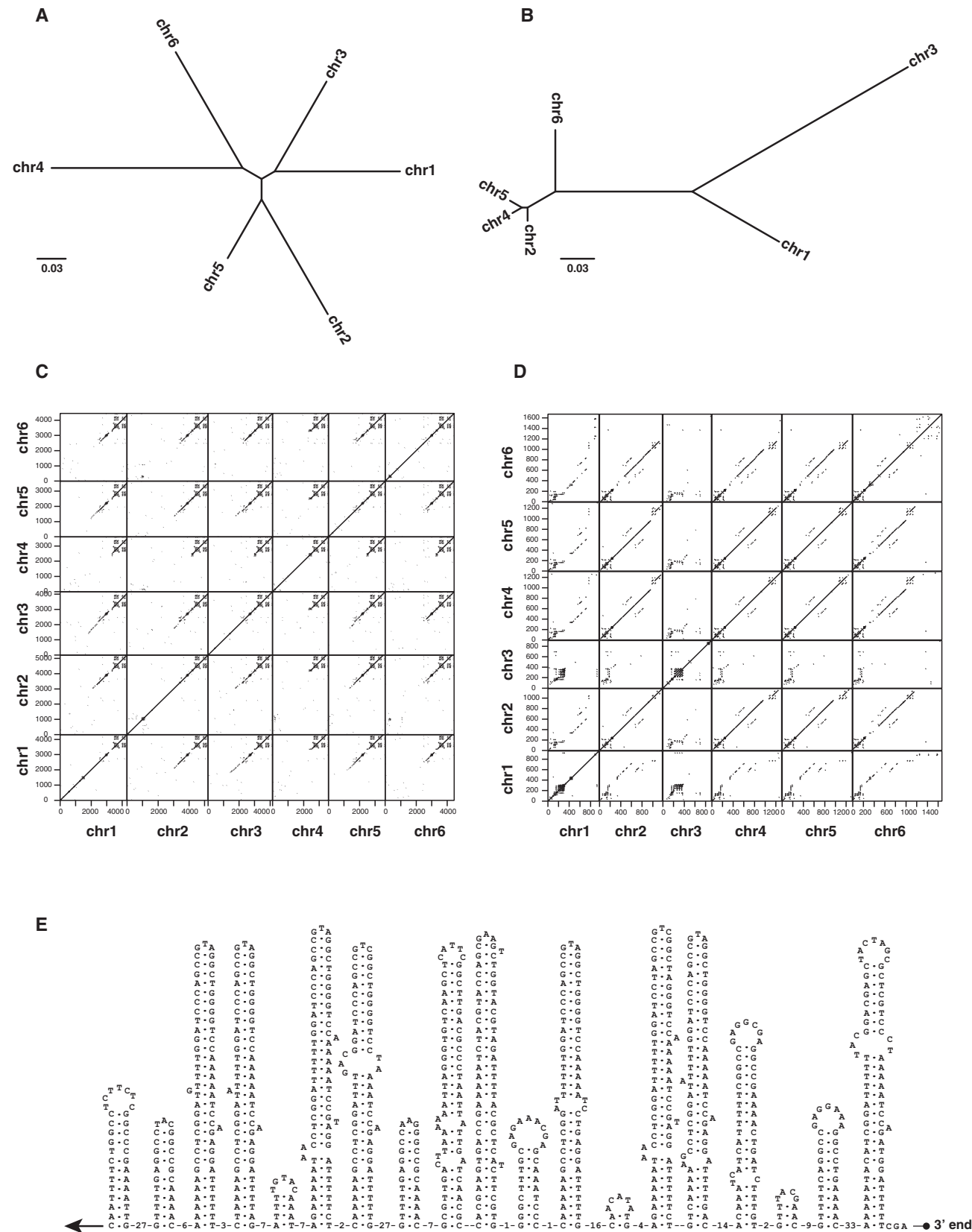


FIG. 4. Sequence conservation and repetition in terminal large noncoding regions of *C. clathrus* mitochondrial chromosomes. (A, B) Sequence similarity in noncoding regions among chromosomes. (A) Noncoding region 1. (B) Noncoding regions 2. Neighbor-joining trees based on P distances are shown. (C, D) Dot plots for the two noncoding regions. The sequence of each chromosome was compared with itself (squares on the diagonal running from the bottom-left corner to the top-right corner of the graph) and between chromosomes (other rectangles) with the EMBOSS Polydot program (Rice et al. 2000). Marks on the graph indicate an exact match of ≥ 10 bp between two sequences. Matches across or outside the diagonal line in each square signify the presence of repeats. (E) Potential secondary structures in the terminal part of *nc1* inferred with the MFold program (Zuker 2003).

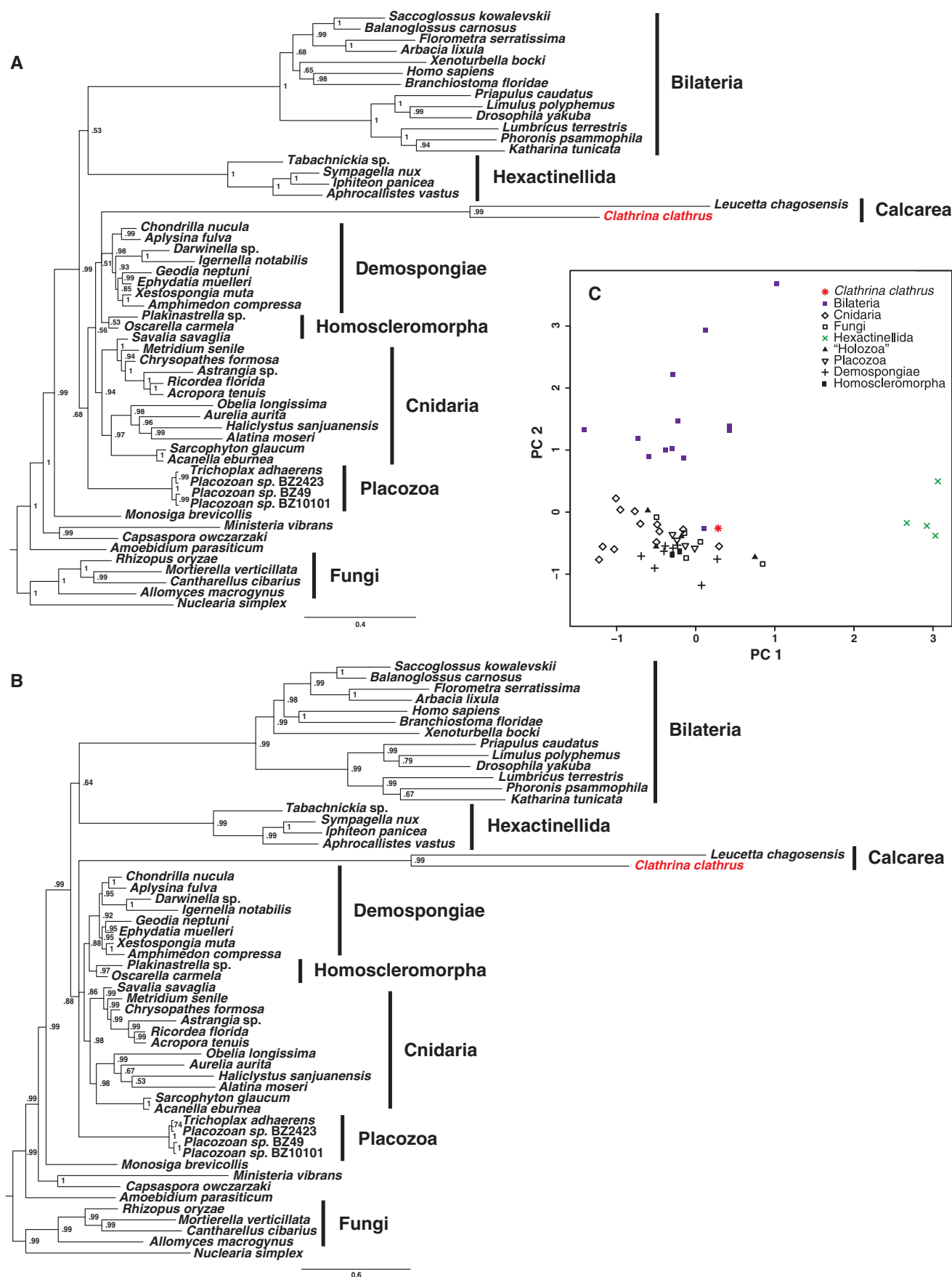


FIG. 5. Phylogenetic analysis of animal relationships using mitochondrial sequence data. Posterior majority-rule consensus trees obtained from the analysis of concatenated mitochondrial amino acid sequences (2473 positions) under the CAT+Γ (A) and CAT+GTR+Γ (B) models in the PhyloBayes program. The number at each node represents the Bayesian posterior probability. (C) Correspondence analysis of amino acid usage in concatenated sequences utilized for phylogenetic inference.

L. chagosensis. In general, the PhyloBayes analyses with the CAT or CAT + GTR model of sequence evolution placed calcareous sponges either in a group with demosponges and homoscleromorphs, or in a polytomy with other nonbilaterian animals (fig. 5). By contrast, ML analyses using a site-homogeneous model of amino acid substitutions produced phylogenies with fast evolving sponges (Calcarea and Hexactinellida) grouping with bilaterian animals (supplementary fig. S6, Supplementary Material online). Importantly, rates of sequence evolution were markedly different among individual lineages, making the results susceptible to the LBA artifact. In addition, we found large deviations in amino acid composition of mitochondrial proteins in bilaterian animals and glass sponges compared with the rest of the animals and outgroups (fig. 5C) that might violate the stationarity assumption made by most models of sequence evolution (Lockhart et al. 1994; Gowri-Shankar and Ratnayake 2006). The use of a site and time-heterogeneous model (CAT-BP) of amino acid replacement (Blanquart and Lartillot 2008) was not able to overcome this bias and produced results similar to the CAT and CAT-GTR models (supplementary fig. S6, Supplementary Material online). Finally, the inferred topology changed little when we eliminated various percentages of sites with the highest rates of sequence evolution from the alignment as described in Goremykin et al. (2010) and Derelle and Lang (2012) (data not shown).

Discussion

A Unique Combination of Unusual Features in *C. clathrus* mtDNA

Early sampling of animal mitochondrial genomes was biased by two different factors: a prevailing scientific interest in some (mostly vertebrate) taxa and technical difficulties associated with other. Although the invention of PCR and design of Metazoa-specific mtDNA primers allowed a broader sampling of animal mt-genomes, the unusual mtDNA architecture and/or rapid evolution, remained an insurmountable obstacle for some taxa. The advent of the next generation sequencing technology overcame these technical limitations and allowed the sequencing of several highly unusual mt-genomes (Shao et al. 2009; Pett et al. 2011; Smith et al. 2012). The mitochondrial genome of the calcareous sponge *C. clathrus* characterized in this study is among the most unusual in sponges and, arguably, among all animals. In this section, we discuss four of its outstanding features: a multipartite and linear genome architecture, a novel genetic code, edited tRNAs, and fragmented rRNAs.

Although it is common to depict animal mtDNA as a single, small, circular molecule, it has been known for more than 20 years that Medusozoa, one of the two major lineages in the phylum Cnidaria, have linear and, in some cases, multipartite genomes (Bridge et al. 1992). Recent studies have characterized several of these genomes and showed that they are composed of 1–8 linear chromosomes (Shao et al. 2006; Kayal and Lavrov 2008; Voigt et al. 2008; Smith et al. 2012; Kayal et al. 2012). In addition, studies of several isopod species have shown that their mtDNA is composed of a

combination of ~14 kb linear monomers and approximately 28 kb head-to-head dimers (Raimond et al. 1999; Marcadé et al. 2007; Doublet et al. 2012). This study revealed that another independent origin of linear genome organization in animal mtDNA occurred within calcareous sponges.

Linear mtDNA is also found in close unicellular relatives of animals *Amoebidium parasiticum* (Burger, Forget, et al. 2003), and potentially in *Capsaspora owczarzaki*, and *Ministeria vibrans* (D.V. Lavrov and B.F. Lang, unpublished data). It is also widespread in plants and fungi, where so-called “polydisperse linear DNA” made of linear concatemers of various sizes appears to be the norm (Nosek et al. 1998; Valach et al. 2011). However, a multipartite linear organization is relatively rare in these groups, with the notable exception of the *A. parasiticum* genome, consisting of several hundred linear chromosomes (Burger, Forget, et al. 2003). By contrast, a multipartite genome organization evolved at least twice among linear mt-genomes in Cnidaria: in the common ancestor of Cubozoa and within Hydrozoa (Voigt et al. 2008; Kayal et al. 2012). Linear and multipartite genome organization requires special mechanisms for the maintenance and transmission of mtDNA (reviewed by Nosek et al. 1998). Based on the presence of two DNA bands in southern blot analysis, we propose that *C. clathrus* mitochondrial chromosomes may resemble that found in *Paramecium aurelia*, where approximately 10% of the molecules are present as linear duplexes of dimer length and where replication of mtDNA is initiated by a cross-linking of the duplex strands at one end of the molecule (Pritchard and Cummings 1981). The mechanism of concerted transmission of all six chromosomes to the daughter mitochondria is unknown. However, a similar problem has been solved several times in organisms that have multichromosomal mitochondrial genomes (Burger, Gray, et al. 2003), including extreme cases of thousands of circular and linear chromosomes (Burger, Forget, et al. 2003; Liu et al. 2005).

The second unusual feature of calcarean sponge mitochondria is a novel genetic code that includes unique reassignments of the UAG codons from termination to tyrosine and of the CGN codon family from arginine to glycine (Calcarean mitochondrial genetic code). Although CGN codons were previously identified as “mutable in their meaning” (Soll and RajBhandary 2006) and may not code for an amino acid in yeast mitochondria (Clark-Walker et al. 1985), this is the first report of their actual reassignment. Similarly, the reassignment of the UAG codon to tyrosine is novel, although changes of the UAG codon to leucine/alanine and glutamine have been found in mtDNA of green plants and nuclear genomes of ciliates, respectively (Knight et al. 2001). Interestingly, a convergent change to tyrosine has been recently inferred for another (UAA) termination codon in a nematode (Jacob et al. 2009).

The third interesting feature of calcarean mitochondrial genomes is fragmented and discontinuous rRNA genes. Fragmented rRNA genes are highly unusual in animal mtDNA and so far have been found only in the oysters *Crassostrea gigas*, *C. hongkongensis*, and *C. virginica*, where

LSU-rRNA is encoded in two pieces (Milbury and Gaffney 2005; Milbury et al. 2010), and in Placozoa, where LSU-rRNA is encoded in two or three pieces (Signorovitch et al. 2007; Burger et al. 2009). Outside of Metazoa, fragmented mitochondrial rRNAs have been described in bacteria, green algae, ciliates, and some other protist lineages (Schnare et al. 1986; Heinonen et al. 1987; Boer and Gray 1988; Denovan-Wright and Lee 1995; Nedelcu 1997; Evguenieva-Hackenberg 2005) as well as monoblepharidalean fungi (Forget et al. 2002; Bullerwell et al. 2003). Extreme rRNA segmentation has been documented for mitochondrial rRNA genes of apicomplexans and dinoflagellates (Feagin et al. 1997; Jackson et al. 2007, 2012; Waller and Jackson 2009). As is also the case in other organisms (Gillespie et al. 1999), there was no obvious RNA processing signal at the inferred ends of *C. clathrus* mt RNA fragments and mechanisms of their processing/assembly remain unknown.

Finally, mitochondrial sequences of calcareous sponges encode tRNA molecules with multiple mismatches in the amino-acyl acceptor stem, which undergo posttranscriptional RNA editing. Several different types of mitochondrial tRNA editing have been reported to date, including C-to-U editing in mitochondria of marsupials (Janke and Pääbo 1993; Borner et al. 1996), trypanosomes (Alfonzo et al. 1999), and plants (Maréchal-Drouard et al. 1993; Fey et al. 2002); insertion editing in slime molds (Antes et al. 1998); 5'-end editing in amoebozoans (Lonergan and Gray 1993a, 1993b) and some fungi (Laforest et al. 1997); or a combination of these (Gott et al. 2010). The 3'-end editing similar to that found in calcareous sponges has been found in some bilaterian animals (Yokobori and Paabo 1995; Tomita et al. 1996; Yokobori and Pääbo 1997; Reichert et al. 1998; Lavrov et al. 2000; Segovia et al. 2011) and a jakobid flagellate (Leigh and Lang 2004). Further studies are needed to determine whether the tRNA editing in *C. clathrus* is a recent acquisition that is specific for this species/group of species, as is typical for animals (Brennicke et al. 1999) or if it is a common feature of all calcinean/calcareous sponges. In either case, calcareous sponges do represent an example of yet another independent acquisition of 3'-end tRNA editing in Metazoa.

The combination of the four unusual mitochondrial features listed above makes *C. clathrus* mtDNA among the most unusual in animals. Elucidation of its structure adds a new dimension to the observed diversity of animal mtDNA by showing that all four classes of sponges (Demospongiae, Hexactinellida, Homoscleromorpha, and Calcarea) have distinct modes and tempos of mitochondrial genome evolution. Furthermore, our previous studies of demosponges and homoscleromorphs (Wang and Lavrov 2008; Gazave et al. 2010), as well as our preliminary mitochondrial data from *Sycon ciliatum* and *Petrobiona massiliana*, two calcareous sponges belonging to the second calcarean subclass Calcaronea, reveal additional variation in mt genome organization within classes of sponges.

Conclusion

Because of their phylogenetic position as the sister group to the rest of the animals as well as their ancient origin and early

diversification, sponges can be expected to harbor a large molecular and genomic diversity, most of which is only starting to be explored. The variation in mitochondrial genome architecture among the four classes of sponges is, in our view, a reflection of this diversity and a telltale sign of sponge nuclear genome diversity. A highly unusual mitochondrial genome of *C. clathrus* with multiple linear chromosomes, several changes in the genetic code, fragmented rRNA genes, tRNA editing, and high rate of sequence evolution pushes the boundaries of what is possible in mtDNA evolution in animals and warrants further studies of molecular mechanisms underlying these changes. Although high rates of sequence evolution in calcareous sponge mtDNA prevented us from resolving their phylogenetic position, they make calcarean mtDNA a promising marker for population studies in this group.

Materials and Methods

DNA Extraction and Sequencing

Specimens of *C. clathrus* were collected by Michael Nickel off Isola del Giglio (Italy) in 2004 and by DVL at the Station marine d'Endoume, Marseille (France) in 2008, and preserved in 8M guanidine hydrochloride solution. Total DNA from both specimens was prepared by phenol-chloroform extraction following proteinase K digestion (Saghai-Marroof et al. 1984). We used two different approaches to determine mtDNA sequences from these specimens.

A PCR-based approach (Burger et al. 2007) was used for the first specimen. Regions of *cob*, *cox1*, and *rnl* were amplified and sequenced with conserved primers developed for animal mtDNA (Burger et al. 2007). Sequences were extended using a modified step-out protocol developed in our lab (Burger et al. 2007) and reamplified by long PCR with the TAKARA LA-PCR kit. Long PCR products were combined in equimolar concentration, sheared into pieces 1–2 kb in size, end-repaired, and cloned using the TOPO Shotgun Subcloning Kit from Invitrogen. White colonies containing inserts were collected, grown overnight in 96-well blocks, and submitted to the DNA Sequencing and Synthesis Facility of the ISU Office of Biotechnology for high-throughput plasmid preparation and sequencing.

A total DNA sequencing approach was used for the second specimen. Paired 100 bp reads of the total genomic DNA were generated with Illumina sequencing by synthesis chemistry. mtDNA sequence was assembled directly from these reads (supplementary fig. S1, Supplementary Materials online). Repeated sequences in the terminal regions were PCR-amplified using primers designed at the ends of the longest contigs and sequenced using Sanger sequencing. In addition to mtDNA, both 18S and 28S rRNA gene sequences were assembled for this specimen of *C. clathrus* and used to confirm its identification by comparison with orthologs from other calcareous sponges (supplementary fig. S2, Supplementary Materials online). Both conventional (Sanger) and Illumina (Solexa) sequencing were conducted at the Iowa State University DNA Facility on an Applied Biosystems 3730xl DNA Analyzer and Illumina HiSeq 2000,

respectively. Sanger sequences were assembled with the STADEN software suite (Staden 1996). Illumina sequences were assembled with ABySS using a k-mer size of 82 (Simpson et al. 2009). All assemblies were submitted to NCBI's GenBank (accessions JX978466–JX978471 and JX996194–JX996196).

Southern Blot Analysis

Southern hybridization was performed following a modified version of the protocol described in Sambrook et al. (1989) on a total DNA preparation from the specimen of *C. clathrus* collected off Isola del Giglio (Italy). Partial sequences of three genes, *cob* (chromosome 4), *cox1* (chr. 5), and *rnl_5* (chr. 1) were independently amplified with specific primers and radiolabeled by random priming. Three samples of total DNA (~1 µg each) and approximately 0.1 ng of each PCR product were run on a 0.8 % agarose gel and capillary transferred to a Hybond-XL nylon membrane (GE Healthcare). After treatment, the membrane was cut in three pieces corresponding to each PCR product, which were exposed separately but simultaneously to corresponding radiolabeled probes during an overnight incubation at 60°C. The results of SH were visualized by exposing an X-ray film to membranes for 40 h.

RNA Extraction and cDNA Synthesis

Total RNA was extracted from the specimen of *C. clathrus* collected at the Station marine d'Endoume using the TRIZOL Reagent (Invitrogen) and treated with DNase. An aliquot of total RNA was circularized using T4 RNA ligase (Fermentas). tRNA, partial rRNAs (both *rns* and *rnl*), and mRNA (*cox1*) sequences were amplified by RT-PCR, with primers designed based on mtDNA sequences (supplementary table S1, Supplementary Material online). Reverse transcription was performed using SuperScript III Reverse Transcriptase (Invitrogen) under conditions described by the manufacturer. PCR amplification of cDNA products was performed with recombinant Taq DNA polymerase (Invitrogen). PCR products for tRNAs were cloned using the TA Cloning Kit from Invitrogen, while PCR amplicons for the mRNA and rRNAs were directly sequenced. At least eight colonies were analyzed per tRNA gene.

Partial mt Sequences of *L. chagosensis*

EST sequences of one specimen of *L. chagosensis* sampled at North Stradbroke Island, Australia, were available from a previous study (Philippe et al. 2009). Mitochondrial protein-coding genes were identified from the EST pool via BLAST searches (Altschul et al. 1990). *Leucetta*-specific primers were designed for *cob*, *cox1*, *cox3*, and *nad1* (supplementary table S1, Supplementary Material online) and used to PCR-amplify and sequence these genes from another specimen of *L. chagosensis* sampled in the Gulf of Aqaba, Egypt.

Gene Annotation, RNA Folding, and Genome Description

Open reading frames in the nucleotide sequences were first characterized using the minimally derived code (with

TGA = Trp). Further deviations from this genetic code were discovered with the GenDecoder v.1.6 program (Abascal et al. 2006). tRNA genes were identified by tRNAscan-SE (Lowe and Eddy 1997); protein genes were recognized by similarity searches in local databases using the FASTA program (Pearson 1994), and in GenBank at NCBI using BLAST network service (Benson et al. 2003). rRNA genes were characterized by the Cmsearch program in the Infernal 1.0.2 package (Nawrocki et al. 2009) based on manually curated alignments of nonbilaterian mitochondrial rRNA sequences, and BLAST searches (Altschul et al. 1990) against other mitochondrial sequences. The secondary structures of these genes were adjusted manually according to the patterns of conservation in the three phylogenetic domains + chloroplasts + mitochondria compiled by the Gutell lab at the Comparative RNA Web Site (<http://www.ma.icmb.utexas.edu>, last accessed 18 December 2012) and drawn with RNAviz (De Rijk et al. 2003).

Phylogenetic Inference

Two data sets were created from mitochondrial sequences of *C. clathrus* and representatives of major lineages of opisthokonts (supplementary table S2, Supplementary Materials online) with and without partial sequences of *L. chagosensis*. Inferred amino acid sequences of individual mitochondrial proteins were aligned with Mafft v6.861b (Katoh and Toh 2008). Conserved blocks within the alignments were selected with Gblocks 0.91b (Castresana 2000) using relaxed parameters (parameters 1 and 2 = 1/2, parameter 3 = 8, parameter 4 = 5, all gap positions in parameter 5). Cleaned alignments were concatenated in two data sets 2,492 and 2,473 amino acids in length that encompassed 53 and 54 species, respectively. Additional alignments were created by removing 10–50% of the positions with the highest rate of evolution (as estimated by PAML [Yang 2007]) from the dataset of 53 species. Finally, a reduced alignment consisting of 39 species and 2,492 amino acid positions was created and used for phylogenetic analysis with a nonhomogenous model of amino acid substitution.

Bayesian inferences were performed with the CAT + Γ 4 and CAT + GTR + Γ 4 mixture models implemented in the program PhyloBayes version 3.2f (Lartillot et al. 2009) and the CATBP model implemented in nhPhyloBayes v.0.2.1 (Blanquart and Lartillot 2008). PhyloBayes analyses consisted of four chains over 15,000 generations using CAT + GTR + Γ model (maxdiff < 0.1), and four chains over 65,000 generations using CAT + Γ model (maxdiff < 0.15). The chains were sampled every 10th tree after the first 100 burn-in cycles. nhPhyloBayes analysis consisted of two chains over 5,000 generation using the CATBP + GTR + Γ model, sampled every 10th tree after the first 500 burn-in cycles. ML analyses were performed in RAXML version 7.3 (Stamatakis 2006), using “the easy and fast way” (100 bootstrap replicates followed by an ML search). We used the PROTGAMMALGF model (Le and Gascuel 2008) for ML searches, which is the best-fitting model for the concatenated matrix according to the Akaike information criterion

(computed with ProtTest 3.2; Darriba et al. 2011). Site rates were calculated using CODEML (Yang 2007) under the same model and positions with the highest substitution rates were sequentially eliminated from the alignment.

The correspondence analysis was performed using the R package *ca* (Greenacre and Nenadic 2007). For all species used in phylogenetic analysis, we determined the total amino acid usage, obtaining a matrix where the rows represent the species and the 20 columns are the respective amino acid frequencies.

Supplementary Material

Supplementary figures S1–S6 and tables S1 and S2 are available at Molecular Biology and Evolution online (<http://www.mbe.oxfordjournals.org/>).

Acknowledgments

This work was supported by the National Science Foundation grant (no. DEB-0828783) to D.V.L.; the National Sciences and Engineering Research Council of Canada grant (no. NSERC 194560-2011) to B.F.L.; and the German Research Foundation (DFG) projects WO896/3 and WO896/6 within the SPP1174 “Deep Metazoan Phylogeny” to G.W. The authors thank Michael Nickel and Dan Jackson for specimens of *Clathrina clathrus* and *Leucetta chagosensis*, respectively, and Katherine Wilson and Katrina Lutap for technical assistance. They also thank Michelle Klautau, two anonymous reviewers, and the associate editor for their comments on an earlier version of this manuscript.

References

- Abascal F, Zardoya R, Posada D. 2006. GenDecoder: genetic code prediction for metazoan mitochondria. *Nucleic Acids Res.* 34: W389–W393.
- Alfonzo J, Blanc V, Estevez A, Rubio M, Simpson L. 1999. C to U editing of the anticodon of imported mitochondrial tRNA(Trp) allows decoding of the UGA stop codon in *Leishmania tarentolae*. *EMBO J.* 18:7056–7062.
- Altschul S, Gish W, Miller W, Myers E, Lipman D. 1990. Basic local alignment search tool. *J Mol Biol.* 215:403–410.
- Antes T, Costandy H, Mahendran R, Spottswood M, Miller D. 1998. Insertional editing of mitochondrial tRNAs of *Physarum polycephalum* and *Didymium nigripes*. *Mol Cell Biol.* 18:7521–7527.
- Belinky F, Rot C, Ilan M, Huchon D. 2008. The complete mitochondrial genome of the demosponge *Negombata magnifica* (Poecilosclerida). *Mol Phylogenet Evol.* 47:1238–1243.
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2003. GenBank. *Nucleic Acids Res.* 31:23–27.
- Birky CJ. 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu Rev Genet.* 35:125–148.
- Blanquart S, Lartillot N. 2008. A site- and time-heterogeneous model of amino acid replacement. *Mol Biol Evol.* 25:842–858.
- Boer P, Gray M. 1988. Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. *Cell* 55:399–411.
- Borchiellini C, Chombard C, Manuel M, Alivon E, Vacelet J, Boury-Esnault N. 2004. Molecular phylogeny of Demospongiae: implications for classification and scenarios of character evolution. *Mol Phylogenet Evol.* 32:823–837.
- Borchiellini C, Manuel M, Alivon E, Boury-Esnault N, Vacelet J, Le Parco Y. 2001. Sponge paraphyly and the origin of Metazoa. *J Evol Biol.* 14: 171–179.
- Borner G, Mörl M, Janke A, Pääbo S. 1996. RNA editing changes the identity of a mitochondrial tRNA in marsupials. *EMBO J.* 15: 5949–5957.
- Brennick A, Marchfelder A, Binder S. 1999. RNA editing. *FEMS Microbiol Rev.* 23:297–316.
- Bridge D, Cunningham CW, Schierwater B, Desalle R, Buss LW. 1992. Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. *Proc Natl Acad Sci U S A.* 89: 8750–8753.
- Brimacombe R. 1995. The structure of ribosomal RNA: a three-dimensional jigsaw puzzle. *Eur J Biochem.* 230:365–383.
- Bullerwell CE, Forget L, Lang BF. 2003. Evolution of monoblepharidalean fungi based on complete mitochondrial genome sequences. *Nucleic Acids Res.* 31:1614–1623.
- Burger G, Forget L, Zhu Y, Gray MW, Lang BF. 2003. Unique mitochondrial genome architecture in unicellular relatives of animals. *Proc Natl Acad Sci U S A.* 100:892–897.
- Burger G, Gray MW, Lang BF. 2003. Mitochondrial genomes: anything goes. *Trends Genet.* 19:709–716.
- Burger G, Lavrov D, Forget L, Lang B. 2007. Sequencing complete mitochondrial and plastid genomes. *Nat Protoc.* 2:603–614.
- Burger G, Yan Y, Javadi P, Lang B. 2009. Group I-intron trans-splicing and mRNA editing in the mitochondria of placozoan animals. *Trends Genet.* 25:381–386.
- Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 17: 540–552.
- Cavalier-Smith T, Allsopp M, Chao E, Boury-Esnault N, Vacelet J. 1996. Sponge phylogeny, animal monophyly, and the origin of the nervous system: 18S rRNA evidence. *Can J Zool.* 74:2031–2045.
- Chen J, Joyce P, Wolfe C, Steffen M, Martin N. 1992. Cytoplasmic and mitochondrial tRNA nucleotidyltransferase activities are derived from the same gene in the yeast *Saccharomyces cerevisiae*. *J Biol Chem.* 267:14879–14883.
- Clark-Walker GD, McArthur C, Sriprakash K. 1985. Location of transcriptional control signals and transfer RNA sequences in *Torulopsis glabrata* mitochondrial DNA. *EMBO J.* 4:465–473.
- Clary DO, Wolstenholme DR. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *J Mol Evol.* 22:252–271.
- Darriba D, Taboada G, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27: 1164–1165.
- De Rijk P, Wuyts J, De Wachter R. 2003. RnaViz 2: an improved representation of RNA secondary structure. *Bioinformatics* 19: 299–300.
- Denovan-Wright EM, Lee R. 1995. Evidence that the fragmented ribosomal RNAs of *Chlamydomonas* mitochondria are associated with ribosomes. *FEBS Lett.* 370:222–226.
- Derelle R, Lang B. 2012. Rooting the eukaryotic tree with mitochondrial and bacterial proteins. *Mol Biol Evol.* 29:1277–1289.
- Dohrmann M, Voigt O, Erpenbeck D, Wörheide G. 2006. Non-monophyly of most supraspecific taxa of calcareous sponges (Porifera, Calcarea) revealed by increased taxon sampling and partitioned Bayesian analysis of ribosomal DNA. *Mol Phylogenet Evol.* 40:830–843.
- Doublet V, Raimond R, Grandjean F, Lafitte A, Souty-Grosset C, Marcade I. 2012. Widespread atypical mitochondrial DNA structure in isopods (Crustacea, Peracarida) related to a constitutive heteroplasmy in terrestrial species. *Genome* 55:234–244.
- Ereskovsky A, Lavrov D, Boury-Esnault N, Vacelet J. 2011. Molecular and morphological description of a new species of *Halisarca* (Demospongiae: Halisarcida) from Mediterranean Sea and a redescription of the type species *Halisarca dujardini*. *Zootaxa* 2768:5–31.
- Erpenbeck D, Voigt O, Adamski M, Adamska M, Hooper J, Wörheide G, Degnan B. 2007. Mitochondrial diversity of early-branching metazoa is revealed by the complete mt genome of a haplosclerid demosponge. *Mol Biol Evol.* 24:19–22.

- Erpenbeck D, Voigt O, Wörheide G, Lavrov D. 2009. The mitochondrial genomes of sponges provide evidence for multiple invasions by repetitive hairpin-forming elements (RHE). *BMC Genomics* 10:591.
- Evguenieva-Hackenberg E. 2005. Bacterial ribosomal RNA in pieces. *Mol Microbiol.* 57:318–325.
- Feagin J, Mericle B, Werner E, Morris M. 1997. Identification of additional rRNA fragments encoded by the *Plasmodium falciparum* 6 kb element. *Nucleic Acids Res.* 25:438–446.
- Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool.* 27:401–410.
- Fey J, Weil J, Tomita K, Cosset A, Dietrich A, Small I, Marechal-Drouard L. 2002. Role of editing in plant mitochondrial transfer RNAs. *Gene* 286: 21–24.
- Forget L, Ustinova J, Wang Z, Huss V, Lang B. 2002. *Hyaloraphidium curvatum*: a linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi. *Mol Biol Evol.* 19:310–319.
- Gazave E, Lapébie P, Ereskovsky AV, Vacelet J, Renard E, Cárdenas P, Borchellini C. 2011. No longer Demospongiae: Homoscleromorpha formal nomination as a fourth class of Porifera. *Hydrobiologia* 687: 3–10.
- Gazave E, Lapebie P, Renard E, Vacelet J, Rocher C, Ereskovsky A, Lavrov D, Borchellini C. 2010. Molecular phylogeny restores the supra-generic subdivision of homoscleromorph sponges (Porifera, Homoscleromorpha). *PLoS One* 5:e14290.
- Gibor A, Granick S. 1964. Plastids and mitochondria: inheritable systems. *Science* 145:890–897.
- Gillespie D, Salazar N, Rehkopf D, Feagin J. 1999. The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum* have short A tails. *Nucleic Acids Res.* 27:2416–2422.
- Goremykin V, Nikiforova S, Bininda-Emonds OR. 2010. Automated removal of noisy data in phylogenomic analyses. *J Mol Evol.* 71: 319–331.
- Gott J, Somerlot B, Gray M. 2010. Two forms of RNA editing are required for tRNA maturation in *Physarum mitochondria*. *RNA* 16: 482–488.
- Gowri-Shankar V, Rattray M. 2006. On the correlation between composition and site-specific evolutionary rate: implications for phylogenetic inference. *Mol Biol Evol.* 23:352–364.
- Gray J. 1867. Notes on the arrangement of sponges, with the description of some new genera. *Proc Zool Soc Lond.* 2:492–558.
- Greenacre M, Nenadic O. 2007. Correspondence analysis in R, with two- and three-dimensional graphics: the ca package. *J Stat Software.* 20: 1–13.
- Haen K, Lang B, Pomponi S, Lavrov D. 2007. Glass sponges and bilaterian animals share derived mitochondrial genomic features: a common ancestry or parallel evolution? *Mol Biol Evol.* 24:1518–1527.
- Heinonen T, Schnare M, Young P, Gray M. 1987. Rearranged coding segments, separated by a transfer RNA gene, specify the two parts of a discontinuous large subunit ribosomal RNA in *Tetrahymena pyriformis* mitochondria. *J Biol Chem.* 262:2879–2887.
- Hooper JNA, Van Soest RWM, Debrenne F. 2002. Phylum Porifera Grant, 1836. In: Hooper JNA, Van Soest RWM, editors. *Systema Porifera: a guide to the classification of sponges*. New York: Kluwer Academic/Plenum Publishers. p. 9–13.
- Jackson C, Gornik S, Waller R. 2012. The mitochondrial genome and transcriptome of the basal dinoflagellate *Hematodinium* sp.: character evolution within the highly derived mitochondrial genomes of dinoflagellates. *Genome Biol Evol.* 4:59–72.
- Jackson C, Norman J, Schnare M, Gray M, Keeling P, Waller R. 2007. Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria. *BMC Biol.* 5:41.
- Jacob J, Vanholme B, Van Leeuwen T, Gheysen G. 2009. A unique genetic code change in the mitochondrial genome of the parasitic nematode *Radopholus similis*. *BMC Res Notes.* 2:192.
- Janke A, Pääbo S. 1993. Editing of a tRNA anticodon in marsupial mitochondria changes its codon recognition. *Nucleic Acids Res.* 21: 1523–1525.
- Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform.* 9:286–298.
- Kayal E, Bentlage B, Collins A, Kayal M, Pirro S, Lavrov D. 2012. Evolution of linear mitochondrial genomes in medusozoan cnidarians. *Genome Biol Evol.* 4:1–12.
- Kayal E, Lavrov D. 2008. The mitochondrial genome of *Hydra oligactis* (Cnidaria, Hydrozoa) sheds new light on animal mtDNA evolution and cnidarian phylogeny. *Gene* 410:177–186.
- Knight R, Freeland S, Landweber L. 2001. Rewiring the keyboard: evolvability of the genetic code. *Nat Rev Genet.* 2:49–58.
- Kohn A, Citarella M, Kocot K, Bobkova Y, Halanych K, Moroz L. 2012. Rapid evolution of the compact and unusual mitochondrial genome in the ctenophore, *Pleurobrachia bachei*. *Mol Phylogenet Evol.* 63: 203–207.
- Laforest M, Roewer I, Lang B. 1997. Mitochondrial tRNAs in the lower fungus *Spizellomyces punctatus*: tRNA editing and UAG 'stop' codons recognized as leucine. *Nucleic Acids Res.* 25:626–632.
- Lang BF, Gray MW, Burger G. 1999. Mitochondrial genome evolution and the origin of eukaryotes. *Annu Rev Genet.* 33:351–397.
- Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25:2286–2288.
- Lavrov D. 2007. Key transitions in animal evolution: a mitochondrial DNA perspective. *Integr Comp Biol.* 47:734–743.
- Lavrov D. 2010. Rapid proliferation of repetitive palindromic elements in mtDNA of the endemic Baikalian sponge *Lubomirskia baicalensis*. *Mol Biol Evol.* 27:757–760.
- Lavrov D, Brown W, Boore J. 2000. A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proc Natl Acad Sci U S A.* 97:13738–13742.
- Lavrov D, Forget L, Kelly M, Lang B. 2005. Mitochondrial genomes of two demosponges provide insights into an early stage of animal evolution. *Mol Biol Evol.* 22:1231–1239.
- Lavrov D, Maikova O, Pett W, Belikov S. 2012. Small inverted repeats drive mitochondrial genome evolution in Lake Baikal sponges. *Gene* 505:91–99.
- Lavrov DV. 2011. Key transitions in animal evolution: a mitochondrial DNA perspective. In: Desalle R, Schierwater B, editors. *Key transitions in animal evolution*. Enfield (NH): Science. p. 35–54.
- Lavrov DV, Frishman N, Haen K, Shao Z, Wang X. 2006. Mitochondrial genomics of sponges—implications for sponge phylogeny and animal mtDNA evolution. 7th International Sponge Symposium. Biodiversity, Innovation, Sustainability: book of abstracts. Armação de Búzios, Rio de Janeiro, Brazil.
- Le S, Gascuel O. 2008. An improved general amino acid replacement matrix. *Mol Biol Evol.* 25:1307–1320.
- Leffers H, Kjems J, Ostergaard L, Larsen N, Garrett R. 1987. Evolutionary relationships amongst archaeobacteria. A comparative study of 23 S ribosomal RNAs of a sulphur-dependent extreme thermophile, an extreme halophile and a thermophilic methanogen. *J Mol Biol.* 195: 43–61.
- Leigh J, Lang B. 2004. Mitochondrial 3' tRNA editing in the jakobid *Seculamonas ecuadoriensis*: a novel mechanism and implications for tRNA processing. *RNA* 10:615–621.
- Liu B, Liu Y, Motyka S, Agbo E, Englund P. 2005. Fellowship of the rings: the replication of kinetoplast DNA. *Trends Parasitol.* 21: 363–369.
- Lockhart PJ, Steel MA, Hendy MD, Penny D. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol Biol Evol.* 11:605–612.
- Loneragan K, Gray M. 1993a. Editing of transfer RNAs in *Acanthamoeba castellanii* mitochondria. *Science* 259:812–816.
- Loneragan K, Gray M. 1993b. Predicted editing of additional transfer RNAs in *Acanthamoeba castellanii* mitochondria. *Nucleic Acids Res.* 21:4402.
- Love G, Grosjean E, Stalvies C, et al. (13 co-authors). 2009. Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature* 457:718–721.
- Lowe T, Eddy S. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955–964.

- Lukic-Bilela L, Brandt D, Pojskic N, Wiens M, Gamulin V, Muller W. 2008. Mitochondrial genome of *Suberites domuncula*: palindromes and inverted repeats are abundant in non-coding regions. *Gene* 412: 1–11.
- Manuel M. 2006. Phylogeny and evolution of calcareous sponges. *Can J Zool.* 84:225–241.
- Manuel M, Borchellini C, Alivon E, Le Parco Y, Boury-Esnault JV. 2003. Phylogeny and evolution of calcareous sponges: monophyly of Calcinea and Calcaronea, high level of morphological homoplasy, and the primitive nature of axial symmetry. *Syst Biol.* 52:311–333.
- Marcadé I, Cordaux R, Doublet V, Debenest C, Bouchon D, Raimond R. 2007. Structure and evolution of the atypical mitochondrial genome of *Armadillidium vulgare* (Isopoda, Crustacea). *J Mol Evol.* 65: 651–659.
- Marck C, Grosjean H. 2002. tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* 8:1189–1232.
- Maréchal-Drouard L, Ramamonjisoa D, Cosset A, Weil JH, Dietrich A. 1993. Editing corrects mispairing in the acceptor stem of bean and potato mitochondrial phenylalanine transfer RNAs. *Nucleic Acids Res.* 21:4909–4914.
- Medina M, Collins AG, Silberman JD, Sogin ML. 2001. Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. *Proc Natl Acad Sci U S A.* 98: 9707–9712.
- Milbury C, Gaffney P. 2005. Complete mitochondrial DNA sequence of the eastern oyster *Crassostrea virginica*. *Mar Biotechnol* (NY). 7: 697–712.
- Milbury C, Lee J, Cannone J, Gaffney P, Gutell R. 2010. Fragmentation of the large subunit ribosomal RNA gene in oyster mitochondrial genomes. *BMC Genomics* 11:485.
- Nawrocki E, Kolbe D, Eddy S. 2009. Infernal 1.0: inference of RNA alignments. *Bioinformatics* 25:1335–1337.
- Nedelcu AM. 1997. Fragmented and scrambled mitochondrial ribosomal RNA coding regions among green algae: a model for their origin and evolution. *Mol Biol Evol.* 14:506–517.
- Normanly J, Ollick T, Abelson J. 1992. Eight base changes are sufficient to convert a leucine-inserting tRNA into a serine-inserting tRNA. *Proc Natl Acad Sci U S A.* 89:5680–5684.
- Nosek J, Tomaska L, Fukuhara H, Suyama Y, Kovac L. 1998. Linear mitochondrial genomes: 30 years down the line. *Trends Genet.* 14: 184–188.
- Pearson WR. 1994. Using the FASTA program to search protein and DNA sequence databases. *Methods Mol Biol.* 25:365–389.
- Pett W, Ryan J, Pang K, Mullikin J, Martindale M, Baxeianis A, Lavrov D. 2011. Extreme mitochondrial evolution in the ctenophore *Mnemiopsis leidyi*: insight from mtDNA and the nuclear genome. *Mitochondrial DNA* 22:130–142.
- Philippe H, Brinkmann H, Lavrov D, Littlewood D, Manuel M, Wörheide G, Baurain D. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biol.* 9:e1000602.
- Philippe H, Derelle R, Lopez P, et al. (20 co-authors). 2009. Phylogenomics revives traditional views on deep animal relationships. *Curr Biol.* 19:706–712.
- Pick K, Philippe H, Schreiber F, et al. (11 co-authors). 2010. Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. *Mol Biol Evol.* 27:1983–1987.
- Pisera A. 2006. Palaeontology of sponges—a review. *Can J Zool.* 84: 242–261.
- Pleše B, Lukić-Bilela L, Bruvo-Madžarić B, Harget M, Imešek M, Bilandžija H, Četković H. 2012. The mitochondrial genome of stygobitic sponge *Eunapius subterraneus*: mtDNA is highly conserved in freshwater sponges. *Hydrobiologia* 687:49–59.
- Price D, Gray M. 1999. Confirmation of predicted edits and demonstration of unpredicted edits in *Acanthamoeba castellanii* mitochondrial tRNAs. *Curr Genet.* 35:23–29.
- Pritchard A, Cummings D. 1981. Replication of linear mitochondrial DNA from *Paramecium*: sequence and structure of the initiation-end crosslink. *Proc Natl Acad Sci U S A.* 78:7341–7345.
- Raimond R, Marcade I, Bouchon D, Rigaud T, Bossy J, Souty-Grosset C. 1999. Organization of the large mitochondrial genome in the isopod *Armadillidium vulgare*. *Genetics* 151:203–210.
- Reichert A, Rothbauer U, Morl M. 1998. Processing and editing of overlapping tRNAs in human mitochondria. *J Biol Chem.* 273: 31977–31984.
- Reiswig H, Mackie G, Reiswig H, Mackie G. 1983. Studies on hexactinellid sponges. III. The taxonomic status of Hexactinellida within the Porifera. *Philos Trans R Soc London B Biol Sci.* 301:419–428.
- Rice P, Longden I, Bleasby A. 2000. EMBOS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16:276–277.
- Robertus JD, Ladner JE, Finch JT, Rhodes D, Brown RS, Clark BF, Klug A. 1974. Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature* 250:546–551.
- Rosengarten R, Sperling E, Moreno M, Leys S, Dellaporta S. 2008. The mitochondrial genome of the hexactinellid sponge *Aphrocallistes vastus*: evidence for programmed translational frameshifting. *BMC Genomics* 9:33.
- Rossi A, De Moraes Russo CA, Solé-Cava AM, Rapp H, Klautau M. 2011. Phylogenetic signal in the evolution of body colour and spicule skeleton in calcareous sponges. *Zool J Linn Soc.* 163: 1026–1034.
- Saghai-Marroof MA, Soliman K, Jorgensen R, Allard R. 1984. Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci U S A.* 81:8014–8018.
- Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory Press.
- Schnare M, Heinonen T, Young P, Gray M. 1986. A discontinuous small subunit ribosomal RNA in *Tetrahymena pyriformis* mitochondria. *J Biol Chem.* 261:5187–5193.
- Schürer H, Schiffer S, Marchfelder A, Mörl M. 2001. This is the end: processing, editing and repair at the tRNA 3'-terminus. *Biol Chem.* 382:1147–1156.
- Segovia R, Pett W, Trewick S, Lavrov D. 2011. Extensive and evolutionarily persistent mitochondrial tRNA editing in velvet worms (phylum onychophora). *Mol Biol Evol.* 28:2873–2881.
- Shao R, Kirkness E, Barker S. 2009. The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, *Pediculus humanus*. *Genome Res.* 19: 904–912.
- Shao Z, Graf S, Chaga O, Lavrov D. 2006. Mitochondrial genome of the moon jelly *Aurelia aurita* (Cnidaria, Scyphozoa): a linear DNA molecule encoding a putative DNA-dependent DNA polymerase. *Gene* 381:92–101.
- Signorovitch A, Buss L, Dellaporta S. 2007. Comparative genomics of large mitochondria in placozoans. *PLoS Genet.* 3:e13.
- Simpson J, Wong K, Jackman S, Schein J, Jones S, Birol I. 2009. ABYSS: a parallel assembler for short read sequence data. *Genome Res.* 19: 1117–1123.
- Smith D, Kayal E, Yanagihara A, Collins A, Pirro S, Keeling P. 2012. First complete mitochondrial genome sequence from a box jellyfish reveals a highly fragmented linear architecture and insights into telomere evolution. *Genome Biol Evol.* 4:52–58.
- Soll D, Rajbhandary U. 2006. The genetic code—thawing the 'frozen accident'. *J Biosci.* 31:459–463.
- Sperling E, Peterson K, Pisani D. 2009. Phylogenetic-signal dissection of nuclear housekeeping genes supports the paraphyly of sponges and the monophyly of Eumetazoa. *Mol Biol Evol.* 26: 2261–2274.
- Sperling EA, Rosengarten RD, Moreno MA, Dellaporta SL. 2012. The complete mitochondrial genome of the verongid sponge *Aplysina cauliformis*: implications for DNA barcoding in demosponges. *Hydrobiologia* 687:61–69.
- Staden R. 1996. The Staden sequence analysis package. *Mol Biotechnol.* 5: 233–241.
- Stamatakis A. 2006. RAxML-VI-HP: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.

- Tomita K, Ueda T, Watanabe K. 1996. RNA editing in the acceptor stem of squid mitochondrial tRNA(Tyr). *Nucleic Acids Res.* 24:4987–4991.
- Valach M, Farkas Z, Fricova D, et al. (11 co-authors). 2011. Evolution of linear chromosomes and multipartite genomes in yeast mitochondria. *Nucleic Acids Res.* 39:4202–4219.
- Van Soest RWM, Boury-Esnault N, Vacelet J, Dohrmann M, Erpenbeck D, De Voogd NJ, Santodomingo N, Vanhoorne B, Kelly M, Hooper JNA. 2012. Global Diversity of Sponges (Porifera). *PLoS One* 7: e35105.
- Voigt O, Eichmann V, Wörheide G. 2012. First evaluation of mitochondrial DNA as a marker for phylogeographic studies of Calcarea: a case study from *Leucetta chagosensis*. *Hydrobiologia* 687:101–106.
- Voigt O, Erpenbeck D, Wörheide G. 2008. A fragmented metazoan organellar genome: the two mitochondrial chromosomes of *Hydra magnipapillata*. *BMC Genomics* 9:350.
- Voigt O, Wülfing E, Wörheide G. 2012. Molecular phylogenetic evaluation of classification and scenarios of character evolution in calcareous sponges (Porifera, class Calcarea). *PLoS One* 7:e33417.
- Waller R, Jackson C. 2009. Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. *Bioessays* 31:237–245.
- Wang X, Lavrov D. 2007. Mitochondrial genome of the homoscleromorph *Oscarella carmela* (Porifera, Demospongiae) reveals unexpected complexity in the common ancestor of sponges and other animals. *Mol Biol Evol.* 24:363–373.
- Wang X, Lavrov D. 2008. Seventeen new complete mtDNA sequences reveal extensive mitochondrial genome evolution within the Demospongiae. *PLoS One* 3:e2723.
- Wolstenholme DR. 1992. Animal mitochondrial DNA: structure and evolution. *Int Rev Cytol.* 141:173–216.
- Wörheide G, Dohrmann M, Erpenbeck D, Larroux C, Maldonado M, Voigt O, Borchellini C, Lavrov D. 2012. Deep phylogeny and evolution of sponges (phylum Porifera). *Adv Mar Biol.* 61:1–78.
- Wörheide G, Hooper JNA. 2003. New species of Calcaronea (Porifera: Calcarea) from cryptic habitats of the southern Great Barrier Reef (Heron Island and Wistari Reef, Capricorn-Bunker Group, Australia). *J Nat Hist.* 37:1–47.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586–1591.
- Yokobori S, Paabo S. 1995. tRNA editing in metazoans. *Nature* 377:490.
- Yokobori S, Pääbo S. 1997. Polyadenylation creates the discriminator nucleotide of chicken mitochondrial tRNA(Tyr). *J Mol Biol.* 265: 95–99.
- Zrzavy J, Mihulka S, Kepka P, Bezdek A, Tietz D. 1998. Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. *Cladistics* 14:249–285.
- Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31:3406–3415.