Molecular Phylogenetics and Evolution 94 (2016) 196–206

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Evolution of mitochondrial gene order in Annelida ☆

Anne Weigert [a,b,*], Anja Golombek [c], Michael Gerth [a], Francine Schwarz [a], Torsten H. Struck [c,1], Christoph Bleidorn [a,d,1]

[a] Molecular Evolution and Animal Systematics, University of Leipzig, Talstr. 33, 04103 Leipzig, Germany
[b] Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany
[c] Zoological Research Museum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany
[d] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany

ARTICLE INFO

Article history:
Received 22 January 2015
Revised 30 July 2015
Accepted 5 August 2015
Available online 20 August 2015

Keywords:
Annelida
Gene order
Mitochondrial genomes
Next Generation Sequencing

ABSTRACT

Annelida is a highly diverse animal group with over 21,000 described species. As part of Lophotrochozoa, the vast majority of annelids are currently classified into two groups: Errantia and Sedentaria, together forming Pleistoannelida. Besides these taxa, Sipuncula, Amphinomidae, Chaetopteridae, Oweniidae and Magelonidae can be found branching at the base of the tree. Comparisons of mitochondrial genomes have been used to investigate phylogenetic relationship within animal taxa. Complete annelid mitochondrial genomes are available for some Sedentaria and Errantia and in most cases exhibit a highly conserved gene order. Only two complete genomes have been published from the basal branching lineages and these are restricted to Sipuncula. We describe the first complete mitochondrial genome sequences for all other basal branching annelid families: *Owenia fusiformis* (Oweniidae), *Magelona mirabilis* (Magelonidae), *Eurythoe complanata* (Amphinomidae), *Chaetopterus variopedatus* and *Phyllochaetopterus* sp. (Chaetopteridae). The mitochondrial gene order of all these taxa is substantially different from the pattern found in Pleistoannelida. Additionally, we report the first mitochondrial genomes in Annelida that encode genes on both strands. Our findings demonstrate that the supposedly highly conserved mitochondrial gene order suggested for Annelida is restricted to Pleistoannelida, representing the ground pattern of this group. All investigated basal branching annelid taxa show a completely different arrangement of genes than observed in Pleistoannelida. The gene order of protein coding and ribosomal genes in *Magelona mirabilis* differs only in two transposition events from a putative lophotrochozoan ground pattern and might be the closest to an ancestral annelid pattern. The mitochondrial genomes of Myzostomida show the conserved pattern of Pleistoannelida, thereby supporting their inclusion in this taxon.

1. Introduction

Annelida is a major phylum within Lophotrochozoa, whose members occupy a broad range of habitats and are especially abundant in marine environments. This group shows a high diversity in life modes, feeding and reproductive strategies, body forms and developmental patterns (Rouse and Pleijel, 2001). Until recently, relationships among annelid groups were poorly understood, but previous phylogenomic analyses resolved a robust annelid backbone and recovered two major groups comprising the major diversity of Annelida: Errantia and Sedentaria (Andrade et al., 2015; Struck et al., 2015; Struck et al., 2011; Weigert et al., 2014). Additionally, five groups, which are morphological extremely diverse from each other could be found outside of Pleistoannelida: Sipuncula, Amphinomidae, Chaetopteridae, Magelonidae and Oweniidae (Weigert et al., 2014).

To extend our knowledge on annelid evolution and phylogenetic relationships among them, investigation and comparison of mitochondrial gene arrangements is a powerful tool, since rearrangements rarely occur independently in different lineages and closely related species often share identical unchanged gene orders (Boore, 1999; Boore and Brown, 1994). In animals, mitochondrial genomes are usually circular molecules (except in e.g. cnidarians (Bridge et al., 1992) and sponges (Lavrov et al., 2013)), generally

☆ This paper was edited by the Associate Editor Dr. M.A. Arnedo.

* Corresponding author at: Molecular Evolution and Animal Systematics, University of Leipzig, Talstr. 33, 04103 Leipzig, Germany. Fax: +49 341 9736789.
E-mail addresses: anne.weigert@uni-leipzig.de (A. Weigert), A.Golombek@gmx.de (A. Golombek), michael.gerth@uni-leipzig.de (M. Gerth), francine.schwarz@yahoo.de (F. Schwarz), torsten.struck.zfmk@uni-bonn.de (T.H. Struck), bleidorn@uni-leipzig.de (C. Bleidorn).
[1] Shared senior authors.

http://dx.doi.org/10.1016/j.ympev.2015.08.008
1055-7903/

around 16 kb in size, possess only limited intergenic sequences apart from one large non-coding region which is correlated with the origin of replication, and encode for 13 protein-coding genes (PCG), 2 ribosomal RNAs and 22 transfer RNAs (Boore, 1999; Clary and Wolstenholme, 1984; Shadel and Clayton, 1997). The 37 genes can be transcribed either on both strands of the genome or on only one strand.

Lophotrochozoa show a high variability in mitochondrial genomes, including gene number and gene arrangements, strand usage for transcription, repetitive and intergenic regions and unusual modes of inheritance (Boore, 1999; Valles and Boore, 2006; Valles et al., 2008). The mitochondrial gene order in annelids is, unlike in other lophotrochozoan groups, fairly conserved for the families for which gene order has been described so far, especially when not taking tRNA translocations into account (e.g. members of Clitellata, Terebelliformia, Orbiniidae and Phyllodocidae). Exceptions are Sipuncula, Echiura, Ampharetidae, Diurodrilidae and Eunicidae, even though they differ only in a few rearranged genes (or blocks of genes) from the putative annelid ground pattern (Bleidorn et al., 2006; Boore, 2004; Golombek et al., 2013; Jennings and Halanych, 2005; Li et al., 2014; Mwinyi et al., 2009; Shen et al., 2009; Valles and Boore, 2006; Zhong et al., 2008). The only taxon so far completely deviating from this pattern are Syllidae (Errantia), which show completely rearranged mitochondrial genomes (Aguado et al., 2015). Nevertheless, up to now gene rearrangements within Annelida have occurred less often than in other Lophotrochozoa (Boore, 2004; Jennings and Halanych, 2005; Noguchi et al., 2000; Osca et al., 2014; Stechmann and Schlegel, 1999). Additionally, for all annelids from which data is available, genes are described only on one strand of the genome.

To further investigate the putatively conserved mitochondrial gene order evolution in Annelida and to draw a comparison to other lophotrochozoan gene orders, it is crucial to cover mitochondrial genomes from all major annelid groups. So far, ~40 complete mitochondrial genomes are available for annelids covering mainly species in Sedentaria (with 12 of them from clitellates and 10 from Siboglinidae) and species in Errantia. Representing the basal branching lineages, only two complete mitochondrial genomes of Sipuncula species are published (Mwinyi et al., 2009; Shen et al., 2009). In summary, while the majority of annelid taxa in both Sedentaria and Errantia are still not represented, the coverage is much better than for the basal branching lineages, for which there is actually almost no data. Additional information from those basal lineages would provide more insights into mitochondrial genome rearrangements within annelids and help to determine the mitochondrial gene order ground pattern of Annelida.

In this study five new mitochondrial genomes from basal branching annelid families were generated using Illumina-based whole genome shotgun sequencing. Together with the already available mitochondrial genomes of Sipuncula, we covered the complete base of the annelid tree with the taxa *Owenia fusiformis* (Oweniidae), *Magelona mirabilis* (Magelonidae), *Chaetopterus variopedatus* and *Phyllochaetopterus* sp. (Chaetopteridae), and *Eurythoe complanata* (Amphinomidae). Using these data, we investigated the evolution of gene order arrangements in annelids. Moreover, we performed phylogenetic analyses to compare relationships within Annelida inferred by mitochondrial data with the current phylogeny based on transcriptomic data (Andrade et al., 2015; Struck et al., 2015; Struck et al., 2011; Weigert et al., 2014). Our data clearly show a higher variability in mitochondrial gene arrangements in the basal branching lineages in comparison to other annelids and provide additional insights into a putative ancestral mitochondrial gene order pattern for Annelida and Pleistoannelida. Interestingly, *Owenia fusiformis* (Oweniidae) and *Magelona mirabilis* (Magelonidae), representing the lineages which together form the sister taxon of all other annelids, are the only annelids described so far with genes transcribed on both strands of the mitochondrial genome. Especially *Magelona mirabilis* shares a gene order pattern with lophotrochozoans outgroups, which we regard as plesiomorphic. The hitherto reported conserved pattern of Annelida is supported as the ancestral condition for Pleistoannelida (Sedentaria + Errantia).

2. Material and methods

2.1. Taxonomic sampling

Representatives from all basal branching annelid groups were selected according to Weigert et al. (2014). Specimens of *Magelona mirabilis* and *Chaetopterus variopedatus* were collected in Morgat (France), *Owenia fusiformis* in Helgoland (Germany), *Phyllochaetopterus* sp. in Southern New England (USA) and *Eurythoe complanata* was obtained from bought live rock of the Indian ocean kept in the aquarium in Leipzig (Germany). Data for additional annelid families and lophotrochozoan groups were extracted from public resources. Species sampling and accession numbers of all sequences are given in Table 1.

2.2. Library construction, sequencing and raw data processing

Genomic DNA was extracted from a single individual by proteinase K digestion followed by the standard phenol–chloroform extraction (Gustincich et al., 1991). For *Magelona mirabilis*, *Chaetopterus variopedatus* and *Eurythoe complanata*, double-indexed libraries with an average insert size of 350 bp were prepared as described in Meyer and Kircher (2010) and sequenced at the Max Planck Institute for evolutionary Anthropology in Leipzig on the Illumina Hi-Seq 2000 as a 96-bp paired-end run. Base calling was conducted with freeIbis (Renaud et al., 2013), adaptor and primer sequences were removed, reads with low complexity as well as false paired indices were discarded. All three libraries were trimmed by applying a filter of 15, i.e., reads with more than five bases below a phred quality score of 15 were removed. For *Owenia fusiformis* and *Phyllochaetopterus* sp. library construction and sequencing as a 100-bp paired-end run on the Illumina HiSeq as well as quality filtering and adapter trimming with the Chastity filter, was performed by Genterprise Genomics in Mainz (Germany). The quality of all sequences was checked with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and de novo assembly was conducted with CLC Genomics Workbench 7.5 (CLCbio, Arhus, Denmark) with the following settings: mismatch cost 3; insertion cost 3; deletion cost 3; length fraction 0.5; similarity fraction 0.8; minimum contig length 200; automatic word size; automatic bubble size; and contig adjustment by mapped reads. Assemblies were screened for possible (cross) contamination by investigating 18S rRNA gene sequences using local Blast. More information on raw data, including number of reads and contigs are given in Supplementary Table S1.

2.3. Mitochondrial genome annotation and comparison

We annotated the five newly sequenced mitochondrial genomes using MITOS under the mitochondrial code for invertebrate mitochondria (Bernt et al., 2013b) and subsequented curated the annotation manually. To detect and analyse the secondary structure and duplication events of all tRNAs, the program ARWEN (Laslett and Canback, 2008) was applied. For pairwise comparison of the mitochondrial gene order of all basal branching annelids to the most likely ground pattern of Pleistoannelida and Lophotrochozoa, we used the program CREx (Bernt et al., 2007), which reconstructs events for reversals, transpositions, reverse transposi-

Table 1
Source of mitochondrial genomes used for phylogenetic analyses. Asterisks indicate incomplete mitochondrial data, bold taxa represent new mitochondrial data generated in this study.

Phylum	Family	Species	Accession
Annelida	Alvinellidae	*Paralvinella sulfincola**	FJ976042
	Ampharetidae	*Eclysippe vanelli*	EU239687
		Auchenoplax crinita	FJ976041
	Amphinomidae	***Eurythoe complanata***	**KT726962**
	Chaetopteridae	***Chaetopterus variopedatus***	**KT726958**
		***Phyllochaetopterus* sp.**	**KT726961**
	Diurodrilidae	*Diurodrilus subterraneus**	KC790350
	Eunicidae	*Marphysa sanguinea*	KF733802
	Glossiphoniidae	*Helobdella robusta**	AF178680
	Hirudinidae	*Hirudo nipponia*	KC667144
	Lumbricidae	*Lumbricus terrestris*	LTU24570
	Magelonidae	***Magelona mirabilis***	**KT726959**
	Maldanidae	*Clymenella torquata*	AY741661
	Megascolecidae	*Perionyx excavates*	EF494507
	Myzostomida	*Myzostoma seymourcollegiorum**	EF506562
		*Endomyzostoma sp.**	FJ975144
	Nephtyidae	*Nephtys* sp.	EU293739
	Nereididae	*Platynereis dumerilii*	AF178678
		Tylorrhynchus heterochaetus	KM111507.1
		Perinereis aibuhitensis	KF611806
		Perinereis nuntia	JX644015
	Orbiniidae	*Orbinia latreillii*	AY961084
		*Questa ersei**	FJ612452
		Scoloplos cf. *armiger**	DQ517436
	Oweniidae	***Owenia fusiformis***	**KT726960**
	Pectinariidae	*Pectinaria gouldii**	FJ976040
	Phascolosomatidae	*Phascolosoma esculenta*	EF583817
	Siboglinidae	*Galathealinum brachiosum**	AF178679
		*Riftia pachyptila**	AY741662
	Sipunculidae	*Phascolopsis gouldii**	AF374337
		Sipunculus nudus	FJ422961
	Terebellidae	*Pista cristata*	EU239688
	Trichobranchidae	*Terebellides stroemii*	EU236701
	Urechidae	*Urechis caupo*	AY619711
		Urechis unicinctus	EF656365
Mollusca	Haliotoidae	*Haliotis tuberculata*	FJ599667
	Mophaliidae	*Katharina tunicata*	NC_001636
	Sepiidae	*Sepia officinalis*	NC_007895
	Solemyidae	*Solemya velum*	NC_017612
Brachiopoda	Cancellothyrididae	*Terebratulina retusa*	NC_000941
	Laqueidae	*Laqueus rubellus*	NC_002322
Bryozoa	Bugulidae	*Bugula neritina*	NC_010197
	Tubuliporidae	*Tubulipora flabellaris*	NC_015646
	Watersiporidae	*Watersipora subtorquata*	NC_011820
Phoronida	Phoronidae	*Phoronis psammophila*	AY368231
Nemertea	Cephalothricidae	*Cephalothrix simula*	FJ594739
	Emplectonematidae	*Emplectonema gracile*	NC_016952
	Lineidae	*Lineus viridis*	FJ839919
	Nectonemertidae	*Nectonemertes* cf. *mirabilis*	NC_017874

tions and tandem duplication random loss. The analysis was performed by applying the common intervals parameter for distance measurement and only taxa with the complete mitochondrial gene order of protein-coding and ribosomal RNA genes were included (the more variable tRNAs were excluded). For Annelida we included the basal branching annelids, the most likely ground pattern of Pleistoannelida, which is realized in many representatives of Errantia and Sedentaria, as well as orders of Pleistoannelida which differ from that pattern (i.e., Echiura, Eunicidae, Ampharetidae). For Lophotrochozoa we included the most likely ground pattern according to Bernt et al. (2013a) and members of several phyla, as representatives of these phyla differing from that pattern. Moreover, the most likely genome rearrangement scenarios between the gene order of each basal branching annelid and the gene orders of either its sister group, Pleistoannelida or Lophotrochozoa were determined.

2.4. Phylogenetic analyses

For phylogenetic analyses, we generated a data set including all annelid taxa of which all 13 protein-coding genes were available (data set 1) and a data set which also included annelid taxa with partial mitochondrial genomes (data set 2). Data set 1 comprises 40 taxa, including 4 nemerteans, 1 phoronid, 3 bryozoans, 2 brachiopods, 4 molluscs and 26 annelids, data set 2 comprises 48 taxa, including the same out group taxa and 34 annelids (Table 1). For Clitellates, we only included 3 representatives out of the 12 available genomes in all analyses. Sequences for all 13 mitochondrial protein-coding genes were first translated into amino acid sequences from the nucleotide sequences using the mitochondrial code for invertebrates and then independently aligned using MAFFT version 7 (Katoh et al., 2002). For each gene alignment columns containing highly diverse amino acids and many gaps were

masked with REAP (Hartmann and Vision, 2008) and single alignments were concatenated into one data set using FASconCAT version 1.0 (Kück and Meusemann, 2010). Data set 1 covers 3654 amino acid positions and data set 2 3630.

For both datasets, we employed IQ-TREE version 1.3.4 (Nguyen et al., 2015) to determine the best fitting partitioning schemes as well as amino acid substitution models for each of the partitions (Supplementary Table S8). We then performed 10 independent maximum likelihood estimations of both partitioned datasets with RAxML version 8.1.3 (Stamatakis, 2014). Bootstrap support was estimated from 1000 pseudoreplicates. Maximum likelihood phylogenetic estimation was further conducted with IQ-TREE, which, with a different stochastic algorithm, produced topologies congruent to those inferred with RAxML.

In addition, we performed Bayesian phylogenetic analysis with PhyloBayes version 3.3. (Lartillot et al., 2009). First, we tested if for our data sets the site-heterogeneous CAT model implemented in PhyloBayes has a better statistical fit then the single-matrix models used in maximum likelihood analyses. To this end, we performed a comparison of the best-fitting single matrix model implemented in PhyloBayes (mtART) and the CAT models (CAT-GTR and CAT-Poisson) by cross-validation. From both of our datasets, we created 10 learning sets and sampled for 1000 generations for all models under a fixed topology (best maximum likelihood tree). For both datasets, CAT-GTR was supported as model that best fits the data and therefore used in subsequent analyses (Supplementary Table S9). MCMC sampling with PhyloBayes was performed by running two independent chains each for >16,000 cycles, discarding the first 6000 as burnin. All summary variables of all runs were plotted to check for stationarity and convergence. Furthermore, the 'tracecomp' function implemented in PhyloBayes was used to ensure convergence of runs (maximal discrepancy of all variables: 0.3, minimal effective sampling size: 50). Convergence of bipartition frequencies was ensured by using the 'bpcomp' function (maxdif < 0.1). Finally, a consensus tree was constructed from all trees of the posterior sample. Posterior probabilities were inferred from clade frequencies of post-burnin-trees.

3. Results

3.1. Genomic features

Annotations, length and strand position of all genes and RNAs are given in Supplementary Tables S4–S8 and circular genomes are illustrated in Figs. 1 and 2 and Supplementary Figs. S1–S3. Mitochondrial genome size varies from 15,239 bp (*Magelona mirabilis*) to 16,204 bp (*Owenia fusiformis*). For each of the five mitochondrial genomes sequenced in this study, all 13 protein-coding genes, two rRNAs and 22 tRNAs could be detected as typical in most other metazoans. In all five genomes, two tRNAs encoding for serine (tRNA-S1 and -S2) and leucine (tRNA-L1 and -L2) were found. As usual for all other annelids investigated to date, all genes and RNAs are organized on a single strand, the '+' strand, with the only exception of rRNA-T and tRNA-P in *Owenia fusiformis* and *Magelona mirabilis*, which encode both tRNAs on the '–' strand. The mitochondrial genomes of *Chaetopterus variopedatus* and *Eurythoe complanata* both contain 2 copies of the tRNA encoding for methionine, which was verified by visual inspection of the secondary structure. All mitochondrial genomes are deposited in Genbank and accession numbers can be found in Table 1.

3.2. Unassigned, non-coding regions and duplication events

Characteristics of each of the five mitochondrial genomes investigated in this study are the larger intergenic regions which can be found besides the numerous smaller non-coding regions below 100 bp. In addition, copies of the tRNA encoding for methionine

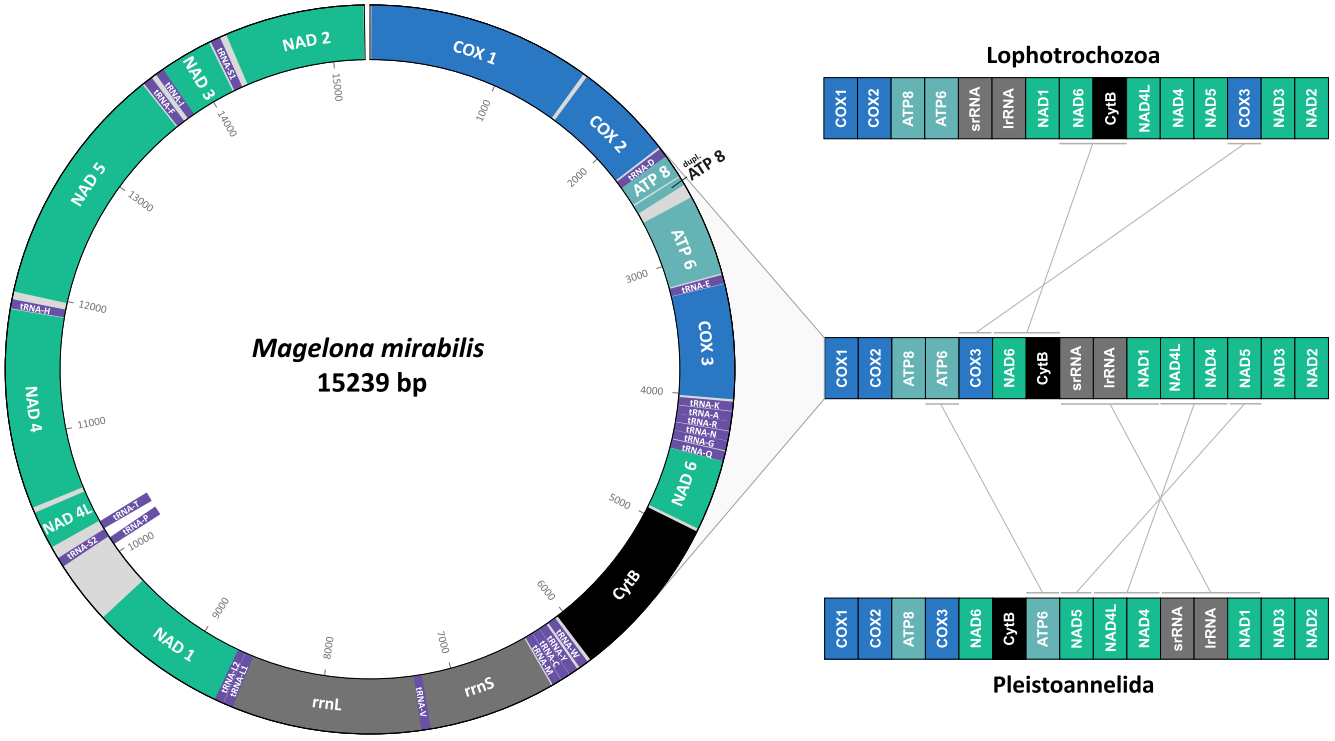

Fig. 1. Gene order of the mitochondrial genome of *Magelona mirabilis* and comparison of mitochondrial genes to the putative ground pattern of Pleistoannelida and Lophotrochozoa. All genes are transcribed on the '+' strand except for tRNA-T and tRNA-P.

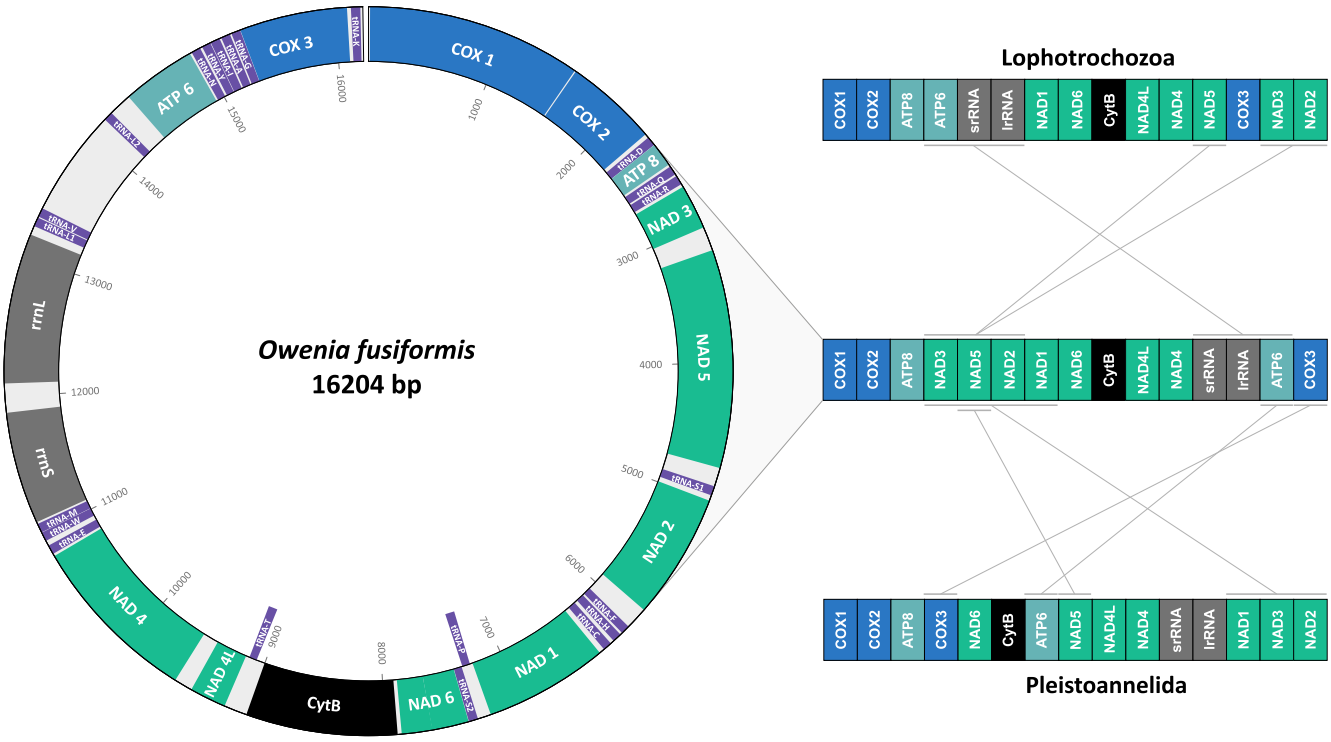

Fig. 2. Gene order of the mitochondrial genome of *Owenia fusiformis* and comparison of mitochondrial genes to the putative ground pattern of Pleistoannelida and Lophotrochozoa. All genes are transcribed on the '+' strand except for tRNA-T and tRNA-P.

can be observed in two species, as well as full or partial gene duplications of COX3, NAD3 and ATP8.

In the mitochondrial genome of *Eurythoe complanata* one large non-coding region of 1121 bp between NAD4 and COX1 could be assigned, as well as two copies of the tRNA encoding for methionine that are directly adjacent to each other (Supplementary Fig. S1). We called the 5′ upstream copy of the "+" strand tRNA-M1 and the 3′ downstream copy tRNA-M2. tRNA-M1 has a shorter TΨC stem with only 3 bases, a shorter TΨC loop with 3 bases and a larger DHU loop with 8 bases, rather than tRNA-M2 with 4 matching bases in the TΨC stem, 4 bases in the TΨC loop and 7 bases in the DHU loop (Supplementary Fig. S4).

For *Chaetopterus variopedatus* two large unassigned regions of 639 bp and 314 bp were found between tRNA-P and CytB (containing a 110 bp long fragment of a presumed duplication of NAD3) and within COX1, respectively (Supplementary Fig. S2). Additionally, two copies of the tRNA encoding for methionine as found in the mitochondrial genome of *Eurythoe complanata* could be assigned. The upstream copy of the "+" strand was named tRNA-M1 and the 3′ downstream copy tRNA-M2. In both copies, the number of matching bases in the TΨC and DHU stem is the same. However, the TΨC loop in tRNA-M2 is shorter than in tRNA-M1 with only 3 bases instead of 5 and the DHU loop of tRNA-M1 is shorter than in tRNA-M2 with only 6 bases instead of 8 (Supplementary Fig. S4).

The annotated mitochondrial genome of *Phyllochaetopterus* sp. contains one region of 208 bp between tRNA-G and ATP6; one region of 350 bp between tRNA-F and COX3 containing a 74 bp long fragment of a COX3 duplication (Supplementary Fig. S3).

For *Magelona mirabilis* two larger intergenic regions can be observed. One of 163 bp between the genes ATP8 and ATP6, containing a 30 bp fragment of a ATP8 duplication, and one of 322 bp between NAD1 and tRNA-P are found (Fig. 1).

Of all mitochondrial genomes analysed in this study the mitochondrial genome of *Owenia fusiformis* contains the most non-coding regions larger than 100 bp. In total, seven unassigned regions could be found: a 173 bp region between NAD3 and NAD5, a 135 bp region between NAD5 and tRNA-S1, a 161 bp region between NAD2 and tRNA-F, a 132 bp region between NAD4L and NAD4, a 209 bp region between the small and large ribosomal RNA, a 167 bp region between tRNA-L2 and ATP6 and the largest non-coding region of 773 bp between tRNA-V and tRNA-L2 (Fig. 2).

3.3. Phylogenetic analyses

In all analyses of the two data sets monophyletic Annelida could not be recovered since *Owenia fusiformis* groups with Nemertea (Fig. 3, Supplementary Figs. S5 and S6). Additionally, in the ML analysis including complete and partial annelid mitochondrial genomes (data set 2), *Magelona mirabilis* groups together with Oweniidae as sister group to Nemertea and Chaetopteridae are sister group to Brachiopoda (Supplementary Fig. S6).

Despite the position of Oweniidae, in the ML and BI analyses including only complete mitochondrial genomes (data set 1) the phylogenetic relationships within Annelida are in agreement with previous molecular analyses based on transcriptomes (Andrade et al., 2015; Struck et al., 2015; Weigert et al., 2014) by recovering monophyletic Errantia (BS = 69, PP = 1), Sedentaria (BS = 26, PP = 0.99) and Pleistoannelida (BS = 25, PP = 0.99), as well as the basal branching annelids (Fig. 3A, Supplementary Fig. S5). However, for the deeper nodes the support values are very low, but the topologies of both analyses are very similar, except for the position of Amphinomida and sister group relationships within Sedentaria.

In the ML and BI analyses including additional partial mitochondrial genomes (data set 2) annelid relationships as described above could not be resolved, except for monophyletic Errantia (without Myzostomida; BS = 74, PP = 1) and monophyletic Pleistoannelida (only BI analysis, PP = 0.89). The trees differ significantly in the positions of Errantia, Sipuncula, Amphinomidae,

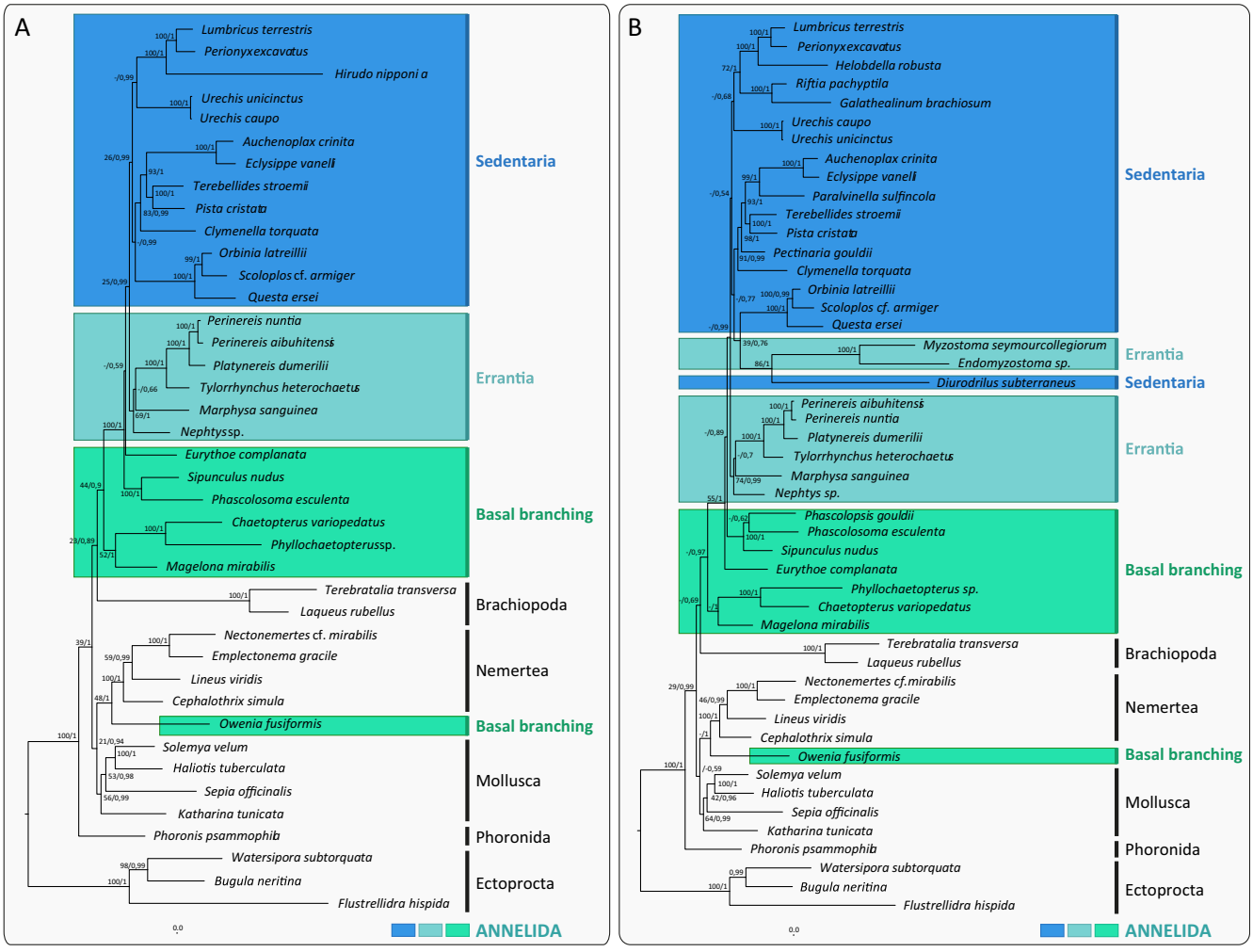

Fig. 3. Phylogenetic relationships of Annelida based on mitochondrial genome data. (A) Consensus tree of the Bayesian analysis using the CAT-GTR model of data set 1 comprising only complete mitochondrial genomes (40 taxa, 3654 amino acid positions). (B) Consensus tree of the Bayesian analysis using the CAT-GTR model of data set 2 comprising complete and partial annelid mitochondrial genomes (48 taxa, 3630 amino acid positions). Bootstrap support values (BS) where the topology of the Maximum Likelihood and Bayesian analyses agree are depicted before the posterior probabilities (PP) or indicated with a hyphen if not so.

Myzostomida, Diurodrilidae, Chaetopteridae, and Magelonidae and the support for deeper nodes is also very low (Fig. 3B, Supplementary Fig. S6). In both analyses *Diurodrilus* sp. and the two myzostomids group within Annelida as sister group to Orbiniidae (Fig. 3B, Supplementary Fig. S6). Amphinomida and Sipuncula branch off early only in the BI analysis, whereas in the ML analysis they group with Errantia rendering Pleistoannelida paraphyletic (Fig. 3B, Supplementary Fig. S6).

3.4. Mitochondrial gene order and rearrangements

The gene order of each of the five mitochondrial genomes of the basal branching annelids differ significantly from each other and show a high variability in contrast to gene orders from annelids belonging to either Errantia or Sedentaria (Fig. 4). The latter groups generally share a conserved mitochondrial ground pattern (Figs. 4 and 5). In each scenario reconstructed with CREx several transpositions, reverse transpositions, reversals or tandem-duplication-random-loss (tdrl) events are necessary to rearrange genes of the basal branching annelids in comparison to the ground pattern of Lophotrochozoa or Pleistoannelida (Fig. 5, Supplementary Table S9). For all rearrangement scenarios, tRNAs were not compared due to their higher variability in location. Nevertheless, the location of tRNA-P and tRNA-T in *Owenia fusiformis* and *Magelona mirabilis* has to be highlighted, since they are the only known genes/RNAs within Annelida which are encoded on the '—' strand. In general, the highest number of similarities in mitochondrial genome organization within annelids can be found between the putative ground pattern of Pleistoannelida and members of Pleistoannelida which differ in that pattern (Eunicidae and Ampharetidae), except for Echiura (Fig. 5).

The gene order of *Eurythoe complanata* is more similar to the one found in Pleistoannelida and closely related families (Sipuncula and Chaetopteridae) than to Lophotrochozoan taxa and the most basal branching annelids *Owenia fusiformis* and *Magelona mirabilis* (except for the pattern found in *Katharina tunicata* and *Chaetopterus variopedatus*). Most similarities can be found in the pattern of Sipuncula, *Phyllochaetopterus* sp., Pleistoannelida (ground pattern and *Marphysa sanguinea*) and *Katharina tunicata* (Fig. 5). The genome differences to its sister group Sipuncula can be reconstructed with three transposition events, to Pleistoannelida with one tdrl and one transposition event, and to *Phyllochaetopterus* sp. with two transposition, one reverse transposition, one reversal and one tdrl event (Fig. 5, Supplementary Table S9).

The fewest events in gene order rearrangement can be observed between *Chaetopterus variopedatus* and *Phyllochaetopterus* sp. with the only difference being a transposition of the CytB-NAD4L-NAD4-NAD5 cluster (Supplementary Table S9). The gene order of *Chaetopterus variopedatus* is more similar to the hypothetical ground pattern of Lophotrochozoa and other lophotrochozoans

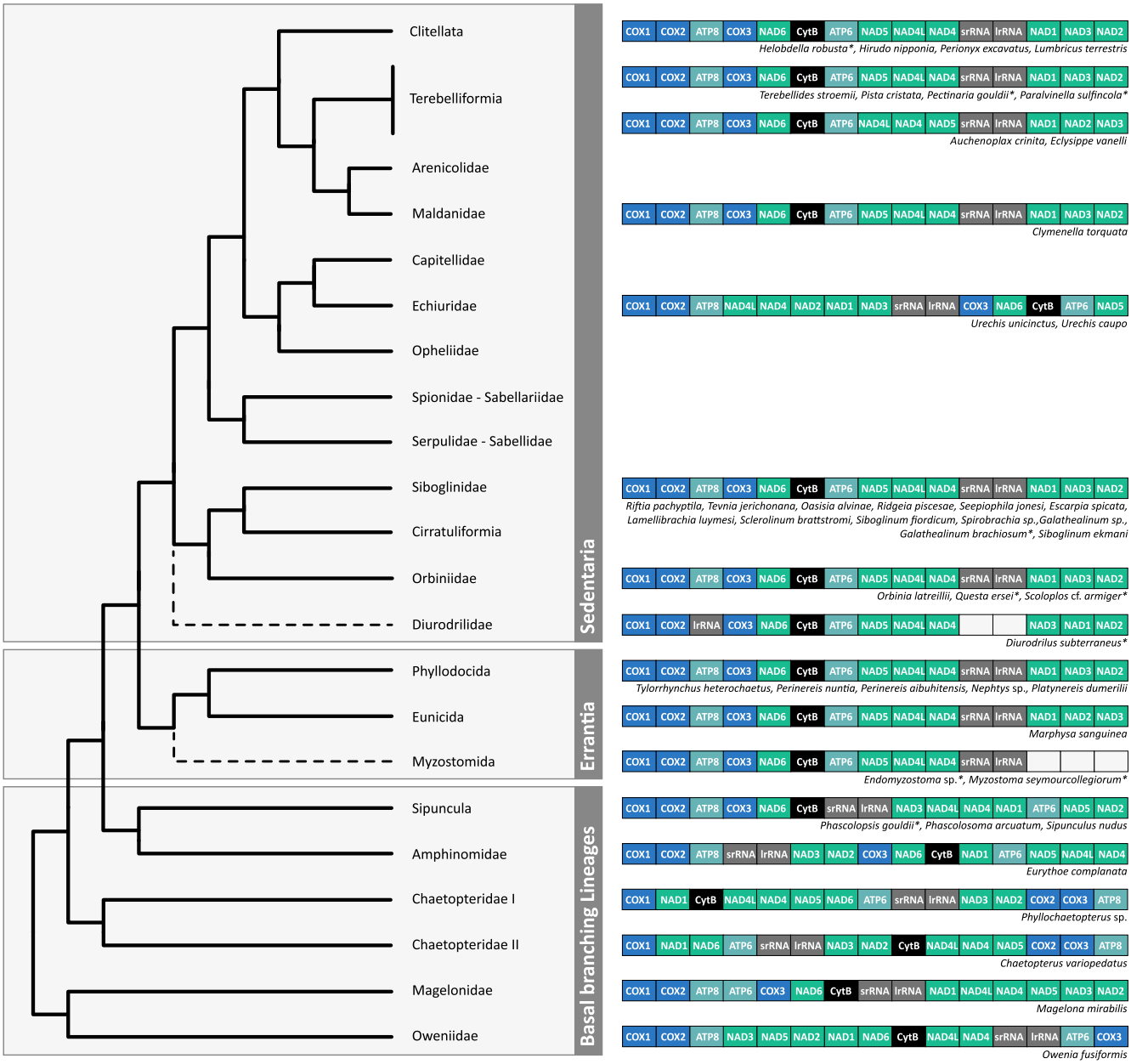

Fig. 4. Relationships within Annelida and different mitochondrial gene order of each taxon. Annelid phylogeny is depicted based on Weigert et al. (2014), sister group relationships of families not represented in Weigert et al. (2014) are obtained from Struck et al. (2007) (Maldanidae and Ampharetidae) and Golombek et al. (2013) (Diurodrilidae). Dashed lines indicate an uncertain phylogenetic position. Only protein-coding genes of available mitochondrial genomes and ones, which were generated in this study, are included. Taxa with partial mitochondrial genomes are marked with an asterisk and missing genes within the mitochondrial gene order are indicated with gray boxes. Missing genes are: ATP8 and srRNA (Diurodrilidae); NAD1, NAD2 and NAD3 (Myzostomida). Genes are not scaled to real length and are indicated by standard abbreviations.

than to any other annelid pattern (except for *Phyllochaetopterus* sp.) differing in only three transposition events (Fig. 5, Supplementary Table S9). For the gene order of *Phyllochaetopterus* sp. the similarity to other Lophotrochozoa is even higher and gene rearrangements can be reconstructed with four transposition events (Supplementary Table S9). From all annelids the two chaetopterids share the most similarities with the amphinomid *Eurythoe complanata*, which is also part of the basal radiation.

Of all basal branching taxa, the mitochondrial pattern of *Magelona mirabilis* and Sipuncula show the most similarities in mitochondrial gene order with the putative ground pattern of Pleistoannelida (Fig. 5). Interestingly, *Magelona mirabilis* shares as many similarities with gene arrangements found in members of Mollusca, Phoronida and the ground pattern of Lophotrochozoa, differing in only two rearrangement events (Supplementary Table S9, Fig. 5). The differences of the mitochondrial gene order to the one of its sister group *Owenia fusiformis* are much higher than to most other Pleistoannelida (Fig. 5, Supplementary Table S9).

The mitochondrial gene order of *Owenia fusiformis* shares the most similarities with the one found in the ground pattern of Lophotrochozoa and *Phoronis psammophila* (Fig. 5). The organization of the mitochondrial genome of *Owenia fusiformis* differs in one transposition, one reverse transposition, one reversal and two tdrl events to Pleistoannelida, and four reversals and two transposition events to Lophotrochozoa (Supplementary Table S9).

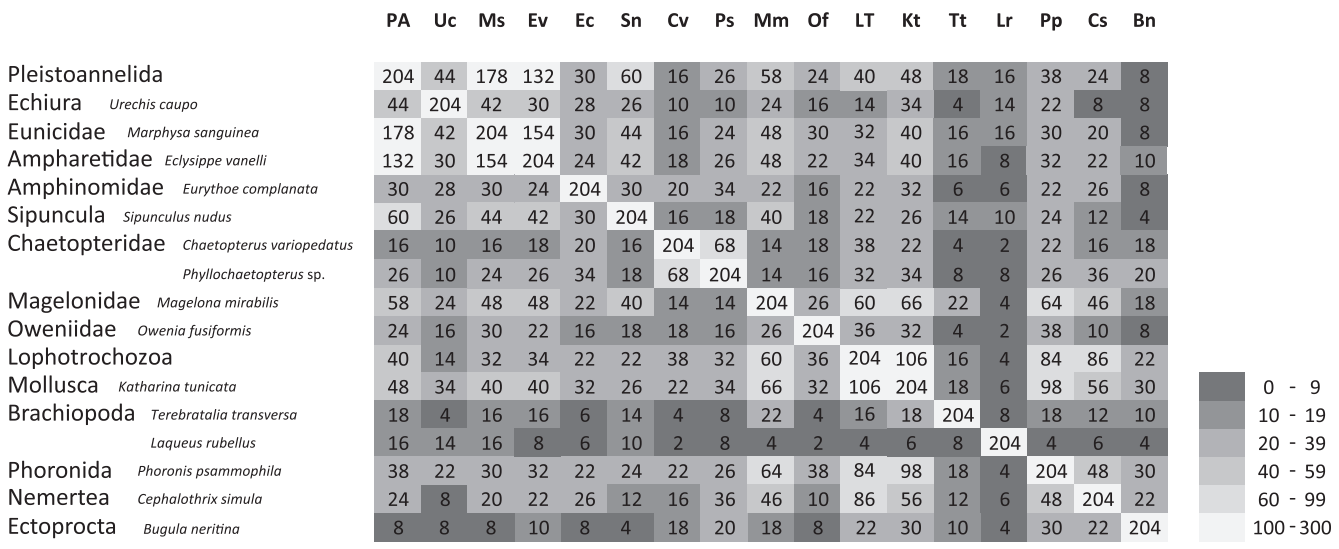

Fig. 5. Results of the pairwise comparisons of mitochondrial gene orders of basal branching annelids with the putative ground pattern of Pleistoannelida and Lophotrochozoa as well as with differing annelid and lophotrochozoan members. Scores of the CREx analysis of each pairwise comparison indicate the similarities of the compared mitochondrial gene orders, where 204 is the highest score and represents identical gene order. Only taxa with the complete mitochondrial gene order of protein-coding genes were included in the analysis. The gene order of the putative ground pattern of Pleistoannelida is identical with the one found in Clitellata (*Helobdella robusta, Hirudo nipponia, Perionyx excavates* and *Lumbricus terrestris)*, Terebelliformia (*Terebellides stroemii*, *Pista cristata*, *Pectinaria gouldii* and *Paralvinella sulfincola,* except Ampharetidae), Maldanidae (*Clymenella torquata*), Siboglinidae (*Riftia pachyptila*), Orbiniidae (*Orbinia latreillii, Questa ersei* and *Scoloplos* cf. *armiger*), Phyllodocida (*Tylorrhynchus heterochaetus, Perinereis nuntia, Perinereis aibuhitensis, Nephtys* sp. and *Platynereis dumerilii*) and Myzostomida (*Endomyzostoma* sp. and *Myzostoma seymourcollegiorum*). Abbreviated taxa: Bn – Bugula neritina, Cs – *Cephalothrix simula*, Cv – *Chaetopterus variopedatus*, Ev – *Eclysippe vanelli*, Ec – *Eurythoe complanata*, Kt – *Katharina tunicata*, Lr – *Laqueus rubellus*, LT – Lophotrochozoa, Ms – *Marphysa sanguinea*, Mm – *Magelona mirabilis*, Of – *Owenia fusiformis*, PA – Pleistoannelida, Pp – *Phoronis psammophila*, Ps – *Phyllochaetopterus* sp., Sn – *Sipunculus nudus*, Tt – *Terebratalia transversa*, Uc – *Urechis caupo*.

4. Discussion

4.1. Genome organization and structural features

Mitochondrial genes in Annelida are generally transcribed only from one strand. However, *Owenia fusiformis* and *Magelona mirabilis,* are the first known annelids where not all of the 37 genes are transcribed from one single strand, since the two tRNAs encoding for proline (tRNA-P) and threonine (tRNA-T) are located on the opposite strand (Figs. 1 and 2). Boore (1999) suggested a "ratchet effect" for the scenario in which by chance all genes were placed on one strand and that transcription therefore would sooner or later be lost for one of the two strands, hindering further inversion without additional transcription elements. This scenario was proposed for the last common ancestor (LCA) of annelids and might be the most parsimonious hypotheses, if basal branching lineages would also show the same transcription direction for all genes (Boore, 1999; Valles and Boore, 2006). However, *Owenia fusiformis* and *Magelona mirabilis* are the exception to the typical annelid pattern. Both share the same tRNAs, which were placed on the opposite strand, presumably in the LCA of both families (which together form the sister taxon of all other annelids based on Weigert et al. (2014)). There are two hypotheses to interpret this pattern: (1) The LCA of Annelida had all genes on one strand and already lost the transcription signal on the other strand. With the inversion event of the two tRNAs in *Owenia fusiformis* and *Magelona mirabilis* necessary elements for transcription were also transposed. (2) The LCA of Annelida still had transcription signals on both strands and it was lost on one strand in the lineage leading to the rest of the Annelida, which forms the sister taxon of the clade comprising Oweniidae and Magelonidae. In this case, the LCA of Annelida likely possessed the same mitochondrial pattern as observed in Oweniidae and Magelonidae (all genes except for tRNA-P and tRNA-T are located on one strand), and both tRNAs were inverted on the opposite strand after the split of Oweniidae/Magelonidae from the rest of the annelid lineages. Nevertheless, if the strand usage and inversion of both tRNAs is a plesiomorphic condition in annelids rather than a synapomorphy for Oweniidae and Magelonidae, there should be traces in other lophotrochozoan groups. The mitochondrial genes in brachiopods are generally encoded on one strand (Helfenbein et al., 2001) and in molluscs there is no general pattern. Different families of molluscs show a high diversity in strand usage and gene order (Osca et al., 2014). Interestingly, all but one mitochondrial genome of Nemertea, which are either sister group to Annelida or very closely related to them (Dunn et al., 2008; Laumer et al., 2015; Weigert et al., 2014), show a similar pattern to that found in *Owenia fusiformis* and *Magelona mirabilis,* where all genes are transcribed from one strand except for the two tRNAs encoding for threonine and proline (e.g. Chen et al., 2012; Chen et al., 2011). This shared strand usage and identical pattern in the two tRNAs between Nemerteans and the basal branching annelid families Oweniidae and Magelonidae favors hypothesis 2 suggesting that this pattern is the ancestral condition for annelids.

Another feature which could not be observed in annelids to date is the conjecture of the ATP8 and ATP6 gene, usually a common order in nearly all animal mitochondrial genomes with exception in members of certain lophotrochozoan phyla. In some species of Mollusca, Brachiopoda, Nemertea and Phoronida this gene boundary is disrupted (Boore, 2006; Chen et al., 2012; Helfenbein and Boore, 2004; Helfenbein et al., 2001; Noguchi et al., 2000). In Platyhelminthes, Acoelomorpha and Acanthocephala, the ATP8 gene is missing (Mwinyi et al., 2010; Steinauer et al., 2005; Valles and Boore, 2006). In Chaetognatha both ATP8 and ATP6 are missing (Papillon et al., 2004). *Magelona mirabilis* is the only annelid described so far which retained ATP8 adjacent to ATP6. This increases the likelihood of the ATP8–ATP6 conjecture as part of the annelid mitochondrial ground pattern, where the loss of this gene boundary might have occurred independently in Oweniidae and the LCA of Pleistoannelida + Chaetopteridae + Sipuncula + Amphinomidae, as observed among various other phyla.

4.2. Phylogenetic relationships based on mitochondrial sequence data

Reconstructing robust annelid relationships with morphological or few molecular markers failed in the past, recovering trees with paraphyletic Annelida or general low support values and lack of resolution (for review see (Struck, 2012). With the advent of new sequencing techniques including sequencing cost reduction, amplification of the relatively small mitochondrial genome became easily feasible and mitochondrial data an increasingly useful tool for investigating phylogenies. Nevertheless, robust relationships of higher ranked annelid groups and deeper splits could not be resolved by incorporating those data, whereas their application in affiliating uncertain taxa to Annelida (e.g. Diurodrilidae, Echiura, Siboglinidae, and Sipuncula) provided additional support and more stable results (Boore and Brown, 2000; Boore and Staton, 2002; Golombek et al., 2013; Jennings and Halanych, 2005; Mwinyi et al., 2009; Shen et al., 2009; Wu et al., 2009).

Our analyses yielded similarly unstable results. By analysing only complete mitochondrial genomes (data set 1), we found similar annelid relationships as proposed in recent molecular analyses (e.g. Weigert et al., 2014), with Pleistoannelida comprising the reciprocal monophyletic Sedentaria and Errantia. The basal branching lineages group as well in the basal part of the tree, except of *Owenia fusiformis*, rendering Annelida not monophyletic (Fig. 3A, Supplementary Fig. S5). However, with the inclusion of additional partial genomes (data set 2), these relationships cannot be reconstructed (Fig. 3B, Supplementary Fig. S6). Pleistoannelida and Sedentaria form non-monophyletic groups, which can likely be explained due to the influence of long branch attraction (LBA) introduced by faster evolving taxa like myzostomids and diurodrilids. Similar problems with mitochondrial genomes of these two taxa have been already reported before (Golombek et al., 2013). Recent studies based on transcriptomes tend to group Myzostomida within Pleistoannelida, either as part or sister group to Errantia or within Sedentaria (Andrade et al., 2015; Weigert et al., 2014) and *Diurodrilus* within Sedentaria (Andrade et al., 2015; Laumer et al., 2015; Struck et al., 2015), which is also supported in our analyses (Fig. 3B, Supplementary Fig. S6). However, our findings are in congruence with a proposed basal branching position of Sipuncula, Amphinomidae, Chaetopteridae, Oweniidae, and Magelonidae. As such it comes without surprise that these taxa are drawn to outgroup taxa in analyses based on only few molecular markers, e.g. mitochondrial genomes, as the presumed divergences date back into the Cambrian (Weigert et al., 2014). The lack of resolution of such datasets for deep metazoan phylogeny has been already demonstrated by Bernt et al. (2013a). Further investigations on phylogenetic relationships in major animal groups as old as annelids should be based on more suitable and a higher amount of molecular and morphological data.

4.3. Mitochondrial genome rearrangements in Annelida

Whereas the reconstruction of ancient relationships with mitochondrial sequences has its limits, comparison of mitochondrial gene order seems promising, since gene rearrangements seldom occur convergently and closely related species often share the same gene order (Boore, 1999). Furthermore, questionable assignments of taxa to certain groups can be investigated by comparing mitochondrial gene orders which thereby serve as an additional marker if morphology or molecular data is highly controversial or lacking, as previously demonstrated for Sipuncula (Mwinyi et al., 2009) and Diurodrilidae (Golombek et al., 2013). In contrast to closely related phyla like molluscs, brachiopods and nemerteans, Annelida were believed to possess a highly conserved mitochondrial gene order, given the available data (Jennings and Halanych, 2005; Valles and Boore, 2006). Our results demonstrate that this hypothesis is clearly restricted to members belonging to Pleistoannelida (Errantia and Sedentaria) and it is parsimonious to assume that this conserved order (even true for most tRNAs) represents the ground pattern for this clade. When comparing the mitochondrial gene order of Myzostomida with basal branching annelids or Pleistoannelida, the position of myzostomids is in congruence with recent molecular analyses as part of Pleistoannelida (Andrade et al., 2015; Weigert et al., 2014) instead of being part of the base of the annelid tree (Struck et al., 2011), since they exhibit the same conserved arrangement of genes as seen in members of Errantia and Sedentaria. In contrast, all represented taxa branching from the base of the annelid tree show a completely different arrangement of genes than observed in Pleistoannelida, but the reconstruction of a putative annelid mitochondrial ground pattern still remains difficult. However, as already mentioned, it seems very likely that all genes were encoded on one strand except for tRNA-P and tRNA-T, as found in the two annelids *Owenia fusiformis* and *Magelona mirabilis* and in the ground pattern of Nemertea (Chen et al., 2012, 2011). Additionally, there are a few conserved blocks of genes in annelids that, with a few exceptions, can also be found in the putative ground pattern of Lophotrochozoa and Bilateria (Bernt et al., 2013a; Lavrov and Lang, 2005) and might be represented in the annelid ground pattern: COX1-COX-2-ATP8, NAD6-CYTB, SrRNA-LrRNA and NAD4L-NAD4. From all basal branching annelids, the gene order of *Magelona mirabilis* is most similar to the putative ground pattern of Lophotrochozoa (Bernt et al., 2013a), differing only in 2 transposition events and exhibiting a high number of identical gene blocks including tRNAs with conserved bilaterian gene blocks (Bernt et al., 2013a): NAD6-CytB-S2 (without tRNA-S2), SrRNA-V-LrRNA-L1-L2-NAD1, NAD4L-NAD4-H-NAD5, NAD2-COX1-COX2-K-ATP8-ATP6-COX3 (tRNA-D instead of tRNA-K, tRNA-E between ATP6-COX3). It is tempting to assume that the pattern of *Magelona mirabilis* is similar to the ancestral pattern for Annelida and this might be also close to the lophotrochozoan ground pattern. The fact that lophotrochozoan taxa have convergently lost and rearranged mitochondrial genes in numerous ways resulted in a blurry picture of the mitochondrial ground pattern of this group. However, our mitochondrial data on early branching annelids is an important contribution to understand evolutionary relationships within Annelida and perhaps of putative sister groups and demonstrates that annelids fall in line with other lophotrochozoan animal groups regarding the high variability in mitochondrial gene rearrangements at least in their basal radiation.

Abbreviations

NAD1-6, 4L	NADH dehydrogenase subunits 1-6 and 4L
COX1-3	cytochrome oxidase subunits 1-3
CYTB	cytochrome b
ATP6/8	ATP synthase subunit 6/8
lrRNA	large ribosomal RNA
srRNA	small ribosomal RNA
mt	mitochondrial
single letters	tRNAs encoding for amino acids: alanine (A), cysteine (C), aspartic acid (D), glutamatic acid (E), phenylalanine (F), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L1 and L2), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S1 and S2), threonine (T), valine (V), tryptophan (W), tyrosine (Y)

Acknowledgments

We are thankful to Conrad Helm (University of Leipzig, Germany) and the staff of the Station Biologique de Roscoff (France) for supplying the specimens and for providing laboratory space. We additionally thank Martin Schlegel (University of Leipzig, Germany) for providing materials and facilities and Sarah Tobergte (University of Osnabrück, Germany) for assistance in the laboratory. This study was supported by the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, by the DFG grant BL787/5-1 to CB, DFG-STR-683/6-1, 6-2, 8-1 & 8-2 to THS and the EU due to ASSEMBLE grant agreement no. 227799 to CB (http://www.assemblemarine.org).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2015.08.008.

References

Aguado, M.T., Glasby, C.J., Schroeder, P.C., Weigert, A., Bleidorn, C., 2015. The making of a branching annelid: an analysis of complete mitochondrial genome and ribosomal data of Ramisyllis multicaudata. Sci. Rep. 5, 12072.

Andrade, S.C.S., Novo, M., Kawauchi, G.Y., Worsaae, K., Pleijel, F., Giribet, G., Rouse, G.W., 2015. Articulating "archiannelids": phylogenomics and annelid relationships, with emphasis on meiofaunal taxa. Mol. Biol. Evol., in press.

Bernt, M., Bleidorn, C., Braband, A., Dambach, J., Donath, A., Fritzsch, G., Golombek, A., Hadrys, H., Juhling, F., Meusemann, K., Middendorf, M., Misof, B., Perseke, M., Podsiadlowski, L., von Reumont, B., Schierwater, B., Schlegel, M., Schrodl, M., Simon, S., Stadler, P.F., Stoger, I., Struck, T.H., 2013a. A comprehensive analysis of bilaterian mitochondrial genomes and phylogeny. Mol. Phylogenet. Evol. 69, 352–364.

Bernt, M., Donath, A., Juhling, F., Externbrink, F., Florentz, C., Fritzsch, G., Putz, J., Middendorf, M., Stadler, P.F., 2013b. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69, 313–319.

Bernt, M., Merkle, D., Ramsch, K., Fritzsch, G., Perseke, M., Bernhard, D., Schlegel, M., Stadler, P.F., Middendorf, M., 2007. CREx: inferring genomic rearrangements based on common intervals. Bioinformatics 23, 2957–2958.

Bleidorn, C., Podsiadlowski, L., Bartolomaeus, T., 2006. The complete mitochondrial genome of the orbiniid polychaete *Orbinia latreillii* (Annelida, Orbiniidae) – a novel gene order for Annelida and implications for annelid phylogeny. Gene 370, 96–103.

Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27, 1767–1780.

Boore, J.L., 2004. Complete mitochondrial genome sequence of *Urechis caupo*, a representative of the phylum Echiura. BMC Genom. 5, 67.

Boore, J.L., 2006. The complete sequence of the mitochondrial genome of *Nautilus macromphalus* (Mollusca: Cephalopoda). BMC Genom. 7, 182.

Boore, J.L., Brown, W.M., 1994. Mitochondrial genomes and the phylogeny of mollusks. Nautilus 108, 61–78.

Boore, J.L., Brown, W.M., 2000. Mitochondrial genomes of *Galathealinum*, *Helobdella*, and *Platynereis*: sequence and gene arrangement comparisons indicate the Pogonophora is not a phylum and Annelida and Arthropoda are not sister taxa. Mol. Biol. Evol. 17, 988.

Boore, J.L., Staton, J.L., 2002. The mitochondrial genome of the sipunculid *Phascolopsis gouldii* supports its association with Annelida rather than Mollusca. Mol. Biol. Evol. 19, 127–137.

Bridge, D., Cunningham, C.W., Schierwater, B., Desalle, R., Buss, L.W., 1992. Class-level relationships in the phylum Cnidaria – evidence from mitochondrial genome structure. Proc. Natl. Acad. Sci. USA 89, 8750–8753.

Chen, H.X., Sun, S.C., Sundberg, P., Ren, W.C., Norenburg, J.L., 2012. A comparative study of nemertean complete mitochondrial genomes, including two new ones for *Nectonemertes* cf. *mirabilis* and *Zygeupolia rubens*, may elucidate the fundamental pattern for the phylum Nemertea. BMC Genom. 13, 139.

Chen, H.X., Sundberg, P., Wu, H.Y., Sun, S.C., 2011. The mitochondrial genomes of two nemerteans, *Cephalothrix* sp. (Nemertea: Palaeonemertea) and *Paranemertes* cf. *peregrina* (Nemertea: Hoplonemertea). Mol. Biol. Rep. 38, 4509–4525.

Clary, D.O., Wolstenholme, D.R., 1984. The *Drosophila* mitochondrial genome. Oxf. Surv. Eukaryot. Genes 1, 1–35.

Dunn, C.W., Hejnol, A., Matus, D.Q., Pang, K., Browne, W.E., Smith, S.A., Seaver, E., Rouse, G.W., Obst, M., Edgecombe, G.D., Sorensen, M.V., Haddock, S.H., Schmidt-Rhaesa, A., Okusu, A., Kristensen, R.M., Wheeler, W.C., Martindale, M.Q., Giribet, G., 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452, 745–749.

Golombek, A., Tobergte, S., Nesnidal, M.P., Purschke, G., Struck, T.H., 2013. Mitochondrial genomes to the rescue – Diurodrilidae in the myzostomid trap. Mol. Phylogenet. Evol. 68, 312–326.

Gustincich, S., Manfioletti, G., Delsal, G., Schneider, C., Carninci, P., 1991. A fast method for high-quality genomic DNA extraction from whole human blood. Biotechniques 11, 298.

Hartmann, S., Vision, T.J., 2008. Using ESTs for phylogenomics: can one accurately infer a phylogenetic tree from a gappy alignment? BMC Evol. Biol. 8, 95.

Helfenbein, K.G., Boore, J.L., 2004. The mitochondrial genome of *Phoronis architecta* – comparisons demonstrate that phoronids are Lophotrochozoan protostomes. Mol. Biol. Evol. 21, 153–157.

Helfenbein, K.G., Brown, W.M., Boore, J.L., 2001. The complete mitochondrial genome of the articulate brachiopod *Terebratalia transversa*. Mol. Biol. Evol. 18, 1734–1744.

Jennings, R.M., Halanych, K.M., 2005. Mitochondrial genomes of *Clymenella torquata* (Maldanidae) and *Riftia pachyptila* (Siboglinidae): evidence for conserved gene order in Annelida. Mol. Biol. Evol. 22, 210–222.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Kück, P., Meusemann, K., 2010. FASconCAT: convenient handling of data matrices. Mol. Phylogenet. Evol. 56, 1115–1118.

Lartillot, N., Lepage, T., Blanquart, S., 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25, 2286–2288.

Laslett, D., Canback, B., 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24, 172–175.

Laumer, C.E., Bekkouche, N., Kerbl, A., Goetz, F., Neves, R.C., Sorensen, M.V., Kristensen, R.M., Hejnol, A., Dunn, C.W., Giribet, G., Worsaae, K., 2015. Spiralian phylogeny informs the evolution of microscopic lineages. Curr. Biol. 25, 2000–2006.

Lavrov, D.V., Lang, B.F., 2005. Poriferan mtDNA and animal phylogeny based on mitochondrial gene arrangements. Syst. Biol. 54, 651–659.

Lavrov, D.V., Pett, W., Voigt, O., Worheide, G., Forget, L., Lang, B.F., Kayal, E., 2013. Mitochondrial DNA of Clathrina clathrus (Calcarea, Calcinea): six linear chromosomes, fragmented rRNAs, tRNA editing, and a novel genetic code. Mol. Biol. Evol. 30, 865–880.

Li, S., Chen, Y., Zhang, M., Bao, X., Li, Y., Teng, W., Liu, Z., Fu, C., Wang, Q., Liu, W., 2014. Complete mitochondrial genome of the marine polychaete, Marphysa sanguinea (Polychaeta, Eunicida). Mitochondr DNA, in press.

Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010.

Mwinyi, A., Bailly, X., Bourlat, S.J., Jondelius, U., Littlewood, D.T.J., Podsiadlowski, L., 2010. The phylogenetic position of Acoela as revealed by the complete mitochondrial genome of *Symsagittifera roscoffensis*. BMC Evol. Biol. 10, 309.

Mwinyi, A., Meyer, A., Bleidorn, C., Lieb, B., Bartolomaeus, T., Podsiadlowski, L., 2009. Mitochondrial genome sequence and gene order of *Sipunculus nudus* give additional support for an inclusion of Sipuncula into Annelida. BMC Genom. 10, 27.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Noguchi, Y., Endo, K., Tajima, F., Ueshima, R., 2000. The mitochondrial genome of the brachiopod *Laqueus rubellus*. Genetics 155, 245–259.

Osca, D., Irisarri, I., Todt, C., Grande, C., Zardoya, R., 2014. The complete mitochondrial genome of *Scutopus ventrolineatus* (Mollusca: Chaetodermomorpha) supports the Aculifera hypothesis. BMC Evol. Biol. 14, 197.

Papillon, D., Perez, Y., Caubit, X., Le Parco, Y., 2004. Identification of chaetognaths as protostomes is supported by the analysis of their mitochondrial genome. Mol. Biol. Evol. 21, 2122–2129.

Renaud, G., Kircher, M., Stenzel, U., Kelso, J., 2013. freeIbis: an efficient basecaller with calibrated quality scores for Illumina sequencers. Bioinformatics 29, 1208–1209.

Rouse, G.W., Pleijel, F., 2001. Polychaetes. Oxford University Press.

Shadel, G.S., Clayton, D.A., 1997. Mitochondrial DNA maintenance in vertebrates. Annu. Rev. Biochem. 66, 409–435.

Shen, X., Ma, X.Y., Ren, J.F., Zhao, F.Q., 2009. A close phylogenetic relationship between Sipuncula and Annelida evidenced from the complete mitochondrial genome sequence of Phascolosoma esculenta. BMC Genom. 10, 136.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Stechmann, A., Schlegel, M., 1999. Analysis of the complete mitochondrial DNA sequence of the brachiopod *Terebratulina retusa* places Brachiopoda within the protostomes. Proc. Roy. Soc. B – Biol. Sci. 266, 2043–2052.

Steinauer, M.L., Nickol, B.B., Broughton, R., Orti, G., 2005. First sequenced mitochondrial genome from the phylum Acanthocephala (*Leptorhynchoides thecatus*) and its phylogenetic position within metazoa. J. Mol. Evol. 60, 706–715.

Struck, T.H., 2012. Phylogeny of annelida. In: Schmidt-Rhaesa, A. (Ed.), Handbook of Zoology Online. De Gruyter, Berlin, Boston.

Struck, T.H., Golombek, A., Weigert, A., Franke, F.A., Westheide, W., Purschke, G., Bleidorn, C., Halanych, K.M., 2015. The evolution of annelids reveals two adaptive routes to the interstitial realm. Curr. Biol. 25, 1993–1999.

Struck, T.H., Paul, C., Hill, N., Hartmann, S., Hosel, C., Kube, M., Lieb, B., Meyer, A., Tiedemann, R., Purschke, G., Bleidorn, C., 2011. Phylogenomic analyses unravel annelid evolution. Nature 471, 95–98.

Struck, T.H., Schult, N., Kusen, T., Hickman, E., Bleidorn, C., McHugh, D., Halanych, K. M., 2007. Annelid phylogeny and the status of Sipuncula and Echiura. BMC Evol. Biol. 7, 57.

Valles, Y., Boore, J.L., 2006. Lophotrochozoan mitochondrial genomes. Integr. Comp. Biol. 46, 544–557.
Valles, Y., Halanych, K.M., Boore, J.L., 2008. Group II introns break new boundaries: presence in a bilaterian's genome. PLoS ONE 3, e1488.
Weigert, A., Helm, C., Meyer, M., Nickel, B., Arendt, D., Hausdorf, B., Santos, S.R., Halanych, K.M., Purschke, G., Bleidorn, C., Struck, T.H., 2014. Illuminating the base of the annelid tree using transcriptomics. Mol. Biol. Evol. 31, 1391–1401.
Wu, Z.G., Shen, X., Sun, M.A., Ren, J.F., Wang, Y.J., Huang, Y.L., Liu, B., 2009. Phylogenetic analyses of complete mitochondrial genome of *Urechis unicinctus* (Echiura) support that echiurans are derived annelids. Mol. Phylogenet. Evol. 52, 558–562.
Zhong, M., Struck, T.H., Halanych, K.M., 2008. Phylogenetic information from three mitochondrial genomes of Terebelliformia (Annelida) worms and duplication of the methionine tRNA. Gene 416, 11–21.