Group II Introns Break New Boundaries: Presence in a Bilaterian’s Genome

Yvonne Vallès1,2, Kenneth M. Halanych3, Jeffrey L. Boore1,2,4*

1 Department of Energy (DOE) Joint Genome Institute and Lawrence Berkeley National Laboratory, Walnut Creek, California, United States of America, 2 Department of Integrative Biology, University of California at Berkeley, Berkeley, California, United States of America, 3 Life Sciences Department, Auburn University, Auburn, Alabama, United States of America, 4 Genome Project Solution, Hercules, California, United States of America

Group II introns are ribozymes, removing themselves from their primary transcripts, as well as mobile genetic elements, transposing via an RNA intermediate, and are thought to be the ancestors of spliceosomal introns. Although common in bacteria and most eukaryotic organelles, they have never been reported in any bilaterian animal genome, organellar or nuclear. Here we report the first group II intron found in the mitochondrial genome of a bilaterian worm. This location is especially surprising, since animal mitochondrial genomes are generally distinct from those of plants, fungi, and protists by fast replication while establishing germ plasm during early development. This intron is found in the mtDNA of an annelid worm, (an undescribed species of Nephtys), where the complete sequence revealed a 1819 bp group II intron inside the cox1 gene. We infer that this intron is the result of a recent horizontal gene transfer event from a viral or bacterial vector into the mitochondrial genome of Nephtys sp. Our findings hold implications for understanding mechanisms, constraints, and selective pressures that account for patterns of animal mitochondrial genome evolution.

INTRODUCTION

Ribozymes are RNA molecules with enzymatic activities [1]. One type is called a self-splicing intron because these are removed from the gene’s transcript without the formation of the spliceosomal protein complex (although some may require other proteins for efficient splicing in vivo) [2]. These are separated into group I and II types depending on the mechanism of splicing [1,3,4]. Although each type folds into a characteristic structure that is necessary for catalysis, there is almost no widely conserved nucleotide sequence even within each type [1,2,3]. These introns are also mobile genetic elements, capable of movement into other genes and, in many cases, they contain one or more genes that encode for proteins (e.g. reverse transcriptase) that enable this mobility [4,5]. Both types of introns have a wide phylogenetic distribution (Table 1), being found in bacteria and the organelles of plants, fungi, protists and animals [3,6,7,8]. Interestingly, group II introns, although completely absent in nuclear eukaryotic genomes, are believed to be the ancestors of spliceosomal introns and therefore have played a central role in eukaryotic genome evolution [9]. In contrast, group I introns are found in phage, viruses and nuclear genomes of fungi and protists [3,10,11,12,13].

The very rare presence of introns in animal mtDNAs challenges current views on organelle evolution. With more than 900 complete animal mtDNA sequences available, only some members of the basal groups, sponges and cnidarians (which have group I introns) and the placozoan Trichoplax adhaerens (which has multiple introns, including one of group II), have been found to possess introns, and all appear to have been acquired secondarily [7,10,12]. Although mitochondrial genomes of protists, fungi, and plants display wide variation in size, structure, and gene content, undergo high rates of recombination, and often contain large amounts of non-coding sequence and both types of self-splicing introns [14,15,16], those of animals have become practically evolutionarily static, with almost all being small, compact, circular molecules with the same 37 genes and lacking introns, all but small tracts of non-coding sequence, and all or nearly all recombination [17]. Understanding the forces responsible for these restricted set of changes in animal mtDNAs, in contrast to those of fungi and plants, has been the subject of much study and debate among evolutionary biologists [18].

RESULTS

We have determined the complete sequence of the mitochondrial genome of Nephtys sp., a carnivorous polychaete inhabiting the intertidal and subtidal zones. This genome is typical of animal mtDNAs in possessing 37 genes on a single circular molecule with few and short non-coding regions [17]. However, contrary to all expectations, the protein coding gene cox1 contains a group II intron (Fig. 1). We confirmed that the intron is a part of the mtDNA rather than a nuclear pseudogene by using polymerase chain reactions (PCR) to amplify the entire mtDNA in two overlapping pieces using inverted primers that anneal within the intron (Table 2). We verified that this intron is, in fact, removed from the mRNA by cloning and sequencing cDNA made from the...
Table 1. Key characteristics of group I and group II self-splicing introns.

|                       | Group I introns                                                                 | Group II introns                                                                 | Nephtys Intron |
|-----------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|----------------|
| Distribution          | Bacteria, virus, phage and organelles of plants, fungi, and some animals (sponges and cnidarians), nuclear genomes of protists and fungi [8,13] | Bacteria, organelles of plants, fungi, protists, and animals (Trichoplax and Nephtys) [6,7] | Nephtys mitochondria |
| Size (depends on the presence/absence of ORFs) | 0.25–3.0 kb [3,41] | 0.38–3.4 kb [6] | 1.8 kb |
| Secondary structure   | Conserved within this group                                                     | Conserved within this group                                                     | Conserved with other group II introns |
| Splicing              | Self-splicing without lariat formation                                          | Self-splicing with lariat formation                                              | Potential for lariat formation |
| ORFs                  | Homing endonucleases [3,5]                                                     | Reverse transcriptase, maturase, DNA binding, and/or endonuclease domains [25] | Reverse Transcriptase and partial maturase |

doi:10.1371/journal.pone.0001488.t001

transcripts of the cox1 gene. We identified it as a group II self-splicing intron by a detailed examination of its sequence and potential secondary structure that revealed these diagnostic features: (1) conserved GUGYG and AY nucleotides at the 5' and 3' intron boundaries, respectively; (2) conserved sequence of domain V, which is the catalytic core of the intron's ribozyme activity; (3) presence of an ORF for a contiguous reverse transcriptase (RT) gene and a partial maturase gene; (4) potential secondary structure with six helical domains radiating from a central core consistent with the highly conserved secondary structure of group II introns (Fig. 1) [2,6,19,20]. This is the first case of any intron found in the mtDNA of any bilaterian animal (Table 1).

Attempting to identify the evolutionary origin of the intron, we incorporated the inferred amino acid sequence of the intron’s ORF (RT and partial maturase) into the alignment of Zimmerly et al. [19], which contains ORFs from other group II introns from bacterial and organelar genomes (both mitochondrial and chloroplast) from plants, fungi and protists (Table 3). In addition we have included the three most similar sequences in a search (using BLAST) of GenBank to the intron’s ORF of Nephtys and of Trichoplax adhaerens in the alignment. A maximum likelihood phylogenetic analysis suggests that Nephtys’s ORF is sister to the cox1 ORF718 of the marine centric diatom Thalassiosira pseudonana among those RT sequences available for comparison (Fig. 2). However, broader taxon sampling of group II introns is needed to reliably infer their evolutionary history.

**DISCUSSION**

The amino acid sequences of the intronic ORFs of Nephtys sp. and T. adhaerens (the only other animal shown to have a group II intron) have only 29% identity, indicating that these introns diverged long ago, presumably long-predating the divergence of these animal groups. Although both the T. adhaerens and Nephtys ORFs are found in the cox1 mitochondrial gene, their positions differ by 108 nucleotides. For these reasons it seems very likely that this intron has integrated into these two mtDNAs in separate events, particularly because the alternative would require the hypothesis of many parallel losses in related lineages.

Thus, this group II intron is most likely the result of recent horizontal gene transfer, presumably from a bacterial or a viral intermediary. This would require that the transferred genetic material was specifically sequestered by the germline in order to be inherited. Interestingly, some bacterial lineages (i.e. Wolbachia) invade the female reproductive tissues of their host (i.e. Drosophila) and live intracellularly inside of the eggs, leading to inheritance of the microbial population in subsequent generations [21]. Furthermore, horizontal gene transfer from these endosymbionts to the host has been documented [22]. In the case of annelids with high regenerative abilities such as Nephtys [23,24], it is also conceivable that the initial horizontal transfer event occurred in tissue that later re-differentiated during regeneration. Because these introns are highly mobile and move throughout populations by both horizontal transfer and vertical inheritance [2,16,25], the mitochondrial host could have acquired the intron from a bacterial endosymbiont. However, whether Nephtys sp. harbors endosymbionts is yet unknown.

It may be that this intron is present only because it was recently acquired and insufficient time has passed for it to be lost. Nonetheless, it is tempting to speculate on whether there are properties of this mitochondrial genome that differ from those of most animals that have allowed it to escape from the presumed selection for small size to ensure rapid replication. Lynch and colleagues [18,26] have advanced a theory to explain the opposite trends in the evolution of plant and animal mtDNAs that links rate of nucleotide substitution (much lower for plants) with the propensity to accumulate non-coding DNA (much higher in plants). Supporting this hypothesis, it is associated that, in contrast to most animals, cnidarians have a very slow rate of mitochondrial sequence change that may account for their adoption of introns [27,28]. However, this does not appear to be consistent in the case of Nephtys sp., since a maximum likelihood analysis including all annelids for which the complete mtDNA sequence is available shows no great differences in branch lengths among them, even though only Nephtys sp. has acquired an intron (Fig. 3).

Further study of mtDNAs of annelids, as well as the other groups that have acquired introns (i.e. sponges, cnidarians and placozoans) may illuminate the extent and patterns of intron gain as well as provide further genome-level data for better understanding the forces shaping mitochondrial genome evolution.

**MATERIALS AND METHODS**

**Nucleic acid extraction and sequencing**

Total genomic DNA was extracted from frozen tissue using a Qiagen DNeasy kit according to supplier’s instructions. The mtDNA was amplified by long PCR (using the Takara polymerase kit) in three overlapping pieces (~8 kb each) using specific and universal primers (Table 2) [29]. Each amplification product was ethanol-precipitated with NaSO4, dried, and resuspended in 100 μl of water. This was sheared into ~1.5 kb fragments using a HydroShear device (GeneMachines), then the fragment ends were repaired using Klenow fragment and T4 polymerase. The product was size-selected using an agarose gel and ligated into pUC18 vector. These clones were introduced into E. coli cells by electroporation. This was plated and grown overnight at 37°C. Colonies were picked and processed to generate reads from each end of randomly selected clones.
Figure 1. Predicted secondary structure of the *Nephtys* sp. group II intron. Potentially conserved secondary structure consisting of a central core from which radiate six domains (I–VI). The RT and partial maturase ORF are encoded within domain IV. EBS and IBS indicate sites where interaction between the intron and exon (respectively) occurs when splicing. Greek symbols designate sequence sites potentially involved in tertiary structure. doi:10.1371/journal.pone.0001488.g001
Total RNA was extracted from tissues samples using Quiazo1 reagent (Quiagen) according to the manufacturers instructions. cDNA was constructed from total RNA using the Superscript III First-Strand Synthesis System for RT-PCR (Invitrogen) following the manufacturers instructions. In order to verify that the presumed intron was removed from the transcript, a portion of the \( \text{cox1} \) mRNA was amplified by PCR using primers matching the mtDNA sequence, and this product was isolated, cloned, and sequenced as above.

### Table 2. Primers used for completion of \( \text{Nephtys} \)’s mtDNA amplification.

| Primers          | Sequences                                  |
|------------------|--------------------------------------------|
| Universal primers | LCO1490 5'- GGTCAACAACTATACAAAAATGG-3'    |
|                  | HCO2198 5'- TAAACTTCAAGGTAGTCACAAAAATCA-3' |
| Specific primers | Neph-cx1F 5'- TGACCACTACGATCTGGAG-3'      |
|                  | Neph-cx1R 5'- TGACCACTACGATCTGGAG-3'      |
|                  | Intron-F 5'- TCAGGCTACACATCTGGG-3'       |
|                  | Intron-R 5'- AACGATGACAGATTTCCC-3'       |
|                  | A.186. 4C-cb-F-216 5'- AAGATTGCTTTTCAACATATTACACCCG-3' |
|                  | A.186. 4C-cb-R-306 5'- GGTCTACAGATCGTACAGAGTGG-3' |

Sequence annotation, alignment and phylogenetic analyses

Phred and Phrap were used to call bases and produce an alignment (~10×) and Consed was used for manual verification of quality. The mtDNA was annotated using DOGMA and MacVector (Accelrys). The intron secondary structure was folded initially with Mfold followed by hand editing.

In order to evaluate the evolutionary history of the introns themselves, the intronic ORFs within \( \text{Nephtys} \) sp. (Accession number EU293739), \( \text{Trichoplax adhaerens} \) [NC_008151], \( \text{Thalassiosira pseudonana} \) (ORF718) [YP_316587.1], \( \text{Schizosaccharomyces octosporus} \) (ORF786) [NP_700371.1] and \( \text{Chlorokybus atmophyticus} \) (ORF845) [NC_009630.1] were incorporated within the Zimmerly et al. alignment using DIALIGN (alignment accession number ALIGN_001217). Gblocks was used to determine the amino acid positions included in the phylogenetic analyses.

We performed a maximum likelihood (ML) analysis following the JTT model for amino acid substitution and executed bootstrap resampling to evaluate branch support in RAxML. Percentage identity between \( \text{Nephtys} \)’s ORF and \( \text{T. adhaerens} \) was calculated using the amino acid sequence included in the analysis with the program WU-blastp.

In order to evaluate the rates of change for \( \text{Nephtys} \) sp. and related mtDNAs, the protein coding genes of six annelids (\( \text{Lumbricus terrestris} \), \( \text{Nephtys} \) sp., \( \text{Clymenella torquata} \), \( \text{Platynereis dumerilii} \), \( \text{Orbinia latreillii} \), and \( \text{Urechis caupo} \)) and two mollusks (\( \text{Nautilus macromphalus} \) and \( \text{Katharina} \)) were incorporated within the Zimmerly et al. alignment using DIALIGN (alignment accession number ALIGN_001217). Gblocks was used to determine the amino acid positions included in the phylogenetic analyses.

### Figure 2. Phylogenetic analysis of 71 group II intron ORFs. A maximum likelihood analysis of the amino acid sequence for 71 ORFs suggests the \( \text{cox1} \) ORF718 of the marine centric diatom \( \text{Thalassiosira pseudonana} \) as sister to the \( \text{Nephtys} \)’s ORF. Red stars indicate a bootstrap support ≥ 90.

Names of taxa are indicated by the capital letter of the genus name, followed by species name and when applicable the intron location (specified in table 3).

doi:10.1371/journal.pone.0001488.t002
Table 3. Mitochondrial, chloroplast and bacterial group II introns included in the phylogenetic analysis (modified from Zimmerly et al.).

| Organism class | Species name | Host gene | Intron |
|----------------|--------------|-----------|--------|
| Placozoans     | Trichoplax adhaerens | cox1 |        |
| Bilaterians    | Nephtys sp. | cox1 |        |
| Yeast          | Kluyveromyces lactis | cox1 | I1     |
|                | Saccharomyces cerevisiae | cox1 | I2     |
|                | Schizosaccharomyces pombe | cox1 | I1     |
|                | Schizosaccharomyces pombe | cob1 | I1     |
|                | Schizosaccharomyces pombe | cox2 | I1     |
|                | Schizosaccharomyces octosporus | cox1 | I4     |
| Other Fungi    | Allomyces macrognus | cox1 | I3     |
|                | Cryphonectia parasitica | No data | No data |
| Green algae    | Bryopsis maxima | rbcL | I1     |
|                | Chlorokybus atmophyticus | atp9 | I2     |
|                | Oocystacea sp. | petD | I1     |
| Liverworts     | Marchantia polymorpha | atpA | I1     |
|                | Marchantia polymorpha | atpA | I2     |
|                | Marchantia polymorpha | atp9 | I1     |
|                | Marchantia polymorpha | cox1 | I1     |
|                | Marchantia polymorpha | cox1 | I2     |
|                | Marchantia polymorpha | cob1 | I3     |
|                | Marchantia polymorpha | cox2 | I2     |
|                | Marchantia polymorpha | SSU rDNA | I1     |
|                | Marchantia polymorpha | Free standing | Orf732 |
| Other green plants | Arabidopsis thaliana | nad1 | I4     |
|                | Glycine max | nad1 | I4     |
|                | Oenothera berteriana | nad1 | I4     |
|                | Solanum tuberosum | nad1 | I4     |
|                | Triticum aestivum | nad1 | I4     |
|                | Vicia faba | nad1 | I4     |
|                | Zea mays | nad1 | I4     |
| Red algae      | Porphyra purpurea | LSU rDNA | I1     |
|                | Porphyra purpurea | LSU rDNA | I2     |
| Brown algae    | Pylaiella littoralis | LSU rDNA | I1     |
|                | Pylaiella littoralis | LSU rDNA | I2     |
|                | Pylaiella littoralis | cox1 | I1     |
|                | Pylaiella littoralis | cox1 | I2     |
|                | Pylaiella littoralis | cox1 | I3     |
| Stramenopile   | Thalassiosira pseudonana | cox1 |        |
| Cryptomonad    | Pyrenomonas salina | cpn60 | I1     |
| Euglenoids     | Astasia longa | Free standing | Orf456 |
|                | Euglena gracilis | psbC | I4     |

Table 3. cont.

| Organism class | Species name | Host gene | Intron |
|----------------|--------------|-----------|--------|
| Bacteria       | Bacillus anthracis | None |        |
|                | Bacillus anthracis | PX01-08/PX01-06 | None |
|                | Bacillus halodurans | PX01-24/ORFX | None |
|                | Bacillus megaterium | None |        |
|                | Calothrix sp. | ORF1 |        |
|                | Clostridium difficile | ORF14 | None |
|                | Escherichia coli | None |        |
|                | Escherichia coli | ORF14 | None |
|                | Lactococcus lactis | Relaxed |        |
|                | Pseudomonas alcaligenes | None |        |
|                | Pseudomonas putida | None |        |
|                | Pseudomonas sp. | None |        |
|                | Serratia marcescens | None |        |
|                | Shigella flexneri | IS629-likeORF |        |
|                | Sinorhizobium mellioti | ORF B |        |
|                | Sphingomonas aromaticivorans | Replication primer |        |
|                | Sphingomonas aromaticivorans | ORF392/ORF416 |        |
|                | Streptococcus pneumoniae | None |        |
|                | Xylella fastidiosa | DNA Methyltransferase |        |

Figure 3. Maximum likelihood analysis of the protein coding genes. The maximum likelihood analysis of the mitochondrial protein coding genes of six annelids shows that branch lengths among them are similar, suggesting that Nephtys does not have an obviously slower rate that might create a propensity for harboring introns.

doi:10.1371/journal.pone.0001488.t003
were aligned with the CLUSTALX [38]. Gblocks [36] was used to determine the positions included in the analyses. Modtest determined the ML model (TVMrcat) that best fit the data. We built the tree (Fig. 3) using the settings given by Modtest [39] (Lset: Base = (0.3156 0.2146 0.1402) Nst = 6 Rmat = (1.1922 12.4843 20.0126 3.9256 12.4843) Rates = gamma Shape = 0.2380 Pinvar = 0) and 1000 bootstrap replicates in PAUP [40].

ACKNOWLEDGMENTS
We thank R. Baker, P. Francisco, T. Hausmann, J. Kuehl, C. Moritz, D. Simon, M. Stoeck, H. Valles and D. Weisblat for valuable comments and help with the analysis. The organism was collected with help from Friday Harbor Marine Labs of the University of Washington. L. Harris of the Los Angeles County Museum was most helpful with organismal identification (undescribed species, Nephtys sp. 3 Harris); this is now accession LACM-AHF POLY 2180 at that institution.

Author Contributions
Conceived and designed the experiments: YV JB. Performed the experiments: YV. Analyzed the data: YV. Contributed reagents/materials/analysis tools: KH JB. Wrote the paper: KH YV JB.

REFERENCES
1. Lehmann K, Schmidt U (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. Crit Rev Biochem Mol Biol 38: 249–303.
2. Lambowitz AM, Zimmerly S (2004) Mobile group II introns. Annu Rev Genet 38: 1–35.
3. Haugen P, Simon DM, Bhattacharya D (2005) The natural history of group I introns. Trends Genet 21: 111–119.
4. Robart AR, Zimmerly S (2005) Group II intron retroelements: function and diversity. Cytogenet Genome Res 110: 589–597.
5. Belfort M, Derbshey V, Parker MM, Cousineau B, Lambowitz AM (2002) Mobile introns: pathways and proteins. In: Geis JL, ed (2002) Mobile DNA II. Washington, D.C.: ASM, Press.
6. Bonen L, Vogel J (2001) The ins and outs of group II introns. Trends Genet 17: 322–331.
7. Dellaporta SL, Xu A, Sagasser S, Jakob W, Moreno MA, et al. (2006) Mitochondrial genome of Trichoplax adhaerens supports placozoa as the basal lower metazoan phylum. Proc Natl Acad Sci U S A 103: 8754–8756.
8. Foley S, Bruttin A, Brussow H (2000) Widespread distribution of a group II intron and its three deletion derivatives in the yin gene of Streptococcus thermophilus bacteriophage. J Virol 74: 611–618.
9. Palmer JD, Logdon JM Jr (1991) The recent origins of introns. Curr Opin Genet Dev 1: 470–477.
10. Beagley CT, Okimoto R, Wolstenholme DR (1998) The mitochondrial genome of the sea anemone Metridium senile (Cnidaria): introns, a paucity of DNA genes, and a near-standard genetic code. Genetics 148: 1091–1108.
11. Medina M, Collin AG, Takaoka TL, Kuehl JV, Boone JL (2006) Naked corals: skeleton less in Scleractinia. Proc Natl Acad Sci U S A 103: 9096–9100.
12. Rot C, Goldfarb I, Ilan M, Huchon D (2006) Putative cross-kingdom horizontal gene transfer in sponge (Porifera) mitochondria. BMC Evol Biol 6: 71.
13. Yamada T, Tamura K, Aimi T, Sengori P (1994) Self-splicing group I introns in eukaryotic viruses. Nucleic Acids Res 22: 2532–2537.
14. Andre C, Levy A, Wallot V (1992) Small repeated sequences and the structure of plant mitochondrial genomes. Trends Genet 8: 128–132.
15. Gray MW, Burger G (1999) Mitochondrial evolution. Science 283: 1476–1481.
16. Palmer JD, Adams KL, Cho Y, Parkinson CL, Qui YL, et al. (2000) Dynamic evolution of plant mitochondrial genomes: mobile introns and introns and highly variable mutation rates. Proc Natl Acad Sci U S A 97: 6960–6966.
17. Boehm JL (1999) Animal mitochondrial genomes. Nucleic Acids Res 27: 1767–1780.
18. Lynch M, Koskella B, Schaack S (2006) Mutation pressure and the evolution of organelle-genomic architecture. Science 317: 1730–1737.
19. Zimmerly S, Hasnain G, Wu X (2001) Phylogenetic relationships among group II intron ORFs. Nucleic Acids Res 29: 1239–1250.
20. Knoop V, Kloska S, Brennicke A (1994) On the identification of group II introns in nucleotide sequence data. J Mol Biol 242: 309–396.
21. Pontes MH, Dale C (2006) Culture and manipulation of insect facultative symbionts. Trends Microbiol 14: 406–412.
22. Kondo N, Nishik N, Iijichi N, Shimada M, Fukatsu T (2002) Genome fragment of Wolbachia endosymbiont transferred to X chromosome of host insect. Proc Natl Acad Sci U S A 99: 14290–14295.
23. Clark ME (1968) Later stages of regeneration in the Polycheate, Nephtys. J Morphol 124: 483–510.
24. Clark ME, Clark KB (1982) Growth and regeneration in Nephtys. Zoolog J Physiol 76: 24–90.
25. Dai L, Toor N, Olson R, Keeping A, Zimmerly S (2003) Database for mobile group II introns. Nucleic Acids Res 31: 424–426.
26. Lynch M, Conery JS (2003) The origins of genome complexity. Science 302: 1401–1404.
27. van Oppen MJ, Willis BL, Miller DJ (1999) Atypically low rate of cytochrome b evolution in the scleractinian coral genus Acropora. Proc Biol Sci 266: 179–183.
28. Sharrer TL, Van Oppen MJ, Romano SL, Worheide G (2002) Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). Mol Ecol 11: 2475–2487.
29. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for the amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3: 294–299.
30. Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8: 186–194.
31. Ewing B, Hiller L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred: I Accuracy assessment. Genome Res 8: 175–185.
32. Gordon D, Aghaiian C, Green P (1998) Consed: a graphical tool for sequence finishing. Genome Res 8: 195–202.
33. Wyman SK, Jamson RK, Boone JL (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20: 3252–3255.
34. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406–3415.
35. Morgenstern B (1999) DALIGN 2: improvement of the segment-to-segment approach to multiple sequence alignment. Bioinformatics 15: 211–218.
36. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution 17: 540–552.
37. Stamatakis A, Ludwig T, Meier H (2005) RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 21: 456–463.
38. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 24: 4876–4882.
39. Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14: 817–818.
40. Swofford DL (2000) PAUP*. Phylogenetic Analysis Using Parsimony *and Other Methods, Version 4.0 b10. Sinauer Associates S, ed. Massachusetts.
41. Lambowitz AM, Belfort M (1993) Introns as mobile genetic elements. Annu Rev Biochem 62: 587–622.