Marine Genomics xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Marine Genomics

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

The mitochondrial genome sequence of a deep-sea, hydrothermal vent limpet, *Lepetodrilus nux*, presents a novel vetigastropod gene arrangement

Yuichi Nakajima [a,*], Chuya Shinzato [b], Mariia Khalturina [b], Masako Nakamura [c], Hiromi Watanabe [d], Noriyuki Satoh [b], Satoshi Mitarai [a]

[a] Marine Biophysics Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan
[b] Marine Genomics Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan
[c] School of Marine Science and Technology, Tokai University, Shimizu, Shizuoka 424-8610, Japan
[d] Japan Agency for Marine-Earth Science and Technology, Yokosuka, Kanagawa 237-0061, Japan

ARTICLE INFO

Article history:
Received 12 January 2016
Received in revised form 8 April 2016
Accepted 8 April 2016
Available online xxxx

Keywords:
Mitogenome
Next-generation sequencing
Okinawa Trough
Vetigastropoda

ABSTRACT

While mitochondrial (mt) genomes are used extensively for comparative and evolutionary genomics, few mt genomes of deep-sea species, including hydrothermal vent species, have been determined. The Genus *Lepetodrilus* is a major deep-sea gastropod taxon that occurs in various deep-sea ecosystems. Using next-generation sequencing, we determined nearly the complete mitochondrial genome sequence of *Lepetodrilus nux*, which inhabits hydrothermal vents in the Okinawa Trough. The total length of the mitochondrial genome is 16,353 bp, excluding the repeat region. It contains 13 protein-coding genes, 22 tRNA genes, two rRNA genes, and a control region, typical of most metazoan genomes. Compared with other vetigastropod mt genome sequences, *L. nux* employs a novel mt gene arrangement. Other novel arrangements have been identified in the vetigastropod, *Fissurella volcano*, and in *Chrysomallon squamiferum*, a neomphaline gastropod; however, all three gene arrangements are different, and Bayesian inference suggests that each lineage diverged independently. Our findings suggest that vetigastropod mt gene arrangements are more diverse than previously realized.

1. Introduction

Since deep-sea hydrothermal vents were discovered along the Galapagos Rift in 1977 (Lonsdale, 1977; Corliss et al., 1979), more hydrothermal vents have been found in the Pacific, Atlantic, Arctic, and Indian Oceans (e.g., Van Dover, 2000; Beaulieu et al., 2013). These vent fields are characterized by extreme physical and chemical conditions, such as high concentrations of hydrogen sulfide and metals, high levels of sulfur dioxide, low levels of oxygen, and low pH; however, they are often densely inhabited by various invertebrates such as tube worms, bivalves, gastropods, crabs, and shrimps (e.g., Van Dover, 2000; Van Dover et al., 2002). Some vent species maintain symbiotic relationships with chemoautotrophic bacteria (reviewed in Van Dover, 2000; Dubilier et al., 2008). Deep cold seeps, wood-falls, and whale-falls also help to support chemosynthetic organisms. Genetic studies of hydrothermal vent species are poorly known.

Mitochondrial genome sequences have facilitated studies of phylogenetics, population genetics, species identification, molecular evolution, recombination, and maternal inheritance (e.g., Brown et al., 1979; Moritz and Brown, 1987; Hebert et al., 2002). Most metazoan mitochondrial genomes consist of 37 genes: 13 encoding subunits of enzymes involved in oxidative phosphorylation, 22 tRNA genes, and two rRNA genes (Boore, 1999). Metazoan mitochondrial genomes are usually circular, double-stranded DNA molecules of ~12–20 kb. Anderson et al. (1981) reported the complete sequence of the human mitochondrial genome, which was the first organellar genome sequenced. Thus far, over 5700 animal mitochondrial genomes have been sequenced (NCBI Organelle Genome Resources, http://www.ncbi.nlm.nih.gov/genome/organelle/), but only three from hydrothermal vent gastropods have been determined (*Chrysomallon squamiferum*, Nakagawa et al., 2014, *Ifremeria nautilei*, Osca et al., 2014, and *Provanna subglabra*, Xu et al., 2015).

The Vetigastropoda is a major clade within the Gastropoda. At least, 49 vetigastropod species (8 families, 22 genera) have been identified at hydrothermal vents and/or cold seeps (Table S1, Desbruyères et al., 2006; Fujikura et al., 2008). The clades Caenogastropoda, Cephalopoda, and Polyplacophora, and three genera of the Clade Vetigastropoda from shallow marine habitats show the same gene arrangement; only *Fissurella volcano* displays a unique arrangement (Williams et al., 2014). However, no mt genome sequences for deep-sea vetigastropods

* Corresponding author.
E-mail addresses: yuichi.nakajima@oist.jp, yuichi.nakajima@outlook.com (Y. Nakajima).

http://dx.doi.org/10.1016/j.margen.2016.04.005
1874-7787/

have been reported, although there is one for a hydrothermal vent species, the scaly-foot gastropod (*C. squamiferum*) (Nakagawa et al., 2014), which belongs to the Clade Neomphalina, a sister group of all other vetigastropods (Warén et al., 2003; Chen et al., 2015). Interestingly, its mt genome arrangement is different from those of all known gastropods (Nakagawa et al., 2014, Genbank Accession No. AP013032).

The Superfamily Lepetodriloidea, pertaining to the Clade Vetigastropoda, contains the families Lepetodrilidae (four genera) and Sutilizonidae (two genera). The Lepetodriloidea occurs in the deep-sea, mainly at hydrothermal vents. The Genus *Lepetodrilus* (Family Lepetodriloidae), is the most diversified taxon among deep-sea vetigastropods (Desbruyères et al., 2006). It occurs at deep-sea hydrothermal vents, cold seeps, wood-falls, and whale-falls. Molecular phylogenetic trees, using DNA barcoding, and focusing on the CO1 gene, identified at least 19 taxa in this genus; however, eight of the 19 are undescribed or cryptic species (Johnson et al., 2008). *Lepetodrilus nux* shows high genetic diversity and connectivity among vent sites, and is an opportunistic colonizer with a continuous reproductive strategy (Nakamura et al., 2014). Here we report the mitochondrial genome sequence of *L. nux*, explore possible evidence of mt gene rearrangements using next-generation sequencing, and provide reference information for the Genus *Lepetodrilus* and the Family Lepetodrilidae.

2. Materials and methods

2.1. Sampling and sequencing

For sequencing of total genomic DNA, specimens of *L. nux* were collected around deep-sea hydrothermal vents in the Iheya North field (27°47.412′N, 126°54.037′E, at 1058 m depth) in the Okinawa Trough. Collection was accomplished during a cruise of the Research Vessel 'Natsushima' (Cruise No. NT13-22), using the remotely operated vehicle (ROV) 'Hyper-Dolphin' (Dive No. HPD#1592). Specimens were preserved in ethanol and transferred to the laboratory. Genomic DNA of *L. nux* was isolated using a proteinase K–phenol–chloroform extraction and purified using ethanol precipitation and a QIAquick PCR purification kit (Qiagen).

Extracted DNA was fragmented and the resulting paired-end library was sequenced using a MiSeq sequencer (Illumina) (300-bp reads) according to the manufacturer's instructions. Sequencing adapter contaminants were trimmed with fastq-mcf in ea-utils ver. 1.1.2-537 (Aronesty, 2011) and sequences of each read pair were merged using fastq-join in ea-utils. Then these sequences were assembled using the Newbler de novo assembler ver. 2.8 (Roche, Margulies et al., 2005). Despite assembling read pairs and annotating the mt genome sequence, we were unable to confirm whether the sequence is circular, as is normal for mt genomes (see Results and discussion). Therefore, we designed primers and conducted polymerase chain reaction (PCR) using a Dice Touch TP350 Thermal Cycler (Takara) to determine the configuration (circular or linear) of mt DNA in the missing portion located in the non-coding region known as the control region. The reaction mixture (25 μL) for PCR contained <100 ng of template genomic DNA, Ex Taq DNA Polymerase (Takara), and forward and reverse primers (5′-GGTGAGGTAATTATGGTGTCAGGATTG-3′ and 5′-GAGGCAAGACGAATCACCTCTAAAGGAG-3′). PCR conditions were: 2 min at 95 °C, followed by 30 cycles of 30 s at 95 °C, 30 s at 58 °C or 60 °C, and 60 s at 72 °C, with an extension of 5 min at 72 °C in the final cycle. Furthermore, nested PCR was conducted using nested primers (5′-GAGATTTAGAGAAAATACATG-3′ and 5′-CCGTCAACTTCTATAAACTTC-3′). We confirmed that mt DNA was circular using agarose gel electrophoresis. The PCR product was purified with Exonuclease I (Takara) and Shrimp Alkaline Phosphatase (Takara), and then directly sequenced using a BigDye Terminator Kit ver. 3.1 (Thermo Fisher Scientific). Products for sequencing were purified by ethanol precipitation, and the sequence was analyzed using an ABI 3130xl sequencer (Thermo Fisher Scientific).

2.2. Mitochondrial genome annotation and comparison of gene arrangements

We edited sequences obtained using ApE — A Plasmid Editor ver. 2.0.45 (http://biologylabs.utah.edu/jorgensen/wayned/ape/). Possible open reading frames in the genome were confirmed with ExPASy's translate tool (Artimo et al., 2012, http://web.expasy.org/translate/). Preliminarily, DOGMA (Wyman et al., 2004, http://dogma.ccbb.utexas.edu/) was used to determine the outline of genes and their positions. Using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), we confirmed all 13 protein-coding genes and two rRNA genes. Ends of rRNA genes were determined from adjacent genes. All tRNA gene sequences were identified using ARWEN ver. 1.2 (Laslett and Canbäck, 2008, http://mbio-serv2.mbioekol.lu.se/ARWEN/). Mitochondrial genome sequences and their annotation information were submitted to GenBank/EMBL/DDBJ under accession number LC107880. A circular display of *L. nux* mitochondrial genomes was created using OGDRAW ver. 1.2 (Lohse et al., 2013, http://ogdraw.mpimp-golm.mpg.de/) and then modified manually. Mitochondrial genomes of vetigastropod species determined to date were used for comparison of nucleotide sequences, gene composition, and arrangement: *F. volcano* (JN790612, Simison W.B.), *Haliotis discus* (KF724723, Yang E.C., Noh S.J., Park J.-H., and Yoon H.S.), *Haliotis laevigata* (KJ472483, Robinson et al., 2016), *Haliotis rubra* (AY588938, Maynard et al., 2005), *Haliotis tuberculata* (FJ599667, Van Wormhoudt et al., 2009), *Lunella* aff. *cinerea* (KF700096, Williams et al., 2014), and *Tegula brunnea* (JN790613, Simison W.B.).

2.3. Construction of a phylogenetic tree

Amino acid sequences of *L. nux* and seven other species from the Clade Vetigastropoda (*F. volcano*, *H. discus*, *H. laevigata*, *H. rubra*, *H. tuberculata*, *L.* aff. *cinerea*, and *T. brunnea*), the Clade Neomphalina, *C. squamiferum* (AP013032, Nakagawa et al., 2014), the Neritimorpha, *Nerita melanotragus* (GU810158, Castro and Colgan, 2010), and the Littorinimorpha, *Oncomelania hupensis* (FJ997214, Li S. and Zhou X.), and *Dendropoma maximum* (HM174253, Rawlings et al., 2010) were used as references. *D. maximum* was classified in a different group from *O. hupensis*, despite belonging to the same Clade, according to recent phylogenetic studies (Williams et al., 2014). *D. maximum* was used as an outgroup species in this study. Phylogenetic trees were constructed to compare mitochondrial gene arrangements in the Vetigastropoda. Concatenated amino acid sequences for all protein-coding genes, except for ATP8, which shows very low identity with other organisms, were used. Amino acid sequences for these 11 species were aligned using MUSCLE in MEGA ver. 6.06 (Tamura et al., 2013). In aligned sequences, ambiguously aligned regions and gaps were removed using the Gblocks server ver. 0.91b (Talavera and Castresana, 2007, http://molevol.cmima.csic.es/castresana/Gblocks_server.html) with default settings. The molecular evolution model was selected with Partition Finder ver. 1.1.1 (Lanfear et al., 2012) implementing a partition scheme that considered mutation rate variation across protein-coding genes. Employing the lowest scores from the Bayesian Information Criterion, the best model for 10 genes (excluding ND2 and CO1) was determined to be MtArt (empirical amino acid substitution matrices in arthropod mitochondria, Abascal et al., 2007) with a proportion of invariable sites (+I), a gamma distribution of rate variation across sites (+G), and observed amino acid frequencies (+F). The best model for ND2 and CO1 was determined to be MtArt+G.

Bayesian inference (BI) for phylogeny was performed with MrBayes ver. 3.2.5 (Ronquist and Huelsenbeck, 2003; Ronquist et al., 2012), using the MtRev (in vertebrate mitochondria, Adachi and Hasegawa, 1996) +I+G+F model for 10 genes and MtRev+G model for ND2 and CO1, because the MtArt model was not implemented in MrBayes. Two independent runs were made using the following conditions: 4 Markov chain Monte Carlo (MCMC) iterations, 50,000,000 generations, a print

Please cite this article as: Nakajima, Y., et al., The mitochondrial genome sequence of a deep-sea, hydrothermal vent limpet, *Lepetodrilus nux*, presents a novel vetigastropod gene arrangement..., Mar. Genomics (2016), http://dx.doi.org/10.1016/j.margen.2016.04.005