Parasitology International 65 (2016) 271–275

Contents lists available at ScienceDirect

Parasitology International

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

Short communication

The mitochondrial genome and ribosomal operon of *Brachycladium goliath* (Digenea: Brachycladiidae) recovered from a stranded minke whale

Andrew G. Briscoe [a], Rodney A. Bray [a], Jan Brabec [b], D.T.J. Littlewood [a,*]

[a] Department of Life Sciences, Natural History Museum, London SW7 5BD, UK
[b] Institute of Parasitology, Biology Centre of the Czech Academy of Sciences, Branišovská 31, 370 05 České Budějovice, Czech Republic

ARTICLE INFO

Article history:
Received 24 December 2015
Received in revised form 1 February 2016
Accepted 12 February 2016
Available online 13 February 2016

Keywords:
Digenea
Balaenoptera acutorostrata
Cetacea
Hologenophore
NGS

ABSTRACT

Members of the Brachycladiidae are known to cause pathologies implicated in cetacean strandings and it is important to develop accurate diagnostic markers to differentiate these and other helminths found in cetaceans. *Brachycladium goliath* (van Beneden, 1858) is a large trematode found, as adults, usually in the hepatic (bile) and pancreatic ducts of various cetaceans. Complete sequences were determined for the entire mitochondrial genome, and phylogenetically informative nuclear genes contained within the ribosomal operon, from a small piece of an individual worm taken from a common minke whale *Balaenoptera acutorostrata* Lacépède, 1804. Genomic DNA was sequenced using an Illumina MiSeq platform. The mtDNA is 15,229 bp in length consisting of 12 protein-coding genes, 22 tRNA genes, 2 rRNA genes and 2 non-coding regions of which the larger is comprised of 4 tandemly repeated units (260 bp each). The ribosomal RNA operon is 9297 bp long. These data provide a rich resource of molecular markers for diagnostics, phylogenetics and population genetics in order to better understand the role, and associated pathology of helminth infections in cetaceans.

1. Introduction

Historical collecting of marine mammals has revealed a rich and relatively unique fauna of metazoan parasites, particularly helminths. Excessive, unregulated and illegal hunting has contributed to the decline of many marine mammals, particularly those species targeted through whaling. Currently, the availability of parasites from large marine mammals relies almost exclusively on strandings and the subsequent swift action of veterinary teams and biologists attending strandings in securing parasite material through post-mortem examinations and autopsies. Helminth infections have been implicated as a cause of cetacean strandings, with reports of nematodes and trematodes affecting central nervous systems, including the brain, and cranial sinuses [1]. In order to best utilise such opportunistically collected material and as part of an on-going commitment to resolving molecular-based phylogenies for Digenea, we sought to resolve the full mitogenome and full nuclear ribosomal RNA operon from a small tissue sample taken from an individual trematode collected from a common minke whale autopsy, using the Illumina MiSeq sequencing platform. Molecular data for such rare parasite specimens are key to (i) better understanding life cycles, through the identification of larval forms, whether free-living or those found within intermediate hosts [2], (ii) non-invasive parasite identification, e.g. from eDNA or faecal samples [3,4], (iii) biogeographical distributions and population ecology [5], and (iv) the development of reliable diagnostic markers to assist in parasite identification and understanding their role in associated pathologies [6].

A large specimen of *Brachycladium goliath* (van Beneden, 1858), collected from a minke whale, was sent to the Natural History Museum London, for identification purposes (full details below). Found predominantly in the bile ducts of their cetacean hosts, it is an impressively sized trematode (50–90 mm in length). The parasite has been recorded (usually as *Lecithodesmus goliath*) from various baleen whales including at least four additional species of *Balaenoptera* (sei, blue and fin whales), the grey whale *Eschrichtius robustus* and the bowhead whale *Balaena mysticetus*; source NHM Host–Parasite database [7]. Records show that its presence in minke whales has been reported previously from the Sea of Okhotsk and the coast of Japan [8,9], as well as the northeast Atlantic [10]; its distribution appears to be global reflecting the global distribution of baleen whales. However, the life cycle of *B. goliath* is not known and only adults have been found to date, often in association with chronic parasitic hepatitis. Members of the Brachycladiidae are close relatives of the Acanthocolpidae and are likely derived evolutionarily from these parasites of piscivorous marine fish [11], but greater molecular sampling is required to establish their position amongst the Digenea. Here we explore the potential for maximising next generation sequencing techniques in rapidly recovering useful molecular phylogenetic and molecular ecological data from these rarely sampled parasites.

* Corresponding author.
E-mail address: t.littlewood@nhm.ac.uk (D.T.J. Littlewood).

http://dx.doi.org/10.1016/j.parint.2016.02.004
1383-5769/

2. Materials and methods

A single specimen of *B. goliath* (van Beneden, 1858) was provided by the Scottish Marine Animal Stranding Scheme (SRUC Veterinary Services, Inverness, Scotland), removed from the liver of an adult female common minke whale, *Balaenoptera acutorostrata* Lacépède, 1804, fixed and stored in 90% ethanol; whale stranded and died at Whitehills, Aberdeenshire 7 September 2014, grid ref NJ66226535. A small piece (1 mm^3) of this large trematode (50 mm total length, fixed) was removed for molecular analysis; the rest of the worm (hologenophore) was stained, mounted, positively identified and accessioned into the NHM Parasitic Worms collection (NHM 2014.9.12.1).

Total genomic DNA was extracted from the 1 mm^3 piece of tissue, using the Bioline Genomic Isolate II extraction kit (Bioline). Following quantification of double stranded DNA with a Qubit fluorometer 2.0 (Invitrogen) an indexed library was constructed using a TruSeq Nano DNA sample preparation kit (Illumina) with 550 bp inserts, as per manufacturer's recommendations, and sequenced on 1/6th of an Illumina MiSeq flowcell. Reads were trimmed using default settings in Geneious v.6.1.4 (Biomatters) and subsequently assembled to the *cox*1 region of the *Paragonimus westermani* mitochondrial genome [12] GenBank accession: AF219379. Unassembled reads were then iteratively mapped and reassembled to the putative *B. goliath cox*1 sequence until the resulting contig could be circularised. Gene boundaries were annotated using MITOS [13] and verified by visualisation of open reading frames and comparison to alignments of platyhelminth mitochondrial genes.

Using nuclear ribosomal RNA gene sequences for *Brachycladium* from GenBank (Accession GQ226035), fragments of the ribosomal RNA operon were identified and assembled using the same iterative process as described for the mitochondrial genome. Exact coding positions of the 18S and 28S nuclear rDNAs, as well as the respective 5′ and 3′ boundaries of the external transcribed spacers, were determined using RNAmmer [14]. The programme ITSx [15] was used to identify the boundaries of ITS1 and ITS2. Subsequently the complete annotation was compared with the fully-annotated human rDNA repeating unit [16], GenBank accession: HSU13369.

Amino acid translations for each of the 12 protein-coding genes were aligned with 22 reference mt genomes taken from a previous study [17] using MAFFT. Aligned sequences were concatenated, verified by eye and subject to phylogenetic analysis using RAxML 8.1.7 [18] with a MtZoa substitution model; trees were rooted against cestode outgroups. All model parameters and bootstrap nodal support values were estimated using RAxML; the number of bootstrap repetitions was estimated with the extended majority rule (MRE) bootstrapping method [19].

3. Results and discussion

We determined the mitochondrial genome of *B. goliath* to be 15,229 bp in length and the complete ribosomal operon to be 9296 bp. The sequences have been deposited in GenBank under accession numbers KR703278 and KR703279 respectively. The mitogenome codes for 12 protein-coding genes, 22 tRNA genes, and the large and small rRNA genes and contains two large non-coding regions (NC1 & NC2) (Table 1); see Fig. 1. The largest of the non-coding regions (NC1) contains a 260 bp feature that is repeated four times (R1–R4). All genes were transcribed in the same direction. The overall AT content is 55.6% with a GC skew of 0.36 and AT skew of −0.49 showing strong compositional bias towards guanine and thymine. Nucleotide composition of the entire mt genome is A = 2175 (14.3%), T = 6292 (41.3%), G = 4928 (32.4%), and C = 1834 (12.0%). Gene order of protein-coding and ribosomal RNA genes follows the basic pattern of other trematodes (except some derived species of *Schistosoma*; [20,21]). In contrast, two tRNA genes (tRNA-Tyr, tRNA-Glu) occupy uniquely altered positions relative to protein-coding genes when compared to all other digenean mt genomes [17]. The differences in the relative positions of tRNA-Tyr and tRNA-Glu as compared to other published mitogenomes of trematodes suggest that denser taxonomic sampling may reveal these as useful phylogenetic or taxonomic markers.

Table 1
Positions and nucleotide sequence lengths of individual genes in the mt genome of *Brachycladium goliath* as well as codons (start and stop), anticodons and the length of predicted proteins.

Gene	Positions of nt	Sequence lengths		Codons		Anticodon
		No. of nt	No. of aa	Start	Stop	
cox1	14–1591	1,578	526	GTG	TAA	
trnT	1556–1618	63				TGT
rrnL	1619–2592	1,006				
trnC	2593–2657	65				GCA
rrnS	2658–3405	825				
cox2	3406–4014	609	203	ATG	TAA	
nad6	4023–4472	450	150	GTG	TAG	
trnL1	4478–4541	64				TAG
trnS2	4539–4610	72				TGA
trnL2	4611–4674	64				TAA
trnR	4677–4738	62				TCG
nad5	4744–6324	1,581	527	GTG	TAG	
trnG	6351–6417	67				TCC
trnY	6468–6533	66				GTA
Non-coding NC1	6534–8037					
Repeat R1	6709–6968					
Repeat R2	6969–7228					
Repeat R3	7229–7488					
Repeat R4	7489–7748					
cox3	8038–8682	645	215	ATG	TAG	
trnH	8702–8764	63				GTG
cytb	8768–9880	1,113	371			
nad4L	9883–10146	264	88	GTG	TAG	
nad4	10107–11384	1,278	426	GTG	TAG	
trnQ	11387–11449	63				TTG
trnF	11456–11518	63				GAA
trnM	11519–11585	67				CAT
atp6	11586–12101	534	178	GTG	TAG	
nad2	12105–12971	867	289	GTG	TAG	
trnV	12973–13040	68				TAC
trnA	13046–13109	64				TGC
trnD	13116–13180	65				GTC
nad1	13181–14083	903	301	ATG	TAG	
trnN	14088–14155	68				GTT
trnE	14166–14232	67				TTC
Non-coding NC2	14233–14530					
trnP	14531–14597	67				TGG
trnI	14600–14662	63				GAT
trnK	14674–14739	66				CTT
nad3	14743–15099	381	127	GTG	TAG	
trnS1	15099–15156	58				GCT
trnW	15161–15226	66				TCA

The ribosomal operon codes for the large and small nuclear rRNA genes as well as two internal transcribed spacers (ITSs) and two external transcribed spacers (ETSs) (Table 2). Of the MiSeq run, 152,049 reads (1.88%) were determined as mitochondrial and 82,226 reads (1.01%) were determined as belonging to the ribosomal operon. There were no problems in reconstructing the full operon with our methodology; a relatively low mismatch threshold (5% error across the entire read length) readily assembled the operon. The pairwise percentage identity of assembled reads (84,559 total) was 99.2%, i.e. only 0.8% of nucleotides in the final assembly differed from the consensus sequence. There are currently few complete ribosomal operons for trematodes available so we restricted our phylogenetic analyses to the mitogenomic data.

The phylogeny, based on an assessment of amino acid sequences from available mitogenomes (Fig. 2), indicated that *Brachycladium* was strongly supported (99% bootstrap) in a clade with paragonimids, heterophyids and opisthorchids; the sister relationship with *Paragonimus* was poorly supported. These results further illustrate the potential use of mitogenomes for trematode phylogentics as promoted by Brabec et al.