International Journal for Parasitology xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

International Journal for Parasitology

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

Succinctus

A linear mitochondrial genome of *Cyclospora cayetanensis* (Eimeriidae, Eucoccidiorida, Coccidiasina, Apicomplexa) suggests the ancestral start position within mitochondrial genomes of eimeriid coccidia

Mosun E. Ogedengbe [a], Yvonne Qvarnstrom [b], Alexandre J. da Silva [b,c], Michael J. Arrowood [b], John R. Barta [a,*]

[a] *Department of Pathobiology, University of Guelph, Guelph, ON, Canada*
[b] *United States Centers for Disease Control and Prevention, Atlanta, GA, USA*
[c] *United States Food and Drug Administration, Laurel, MD, USA*

ARTICLE INFO

Article history:
Received 9 December 2014
Received in revised form 9 February 2015
Accepted 9 February 2015
Available online xxxx

Keywords:
Mitochondrial genome
Eimeriidae
Cyclospora cayetanensis
Apicomplexa
Linear genome
Genomics

ABSTRACT

The near complete mitochondrial genome for *Cyclospora cayetanensis* is 6184 bp in length with three protein-coding genes (*Cox*1, *Cox*3, *Cyt*B) and numerous lsrDNA and ssrDNA fragments. Gene arrangements were conserved with other coccidia in the Eimeriidae, but the *C. cayetanensis* mitochondrial genome is not circular-mapping. Terminal transferase tailing and nested PCR completed the 5′-terminus of the genome starting with a 21 bp A/T-only region that forms a potential stem-loop. Regions homologous to the *C. cayetanensis* mitochondrial genome 5′-terminus are found in all eimeriid mitochondrial genomes available and suggest this may be the ancestral start of eimeriid mitochondrial genomes.

Cyclospora cayetanensis is a protistan disease agent of humans that has been responsible for waterborne and large scale foodborne outbreaks worldwide; this apicomplexan parasite is now recognised as an emerging intestinal pathogen of public health importance (Karanja et al., 2007). In developed countries, the protist has been incriminated in diarrhoeal illnesses linked to imported foods (fresh uncooked vegetables and soft skinned fruits) and has been associated with diarrhoea acquired through increased travel to endemic tropical regions (Shields and Olson, 2003). Particularly in the U.S., the two most recent diarrhoeal outbreaks caused by *C. cayetanensis* were solely associated with the consumption of fresh produce and affected 631 and 304 persons in 2013 and 2014, respectively (http://www.cdc.gov/parasites/cyclosporiasis/outbreaks/index.html). *Cyclospora cayetanensis* was first reported in stools of individuals suffering from protracted intermittent watery non-bloody diarrhoea (CDC Report, 1991). Ortega et al. (1993) first identified the agent as an apicomplexan protist on the basis of oocyst morphology following sporulation and later on intracellular developmental stages. The organism was concluded to belong to the coccidian genus *Cyclospora* (Schneider, 1881), family Eimeriidae (Minchin, 1903) based on morphology of the sporulated oocyst that contains two sporocysts, each possessing a Stieda body and containing two sporozoites.

A single apicomplexan mitochondrial (mt) genome copy is approximately 6–7 kb in length (Gray et al., 2004). A variety of mt genome forms have been described amongst these parasites including circular-mapping mitochondrial genomes (e.g., circular in haemosporinids (Wilson and Williamson, 1997; Feagin et al., 2012) or linear concatemers (multiple mt genome copies joined end to end) in coccidia (Hikosaka et al., 2011)) as well as linear genomes with terminal inverted telomeric repeats in piroplasms (Hikosaka et al., 2010, 2012). Regardless of their structural forms, apicomplexan mt genomes usually possess three protein coding genes encoding cytochrome *c* oxidase subunit I (*Cox*1 cytochrome *c* oxidase subunit III (*Cox*3) and cytochrome b (*Cyt*B) as well as fragmented ssrDNA and lsrDNA. The present study reports on the sequence and structure of the mt genome of *C. cayetanensis*.

Stool samples that were positive for the presence of *C. cayetanensis* by UV fluorescence microscopy were selected for the molecular studies. The samples were collected and used in accordance with the United States Centers for Disease Control and Prevention (CDC)

* Corresponding author at: Department of Pathobiology, University of Guelph, 50 Stone Road East, Guelph, ON, Canada. Tel.: +1 519 824 4120x54017; fax: +1 519 824 5930.
E-mail address: jbarta@uoguelph.ca (J.R. Barta).

http://dx.doi.org/10.1016/j.ijpara.2015.02.006
0020-7519/

Institutional Review Board (IRB) protocol entitled "Use of Human Specimens for Laboratory Methods Research". Three samples were used for the whole mt genome sequencing: two samples from different time points (2011 and 2013) from an endemic area of southeastern Asia, plus a sample collected during the 2013 outbreaks in the USA. Partial *Cox*1 and *Cox*3 genes were sequenced from five additional samples: two samples from the same southeastern Asian location (2012 and 2013); and, three samples from two different outbreaks in the USA during 2013. DNA was extracted using the Universal Nucleic Acid Extraction (UNEX) method as described by Shields et al. (2013) with some adjustments. Approximately 0.5 ml of stool was added to a matrix E bead beating tube (MP Biomedicals, Santa Ana, CA, USA) together with 60 µl of proteinase K (QIAGEN, Valencia, CA, USA) and 600 µl of UNEX buffer (Phthisis Diagnostics, Charlottesville, VA, USA). The tube was incubated at 56 °C for 15 min to allow for proteinase K activity. The mixture was homogenised in a FastPrep-24 tissue and cell disruptor instrument (MP Biomedicals) at a speed of 6.0 m/s for 1 min. The sample was then centrifuged at maximum speed (>13,000*g*) for 1 min to pellet the debris. The supernatant was collected and passed through a DNeasy mini spin binding silica column (QIAGEN). Following two wash cycles using ethanol-containing wash buffers, the DNA was eluted from the column in 80 µl of AE buffer (QIAGEN). The eluted filtrate was further purified by passing through a Zymo-Spin IV-HRC column (Zymo Research Corp., Irvine, CA, USA). DNA samples were confirmed positive for *C. cayetanensis* by real-time PCR as described by Verweij et al. (2003).

Initial attempts to amplify near-complete mt genomes using methods that had worked reliably with various *Eimeria* spp. (see Ogedengbe et al., 2013, 2014) failed repeatedly with *C. cayetanensis*. Thereafter, shorter regions of the mt genome were amplified with primers targeting conserved regions of other apicomplexan mitochondrial genomes. The resulting PCR products were purified, sequenced directly using internal sequencing primers as necessary, and readily assembled into a partial mt genome using the de novo assembler within Geneious (www.geneious.com). Repeated attempts were made to complete the mt genome by amplifying across the 'gap' (assuming a circular or linear concatenated genome) with three pairs of additional amplification primers (Table 1). These efforts failed repeatedly with *C. cayetanensis* but succeeded in the case of coccidia in the genera *Eimeria, Caryospora* and *Isospora* (e.g., Lin et al., 2011; Ogedengbe et al., 2013, 2014. 2015; Ogedengbe and Barta, 2015), all of which possess circular-mapping mt genomes; *Eimeria tenella* was used as positive control for all such PCRs due to its linear concatenated genome.

Under the assumption, based on PCR results, that the mt genome of *C. cayetanensis* was linear, bulk cellular DNA was tailed using a terminal deoxynucleotidyl transferase (TdT) tailing method previously described by Hikosaka et al. (2012) with minor modifications. Briefly, sample DNA (75 ng) in 5 µl of nuclease-free water was denatured at 94 °C for 5 min and then immediately used in a 25 µl of 3′-tailing reaction consisting of 0.0125 µmol of dCTP (Clonetech Laboratories, Inc., Mountain View, CA, USA), 7 U of TdT enzyme (Clonetech) and 0.02% BSA (w/v) in 1× TdT buffer (Clonetech) for 30 min at 37 °C. At the conclusion of the tailing reaction, the TdT enzyme was heat inactivated at 65 °C for 10 min (see Hikosaka et al., 2012). The poly-C-tailed genomic DNA was then used as template for a pair of nested PCRs. In the first PCR, 2 µl of the end-labelled genomic DNA (i.e., 6 ng) was used as template in a reaction mixture containing 1.25 units of Platinum Taq Polymerase (Invitrogen, Carlsbad, CA, USA), 1× PCR buffer (Invitrogen), 2.5 mM $MgCl_2$, 200 µM dNTPs and 0.4 µM each of the required primers. For the amplification of the 5′-end of the genome, an mt genome-specific primer 'q_Eim_CytB_398R' (5′-CCCCAGWARCTCATYTGACCCCA-3′) was used with a poly-G-containing anchor primer 'Telo_F_polyG' (5′-GGCCACGCGTCGACTAGTACGGGGGGGGGGGGGGG-3′). For the amplification of the 3′-end of the genome, an mt genome-specific WG-MT_5416F (5′-GGTCCAGATAAGCGATCTCATG-3′) was paired with the same poly-G-containing anchor primer ('Telo_F_polyG'). Cycling conditions were initial denaturation at 95 °C for 2 min followed by 40 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 45 s and final extension step at 72 °C for 7 min. Following the primary PCR, 1 µl from each of the first PCRs was used as template for amplification reactions of 5′- and 3′-ends. The 5′-end was amplified using amplification primer Telo_F (5′-GGCCACGCGTCGACTAGTAC-3′) with mt genome-specific WG-MT_63R (5′-CTGGTATGGATGGATAACACT-3′) under the following cycle conditions: initial denaturation at 95 °C for 2 min, 40 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 20 s, and a final extension step at 72 °C for 7 min. Similarly, the 3′-end was amplified with the primer Telo_F combined with mt genome-specific WG-MT_5813R (5′-AGGTGCTCAGGGTCTTACCG-3′) under slightly different cycle conditions: initial denaturation at 95 °C for 2 min, 40 cycles of 94 °C for 30 s, 63 °C for 30 s, 72 °C for 40 s and final extension step at 72 °C for 7 min. After electrophoresis, amplification products were excised, purified and submitted for sequencing using the appropriate *Cyclospora*-specific primer used in the final amplification. The sequence data resulting from the mt genome tailing and subsequent nested PCR products was mapped onto the assembled partial *C. cayetanensis* mt genome

Table 1
Amplification primers used to obtain the mitochondrial genome of *Cyclospora cayetanensis* and confirm its linear nature.

Fragment	Primer ID	Nucleotide sequence (5′–3′)	Size in bp (position within mitochondrial genome[a])	References
Fragment 1	Cocci_MT-WG-F	TACACCTAGCCAACACGAT	1807 (23–1829)	Ogedengbe et al. (2014)
	Cyclo_COI_473R	ATACCCGCAAGAGCTAAACC		Ogedengbe et al. (2013)
Fragment 2	qPCR400-F	GDTCAGGTRTTGGTTGGAC	803 (1723–2525)	Ogedengbe et al. (2013)
	COI_1202R	CCAAKRAYHGCACCAAGAGATA		Ogedengbe et al. (2013)
Fragment 3	Cyclo_COI-1085F	CTCCGCTCTAGATGTTGCTT	2030 (2442–4471)	This study
	Cyclo_CO3_113R	TCACCATTCTTGCTCACTGT		This study
Fragment 4	WG-MT_4140F	AGAAAACCTAAAATCATCATGT	961 (4191–5151)	Ogedengbe et al. (2014)
	Eim_CO3_799R	AAGTGAGTTCGCATGTTTAC		This study
Fragment 5	Cyclo_CO3_219F	AGCTTCTTCTGGGGTGCATAC	1607 (4578–6184)	This study
	Cocci_MT-WG-R	GCAGCTGTAGATGGATGCTT		Ogedengbe et al. (2014)
PCR confirmation of linear genome	WG_MT_344R	GTAGGAATCTRAATTCCCAACC	431 (23–453)	Ogedengbe et al. (2013)
	Cocci_MT_WG_F (positive control)	TACACCTAGCCAACACGAT		Ogedengbe et al. (2014)
	WG_MT_344R	GTAGGAATCTRAATTCCCAACC	No product	Ogedengbe et al. (2013)
	WG_MT_6219F	GCATCCATCTACAGCTGCGG		Ogedengbe et al. (2013)
	WG_MT_344R	GTAGGAATCTRAATTCCCAACC	No product	Ogedengbe et al. (2013)
	WG_MT_5416F	GGTCCAGATAAGCGATCTCATG		Ogedengbe et al. (2013)

[a] Positions reported are relative to the complete mitochondrial (mt) genome of *C. cayetanensis* (GenBank KP658101); complete 5′-end was generated using terminal deoxynucleotidyl transferase tailing followed by nested PCR and sequencing but 3′-terminus has not been similarly completed.

Please cite this article in press as: Ogedengbe, M.E., et al. A linear mitochondrial genome of *Cyclospora cayetanensis* (Eimeriidae, Eucoccidiorida, Coccidiasina, Apicomplexa) suggests the ancestral start position within mitochondrial genomes of eimeriid coccidia. Int. J. Parasitol. (2015), http://dx.doi.org/10.1016/j.ijpara.2015.02.006