International Journal for Parasitology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

International Journal for Parasitology

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

Next Generation Sequencing uncovers within-host differences in the genetic diversity of *Cryptosporidium gp60* subtypes

Alireza Zahedi [a], Alexander William Gofton [a], Fuchun Jian [b], Andrea Paparini [a], Charlotte Oskam [a], Andrew Ball [c], Ian Robertson [a], Una Ryan [a,*]

[a] *Vector and Waterborne Pathogen Research Group, School of Veterinary and Life Sciences, Murdoch University, Perth, Australia*
[b] *Henan Agricultural University, Henan, China*
[c] *WaterNSW, Sydney, Australia*

ARTICLE INFO

Article history:
Received 10 February 2017
Received in revised form 23 March 2017
Accepted 31 March 2017
Available online xxxx

Keywords:
Cryptosporidium
Within-host diversity
Glycoprotein 60 (*gp60*)
Sanger
Next Generation Sequencing (NGS)

ABSTRACT

The extent of within-host genetic diversity of parasites has implications for our understanding of the epidemiology, disease severity and evolution of parasite virulence. As with many other species, our understanding of the within-host diversity of the enteric parasite *Cryptosporidium* is changing. The present study compared Sanger and Next Generation Sequencing of glycoprotein 60 (*gp60*) amplicons from *Cryptosporidium hominis* ($n = 11$), *Cryptosporidium parvum* ($n = 22$) and *Cryptosporidium cuniculus* ($n = 8$) DNA samples from Australia and China. Sanger sequencing identified only one *gp60* subtype in each DNA sample: one *C. hominis* subtype (IbA10G2) ($n = 11$), four *C. parvum* subtypes belonging to IIa ($n = 3$) and IId ($n = 19$) and one *C. cuniculus* subtype (VbA23) ($n = 8$). Next Generation Sequencing identified the same subtypes initially identified by Sanger sequencing, but also identified additional *gp60* subtypes in *C. parvum* and *C. cuniculus* but not in *C. hominis*, DNA samples. The number of *C. parvum* and *C. cuniculus* subtypes identified by Next Generation Sequencing within individual DNA samples ranged from two to four, and both *C. parvum* IIa and IId subtype families were identified within the one host in two samples. The finding of the present study has important implications for *Cryptosporidium* transmission tracking as well as vaccine and drug studies.

1. Introduction

Humans and animals often become co-infected with different species and genotypes of the same parasite genus, resulting in within-host parasite interactions (Holmes and Price, 1986; Read and Taylor, 2001; Choisy and de Roode, 2010; Seppälä and Jokela, 2016). Importantly, the presence of co-infecting parasite species/genotypes within a host can potentially modify parasite fitness by allowing them to adapt to different selection pressures and can drive the evolution of parasite virulence and alter host susceptibility to other parasites, infection duration, disease severity, transmission risks, clinical symptoms and consequently treatment and prevention strategies (Vaumourin et al., 2015; Grinberg and Widmer, 2016; Seppälä and Jokela, 2016). Therefore within-host parasite interactions have important repercussions for human or animal health. For instance, parasite co-infections within a single host can result in gene exchange via recombination. This can drive parasite evolution by making the parasites more resistant to drugs. Interactions among co-infecting parasite species, genotypes and subtypes of the same parasite genus can also modify co-evolutionary dynamics between the host and parasites. In addition, parasite interactions can help with maintaining genetic variation in parasite traits such as infectivity and virulence which are crucial components of pathogen fitness and are important to better understand disease dynamics and the changing epidemiology of parasitic diseases (Seppälä et al., 2012; Vaumourin et al., 2015; Seppälä and Jokela, 2016).

Cryptosporidium spp are ubiquitous enteric parasites that infect a broad range of hosts including humans and animals (Xiao, 2010). They are a major contributor to moderate-to-severe diarrhoeal disease in developing countries and are second only to rotavirus as a cause of moderate-to-severe diarrhoea in children younger than 2 years (Kotloff et al., 2013). Of the 31 species currently recognised, *Cryptosporidium parvum* and *Cryptosporidium hominis* account for the majority of human infections and typed outbreaks (Xiao, 2010; Ryan et al., 2017), with the exception of *Cryptosporidium cuniculus* which was responsible for a waterborne outbreak in the UK (Puleston et al., 2014). The parasite is transmitted via the faecal-oral route through human to human, animal to human

* Corresponding author. Fax: +61 8 9310 4144.
E-mail address: Una.Ryan@murdoch.edu.au (U. Ryan).

http://dx.doi.org/10.1016/j.ijpara.2017.03.003
0020-7519/

and animal to animal contact, and via contaminated water; therefore hosts are exposed to multiple sources of potentially genetically diverse oocysts (Xiao, 2010; Grinberg and Widmer, 2016). Once ingested, sporozoites excyst from the oocyst, invade the host cells and undergo subsequent rounds of asexual and sexual reproduction.

Currently, the only available drug for human infections (nitazoxanide - Romark Laboratories, Florida, USA), has variable efficacy (Abubakar et al., 2007; Amadi et al., 2009) and an effective vaccine has yet to be developed (Mead, 2014; Ryan et al., 2016). Halofuginone lactate (Halocur; Intervet, New Zealand) is commercially available against cryptosporidiosis in dairy calves, with variable efficacy (Trotz-Williams et al., 2011; Almawly et al., 2013). Therefore, *Cryptosporidium* control currently relies mainly on improved sanitation and understanding its transmission dynamics.

Analysis of the extent of within-host genetic diversity in *Cryptosporidium* has been hampered due to the difficulties in culturing this parasite, with clonal lineages derived from individual sporozoites unavailable (Grinberg and Widmer, 2016). Of the multilocus sequencing typing strategies employed to examine within-host genetic diversity, sequence analysis of the glycoprotein 60 (*gp60*) gene is the most common (Xiao, 2010), as it is the most polymorphic locus in the genome (Abrahamsen et al., 2004). Despite the importance of within-host genetic diversity for our understanding of cryptosporidiosis epidemiology, relatively little is known (Cama et al., 2006; Waldron and Power, 2011; Jeníková et al., 2011; Shrestha et al., 2014; Ramo et al., 2014, 2016). Most studies have relied on conventional PCR and Sanger-based genotyping methods, and automated fragment analysis, however a major limitation of these approaches is their inability to resolve complex DNA mixtures and detect low-abundance intra-isolate variants (Paparini et al., 2015; Grinberg and Widmer, 2016).

Next Generation Sequencing (NGS) of amplicons offers the advantage of massive parallelization of sequencing reactions to more effectively identify low-abundance genotypes in mixed infections. To date, only one study has examined the extent of intra-isolate diversity of *Cryptosporidium* at the *gp60* locus using NGS (Grinberg et al., 2013). In that study, NGS analysis of two *C. parvum* samples from one geographic location (New Zealand) revealed much higher levels of intra-isolate diversity compared with Sanger sequencing. In the present study, we examined intra-host genetic diversity of a much larger number of *Cryptosporidium* samples ($n = 41$) from three different species (*C. hominis*, *C. parvum* and *C. cuniculus*) and from two distinct geographic regions (Australia and China), using both NGS and conventional Sanger sequencing at the *gp60* locus to better understand the epidemiology of this important parasite. Animal faecal samples in Australia were collected from watersheds within the WaterNSW (New South Wales) area of operations and included two dairy farms and faecal samples collected from the surrounding bushland. Faecal samples from China were collected from a cattle breeding centre and two dairy farms, all located in Henan province.

2. Materials and methods

2.1. Sample collection and processing

A total of 41 DNA samples positive for *Cryptosporidium*, belonging to three *Cryptosporidium* spp. (as determined by Sanger sequencing – see Section 2.3), were analysed in the present study; *C. parvum* ($n = 22$) from cattle (*Bos taurus*), *C. hominis* ($n = 11$) from Eastern Grey kangaroos (*Macropus giganteus*), and *C. cuniculus* ($n = 8$) from rabbits (*Oryctolagus cuniculus*).

2.2. DNA isolation

Upon collection, faecal samples were stored at 4 °C until analysed. Following five cycles of freeze–thaw, genomic DNA was extracted from 250 mg of each faecal sample using a Power Soil DNA Kit (MO BIO, Carlsbad, California, USA). Extraction blanks (no faecal sample) were used in each extraction group. Purified DNA was stored at −20 °C prior to PCR. DNA extraction and post-DNA extraction procedures were performed in separate dedicated laboratories.

2.3. Sanger sequencing

All samples were initially identified to species level at the 18S locus using nested PCR amplification and Sanger sequencing of a fragment of the 18S locus as previously described (Silva et al., 2013). Samples were then subtyped at the *gp60* locus using a nested PCR to amplify an approximately 400 bp product using the primers AL3531 (5′-ATAGTCTCCGCTGTATTC-3′) and AL3533 (5′-GAGATATATCTTGGTGCG-3′) for the primary PCR, and AL3532 (5′-TCCGCTGTATTCTCAGCC-3′) and LX0029 (5′-CGAACCACATTACAAATGAAGT-3′) for the secondary PCR (Sulaiman et al., 2005). Each 25 μl PCR mixture contained 1 μl of genomic DNA, 1× Go Taq PCR buffer (KAPA Biosystems, South Africa), 3.75 mM $MgCl_2$, 400 μM of each dNTP, 0.4 μM of forward and reverse primers and 1 U of Kapa DNA polymerase (MO BIO). The PCR cycling conditions were modified and consisted of an initial denaturation at 94 °C for 3 min and then 40 cycles of 94 °C for 45 s, 54 °C for 45 s and 72 °C for 1 min, followed by a final extension step at 72 °C for 7 min. PCR contamination controls were used including negative controls. PCR setup and DNA handling procedures were performed in separate and dedicated exclusion hoods; PCR and post-PCR procedures were performed in separate dedicated laboratories.

Gel electrophoresis was used to separate the amplified DNA fragments from the secondary PCR products at the *gp60* locus, which were subsequently purified for sequencing using an in-house filter tip method as previously described (Yang et al., 2013). Purified PCR products were sequenced independently in both directions using an ABI Prism™ Dye Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, California, USA) according to the manufacturer's instructions and with a 54 °C annealing temperature. Sanger sequencing chromatogram files were imported into Geneious Pro 8.1.6 (Kearse et al., 2012), and the nucleotide sequences of each gene were analysed and aligned with reference sequences from GenBank using Clustal W (http://www.clustalw.genome.jp).

2.4. Next Generation Sequencing (NGS)

Partial *Cryptosporidium gp60* gene sequences were amplified for NGS on the MiSeq (Illumina) platform using the same assay described for Sanger sequencing (Sulaiman et al., 2005), with the exception that secondary PCR primers were modified to contain MiSeq adapter sequences on the 5′ end, as per standard protocols for the MiSeq platform (Illumina Demonstrated Protocol: Metagenomic Sequencing Library Preparation). PCRs were performed in 25 μl volumes containing PCR buffer (KAPA Biosystems), 2 mM $MgCl_2$, 0.01 mg of BSA (Fisher Biotech, Australia), 1 mM dNTPs (Fisher Biotech), 0.4 μM of each primer and 0.5 U of KAPA Taq DNA Polymerase (KAPA Biosystems). Primary PCRs used 2 μl of DNA as a template and secondary reactions contained 1–2 μl of the primary product as a template. All PCRs contained no-template controls and extraction reagent blank controls. All PCRs were performed with an initial denaturation at 95 °C for 5 min, fol-

Please cite this article in press as: Zahedi, A., et al. Next Generation Sequencing uncovers within-host differences in the genetic diversity of *Cryptosporidium gp60* subtypes. Int. J. Parasitol. (2017), http://dx.doi.org/10.1016/j.ijpara.2017.03.003