International Journal for Parasitology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

International Journal for Parasitology

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

Invited Review

Cryptosporidium within-host genetic diversity: systematic bibliographical search and narrative overview

Alex Grinberg [a,*], Giovanni Widmer [b]

[a] *Massey University, Institute of Veterinary, Animal and Biomedical Sciences, Infectious Disease Group, Private Bag 11 222, Palmerston North, New Zealand*
[b] *Tufts University, Cummings School of veterinary Medicine, Department of Infectious Disease and Global Health, 200 Westboro Road, North Grafton, MA 01536, USA*

ARTICLE INFO

Article history:
Received 19 January 2016
Received in revised form 29 February 2016
Accepted 6 March 2016
Available online xxxx

Keywords:
Cryptosporidium
Protozoa
Genetic diversity
Genotyping
Genomics
Population genetics
Recombination

ABSTRACT

Knowledge of the within-host genetic diversity of a pathogen often has broad implications for disease management. *Cryptosporidium* protozoan parasites are among the most common causative agents of infectious diarrhoea. Current limitations of in vitro culture impose the use of uncultured isolates obtained directly from the hosts as operational units of *Cryptosporidium* genotyping. As the validity of this practice is centred on the assumption of genetic homogeneity of the parasite within the host, genetic studies often take little account of the within-host genetic diversity of *Cryptosporidium*. Yet, theory and experimental evidence contemplate genetic diversity of *Cryptosporidium* at the within-host scale, which is not easily identified by end-point genotyping methods ill-suited for the resolution of DNA mixtures. We performed a systematic bibliographical search of the occurrence of within-host genetic diversity of *Cryptosporidium* parasites in epidemiological samples. Our results indicate that genetic diversity at the within-host scale, in the form of mixed species or intra-species diversity, has been identified in a large number ($n = 55$) of epidemiological surveys of cryptosporidiosis in variable proportions, but has often been treated as a secondary finding and not analysed further. As in malaria, there are indications that the scale of this diversity varies between geographical regions, perhaps depending on the prevailing regional transmission pathways. These results provide a significant knowledge base from which to draw alternative population genetic structure models, some of which are discussed in this paper.

1. Introduction

In many infectious diseases there is a presence of distinct microbial lineages within the host, and the within-host genetic diversity of the pathogen is often associated with its ability to adapt to selective pressures exerted during the infection. For instance, in hepatitis C and HIV infections the presence of viral subpopulations enhances evasion from the immune response and resistance to chemotherapy (Wolinsky et al., 1996; McMichael and Phillips, 1997; Farci et al., 2000; Briones et al., 2006). It has been recently reported that approximately 20% of people infected with *Mycobacterium tuberculosis* in KwaZulu-Natal, South Africa, harboured genetically heterogeneous infections, and microbial heterogeneity was associated with increased odds of treatment failure (Cohen et al., 2016). Examples in the protozoan world also abound. Polyclonal infections with *Eimeria tenella* (poultry) and *Theileria annulata* (cattle) seem to be common (Weir et al., 2011; Blake et al., 2015), and mixed infections with *Plasmodium falciparum* clones were already observed in people during the 1990s (Viriyakosol et al., 1995). Furthermore, in some regions most malaria patients may harbour multiple clones (Arnot, 1998; Apinjoh et al., 2015). Other studies have found associations between the within-host genetic diversity of *Plasmodium* spp. and the natural history of malaria, and it is currently recognised that this diversity may influence the evolution of virulence, hamper chemotherapeutic control and potentially promote the emergence of vaccine escape variants (de Roode et al., 2004; Kwiek et al., 2007; Juliano et al., 2010a,b; Tyagi et al., 2013). Thus, knowledge of the genetic diversity of a pathogen within the host may have broad implications for disease management.

Cryptosporidium is a genus of sexually reproducing protozoan parasites of amphibians, fish, reptiles, birds and mammals. Parasites belonging to this genus are major contributors to the burden of paediatric diarrhoea in many developing countries (Kotloff et al., 2013). The intestinal species *Cryptosporidium parvum* and *Cryptosporidium hominis*, in particular, are among the most common and cosmopolitan causes of diarrhoea in people, and the latter species is also a leading cause of diarrhoea in young farmed ruminants (Al Mawly et al., 2015). Despite the availability of the sequences of

* Corresponding author. Tel.: +64 6 3569099; fax: +64 6 3505635.
E-mail address: a.grinberg@massey.ac.nz (A. Grinberg).

http://dx.doi.org/10.1016/j.ijpara.2016.03.002
0020-7519/

the *C. parvum* and *C. hominis* genomes (Abrahamsen et al., 2004; Xu et al., 2004), our understanding of cryptosporidiosis is still incomplete. Efficient anti-*Cryptosporidium* drugs and vaccines are lacking and there is poor knowledge on the role of genetic exchange in the evolution of the parasites' virulence. One of the main reasons for the poor progress in many fields is the limitations of in vitro culture. Despite recent progress (Morada et al., 2015), methods for *Cryptosporidium* culture are still inefficient and clonal lineages derived from individual sporozoites, which are the basic cells carrying the parasite's haploid genome, are not available. As a consequence, *Cryptosporidium* strains remain loosely defined (Cama et al., 2006a) and most research is currently performed using isolates composed of parasites obtained from single infected hosts, without previous culture. The use of uncultured isolates as operational units of *Cryptosporidium* genotyping is also an established practice (Ryan et al., 2014), the working assumption being that a broadly phylogenetically homogeneous parasite infects the host. However, as in malaria, where individuals can harbour several genetically distinct parasites as a result of multiple mosquito bites, or single bites from mosquitoes bearing multiple clones (Talisuna et al., 2007), *Cryptosporidium* within-host diversity may originate from multiple exposures, or a single infection with genetically diverse oocysts. In fact, studies have identified water as a major route of *Cryptosporidium* transmission in some areas (Xiao et al., 2004), and it has been demonstrated that water sources can contain multiple *Cryptosporidium* spp. originating from disparate locations or hosts (Ruecker et al., 2005, 2013; Nichols et al., 2006; Feng et al., 2009). Thus, one would expect to find heterogeneous parasites within hosts exposed to such sources, as well as in secondary infections.

There is long-standing evidence for the existence of within-host genetic diversity in *Cryptosporidium* (Widmer, 1998; Tanriverdi et al., 2003, 2008), but most epidemiological studies of cryptosporidiosis used PCR-based genotyping methods able to resolve the dominant sequence but insensitive to minority variants and ill-suited for resolution of complex DNA mixtures (Suzuki and Giovannoni, 1996; Rochelle et al., 2000; Reed et al., 2002; Liu et al., 2008; Paparini et al., 2015). Furthermore, when identified in the field, the *Cryptosporidium* within-host diversity has been often deemed marginal, defined as 'mixed infections' and not analysed, and only a few groups have pursued further analyses at the within-isolate scale (Cama et al., 2006b; Jeníková et al., 2011; Shrestha et al., 2014). As a consequence, the extent of within-host *Cryptosporidium* diversity in nature remains unknown.

There are practicalities in working with the assumption of a within-host genetic homogeneity of *Cryptosporidium*, and genotyping of uncultured isolates has significantly aided our understanding of the epidemiology of cryptosporidiosis. Nevertheless, from the study of other pathogens we learn that genetic analyses at the within-host level could enhance our ability to answer many unresolved questions on cryptosporidiosis. Here, we present the results of a systematic bibliographical search of the epidemiological evidence for *Cryptosporidium* within-host genetic diversity. We also assemble current understandings of aspects of the parasites' life cycle pertaining to genetic exchange, to hypothesise alternative population genetic structures that challenge the assumption of within-host genetic homogeneity. Finally, the gaps between actual and desired genotyping practices for the study of the genetic structure of *Cryptosporidium* populations are briefly addressed.

2. Within-host *Cryptosporidium* genetic diversity: epidemiological evidence

We aimed to estimate the extent of *Cryptosporidium* within-host genetic diversity in nature based on information published in the scientific literature. We noted that in many papers the genetic heterogeneity revealed in the isolates was not the main focus of the discussion, and when identified, it was defined as 'mixed infections'. Therefore, we interrogated the PubMed database (http://www.ncbi.nlm.nih.gov/PubMed) using the descriptor '*Cryptosporidium*' (all fields); and 'mixed' (all fields). Filters applied were 'species' (values: 'humans'; and 'other animals'); 'text availability' (value: 'abstract'); and 'languages' (value: 'English'). The search was limited to the decade of 2005–2015. This strategy returned all the papers in English, containing the terms '*Cryptosporidium*' and 'mixed', referring to infections in humans and animals for which an abstract was available in PubMed, published between 2005 and 2015. The database was accessed repeatedly from 31 August 2014. Relevant papers retrieved by us elsewhere were also used.

The articles were downloaded and assessed for the type of intra-isolate diversity reported (mixed-species/intra-species diversity). Each article was scored by its evidentiary value as follows: Evidentiary Score 1 (ES1), studies in which spurious intra-isolate genetic diversity was unlikely by virtue of the type of genotyping method used, or in which the possibility of genotyping error was controlled; ES2, studies in which spurious intra-isolate genetic diversity was possible, and this was acknowledged or discussed by the authors; ES3, studies in which spurious intra-isolate genetic diversity was possible, but not discussed.

Of 144 articles initially retrieved, 55 reported intra-isolate genetic diversity, worldwide. Other papers ($n = 89$) reported mixed infections with *Cryptosporidium* and other organisms, or analysed *Cryptosporidium* mixtures in water, or involved in vitro experiments and were discarded. Twelve articles reported intra-isolate genetic diversity in humans and 41 in animal species. Two articles reported mixed infections in both humans and animals. Forty articles reported mixed species and 13 intra-species diversity, whilst two articles reported both occurrences. In many articles the diversity was observed in a non-negligible proportion of the analysed isolates, and in one it was reported in 42% of the isolates. A total of 25/55 (45%) articles received an ES1 score, five received ES2 and 25 received ES3. Two articles could not be scored due to missing information. The details of the surveyed articles are presented in Supplementary Table S1.

3. Origins of *Cryptosporidium* within-host genetic diversity

The bibliographic search indicated a substantial number of epidemiological studies reporting *Cryptosporidium* intra-isolate genetic diversity. Hence, to establish the origins of this diversity it is necessary to review aspects of the parasites' life cycle pertinent to genetic exchange. We note as a limitation, that much of what we assume today about this life cycle is based on inferences from electron microscopic studies performed before different species could be differentiated by genotyping (Tzipori, 1986; Tzipori and Griffiths, 1998; Tzipori and Ward, 2002).

Infections with *C. parvum* (or as a generalisation, with any intestinal *Cryptosporidium* sp.) are acquired through the ingestion of oocysts, and the life cycle culminates with excretion in the faeces of millions of highly resistant oocysts containing four haploid sporozoites. With some exceptions (Karanis et al., 2008; Koh et al., 2013), it is generally assumed that *Cryptosporidium* only replicates inside the host. In the gut, sporozoites excyst from the oocyst, invade the host cells and undergo subsequent rounds of conservative cell divisions in an intracellular but extracytoplasmatic position within the enterocytes, followed by differentiation into micro- or macrogametocytes and mating of gametes, resulting in the formation of a transient diploid zygote. It is not known

Please cite this article in press as: Grinberg, A., Widmer, G. *Cryptosporidium* within-host genetic diversity: systematic bibliographical search and narrative overview. Int. J. Parasitol. (2016), http://dx.doi.org/10.1016/j.ijpara.2016.03.002