International Journal for Parasitology 43 (2013) 971–982

Contents lists available at ScienceDirect

International Journal for Parasitology

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

Invited Review

Getting to the guts of the matter: The status and potential of 'omics' research of parasitic protists of the human gastrointestinal system

Aaron R. Jex [a,*], Anson V. Koehler [a], Brendan R. Ansell [a], Louise Baker [a], Harin Karunajeewa [b], Robin B. Gasser [a]

[a] Faculty of Veterinary Science, The University of Melbourne, Parkville, Victoria 3010, Australia
[b] Western Health, Footscray, Victoria 3011, Australia

ARTICLE INFO

Article history:
Received 3 May 2013
Received in revised form 7 June 2013
Accepted 7 June 2013
Available online 18 July 2013

Keywords:
Diarrhoeal pathogens
Genomics
Systems biology
Giardia
Cryptosporidium
Entamoeba

ABSTRACT

Parasitic protists are a major cause of diarrhoeal illnesses in humans globally. Collectively, enteric pathogens exceed all other forms of infectious disease, in terms of their estimated global prevalence and socio-economic impact. They have a disproportionately high impact on children in impoverished communities, leading to acute (diarrhoea, vomiting, dehydration and death) and chronic disease (malabsorption, malnutrition, physical and cognitive stunting and predisposition to chronic, non-communicable disease) consequences. However, historically, investment in research and disease control measures has been disproportionately poor, leading to their current classification as neglected pathogens. A sound understanding of their biology is essential in underpinning detection, treatment and control efforts. One major tool in rapidly improving our knowledge of these parasites is the use of biological systems, including 'omic' technologies. In recent years, these tools have shown significant success when applied to enteric protists. This review summarises much of this knowledge and highlights the significant remaining knowledge gaps. A major focus of the present review was to provide a perspective on a way forward to address these gaps using advanced biotechnologies.

1. Introduction

Globally, diarrhoeal pathogens (viruses, bacteria and parasites) are among the most significant causes of human morbidity and mortality, exceeding nearly all other forms of infectious disease. Each year such pathogens cause an estimated 1.7 billion cases of infectious diarrhoea in children in low to middle-income countries alone (Fischer Walker et al., 2012) and, after perinatal disorders, are the second leading cause of death in this group (~25% of all under-5 years childhood mortality globally (Anon., 2004)). Recent estimates of the loss of disability adjusted life-years (DALYs) attributable to infectious diarrhoea exceed 300 million (Ricci et al., 2006). Although these figures place enteric pathogens amongst the commonest and most devastating causes of human morbidity and mortality worldwide (Anon., 2004), it is highly likely that they still represent a significant underestimate, as many enteric infections go unreported, undiagnosed and/or untreated in endemic regions of the world. Emerging evidence now suggests that early childhood diarrhoea has a previously unrecognised, but profound, impact on long-term physical and cognitive development (Kosek et al., 2003; Ricci et al., 2006). Particularly troubling is growing evidence that in impoverished populations recurrent diarrhoeal disease in the first 2 years of childhood contributes to an estimated (average) 10 cm growth and 10 IQ point shortfall by the time a child is 7–9 years of age (Guerrant et al., 2012). In part, this stunting is the result of pathophysiological changes in the gastrointestinal tract itself, with frequent/chronic diarrhoea leading often to permanent atrophy of the intestinal villi, and causing long-term changes to the gastrointestinal microfauna, leading to diminished nutrient/fluid absorption and pro-longed gastrointestinal dysfunction (Guerrant et al., 2012). Moreover, evidence now suggests that high burdens of disease due to gastrointestinal pathogens in early life can, paradoxically, predispose humans to increased risks of obesity and associated metabolic syndromes in adult life (Guerrant et al., 2012). The combination of (i) acute gastroenteric illness, (ii) indirect effects of chronic or recurrent disease and (iii) predisposition to chronic metabolic disease in later adult life, have been described as the "triple burden" of gastrointestinal pathogens (Guerrant et al., 2012). Taken together, their effects on human capital and the future health-care burden are likely to have profound consequences that go beyond health and impact significantly on potential for economic development in these countries.

To an overwhelmingly disproportionate extent, the long-term consequences of diarrhoeal diseases are linked to the frequency and chronicity of infections during childhood (being most

* Corresponding author. Tel.: +61 3 9731 2294; fax: +61 3 8344 7374.
E-mail address: ajex@unimelb.edu.au (A.R. Jex).

0020-7519/$36.00
http://dx.doi.org/10.1016/j.ijpara.2013.06.005

significant in children experiencing ⩾3 episodes of enteritis per year, when each episode lasts ~2–3 weeks (Guerrant et al., 2012)). In most instances, chronic cases are caused by parasitic protists, including *Cryptosporodium parvum* and *Cryptosporidium hominis* (see Jex et al., 2010), the *Giardia intestinalis* (syn. *Giardia duodenalis* and *Giardia lamblia*) complex (Cacciò and Ryan, 2008) and *Entamoeba histolytica* (see Haque et al., 2003). Indeed, *G. intestinalis* is considered among the commonest parasites of humans globally, irrespective of socioeconomic status, being estimated as the cause of >200 million cases of symptomatic illness and >1 billion total infections annually (Feng and Xiao, 2011). Increasingly, the contributions of other pathogenic, enteric protists to human disease are being recognised. These taxa include *Enterocytozoon bieneusi* (see Mathis et al., 2005), *Cyclospora cayetanensis* (see Helmy, 2010; Ortega and Sanchez, 2010), *Blastocystis hominis* (see Stensvold et al., 2009; Roberts et al., 2013) and *Dientomoeba fragilis* (see Stark et al., 2010). The pathogenicity of some of these (e.g., *Cyclospora* and *Enterocytozoon*), although relatively recently recognised, now appear to be well established (Mathis et al., 2005; Helmy, 2010; Ortega and Sanchez, 2010). In contrast, the pathogenic status of *Blastocystis* and *Dientamoeba* remain somewhat controversial (Stensvold et al., 2009; Stark et al., 2010; Roberts et al., 2013). However, intriguingly, these latter protists appear to favour a role in manipulating and modifying the healthy gastrointestinal microfauna, and therefore have potential for contributing to long-term health problems including autoimmune disorders such as irritable bowel disease (IBD) (Johnson et al., 2004; Stark et al., 2007). Many enteric pathogens are now considered to be seriously neglected and in urgent need of increased attention as agents of disease (Savioli et al., 2006). Although some drugs are available for some enteric protists (e.g., *Giardia*; seeBusatti et al., 2009), many of these pathogens (e.g., *Cryptosporidium*; see Jex et al., 2010) remain essentially untreatable, aside from standard supportive therapy (including rehydration and anti-inflammatory treatment).

The genomics age, heralded by the advent of next generation sequencing technologies (Mardis, 2008), has the potential to stimulate advances in our understanding of many critically important pathogens, providing insights into their epidemiology, evolution, and molecular and cellular biology. These technologies provide real prospects for making tangible impacts on global health through improved resources for developing molecular diagnostic tools and identifying novel targets for urgently needed new drugs. Indeed, an expert panel of clinicians and research scientists (Daar et al., 2002) has ranked improved tests for infectious diseases, genomics research and drug discovery as first, fourth and sixth, respectively, among a list of the top 10 biotechnologies based on their potential for improvement of public health in impoverished countries, with improved diagnostic tools for enteric pathogens alone estimated to deliver a benefit to the global health burden equivalent to 50% of that estimated for malaria (Ricci et al., 2006). Acknowledging the importance of this field, we review the current status of 'omics' research (with a primary focus on genomics) of parasitic protists of the human gastrointestinal tract, identify major knowledge gaps and provide a perspective on the future directions of biological research of these most important pathogens.

2. Major pathogens: *Cryptosporidium*, *Giardia* and *Entamoeba*

Among protists parasitising the gastrointestinal tract, *C. hominis* and *C. parvum* (see Jex et al., 2010), *G. intestinalis* (i.e., assemblages A and B: Cacciò et al., 2005; Monis et al., 2009) and *E. histolytica* (see Stauffer and Ravdin, 2003) are best characterised at the biological and genomic levels. Recent estimates indicate that these three pathogens cause a global incidence of gastrointestinal disease of >1 billion, with this value likely representing an underestimation due to limited diagnosis and/or under-reporting in endemic regions of the world (Ricci et al., 2006). Extensive investigation of the genomics and molecular biology of these parasites within the last decade has led to substantial insights into their biology and the diseases that they cause, providing genuine prospects for new treatments to underpin their effective control.

2.1. Cryptosporidium

A major advance in our understanding of the molecular biology of *Cryptosporidium* has arisen from the sequencing of the genomes of *C. parvum* and *C. hominis* (see Abrahamsen et al., 2004; Xu et al., 2004). The genomes of these closely related species (98% identity genome-wide) are each ~9 million bases (Mb) in size and encode ~4000 genes (Abrahamsen et al., 2004; Xu et al., 2004). The genome of *C. parvum* is essentially fully assembled (13 scaffolds representing eight chromosomes; see www.cryptodb.org), whereas the *C. hominis* genome still has some gaps (90 scaffolds; see www.cryptodb.org). Energy generation (i.e., ATP) in these parasites is dependent upon anaerobic glycolysis, with no evidence of a mitochondrial genome or any other components of cellular or aerobic respiration (Abrahamsen et al., 2004; Xu et al., 2004). Similarly, there is no evidence of pathways for fatty acid/protein digestion, the nitrogenous cycles or the shikimate pathway (Xu et al., 2004). The reduction of these pathways leads to significant restrictions in the biosynthetic abilities of *Cryptosporidium*, likely leading to a major dependency on the host cell. For example, these parasites appear largely unable to synthesize most nucleotides or amino acids and, instead, seem to have a variety of pathways for salvaging small molecules, and the conversion between and among amino acids or nucleotides (Abrahamsen et al., 2004; Striepen and Kissinger, 2004; Xu et al., 2004). Bottlenecks in these pathways have been a significant focus (Jex et al., 2011b), with the most conspicuous example being the conversion of adenosine monophosphate to guanosine monophosphate, which is exclusively mediated by inosine 5′ monophosphate dehydrogenase (IMPDH) (Abrahamsen et al., 2004). Such bottlenecks have significant prospects as drug targets (Jex et al., 2011b), and are clearly worthy of further exploration and assessment for this genus.

Structural and surface molecules are also of major interest in relation to the *Cryptosporidium* genome. Aside from the apical complex, which is comparable in composition in *Cryptosporidium* to that reported for other apicomplexans (Xu et al., 2004), the variant-specific surface proteins (VSPs) are among most extensively studied groups of proteins encoded by an apicomplexan, especially for species of *Plasmodium*, in which they contribute to immunoevasion (Reeder and Brown, 1996). Although *Cryptosporidium* does not appear to encode VSPs (Abrahamsen et al., 2004), several large groups of cell-surface glycoproteins are present. Unlike *Plasmodium* VSPs, *Cryptosporidium* surface glycoproteins are involved in cell–cell adhesion and other mechanisms associated with infection and invasion (Strong and Nelson, 2000). Intriguingly, the 15 kDa glycoprotein (GP15), which assists with adhesion of the *Cryptosporidium* sporozoite to the host epithelial cell (Strong and Nelson, 2000), has shown some potential as a vaccine candidate, with the recombinant GP15 protein stimulating key aspects of the host immune response (i.e., interferon-γ production) in cultures of peripheral mononuclear blood cells collected from seropositive patients (Preidis et al., 2007).

Current genomic research of *Cryptosporidium* is now focused on exploring some key biological questions. Significant epidemiological research, particularly underpinned by PCR-based technologies (Jex et al., 2008), has revealed high levels of host-specificity and host-adaptation for species of *Cryptosporidium* (Xiao et al., 2004). A key appears to be understanding why *C. hominis* appears to be