The genetic basis of host range and virulence in human-infective *Cryptosporidium spp*.

Johanna L. T. Nader, B.A., MSc.

A Thesis submitted to the University of East Anglia in partial fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

University of East Anglia

Norwich Medical School

June 2017

DECLARATION

I declare that the content of this thesis entitled “The genetic basis of host range and virulence in human-infective *Cryptosporidium spp*.” was undertaken and completed by myself, unless otherwise stated, and has not been submitted in support of an application for another degree of qualification in this or any other university or institution.

.....

Johanna Lucia Thorbjørnsrud Nader

University of East Anglia

June 2017

ACKNOWLEDGEMENTS

No one who achieves success does so without acknowledging the help of others. The wise and confident acknowledge this help with gratitude.

-- Alfred North Whitehead (1861-1947)

I would like to hereby express my sincerest gratitude to everyone who supported and inspired me during the course of my PhD. The academic developments and achievements I made would not have been possible without the valuable guidance and teaching I received from my supervisors and their associates. I thank first and foremost my supervisor Dr. Kevin Tyler, who supervised every aspect of my PhD with an infinite amount of knowledge, insight, and patience. I also thank my secondary supervisors Dr. Paul Hunter and Dr. Maha Bouzid for their continuous support in helping me drive my research forward, and for encouraging me to approach my work with innovation and creativity. Special recognition is extended to our AQUAVALENS collaborators at the *Cryptosporidium* Reference Unit in Wales (Dr. Rachel Chalmers, Dr. Kristin Elwin, and Dr. Guy Robinson), who generously gave me their expert advice on all matters cryptosporidiosis-related and a seemingly endless supply of *Cryptosporidium spp*. DNA and oocysts. I would also like to thank our AQUAVALENS collaborators at the University of Vienna in Austria (Dr. Silvia Cervero-Arago, Dr. Alexander Kirschner, and Stefan Jakwerth), who taught me many of the practical skills I needed to complete the laboratory-based component of my PhD, and who warmly hosted me during my extended visit to the Institute for Hygiene and Applied Immunology. Tremendous gratitude is especially reserved for Dr. Cock van Oosterhout, without whom I would have never understood the true value or meaning of a polymorphism, and whose immeasurable help throughout my PhD cannot be acknowledged and appreciated enough. I am equally indebted to Anne Conde, whose genuine kindness and positive outlook on life acted as a constant source of comfort for me throughout my entire PhD. Infinite thanks also go to Mike van der Es the AQUAVALENS finance guru, without whom I would not have been able to finance my consumables shopping sprees and wonderful conference trips abroad. I finally thank all of my remarkable friends, comrades, and academic peers who kept me entertained and level-headed during my time in Norwich. I am particularly thankful to my dear friends and partners-in-crime Cecilie Tsesmetsis and Kirsten Duda, whose unconditional support, personal integrity and boundless enthusiasm are invaluable to me.

I dedicate this thesis to my parents, Ingrid and Thomas, whose individual accomplishments and perspectives act as a constant source of inspiration and motivation for everything I aspire to achieve. Takk for alt dere har gjort for meg - Ich wäre nichts ohne eure endlose Ermutigung und Unterstützung.

ABSTRACT

Cryptosporidium is a leading cause of diarrhoeal disease worldwide. 26 species and more than 70 genotypes have been characterised to date, but only two species, *Cryptosporidium hominis* and *Cryptosporidium parvum*, account for a majority of human infections. Despite a close genomic relationship between these species (>96.0% nucleotide identities) the former propagates in a predominantly anthroponotic transmission cycle, while the latter exhibits a broad zoonotic host range and appears well-adapted to both humans and cattle. Intra-*C. parvum* host range differences have also been demonstrated through the existence of an anthroponotic subtype commonly referred to as 'IIc'. Using a novel dataset of diverse *C. hominis* and *C. parvum* whole genome sequences to perform comparative genomics analyses between zoonotic and anthroponotic subtypes, a number of divergent loci are herein identified that are subject to heightened polymorphism and selective pressure in both *C. hominis* and *C. parvum* subtype IIc. Glycoproteins, multi-copy gene families and secreted proteins are predominant, and the subtelomeres houses a significant majority of these highly-variable putative virulence factors. A concatenated phylogeny of novel WGS reveals the existence of an anthroponotic subspecies within the *C. parvum* species clade, warranting re-classification of zoonotic versus anthroponotic strains as *C. parvum parvum* and *C. parvum anthroponosum*, respectively. The concatenated phylogeny also demonstrates a hybrid-type IIcA5G3j isolate, which contains a IIc-type GP60 locus but an otherwise predominantly *C. parvum parvum*-type genome. Recombination analyses between this hybrid strain and two common zoonotic (IIa-type) and anthroponotic (IIc-type) *C. parvum* WGS and a *C. hominis* (Ib-type) WGS shows extensive genome-wide recombination between all of these strains, and divergence dating using recombination data reveals a relatively recent estimated time of divergence (<1,000 years ago) between these species and subspecies. Population genetics analyses of *C. hominis* GP60 sequences reveals inter-family disparities in allelic divergence and selective pressure, with subtype famailies Ia and Id exhibiting significantly higher rates of diversification. Evidence of a recent population expansion event is also identified for these two subtypes in *C. hominis* populations from Central Asia and Sub-Saharan Africa. An updated look at whole genome divergence between a 'virulent' Ib-type and 'avirulent' Ia-type *C. hominis* WGS reveals

structural changes at the subtelomeres resulting in subtype-specific coding sequences, and identifies the location, type and degree of polymorphism and selective pressure that accompanies lineage-specific strain emergence and evolution. Finally, the development of a novel In Situ PCR/Fluorescence In Situ Hybridization technique for sensitive and specific detection of *C. p. parvum* oocysts is hereby presented. This demonstrates how in silico findings from preceding chapters identified a suitable subspecies-specific target, and were subsequently translated into an applied molecular method for discrimination of lineages at a subspecific level. Taken together, the results from this thesis have contributed significantly towards increasing our understanding of the genomic factors driving evolution and adaptation in human-infective *Cryptosporidium spp*.

TABLE OF CONTENTS

Chapter 0

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

APPENDICES

A	Adenine
AA	Amino Acid
ABS	Absorbance
AD	Allelic Diversity
AIDS	Acquired Immune Deficiency Syndrome
Ala	Alanine
Arg	Arginine
ART	Antiretroviral Therapy
Asn	Asparagine
Asp	Aspartic Acid
AU	Australasia
BioEdit	Biological Sequence Alignment Editor
BLAST	Basic Local Alignment Search Tool
C	Cytosine
CA	Central Asia
CD	Coding Sequence
CDC	Centers for Disease Control and Prevention
COWP	*Cryptosporidium* Oocyst Wall Protein
CP47	47 kDa Glycoprotein
CRU	*Cryptosporidium* Reference Unit
Cys	Cysteine
DAPI	4',6-diamidino-2-phenylindole fluorescent stain
DN	Rate of non-synonymous substitutions
DNA	Deoxyribonucleic Acid
DnaSP	DNA Sequence Polymorphism software package
DS	Rate of synonymous substitutions
Dxy	Absolute genetic divergence
EA	Eastern Asia
EMBOSS	European Molecular Biology Open Software Suite
EU	Europe
6-FAM	6-Carboxyfluorescein
FAO	Food and Agriculture Organization of the United Nations
FISH	Fluorescence In Situ Hybridization

FITC	Fluorescein isothiocyanate
Fst	Fixation index
G	Glycine
GBD	Global Burden of Disease
GenBank	National Institutes of Health Genetic Sequence Database
Gln	Glutamine
Glu	Glutamic Acid
Gly	Glycine
GP21	21 kDa Glycoprotein
GP60	60 kDa Glycoprotein
GP63	63 kDa Glycoprotein
GPI	Glycosylphosphatidylinositol
HAART	Highly Active Antiretroviral Therapy
HE	Haematoxylin and Eosin
His	Histidine
HIV	Human Immunodeficiency Virus
HKY85	Hasegawa-Kishono-Yano nucleotide substitution model
HSP	Heat Shock Protein
HTLV	Human T-cell Lymphotropic Virus
Ile	Isoleucine
INDEL	Insertion/Deletion
IS-PCR	In Situ Polymerase Chain Reaction
ITS	Internal Transcribed Spacer
KA	Rate of non-synonymous substitutions
KCl	Potassium chloride
KH2PO4	Monopotassium phosphate
KS	Rate of synonymous substitutions
Leu	Leucine
LOC	Lab-on-a-chip
LOD	Limit of Detection
Lys	Lysine
MCL	Maximum Composite Likelihood
MEDLE	Telomerically-encoded secreted *Cryptosporidium* protein family
MEGA	Molecular Evolutionary Genetics Analysis
Met	Methionine
ML	Maximum Likelihood

MLFT	Multilocus Fragment Typing
MLG	Multilocus Genotypes
MLST	Multilocus Sequence Typing
MSC6-7	Serine repeat antigen
Mu (μ)	Per base mutation rate per generation
N	No. non-synonymous sites
Na2HPO4	Disodium phosphate
NaCl	Sodium chloride
NAF	North Africa/Middle East
NAM	North America
NCBI	National Center for Biotechnology Information
NNI	Nearest-Neighbour-Interchange
ORF	Open Reading Frame
PAS	Periodic Acid Schiff
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
Phe	Phenylalanine
Pi(π)	Nucleotide diversity
Pi(π)i	InDel diversity
Pro	Proline
PubMed	National Institutes of Health Public MEDLINE Database
RCF	Relative Centrifugal Force
RDP	Recombination Detection Program
RPM	Revolutions Per Minute
Rrna	Ribosomal ribonucleic acid
S	No. sysnonymous sites
SAAG	Single Amino Acid Gap
SAAP	Single Amino Acid Polymorphism
SAM	South America
Ser	Serine
SNP	Single Nucleotide Polymorphism
SPSS	Statistical Package for the Social Sciences
SSA	Sub-Saharan Africa
SSU	Small Subunit
SVSP	Subtelomere-encoded Variable Secreted Proteins
T	Thymine

TGA	Thousands of Generations Ago
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
UNAIDS	Joint United Nations Programme on HIV/AIDS
Val	Valine
WGA	Whole Genome Alignment
WGS	Whole Genome Sequence
WHO	World Health Organization
ZN	Ziehl Neelsen

Establishing the genetic basis of host specificity has been one of the leading aims for Cryptosporidiologists since the first complete reference genomes became available just over a decade ago (Abrahamsen *et al.* 2004; Xu *et al.* 2004). Since their release, there has been strong concordance within the scientific community regarding a ~95.0-97.0% consensus between the genomes of the primary human-infective species, *Cryptosporidium parvum* and *Cryptosporidium hominis*, in terms of global genetic composition (Mazurie *et al.* 2013). Initial comparative analyses also pointed towards complete synteny of chromosome size and arrangement between these two closely-related, yet phenotypically-diverse, species. Despite the interesting phylogenetic inferences that can be drawn from currently-available data regarding genetic distance and evolutionary relatedness between and within *C. hominis* and *C. parvum*, they do not sufficiently explain the phenotypic discrepancies that exist between them, most notably the narrow anthroponotic host range in the former, and expansive zoonotic host range in the latter. The available studies also do not disclose any genotypic explanation for the existence of a presumptively anthroponotic *C. parvum* subtype IIc, with the single strain-by-strain analysis performed to date only identifying limited genome-wide variation between this strain and its zoonotic IIa *C. parvum* counterpart (Widmer *et al.* 2012). Although the initial release of the reference genomes will forever be hailed for its revolutionary impact within the *Cryptosporidium* community, recent access to new and improved genomes have proven that current suppositions are outdated and erroneous, in part due to the poor sequence coverage and quality of the original reference TU502 *C. hominis* genome, but also as a result of the superficial and nucleotide-biased view with which the first genome-wide analyses were performed. A transition to proteomics-based comparisons, in conjunction with accurate synteny evaluation through improved end-to-end chromosome alignments, and in-depth views into the nature and degree of variation at recombination hotspots like the subtelomeres, will help project *Cryptosporidium* genomics into a new era of understanding. Whole-genome comparisons in this context are specifically aimed at characterizing genetic affinities and disparities between and within distinct phenotype groups (i.e. anthroponotic versus zoonotic), unlocking the genetic basis of divergent features such as host range and virulence, and providing a new

stockpile of supervirulence candidates for potential translation of molecular targets into novel diagnostics and prevention. From a species classification standpoint, the objective is to achieve a validated method for classifying strains that builds upon and improves currently-employed phylogenetic inferences, paving the way for novel genotyping alternatives to be implemented throughout academic and public health spheres. Population genetics of the hypervariable host-parasite interactor GP60 will complement this by evaluating the allelic divergences and evolutionary pressures accompanying strain emergence and evolution worldwide, providing insight into both the geographic spread and chronology of GP60 subtype emergence and describing novel epidemiological trends across subtype families. This thesis thereby aims to contribute significantly towards determining the impactful yet subtle genetic mechanisms underlying phenotypic differences between closely related strains and species, identifying the key players in host-parasite interactions, and providing new insight into a parasitic disease which has largely been neglected and underestimated by the global research and public health community. The following opening section aims at providing a brief introduction to three important themes interlaced throughout the main chapters: classification, virulence, and incentive. Understanding both the history and present situation of *Cryptosporidium* isolate classification highlights the inaccuracy and inefficiency of common phylogenetic typing tools that this thesis aims to describe and correct, and explains why previous studies may have missed the extent of genomic recombination that is herein illustrated and discussed. Virulence and divergence are presented within the context of host-parasite interactions and pathogenic impact, in an attempt to highlight the importance of these key features in differentiating phenotypically-diverse *C. hominis* and *C. parvum* subtypes, and to pre-emptively justify a methodological approach that focuses on subtelomeres, hypervariability and selective pressure. Brief introductions to the current understanding of genomic differences and recombination in human-infective *Cryptosporidium spp.* further highlight what we know versus what knowledge might be missing, creating a platform upon which all four in silico chapters will build. The final introductory section presents the all-important public health incentive for conducting *Cryptosporidium* research, through systematic review of published human prevalence studies. This review presents and discusses global prevalence data in the general public, as well as at-risk groups such as immunocompromised individuals, and illustrates the relative proportion of human infections worldwide due to *C. hominis* versus *C. parvum*.

Systematic review of GP60 subtype data further illustrates our current understanding of intra- and inter-species differences in host range and global distribution of GP60 subtype families, which inspired the idea for a GP60 population genetics study and intra-*C. hominis* comparative genomics piece. Introductory themes throughout chapter 0 are left intentionally vague and generalist, as tailored introductions have been prepared for each chapter individually. This has been done with the intension of presenting the results of this thesis within a more contextual and comprehensive 'publication-style' sphere.

0.1 *Cryptosporidium* classification – Past and Present

(A) E. E. Tyzzer, 1875-1965 (https://web.stanford.edu/group/parasites/ParaSites2001/crypto/)
(B) Stomach section of a mouse infected with *C. muris* (Tyzzer, 1910)

In 1907, an American scientist by the name of Ernest Edward Tyzzer peered down his oil emersion lens at the gastric mucosa of a common mouse and discovered for the first time a new sporozoan type coccidian he named *Cryptosporidium muris*. For most of the twentieth century that followed, Cryptosporidiosis was thought of as a disease of primarily zoonotic and benign importance, and its potentially devastating impact on humans was only revealed when a new global population of immunocompromised individuals emerged as a result of the AIDS epidemic in the 1980s (Ma & Soave, 1983; Current, *et al.* 1983; Cohen, *et al.* 1984; Casemore, *et al.* 1985; Tzipori, 1988). The approach for species classification in this pre-molecular era was frequently based on the host type from which a strain was isolated, with individual isolates being designated as *C. parvum* "human" genotype, *C. parvum* "dog" genotype, *C. parvum* "bovine" genotype, and so on (Morgan *et al.* 2000; Xiao *et al.* 2001), in addition to the *C. muris* mouse strain originally characterized by Tyzzer almost a century earlier. The absence of a detailed epidemiological or biological understanding of this parasite in the decades that followed meant that until recently, the primary human-infective forms of *Cryptosporidium* were merely classed as genotypes belonging to a single species, *C. parvum* Type I (anthroponotic) and Type II (zoonotic) (Feng *et al.* 2000; Widmer *et al.* 2000).

Since then, extended morphological characterization and later genomic verification have indicated that *C. parvum* Type I and Type II are in fact two genetically distinct species (Morgan-Ryan *et al.* 2002), *C. hominis* and *C. parvum* respectively, and that the reported human-infective strains can be further grouped into a range of strains based on subtle genetic differences within each species. The *Cryptosporidium* genotyping and subtyping approaches have subsequently expanded, in an attempt to identify strain discriminatory markers exhibiting a degree and nature of variability that is considered sufficient and reliable enough to represent genome-wide differences between and within infective *Cryptosporidium spp*. 26 *Cryptosporidium* species and 70 genotypes have been characterized to date (Qi *et al.* 2015), and access to a wider range of whole-genome sequences is beginning to increase. This creates a platform for the advancement of molecular tools to better understand the way in which strains and species emerge and evolve in this diverse genus, while simultaneously opening the door for in depth studies of the structural and compositional genomic changes accounting for differential patterns of behaviour and infectivity.

In the early days of molecular diagnostics, identification was confined to a limited genotypic classification of *Cryptosporidium spp.* and an absence of associated genomic data. As a result, molecular diagnostics were largely dependent on mainstream scientific targets such as the SSU rRNA (Awad-el-Kariem *et al.* 1994; Johnson *et al.* 1995), heat shock protein (HSP; Stinear *et al.* 1996), and internal transcribed spacer 1 (1TS1) genes (Cai *et al.* 1992). Unknown targets produced by random amplification of polymorphic DNA (Morgan *et al.* 1996; Benigno *et al.* 1996) and plasmid-based genomic library construction also started to emerge, attempting to enhance the sensitivity of typing techniques as well as broaden the specificity to discriminate between phenotypically-diverse targets (Laxer *et al.* 1991). The PCR toolbox soon expanded to contain a number of oocyst protein (Wagner-Wiening & Kimmig 1995; Laberge *et al.* 1996) and oocyst wall protein (Ranucci *et al.* 1993; Meyer & Palmer 1996) targets, but these were still overwhelmingly catered towards *C. parvum* specificity and non-specific to less frequent causes of human disease such as *C. meleagridis*, *C. felis*, *C. canis*, and *C. muris*.

One of the first improved *Cryptosporidium*-specific discriminatory markers to be implemented as a typing tool was based on a 70-kDa heat shock protein (HSP70) gene

(Gobet & Toze, 2001). Although successful at differentiating between *C. hominis* and *C. parvum* based on a simple PCR banding size approach, a number of non-human genotypes could not be targeted, and the marker had too great sequence conservation to be of any use for detailed subtype analyses. The 18S small-subunit ribosomal ribonucleic acid (18SSU rRNA) locus, which is the most commonly used diagnostic marker across eukaryotic and prokaryotic pathogens (Wang *et al.* 2014), similarly proved to have insufficient discriminatory power between subtypes and failed to even identify the presence of separate species in mixed infections (Chalmers *et al.* 2005). The search for an improved substitute that contained a suitable balance between intra-strain conservation and inter-strain variability led to the identification of a highly-polymorphic 60 kDa glycoprotein commonly referred to as GP60 (Strong *et al.* 2000). The primary diagnostic benefit of this locus derives from the presence of several tandem repeats, one based on trinucleotides encoding a stretch of serines and another comprised of a repetitive KTVVR-type motif (Xiao, 2010), whose amplicon size initially appeared to correspond to genome-wide genetic changes between phenotypically-diverse isolates. This feature gave GP60 instantaneous popularity for its ability to produce interesting and expansive phylogenetic characterizations for a wide range of human- and animal-isolated parasites (Jex *et al.* 2007; Stensvold *et al.* 2014). These highly-repetitive regions of DNA, often referred to as microsatellites, are hypothesized to frequently contract and expand at a constant rate when one strain evolves into another (Ellegren, 2004), providing a cheap and reliable target for studying population genetics. Although GP60 has been used in more than 150 studies for the purpose of molecular characterization of *Cryptosporidium spp*. to date, the high mutation accumulation rate that characterizes the regions upstream and downstream of the repeats, as well as the recombinatory nature of this locus (Li *et al.* 2013), have started to bring into question whether this gene is reliable enough to provide phylogenies that accurately represent the genome-wide variability between strains. There is also the constant uncertainty regarding whether any true associations can be drawn between the nature of repeat divergence at a single genetic locus and the phenotypic manifestations of an infecting strain. In light of the potential drawbacks of GP60, a wave of alternative markers has since been introduced, and the current approach most western diagnostic laboratories rely on is a multilocus fragment size-based typing (MLFT) method that employs a combination of these markers to obtain the most reliable subtyping of infective species. However, these markers display varying discriminatory

powers between *C. hominis* and *C. parvum* (Robinson & Chalmers, 2012), and there is an absence of MLST validation at reference laboratories located outside of the European geographic area. A lack of access to diverse WGS also means that the phenotypic correlation between single marker targets and whole genome divergence remains poorly validated. Current data does not endorse or reject any combination of these markers wholly, and the absence of a consensus approach to *Cryptosporidium* genotyping, or a single marker for confident subtype discrimination, highlights the need to perform large-scale genomic comparisons to fill the current gap in *Cryptosporidium* molecular diagnostics and phylogenetic characterization.

0.2 The road to virulence in cryptosporidiosis – Divergence, selective pressure, and the telomeres

Aside from the development of improved phylogenetic typing tools, whole-genome comparisons in parasitic species are primarily aimed at identifying virulence factors which are essential for parasite survival and host pathogenesis, and through this understanding the genetic basis of host specificity. The genetic changes that accompany host-driven adaptation most significantly impact the protein repertoire of first line defences which a parasite utilizes upon encountering novel host immune cells and response mechanisms. As such, it is often possible locate virulence factors by identifying loci that are subject to a particularly high degree of divergence or selective pressure between species. Although the exact genetic changes resulting in specific virulence features are driven by individual host and parasite combinations, there are certain genetic changes that are ubiquitous in evolutionary adaptation across most parasitic species. Understanding the way in which these generalized genetic changes manifest themselves in genome data, and identifying the shortcomings of currently-available software tools for detecting them, provides a way to develop a customized approach to obtaining a thorough, accurate, and reliable understanding of genome-wide divergence between closely-related strains and species. One of these changes is the increased occurrence of indels, or single nucleotide insertion and deletion events, which result in frameshifts that have the potential to drastically alter or altogether eradicate coding sequences, and thereby impact the activity and function of encoded proteins (Williams & Wernegreen, 2013). The implications of indels in host-driven divergence is compounded by the increased mutation accumulation rate observed

in proximal genetic regions, seemingly a form of knock-on effect as a consequence of their occurrence (Jovelin & Cutter, 2013). What complicates the identification of genome-wide indel inflictions is that automated software tools are often customized to search for specific changes in SNP rates, rather than the protein-level impact of single nucleotide changes, and as a result these are often overlooked. Despite these shortcomings, an increase in indel mutations as a result of host species switching has been well-established in a number of pathogenic microorganisms, such as *Campylobacter jejuni* (Thomas *et al.* 2014), *Streptococcus agalactiae* (Rosinski-Chupin *et al.* 2013), parasitic nematodes (Wang *et al.* 2009), and *Mycobacterium spp.* (Streenu *et al.* 2006). Monitoring the occurrence of frameshift-causing indels at the genomic level can hence be used as evidence for a presumed case of host-driven adaptation, and is an important component towards understanding the genetic basis of host specificity and virulence in *Cryptosporidium*.

Another mechanistic consequence of speciation is the overrepresentation of sequence divergence, loss, and gain at subtelomeric regions (Monerawela *et al.* 2015). The enhanced gene turnover rates at these regions is facilitated by a high frequency of recombination (Jiang *et al.* 2011), often resulting in extensive expansion or contraction of subtelomeric gene copy number and an elevated rate of divergence (Anderson *et al.* 2015). Subtelomeres are thus often observed as the source of virulence and pathogenicity factors in microbial parasites, and frequently feature at the core of genome-wide divergence studies. In *Plasmodium*, the rapid rate of divergence at the subtelomeres has resulted in the evolution of an expansive and highly-effective virulence gene family collectively known as *var* genes, whose hypervariability and variegated expression enables the parasite to evade the host's antibody-driven immune response (Witmer *et al.* 2012). A number of other subtelomeric gene families also promote virulence in this protozoan parasite, including a repetitive interspersed DNA element family (Weber 1988), a variable open reading frame family (Niang *et al.* 2014), an interspersed helical domain gene family (Oberli *et al.* 2014), and a Maurer's cleft membrane gene family (Sam-Yellowe *et al.* 2004). In *Trypanosoma*, a vast network of subtelomerically-encoded variable surface glycoproteins allows this parasite to employ antigenic variation (Jehi *et al.* 2014). This fascinating feature of continuously shifting recognition molecules on the parasite's cell surface almost uniformly accounts for the success with which trypanosomes evade host immune responses during the bloodstream form, and the high rates of morbidity and

lethality it can inflict as a result. *Theileria* similarly contains a range of subtelomere-encoded variable secreted proteins (SVSP) whose rapid inter-strain divergence, variable gene copy number and significant representation in the macroschizont secretome makes them likely virulence factor candidates (Hayashida *et al.* 2012). Despite this wealth of knowledge in other protozoan parasites, very little is known about the significance of subtelomeric regions in *Cryptosporidium* host specificity and virulence; in fact, only few studies are currently published that even mention subtelomeres in conjunction with Cryptosporidial genetics (Bouzid *et al.* 2013a; Li, *et al.* 2013; Bouzid *et al.* 2013b; Guo *et al.* 2015; Liu *et al.* 2009; Liu & Abrahamsen, 1999). Novel whole genome comparisons using improved WGS assemblages have begun to expose physical evidence of genome-wide incongruence, with significant subtelomeric sequence insertion and deletion (indels) events identified between *C. hominis* and *C. parvum* (Guo *et al.* 2015). Identifying significant changes in subtelomeric consistency between phenotypically-diverse strains and species is thus a key component to understanding how and why certain human-infective *Cryptosporidium* subtypes and species exhibit differential patterns of behaviour, host specificity, and virulence.

Although the subtelomeres require special focus in discerning virulence factors in protozoan parasites, proteins involved in host-parasite interactions and pathogenesis are by no means limited to this region of the genome. A number of internally-encoded proteins have been implicated as equally contributing to the success and severity with which protozoa cause disease. This includes the metalloprotease GP63 of *Leishmania*, involved in cell signal modulation and evasion of complement-driven lysis (Shio *et al.* 2015); a family of rhoptry kinases in Toxoplasma that facilitate invasion and host-cell gene expression in supervirulent strains (Sibley *et al.* 2009); a family of cathepsin L-like peptidases in *Eimeria* (Liu *et al.* 2014); and GP60, a GPI-anchored secreted sporozoite antigen that displays hypervariability between *Cryptosporidium* strains and species (Strong *et al.* 2000). In addition, protozoan parasites as a group seem to contain an impressive repertoire of glycoconjugates encoded at numerous loci spanning the genome, whose sugar-saturated outer coating and frequently-secreted nature implicates them as key host-parasite interactors at various stages of parasite development (Ivens *et al.* 2005; Ramly *et al.* 2013; Tomita *et al.* 2013; Bhalchandra *et al.* 2013). Putative virulence factors have a very diverse infrastructure in terms of localization, function, and activity across the

tree of Apicomplexan life, and exhibit a highly taxon-specific nature at many of the encoding loci (Kissinger & DeBarry, 2011). This makes it difficult to develop a generalized pipeline for deducing virulence-associated protein function based on sequence similarity between putative virulence factors in one Apicomplexan genus and established virulence factors in another. Whole-genome comparisons between phenotypically-diverse strains are consequently one approach to overcome this, generating a list of tentative candidates based on all areas of the genome displaying higher-than-average divergence rates, as well as those that are under a positive versus negative degree of selective pressure. This provides a solid virulence platform upon which molecular studies can build.

0.3 Justification and rationale – the global significance of exploring human-infective *Cryptosporidium spp.* genomics

The importance of performing this genome-wide comparative analysis and population genetics study has to be explained in the context of the known epidemiological features that classify *C. hominis* and *C. parvum* as distinct species types, as well as the big picture of global disease burden imposed by these on human populations. The primary pathogenic distinction between *C. hominis* and *C. parvum* is an almost entirely anthroponotic host range in the former, and an extensive zoonotic host range in the latter, characterizations which have repeatedly been reported and were independently validated in this systematic review (Figure 0.3.1). Despite this widely-accepted characterization, the genomes of these two species have perpetually been described as largely homogeneous (Mazurie *et al*. 2013), with current data providing only limited genotypic insight into their unique patterns of host specificity through SNP-based divergences, missing subtelomeric sequence data and occasional recombination at the GP60 locus. What adds to the mystery of the genetic basis of host specificity is the existence of a presumptively anthroponotic *C. parvum* strain, described as GP60 subtype family IIc (Widmer *et al*. 2012), which complicates the generalizability of the association between host range and species as it currently stands. The IIc isolates are closely related to other *C. parvum* subtype families genetically, but they appear to predominantly propagate in an overwhelmingly anthroponotic transmission cycle, similar to the *C. hominis* host range phenotype. This alludes to the potential existence of more complex differences at the genomic level which previous analytical approaches may have missed,

and which the lack of access to diverse WGS have historically hampered. Re-visiting the phenotypic differences of human-infective *Cryptosporidium spp.*, and illustrating global prevalence rates and disease burden in high risk populations, provides a core justification for studying genomic differences in human disease.

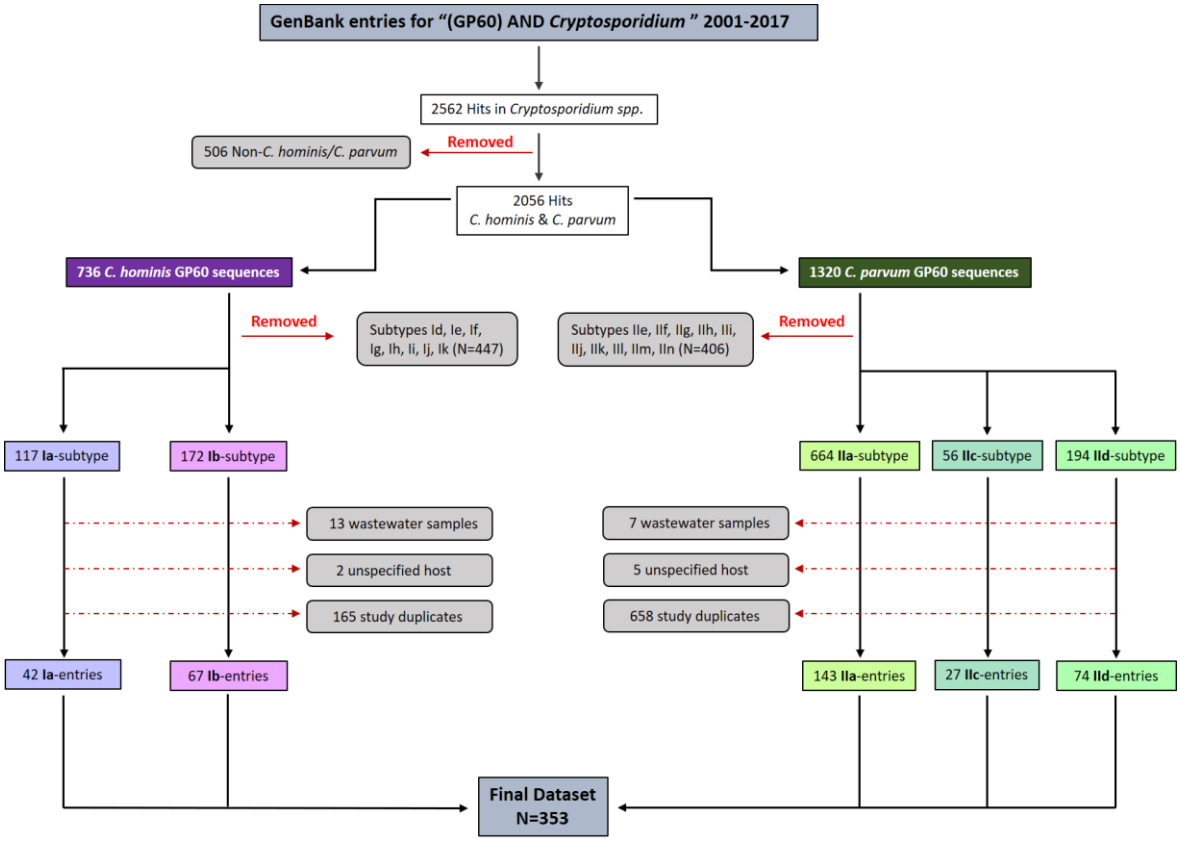


Figure 0.3.1. Flow diagram of inclusion/exclusion procedure during systematic review of GenBank GP60 entries, to define host range for the GP60 subtype families used in this study: *C. hominis* subtypes Ia and Ib, and *C. parvum* subtypes IIa, IIc, and IId

The process of data extraction, exclusion, and inclusion during systematic review of GP60 sequences is outlined, describing the approach by which host range was characterised for GP60 subtype families Ia, Ib, IIa, IIc, and IId. The initial GenBank (https://www.ncbi.nlm.nih.gov/genbank/) search was based on the terms "GP60", "GP40", and "GP15" (GP60 has previously been referred to as GP15, GP40, and GP40/15), and hits were subsequently filtered to remove all non-*C. parvum*/*C. hominis* related entries. These hits were further cleaned to remove all GP60 sequences not belonging to families Ia, Ib, IIa, IIc or IId, as were all entries of non-human (e.g. environmental) or unspecified origin. The last phase involved removing all study duplicates, which resulted in a final dataset of 353 GenBank entries for inclusion in this study. Data extracted from each GenBank entry included nucleotide sequences (to confirm subtype family), date of submission, geographic origin of isolates, any associated publications and host species.

The systematic review of GP60 sequences submitted to GenBank revealed 1,203 submissions for the subtypes used in this comparative genomics study (Ia, Ib, IIa, IIc, IId). After the dataset was cleaned to remove environmental samples, entries with unspecified host and study duplicates, a total of 353 remained and were included in the host range analysis. Submission quantities were not evenly distributed across the subtypes, with a significant proportion belonging to *C. parvum* subtype IIa (N=143), followed by *C. parvum* subtype IId (N=74), *C. hominis* subtype Ib (N=67), *C. hominis* subtype Ia (N=42), and *C. parvum* subtype IIc (N=27) (Figure 0.3.1). GenBank submissions included GP60 sequences from 48 countries, but the geographic origin of samples was equally unbalanced across the world. More than one-fourth (26.9%) of entries referring to samples from the United Kingdom and Australia, and when the dataset was further stratified into geographic regions it became evident that a vast majority of samples originated in the European region (40.5%), followed by Australasia (14.7%), Eastern Asia (10.7%), North Africa (10.7%), South America (9.2%), Central Asia (5.5%), Sub-Saharan Africa (4.9%) and North America (3.7%) (Figure 0.3.2).

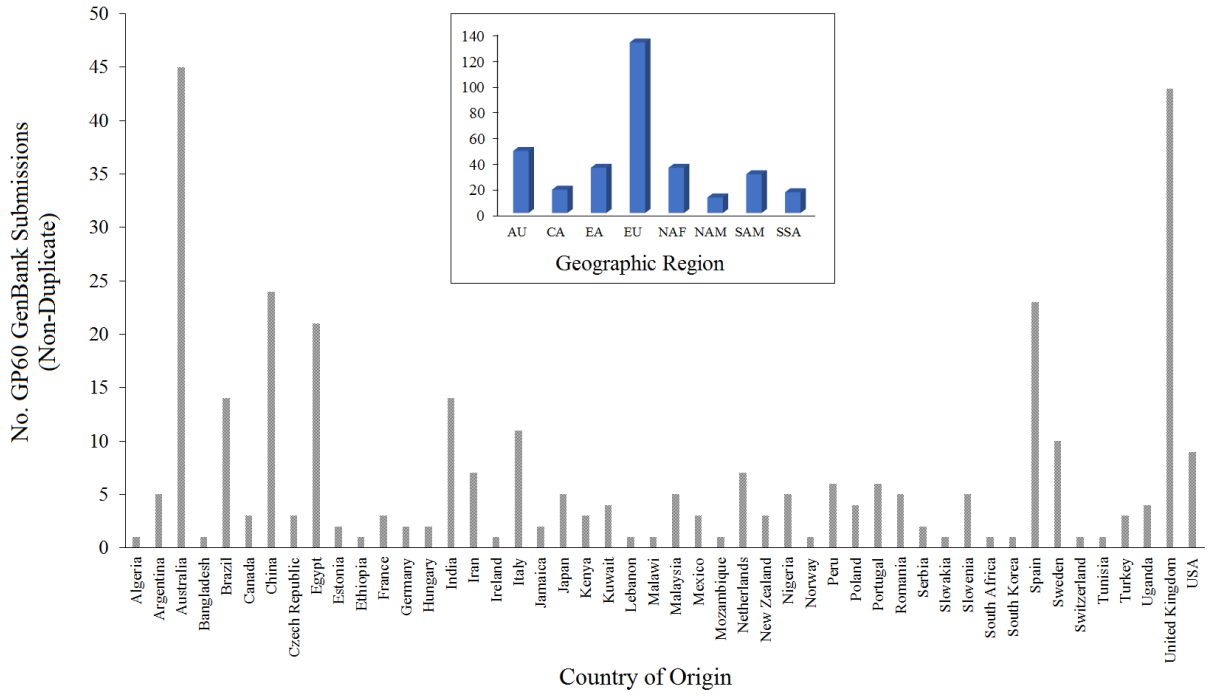


Figure 0.3.2. Geographic origin (country and region) of GP60 GenBank submissions included in systematic review of host range for *C. hominis* (GP60 subtypes Ia and Ib) and *C. parvum* (subtypes IIa, IIc, and IId)

The country and region of origin of each GP60 sequence retrieved for this systematic review is here summarised. Regions are defined as Australasia/Oceania (AU), Central Asia (CA), Eastern Asia (EA), Europe (EU), North Africa/The Middle East (NAF), North America (NAM), South America (SAM), and Sub-Saharan Africa (SSA). 326 of the 353 (92.4%) GenBank submissions included in this host range characterisation reported a country of origin, or referred to an associated publication which provided isolate origin information, while the remaining 27 (7.6%) did not and so were excluded from this figure.

The systematic review of the *Cryptosporidium spp.* subtypes for which WGS were available in this study (Ia, Ib, IIa, IIc, and IId) provided an overview of host specificity that accurately reflects and confirms host range associations as they currently stand (Figure 0.3.3). Collated data from a range of epidemiological and clinical studies clearly separate certain subtypes into zoonotic versus anthroponotic categories. *C. hominis* subtypes Ia and Ib, and *C. parvum* subtype IIc, were isolated from humans in more than three-fourths (>75.0%) of published studies reviewed, corroborating previous reports of predominantly anthroponotic transmission cycles. *C. parvum* subtypes IIa and IId, on the other hand, were reported in approximately equal proportions from human, cattle, and 'other' host types (~30% each). This demonstrates the true breadth of zoonotic host range in these subtypes, together identified in more than twenty separate host species, which appear well-adapted to both human and cattle hosts. These trends describe *C. hominis* and *C. parvum* as specialist and generalist respectively in terms of their ability to thrive in various host environments. Their close overall genomic relationship implies that a subset of targeted divergences and rearrangements from a recent common ancestor could have driven this evolutionary adaptation and speciation event, supporting the rationale for a multi-faceted and extensive study to describe the underlying genetic features. Interestingly, *C. hominis* subtype Ib exhibited a slight zoonotic tendency, having been isolated from 8 non-human host types on at least fourteen separate occasions. This observation places this subtype's host range phenotype closer to that of *C. parvum* subtypes IIa and IId than indeed *C. parvum* subtype IIc is situated. This provides some evidentiary support for the claim that certain subtypes may be more virulent and infectious than others (Cama *et al.* 2008), based on a direct association between broader host range and greater pathogenicity as observed in other pathogens (Sree & Varma, 2015; Stephenson & Foley, 2016). This adds another layer of complexity to what this comparative genomics and population genetics study aims to achieve, as similar to the IIc-phenomenon it shows a tendency of certain subtypes to contradict intra-species phenotypic expectations. Correlating genetic differences with these types of phenotypic characteristics therefore requires expansive and detailed genomic and population genetics analyses, and provides a strong scientific rationale for expanding the scope of this study to elucidate genetic differences not simply on a species-by-species basis, but between individual subspecies and subtype families for which WGS are available.

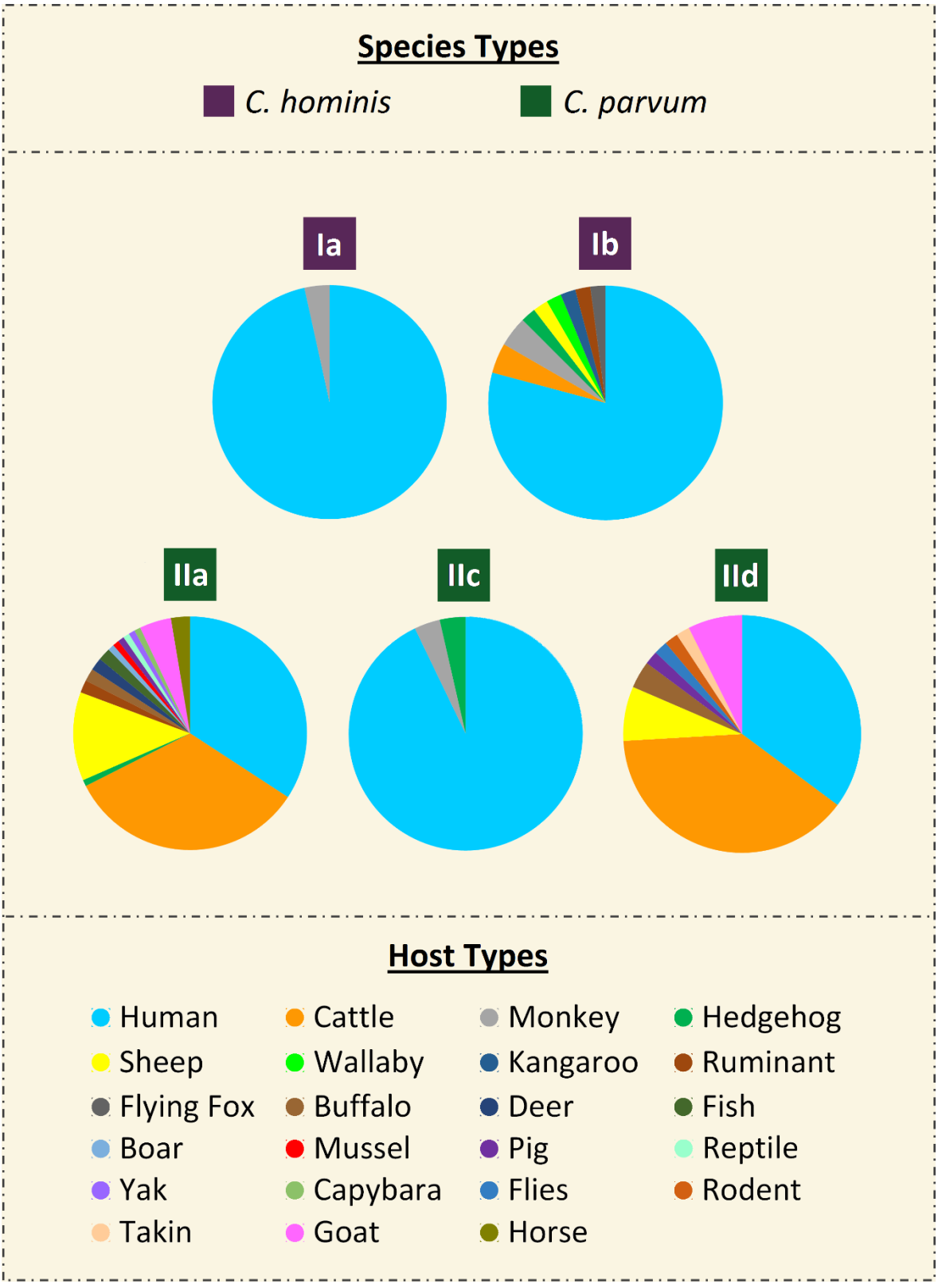


Figure 0.3.3. Host range patterns of *Cryptosporidium* GP60 subtype families Ia, Ib, IIa, IIc, and IId

Host range patterns were determined through a systematic review of peer-reviewed studies describing molecularly-discriminated *Cryptosporidium* subtypes Ia, Ib, IIa, IIc, and IId isolated from animal species on NCBI's PubMed database (http://www.ncbi.nlm.nih.gov/pubmed). The presence of a subtype through host faecal examination and molecular screening was given a value of 1.0 per study, and comparative figures are hence not representative of overall or relative prevalence of infection or subtype in any given host population. Frequency of subtype distribution was estimated from 48, 29, 114, 54, and 28 peer-reviewed articles for subtype families Ib, Ia, IIa, IId, and IIc, respectively.

A further important justification and rationale for this comparative genomics study lies in the global burden of disease that can be attributed to *C. hominis* and *C. parvum*, in humans generally and in at-risk groups such as immunocompromised individuals. For this purpose, a systematic review of all human prevalence studies published in the new millennium (2000-2017) was undertaken. The objective was to capture the relative burden of human cryptosporidiosis between different geographic regions and study groups (e.g. children, immunocompromised, etc.), and to illustrate differential distribution patterns of *C. parvum* and *C. hominis* as well as their respective GP60 subtypes. A total of 754 peer-reviewed journal articles were extracted and included in this systematic review, representing 128 countries worldwide (Figure 0.3.4). Most studies originated in Sub-Saharan Africa (19.9%), with similar proportions stemming from Central Asia (15.9%), Europe (15.7%), North Africa/The Middle East (16.9%) and South America (14.5%), and slightly fewer originating in Eastern Asia (11.3%) (Figure 0.3.5). A notable fewer proportion of prevalence studies took place in Australasia (3.3%) and North America (2.4%). Country-based differences were also observed, with a significant number of studies originating from India (N=63) and Iran (N=40), and half of the studies reviewed overall originated from only 14 countries (Australia, Brazil, China, Egypt, Ethiopia, India, Iran, Malaysia, Nigeria, Saudi Arabia, Spain, Thailand, Turkey, United Kingdom). It is also important to note that no human prevalence data could be identified for more than one-third of countries worldwide (66 of 194; 34.0%), with islands in Oceania and the Caribbean accounting for >30.0% of the missing data (Figure 0.3.4). This highlights a significant knowledge gap for human cryptosporidiosis that is more pronounced in remote, resource-poor areas. As a consequence, the current description of human *Cryptosporidium spp.* distribution and burden worldwide remains incomplete and skewed. A substantial margin of uncertainty remains to be corrected and filled in order to improve our understanding of the burden cryptosporidiosis poses on human populations worldwide, and to better understand the epidemiogical features that define and separate human-infective species and subtypes.

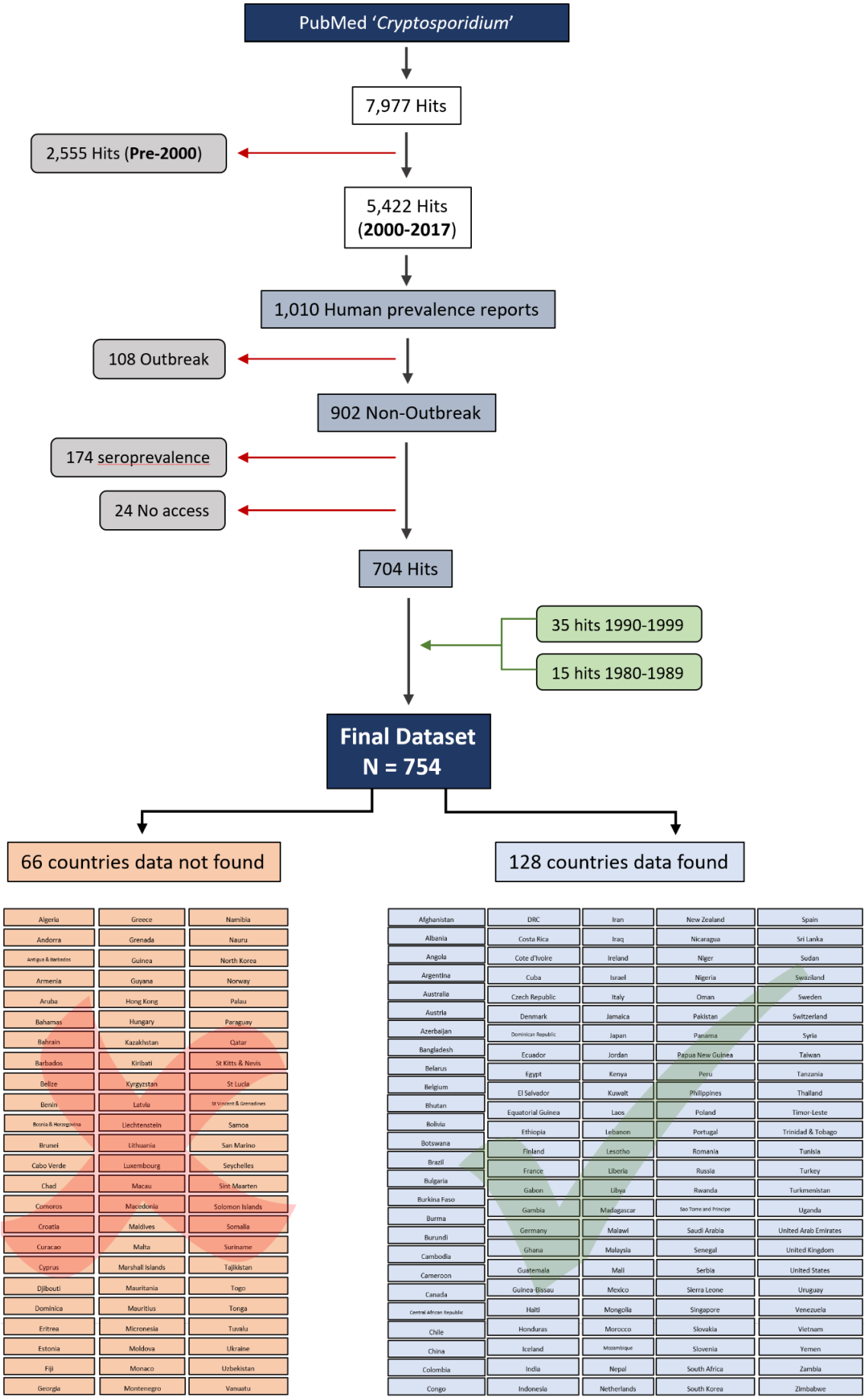


Figure 0.3.4. Flow diagram of study selection procedure and criteria for systematic review of human cryptosporidiosis prevalence studies, 2000-2017.

Figure 0.3.4. Flow diagram of study selection procedure and criteria for systematic review of human cryptosporidiosis prevalence studies, 2000-2017.

The human cryptospodiosis database was constructed using data from all peer-reviewed journal publications describing non-outbreak associated prevalence studies in humans, published between 2000-2017. A total of 7,977 PubMed (https://www.ncbi.nlm.nih.gov/pubmed) hits for the search term "*Cryptosporidium*" were obtained, and initially filtered to remove all pre-2000 publications (N = 2,555). Abstracts for the 5,422 hits that remained were individually reviewed to extract all human cryptosporidiosis prevalence studies (N = 1,010), and subsequently cleaned to remove all outbreak and seroprevalence studies, as well as studies where a full version of the publication could not be accessed. This resulted in a dataset of 704 peer-reviewed studies spanning 115 countries worldwide. In order to enhance global coverage, pre-2000 published human prevalence studies were included for countries where post-2000 data was not available, which added another 50 hits from 13 countries (Austria, Azerbaijan, Belarus, Bolivia, Burma, Burundi, Congo, Costa Rica, El Salvador, Honduras, Liberia, Trinidad and Tobego, and Turkmenistan). The final dataset consisted of 754 peer-reviewed journal publications, and spanned a total of 128 (66.3%) of the 194 countries worldwide. The database was subsequently constructed using manually-extracted data from each of these studies, and included: date of publication, date(s) of study, study group(s) (including all descriptive characteristics pertaining to age, pre-existing chronic conditions, disease manifestations, and high-risk exposures such as animal contact), location (including where possible country, city, and rural versus urban environment), overall prevalence, prevalence for individual study groups (e.g. HIV-positive versus HIV-negative study participants), proportion of prevalence attributable to known species/genotypes, and the proportion of *C. hominis* and *C. parvum* cases caused by individual subtype families.

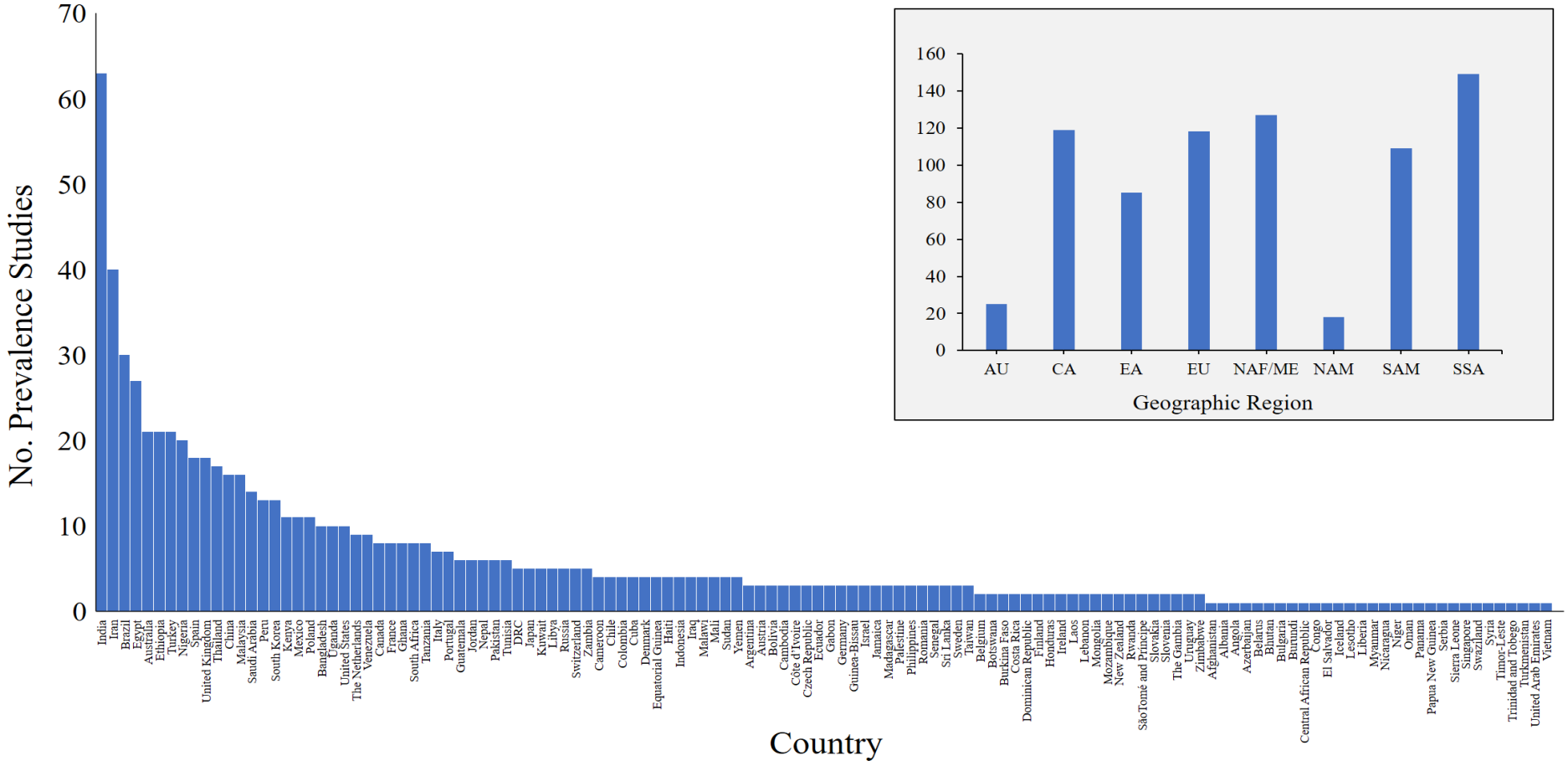


Figure 0.3.5. Geographic origin (country and region) of human cryptosporidiosis prevalence studies, 2000-2017

The geographic origin of peer-reviewed studies included in the human cryptosporidiosis database is here presented, describing the relative proportion of studies originating from specific countries and regions worldwide. Regions are defined as Australasia/Oceania (AU), Central Asia (CA), Eastern Asia (EA), Europe (EU), North Africa/The Middle East (NAF/ME), North America (NAM), South America (SAM), and Sub-Saharan Africa (SSA).

A stratification of the study groups in which cryptosporidiosis rates are frequently investigated generates further information about populations at a higher risk of infection or severe pathogenesis (Figure 0.3.6). Children made up the largest proportion of study groups targeted in prevalence studies between 2000-2017 (35.9%), of which ~65% involved symptomatic patients. This heightened focus reflects the high burden of cryptosporidiosis and the associated morbidity and mortality in children that has previously been reported (Sow *et al*. 2016). Another high-risk group that featured predominantly was HIV-positive individuals, making up 26.0% (196 of 754) of the human prevalence studies reviewed. This targeted focus similarly reflects the significant burden that cryptosporidiosis is known to inflict on persons infected with HIV and other immunocompromising conditions. Symptomatic populations (including adults only or cases of all age groups) was the final major study group targeted by the prevalence studies reviewed, making up 24.7% (186 of 754). The remainder were made up of immunodeficient individuals suffering from non-HIV related conditions (9.0%), asymptomatic cases (7.0%), and humans reporting close contact to animals or farming facilities (2.1%).

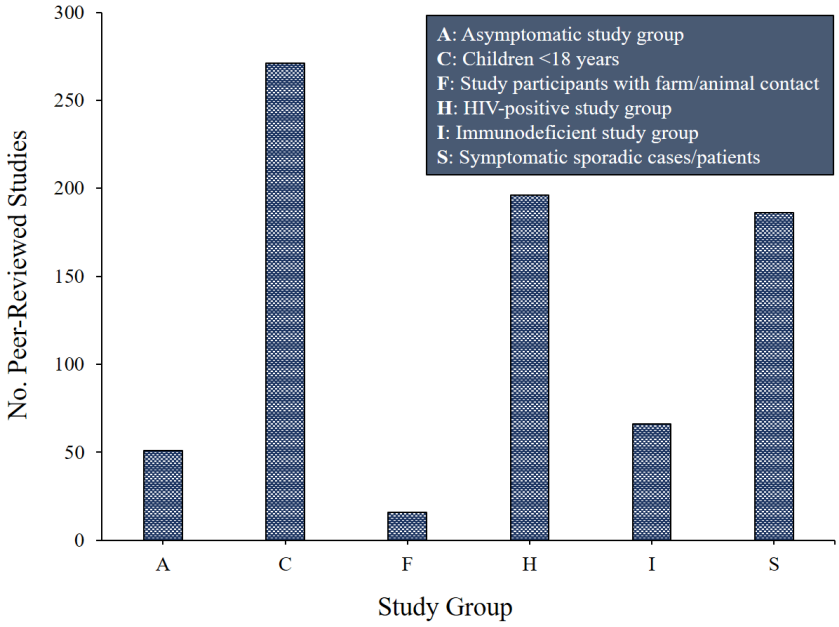


Figure 0.3.6. Types of study groups included in prevalence database of human cryptosporidiosis, 2000-2017

The types of study populations included in human cryptosporidiosis prevalence studies (2000-2017) are here illustrated. These characterisations are based on all pre-existing conditions/exposures/criteria present in the study population and used by the reporting authors to define their study groups, and include: asymptomatic individuals (A), children <18 years (C), individuals with exposure to agricultural settings/animals (F), HIV-positive (H), other immune-compromising conditions (e.g. organ transplant recipients) (I), and symptomatic sporadic cases/patients (S). Studies with multiple pre-existing criteria/conditions were included in multiple groups where applicable (e.g. studies focused on HIV-positive children were included in both "C" and "H").

Global prevalence rates in affected human communities and populations highlight the persistence of *C. hominis* and *C. parvum* across a wide range of geographically- and socioeconomically-diverse localities (Figure 0.3.7). In the new millennium (2000 – present), cryptosporidiosis was diagnosed in symptomatic (broadly referring here to groups exhibiting some form of gastrointestinal manifestations) individuals in more than 93% (699 of 754) of prevalence investigations reviewed. Prevalence rates for cryptosporidiosis ranged from <0.1 to 100.0% of study participants, varying significantly across geographic regions and between immunocompromised versus competent patients. In immunocompetent patients with gastrointestinal symptoms, countries of North America, Europe, North and East Asia and Australasia showed relatively low frequencies of cryptosporidiosis (<1.0-5.0%), while large parts of South America, Sub-Saharan Africa, the Middle East and Central/South Asia had higher average prevalence rates ranging from 5.0 to >30.0%. These more heavily affected regions are simultaneously areas with disproportionately higher rates of malnutrition, particularly in children <5 years (FAO 2013), which re-emphasizes the significant impact of cryptosporidiosis on populations suffering from compromised nutritional states (Sallon *et al*. 1988; Costa *et al*. 2011).

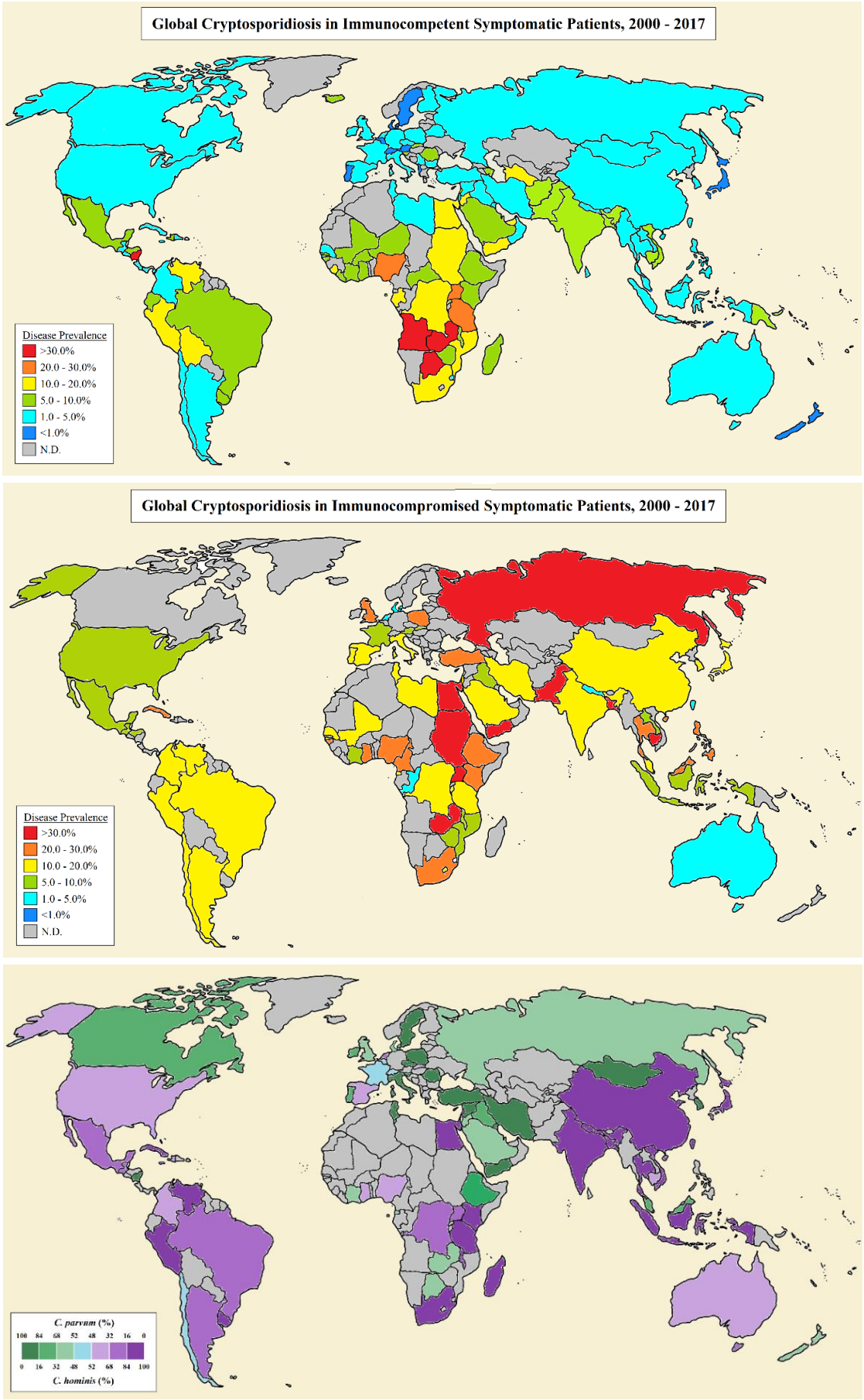


Figure 0.3.7. Global prevalence maps of human cryptosporidiosis in immunocompetent versus immunocompromised groups, and the relative compositions of primary human-infective *Cryptosporidium spp.* in human populations, 2000-2017

Figure 0.3.7. Global prevalence maps of human cryptosporidiosis in immunocompetent versus immunocompromised groups, and the relative compositions of primary human-infective *Cryptosporidium spp.* in human populations, 2000-2017

The prevalence and geographic distribution of cryptosporidiosis cases in immunocompetent and immunocompromised human study groups (non-outbreak related investigations), and relative proportion of infections attributable to *C. hominis* versus *C. parvum*, were assessed through systematic review of all relevant prevalence studies published between 2000-2017 on NCBI's PubMed Database (http://www.ncbi.nlm.nih.gov/pubmed). A total of 754 peer-reviewed publications providing data from 128 countries (66.3% global coverage) were selected, having fulfilled the pre-determined inclusion criteria as specified for this extensive systematic review (detailed methodology outlined in Figure 0.3.4.)

The global picture of burden changes instantaneously when immunodeficiency is taken into focus, with the heat-mapped figures demonstrating a more than five-fold increase in prevalence in all but five of the countries investigated (Australia, Denmark, Nepal, Taiwan, and The Netherlands). The association between immune-compromising conditions, most notable HIV positivity, and substantial rates of cryptosporidiosis has repeatedly been reported (O'connor *et al*. 2011), and this systematic review reveals exactly how pervasive this association is across the world. More than 60% of studies across all geographic regions exhibited greater than 10.0% prevalence rates in immunocompromised symptomatic patients, and in almost half of the countries in Sub-Saharan Africa (47.8%; 11 of 23 countries) at least one-fifth of immunocompromised patients had cryptosporidiosis. This is largely linked to the high rates of HIV infection found in this region, with Sub-Saharan Africa accounting for more than 70.0% of individuals infected with HIV worldwide (GBD HIV Collaborators, 2015). The increased availability of highly active antiretroviral therapy (HAART) to treat HIV, particularly in the last six years (UNAIDS 2013), means a number of Cryptosporidiosis studies focused on HIV-positive groups have begun to report decreases in prevalence rates in HAART naïve versus HAART experienced patients (Figure 0.3.8).

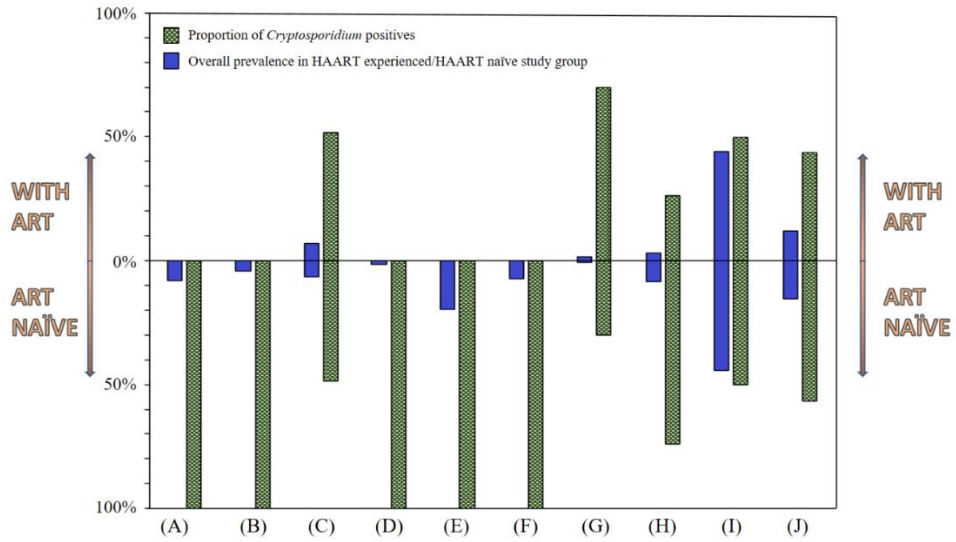


Figure 0.3.8. Prevalence rates for *Cryptosporidium* infection in ART-experienced versus ART-naïve patients

Human cryptosporidiosis studies comparing prevalence rates between ART-experienced and ART-naïve patients are here described. The figure illustrates data for both the overall prevalence of cryptosporidiosis in the study population, as well as the relative proportion of cases found in ART-experienced versus ART-naïve individuals. References: Bachur *et al*. 2008; (B) Mengist *et al.* 2015; (C) Ukwah *et al.* 2017; (D) Missaye *et al.* 2013; (E) Adamu *et al.* 2013; (F) Teklemariam *et al.* 2013; (G) Taye *et al.* 2014; (H) Hiros *et al.* 2015; (I) Nsagha *et al.* 2016; (J) Shimelis *et al.* 2016

Despite this initial success towards lowering disease burden, HAART coverage in many of the most resource poor groups remains low, and studies as recent as 2016 continue to report rates well above 30.0% in patients with HIV (Abdel-Hafeez *et al*. 2012; Masarat *et al*. 2012; Noor *et al*. 2012; Wanyiri *et al*. 2014; Nsagha *et al*. 2016). In addition, a few of the HAART-experienced versus HAART-naïve prevalence studies were unable to demonstrate a direct benefit of antiretroviral therapy on cryptosporidiosis, with negligible differences in prevalence rates between HAART naïve and experienced groups (Nsagha *et al*. 2016; Shimelis *et al*. 2016; Ukwah *et al*. 2017), and some studies even identifying higher rates in the HAART experienced than naïve individuals (Taye *et al*. 2014). However, the higher incidence rates in HAART experienced individuals could reflect selective administration of these drugs to those patients suffering most severely from AIDS-related conditions, and that immune functioning is already so compromised in these patients that HAART is ineffective at protecting against infections such as cryptosporidiosis. These consistently high prevalence rates and potential inefficiency of HAART at resolving cryptosporidiosis in the most at-risk group of HIV individuals are worrying, as are reports of the only available treatment option Nitazoxanide repeatedly failing to reach optimal levels of symptomatic and infection relief in both immunocompromised *and* immunocompetent individuals (Doumbo *et al*. 1997; Rossignol *et al*. 2001; Elaine &

Sanford, 2008). This means the search for an effective approach towards treatment and prevention continues. As new studies continue to appear that highlight the true morbidity and mortality attributed to cryptosporidiosis worldwide (Sow *et al*. 2016), the need to prevent and cure this neglected disease continues to rise up the international public health agenda. The identification of new diagnostic and treatment targets through genomic comparisons of phenotypically-diverse *Cryptosporidium spp*. is therefore an important step significant towards improving diarrhoeal health and thereby productivity worldwide.

The global distributions and relative proportions of *C. hominis* and *C. parvum* GP60 subtypes identified in human populations provided a further summary of the intra-species genetic heterogeneity known to date. GP60 subtype families were sub-categorised based on the defined region of isolate origin, to analyse potential associations between certain subtypes and/or phenotypes and global distribution. This was specifically aimed at demonstrating geographic disparities in the diversity and relative proportions of subtype families, and correlating these with described differences in host range (i.e. zoonotic versus anthroponotic sources of infection). Intra-*C. parvum* distribution patterns revealed one such association, with the anthroponotic *C. parvum* subtype IIc exhibiting regional dominance over zoonotic subtypes such as IIa and IId in certain locations, and vice versa. In the more economically-prosperous regions of Australasia, Europe, North America, and The Middle East, the anthroponotic:zoonotic subtype ratios strongly favoured zoonotic *C. parvum* subtypes, at rates of 0.033, 0.034, 0.071, and 0.068, respectively. In the more resource-poor areas of Central Asia, South America, and Sub-Saharan Africa, on the other hand, this trend was entirely reversed, with anthroponotic:zoonotic subtype ratios strongly favouring anthroponotic *C. parvum* subtype IIc, at rates of 5.5, 6.125, and 5.429. What this seems to imply is that zoonotic sources of infection, such as farm animals and household pets, may account for a significant majority of *C. parvum*-related cryptosporidiosis in more developed countries, while *C. parvum* isolates in less developed areas predominantly propagate within an anthroponotic transmission cycle. The potential of zoonotic sources being more responsible for human cases in developed regions is further supported by the frequent reports of *Cryptosporidium* outbreaks linked to petting zoos and agricultural practices in developed countries such as the United States (Reif *et al*. 1989; Preiser *et al*. 2003; Smith

et al. 2004; Kiang *et al*. 2006; Starkey *et al*. 2007; CDC 2011; Webb *et al*. 2014; Drinkard *et al*. 2015; Suler *et al*. 2016; Hancock-Allen *et al*. 2017), the United Kingdom (Evans & Gardner, 1996; McGuigan 2005; Gait *et al*. 2008; Hoek *et al*. 2008; McGuigan *et al*. 2010; Gormley *et al*. 2011; Utsi *et al*. 2016), Sweden (Kinross *et al*. 2015), Norway (Robertson *et al*. 2006; Rimšelienė *et al*. 2011; Lange *et al*. 2014), Ireland (Sayers *et al*. 1996; Zintl *et al*. 2006), and Italy (Cacciò *et al*. 2013). Results from this systematic review thereby confirms and enhances the frequent reports of variable phenotypes and epidemiological specifications separating zoonotic from anthroponotic *C. parvum* GP60 subtype families, and motivates the comprehensive whole genome comparison and subsequent recombination analysis described in chapters 1 and 2. Access to a dataset of novel *Cryptosporidium spp*. WGS provides a unique opportunity to further investigate the extent to which genetic variation drives species and subspecies development in this species, and will assess the degree to which current GP60-based subtyping approaches accurately correlate with genomic divergence.

The intra-regional diversity of GP60 subtypes was also variable, with Sub-Saharan Africa reporting the broadest range of GP60 subtypes (14 families total) followed by Central Asia (11 families), compared to 9 or fewer subtype families reported in the remaining geographic regions. This indicates a trend that associates particularly high regional rates of cryptosporidiosis with a greater GP60 subtype diversity, as both Sub-Saharan and Central Asian countries exhibited some of the highest prevalence rates in both immunocompetent and immunocompromised individuals (Figure 0.3.7). Inter-species differences between *C. hominis* subtypes were similarly striking across regions, and this systematic review provides the first accumulation of comprehensive prevalence data to illustrate a global overview of regional subtype composition. As *C. hominis* predominated as the cause of cryptosporidiosis in the highest prevalence regions (Figure 0.3.7), and a greater subtype family diversity was determined across a majority of geographic regions (Figure 0.3.9A and B), this species was selected for the focus of population genetics analyses at the GP60 locus in chapter 3. The tentative differences in host range patterns and distributions between subtypes Ia and Ib also explains their focus in the intra-*C. hominis* whole genome comparison in chapter 4. Understanding the dynamics of population genetics at this hypervariable locus, particularly within the context of evolutionary selection and genetic differentiation, aims to evaluate how these features

may link to host environments and behaviours, or help explain epidemiological differences in subtype distribution and pathogenicity. A detailed examination of the complex genetic features that separate phenotypically-diverse *Cryptosporidium* isolates at both the allelic and genomic level is fundamental to our understanding of how this parasite infects, evolves, and evades. This study therefore hopes to provide some of the vital in silico information about virulence that is required in order to translate and apply scientific findings into public health practice, and start to make significant strides towards lowering the disease burden and impact of this still neglected tropical disease.

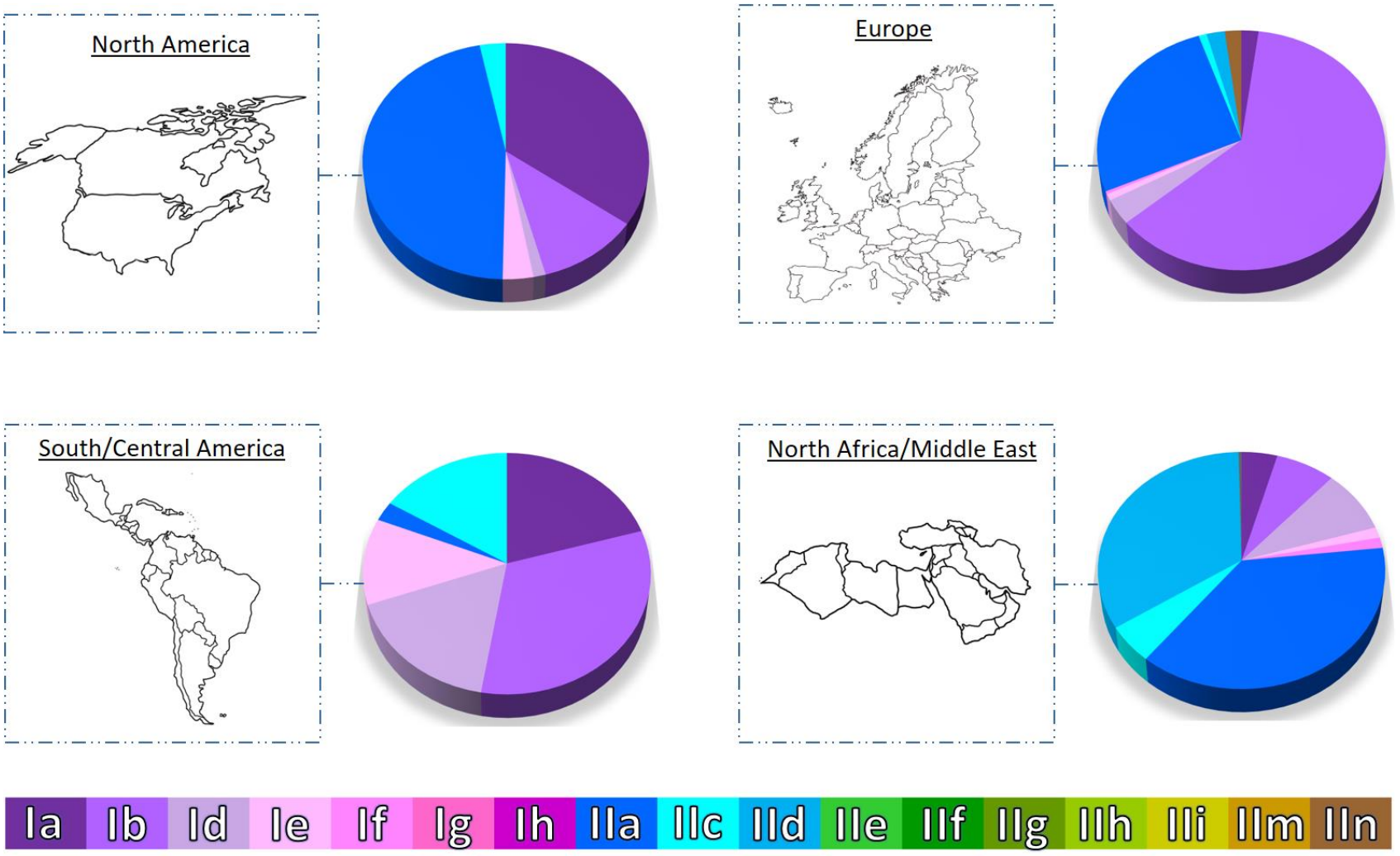


Figure 0.3.9.A. Global distribution and relative proportion of *C. hominis* and *C. parvum* subtypes in the Americas, Europe, and North Africa/The Middle East
Prevalence rates and geographic origins of GP60 subtype families within anthroponotic *C. hominis* (Ia, Ib, Id, Ie, If, Ig, Ih) and zoonotic *C. parvum* (IIa, IIc, IId, IIe, IIf, IIg, IIh, Ii, IIm, IIn) species groups were extracted as part of the systematic review of 757 peer-reviewed journal articles on human cryptosporidiosis, and are shown here for the geographic regions of North America (N=8), South America (N=9), Europe (N=21), and North Africa/The Middle East (N=13).

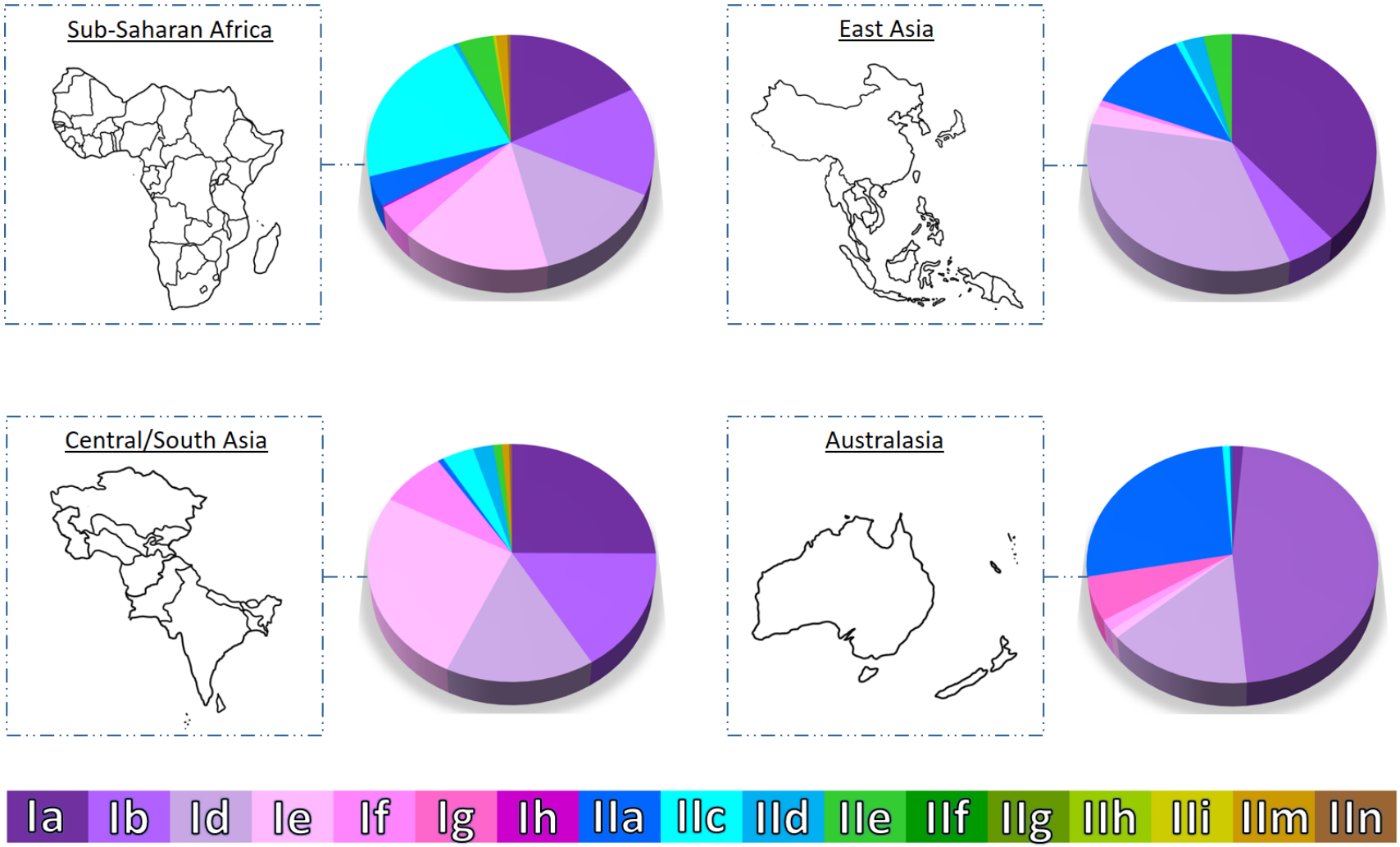


Figure 0.3.9.B. Global distribution and relative proportion of *C. hominis* and *C. parvum* subtypes in Sub-Saharan Africa, Central/South Asia, East Asia and Australasia
Prevalence rates and geographic origins of GP60 subtype families within anthroponotic *C. hominis* (Ia, Ib, Id, Ie, If, Ig, Ih) and zoonotic *C. parvum* (IIa, IIc, IId, IIe, IIf, IIg, IIh, IIi, IIm, IIn) species groups were extracted as part of the systematic review of 757 peer-reviewed journal articles on human cryptosporidiosis, and are shown here for the geographic regions of Sub-Saharan Africa (N=21), Central/South Asia (N=10), East Asia (N=7) and Australasia (N=12).

Characterizing evolutionary histories and protein sequence divergence in *Cryptosporidium spp.* WGS: the genetic basis of host range and virulence in human-infective strains

1.1 INTRODUCTION

(i) Limitations of the single marker *Cryptosporidium spp.* phylogenetics

Phylogenetic analyses and trees have long been the primary approach towards visualizing and evaluating genetic distances between species. Although traditionally based on morphological characteristics (Wheeler, 1917; Chamberlain, 1915; Van Cleave, 1949; Sinclair & Dunn, 1961), the development of genetic sequencing and computer programming have enabled detailed insight into the genetic distances, and chronological spacing, that separate organisms between and within all existing taxonomic classifications. The classical approaches to designing phylogenies can be categorized as either phenetic (distance matrix) or cladistic (maximum parsimony versus maximum likelihood) in nature, with the former drawing from overall similarities between compared organisms, and the latter basing predictions on shared characteristics to the most recent common ancestor. The introduction of nucleotide sequencing in the 1970s facilitated a massive improvement to phylogenetic interpretations (Salser *et al.* 1972, Fry, *et al.* 1973; Maizels, 1973), through statistical evaluation of the rate of evolutionary divergence, and thereby the development of mathematical models which evaluate and weight patterns of genetic changes. These models, which are based on pre-calculated assumptions and predictions of rate heterogeneity, DNA substitution, and amino acid replacement, have expanded the functionality of phylogenetics beyond a simple understanding of how certain traits have evolved over time (Liò & Goldman, 1998). It is now possible to identify unique genetic components of non-ancestral origin (Peng *et al.* 2012), designate a timescale for the emergence of closely related phenotypically-diverse strains (Schuh *et al.* 2015), and trace the path of infectious diseases back to the original

source of infection, a vital component for battling outbreaks and understanding the rate at which an infectious agent is capable of mutating and spreading geographically (Emmett *et al.* 2015; Weisberg *et al.* 2015). The construction of extensive, detailed, and most notably accurate phylogenies is thus a vital component towards understanding the evolutionary relationships between pathogenic species, along with inferences about the environmental changes that historically drove phenotypic diversification.

The phylogenetic relationships between *Cryptosporidium spp.* have been explored in more than 57 peer-reviewed publications (NCBI PubMed; https://www.ncbi.nlm.nih.gov/pubmed). A common feature of a majority of these studies is the application of only a handful of markers (18S rRNA, HSP70, Actin, COWP, and GP60), with the exception of a few studies which employed alternative markers such as ITS1, ITS2, and 5.8S rRNA (Morgan *et al.* 1999), CP47 and MSC6-7 (Ramo *et al.* 2015), HSP90 (Feng *et al.* 2009) and a concatenated sequence of 10 arbitrarily selected loci (Bouzid *et al.* 2010). What this means is that there is a significant bias in terms of the genetic targets used to explore phylogenetic relationships between *Cryptosporidium* strains and species, and questions surrounding how well these associations actually represent genome-wide divergence have not sufficiently been addressed. Studies using a common marker have frequently used different tree-building algorithms and nucleotide substitution rates, which can make it difficult to interpret the true phylogenetic distances between strains and species, and there has been a historical irregularity in the specific nomenclature used to define isolates included in phylogenetic analyses, with studies repeatedly referring to the host from which parasites were extracted rather than the species or subtype group to which each strain is likely to belong based on genetic similarities. The lack of a reliable typing method that consistently correlates marker variation with whole genome divergence makes it difficult to classify and understand the genomic diversity of *Cryptosporidium*. The development of new phylogeny-based approaches is therefore key to understanding how phenotypically-diverse strains and species have emerged and evolved over time.

(ii) A brief history of divergence between *C. hominis* and *C. parvum*

The ~9.14Mbp *Cryptosporidium parvum* genome is comprised of 8 chromosomes ranging in size from 0.88 to 1.34Mbp, and has a highly compact coding sequence composition (73.2-77.6%). Genomic comparisons between the original *C. parvum* IIaA15G2R1 (Abrahamsen *et al.* 2004) and *C. hominis* IaA25R3 (Xu *et al.* 2004) reference genomes currently provide an overview of chromosome-wide hotspots for single nucleotide polymorphism (SNP) elevation and increased levels of divergence and selective pressure between orthologous coding sequences, and have established a number of species-specific genes and duplication events (Mazurie *et al.* 2013; Bouzid *et al.* 2013). This has highlighted telomeric clustering of hyperpolymorphism (Bouzid *et al.* 2013), and identified a number of mucin-like and GPI-linked putative virulence mediators (Priest *et al.* 2001; O'Connor *et al.* 2009; Chatterjee *et al.* 2010; Jenkins *et al.* 2014). The assessment of protein-level consequences of genomic diversity has been limited, however, as has the largescale assessment of intra-species divergence for key host-parasite interactors. In addition, the only available study providing a genomic comparison between zoonotic and anthroponotic *C. parvum* subtype WGS (IIa and IIc, respectively) identified only a limited number of shared SNPs between the anthroponotic *C. parvum* and *C. hominis* subtype WGS (Widmer *et al.* 2012), and the genetic mechanisms or rates of intra-*C. parvum* suspeciation events were not explored. Novel whole genome comparisons using improved WGS assemblies have begun to explore physical evidence of genome-wide incongruence, with significant sequence insertion and deletion (indels) events identified between *C. hominis* and *C. parvum* (Guo *et al.* 2015), and a well-established case for interallelic recombination at the site of the hypervariable 60-kDa GP60 subtyping locus (Li *et al.* 2013). Expanding cross-comparisons to include multiple WGS across a range of anthroponotic and zoonotic *C. parvum* and *C. hominis* strains will help to explore these phenotype-associated features, and is an essential component to understanding the genetic basis of divergence between human-infective strains.

The first complete reference genomes for *C. hominis* and *C. parvum* were both published in 2004 (Xu *et al.* 2004; Abrahamsen *et al.* 2004). The former, named TU502, was developed from oocysts isolated from an infected Ugandan child, while the latter, Iowa II, was so named because the utilised oocysts were obtained during a largescale outbreak in

the US state earlier that year. Although these genomes are still considered one of the most significant advancements within the field of *Cryptosporidium* genetics, providing the very first insights into the genomic basis of host specificity, the sequence coverage and quality between the two genomes differs significantly. This has limited the extent to which genomic comparisons between these two WGS could be used to explore genome-wide drivers and consequences of speciation, particularly at highly-repetitive and hypervariable regions like the subtelomeres. The first publicly-available releases of these genomes contained 69 non-specific nucleotides (N's) and 18 contigs for *C. parvum* Iowa II, and 2,449 N's and 1,422 contigs for *C. hominis* TU502 (GenBank Accession: *C. parvum* AAEE01000000; *C. hominis* AAEL000000). Equally problematic is the complete lack of synteny between the assemblies (NCBI Map Viewer; Wheeler *et al.* 2003), which is most severe at subtelomeric regions, and has complicated the process of whole genome comparisons through automated analyses (Figure 1.1.1).

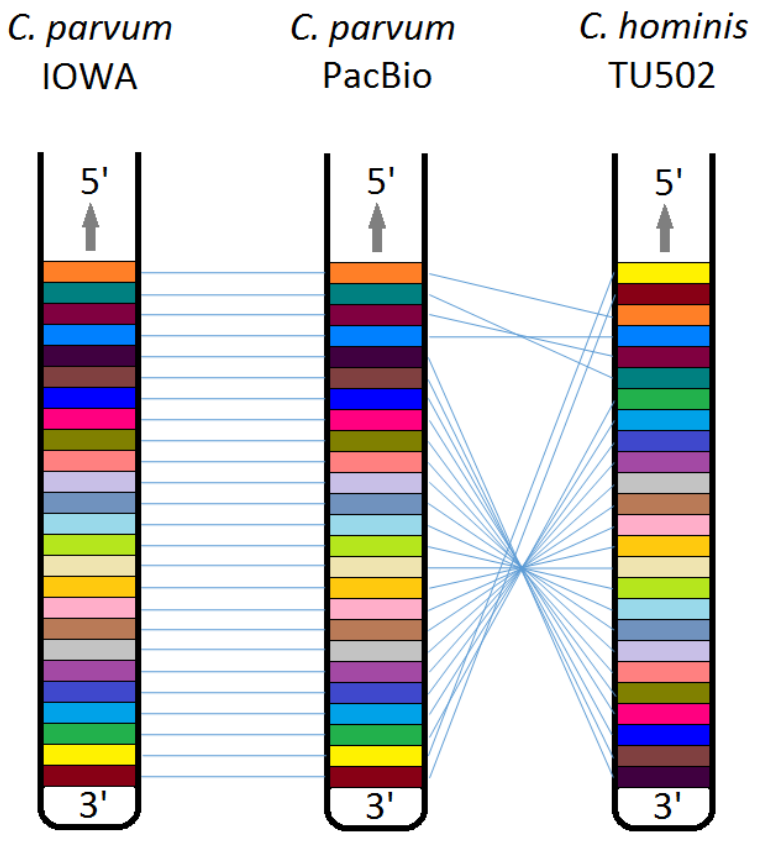


Figure 1.1.1. Degree of inconsistent synteny between *C. parvum* Iowa II and *C. hominis* TU502 reference WGS
The lack of coding sequence synteny between reference *C. parvum* Iowa II and *C. hominis* TU502 whole genome sequences is demonstrated based on the annotated order of the first 25 genes extending from the 3'-end telomeric repeat on chromosome 8. A novel PacBio-produced *C. parvum* WGS is included for synteny verification purposes (source: Dr. James Cotton, The Wellcome Trust Sanger Institute). Gene order was extracted from NCBI's Map Viewer Tool (http://www.ncbi.nlm.nih.gov/mapview/; Wheeler *et al.* 2003).

First insights from comparative genomics studies using these reference WGS focused largely on global summary figures (Widmer & Sullivan 2012) such as nucleotide sequence identities (~95-97%), GC content (~30%), genome size (9.1-9.2Mb) and protein coding gene content (~4,000). Shortly thereafter an attempt was made to correlate host range differences with genomic divergence, in a study that cross-compared an anthroponotic *C. parvum* WGS (TU114) against the two reference genomes (Widmer *et al.* 2012). The results of this study highlighted genomic areas of elevated SNP frequency and re-emphasized the notion of polymorphic clustering around the telomeres, simultaneously describing patterns of mutations between chromosomes and reporting basic functional properties (molecular weight, localization signals, and putative protein function) of highly-diverged proteins. However, very few correlations between host range phenotype and genetic divergence were identified.

In the second study that followed, a significant improvement to comparative genomics was achieved by utilising a re-assembled version of *C. hominis* TU502 (Mazurie *et al.* 2013). The primary benefit of this study was inclusion of protein level comparisons to delineate divergence, facilitating an assessment of selective pressure through dN/dS characterization of gene orthologs. This study confirmed the previously-reported global average gene identities of ~97%, but additionally identified highly diverged amino acid sequences (<90.0% AA IDs) in conjunction with those under positive selective pressure (dN/dS > 1.1). However, this study erroneously reported the absence of any significant insertion, deletion, or rearrangements, and did not consider the impact of frameshift-causing indels. Despite this, the authors provided the first tentative link between site-specific divergence and host range, and the realization that speciation likely occurs rapidly, and as a result of direct interactions with novel environmental conditions and most importantly host immune systems. A review of the available data has highlighted a number of gaps in *Cryptosporidium* genomics analyses and interpretations, but simultaneously provides a good theoretical basis of reference *C. hominis*/*C. parvum* subtype WGS divergence upon which novel and more expansive comparisons can build. This chapter will present phylogenetic and comparative genomics results for a novel collection of more diverse subtype WGS, in a way that expands on our understanding of the way intra- and inter-species evolution is structured. It also aims to provide greater depth of knowledge on the common features driving anthroponosis in human-infective

species, and thereby enhance our understanding of the mechanisms underlying phenotypic divergence within and between *Cryptosporidium hominis* and *Cryptosporidium parvum*.

1.2 METHODS

(i) Whole genome sequences for phylogenetic analyses – *empirical data*

16 novel *Cryptosporidium spp.* whole genome sequences were generously provided by the *Cryptosporidium* Reference Unit (Rachel Chalmers; Singleton Hospital, Wales, United Kingdom) for use in this comparative genomics study (Table 1.2.1). These included one *C. cuniculus* (GP60 subtype VbA37), one *C. hominis* (IaA14R3), two *C. meleagridis* (IIIgA23G3), ten *C. parvum* (IIaA15G2R1, IIaA17G1R1, IIaA18G2R1, IIaA19G1R2, IIdA22G1, IIcA5G3a, IIcA5G3j, and IIcA5G3k), and two *C. ubiquitum* (XIIb) whole genome sequences. A detailed methodology describing the generation of these novel WGS can be found in the associated publication (Hadfield *et al*. 2015). As a supplement these previously-available genomes were additionally included: RN66 *C. muris* WGS (Lorenzi, H., unpublished; BioProject Accession: PRJNA29965), TAMU-09Q1 *C. baileyi* WGS (Da Silva, JC., unpublished; BioProject Accession: PRJNA222835), UKMEL1 *C. meleagridis* (IIIbA22G1R1) WGS (Widmer, G., unpublished; BioProject Accession: PRJNA222838), and UKH1 *C. hominis* (IbA10G2) WGS (Widmer, G., unpublished; BioProject Accession: PRJNA222837). FASTA files containing whole genome sequence data for these isolates were retrieved from The *Cryptosporidium* Genomics Resource CryptoDB (http://cryptodb.org/cryptodb/). Assembly statistics (including scaffold number, mean scaffold length, and N50) are additionally outlined for each of the twenty included WGS in Appendix I.

Table 1.2.1. Summary of 20 *Cryptosporidium spp.* and GP60 subtypes of whole genome sequences included for phylogenetic analyses and comparative genomics

SPECIES	GP60 SUBTYPE	STANDARD ID	SOURCE	ACCESSION	WGS SIZE (bp)	T (%)	A (%)	C (%)	G (%)	N (%)
C. baileyi	N/A	TAMU-09Q1	Da Silva, J. C. [1]	CryptoDB.org	8,502,994	38.0	37.7	12.2	12.1	0
C. cuniculus	VbA37	UKCU2	This study		7,853,082	34.0	34.1	15.9	16.0	0
C. hominis	IaA14R3	UKH4*	Hadfield, S.[2]	PRJNA253838	9,502,496	34.4	34.3	15.6	15.5	0.2
C. hominis	IbA10G2	UKH1‡	Widmer, G.[3]	CryptoDB.org	9,141,398	34.8	35.0	15.0	15.2	0
C. meleagridis	IIIbA22G1R1	UKMEL1	Widmer, G.[3]	CryptoDB.org	8,973,224	34.5	34.6	15.4	15.5	0
C. meleagridis	IIIgA23G3	UKMEL3	This study		12,181,424	30.6	30.6	18.4	18.3	2.1
C. meleagridis	Novel GP60	UKMEL4	This study		9,663,721	33.9	33.9	15.9	15.7	0.6
C. muris	N/A	RN66	Lorenzi, H.[4]	AAZY02000000	9,245,251	35.7	35.8	14.2	14.2	0.07
C. parvum	IIaA15G2R1	UKP6*‡	Hadfield, S.[2]	PRJNA253846	9,112,937	34.4	34.4	15.5	15.4	0
C. parvum	IIaA17G1R1	UKP7	Hadfield, S.[2]	PRJNA253847	9,035,974	34.8	34.8	15.2	15.0	0.2
C. parvum	IIaA18G2R1	UKP3	Hadfield, S [2]	PRJNA253840	9,172,817	34.8	34.7	15.2	15.1	0.2
C. parvum	IIaA19G1R2	UKP2	Hadfield, S.[2]	PRJNA253836	9,759,982	34.2	34.1	15.9	15.8	0.07
C. parvum	IIdA22G1	UKP8*	Hadfield, S.[2]	PRJNA253848	9,377,202	34.4	34.5	15.4	15.5	0.2
C. parvum	IIcA5G3a	UKP13*‡	This study		8,416,777	33.5	33.4	16.5	16.5	0
C. parvum	IIcA5G3a	UKP14	This study		8,594,935	32.9	32.9	17.2	17.1	0
C. parvum	IIcA5G3a	UKP15	This study		10,134,261	34.4	34.5	15.4	15.6	0
C. parvum	IIcA5G3j	UKP16‡	This study		9,496,690	34.4	34.3	15.8	15.6	0
C. parvum	IIcA5G3k	UKP12	This study		8,038,639	34.2	34.2	15.8	15.8	0
C. ubiquitum	XIIb	UKUB1	This study		9,060,260	34.5	34.5	15.4	15.4	0.2
C. ubiquitum	XIIb	UKUB2	This study		9,069,162	34.5	34.5	15.4	15.5	0.1

* Included in whole genome comparative genomics
‡ Included in whole genome recombination analysis (chapter 2)
[1] University of Maryland Institute for Genome Sciences, Baltimore, Maryland, United States (Unpublished genome, CryptoDB.org)
[2] Hadfield *et al*. 2015
[3] Tufts University School of Veterinary Medicine, Medford, Massachusetts (Unpublished genome, CryptoDB.org)
[4] Craig Venter Institute, Rockville, Maryland, United States (Unpublished genome, CryptoDB.org)
Regarding methods used for oocyst sourcing, purification and preparation; DNA extraction and characterization; and next generation sequencing and analyses, a detailed description can be found in the relevant publication (Hadfield, *et al.* 2015).

The rationale behind strain inclusion for the novel whole genome sequences was determined by the lead investigators running the Hadfield *et al.* study. The overarching aim was to build a comprehensive genomic resource of the primary human-infective species (*C. hominis*, *C. parvum*, and *C. meleagridis*), specifically subtypes for which phenotypic data has previously been described, as well as infrequent human pathogens for which little to no data is currently available (*C. ubiquitum*, *C. cuniculus* and *C. viatorum*). *C. parvum* in particular has exhibited well-established phenotypic and epidemiological differences between many of this species' GP60 subtype families (Figure 0.3.3). There is a significant degree of GP60 divergence between the *C. parvum* subtype families used in this study (Figure 1.2.1), reaching nucleotide divergences as high as 18.5-

26.4% between subtypes IIa, IId, and IIc. Intra-*C. hominis* GP60 divergence between the two subtypes (Ia and Ib) used in this study is equally significant, exhibiting 25.0% nucleotide divergence. Descibing intra-species GP60 divergence provides an important justification for this study, as it highlights the importance of investigating the degree to which GP60 variability correlates with whole genome divergence, and evaluating whether hypervariability between WGS is limited to this locus. The specific *C. hominis* and *C. parvum* GP60 subtype sequence data used in this study also display broad geographic distributions and host range patterns, providing an important "phenotypic incentive" for comparing the included WGS, with the exception of IIaA19G1R2 which has only recently been discovered in the United Kingdom (Figure 1.2.2).

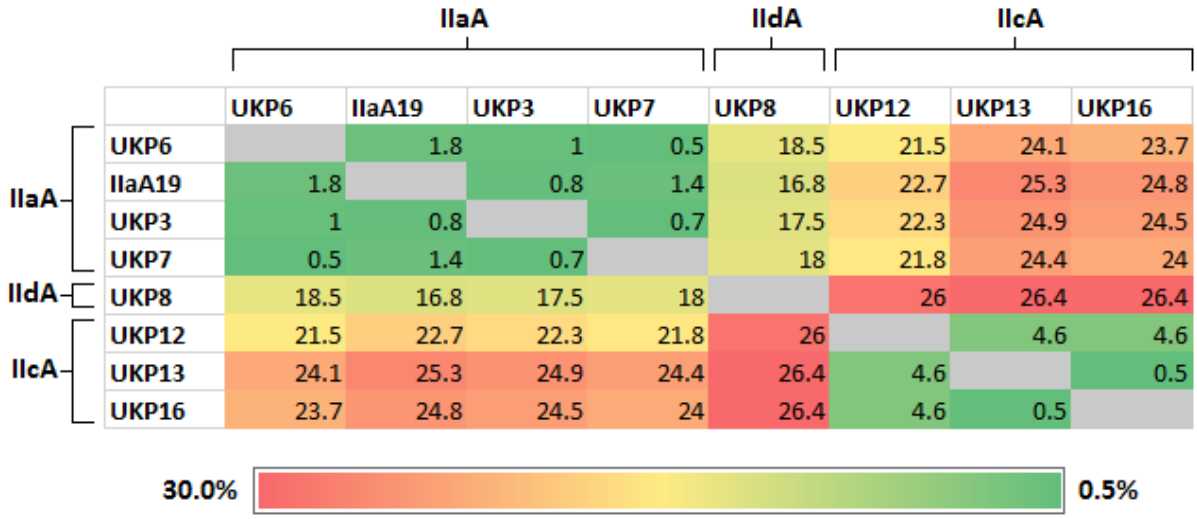


Figure 1.2.1. GP60 nucleotide divergence between whole genome *C. parvum* subtypes used in this study

GP60 sequence divergence was assessed between *C. parvum* strains belonging to varying GP60 subtype families IIa, IId, and IIc. Nucleotide sequences comprising the standard GP60 typing sequence (5'-TCCGCTGTATTCTCAGCCCCA-3' to 5'-AGCAGAGGAACCAGCATCCTT-3) were compared between strains using the EMBOSS Needle pairwise sequence alignment tool (The European Molecular Biology Open Software Suite; Rice *et al.* 2000). Values representing nucleotide divergence between strains in terms of % non-identities are shown.

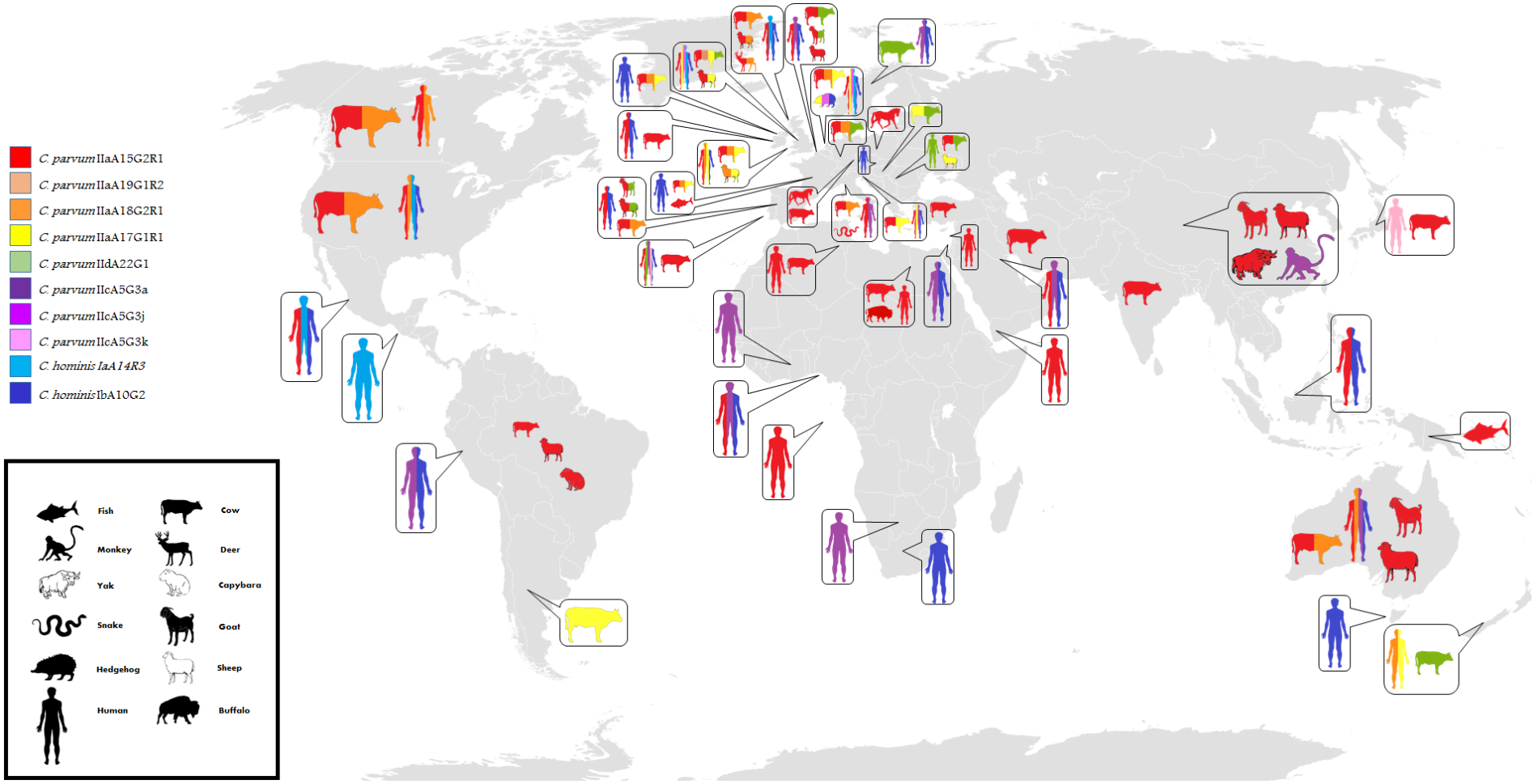


Figure 1.2.2. Global distributions and host types of *C. hominis* and *C. parvum* WGS used in the present study show variable epidemiological characteristics between GP60 subtypes

A systematic review of recorded cases of cryptosporidiosis due to specific GP60 subtypes was conducted, by searching NCBI's PubMed and GenBank databases for "IIaA15G2R1", "IIaA19G1R2", "IIaA18G2R1", "IIdA22G1", "IIcA5G3a", "IIcA5G3j", "IIcA5G3k", "IaA14R3", and "IbA10G2" (http://www.ncbi.nlm.nih.gov/pubmed). Relative proportions of prevalence are not represented. The maps provide a basic overview of the geographic distribution of GP60 subtypes explored in this study, along with the host species from which *Cryptosporidium* was isolated, based on a reported case of cryptosporidiosis within a sampled population.

(ii) Concatenated *Cryptosporidium* phylogeny

The methodology for the phylogeny relied on concatenation and alignment of 61 neutrally-evolving, genome-wide protein coding sequences, followed by construction of a maximum likelihood tree using the Tamura-Nei distance model, Nearest-Neighbour-Interchange (NNI) maximum likelihood (ML) heuristic method, and 2,000 bootstrap replicates, all courtesy of MEGA version 5.2 (Molecular Evolutionary Genetics Analysis version 5.2; Tamura *et al.* 2011). But the theory behind this approach requires a bit more explanation. The use of concatenated gene sequences to infer species phylogenies has become a preferred alternative to the single-gene approach, producing phylogenies that consistently result in a greater degree of accuracy and resolving power (Gadagkar *et al.* 2005; Devulder *et al.* 2005; Zhao *et al.* 2005). The thought behind this is that the use of a greater number of loci increases the probability of achieving a genome-wide genetic representation, thereby maximising the phylogenetic signal. But what compromises this approach is that different combinations of targets often result in completely different phylogenies (Kubatko & Dregnan, 2007; Roch & Steel, 2015; Warnow, 2015), largely due to the varying degrees of selective pressure (i.e. negative versus neutral versus positive) and divergence (% nucleotide identities) acting throughout the genome. Loci under negative (or purifying) selective pressure often classify isolates as more closely related than they actually are, while loci under positive selective pressure are often subject to such a high mutation accumulation rate that they place strains in completely the wrong proximity and orientation to their true "nearest neighbour". There is also often disagreement between phylogenies resulting from nucleotide- versus protein-based sequence concatenations. A thorough review of these previous concatenation approaches led to a new, experimental methodological technique for selection and inclusion of concatenated loci in this study; one that aimed to correct for previous shortcomings.

The inclusion process for protein coding sequences was based on fulfilment of three criteria: a neutral selective pressure between *C. parvum* (IIaA15) and *C. hominis* (IaA14R3) orthologous coding sequences, a magnitude of nucleotide consensus between 93.0% and 96.0% identities, and an even spread of targets across the genome. Selective pressure in this context was determined by dN/dS, a ratio which evaluates the average rate of

nucleotide substitutions at every non-synonymous versus synonymous site (Yang & Bielawski, 2000). The formal way to interpret dN/dS values is to view anything less than 1 as subject to negative selective pressure, anything greater than 1 as positive selective pressure, and anything equalling 1 as being under neutral selective pressure. However, the use of dN/dS to describe selective pressure does its limitations. The first complication is that certain nucleotide exchanges occur more frequently in some lineages than in others. This can introduce a strong bias that sways the nature of substitution in one direction or another, conveying an incorrect degree of selective pressure for a locus between closely-related strains. Another issue is that dN/dS values can vary significantly across coding sequence data, a phenomenon which often affects virulence proteins that contain active sites and domains whose function is involved in host-parasite interactions (Figure 1.2.3). dN/dS values also cannot distinguish between substitutions that result in amino acids that are biochemically similar versus different to one another, resulting in many proteins appearing to be under selective pressure when in reality their biological purpose and function may not have significantly changed. In light of these limitations, it seems more reasonable to assume that neutrally-evolving loci would have a dN/dS value somewhere in the range of 0.5, with values nearing 0.0 and 1.0 indicating negative and positive selective pressure respectively. The specific dN/dS values required for inclusion in this study were hence set at 0.3 to 0.6. In order to additionally ensure that coding sequences did not contain hidden pockets of positive selective pressure, often the case when specific protein domains become subject to a high degree of divergence, a minimum nucleotide percent identities of 93.0% was enforced. The upper cap of 96.0% was implemented to ensure that sequence variability was sufficient to provide for comprehensive phylogenetic discrimination between strains.

Figure 1.2.3. dN/dS limitations exemplified for a hypervariable *C. hominis*/*C. parvum* coding sequence

The limitations of using dN/dS to describe positive versus negative selective pressure for a single coding sequence is shown. Telomeric *C. parvum* gene cgd8_20 and its *C. hominis* ortholog Chro.80011 were selected due to a low overall dN/dS value (0.440) and simultaneously high degree of protein sequence divergence (<87.0% amino acid identities). Nucleotide and protein coding sequences were aligned using Clustal Omega multiple sequence alignment tool (http://www.ebi.ac.uk/Tools/msa/clustalo/), and dN/dS was consequently calculated using the web-application for Pal2Nal (Suyama *et al.* 2006). Sections displaying positive versus negative selective pressure are highlighted in red versus grey respectively. Predicted protein domains were identified by running nucleotide sequences through NCBI's conserved domain finder tool (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), with a majority of predicted domains overlapping with the central segment exhibiting positive selective pressure.

An even distribution of genomic loci was also implemented, so that the concatenated sequence would be an reliable representation of the entire *Cryptosporidium* genome. This meant that genes were specifically chosen to represent equal genomic distribution between and across (both telomeric and internal) the eight annotated chromosomes (Figure 1.2.4). The process for achieving this was significantly simplified by a pre-existing database containing global dN/dS values and percent nucleotide identities for all annotated *C. hominis* (subtype IaA14R3) versus *C. parvum* (subtype IIaA15G2R1) genes (Mazurie *et al.* 2013; supplementary table 3). Genes exhibiting the correct range of dN/dS

(0.3-0.6) and percent identities (93.0-96.0) were extracted from the data set, and then sorted according to the unique *C. hominis* numeric identifier (Chro.XXXXX). As the five-digit number associated with each identifier indicates the relative position of each gene, it was possible to semi-randomly select genes with a decently widespread genomic distribution across the genome. This approach made it possible to select eight loci from each of the eight chromosomes, three of which would eventually come to be excluded due to an unexpectedly high degree of divergence between closely related *Cryptosporidium parvum* strains (Chro.50435; Chro.60602; Chro.80266).

Figure 1.2.4. Genomic chromosome location of neutrally-evolving loci used for concatenated *Cryptosporidium spp*. phylogeny
The genome-wide localization of neutrally-evolving coding sequences (CDs) chosen for concatenation is shown. Horizontal markings indicate the location of CDs relative to the 5'- end and 3'- end telomeric repeats (5'-AAACCT-3') of chromosomes 1 to 8, but are not in any way representative of the size or orientation of individual coding sequences. Chromosome size was ascertained through manual reassembly of contigs from novel *C. parvum* IIaA15G2R1 (UKP6) genome sequence data, and by simple counting of nucleotide number from 5' telomere repeat to 3' telomere repeat. Specific location was similary determined through manual searches, which generated the exact nucleotide position along whole chromosome contiguous telomere-to-telomere sequence data.

(iii) Superaligning Supersequences and Building Trees

The primary and most time-consuming component of the concatenation approach was to seek and extract the 61 selected loci from the 21 *Cryptosporidium* whole genome sequences. Sequence data for the *C. parvum* IIaA15G2R1 reference strain was initially obtained by searching through the CryptoDB online *Cryptosporidium* genome database (Puiu, *et al.* 2004; http://www.cryptoDB.org). In order to confirm sequence integrity, the extracted genes were compared against the novel *C. parvum* IIaA15G2R1 genome (UKP6). Orthologous coding sequences in all included *Cryptosporidium spp*. were identified in whole genome datasets by setting up local nucleotide databases (FASTA format) in the BioEdit Sequence Alignment Editor, and running gene sequences against these using the local BLAST search tool (BioEdit v7.2.5; Hall, 1999). Individual gene sequences were subsequently extracted manually from each WGS. Distantly related and highly diverged sequences in *C. baileyi* and *C. muris* were additionally confirmed by backwards BLASTing hits through the BLASTn nucleotide database (NCBI Basic Local Alignment Search Tool; Altschul *et al.* 1990).

Nucleotide coding sequences were accumulated in the alignment editor of MEGA v 5.2 (Molecular Evolutionary Genetics Analysis version 5.2; Tamura *et al.* 2011) and automatically translated into amino acid sequences. Stop codons were removed to achieve the contiguous, concatenated amino acid sequences, which were aligned using the ClustalW alignment algorithm for codons (default parameters). The alignment was extracted as a MEGA format file, which was used to construct the maximum likelihood phylogenetic tree (Dayhoff substitution model, Nearest-Neighbour-Interchange (NNI) heuristic method). Phylogenetic confidence was inferred with the bootstrap method, using 2,000 bootstrapping replicates. A bootstrap cut-off of >95% was interpreted as indicating accurate and reliable branch topology.

Summary statistics for genome-by-genome comparisons were obtained by calculating sequence consensus and dN/dS values between the individual concatenated sequences. This was particularly important for the *C. hominis* IaA14R3 and *C. parvum* IIaA15G2R1 reference strains, as it would validate whether the designated approach had achieved the desired amino acid divergence and dN/dS rates at a concatenate-wide level (93.0-96.0%

and 0.3-0.6 respectively). Sequence identities and the relative proportions of SAAPs (single amino acid polymorphisms) and SAAGs (single amino acid gaps) were calculated using the EMBOSS Stretcher pairwise sequence alignment tool (The European Molecular Biology Open Software Suite; Rice *et al.* 2000). dN/dS values were manually ascertained by comparing nonsynonymous to synonymous substitution rates between orthologous coding sequences (PAL2NAL; Suyama *et al.* 2006).

(iv) Whole Genome Protein Comparisons - Included species and subtypes

Strains were selected for WGS protein divergence characterization based on GP60 similarity to the original zoonotic *C. parvum* and anthroponotic *C. hominis* reference genomes (IIaA15G2R1 and IaA25R3 respectively), namely UKP6 (IIaA15G2R1) and UKH4 (IaA14R3). To explore the association between protein-level genome changes and host range specification, two further closely-related but phenotypically-diverse whole genome *C. parvum* strains were included; zoonotic strain UKP8 (IIdA22G1) and the most common anthroponotic strain UKP13 (IIcA5G3a). These were specifically chosen based on equidistant GP60 nucleotide variability from *C. parvum* IIaA15G2R1 (Figure 1.2.1), but non-correlated phylogenetic distances based on the concatenated phylogeny. Parallel comparisons of UKP13 and UKH4 against UKP6 was specifically aimed at determining whether anthroponotic host range is associated with site-specific shared protein sequence divergence. The tertiary comparison between UKP8 and UKP6, both zoonotic, served as a positive control.

(v) Whole Genome Protein Comparisons - Analytical approach

Whole genome datasets were annotated into protein coding sequences (EMBOSS GetOrf), uploaded as standalone protein databases (BioEdit v.7.2.5) and sequences were compared locally using the BLOSUM62 substitution matrix and a cut-off expectation value of 1.0E-10. Tabulated results were visually scrutinized and all putatively divergent protein sequences (<90.0% amino acid identities) were manually extracted. Amino acid coding sequences were re-checked (ORF Finder), underwent functional characterization (UniPROT BLASTp; E-threshold < 10^{-5}; http://www.uniprot.org/blast/), and protein localization (Wolf PSORT; Horton *et al.* 2007), and were backwards blasted (BioEdit) to

identify putative paralogs and classify intra-species protein families. Significant stretches of sequence inconsistencies (indels > 3,000bp) between whole genomes were identified by global nucleotide alignment (EMBOSS Stretcher) of individual contigs from de novo assemblies. Telomeres were additionally re-assembled (30,000bp extending from the 5'-AAACCT-3' telomeric repeats) manually by identifying overlapping contig sequence data, with the exception of the 5' end of chromosomes 1 and 7, and the 3' end of chromosome 8, as these lack sequence evidence of telomeric repeats. The purpose of targeted telomere reassembly was to obtain detailed insight into gaps, insertions, and rearrangements in these highly recombinant regions, through graphical visualization of alignments (Artemis Comparison Tool; Carver *et al.* 2005).

The mechanistic drivers of divergence between orthologous protein sequences were also characterized. Single nucleotide insertions/deletions (indels) resulting in frameshift mutations were determined through alignment of highly-diverged (<90.0% AA IDs) orthologous proteins, which revealed instances where nucleotide changes had shifted the start/stop codon upstream or downstream, or altogether eradicated the open reading frame (EMBOSS Stretcher). The proportion of divergence due to amino acid gaps (SAAGs) versus polymorphisms (SAAPs) was similarly obtained. The direction of selective pressure between conserved orthologous coding sequences was evaluated through dN/dS calculation, using the CodeML analysis component of PAML (PAL2NAL; Suyama *et al.* 2006), and categorised as either negative (<1.0) or positive (>1.0). A student's T-test (significance cut-off P<0.05), assuming heteroscedasticity of recorded values, was used to assign any statistical significance to dN/dS trends across comparisons.

1.3 RESULTS

(i) Concatenated phylogeny of *Cryptosporidium spp.* WGS reveals intra-species genomic divergence and classification of a new *C. parvum* subspecies

A novel approach to phylogenetic evaluation of *Cryptosporidium spp.* was achieved through gene-to-gene concatenation of 61 neutrally-evolving loci across 21 whole genome isolates (Figure 1.3.1 A). Genome-wide representation of the evolutionary relationships was ensured through an even genomic distribution of targets (Figure 1.3.1 B), as well as a well-balanced neutrality of selection pressure (dN/dS 0.3-0.6) and divergence (0.4-0.7%) between orthologs of the reference *C. hominis* TU502 and *C. parvum* Iowa II subtypes (Figure 1.3.1 C). The proposed branching patterns and genetic distances of the concatenated phylogeny are confidently supported (bootstrap values > 99.0), lending statistical credibility to this maximum likelihood tree as a new phylogenetic gold standard for evolutionary interpretation of *Cryptosporidium spp*.

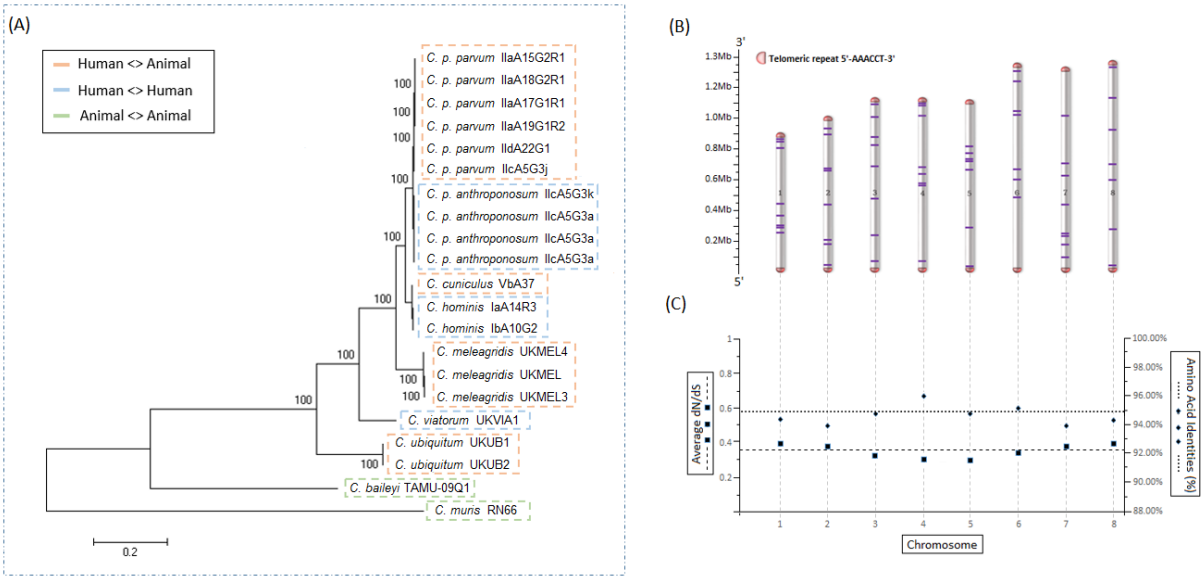


Figure 1.3.1. Concatenated phylogeny of neutrally-evolving loci reveals true evolutionary relationship between zoonotic and anthroponotic *Cryptosporidium spp.*

Concatenated phylogeny of *Cryptosporidium spp.* was constructed using 61 neutrally-evolving loci from 20 whole genome sequences. Loci were included for concatenation based on genome-wide distribution (B), dN/dS values of 0.3-0.6 and amino acid consensus of 93.0-96.0% identities between orthologous *C. parvum* (IIaA15G2R1) and *C. hominis* (IaA14R3) genes. Chromosome-wide averages between concatenated loci confirm consistency of the prescribed parameters (C), ascertained through EMBOSS Stretcher pairwise sequence alignment of concatenated sequences (The European Molecular Biology Open Software Suite; Rice *et al*. 2000) and dN/dS calculation (PAL2NAL; Suyama *et al*. 2006). ClustalW alignment and maximum-likelihood tree construction was performed using the Molecular Evolutionary Genetics Analysis tool (MEGA version 5.2; Tamura *et al*. 2011). Phylogeny was calculated using the Dayhoff substitution model and Nearest-Neighbour-Interchange (NNI) heuristic method, and 2,000 bootstrap replications were implemented to infer confidence values on the phylogenetic tree.

The phylogenetic reconstruction of whole genome sequences provides an improved understanding of the evolutionary basis of phenotypic intra- and inter-species differences of *Cryptosporidium spp*. Primary human-infective isolates form a distinct *C. hominis/cuniculus/C. parvum* superclade, indicative of their shared global nucleotide consensus ($\pi<0.04$) and genetic distinction from less common human-infectious species such as *C. meleagridis*, *C. viatorum*, and *C. ubiquitum* ($0.9<\pi<0.27$). The limited genetic distance between these two species (<0.02 nucleotide substitutions per site) illustrates the recent and rapid nature of the genetic change that distinguishes anthroponotic *C. hominis* from zoonotic *C. parvum*. The concatenated phylogeny further provides a preliminary genotypic association between phenotypically-diverse *C. parvum* strains. Classically zoonotic subtypes IIa and IId form a separate subclade from the presumptively anthroponotic IIc subtypes within the *C. parvum* clade. The taxonomic inference that can be drawn from this phylogenetic trait is the existence of two *C. parvum* subspecies, proposed to be renamed *C. parvum parvum* and *C. parvum anthroponosum* to allow for phenotypic distinction between zoonotic and anthroponotic strains respectively.

The concatenated phylogeny further provides insight into rates and mechanisms of divergence at neutrally-evolving sites between distantly related species. The single-host WGS species *C. muris* and *C. baileyi* reveal a 1.75 and 2.0-fold increase in divergence respectively compared to the next closest *C. ubiquitum* branch of the concatenated tree (Figure 1.3.2 B), corresponding to an overall divergence of 0.63 and 0.73 compared to the *C. parvum* IIaA15G2R1 reference concatenation. This highlights the degree with which adaptive evolution to non-human, narrow host range has influenced changes to amino acid coding sequences in these species over time, even at loci showing little evidence of selective pressure between human strains. Evaluation of selective pressure across strains demonstrates a steady decrease associated with increasing divergence, although this is likely an artefact of the inability of dN/dS to accurately assess positive versus negative selection between distantly related orthologous sequences (Hurst 2002). In addition, the analysis of gaps versus amino acid polymorphisms (SAAPs) featured here revealed an interesting trend, indicating that gaps play a more prominent role than SAAPs in driving initial variability between closely related, phenotypically-similar strains. This observation corresponds with previous reports of significant microsatellite variation between closely-related human-infective strains in *Cryptosporidium* (Widmer *et al.* 2004).

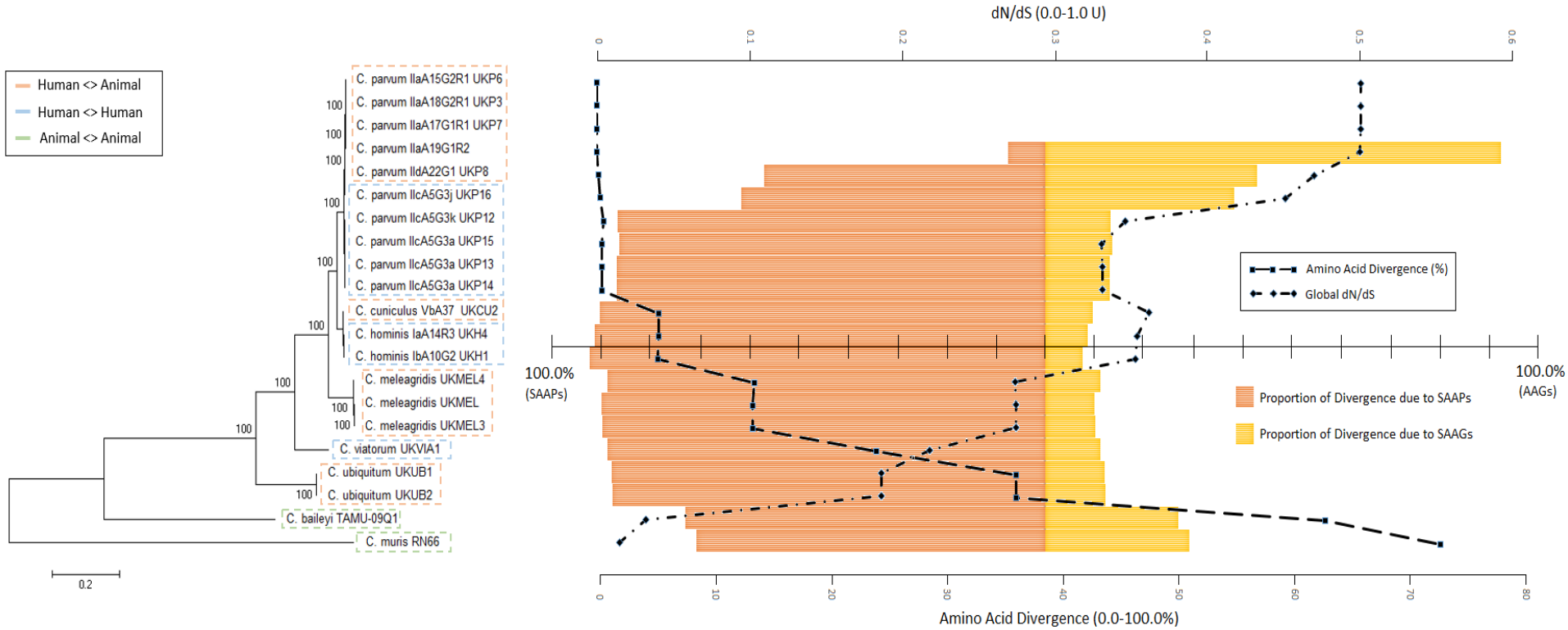


Figure 1.3.2. Divergence (AA % IDs) and selective pressure (dN/dS) of *Cryptosporidium spp*. novel concatenations relative to a IIaA15G2R1 reference-type *C. parvum* WGS (UKP6)

Cryptosporidium spp. WGS included in the concatenated phylogeny are described here based on protein sequence similarity (amino acid % IDs) and selective pressure (dN/dS) compared to a *C. parvum* reference-type genome (UKP6: GP60 subtype IIaA15G2R1), as well as the proportion of divergence due to SAAPs (single amino acid polymorphisms) or SAAGs (single amino acid gaps). The phylogeny annotates WGS according to the described mode of transmission for a given subtype family. The figure on the right demonstrates data points for divergence and selective pressure for the WGS with which they are aligned on the phylogeny. Patterns of divergence and the underlying genetic cause (SAAPs/SAAGs) reveal differential trends for WGS based on phylogenetic distance from the *C. parvum* reference.

(ii) Amino acid composition of neutrally-evolving concatenations reveals potential protein consequence of host adaptation between human and non-human infective strains

The adaptive changes to protein coding sequences are further highlighted by the relative amino acid composition of the concatenated sequences globally (Figure 1.3.3). Of note in this context are the two predominantly-zoonotic species included in the concatenated phylogeny, *C. muris* (murine host type) and *C. baileyi* (fowl host type), which only infrequently are isolated from human hosts (Palmer *et al.* 2003; Ditrich *et al.* 1991). Based on 5-95% confidence intervals around the mean amino acid composition (2 SD) of the concatenated sequences, a significant decrease in aspartic acid (D), glutamic acid (E), glycine (G), asparagine (N), proline (P), and serine (S) was identified for *C. muris*, and glycine (G) and arginine (R) for *C. baileyi*. Alternatively, a significant increase was observed in histidine (H), isoleucine (I), leucine (L), methionine (M), glutamine (Q), valine (V), and tyrosine (Y) composition in *C. muris*, and lysine (K) and asparagine (N) in *C. baileyi*. The prominent decrease in acidic amino acids (D, E) coupled with the sharp increase in aliphatic (I, L) and basic (H) amino acids in *C. muris*, could be attributed to a differing propensity for infection localization. *C. muris* preferentially infects the stomach, while *C. baileyi* and the human-infective species largely infect the small intestine (Xiao *et al.* 1999). Adaptation to a more acidic environment for the former, or a more basic environment for the latter, could in part resolve the observed shifts in *C. muris*' amino acid composition, and simultaneously explain why *C. baileyi* has largely retained a similar protein coding profile to the distantly related human-infective strains (Appendix II). However, the exact evolutionary origin of changes in amino acid composition is poorly understood (Hormoz, 2013), and the focus on neutrally-evolving *Cryptosporidium spp*. genes in this context makes it difficult to infer whether these trends represent genome-wide changes that could have resulted from host range or environmental changes. It is also possible that a number of the observed changes in amino acid frequencies are merely the result of expansions or contractions in repeat-rich sequences, or the consequence of a progressive loss or gain in protein sequence complexity that resulted from zoonotic to anthroponotic speciation events. This was investigated further by analysing the putative protein functions and predicted localizations for each of the genes included in the concatenated phylogeny (Appendix III). Only a small fraction (13.1%; 8 of 61) of the neutrally-evolving genes were predicted as having an extracellular localization,

with the majority localizing to the cytoplasm (34.4%; 21 of 61), plasma membrane (27.9%; 17 of 61) and the nucleus (24.6%; 15 of 61). This implies that the list of concatenated genes includes only few predicted host-parasite interactors, and decreases the possibility that environmental factors might have driven the observed changes in amino acid frequencies. The composition of amino acids also appears relatively stable across the protein sequences analysed, which further points towards subtle progressive changes in the make up of certain over- and under-represented amino acid types during the evolution of these species, and may have resulted in the observed pattern described in figure 1.3.3.

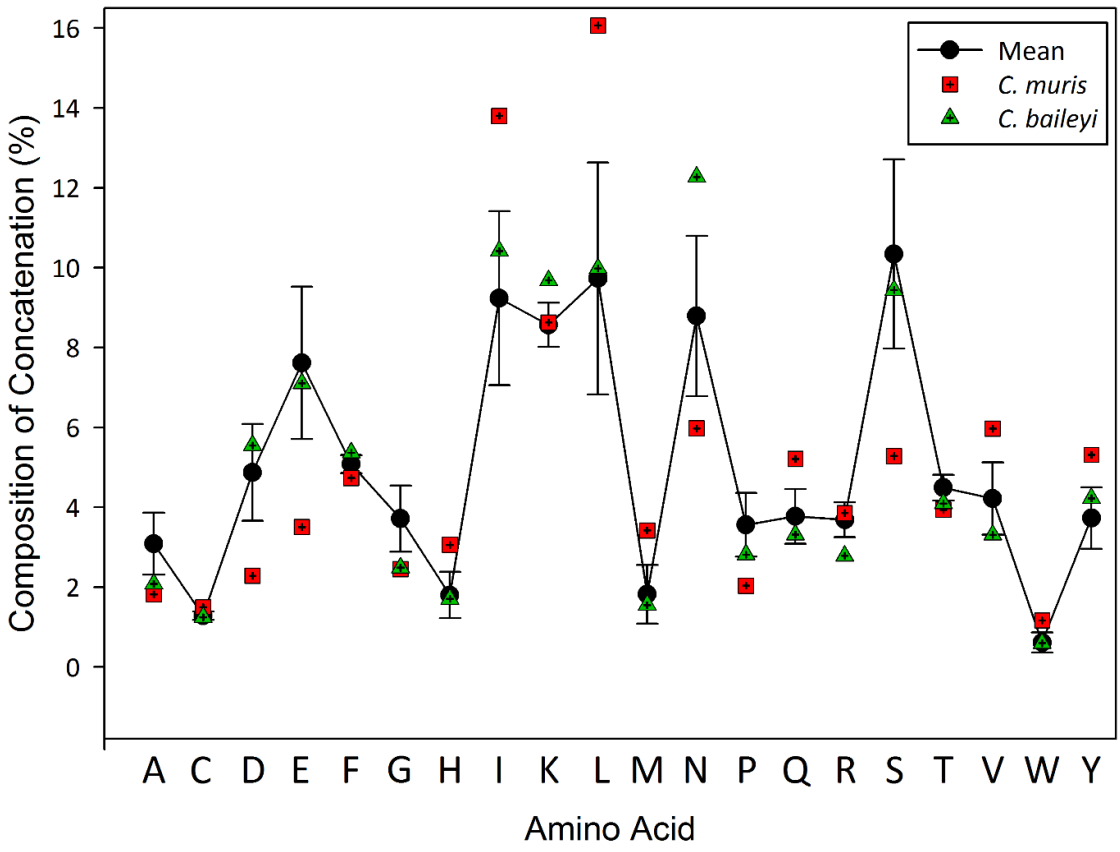


Figure 1.3.3. 5-95% Confidence Intervals for mean amino acid composition of concatenations (2 SD), combined with specific amino acid composition values for *C. baileyi* and *C. muris*

The degree of amino acid composition divergence between zoonotic *Cryptosporidium spp*. (*C. baileyi* and *C. muris*) and human-infective species is here illustrated. Mean amino acid composition across all *Cryptosporidium spp*. concatenations, along with amino acid compositions for *C. muris* and *C. baileyi* only, were calculated using MEGA (v 7.0; Kumar *et al*. 2016) and plotted graphically using SigmaPlot v 13.0 (Systat Software, San Jose, CA). 5-95% confidence intervals were calculated for mean amino acid composition data points, and are indicated by error bars on the figure.

(iii) Novel concatenated phylogeny revisits applicability of classic phylogenetic markers for classification of human-infective *Cryptosporidium spp.*

Three *Cryptosporidium* species-types account for more than 90.0% of all human infections: *C. parvum*, *C. hominis*, and to a lesser extent *C. meleagridis* (Qi *et al.* 2015). The concatenated phylogeny reveals that strains belonging to these species form a branch that is phylogenetically distinct from other strains, resulting in an apparent separation of clades based on a propensity to cause disease in humans. The close evolutionary distance between these species has made it difficult in the past to identify genotyping makers that accurately reflect whole-genome divergence through conventional means (ribosomal RNA, heat shock protein, etc.), which simultaneously makes it difficult to identify appropriate genetic targets for strain tracing and outbreak source tracking. The concatenated phylogeny specifically highlights the shortcomings of the five most widely-used typing markers at representing evolutionary divergence, demonstrating that single markers often display insufficient discriminatory power (D, E) or inaccurate genetic distances and relationships between strains (B, C, F).

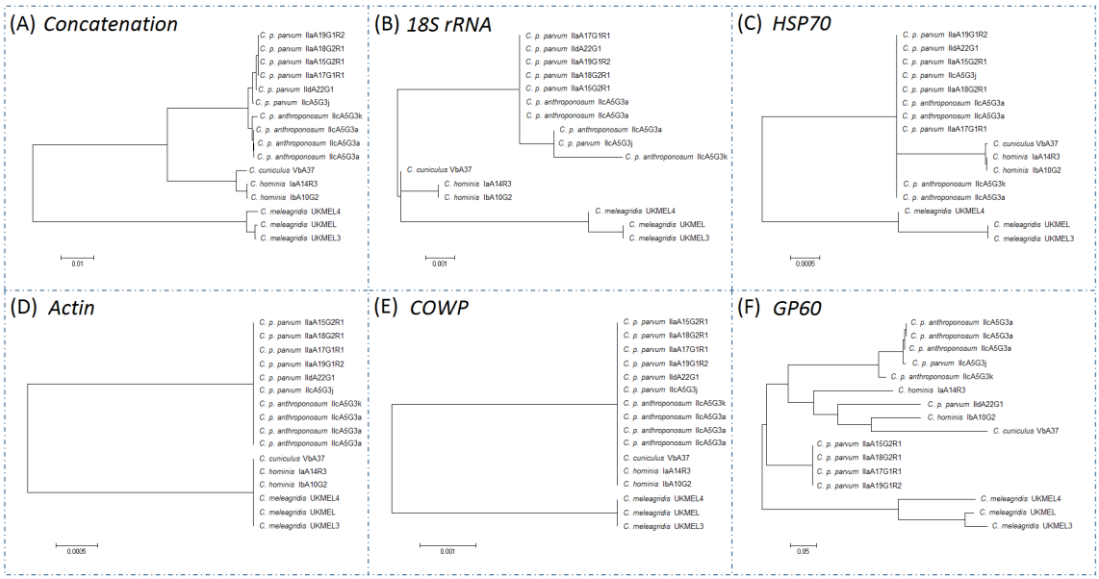


Figure 1.3.4. Comparative phylogenetics between the concatenated phylogeny and standard single-locus subtyping markers for human-infective *Cryptosporidium* genotyping and classification

Nucleotide sequences for five common subtyping loci were extracted from *C. hominis*, *C. parvum*, and *C. meleagridis* WGS using standard primers: (B) 18S rRNA (F 5'-GGAAGGGTTGTATTTATTAGATAAAG-3'; R 5'-AAGGAGTAAGGAACAACCTCCA-3'), (C) HSP70 (F 5'-TCATGTGTTGGTGTATGGAGA-3'; R 5'-CAACAGTTGGACCATTAGATCC-3'),(D) Actin (F 5'- CAAGCATTGGTTGTTGAT-3'; R 5'-TTTCTGTGTACAATTGATG-3'), (E) COWP (F 5'-ACCGCTTCTCAACAACCAT-3'; R 5'-TTGAGTGGGAACAGGTGCG-3') and (F) GP60 (F 5'-TCCGCTGTATTCTCAGCC-3'; R 5'- GCAGAGGAACCAGCATC-3'). Nucleotide sequences were aligned, translated into amino acids, and converted into Maximum Likelihood (ML) phylogenetic trees using MEGA v5.2 (Molecular Evolutionary Genetics Analysis version 5.2; Tamura *et al.* 2011). The ML phylogeny was calculated using the Dayhoff substitution model and Nearest-Neighbour-Interchange (NNI) heuristic method.

The limitations of traditional single locus targets, most notably the erroneous characterization of isolates based on the infamous GP60 subtyping locus (Figure 1.3.4. F), means there is still a vacancy for a universally-applicable typing tool that efficiently discriminates strains in a phenotype-associative manner. The generation of novel WGS and subsequent production of the concatenated phylogeny has provided a phylogenetic gold standard for isolate discrimination, but replicating this approach to genetically characterize each instance of cryptosporidiosis is unfeasible. Based on the phylogenetic grouping of human-infective *Cryptosporidium spp.* WGS in this study, three polymorphic single-gene alternatives were identified that mimic the relative positioning of strains, most importantly separating anthroponotic versus zoonotic isolates of *Cryptosporidium parvum* (Figure 1.3.5). Further significant benefit can be derived from the use of any of these three markers to definitively characterise "true" *C. p. anthroponosum* isolates versus cryptic examples that contain predominantly IIa-type genomes but a IIc-type GP60 (e.g. IIcA5G3j).

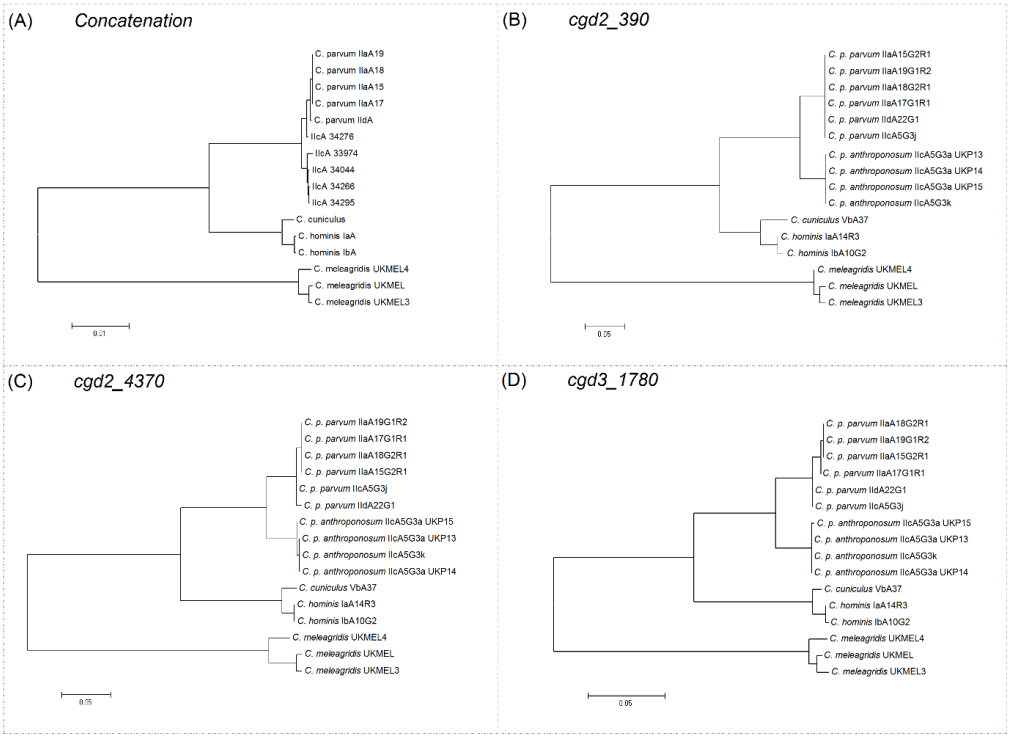


Figure 1.3.5. Single marker alternatives for phenotype-associative phylogenetic classification of *Cryptosporidium spp.* WGS
Human-infective WGS were compared using standalone BLASTn databases (BioEdit v7.2.5; Hall, 1999) to identify divergent (<90.0% amino acid identities) loci, and tabulated results were visually scrutinized to identify loci capable of segregating isolates in a manner resembling the concatenated phylogeny. Candidate loci were subsequently extracted from *C. hominis*, *C. parvum*, and *C. meleagridis* WGS to obtain full length gene sequences for (B) cgd2_390, (C) cgd2_4370, and (D) cgd3_1780, based on annotated coding sequences (Puiu, *et al.* 2004; http://www.cryptoDB.org). Nucleotide sequences were aligned and converted into Maximum Likelihood phylogenetic trees using MEGA v5.2 (Molecular Evolutionary Genetics Analysis version 5.2; Tamura *et al.* 2011). Phylogeny was calculated using the Dayhoff substitution model and Nearest-Neighbour-Interchange (NNI) heuristic method.

(iv) Common protein sequence divergence reveals convergent evolution of anthroponotic *Cryptosporidium spp.* WGS lineages

The level of genetic divergence and the rate of evolution does not appear to be uniform across the genome. Evaluation of genome-wide protein coding sequences identified 165 significantly diverged orthologous loci (<90.0% amino acid IDs) between *C. hominis* IaA14R3 and *C. parvum parvum* IIaA15G2R1 (Figure 1.3.6.A). Forty-six (27.9%) of these variable protein coding sequences are localized to the first and last 20,000bp extending from the 5'- and 3'-end telomeric repeats (i.e. the subtelomeric regions, representing ~3.5% of the genome). The subtelomeric regions are thus significantly more diverged than the rest of the genome (randomization test: $p<0.0001$). Evidence of host-adapted evolution through positive selective (dN/dS-values >1.0) was identified for 11.5% of these variable proteins, with 36.8% of these dN/dS-values affecting subtelomeric loci, which is a significant enrichment of sites under positive selection in the subtelomeric region (binomial test p-value 1.95×10^{-6}). In addition to positive selection, species-specific genome expansions and contractions (>500bp) were identified at the 5'-end of chromosome 5, at the 3'-end of chromosomes 2, 3, and 6, and 7, and approximately 170,000bp internally from the 5'-end of chromosome 8. Functional annotation of highly diverged loci indicates a predominance of mucins and other non-specific glycoproteins (N=34 (20.6%); Chi-sq. test: $\chi^2=196.21$; d.f.=1; $p<0.0001$), and an overrepresentation of secreted proteins (N=98 (59.4%); Chi-sq. test: $\chi^2=322.41$; d.f.=1; $p<0.0001$). The signature of selection, the contraction and expansion of subtelomeric regions, as well as the gene content of this region are all consistent with adaptive coevolution of a pathogen that employs rapidly evolving proteins located in the subtelomeric regions to enable cryptosporidiosis in a variety of hosts (Figure 1.3.6.B).

Analysing and comparing the 165 significantly diverged orthologous loci between *C. hominis* and *C. parvum anthroponosum* made it possible to study parallel evolution in the putatively anthroponotic WGS subtypes. *C. parvum anthroponosum* IIcA5G3a was highly diverged (>10.0%) at the amino acid level from *C. parvum parvum* IIaA15G2 at 54 of the 165 *C. hominis* divergent loci, and a further 42 showed a moderate amino acid divergence (3.0 – 10.0%) (Figure 1.3.6.D). These values are significantly higher than the genome-wide average of <1% amino acid divergence (binomial test: $p<0.0079$ for 96 out of 165 loci).

The impact of host-driven adaptation on protein level divergence is further substantiated by a tertiary comparison with a zoonotic *C. p. parvum* WGS subtype, IIdA22G1. This strain displays high amino acid divergence (>10.0%) at only 13 of these loci, which is significantly fewer than the 54 highly diverged loci in common between the two anthroponotic subtypes (Chi-sq. test: χ^2=33.881; d.f.=1; p<0.0001) (Figure 1.3.6.G). The genomic similarities in hypervariability between anthroponotic *Cryptosporidium* subtypes and obvious disparity from intra-zoonotic divergence provides evidence for host-specific adaptations and convergent evolution at the protein level.

In addition to the evidence of selective pressure being exerted on a common group of proteins, some genomic features are shared between *C. parvum anthroponosum* and *C. hominis* to the exclusion of the *C. parvum parvum* lineages. For instance, they lack a portion of the *C. p. parvum*-specific sequence data on chromosomes 5 and 7 (Figure 1.3.6.D). The absence of this sequence data has most notably resulted in two missing genes from the six-gene-strong *C. parvum parvum*-specific multi copy MEDLE gene family in *C. parvum anthroponosum,* and five missing genes in *C. hominis*. In addition, *C. parvum anthroponosum* uniquely lacks a 3,360bp stretch of sequence data at the 3'-end of chromosome 1. These vast structural chromosomal alterations, which again are not present in zoonotic *C. parvum* strain IIdA22G1 (Figure 1.3.6.G), highlight the potential relevance of telomeric sequence data loss as a major recombinatory process accompanying host species narrowing and adaptation in *Cryptosporidium*.

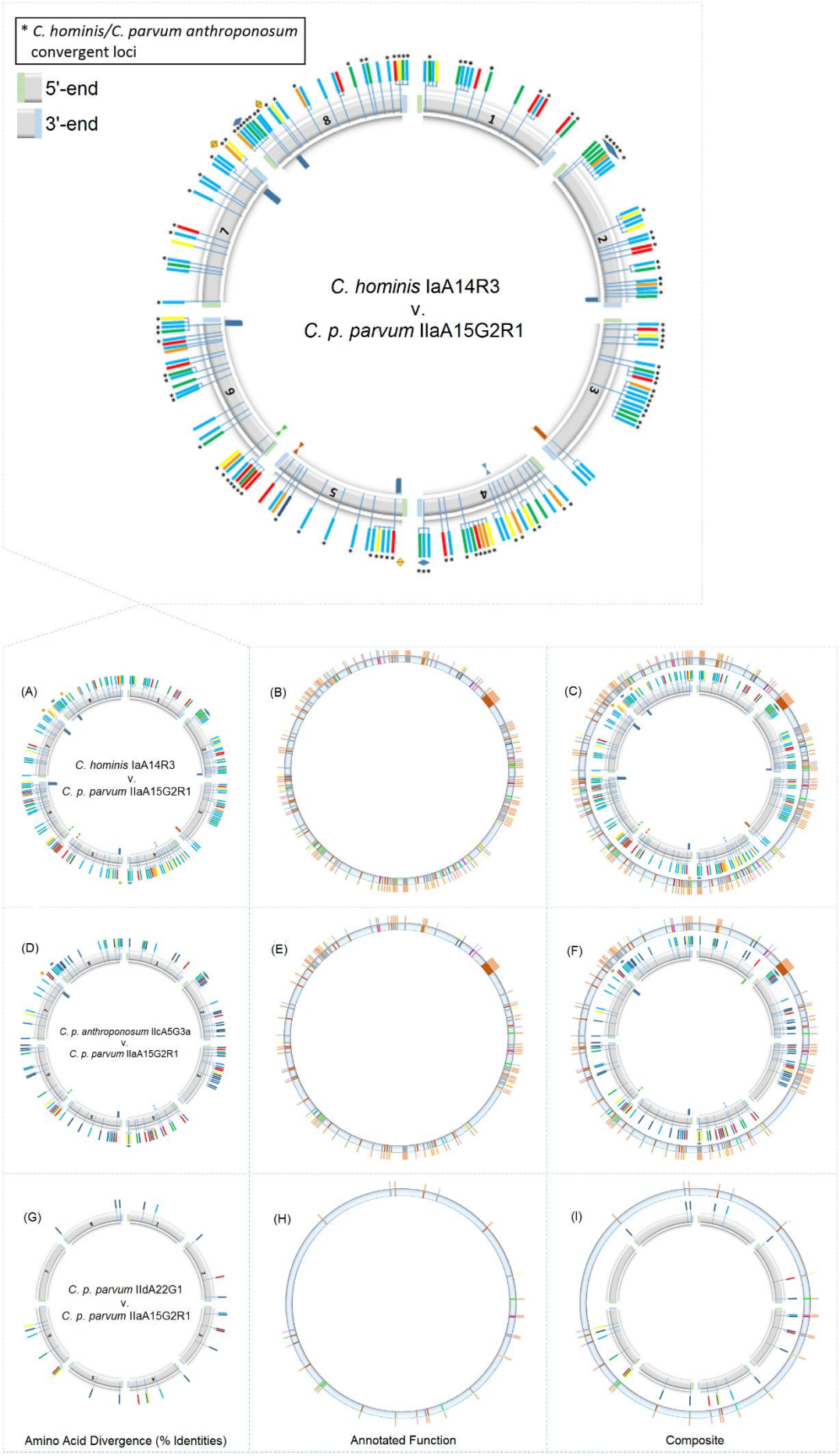


Figure 1.3.6. Comparative genomics of protein divergence reveals whole-genome consequence of host-driven adaptation in anthroponotic *Cryptosporidium spp*.

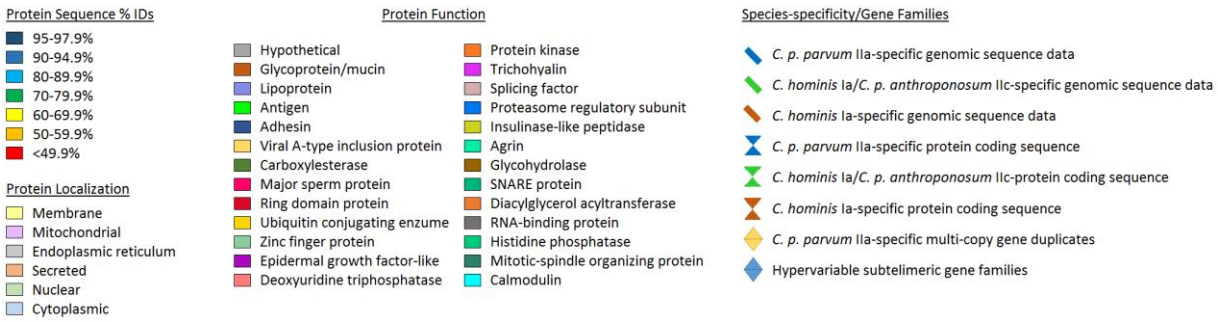


Figure 1.3.6. Comparative genomics of protein divergence reveals whole-genome consequence of host-driven adaptation in anthroponotic *Cryptosporidium spp*.

Whole-genome protein coding sequences were annotated from nucleotide sequence data (EMBOSS GetOrf, Rice *et al.* 2000) for *C. hominis* (IaA14R3) and three *C. parvum* isolates (zoonotic *C. p. parvum* strains IIaA15G2R1 and IIdA22G1, and anthroponotic *C. p. anthroponosum* IIcA5G3a) and cross-compared using standalone protein BLAST databases (BioEdit v7.2.5; Hall 1999). Columns display chromosome-wide protein divergence and structural alterations (A, D, G), putative protein function and localization (B, E, H) and composite representations (C, F, I). Parallel alignments and visual examination of whole chromosome sequence data revealed significant inconsistencies between whole genome sequence content, particularly featured adjacent to 5'-AAACCT-3' repeats (Artemis Comparison Tool; Carver *et al.* 2005), as well as frameshift mutations resulting in species-specific generation or loss of whole coding sequences (ORF Finder, NCBI; http://www.ncbi.nlm.nih.gov/projects/gorf/). Putative protein localization was inferred using Wolf PSORT (Protein Localization Predictor; Horton *et al.* 2007) and predicted protein function was determined by BLASTing protein sequences against The Universal Protein Resource UniPROT (E-threshold $<10^{-5}$; http://www.uniprot.org/blast/).

(v) Frameshift-causing indels and dN/dS characterization reveal micro-level genetic drivers of protein divergence

Descriptive characterization of protein divergence reveals a striking and recurring feature of zoonotic versus anthroponotic orthologous comparisons – frameshift-causing indels. These non-triplet genetic insertions and deletions were identified as the genetic mechanism behind protein divergence in 60 (36.4%) and 25 (23.6%) of variable *C. hominis* IaA4R3 and *C. p. anthroponosum* amino acid coding sequences respectively (Figure 1.3.7). The ability of frameshift-causing indels to dramatically alter genome composition is a well-documented consequence of host-adapted speciation in microbial pathogens. The fact that these types of genetic inflictions are present in both of the anthroponotic strains examined in this study, and that 84.0% of *C. p. anthroponosum* indels simultaneously affect the protein reservoir of *C. hominis*, provides an indication of the genomic neo-, sub-, and non-functionalization of proteins that accompanies narrowing of host range to a single, human host. Although a common subset of loci is subject to similar types of genetic changes between *C. hominis* and *C. p. anthroponosum*, they differ in terms of magnitude (size of non-triplet indel, degree of selective pressure) and location (specific

position within coding sequence), pointing to convergent evolution rather than recombination as the underlying cause of divergence. Exploration of selective pressure through dN/dS values indicating a positive (>1.0) versus negative (<1.0) direction identified evidence of positive selection in 19 (11.5%) and 26 (24.5%) of divergent *C. hominis* IaA4R3 and *C. p. anthroponosum* loci, respectively. The association of positive selective pressure with anthroponotic host-parasite adaptation is further exemplified by the significantly higher dN/dS levels in *C. p. anthroponosum* divergent proteins than in *C. p. parvum* IIdA22 (Student's *t*-test p-value <0.05).

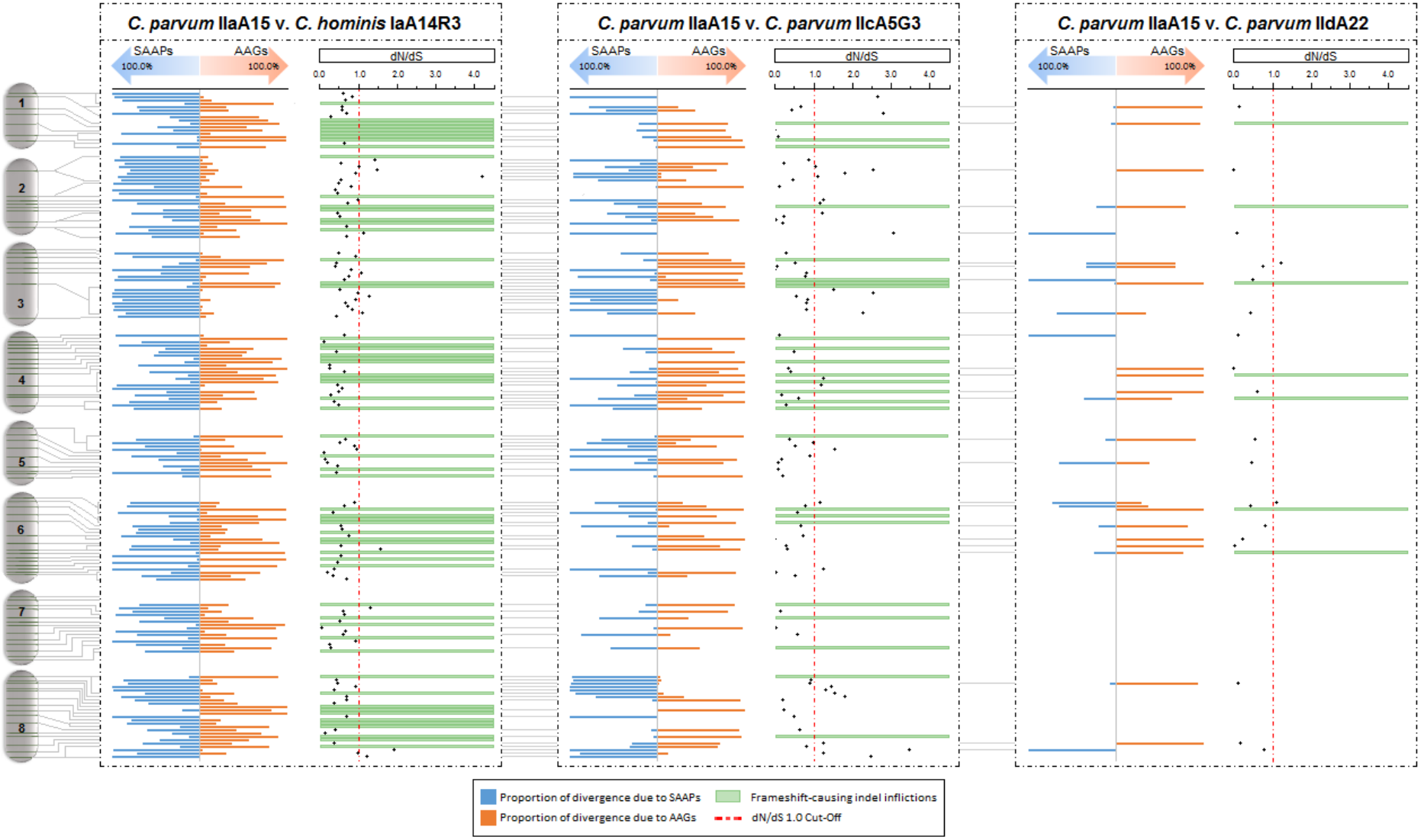


Figure 1.3.7. Characterization of amino acid divergence for variable loci reveals significant positive selective pressure and widespread frameshift inflictions between anthroponotic and zoonotic *Cryptosporidium spp*.

Figure 1.3.7. Characterization of amino acid divergence for variable loci reveals significant positive selective pressure and widespread frameshift inflictions between anthroponotic and zoonotic *Cryptosporidium spp*.

Amino acid divergence of hypervariable loci between anthroponotic *C. hominis* IaA14R3 and zoonotic *C. p. parvum* IIaA15G2R1 (<90.0% amino acid identities) and between anthroponotic *C. p. anthroponosum* IIcA5G3a and zoonotic *C. p. parvum* IIaA15G2R1 (<98.0% amino acid identities) is characterized and described. A tertiary comparison between two *C. parvum* zoonotic subtypes, *C. p. parvum* IIaA15G2R1 and *C. p. parvum* IIdA22G1, at divergent loci (<98.0% amino acid identities) is included. The proportion of non-consensus amino acid sites that can be attributed to single amino acid polymorphisms (SAAPs) versus amino acid gaps (AAGs) was determined through protein coding sequence alignment (The European Molecular Biology Open Software Suite; Rice *et al.* 2000). Protein sequence divergence resulting from frameshift-causing insertions or deletions (indels) was evaluated based on non-triplet nucleotide gaps between genetic coding sequences. dN/dS calculation was performed between true orthologous protein sequences, excluding those where frameshifts have resulted in an altered open reading frame (PAL2NAL; Suyama *et al.* 2006). Grey lozenges represent chromosomes 1-8 in a 5'->3' orientation, with extended grey lines representing relative chromosome location of hypervariable coding sequences and their respective divergence (% AA IDs) and selective pressure (dN/dS). Degree of selective pressure is detailed based on relative positioning above (positive) or below (negative) the 1.0 dN/dS cut-off value.

In addition to assessing selective pressure and diversity for highly-variable loci, universal trends of divergence across coding sequences were examined. For this purpose, an aligned concatenation of all coding sequences (5' -> 3'-end orientation) on chromosome 1 was used as an example. The results for dN/dS and nucleotide diversity (π) largely mirrored trends identified for highly-polymorphic loci across the intra- and inter-species comparisons. Average rates of π and dN/dS were 0.00032 and 0.089, 0.0022 and 0.334, and 0.0236 and 0.119 for comparisons between *C. p. parvum* (subtype IIa) and *C. p. parvum* (subtype IId), *C. p. anthroponosum* (IIc) and *C. hominis* Ia, respectively. The universally-low manifestation of divergence and selective pressure for the intra-*C. p. parvum* zoonotic IIa/IId comparison indicates a high degree of conservation and excess of synonymous mutations, describing a relationship of purifying selection between these two subtype families (Figure 1.3.8). The high rate of divergence yet moderately low selective pressure between zoonotic *C. p. parvum* IIa and anthroponotic *C. hominis* Ia is characteristic of the tail-end of a speciation event, as an already perfected host-parasite adaptation removes the incentive for further largescale positive selective pressure. The relationship between zoonotic *C. p. parvum* IIa and anthroponotic *C. p. anthroponosum* IIc, on the other hand, has signature features of diversifying selection, exhibiting less than one-tenth the average nucleotide diversity yet more than double (dN/dS 0.334 versus dN/dS 0.119) the selective pressure observed between *C. p. parvum* IIa and *C. hominis* Ia. It is important to also note that the significantly lower dN/dS between *C. p. parvum* and *C. hominis* in combination with the high rate of frame-shift causing InDels may be a signal of strong purifying selection, whereby an enhanced rate of deleterious non-synonymous mutations could have acted to shed certain superfluous genes in the more specialized *C. hominis*. A naturally-smaller or recently-reduced population size of *C. hominis* could

further explain the potential signature of purifying selection that was observed. This rapid accumulation of non-synonymous mutations is a key feature of recent evolutionary adaptation towards a novel host. Composite patterns of nucleotide diversity and selective pressure therefore demonstrate not only the genotypic basis of a specialist host range in *C. parvum anthroponosum*, but simultaneously provides a timeline that places this subspeciation event at a more recent historical divergence date than the speciation that occurred between *C. parvum* and *C. hominis*.

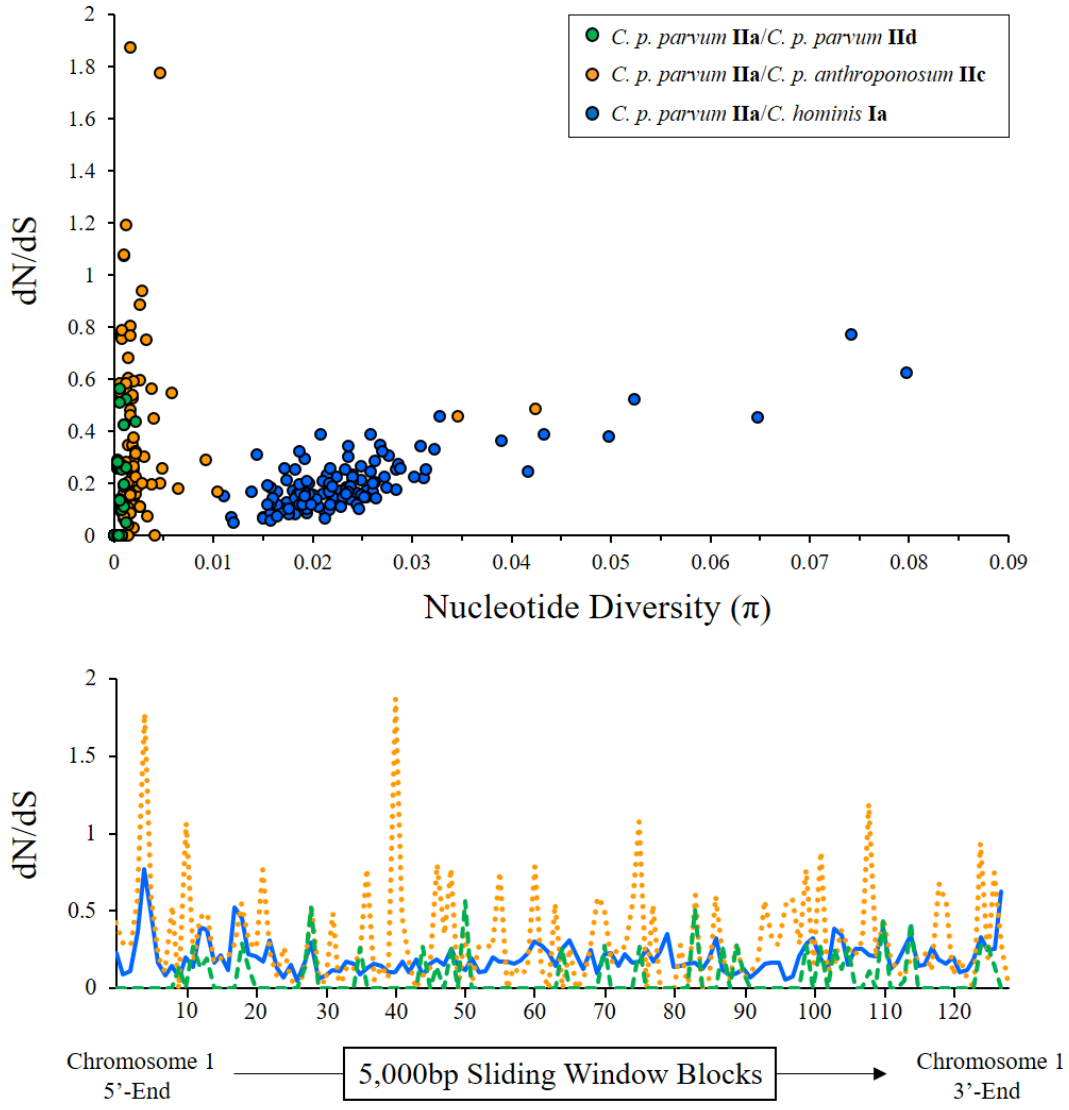


Figure 1.3.8. Sliding window nucleotide diversity (π) and selective pressure (dN/dS) across chromosome 1 coding sequences, for human-infective anthroponotic and zoonotic *Cryptosporidium spp.* WGS

5,000bp sliding window blocks were used to assess selective pressure (dN/dS) and nucleotide diversity (π) for chromosome 1 coding sequences between zoonotic *C. p. parvum* IIa and three other human-infective subtypes (zoonotic *C. p. parvum* IId, anthroponotic *C. p. anthroponosum* IIc, and *C. hominis* Ia). Coding sequences >300bp were initially extracted from chromosome 1 whole chromosome sequence data using Emboss GetOrf (EMBOSS GetOrf, Rice *et al.* 2000). These were subsequently aligned using the ClustalW alignment function for codons in MEGA v5.2 (Molecular Evolutionary Genetics Analysis version 5.2; Tamura *et al.* 2011), and species-specific coding sequences were manually removed. All subsequent sliding window analyses were conducted using 'polymorphism and divergence' functions embedded within the DnaSP software suite v. 5.10.1 (Librado & Rozas, 2009).

1.4 DISCUSSION

In this chapter I have provided the first genome-wide representative phylogenetic evidence of two distinct subspecies within the *C. parvum* species-group, using a novel concatenated approach based on 61 neutrally-evolving loci, and described some of the key potential drivers of anthroponosis versus zoonosis in cryptosporidiosis. Concatenated alignments of neutrally-evolving loci produced a novel phylogenetic representation of *Cryptosporidium* species; one that efficiently correlates phenotypic differences with genetic distance for closely-related human-infective strains based on non-recombinant regions of WGS (Figure 1.3.1). The identifiable sub-grouping of anthroponotic versus zoonotic strains within the *C. parvum* clade confirms the existence of two phenotypically distinct *C. parvum* lineages (as proposed by Widmer *et al*., 2012). Further comparisons between the common GP60 subtyping locus and the concatenated phylogeny reveals an inconsistency in the genome-wide divergence representation of this marker, and highlights the existence of a cryptic IIcA5G3j subtype that houses a IIc-type GP60 locus in a IIa-type genome (Figure 1.4.1). These distinct genotypic features could be attributed to the unique isolation of this subtype from hedgehog species (Sangster *et al*. 2016), revealing novel insight into the recombinatory processes that drive host adaptation between human and animal transmission shifts. This analysis suggests that these lineages be nominated subspecies rather than species due to the observation of relatively recent hybrid formation between the lineages exemplified in the WGS for IIcA5G3J. The subspecies should be named *C. parvum parvum* and *C. parvum anthroponosum* to improve the evolutionary classification within this species.

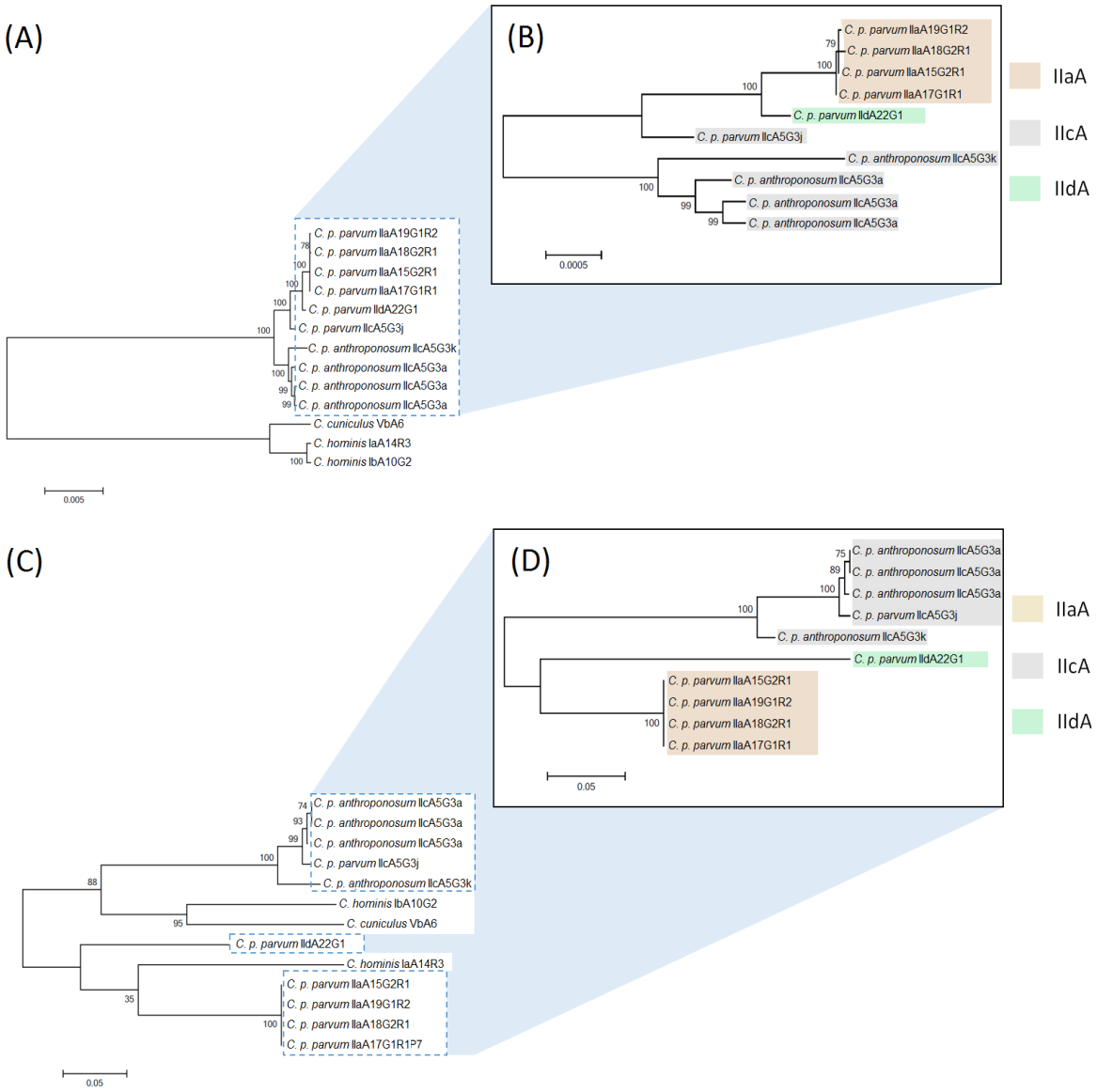


Figure 1.4.1. Concatenated phylogeny of human-infective *Cryptosporidium spp.* confirms new *C. parvum* subspecies (*Cryptosporidium parvum anthroponosum*) and revisits applicability of GP60 typing marker

The concatenated phylogeny of 61 neutrally-evolving amino acid sequences in *C. parvum* and *C. hominis* (A) and *C. parvum* only (B) WGS reveals the genetic relationship between and within the dominant human-infective *Cryptosporidium spp*. The *C. parvum anthroponosum* isolates form a distinct and separate clade from the neighbouring *C. parvum parvum* sister taxon, warranting re-classification of this anthroponotic group as a novel *C. parvum* subspecies. Construction of a comparative phylogenetic tree using standard GP60 typing sequences (5'-TCCGCTGTATTCTCAGCC-3' to 5'-GCAGAACCAGCATC-3') revealed strong discordance with the concatenated phylogeny for *C. hominis* and *C. parvum* WGS (C), and incorrect evolutionary grouping of *C. parvum* isolates (D). Phylogenetic positioning of *C. parvum* isolate IIcA5G3j, which contains a IIc-type GP60 sequence but shares greater genome-wide genetic similarity to the zoonotic *C. parvum parvum* subspecies, reveals incongruence between the concatenated versus single locus phylogenetic approach. Confidence values on the maximum likelihood trees (MEGA 5.2; Tamura *et al.* 2011) were generated using 2,000 bootstrap replications.

Previously-reported genotyping markers were also assessed within the context of this study, and revealed that there is a lack of discriminatory consensus between commonly used targets. An evaluation of the five most popular markers revealed that phylogenetic groupings and genetic distances were consistently inconsistent with one another for discriminating the novel WGS (Figure 1.3.4). None of the markers produced phylogenies that successfully mirrored the concatenated phylogeny, and the Actin gene proved particularly problematic when it grouped *C. hominis* with *C. meleagridis*, and further showed no discrimination between *C. parvum anthroponosum* and *C. parvum parvum*; this demonstrates the Actin marker as lacking discriminatory power at the subspecies level, as well as producing misleading genotypic descriptions of human-infective *Cryptosporidium spp*. GP60 produced a phylogeny that exhibited the weakest correlation with whole genome-divergence, creating entirely unexpected and unexplained phylogenetic patterns between *C. hominis*, *C. parvum*, *C. meleagridis*, and *C. cuniculus*. However, this is likely explained by the previously-demonstrated occurrence of recombination at this locus (Li *et al*. 2013; Guo *et al*. 2015). The hypervariability and recombinant nature of GP60 means it will likely never be effective as a genotyping marker aimed at describing genomic proximities, but rather continue to receive focus as a potential driver of host-parasite interactions in cryptosporidiosis. In this context, the proximity of the *C. hominis* GP60 subtype Ia to anthroponotic *C. parvum* WGS, versus the proximity of *C. hominis* GP60 subtype Ib to zoonotic *C. parvum* IId, and zoonotic *C. cuniculus*, could be of importance (and will be discussed in chapter 3).

Although it could be argued that each marker is useful in its own targeted way, such as the inter-species separation achieved by 18S rRNA, there have been repeated calls from within the *Cryptosporidium* community for novel genotyping methods that accurately describe genome-wide patterns of divergence (Cacciò & Widmer, 2013; Sikora *et al*. 2016). This study proposes three single-marker alternatives, cgd2_390, cgd2_4370, and cgd3_1780 (Figure 1.3.5), which produce comparable phylogenies to the concatenated sequence. Most importantly, they separate the *C. parvum parvum* strains from the *C. parvum anthroponosum* strains, consistently discriminate *C. hominis* from *C. cuniculus* within the *C. hominis*/*C. cuniculus* clade, and repeatedly differentiate UKMEL4 from the other two *C. meleagridis* strains. This provides a strong incentive for moving these markers forward to the next level of validation, by testing whether their efficiency is

retained on larger and more epidemiologically-diverse *Cryptosporidium* sequence datasets. An important component to this will be the improved epidemiological characterisation of *Cryptosporidium*, particularly in-depth classification of the true host ranges and pathogenicity associated with various GP60 subtypes, as it will highlight how well these new markers establish a link between phenotype and genotype. Only then will the *Cryptosporidium* community likely consider replacing outdated and inefficient genotyping approaches, and give these novel markers their scientific endorsement.

The combined divergence and selective pressure analyses performed during this whole genome analysis provides one of the first comprehensive overviews of the micro- and macro-level changes accompanying host specialisation (or generalisation) in *Cryptosporidium spp*. It touched upon many of the important themes and results commonly associated with speciation and subspeciation events, and simultaneously identified a significant new toolset of putative virulence factors. Species-specific gene duplication events were one of the major findings of this study, confirming previous reports on the existence of a *C. p. parvum*-specific MEDLE gene family (Guo *et al*. 2015), but also expanding on this to include specific genomic locations and further examples of *C. p. anthroponosum* and *C. hominis*-specific genes throughout the genome (Figure 1.3.6). It is important, however, to note the variability in the quality and coverage of the whole genome sequences included in this comparison (Appendix I). Of particular importance here is the significant differences in whole genome and scaffold sizes between the various human-infective isolates, which highlights a number of shortcomings due to potentially missing sequence data and a compromised sequence quality and coverage (e.g. an excess of non-specific (N) nucleotides). Access to improved WGS generated from more sophisticated sequencing and assembly platforms will therefore help to expand the comparative genomics analyses between and within the subtypes included in this study, the importance or which is clearly highlighted by the inconsistent assembly statistics described. Despite this limitation, the findings of this study provide numerous examples of gene duplication and deletion events that could be validated across same-subtype family WGS. This genetic mechanism has been observed during phenotypic shifts towards generalisation and specialisation, respectively, in other host-parasite systems (Cole *et al*. 2001; Chain *et al*. 2004; Cuomo *et al*. 2007; Hu *et al*. 2014), and this in silico finding

therefore may provide an important insight into the evolutionary processes driving host range in *C. hominis* and *C. parvum*.

Divergence can tell us a lot about the extent, rate, and consequences of evolution between closely-related species. In this specific study, the functional characterisations and genomic locations of divergence went a step further, providing not only a quantitative analysis of the variability but also beginning to highlight the true nature of pathogenicity in human-infective *Cryptosporidium*. The first divergence feature worth noting is the limited genome-wide proportion of coding sequences which exhibited significantly variable nucleotide identities between *C. p. parvum* IIa and *C. hominis* Ia (Figure 1.4.2). This targeted elevation of divergence and selective pressure, affecting less than 5.0% of all coding sequences, points toward a recent, rapid, and highly effective speciation event that resulted from a perhaps unexpected and strong environmental pressure to evolve and adapt. It also indicates that the genetic factors most vital for successful pathogenicity and proliferation within the host, i.e. the infamous virulence factors, can likely be found among the 165 highly-diverged (<90.0% amino acid IDs) genes identified in this study (Figure 1.3.6).

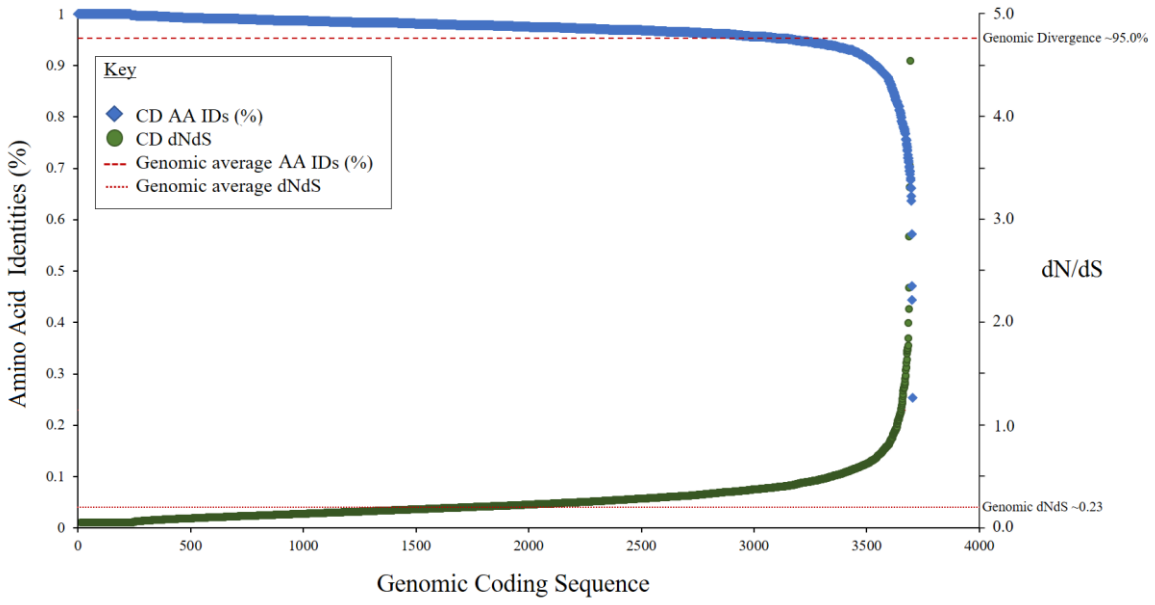


Figure 1.4.2. Genome-wide dN/dS and amino acid identities (%) between *C. parvum* IIa and *C. hominis* Ia coding sequences
Amino acid identities (%) and dN/dS values between complete coding sequences for reference *C. parvum* (Iowa II) and *C. hominis* (TU502) whole genome sequences were obtained from Mazurie *et al*. (2013; supplementary table 3). Individual values for AA IDs (%) and dN/dS are presented, along with the genomic average for these values as determed from whole genome alignments analysed using DnaSP (v 5.10.1; Librado & Rozas, 2009).

The functional annotation of highly-diverged genes further adds to the big picture of virulence and infectivity, with two predominant themes being secretion and glycoproteins. Bacterial and protozoan pathogens rely heavily on extracellular proteins to both defend and attack, and have evolved complex virulence secretion systems that efficiently interact with, modify, and manipulate host cells (China & Goffaux 1999; Souza 2006; De Buck *et al*. 2007; Büttner & Bonas, 2010; Voth *et al*. 2012; Neves *et al*. 2014; Atayde *et al*. 2015). Previous studies have suggested links between secreted proteins in *Cryptosporidium* and host epithelial cell damage, and postulated that secreted enterotoxins may significantly impact the pathogenic manifestations of disease, but the molecular pathways and genetic factors remain poorly understood (Guarino *et al*. 1995; Gookin *et al*. 2002). The significant over-representation of secreted proteins in this study therefore could indicate a crucial role of these in human *Cryptosporidium* infection. Glycoproteins and specifically mucins were also significantly represented amongst the list of putative virulence candidates, and have been implicated as powerful virulence factors in a number of other pathogenic protozoa, such as the variant surface glycoproteins in Trypanosomes (Horn, 2014). A subset of these highly-variable cryptosporidial mucins, located at the 5'-subtelomere on chromosome 2, have already been investigated in a number of epidemiological and in vitro settings (Cervallos *et al*. 2000; O'connor *et al*. 2009a; O'connor *et al*. 2009b; Paluszynski *et al*. 2014). The discovery that these mucins are not only highly variable across a range of geographic and clinical settings, but also likely to play an indispensable role in attachment and invasion, makes their involvement in virulence increasingly credible, and places them high on the candidate list of putative drug targets and potential markers of disease. These virulence-associated features also warrant further investigation of some of the currently unexplored highly-variable mucins and glycoproteins identified through this genomic study.

Genetic exchange as a driving force for phenotypic diversity in human-infective *Cryptosporidium spp.* - and the historical context of divergence

2.1 INTRODUCTION

C. hominis and *C. parvum* have consistently been reported as exhibiting a high average global consensus of ~95.0-97.0% nucleotide identities (Widmer & Sullivan, 2010; Mazurie, *et al*. 2013), yet the former has predominantly been reported from human cases, while the latter permanently traverses human-livestock-wildlife transmission boundaries (Fayer & Xiao, 2007). The high level of genomic conservation between these species and subtypes indicates that phenotypic variation may reflect elevated genetic exchange at virulence-associated loci, rather than slow evolutionary progression of divergence; a view which is mirrored in other pathogenic protozoa (Freitas-Junior *et al*. 2000). With the emerging view that generalist pathogens may evolve novel virulence factors through recombination and genetic introgression (McMullan *et al*. 2015), whole genome comparisons between pathogenic strains, subtypes or races are particularly insightful. Such studies help to identify significant hybridisation events and evolutionary changes to protein repertoires, and hence, they are key to understanding the genetic underpinnings and chronological context of host-pathogen coevolution.

(i) Sex or no sex – the inevitable debate of Cryptosporidial sexual affairs

Genetic exchange is one of the most important drivers of immune evasion and diversification in human pathogenic protozoa. Although a basic probability for recombination exists in all DNA-replicating eukaryotes, the occurrence of the more impactful meiotic recombination is dependent on an important, and less common, biological prerequisite: sex. The potential for sexual reproduction in the life cycle of *Cryptosporidium* has been well-documented and accepted, ever since the first macro- (female) and microgametes (male) were visualized and described (Figure 2.1.1) many decades ago (Barker & Carbonell, 1974; Bird & Smith, 1980; Sanford & Josephson, 1982). For a long time, however, the frequency at which this potential was actually fulfilled, as well as the extent to which separate strains, subtypes, and species were capable of interbreeding, remained elusive. In addition, the lack of effective in vitro culturing techniques meant that the occurrence of sexual reproduction had not previously been observed "real-time" in *Cryptosporidium* (Malik *et al*. 2007). Despite these gaps and limitations, a partially sexual reproductive nature has largely been accepted and demonstrated within the scientific community (Chen *et al*. 2002; Hijjawi *et al*. 2004; Leitch & He, 2012), and where sex occurs, the potential for meiotic recombination follows. This means that the minimum evidentiary support required for investigating genetic exchange as an operator for diversity in *Cryptosporidium* has, at least in theory, been fulfilled (Weedall & Hall, 2015).

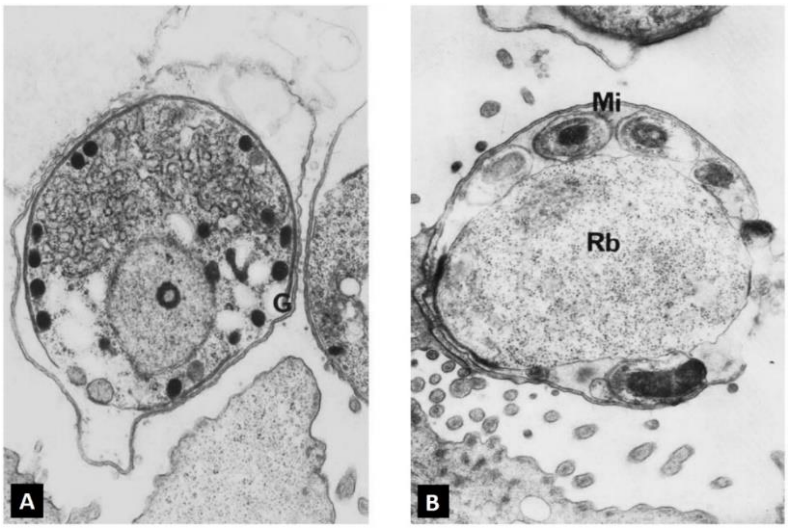


Figure 2.1.1. Electron microscope image of a *Cryptosporidium spp*. macrogametocyte and microgametocyte Image of a macrogametocyte (A) containing macrogametes (spherical bodies) and a microgametocyte (B) containing microgametes (oval bodies) at a magnification of x20,000 and x24,000, respectively (Bird & Smith, 1980).

The first indication of sexual recombination in *Cryptosporidium* came with the introduction of molecular techniques to differentiate epidemiologically-unique strains. Genetic similarities between geographically-distinct *C. parvum* isolates through an RFLP-based approach gave rise to the first supported claim of genetic exchange between human-infective strains (Widmer *et al*. 1998). Experimental genetic crosses in laboratory mice subsequently produced the first instances of recombinant progeny (Feng *et al*. 2002; Tanriverdi *et al*. 2007), and extended microsatellite analyses of human and cattle *C. parvum* isolates showed that genetic exchange does indeed happen in nature (Mallon *et al*. 2003), providing much-needed experimental validation for the underlying biological mechanisms involved. With an increasing portfolio of genomic data, and in particular the extended genetic characterization of the hyperpolymorphic GP60 locus (formerly known as Cpgp40/15), the true extent of genetic exchange within human *Cryptosporidium spp.* has started to emerge. GP60 alleles, which frequently defy genotypic exclusivity and seem to readily "hop" between subtypes, have since been the primary feature of recombination studies in this genus (Leav *et al*. 2002; Peng *et al*. 2003; Gatei *et al*. 2007; genetic exchange, however, more recent attempts to identify and enumerate recombination signals at the genomic level have remained largely limited to GP60 and less than a handful of other loci (Guo *et al*. 2015; Tang *et al*. 2016). Thus, there still remains a large knowledge gap regarding the role of genetic exchange in driving pathogenic changes in cryptosporidiosis. Delineating the presence, features, and consequences of genetic exchange is therefore key to understanding its role in driving phenotypic diversity between genetically-similar, human-infective, *Cryptosporidium* strains and species.

(ii) Mechanistic Capacity for Genetic Exchange in *Cryptosporidium*

Genetic exchange is an essential interaction between DNA molecules that promotes the transposition or interchange of genetic material between or within a species. While recombination primarily serves as a complex DNA repair mechanism and telomere maintenance facilitator, vital for preserving genome stability and integrity (Thompson & Schild, 1999; Heyer 2015; Tacconi & Tarsounas, 2015), it is also one of the most effective tools for promoting allelic diversity and contributing to new or altered phylogenetic profiles. This latter impact of recombination is often particularly pronounced in

pathogenic organisms, where exchange or transfer or genetic data between variably- or differentially-virulent strains often serves to impact the survival and/or pathogenic potential within a given host.

In many other pathogenic protozoa, like in *Cryptosporidium*, reproduction was referred to as predominantly, if not wholly, asexual. Scientific acceptance of this fact was largely a by-product of the comprehensive "clonal theory of parasitic protozoa" presented in the early 1990s (Tibayrenc *et al*. 1990), a controversial argument of exclusive clonality made in the absence of modern molecular techniques and genomics analyses. As a result of this, the presence, nature, and molecular drivers of genetic exchange have only recently begun to be fully explored and understood, and with this, the contribution towards fitness-driven genetic diversity and immune-evasive tactics. Over the past decade, molecular exploration through means such as transfection and hybrid cloning studies, in compilation with in silico characterization of genotype diversity and distribution, have implicated recombinatory processes as effective drivers of virulence in some of the most infamous human parasites, including (but not limited to): *Trypanosoma cruzi* (Gaunt *et al*, 2003; Baptista Rde *et al*. 2014; Weatherly *et al*. 2016), *Leishmania* (*Viannia* and *Leishmania*) species (Kuhls *et al*, 2013; Genois *et al*. 2015), *Toxoplasma Gondii* (Grigg & Suzuki, 2003; Minot *et al*. 2012), *Giardia duodenalis* (Cooper *et al*. 2007; Kosuwin *et al*. 2010), and of course *Plasmodium* (Dias *et al*. 2011; Sutton *et al*. 2011; Jiang *et al*. 2011; Kirkman & Deltsch, 2014) and African *Trypanosome spp.* (Conway *et al*. 2002; Hartley & McCulloch, 2008; Gibson, 2015; McCulloch *et al*. 2015).

Genetic recombination studies in protozoa have largely been focused on *Plasmodium falciparum* and *Trypanosoma brucei*, the primary parasitic causes of severe malaria and African sleeping sickness respectively. This bias, although considerable, is defendable due to the lethal biological ability of these organisms to evade immune responses and rapidly develop resistance to antimicrobials, a characteristic that directly contributes to the devastatingly high mortality rates that persist across much of Sub-Saharan Africa and parts of Southeast Asia (Odiit *et al*. 1997; Stich *et al*. 2002; Nadjm & Behrens, 2012; Kumar *et al*. 2012). The role of recombination through genetic exchange in facilitating this antigenic variation and drug resistance has been well-documented, with a number of the associated recombination hotspots and putative breakpoints identified to date, but the

underlying mechanisms are still largely based on in silico inferences or translated from well-characterised model organisms (Morrison *et al*. 2009; Jiang *et al*. 2011; Sander *et al*. 2014). The physical modus operandi of recombination in these species is predominantly facilitated by DNA double-strand break repair (DSBR) systems (Glover *et al*. 2011; Lee *et al*. 2014), which subsequently lead to genetic transfer or chromosomal rearrangements through homologous, micro-homologous or non-homologous end joining (MMEJ and NHEJ, respectively; Vink *et al*. 2012).

The underlying genetic features involved in these complex processes have been relatively well-defined in *Plasmodium* and *Trypanosoma spp*., initially through homology comparisons against known recombination proteins in model organisms such as yeast. In *Saccharomyces cerevisiae*, the enzymes involved in recombination have been described in much detail (Krogh & Symington, 2004), and by identifying homologous protein coding sequences in other species it is possible to identify whether a likely capacity for recombination exists. RAD51, the eukaryotic homolog of the essential recombination protein RecA in *E. coli*, proved vital for the homologous recombination process in yeast when mutants proved defective at the process (Bhattacharyya *et al*. 2004). Homologs have since been identified in both *Plasmodium* (Bhattacharyya & Kumar, 2003) and *Trypanosoma spp.*, with RAD51-defective *T. brucei* mutants displaying an impaired VSG switching mechanism (McCulloch & Barry, 1999), and PfRad51(K143R) *Plasmodium* mutants associated with a significantly diminished parasite burden within the host (Roy *et al*. 2014). MRE11, which partners with RAD50 and XRS2 to form the MRX complex, is another vital player in DBS repair (Tsubouchi & Ogawa, 1998), and null mutants have demonstrated significant hypersensitivity to DNA damaging agents in *Plasmodium* (Badugu *et al*. 2015) and *Trypanosoma*, although in the latter an associated impairment to VSG switching did not take effect (Robinson *et al*. 2002).

In *Cryptosporidium*, the limitations of in vitro cultivation technologies have delayed progress in understanding the true extent, nature, and impact of genetic exchange in this parasite. Despite this, a solid platform of genetic diversity and putative genetic exchange has already been laid through in silico exploration, and with an increasing number of WGS becoming available there is much interest and incentive to expanding this platform, even in the absence of experimental validation. As a pre-requisite to this WGS genetic

exchange study, an extensive search of Cryptosporidial homologs to characterized recombination proteins in *Saccharomyces cerevisiae*, of which there are around forty (Krogh & Symington, 2004), was performed. This revealed significant hits to more than half of these (20 of 38; E-value < E-04), and an incredible degree of conservation in homologs of the important RAD51 and MRE11 mediators discussed above (Table 2.1.1). Although the absence of a number of these proteins (18 of 38; 47.4%) could seem significant, their absence was confirmed in other protozoa in which recombination has been experimentally confirmed, and could point to drastic divergence of homologs (E-value hits > E-04) rather than complete genomic eradication. This demonstrates a potential mechanistic capacity for genetic exchange in *Cryptosporidium*, providing additional incentive for performing the extensive cross-strain and species genetic exchange analyses for the novel WGS unique to this study.

Gene Name	Organism	UniProt ID	Function	Length (AA)	Ortholog in *C. parvum*	Length (AA)	%IDs versus *C. parvum*	E-value	Score
DMC1	*S. cerevisiae*	P25453	Meiotic recombination, synaptonemal complex formation and cell cycle progression	334	cgd7_1690	342	46.3%	9.00E-87	812
DNL4	*S. cerevisiae*	Q08387	DNA joining activity	944	cgd3_3820	825	21.3%	4.00E-21	296
EXO1	*S. cerevisiae*	P39875	DNA exonuclease involved in mismatch repair	702	cgd6_1580	482	29.7%	7.00E-45	503
YKU70	*S. cerevisiae*	P32807	Non-homologous end joining (NHEJ) DNA double strand break repair	602	None	N/A	N/A	N/A	N/A
YKU80	*S. cerevisiae*	Q04437	Non-homologous end joining (NHEJ) DNA double strand break repair	629	None	N/A	N/A	N/A	N/A
LIF1	*S. cerevisiae*	P53150	Non-homologous repair of DNA double-strand breaks	421	None	N/A	N/A	N/A	N/A
MEI4	*S. cerevisiae*	P29467	Meiotic induction of recombination, viable spore production and chromosome synapsis	408	None	N/A	N/A	N/A	N/A
MER1	*S. cerevisiae*	P16523	Chromosome pairing and genetic recombination	270	None	N/A	N/A	N/A	N/A
MER2	*S. cerevisiae*	P21651	Meiotic chromosome segregation	314	None	N/A	N/A	N/A	N/A
MMS4	*S. cerevisiae*	P38257	Repair of meiosis-specific double strand breaks	691	None	N/A	N/A	N/A	N/A
MRE2	*S. cerevisiae*	Q00539	Suppressor of mitochondrial splicing deficiencies	523	cgd1_2730	693	20.0%	7.00E-12	197.5
MRE11	*S. cerevisiae*	P32829	DNA double-strand break repair (DSBR)	692	cgd1_1420	513	34.3%	3.00E-67	716.5
MSH4	*S. cerevisiae*	P40965	Crossover facilitation during meiosis	878	cgd8_3950	848	23.4%	8.00E-35	477
MSH5	*S. cerevisiae*	Q12175	Crossover facilitation during meiosis	901	cgd8_370	1242	20.8%	9.00E-32	444.5
MUS81	*S. cerevisiae*	Q04149	Repair of meiosis-specific double strand breaks	632	cgd1_3420	734	20.9%	6.00E-08	245.5
NEJ1	*S. cerevisiae*	Q06148	Involved in non-homologous end joining (NHEJ)	342	None	N/A	N/A	N/A	N/A
RAD6	*S. cerevisiae*	P06104	Double-strand break repair via homologous recombination	172	cgd3_2670	123	52.8%	1.00E-43	359.0
RAD18	*S. cerevisiae*	P10862	Postreplicative repair (PRR) of damaged DNA	487	cgd2_880	633	32.9%	7.00E-06	106.5
RAD50	*S. cerevisiae*	P12753	DNA double-strand break repair (DSBR)	1312	cgd1_2410	1062	29.0%	2.00E-68	1221.5
RAD51	*S. cerevisiae*	P25454	Recombination and DNA damage repair	400	cgd5_410	347	51.5%	1.00E-107	982.5
RAD52	*S. cerevisiae*	P06778	DNA double-strand break (DSB) repair and recombination	471	None	N/A	N/A	N/A	N/A
RAD54	*S. cerevisiae*	P32863	DNA repair and recombination	898	cgd1_1180	877	33.9%	2.00E-129	1247.5
RAD55	*S. cerevisiae*	P38953	Recombination and recombinational DNA repair pathways	406	None	N/A	N/A	N/A	N/A
RAD57	*S. cerevisiae*	P25301	DNA recombinase assembly	460	cgd2_4070	304	18.6%	2.7e-8	135.5
RAD59	*S. cerevisiae*	Q12223	Repair of double-strand breaks in DNA	238	None	N/A	N/A	N/A	N/A
REC102	*S. cerevisiae*	Q02721	Formation of double-strand breaks (DSBs) that initiate meiotic recombination	264	None	N/A	N/A	N/A	N/A
REC104	*S. cerevisiae*	P33323	Transcriptional regulator required for expression of early meiotic genes	182	None	N/A	N/A	N/A	N/A
REC114	*S. cerevisiae*	P32841	Production of double-strand breaks (DSBs) in meiotic recombination initiation	428	None	N/A	N/A	N/A	N/A
RFA1	*S. cerevisiae*	P22336	Component of RPA, a single-stranded DNA-binding heterotrimeric complex	621	cgd7_4620	688	21.4%	9.00E-38	444.5
RFA2	*S. cerevisiae*	P26754	Component of RPA, a single-stranded DNA-binding heterotrimeric complex	273	cgd2_4080	362	17.8%	3.00E-04	105.5
RFA3	*S. cerevisiae*	P26755	Component of RPA, a single-stranded DNA-binding heterotrimeric complex	122	None	N/A	N/A	N/A	N/A
SAE2	*S. cerevisiae*	P46946	DNA double-strand break processing	345	None	N/A	N/A	N/A	N/A
SGS1	*S. cerevisiae*	P35187	ATP-dependent DNA helicase	1447	cgd6_4420	990	26.5%	9.00E-96	977.0
SKI8	*S. cerevisiae*	Q02793	Double-strand break (DSB) formation during meiotic recombination	397	None	N/A	N/A	N/A	N/A
SPO11	*S. cerevisiae*	P23179	Meiotic DNA double-strand break formation	398	cgd8_1350	360	22.3%	3.00E-12	185.5
SRS2	*S. cerevisiae*	P12954	ATP-dependent DNA helicase	1174	cgd8_3210	917	23.6%	1.00E-38	532.5
TOP3	*S. cerevisiae*	P13099	ATP-independent breakage of single-stranded DNA	656	cgd5_1180	814	34.6%	1.00E-96	955.0
XRS2	*S. cerevisiae*	P33301	Regulation of 5'-3' exonuclease degradation of double strand breaks	854	None	N/A	N/A	N/A	N/A

Table 2.1.1. Putative Recombination Proteins in *Cryptosporidium parvum*

Essential mediators of recombination in yeast (*S. cerevisiae*) were reviewed (Krogh & Symington, 2004) and described using the online protein database resource UniProt (http;//www.uniprot.org). Putative orthologs in *C. parvum* were identified and characterised by forward and backward blasting nucleotide (BLASTn) and amino acid (BLASTp) sequences using standalone coding sequence databases of genome data for *C. parvum* strain Iowa II and *S. cerevisiae* strain ATCC 204508 / S288c (GenBank Accession no. GCA_000146045.2 R64 and AAEE00000000.1, respectively). Percent identities (%IDs), E-values, and BLAST scores refer to consensus regions of amino acid coding sequences, and an E-value cut-off of E-04 was used to infer a significantly homologous match. Shaded regions highlight highly-conserved protein sequences (E-value < E-50) indicative of a possible maintenance of annotated function.

2.2 METHODS

(i) Recombination Analysis

Whole genome recombination analyses were performed as an adjunct to findings from the concatenated phylogeny. Based on the phylogenetic characterisation of neutrally-evolving loci, one of the IIc-type *C. parvum* GP60 strains, IIcA5G3j, exhibited a closer genomic proximity to the zoonotic (IIa and IId) WGS than the anthroponotic IIc-type GP60 strains. This warranted investigating the presence, location, and chronology of potential recombination events between three representative whole genome sequences, and how these may have contributed to or influenced the observed genetic disparities. Due to a superior sequence coverage and quality, the WGS lineages selected for extensive recombination analyses were UKP16, UKP6, and UKP13 (IIcA5G3j, IIaA15G2R1, and IIcA5G3a respectively). *C. hominis* was included as an outgroup to enhance recombination signals at the intra-*C. parvum* level, and *C. hominis* strain UKH1 was again chosen due to its superior sequence coverage and assembly quality compared to other *C. hominis* WGS. Including a *C. hominis* strain also served to identify regions of uniquely elevated genetic similarity with UKP13 (IIcA5G3a), as they share a narrow anthroponotic host specificity.

Whole chromosome sequences were constructed, edited, and aligned using the same manual approach described for genomic protein comparisons (Section 1.2). Potential recombination signals due to introgression, or repeated back-crossing with an ancestral strain, were detected and dated by cross-comparing triplet WGS, chromosome-by-chromosome, using the HKY85 nucleotide substitution model (50 bp sliding windows) imbedded in the HybridCheck recombination analysis tool (Ward & Oosterhout, 2015). Recombination signals were verified and expanded to include potential parentage and recombinant fragment breakpoints using RDP4 (Martin *et al*. 2015). Automated detection algorithms RDP, GENECONV, Bootscan, Maxchi, and Chimaera were run with default values, with RDP p-values of $<10^{-5}$ used to infer confidence of recombination signals.

(ii) Divergence dating

Estimating the chronology surrounding major speciation and sub-speciation events is a useful way to understand the potential impact of historical events on driving and influencing host-parasite interactions. The generally high conservation and sequence similarity between *C. hominis* and *C. parvum* genomes, and more pronouncedly the *C. parvum parvum* versus *C. parvum anthroponosum* genomes, indicates that the divergence of these species and subtypes may have occurred relatively recently. This makes it possible that the phenotypic variances observed within human-infective *Cryptosporidium spp.* were affected by some of the major geographic and demographic shifts that accompanied the industrialization and modernization of society over the past millennium. The algorithm HybridCheck implements to date recombination (introgression) events calculates chronology in terms of generation numbers, and is based on prokaryotic propagation models whereby parent to offspring development occurs after a single round of DNA replication (i.e. one generation = one replication). Specifically, it bases divergence estimates on the quantitative presence of mutations in introgressed blocks between two sequences, taking into account a baseline mutation rate (μ) for the given species, as well as different substitution models with distinct base frequency and mutation rate assumptions (Ward & Van Oosterhout, 2015). In this study, the HKY85 substitution model was selected, as it assumes unequal base frequencies between sequences, a relevant prerequisite for the considerably AT-rich *Cryptosporidium* genome (Appendix II). The default mutation rate of 10^{-7} per generation (0.0000001) was retained, as it fell within the conservative estimated mutation accumulation rate observed between two outbreak WGS that were sampled seven days apart (Table 2.2.1).

Table 2.2.1. Whole genome comparison of two outbreak strain WGS reveals estimated mutation accumulation rates per generation for *Cryptosporidium spp*.

Whole genome sequences for UKP4 and UKP6 WGS (excluding telomeric repeats) were constructed in MEGA version 7.0 (Kumar *et al*. 2015) and aligned using the ClustalW algorithm to produce whole genome alignments (WGA). Polymorphisms were subsequently characterised in DnaSP v 5.10.1 (Librado & Rozas, 2009), and mutation rates per generation are estimated based on the assumption that infections were successive cases of the same outbreak isolate.

UKP4 v UKP6 Whole Genome Comparison	
Sampling separation	7 days
No. of sites in WGA (bp)	9086411
No. of SNPs	10
Nucleotide diversity	0.0000011
No. of indel sites	78
No. of indel events	35
Total no. of polymorphisms (SNPs + indel Events)	45
Per base SNP mutation rate per generation (μ)	9.50E-08
Per base indel rate per generation (μ)	3.32E-07
Combined mutation rate per generation (μ)	4.27E-07

Applying this method to a protozoan like *Cryptosporidium*, where multiple rounds of nuclear division occur in a single life cycle, therefore meant it would be important first to characterise and quantify asexual versus sexual stages, and then assign tangible units (seconds, minutes, or hours) for the parental (oocyst) to offspring (oocyst) generation time. One common misconception surrounding *Cryptosporidium* reproduction is that one oocyst (or four sporozoites) into the host, equals four oocysts (or 16 sporozoites) out of the host, but this does not take into account that there are multiple rounds of DNA replication that take place within a single cycle. Breaking down the individual nuclear division stages, specifically outlining the input and output of sexual and asexual replication cycles, generated a much better idea of the numbers involved (Figure 2.2.1) and was crucial for calculating a reliable estimate of generation time in *Cryptosporidium*.

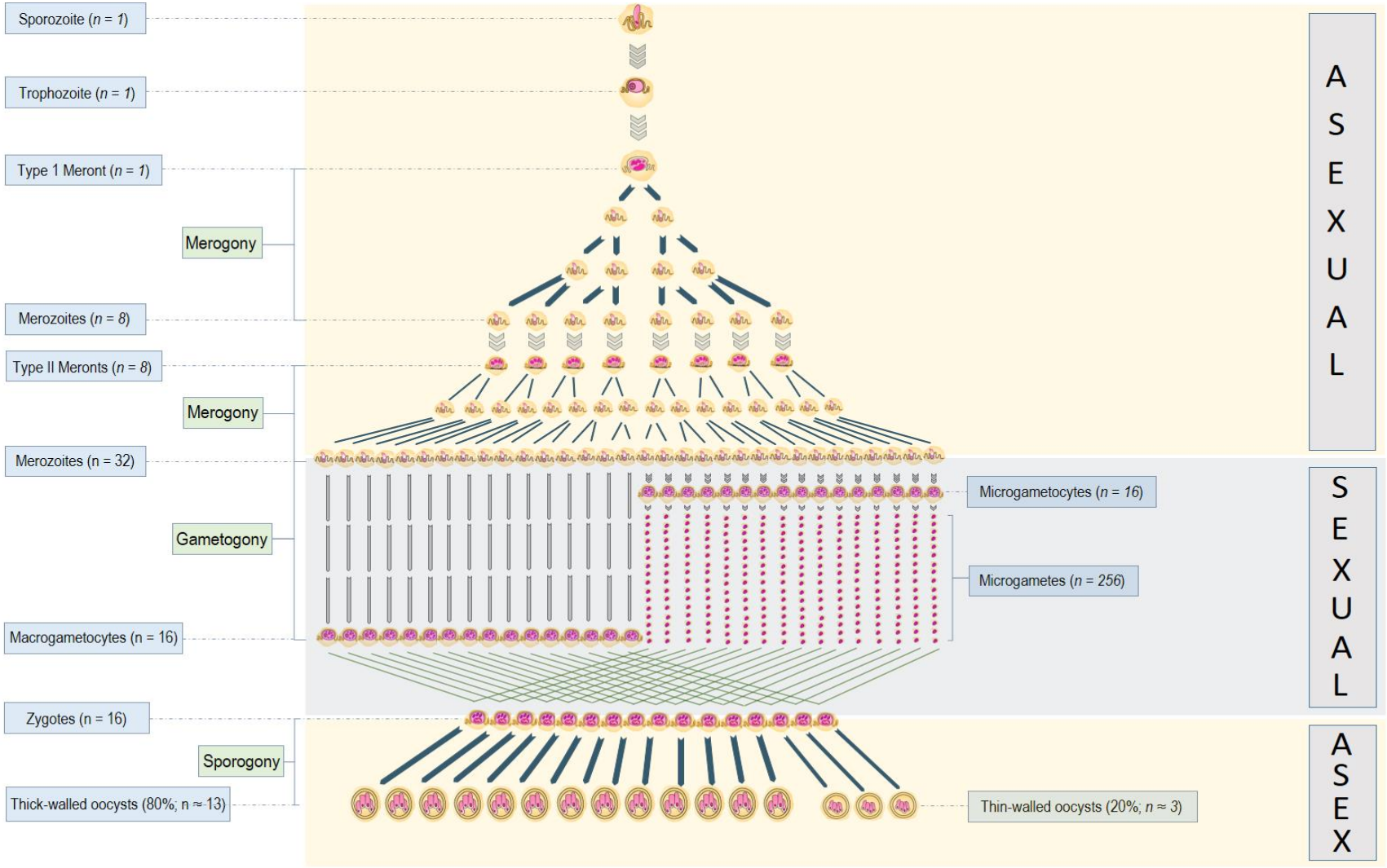


Figure 2.2.1. Life forms and population expansion of *Cryptosporidium spp*. during a single in-host replication cycle

Schematic quantitative representation of the sexual and asexual stages of parasite propagation during in-host replication. Offspring resulting from nuclear division and fertilization events at each life cycle stage are estimated relative to a single sporozoite, and so represent one-fourth of the overall quantities that would result from an ingested oocyst. Adapted from Kosek et al. 2001, Tzipori & Ward, 2002, and Leitch & He, 2012.

The method used to assign units of time to each generation of oocysts was based on a back-extrapolation of "oocysts out" versus "oocysts in", taking into account challenge dose, duration, and total excretion from past infectivity studies, as well as the proportional relationship between offspring and parental oocysts. Important to calculating this proportion was initially an understanding of the sex ratio between gametocytes produced during gametogony, as an unequal rate directly influences the number of progeny that are produced. A single published source makes reference to this rate in *Cryptosporidium*, estimating the proportion of male gametocytes at 0.5 (West *et al*. 2000), or a 1:1 ratio between males (microgametocytes) and females (macrogametocytes). Similar rates have been observed in a number of other intestinal protozoa, including *Eimeria* (Chauve *et al*. 1994), *Schellackia* (Paperna & Finkelman, 1996), and *Toxoplasma* (Omata *et al*. 1997). Another important consideration was the number of oocysts that are excreted (thick-walled) versus recycled (thin-walled) within the host. This number has previously been estimated at around 80% and 20%, respectively (Ridley & Olsen, 1991). As only 20% of new oocysts produced are involved in autoinfection, propagation rates needed to be adjusted for each subsequent cycling calculation accordingly. The quantitative life cycle figure highlights this, demonstrating how a single sporozoite can produce up to 13 thick-walled and 3 thin-walled oocysts (Figure 2.2.1). Hence, one ingested oocyst (4 sporozoites) can result in 52 excreted oocysts and 12 autoinfective ones, and these 12 autoinfective oocysts can in turn result in 624 excreted oocysts and 144 autoinfective ones, which in turn can result in 7,488 excreted and 1,728 autoinfective oocysts, and so on.

Table 2.2.2. Oocyst infectivity and intensity rates in human volunteers summarized from peer-reviewed publications

Study	Challenge organism	Challenge dose	Onset of Excretion (days)	Duration of Excretion (days)	Total no. of oocysts excreted	Estimated no. of oocyst generations	Estimated no. of days/generation
Okhuysen *et al*. 2002	*C. parvum*	100	7.5	3.5	1.8×10^6	3-4	2-4
	C. parvum	300	5	3	3.5×10^6	3-4	2-3
	C. parvum	1,000	4	11	3.1×10^8	4-5	3-4
	C. parvum	3,000	5	6	2.1×10^7	~3	3-4
Chappell *et al*. 2011	*C. meleagridis*	10,000	8	3	4.5×10^8	~3	3-4

In light of these biological numerical considerations, a factor of 12 autoinfective offspring per 1 parental oocyst was used to estimate the rate of expansion of oocyst populations in vivo. Using this value in correlation with excretion data from past infectivity studies revealed a population expansion of 3-5 new generations, and an estimated life cycle duration of 2-4 days, or 48-96 hours, per infection (Table 2.2.2). Although this estimate deviates significantly from the frequently-reported generation time of 12-14 hours for *Cryptosporidium* (Upton, 2008; Oram, 2014; Fleming 2015), not a single one of these resources makes reference to the source of this estimation, which makes it difficult to verify or investigate how these numbers were achieved. In addition, an early cell culture experiment of *Cryptosporidium* development in human and animal cell lines showed that the sexual forms (macro- and microgametes) did not develop until 48 hours after inoculation, and that oocysts only appeared after 72 hours (Current & Haynes, 1984). These estimates nicely complement the 48 to 96-hour generation estimate calculated in this study, lending support to the use of this range to generate estimated divergence dates for human-infective *Cryptosporidium spp*.

2.3 RESULTS

(i) Recombination events are associated with host specialization in closely-related *Cryptosporidium spp.*

The role of recombination in driving phenotypic diversity was investigated through whole genome alignments and comparisons across three *C. parvum* (IIaA15G2R1, IIcA5G3j, IIcA5G3a) and a single *C. hominis* (IbA10G2) GP60 subtype strain. A total of 104 unique recombination events were detected across the whole genomes compared through the RDP4 recombination analysis software (Appendix IV), confidently mirrored by the complementary recombination signal analysis performed using HybridCheck (Figure 2.3.2). Based on observed outliers from the 5-95% confidence intervals encompassing the expected mean number of recombination events between *C. p. parvum* and *C. p. anthroponosum*, the distribution of recombination events across chromosomes was skewed, with a disproportionately higher number of individual events estimated for chromosome 6 (25.9% of total events), and a disproportionately lower number of events calculated for chromosomes 3, 5, and 7 (Figure 2.3.1). Similarly, an unequal distribution of recombination events across subtype strains was identified, with a greater number of events observed for *C. parvum anthroponosum* subtype IIcA5G3a than for any other strain (37.5%). Another prominent feature was a distinct clustering of recombination events within 20,000bp of putative virulence factors. The working definition of virulence here encompasses subtelomeric gene families, hypervariable low-complexity proteins and large secreted glycoprotein-type proteins, particularly those associated with significant dN/dS values (>1.0), or any combination of the aforementioned features. Accordingly, more than 90.0% of all recombination events occurred proximal to loci suspected to drive virulence or play a major role in host-parasite interactions in human cryptosporidiosis. From this, a working hypothesis can be derived that the occurrence of recombination in *Cryptosporidium spp.*, although infrequent, largely results in novel gene configurations and chromosome arrangements that directly benefit the parasite's ability to adapt and survive within a novel host environment.

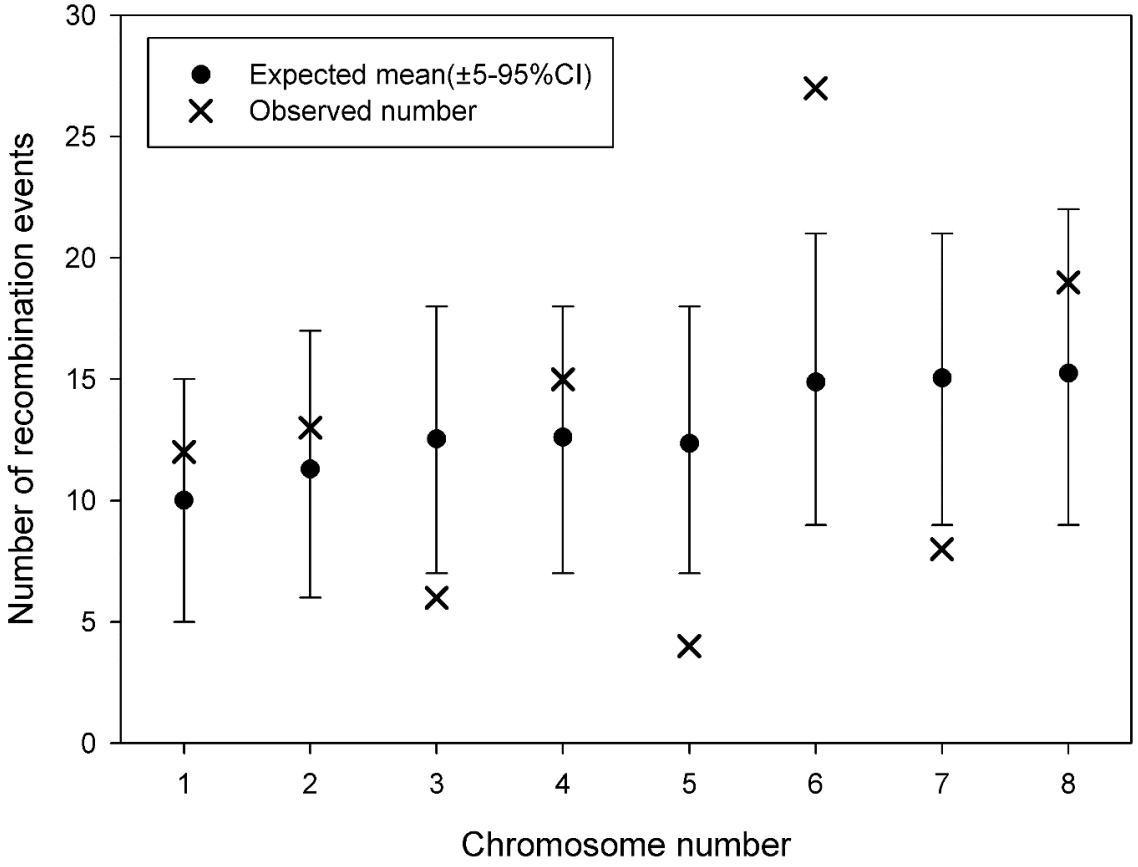


Figure 2.3.1. 5-95% Confidence Intervals for Mean no. of Expected Recombination Events per Chromosome between *C. p. parvum* and *C. p. anthroponosum* WGS

Binomial distribution of expected and observed number of recombination events for each chromosome is shown, based on the number of recombination events identified in *C. p. parvum* (IIaA15G2R1 and IIcA5G3j) and *C. p. anthroponosum* (IIcA5G3a) WGS, with 5-95% confidence intervals indicated by error bars around the expected mean number of events. Statistically significant outliers are indicated by datapoints that fall outside of the error bars. All calculations were performed using Microscoft Excel (2013), and the figure was created using SigmaPlot (v 13.0; Systat Software, San Jose, CA).

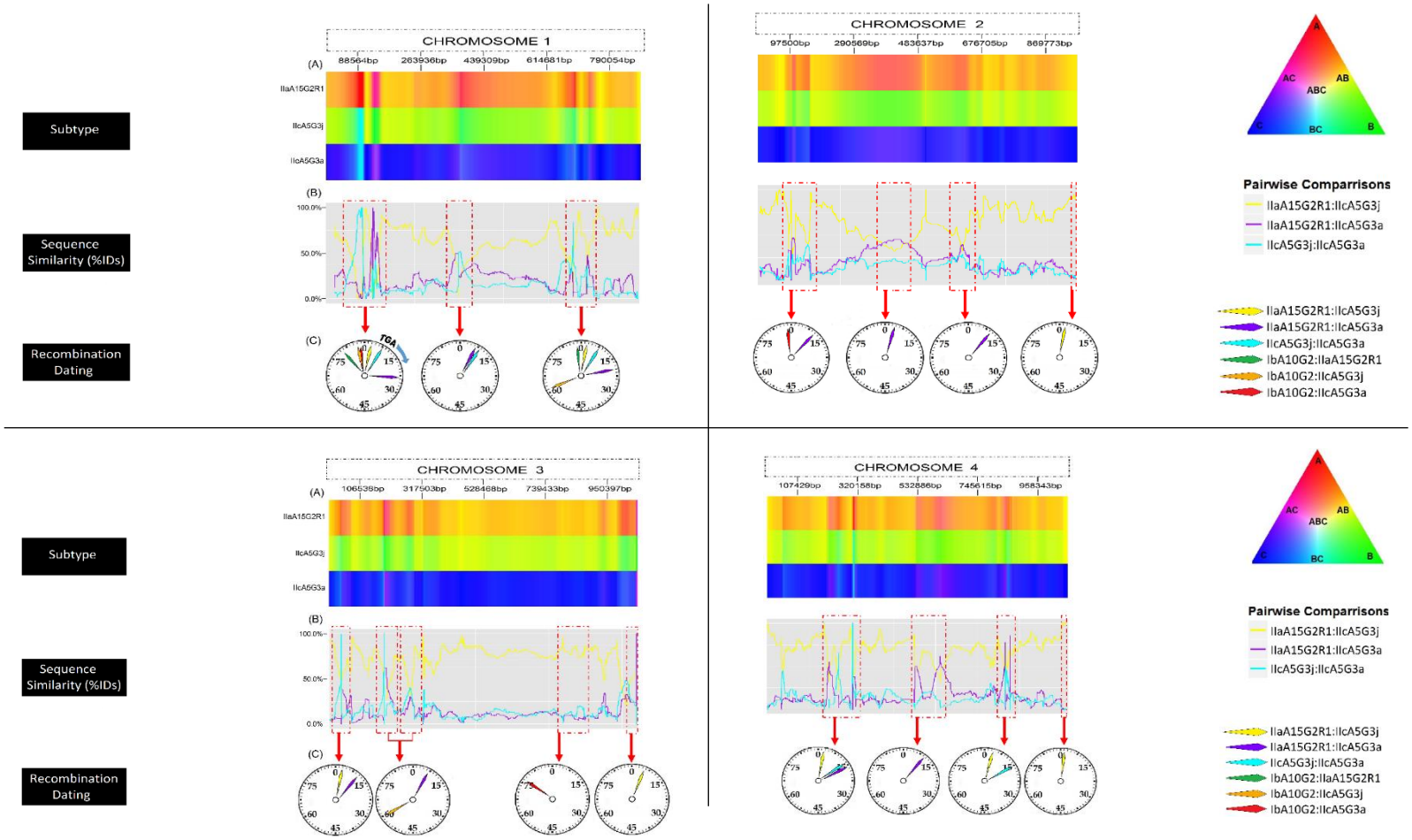


Figure 2.3.2. Recombination signal analysis reveals localization and chronology of major recombination events between phenotypically diverse *Cryptosporidium* subtypes, chromosomes 1-4

Whole chromosome sequences for *C. parvum* GP60 subtypes IIaA15G2R1, IIcA5G3j, and IIcA5G3a, and *C. hominis* GP60 subtype IbA10G2, were aligned and compared in four three-way comparisons using HybridCheck recombination analysis tool (Ward & Van Oosterhout 2015). The detected recombination signals for the intra-*C. parvum* comparison are represented here. Red-green-blue colours represents subtype-unique sequence polymorphisms, with shared polymorphisms between GP60 subtypes highlighted by shaded regions in yellow, purple, and turquoise (A). Recombination events are subsequently described in the form of sequence similarity (% IDs) based on two-way comparisons across all three subtypes (B), with introgression blocks clearly indicated by red arrows. Divergence dates were estimated using the HKY85 mutation model imbedded within HybridCheck, and are represented graphically for all twelve recombination analyses, in terms of thousands of generations ago (C).

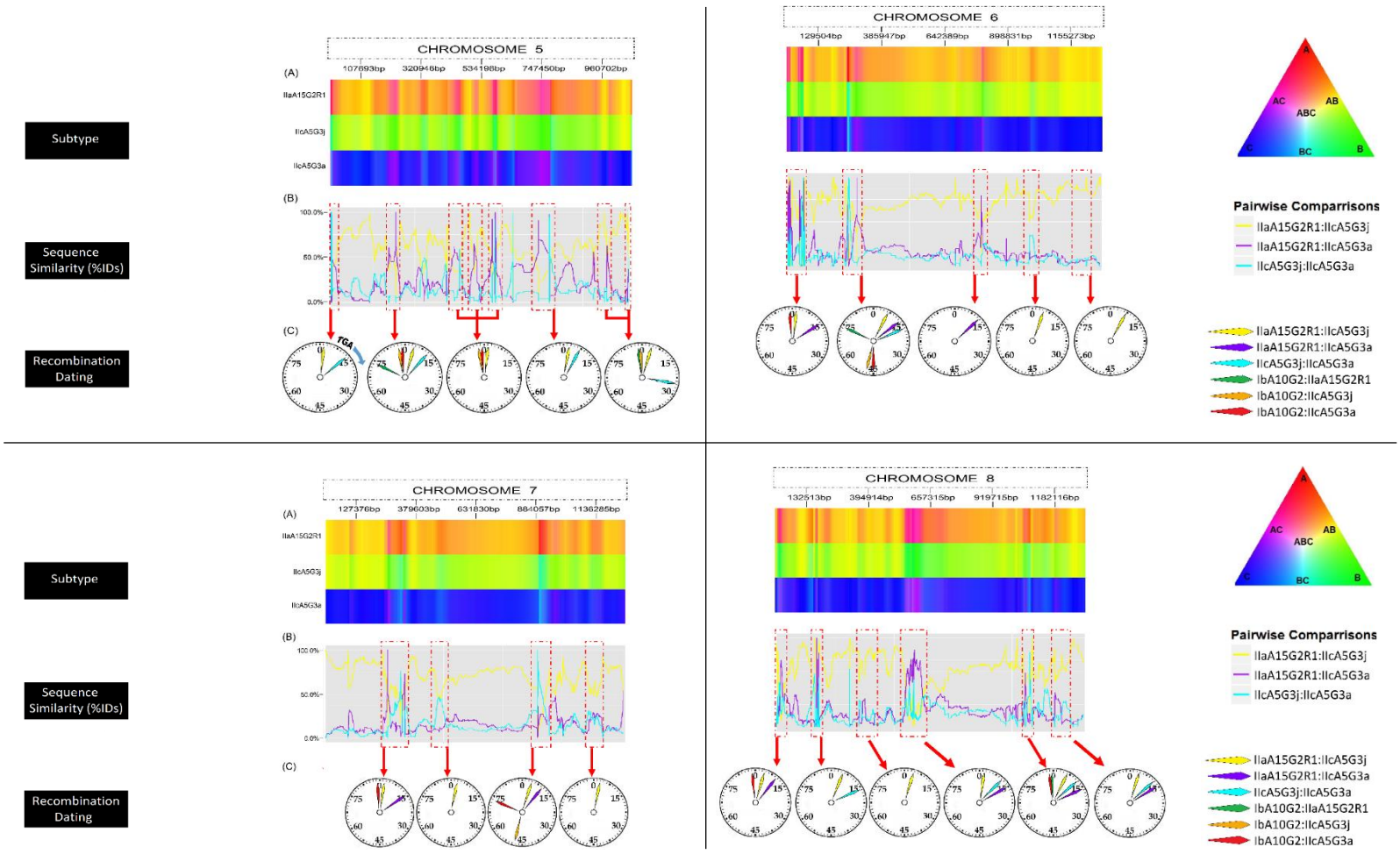


Figure 2.3.2. Recombination signal analysis reveals localization and chronology of major recombination events between phenotypically diverse *Cryptosporidium* subtypes, chromosomes 5-8

The consequential nature of recombination was further explored in this study by phylogenetic reconstruction of highly significant (E-value < E-40) recombination events (Appendix V). These were further summarized into a proportional break-down of the phylogeny-types represented (Figure 2.3.3). By comparing the resulting phylogenies to the concatenated phylogeny of neutrally-evolving loci (Figure 2.3.3. A), it was possible to evaluate how site-specific recombination impacted genotypic sequence integrity in these strains. This detailed phylogenetic analysis revealed that a large proportion of major recombination events in these strains result in a "*C. parvum anthroponosum*-outgroup" type topology (45.5% of highly significant events), with *C. parvum parvum* strains IIaA15G2R1 and IIcA5G3j forming a separate clade at the top and *C. hominis* subtype IbA10G2 at an intermediate position (Figure 2.3.3. C). This describes an excess of recombination between *C. p. anthroponosum* and a yet undescribed divergent *Cryptosporidium* WGS. That is to say, the resulting phylogenies of highly significant recombination events complement the RDP4 findings that *C. parvum anthroponosum* contains the greatest proportion of recombinant fragments amongst the strains compared, but that the origin of this genetic exchange cannot be traced to any of the three remaining WGS. Although the potential exchange of genetic material between *C. p. anthroponosum* and an unknown *Cryptosporidium* WGS might explain the excess of this phylogeny-type, it is important to also discuss the possibility that certain *C. p. anthroponosum* alleles may have failed to coalesce with their orthologous counterparts in other anthroponotic *Cryptosporidium spp*. Genetic loci are further known to be subject to variable evolutionary rates, and an excess or lack of gene flow between isolated populations may have impacted the rate and nature of allelic divergence in some *Cryptosporidium spp*. populations more so than others. Fluctuations in the effective population sizes of certain species and subspecies through time could have further introduced a bias of allelic divergence at certain loci. It is also possible that a broader network of anthroponotic *C. parvum* subspecies exists than is currently known, with multiple lineages that can be traced to a number of subspeciation events, and thereby different most recent common ancestral gene sequences (Nichols, 2001). It is therefore important to note that a number of external factors and processes beyond population genetics alone could account for the observed *C. p. anthroponosum*-specific divergence and the resulting unexpected phylogenies at certain loci.

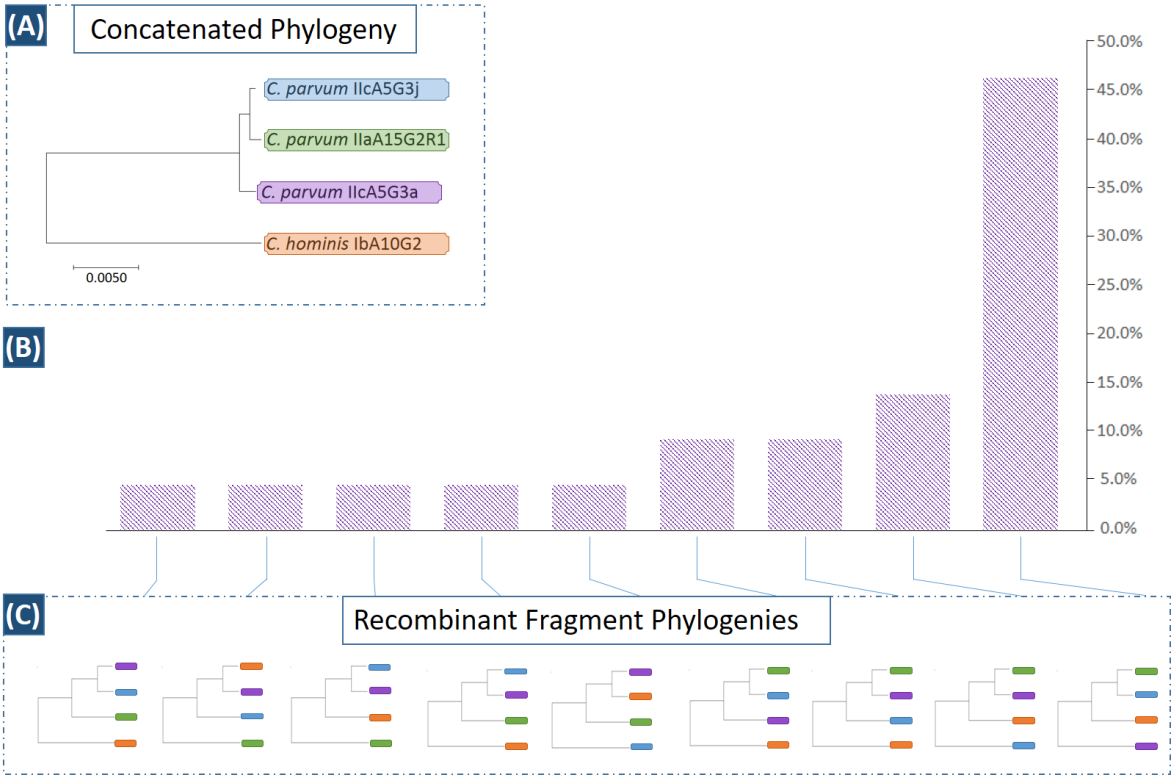


Figure 2.3.3. Relative distribution of recombinant fragment phylogeny-types amongst highly significant recombination events
Phylogenetic reconstructions of highly significant recombination events (E-value < E-40; Appendix V) were organised based on the type of phylogenetic grouping and ordering of strains they represented (C). Sequence data extracted from WGS for *C. parvum parvum* IIcA53j, *C. parvum parvum* IIaA15G2R1, *C. parvum anthroponosum* IIcA5G3a, and *C. hominis* IbA10G2 are represented by blue, green, purple, and orange bars respectively. Phylogeny-types were quantitatively broken down, with percentages representing the proportion each type comprised of the total 22 recombinant fragment phylogenies (B). The concatenated phylogeny for the four included *Cryptosporidium* strains is additionally shown for comparative purposes (A). Note: Trees do not represent accurate branching or nucleotide substitution rates; only a generalised pattern of strain positioning is intended.

(ii) Species diversification is a modern phenomenon in human-infective *Cryptosporidium spp.*

The estimated divergence dates calculated in this study provide the first chronological description for genetic introgression between human-infective *Cryptosporidium spp.* (Figure 2.3.2.C). HybridCheck divergence time analyses revealed a chromosome-wide consensus that places a majority of introgression events between *C. parvum parvum* (IIaA15G2R1 and IIcA5G3j) and *C. parvum anthroponosum* (IIcA5G3a) strains at approximately 10-15 thousand generations ago (TGA), while genetic recombination between the two more closely related *C. parvum parvum* appears to be more recent (~3-5 TGA). Recombinant fragments in *C. hominis* that could have originated from any of the *C. parvum* strains analysed appear to be rare (~6.8% of all recombination events measured),

and as expected occurred more anciently (i.e. ~75-100 TGA) than the intra-*C. parvum* events (Fig. 2.3.2C). Translating generation periodicity into tangible units of time is a difficult task for *Cryptosporidium*, due to its lengthy environmental cyst stage and the lack of existing human transmission studies. Based on transmission rates in healthy adult volunteers, however, this study determined a likely range for generation time to be between 48 and 96 hours (Table 2.2.1). In the absence of experimental data on the environmental versus host transmission rates and durations for *Cryptosporidium*, this study chose to present divergence dating estimates based on the assumption that there is a steady rate of transmission for recombining lineages within a host population *only*. This dates divergence to circa 16-55 years ago for zoonotic *C. parvum* strains IIaA15G2R1 and IIcA5G3j, 55-164 years ago for zoonotic vs. anthroponotic *C. parvum* strains, and 410-1096 years ago for the *C. hominis* vs. *C. parvum* major divergence events (Figure 2.3.4). This estimates that the evolutionary split between the four human-infective *Cryptosporidium spp.* included in this study all occurred quite recently, within or at the dawn of the past millennium.

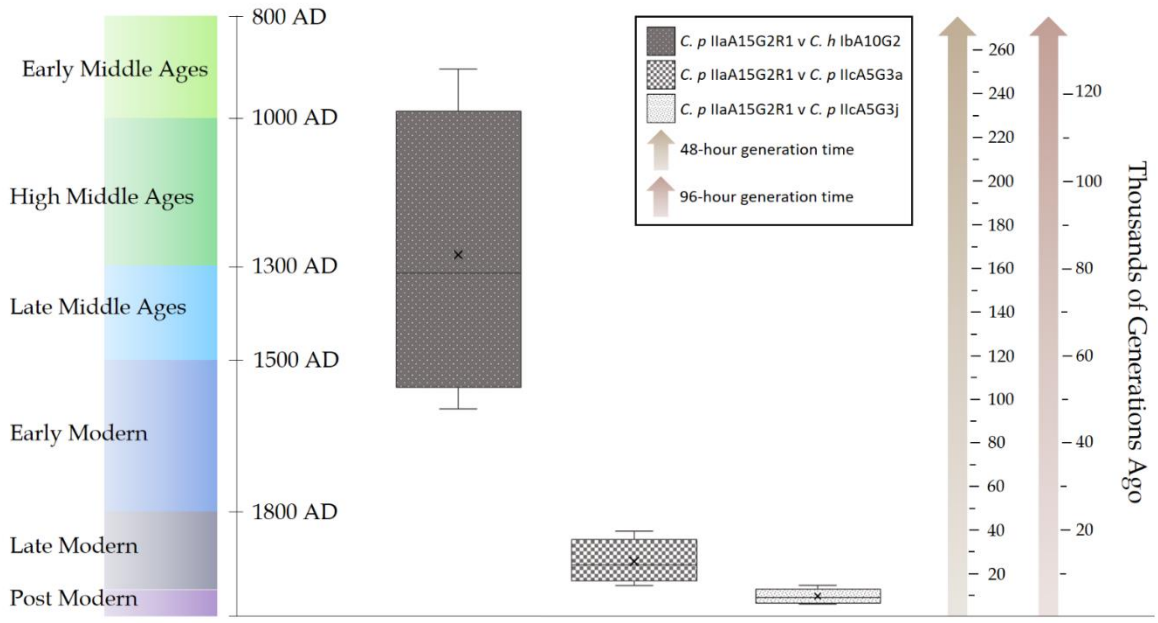


Figure 2.3.4. Divergence dating estimates for three human-infective *Cryptosporidium spp.* WGS comparisons

The range of estimated divergence dates (thousands of generations ago) between the *C. parvum* reference genome (IIaA15G2R1) and three other human-infective *Cryptosporidium* WGS (*C. parvum* IIcA5G3a, IIcA5G3j, and *C. hominis* IbA10G2) is shown, as calculated according to the HKY85 substitution model (HybridCheck; Ward *et al.*, 2015). Minimum, mean, and maximum generation numbers were converted into units of time (years) for both 48- and 96-hour life cycle estimates, and are represented here by box-and-whisker plots. Estimates are portrayed in terms of maximum (upper whisker) and minimum (lower whisker) values, with horizontal box lines indicating (from top to bottom) upper quartiles, medians, and lower quartiles. X-symbols represent mean values.

2.4 DISCUSSION

The role of recombination in driving phenotypic disparities between human-infective *Cryptosporidium* strains has long been central to understanding pathogenicity in cryptosporidiosis, but only a limited number of legitimate examples and explanations have been provided to date. The current study provides the first extensive insight into the **What**, **When**, **Where**, and **Why** surrounding recombination in *Cryptosporidium*: **What** strains appear to recombine with one another, **When** these recombination events are estimated to have taken place, **Where** in the genome these events take their effect, and **Why** these changes appear to have had an evolutionary impact and potential benefit in these species. The discovery that all four of the included GP60-subtype strains appear to have recombined with one another at some point in history, taking into consideration that dozens of further GP60 human subtypes exist, indicates a potentially massive implication of recombination in driving genetic diversity within these species. The overwhelming bias towards recombination in *C. parvum* subtype IIcA5G3a, the most frequent anthroponotic *C. parvum* subtype identified in human cases, further provides a strong case for the direct role of recombination in producing a host-specialised lineage within an otherwise predominantly zoonotic species. The high frequency and distribution of these non-zoonotic type recombinant fragments in *C. parvum* IIcA5G3a also provides genetic verification for the classification of this strain as a separate *C. parvum* subspecies (Figure 2.3.3). Likewise, the presence of recombination in *C. parvum* subtype IIcA5G3j, although significantly less frequent, helps to provide a genotypic basis for the predominantly anthroponotic and sporadically zoonotic (hedgehogs) nature of associated infections.

Recombination has long been viewed as one of the most effective evolutionary mechanisms at influencing virulence and host specificity within parasitic organisms. The discovery that more than 90.0% of recombination events in the analysed *Cryptosporidium* WGS appear to impact regions within close proximity to putative virulence factors is a strong indication of the functional significance of recombination - aimed at enhancing survival within the host environment and increasing the pathogenic impact. This direct incentive is mirrored in the pathogenically-savvy malaria and trypanosome parasites, where recombination has been shown to affect both host range (Rathore *et al*. 2003) and

pathogenicity (Ferreira *et al*. 2004; Kyes *et al*. 2007; Horn & McCulloch, 2010; McCulloch *et al*. 2015). The higher frequency of recombination within 25,000bp of telomere ends, accounting for around half (49 of 104 events, ≈ 47.1%) of the identified occurrences (Appendix IV), is also a common occurrence in both *Plasmodium* and *Trypanosoma*, primarily because this is where many of their most important virulence factors, such as the famous antigenic variation genes, are situated (Pays *et al*. 1985; Freitas-Junior *et al*. 2000). As the underlying genetic factors for virulence in *Cryptosporidium* are largely still unknown and uncharacterised, the parallels that can be drawn with recombination traits in other protozoan parasites form a strong basis for future exploration of virulence and pathogenicity in this genus. This study thus provides a rationale for focusing the search on recombination-affected loci, particularly at the subtelomeres, and perhaps those that additionally exist in multi-variant forms.

The result that perhaps bears the greatest novelty in this study is the unique approach used to estimate *Cryptosporidium* generation times, as well as the chronological characterisation of speciation events through divergence dating. No other study could be found that provides in detail: A.) a quantitative breakdown of the *Cryptosporidium* life cycle, or B.) an estimation of how long ago some of the main human-infective species and subspecies splits are likely to have occurred. If accurate, this study places divergence within a relatively modern context, estimating that divergence from the most recent common ancestor occurred sometime within the past 1,000 years between and within human-infective *Cryptosporidium spp*. This timescale is more recent but comparable to lineage-specific divergence estimates in *Plasmodium falciparum* (Figure 2.4.1), another apicomplexan parasite, which has a similar life cycle duration of ~48 hours within the host. It does, however, significantly contrast from estimates for *Toxoplasma gondii* and *Babesia microti*, both of which are estimated to have diverged ~10,000 years ago at the intra-species level. However, their complex life cycle, involving multiple host and environmental settings, points to a significantly longer (weeks, rather than days) life cycle duration. This could explain the significant disparity in divergence dating between the two. The putative timeframes for major recombination events in this study also revealed a potential link between the evolution of human-infective *Cryptosporidium spp.* and historical changes in human behaviour. This type of real life co-evolution of host and parasite is rarely witnessed in nature, and it demonstrates the potential impact that

demographic and socioeconomic change can have on restructuring pathogenic phenotypic diversity. The chronological assessment of recombination events indicates that the main subspeciation event between *C. parvum* IIaA15G2R1 (zoonotic) and *C. parvum* IIcA5G3a (anthroponotic) took place around the 1800s, a time during which developed societies were constantly evolving and adapting in order to keep up with rapid changes in industrialization and modernization. The massive influx from agricultural and rural communities to the overcrowded and less sanitary urban cities could conceivably have forced *Cryptosporidium* into a new anthroponotic lifestyle, and a single historical switch from broad to narrow-host-range could explain why divergence is primarily limited to a subset of loci in these species, and why the remaining global nucleotide consensus has remained so high. Although preliminary, these results provide the first time-linked theory for how *C. hominis* and *C. parvum* once came to be. Expanded investigation of recombination in more diverse *Cryptosporidium* WGS will help to further characterise the type, rate, and biological impact of recombination and divergence within human-infective species, and thereby continue to increase our understanding of the way strains and species evolve and adapt within this neglected protozoan genus.

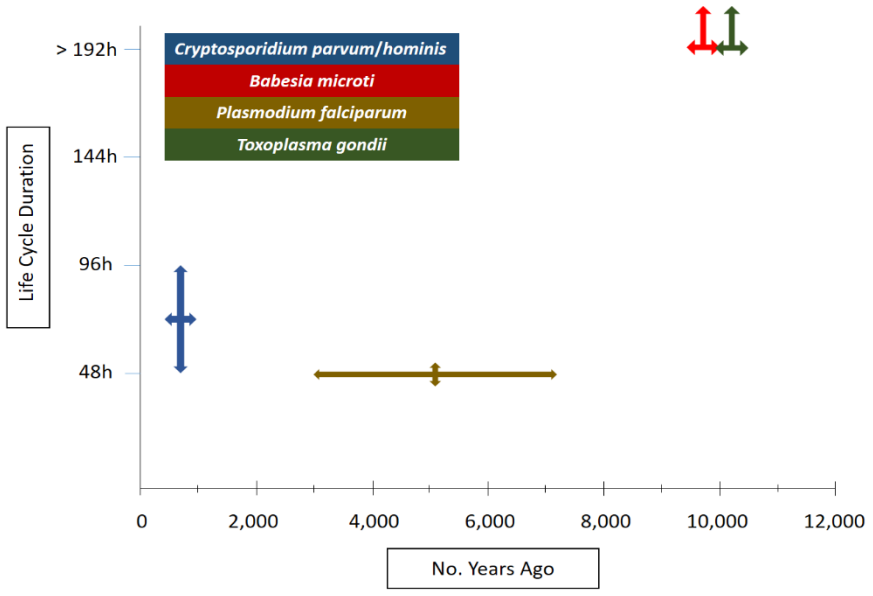


Figure 2.4.1. Estimated dates of divergence for the most recent common ancestor in human protozoan parasites

Divergence dates relating to the most recent common ancestor for closely-related lineages within three other human protozoa (*Plasmodium falciparum*, *Babesia microti*, and *Toxoplasma gondii*) were extracted through PubMed literature searches (https://www.ncbi.nlm.nih.gov/pubmed), and correlated with reported life cycle durations (i.e. generation time between infectious life forms). These were reported at 3,000-7,000 years ago for extant forms of *Plasmodium falciparum*, and a life cycle duration of 48 hours (Volkman *et al*. 2001; de Koning-Ward *et al*. 2015), respectively. In *Babesia microti*, a divergence time of ~10,000 years has been estimated for two geographically-diverse strains of a limited <2% genomic diversity (Lemieux *et al*. 2015). The life cycle for *Babesia* has not been assigned time units, but is estimated at a minimum of one to several weeks, as it is complex and involves two different hosts (Homer *et al*. 2000). *Toxoplasma gondii* was also estimated at exhibiting an estimated divergence time of around 10,000 years between three common clonal lineages (Su *et al*. 2003), and like *Babesia* has complex multi-host and multi-environment requirements for life cycle completion; hence the life cycle was likewise given a minimum but not maximum estimated duration (Weiss & Kim, 2007).

Strain evolution and diversification of the GP60 locus during the global expansion of *Cryptosporidium hominis*

3.1 INTRODUCTION

Cryptosporidium hominis is the leading cause of human cryptosporidiosis in many parts of the world. Although ten unique GP60 subtype families have been classified to date (Ia, Ib, Id, Ie, If, Ig, Ih, Ii, Ij and Ik), very little is known about the genetic diversity that defines and differentiates these subtypes, and comprehensive data on associated distribution and host range patterns remain limited. This chapter has attempted to tackle this gap from a combined epidemiological and population genetics perspective, utilizing systematic review of global GP60 sequence submissions to describe diversity between and within subtype families. Segregating sequences into subtypes and geographic regions provided a unique opportunity to understand how selective pressure and diversity define divergence across subtype families and regions, and genetic differentiation analyses aimed to illustrate whether any evidence of recent population expansion or contraction events exists. Allele frequencies and diversities worked to define how well GP60 allele populations are structured across *C. hominis* subtypes, and assess whether certain allele types significantly dominate human infections. The potential overlap between alleles and regions (i.e. phylogeography) was used to further examine the potential role of genetic migration/interbreeding, and the potential impact of host behaviour on the diversification of certain subtype families was discussed. By exploring and analysing population genetics across/between geographic regions and subtype families, and correlating genetic data with epidemiological features, this chapter provides a new understanding of the processes that have accompanied strain evolution and diversification of GP60 within *C. hominis*.

(i) *Cryptosporidium hominis* – The origin of a new species

Cryptosporidium parvum as a universal descriptor of the human *Cryptosporidium* pathogen first came into use in the mid-1980s (Upton & Current, 1985), when it was characterised morphologically as being significantly smaller than the original mouse species discovered in 1907, *Cryptosporidium muris*. The derivation of the genus name is based on a combination of descriptive Latin terms: 'kruptos', the Latin word for hidden, and 'spora', meaning spore. The species name 'parvus' refers to its small size, although *C. parvum* has since been absolved of its miniature reputation, ranking as the 11th largest out of 18 oocyst species in a recent review (Australian Society for Parasitology, 2010). As molecular advancements towards genotyping began to emerge throughout the 90's, distinct polymorphisms within common targets such as the 18S ribosomal RNA gene (Kilani & Wenman, 1994) and non-telomeric repetitive sequences (Bonnin *et al*. 1996) were being identified, leading to the re-classification of *C. parvum* into two genotypes: type I (human), and type II (animal), correlating with their respective transmission cycles (Peng *et al*. 1997). Extended multilocus analyses and in vitro growth differences eventually added to their divergent phenotypic profiles, widening the evolutionary and biological gap of these two genotypes, and by 2002 *C. hominis* the species was finally recognised (Morgan-Ryan *et al*. 2002).

Historically, the first attempts at characterising differential epidemiological features of human *Cryptosporidium spp*. were predominantly focused on *C. hominis* versus *C. parvum*, largely because the early genetic markers only provided limited discriminatory power between subtypes. After the introduction of GP60-based subtyping, however, the absence of clonality within these species quickly became evident, and in recent years the available information on subtype distribution and virulence has rapidly advanced (Putignani & Menichella, 2010). To date, *C. hominis* has been identified as comprising ten unique GP60 subtype families, Ia, Ib, Id, Ie, If, Ig, Ih, Ii, Ij, and Ik, and has been reported as being the more frequently isolated human pathogen in the Americas, Australia in Africa (Caccio *et al*. 2005). Recently, an expansive study of more than 10,000 children under 24 months of age across Sub-Saharan Africa and South Asia, where the clinical consequences of diarrhoeal disease are often most severe, found that more than three-quarters (>75%) of Cryptosporidiosis associated with mild-to-severe diarrhoea was caused by *Cryptosporidium hominis* (Sow *et al*. 2016). A distribution pattern of the constituent

subtypes themselves has also begun to emerge – although these studies are still too sporadic to draw actual geographic boundaries or measure predominance of subtypes on a global level. No exhaustive review currently exists that attempts to present globally-reported *C. hominis* data in a combined fashion. The most recent and most comprehensive summary put forth to date provides a useful overview of *C. hominis* burden and distribution from a public health perspective (Caccio & Chalmers, 2016). However, this study included data from only 14 European countries (less than one-third of the 51 European countries in existence), did not include subtype data for non-human sources of *C. hominis* infection, and refrained from elaborating on the sequence variability between the reported and submitted GP60 sequences to date, nor the potential genetic factors responsibility for this variability. Compiling all of the available data into a summarised format is thus an important first step towards better understanding global distributions and host ranges associated with variable *C. hominis* GP60 subtype families. Through this, we get closer to being able to establishing if, how, and why certain *C. hominis* GP60 subtype families inflect a greater degree and severity of cryptosporidiosis on human populations worldwide.

(ii) Population genetics at the GP60 locus - what we already know

A number of studies are currently available that describe and discuss GP60 allelic diversity between and within human-infective *Cryptosporidium spp*. populations. One of the earliest reports on GP60 population genetics focused on *C. parvum* isolates from an area of Scotland, and used GP60 in combination with a number of other markers to determine the most common multilocus genotypes (MLGs) infecting humans, sheep, and cattle (Morrison *et al*. 2008). Analysis of Nei's genetic distance and F_{st} found that geographic sub-structuring was largely absent, but genetic sub-structuring indicated that a limited number of MLGs were genetically-isolated from the predominant genotypes, and thereby illustrated the potential for genetically-distinct populations across the areas compared. Another early UK-based study found that less common GP60 subtypes were associated with foreign travel, and that limited heterogeneity could be found for regional (originating in the UK or Europe) isolates, with a vast majority (>91.0%) of these samples belonging to GP60 subtype IbA10G2 (Chalmers *et al*. 2008). The existence of specific subpopulations that are associated with host types and geographic regions was also

reported in an Italian study focused on *C. parvum* in livestock (Drumo *et al*. 2012). A different multilocus study on *C. parvum* isolates from Irish calves reported significantly higher allelic diversity for GP60 compared to a number of other genotyping loci, with regional differences in diversity observed, but found an overall panmictic population structure rather than distinct subpopulations within *Cryptosporidium parvum* (De Waele *et al*. 2013). This lack of genetic sub-structuring appears to reflect random mating within the population, and differs from the distinct and isolated sub-populations described for Scottish and Italian *C. parvum* populations. A similar observation of panmictic population structure within *C. parvum* was identified for human and bovine isolates in the Upper Midwest of the United States, where an MSLT approach including GP60 found little evidence to support sub-structuring due to host or geographical associations (Herges *et al*. 2012). The nature and degree of selective pressure at the GP60 locus was further evaluated for human samples originating in Santiago, Spain, revealing that balancing selection and an excess of synonymous mutations likely drive allelic divergence, and that similar degrees of genetic differentiation separate GP60 subtypes within *C. hominis* and *C. parvum* populations (Abal-Fabeiro *et al*. 2013). Balancing selection at GP60 was also more recently identified for human samples collected from New Zealand between 2010 and 2015 (Garcia-R & Hayman, 2017). This study also found evidence of a recent population expansion in certain sample groups, an observation which was complemented by increased cryptosporidiosis notification rates in the same areas and timeframes. This demonstrates how GP60 could be used to monitor and predict demographic or geographic shifts in human infection rates, or to evaluate how sudden changes in selective pressure may indicate modifications to host-parasite immune interactions and infectivity. Despite this progress, access to studies that have attempted to extensively describe the nature of selection and divergence at the GP60 locus in the context of *Cryptosporidium* population diversity and expansion remains limited, and the data that does exist can at times seem inconsistent and contradictory. A majority of studies tend to be somewhat geographically limited, so as to prevent estimation of the global spread of certain allele-types or compare population genetics between geographically-isolated regions. Divergence is also often analysed between sequences samples within a finite timeframe rather than monitored long-term, making it difficult to identify traits of divergence that may correlate with changes in transmission patterns or pathogenicity. This means population genetics studies need to undergo an expansion, comparing and

contrasting isolates across a wider range of geographic locations. This should generate a better picture of the genetic mechanisms driving divergence at the GP60 locus on a more global scale, thereby identifying certain populations that may be defined by distinctive or unique signals of selection or differentiation in a way that better explains the spread or physical manifestations of disease in human cryptosporidiosis.

C. hominis continues to cause outbreaks across the world, despite regular decontamination attempts through means such as filtration and UV treatment of water supplies, and continues to be implicated as one of the primary causes of diarrhoea-associated morbidity and mortality in much of the developing world (Sow *et al*. 2016). Hence, it is crucial to understand the true nature of this parasitic species epidemiologically, expand on this to characterise the genetic differences driving divergence for host-parasite interactors, and combine these features to form steadfast links between clinical outcomes and specific allele-types. Not only will this help to establish a validated and reliable genotyping approach for *C. hominis* isolates worldwide, but it will help pave the way towards virulence factor elucidation and possibly identify new cellular targets for treatment and prevention. This chapter aims to describe GP60 from both an epidemiological and population genetics perspective in *Cryptosporidium hominis*, defining the distribution patterns and host types for the various subtypes worldwide, and analysing how intra- and inter-population dynamics are shaped from an evolutionary perspective. Through this, a better understanding will be gained of the manner in which selection and divergence act on geographically-diverse populations of *C. hominis* across the world. The potential implication of host behaviour on how diversification of allele-types manifests itself will also be discussed and explored.

3.2 METHODS

(i) Systematic review of *C. hominis* GP60 sequences

NCBI's online archives PubMed (https://www.ncbi.nlm.nih.gov/pubmed) and GenBank (https://www.ncbi.nlm.nih.gov/genbank/) were used to identify and extract data for this study (1999 to present). Search terms included "*Cryptosporidium*", "*Cryptosporidium hominis*", "*C. hominis*", and "GP60". Data was initially extracted from all GenBank submissions described specifically as GP60 in *Cryptosporidium hominis* (search phrase "(GP60) AND "*Cryptosporidium hominis*"[porgn:__txid237895]"). Each submission was traced to the original publication (except for unpublished submissions) to identify missing descriptive data (e.g. host), and to confirm listed details from GenBank records. This was then supplemented to include all non-GenBank submitted data for GP60-sequenced isolates in *C. hominis* through systematic review and manual extraction of all relevant PubMed hits (search term "*C. hominis*" or "*Cryptosporidium hominis*"). Additional *C. hominis* GP60 sequences on GenBank that were missed due to ambiguous labelling (simply *Cryptosporidium spp.*) or mislabelled as *C. parvum* were identified by homology searches against representative sequences from each *C. hominis* subtype family (Ia, Ib, Id, Ie, If, Ig, Ih, Ii, Ij, and Ik) using BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch). The data collected during this systematic review included (where detailed in the associated publication or GenBank entry): the year of submission, country of origin, host type, GP60 subtype family (e.g. Ia, Ib, etc.), specific GP60 subtype (e.g. IbA10G2), and the associated DNA sequences. Submissions that were eligible for inclusion belonged to the *C. hominis* species-type, were ones where GP60 had been sequenced from DNA extract derived directly from collected stool samples, and for which direct submissions had been made to GenBank, or for which the GP60 subtype had been specified in the original publication. Sequences that derived from outbreak investigations or had been isolated from environmental samples (e.g. sewage, wild animal droppings) were excluded. GP60 sequences were also categorised based on the geographic region of isolate origin, in order to achieve detailed population genetics analyses within a geographic context. These regions were defined as Australasia, Central Asia, Europe, East Asia, North Africa/Middle East, North America, South America, and Sub-Saharan Africa (Figure 3.2.1).

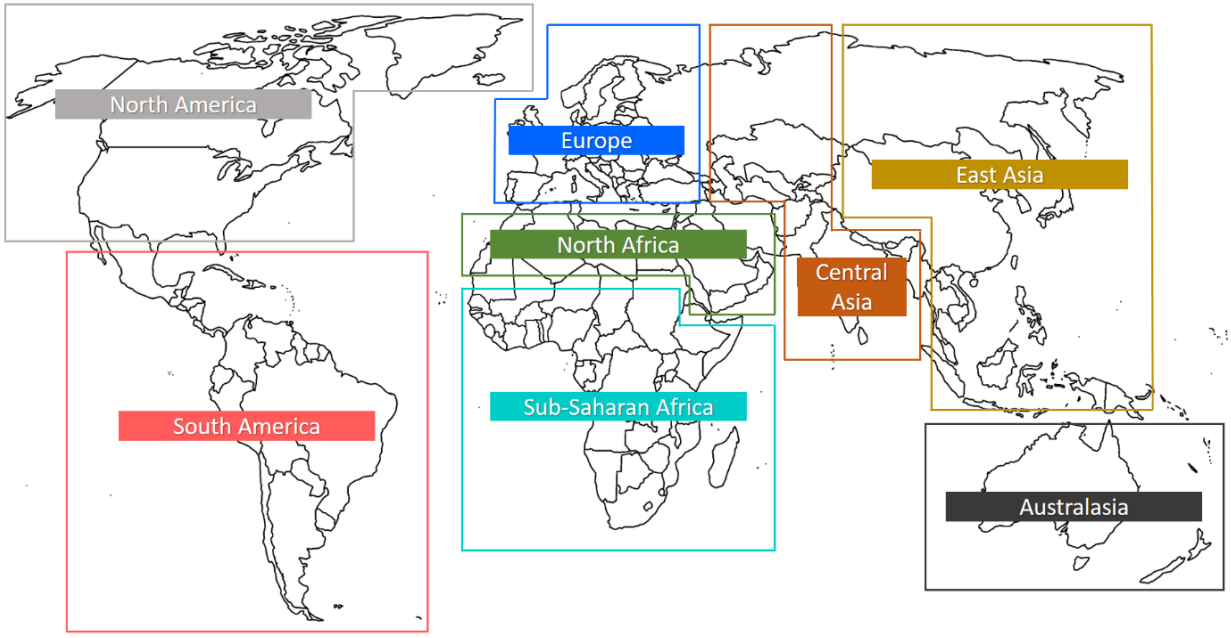


Figure 3.2.1. Boundary descriptions for geographic regions defined in this study as Australasia, Central Asia, East Asia, Europe, North Africa, North America, South America, and Sub-Saharan Africa

(ii) Multiple sequence alignment and phylogenetic analysis

Consensus sequences across all subtype families were prepared prior to alignment. The selected fragment was positioned along the GP60 coding sequence between 5'-GATGTTTCTGTT-3' and 5'-TGTGGTACTTCA-3' (**Ia**), 5'-GATGTTTCTGTT-3' and 5'-GTGCGGTACTTC-3' (**Ib**), 5'-GATGTTTCTGTT-3' and 5'-GTGCGGTACTTC-3' (**Id**), 5'-GATGTTTCTGTT-3' and 5'-GTGCGGTACTTC-3' (**Ie**) and 5'-GATGTTTCTGTT-3' and 5'-TGTGGTACTTCA-3' (**If**), 5'-GATGTTTCTGTT-3' and 5'-TGCGGTACTTCA-3' (**Ig**), 5'-GATGTTTCTGTT-3' and 5'-TGCGGTACTTCA-3' (**Ih**), 5'-GATGTTTCTGTT-3' and 5'-TGCGGTACTTCA-3' (**Ii**), 5'-GATGTTTCTGTT-3' and 5'-TGCGGTACTTCA-3' (**Ij**), and 5'-GATGTTTCTGTT-3' and 5'-TGCGGTACTTCA-3' (**Ik**). All sequence editing, alignment, and phylogenetic analysis was performed in using the Molecular Evolutionary Genetics Analysis tool (MEGA v 7.0; Kumar *et al*. 2016). Sequence formatting and alignment was performed in the alignment explorer window, according to the ClustalW algorithm for codons (Thompson *et al*. 1994). Pairwise phylogenetic distances were calculated using the Maximum Composite Likelihood (MCL) method, estimating the Maximum Likelihood (ML) topology using the Tamura-Nei model. All positions with less than 95% site coverage were excluded from the final estimation of distances, meaning less than 5% alignment gaps, missing data, and ambiguous bases were permitted at any position of the alignment. Phylogenetic distances were subsequently exported to FigTree graphical design software

(http://tree.bio.ed.ac.uk/software/figtree/) for formatting and preparation of the final phylogeny. The final tree is scaled according to the number of nucleotide substitutions per site.

(iii) Population genetics analysis of *C. hominis* GP60 sequences

GP60 sequences belonging to subtype families Ig, Ih, Ii, Ij, and Ik were so few and genetically homogeneous that they were excluded from further population analyses, leaving only subtypes Ia, Ib, Id, Ie, and If. As most of the bioinformatics approaches that were used in this study to describe population genetics require a minimum of 4 sequences, and/or a minimum of one single nucleotide polymorphism (SNPs) between sequences, the subtype families that did not meet these criteria (in this instance Ig, Ih, Ii, Ij and Ik) could not be included. Intra-subtype levels of allelic diversity, gene flow, and genetic differentiation among aligned GP60 consensus sequences were assessed using the DnaSP program v. 5.10.1 (Librado & Rozas, 2009). Signals of selection and predicted gene flow within subtype families were determined using Tajima's D (Tajima, 1989) and Fu & Li's D (Fu & Li, 1993) tests, for geographically-defined populations of four or more sequences, estimating values as moderately significant ($P < 0.05$), significant ($P < 0.01$), or highly significant ($P < 0.001$). Sequence diversity (π) and allelic diversity (AD) within these populations were determined using Nei's method (Nei, 1987) for populations of two or more sequences. Genetic differentiation between geographically-isolated populations ($N > 2$) was described using F_{st} (Hudson *et al*. 1992), providing a model of population structure that defined differentiation as negligible ($F_{st} < 0.15$), moderate ($0.15 < F_{st} < 0.25$), and high ($F_{st} > 0.25$) (Wright, 1978). Absolute diversity between these defined allele populations was further assessed using D_{xy}, according to Nei's model (Nei, 1987). Phylogeographic analyses were performed based on ratios between allelic frequencies (G_{st}) and similarities (N_{st}) between regional populations (Lowe *et al*. 2004). The greater Nst > Gst, the more likely two *C. hominis* allele sequences (i.e. genetic variations of the same gene) are to co-occupy two geographic regions under comparison; alternatively, Nst < Gst indicates that alleles are less likely to coincide geographically. Nst = Gst demonstrates an absence of phylogeographic associations, and describes a scenario when no trend in either direction could be identified between two sequences. Selective pressure between *C. hominis* alleles was assessed using a ratio of non-synonymous (Ka) to

synonymous (Ks) substitution rates, with Ka/ks > 1 indicating a significant signature of adaptive evolution (Kryazhimskiy & Plotkin, 2008). Instances of Ks = 0 were overcome by implementing the equation Ka/((Ks + 1)/S), with S representing the total number of synonymous sites (Stoletzki & Eyre-Walker, 2011). Evolutionary distances were also used to investigate intra-subtype population diversity, determined according to the P-distance method (Nei & Kumar, 2000), using the pairwise distance calculator in MEGA v 7.0 (Kumar *et al*. 2016). An independent student's *t*-test was used to assess statistical significance of variances between compared populations (sig. p-value < 0.05), and the *F*-test was used to assign significance to regression models for correlated variables (sig. p-value < 0.05), both in SPSS v 22.0 (IBM Corp., 2013). All figures and tables were produced and edited in Microsoft Excel and Microsoft PowerPoint (2013).

3.3 RESULTS

(i) Systematic review of GP60-subtyped *C. hominis* isolates reveals novel geographic disparities and highlights significant information gaps worldwide

The purpose of the systematic review conducted in the study served three primary purposes: (1) to describe geographic distributions and allelic diversities of *C. hominis* GP60 subtypes, (2) to identify host range patterns, and (3) to highlight areas across the world where *C. hominis*-specific GP60 data is scarce or non-existent (i.e. "data deficient regions"). A total of 2,972 individual isolates, from 103 peer-reviewed articles and 25 unpublished GenBank submissions, were identified and included in this study (Appendix VI). The geographic origins of these isolates spanned 71 countries (36.2% of the total 196 countries recognised worldwide), but half of these isolates (50.2%) derived from only three countries, Australia, Spain, and The Netherlands (Figure 3.3.1), and more than half of the 71 countries (56.3%) had GP60-subtyped less than ten individual *C. hominis* isolates. A total of seventeen separate host types were identified across these geographic localities. Aside from the vastly predominant human host type (95.6%), the other host types identified (130 of 2972) included donkeys (2.1%), rhesus monkeys (0.6%), cattle (0.6%), sheep (0.4%), baboons (0.2%), horses (0.1%), kangaroos (0.07%), and single instances (0.03% each) of a chimpanzee, cynamolgus monkey, deer, fish, flying fox, gibbon, goat, squirrel monkey, and a wallaby. These 130 non-human isolates derived from only ten countries: China (65.4%), the United Kingdom (10.0%), New Zealand (9.2%), Australia (6.9%), Kenya (3.8%), Papua New Guinea (1.5%), Algeria (0.8%), Brazil (0.8%), India (0.8%), and the United States (0.8%).

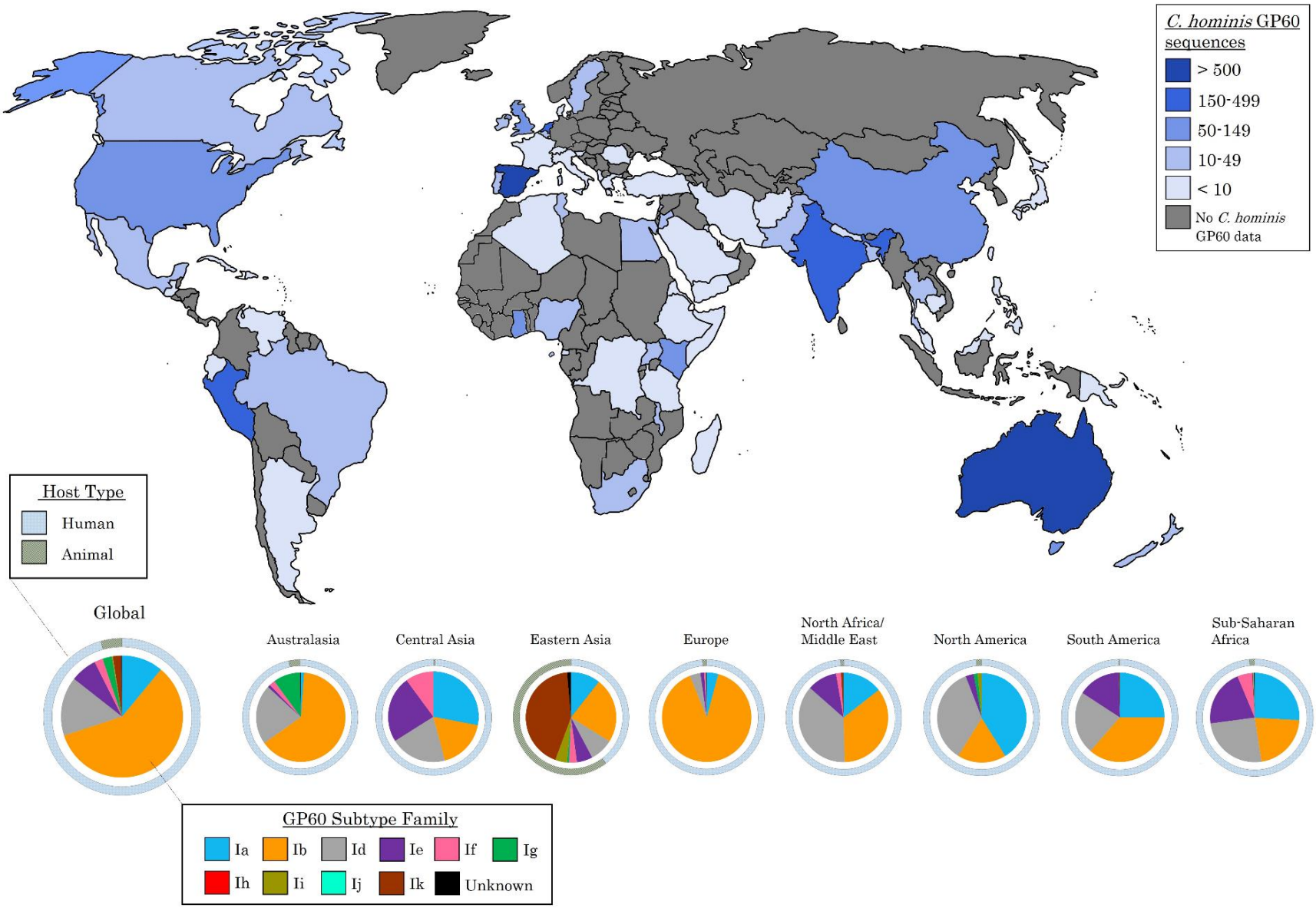


Figure 3.3.1. Distribution and host range of *C. hominis* GP60 subtype families, and quantitative evaluation of available isolate sequence data worldwide

Figure 3.3.1. Distribution and host range of *C. hominis* GP60 subtype families, and quantitative evaluation of available isolate sequence data worldwide

A total of 2,972 individual GP60-subtyped, non-outbreak related, *C. hominis* isolate sequences were identified and included, through exhaustive literature review of all published *Cryptosporidium* studies (PubMed search: https://www.ncbi.nlm.nih.gov/pubmed), and by homology searches of GenBank sequences to identify unpublished submissions (https://www.ncbi.nlm.nih.gov/genbank/). A total of 103 published and 25 unpublished sources were identified, covering 71 countries and 17 unique host types. Countries were assigned heat-mapped values based on number of *C. hominis* isolates GP60-sequenced to date. The proportions of GP60 subtype families identified, as well as the host type from which isolates originated, were characterised globally and by defined geographic regions. Unknown GP60 subtypes refer to isolates exhibiting <95.0% sequence identities with any of the 10 defined subtype families (Ia, Ib, Id, Ie, If, Ig, Ih, Ii, Ij, and Ik). Pie chart data describing regional distributions of subtype families are also presented in more detail in Figure 3.3.2. Individual data points are outlined and referenced in Appendix VI.

To overcome the shortcomings of a widespread geographical bias and absence of *C. hominis* GP60 subtype data, locations of origin were characterised by continental regions rather than on a country-by-country basis. This made it possible to draw more powerful associations between distributions, host types, and allelic diversities. On a global level, GP60 subtype family Ib was the most predominant, representing 59.0% of all subtyped *C. hominis* isolates, followed by Id (15.6%), Ia (11.0%), Ie (7.0%), Ig (2.5%), If (2.3%), Ik (2.2%), Ii (0.2%), Ih (0.03%) and Ij (0.03%). There were also 8 isolates (0.3%) that were classified as “unknown”, as they exhibited less than 95.0% overall sequence identities to any of the hitherto described subtype families. Individual subtypes showed clear phenotypic distinctions in terms of host range and distribution. Subtype Ib isolates were overwhelmingly derived from Europe (56.8%) and Australasia (26.2%), and were identified in ten different host types. In contrast to this, subtype Ia, which constitutes the original *C. hominis* TU502 reference strain subtype, was most frequently identified in Sub-Saharan Africa (26.3%), South America (20.8%), and Central Asia (18.7%) (Figure 3.3.1 and 3.3.2), and was found in only two host types, humans (99.1%) and three rhesus monkeys (0.9%). Subtype Id, the second most prevalent globally, predominantly originated from Australasia (33.4%) compared to the seven other geographical regions (2.5-18.1%), and like Ib had a more expansive host range including one isolate each from cattle, a fish, a gibbon, a goat, a kangaroo, and a rhesus monkey (0.2% each). Subtype Ie resembled subtype Ia in terms of distribution patterns, predominating in Sub-Saharan Africa (33.1%), Central Asia (25.0%), and South America (20.2%), and was also only identified in human (99.5%) and rhesus monkey (0.5%) hosts. Subtype If was not identified in North or South American isolates, predominating in Central Asia (32.8%), Sub-Saharan Africa (26.9%), and Australasia (22.4%), and was found in humans (95.5%), baboons (3.0%), and rhesus monkeys (1.5%). Subtype Ig was entirely anthroponotic and found almost exclusively in

Australasia (95.9%) in addition to one single isolate each from Eastern Asia (1.4%), Europe (1.4%) and North America (1.4%). Only a single instance of subtypes Ih and Ij could be identified, the first in a human, and the latter in a baboon, both from Sub-Saharan Africa. Subtypes Ii and Ik were identified as predominantly zoonotic in the study. Although both were identified in humans (28.6% and 1.6% of isolates, respectively), subtype Ii was also found in donkeys (28.6%), a horse (14.3%), a rhesus monkey (14.3%) and a cynomolgus monkey (14.3%), and subtype Ik was predominantly identified in donkeys (93.8%) and less frequently in horses (4.7%).

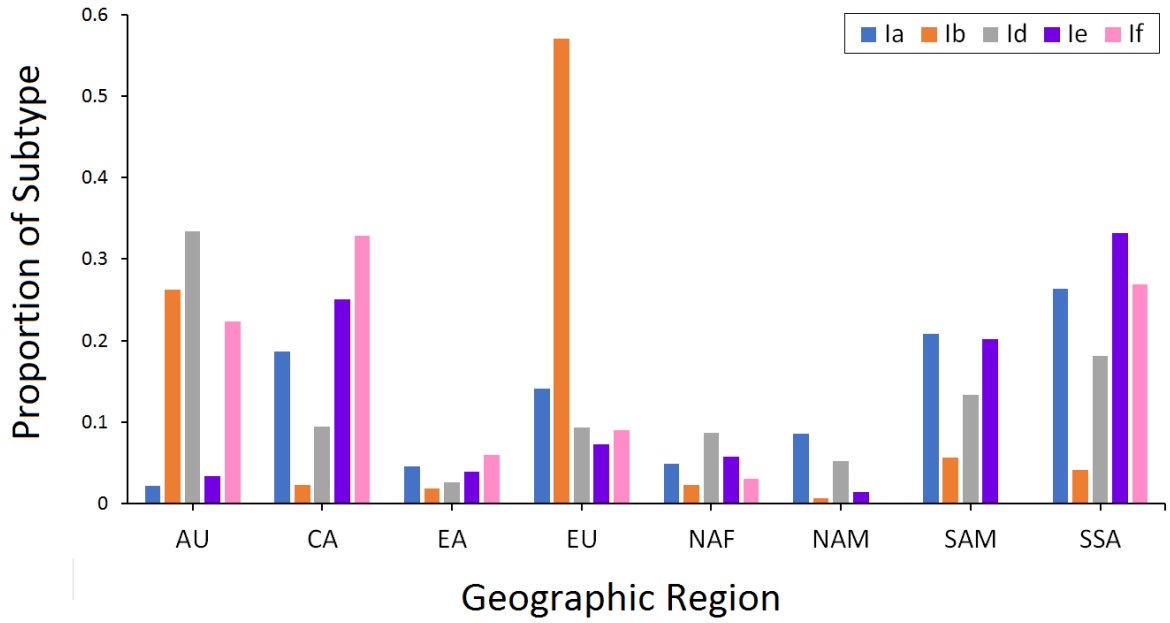


Figure 3.3.2. Geographic distribution of *C. hominis* GP60 subtype families

The geographic origins of *C. hominis* GP60 sequences is shown, describing the proportion of each subtype family that originated in a given region. Included subtype families here are Ia (N=327), Ib (N=1752), Id (N=464), Ie (N=208) and If (N=67), and include only non-study duplicate sequences. The remaining subtypes were excluded due to insufficient sequence data, or because they originated from less than 50.0% of the included geographic regions.

(ii) *C. hominis* GP60 sequences reveals phenotypic disparities that may be subject to selection

A total of 631 GP60 *C. hominis* sequences were initially identified and retrieved from GenBank, all of which were submitted between the years of 2002 and 2016. These sequences were all accurately annotated as *C. hominis*, and so were extracted from the GenBank depository directly using the search phrase "*Cryptosporidium hominis*" and filtering out non-GP60 sequences. 72 of these were immediately excluded for phylogenetic interpretation, as they constituted environmental wastewater samples (N = 19) or clones obtained through laboratory-based animal propagation systems (N=53). A

further 79 isolates were inaccurately labelled as either *Cryptosporidium parvum* or simply *Cryptosporidium sp*. (submitted to GenBank between 1999 and 2016), and so had to be identified through homology searches of representative sequences from each *C. hominis* GP60-subtype family using BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and manually extracting significant sequence hits (>95.0% sequence identities). A further refinement made prior to phylogenetic analyses was the removal of study duplicates, i.e. 100% identical sequences and corresponding host types identified within a single study. This was done to improve the interpretability of a dataset that already contained a high number of replicate sequences, and because the inclusion of duplicate study sequences would not have added any improvements to the resolution of phylogenies or genetic distances at the intra- or inter-subtype level comparisons. This resulted in a final dataset of 480 individual GP60 *C. hominis* sequences for phylogenetic and diversification analyses (Figure 3.3.3).

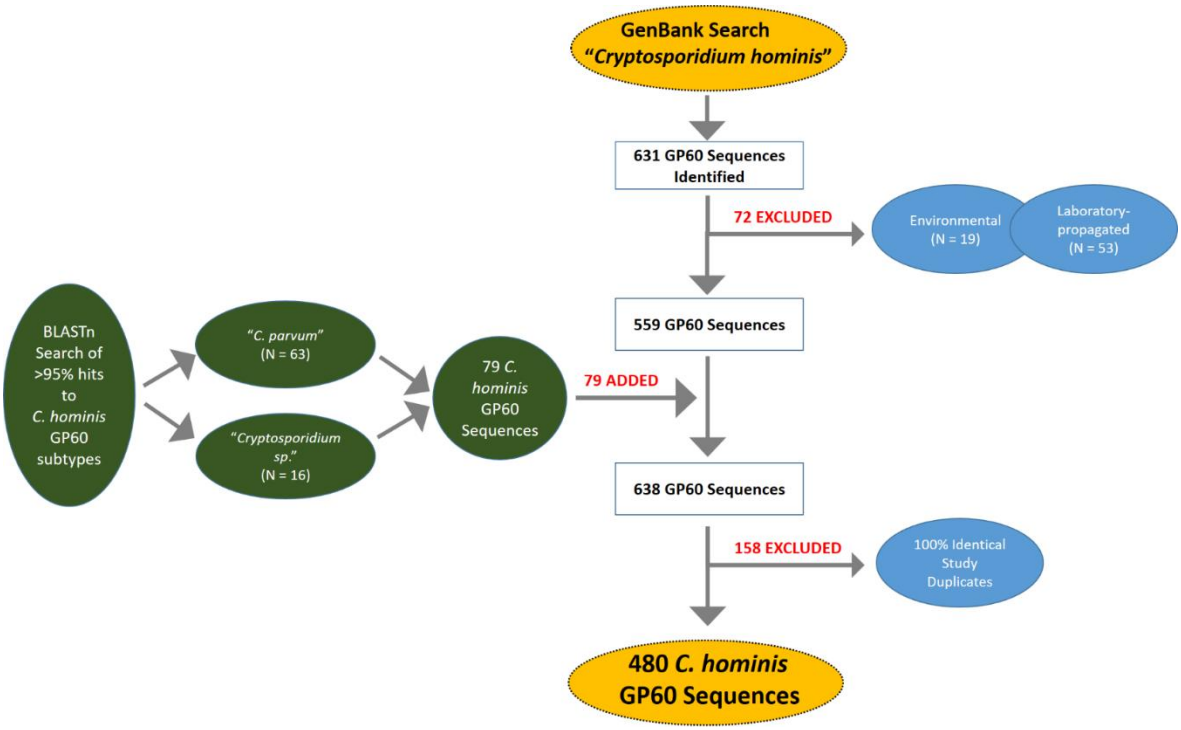


Figure 3.3.3. Flow chart of *C. hominis* GP60 sequence systematic review and extraction for evolutionary and divergence analyses

C. hominis GP60 sequence data was initially identified by retrieving all hits for the search term "*Cryptosporidium hominis*" in combination with "GP60", "GP15", "GP40", and "GP40/15" on GenBank (http://www.ncbi.nlm.gov/genbank/); these hits were subsequently filtered to remove all environmental or laboratory-propagated samples. Each known *C. hominis* GP60 subtype family sequence (Ia, Ib, Id, Ie, If, Ig, Ih, Ii, Ij, and Ik) was subsequently blasted against NCBI's BLASTn database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify *C. hominis* GP60 sequences that had been mislabelled as "*C. parvum*" or "*Cryptosporidium sp*". This resulted in a preliminary dataset of 638 sequences, which was subsequently filtered to remove all 100% identical study duplicates. The final dataset consisted of 480 *C. hominis* GP60 sequences.

C. hominis GP60 subtype families clustered into two very distinct superclades: The B-E-F clade, and the A-D-G-H clade. Subtypes I and K additionally formed an isolated outgroup clade that was completely separate from any of the other subtype families, as did subtype J, however this subtype appeared to be a distant outgroup within the A-D-G-H superclade (figure 3.3.4) and so could be characterized as evolutionarily, albeit distantly, part of this clade. Subtypes E and F, and subtypes A and G, further formed minor clades within their respective superclades, while subtypes B and D remained phylogenetically distinct. Subtype Ib was represented by the largest proportion of non-duplicate study isolates, but only exhibited the third largest intra-family diversity (range in nucleotide diversity: π = 0-0.021, mean = 0.0045 ± 0.00005). In contrast, subtype Ia, the second most prevalent in the phylogeny, had the largest range in nucleotide diversity between strains out of all of the included families (π = 0.00-0.085, mean = 0.0083 ± 0.000161).

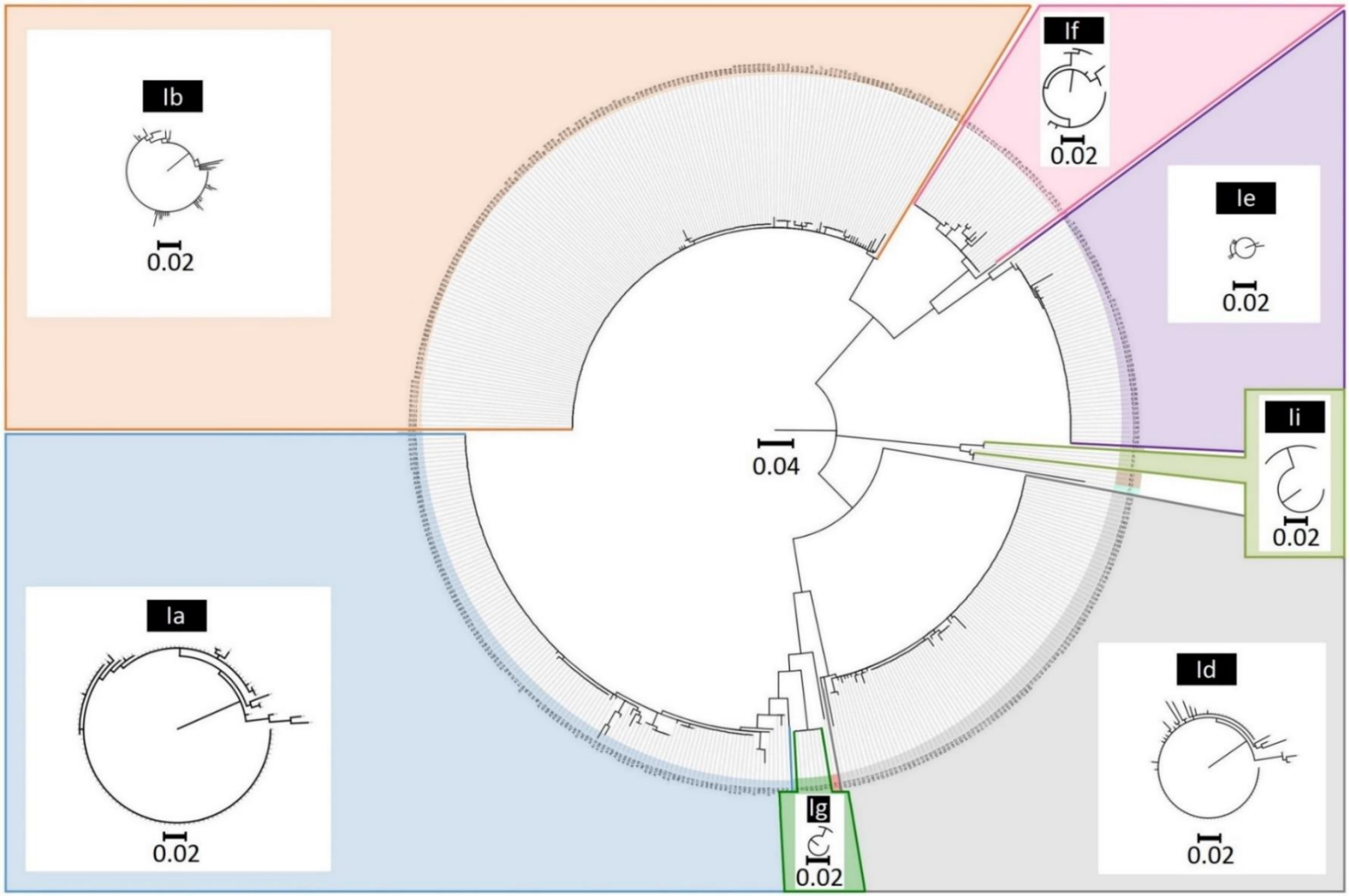


Figure 3.3.4. GP60 phylogeny of *C. hominis* isolates reveals evolutionary relationships between subtype families and differential degrees of intra-family divergences

The evolutionary history between *C. hominis* GP60 sequences was inferred using the Maximum Likelihood method, estimating genetic distances based on the Tamura-Nei model. 480 sequences were included in the final dataset, and a 360bp consensus sequence across all entries was used. The tree with the highest log likelihood (-2963.9104) is shown. All positions with less than 95% site coverage were excluded from the final phylogeny estimation, meaning less than 5% alignment gaps, missing data, and ambiguous bases were permitted at any position of the alignment. All phylogenetic analyses were conducted in MEGA v 7.0 (Kumar et al. 2016); genetic distances were subsequently exported in Newick format, and constructed into a circular phylogeny using the FigTree phylogenetic graphical design software (http://tree.bio.ed.ac.uk/software/figtree/). Subtypes Ih, Ij, and Ik were excluded from within-subtype phylogenetic characterisation, as Ih and Ij were represented by single sequences only, and because Ik showed insufficient sequence divergence ($\pi = 0$) for isolate discrimination, and are highlighted in red, turquoise, and brown respectively.

Diversity and selective pressure of global isolates within their respective GP60 subtype families was assessed for human, non-study duplicate isolates (N=427), and was only possible where intra-subtype isolates displayed some degree of SNP-based divergence; this led to the exclusion of subtype families Ig, Ih, Ii, Ij, and Ik. Subtypes Ia, Ib, Id, Ie, and If possessed a number of SNPs at the intra-family level (see Table 3.3.1), and were taken forward in the subsequent analysis. Subtypes Ia and Id exhibited the greatest degree of allelic diversity, with Ka/Ks ranges of 0.000-4.484 (mean(±SEM) = 0.368 ± 0.006) and 0.000-4.138 (0.298 ± 0.008), and evolutionary distances ranging between π=0.000-0.085 (0.008 ± 0.0001) and π=0.000-0.057 (0.005 ± 0.0001), respectively (Figure 3.3.5; Table 3.3.1). Both of these subtypes had significantly negative Tajima's D-values (P<0.01 and 0.001, respectively) (Table 3.3.2). A negative Tajima's D indicates a relative excess of low frequency SNPs, which could be the result of a recent population size expansion and/or a selective sweep. Subtype If also had a comparably high Ka/Ks range (0.000-3.500, (0.480 ± 0.028)), but a substantially lower evolutionary distance (0.000-0.028), and a Tajima's D-value that was not significantly deviating from zero (P>0.10), consistent with a population that is in a mutation-drift equilibrium. The two remaining subtypes, Ib and Ie, showed a substantially lower range of selective pressure (Ka/Ks ranges 0.000-1.573 (0.319 ± 0.004) and 0.000-1.256 (0.172 ± 0.007), respectively) and diversity (π ranges 0.000-0.021 (0.005 ± 0.00005) and 0.000-0.018 (0.002 ± 0.00008), respectively), and were associated with non-significant Tajima's D-values (P>0.01) (Table 3.3.2). These data indicate an absence of significant selective pressure (neither positive or negative) and no evidence of a recent population size expansion of these subtypes.

Table 3.3.1. Population genetics summary statistics for *C. hominis* GP60 subtypes

Subtype	No. human samples (study duplicates removed)	No. synonymous SNPs	No. non-synonymous SNPs	Ka/Ks (mean ± SEM)	Nucleotide variation π (mean ± SEM)
Ia	128	12	31	0.000-4.484 (0.368 ± 0.006)	0.000-0.0850 (0.008 ± 0.0001)
Ib	130	7	13	0.000-1.573 (0.319 ± 0.004)	0.000-0.021 (0.005 ± 0.00005)
Id	91	21	40	0.000-4.138 (0.298 ± 0.008)	0.000-0.057 (0.005 ± 0.0001)
Ie	52	4	8	0.000-1.256 (0.172 ± 0.007)	0.000-0.018 (0.002 ± 0.00008)
If	26	4	10	0.000-3.500 (0.480 ± 0.028)	0.000-0.028 (0.008 ± 0.0004)
Ig	8	0	0	N/A	N/A
Ih	1	N/A	N/A	N/A	N/A
Ii	3	0	0	N/A	N/A
Ij	1	N/A	N/A	N/A	N/A
Ik	3	0	0	N/A	N/A

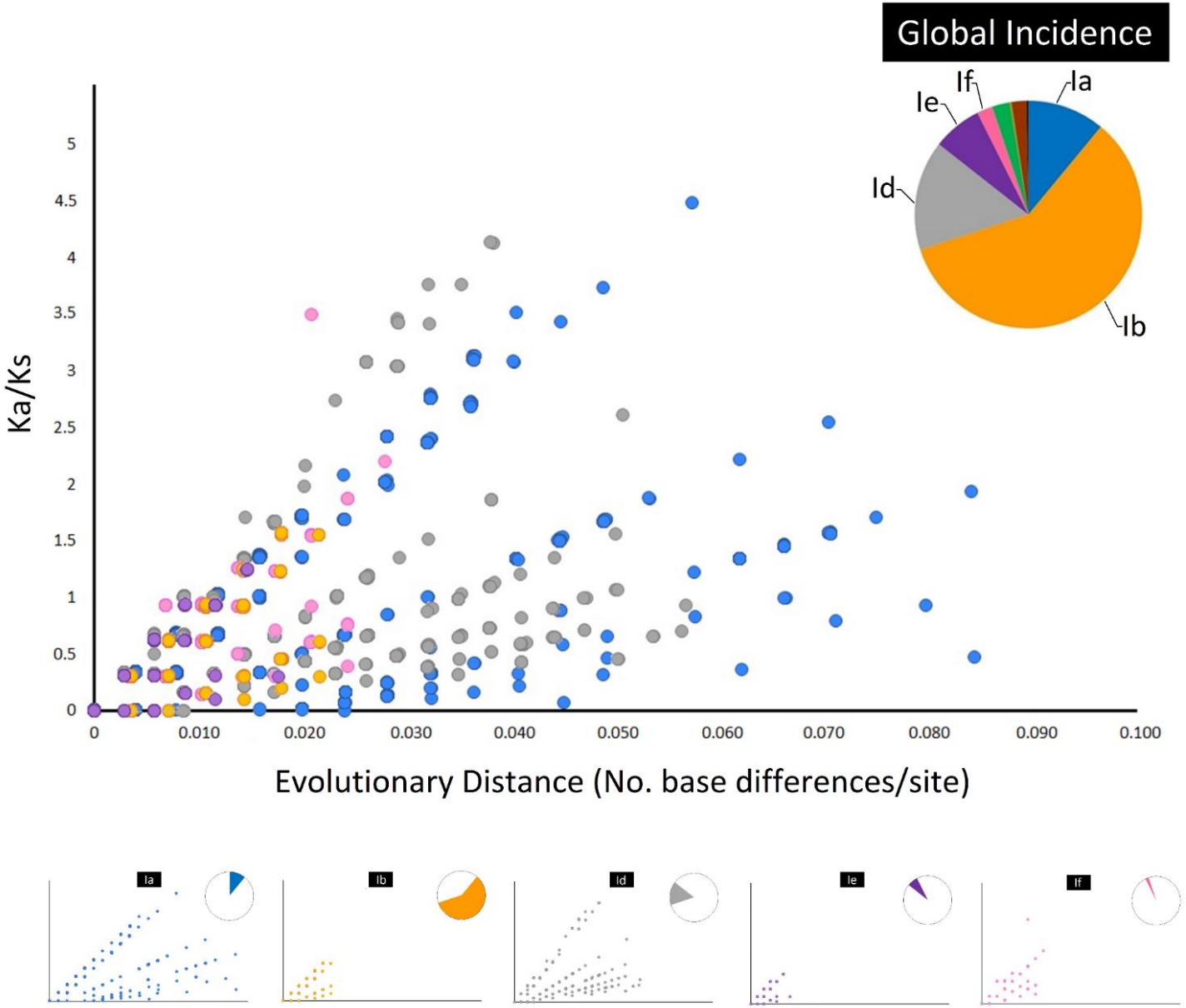


Figure 3.3.5. Population genetics of *C. hominis* GP60 alleles show differential patterns of diversity and selective pressure between subtype families Ia, Ib, Id, Ie, and If

427 *C. hominis* GP60 isolates (non-study duplicates) of human origin were included, characterised as subtype families Ia (N=128), Ib (130), Id (91), Ie (52), and If (26). Consensus sequences (360bp) were aligned separately for each subtype family using the codon-based ClustalW algorithm imbedded in MEGA v 7.0 (Kumar *et al.* 2016). Aligned sequences were subjected to intra-subtype Ka/Ks analysis using DnaSP v 5.10.1 (Librado & Rozas, 2009), and evolutionary distances (p-distance method) were calculated using MEGA v 7.0 (Kumar *et al*. 2016). Results are presented here both as a composite (main figure) and by individual subtype families (minor figures).

Intra-subtype comparisons between strains of shared geographical origin revealed that allelic diversity had interesting global disparities within each of the subtype families analysed (Appendices VI-XI). Subtype Ia strains originating in Central Asia, North Africa, and Europe showed the greatest degree of GP60 divergence, with little to no variation between isolates from Sub-Saharan Africa, Australasia, Eastern Asia, and the Americas. Subtype Ib isolates were substantially more variable if they originated from Australasia than any of the other defined regions, once again highlighting inverse associations between prevalence and diversity, as the overwhelmingly predominant European Ib isolates showed a negligible degree of variance. Subtype Id, on the other hand, had a strong correlation between isolate prevalence and diversity across most of the geographic regions (Regression: R-squared: 0.851, $F = 11.77$, $p < 0.05$). The only outliers were South American isolates, which were the fourth most prevalent and appeared monomorphic at the GP60 locus. Subtype Ie had similar patterns to Ia, with Central Asian isolates exhibiting the greatest degree of variation, and minor diversity in Australasia and South America, while the remaining regions showed overwhelming genetic identity. Subtype If was interesting in that showed allelic diversity primarily in Australasian isolates. A single instance of allelic diversity was additionally revealed for isolates from Sub-Saharan Africa, with the remaining four geographic regions (Central Asia, Eastern Asia, Europe, and North Africa) showing no variation between isolates. Although Australasian samples were the most common (38.5%), a significant number of isolates also came from other regions e.g. Central Asia (23.1%) and Europe (15.4%), so the heightened allelic diversity within Australasian strains could indicate a unique external driver for selective pressure acting upon infectious pathogens of this subtype in this region of the world.

(iii) Genetic differentiation of GP60 *C. hominis* alleles show geographic and intra-subtype family disparities

In order to evaluate whether any regionally-defined GP60 sequences exhibited a deviation from mutation drift equilibrium (neutrality), thereby evidencing a potential recent population contraction or expansion event, Tajima's D was used to assess genetic differentiation. The relatively elevated frequencies of rare alleles for within-subtype sequence comparisons, as indicated by significantly negative ($P<0.05$) Tajima's D values, were identified only in three geographic areas: Central Asia, Europe, and Sub-Saharan

Africa (Table 3.3.2). This is consistent with a recent selective sweep and/or recent population expansion (e.g. after a founder event or bottleneck). In Central Asia, three subtype family showed highly significant values for Tajima's D<0 (Ia, Id, and Ie), while in Sub-Saharan Africa the same trend was identified for subtypes Ia and Id, and in Europe subtype Id isolates seemed to harbour an excess of uncommon alleles. These negative Tajima's D values were most significant for subtype Ia in Central Asia (P<0.001) and for subtype Id in Sub-Saharan Africa (P<0.001). All of the remaining Tajima's D-values were non-significantly different from zero, indicating that the populations for this gene are in a mutation drift equilibrium.

Table 3.3.2. *C. hominis* GP60 intra-subtype family sequence samples size (N), allele number, allelic diversity (AD), nucleotide diversity (π), Tajima's D, and Fu & Li's F values, for human-derived *Cryptosporidium hominis* isolates worldwide. Results are segregated based on subtype and subcategorised by geographic region.

Ia						
Region	Seq (N)	No. Alleles	AD (±SE)	π (±SE)	Tajima's D	Fu & Li's F
AU	3	3	1.000 (±0.272)	0.01434 (±0.00421)	‡	‡
CA	39	9	0.372 (±0.100)	0.00533 (±0.00192)	-2.53736***	-4.29086**
EA	7	3	0.524 (±0.209)	0.00207 (±0.00093)	-1.23716	-1.37408
EU	15	7	0.819 (±0.082)	0.01081 (±0.00448)	-1.68507	-2.21738
NAF	9	8	0.972 (±0.064)	0.01811 (±0.00426)	-1.06317	-1.40289
NAM	14	4	0.659 (±0.090)	0.00382 (±0.00128)	-1.04237	-1.72592
SAM	18	5	0.719 (±0.091)	0.00326 (±0.00063)	-0.65623	-0.02902
SSA	23	5	0.628 (±0.072)	0.00709 (±0.00425)	-2.19418**	-3.80205**
Ib						
Region	Seq (N)	No. Alleles	AD (±SE)	π (±SE)	Tajima's D	Fu & Li's F
AU	24	11	0.862 (±0.053)	0.00855 (±0.00120)	-1.76358	-1.76358*
CA	9	2	0.222 (±0.166)	0.00073 (±0.00055)	-1.08823	-1.08823
EA	9	2	0.222 (±0.166)	0.00222 (±0.00166)	-1.51297	-1.51297
EU	47	3	0.198 (±0.074)	0.00120 (±0.00047)	-0.98279	-0.98279
NAF	12	5	0.667 (±0.141)	0.00387 (±0.00093)	0.47201	0.47201
NAM	9	2	0.556 (±0.090)	0.00550 (±0.00089)	1.94806	1.94806
SAM	11	2	0.182 (±0.144)	0.00180 (±0.00142)	-1.59996	-1.59996
SSA	9	3	0.639 (±0.126)	0.00685 (±0.00146)	0.49704	0.49704
Id						
Region	Seq (N)	No. Alleles	AD (±SE)	π (±SE)	Tajima's D	Fu & Li's F
AU	12	8	0.894 (±0.078)	0.01618 (±0.00401)	-0.43316	-0.44922
CA	20	9	0.705 (±0.111)	0.00700 (±0.00222)	-2.26073**	-3.68715**
EA	7	2	0.286 (±0.196)	0.00078 (±0.00054)	-1.00623	-1.10146
EU	13	6	0.641 (±0.150)	0.00551 (±0.00211)	-2.15214**	-2.96683**
NAF	4	1	0	0	N/A	N/A
NAM	3	2	0.667 (±0.314)	0.00179 (±0.00084)	‡	‡
SAM	10	1	0	0	N/A	N/A
SSA	22	8	0.545 (±0.128)	0.00640 (±0.00272)	-2.47774***	-3.97522**
Ie						
Region	Seq (N)	No. Alleles	AD (±SE)	π (±SE)	Tajima's D	Fu & Li's F
AU	5	4	0.900 (±0.161)	0.00402 (±0.00110)	-0.17475	-0.17531
CA	13	4	0.423 (±0.164)	0.00309 (±0.00157)	-2.14681*	-2.69137**
EA	7	2	0.571 (±0.119)	0.00164 (±0.00034)	1.34164	1.10146
EU	6	1	0	0	N/A	N/A
NAF	4	2	0.500 (±0.265)	0.00144 (±0.00076)	-0.61237	-0.47871
NAM	3	2	0.667 (±0.314)	0.00192 (±0.00090)	‡	‡
SAM	6	2	0.333 (±0.215)	0.00192 (±0.00124)	-1.13197	-1.19511
SSA	8	2	0.250 (±0.180)	0.00072 (±0.00052)	-1.05482	-1.20353
If						
Region	Seq (N)	No. Alleles	AD (±SE)	π (±SE)	Tajima's D	Fu & Li's F
AU	10	6	0.889 (±0.075)	0.01077 (±0.00182)	-0.06612	-0.17822
CA	6	1	0	N/A	N/A	N/A
EA	1	1	0	N/A	N/A	N/A
EU	4	1	0	N/A	N/A	N/A
NAF	1	1	0	N/A	N/A	N/A
NAM	0	N/A	N/A	N/A	N/A	N/A
SAM	0	N/A	N/A	N/A	N/A	N/A
SSA	4	2	0.500 (±0.265)	0.01020 (±0.00541)	-0.80861	-0.77723

‡ Tajima's D and Fu & Li's F values require a minimum of 4 input sequences
***** Sig. P-value < 0.05, ****** Sig. P-value < 0.01, ******* Sig. P-value < 0.001
N/A No intra-subtype diversity or sequence data

(iv) Intra-family allelic diversity reveals globally-dominant allele types for GP60 subtypes

Allelic diversity was characterised further within *C. hominis* GP60 subtypes from a phylogenetic perspective, to identify whether subtype families contained dominant allele types causing human infections worldwide. This provided better insight into the global distribution and population structure of alleles at the intra-subtype level (Appendices XIII-XVII). A striking similarity between all of the five subtype families included in this analysis was that alleles predominantly grouped into two distinct allotypes when gaps and microsatellites were excluded from phylogenetic constructs (Appendices XIII-XVII, panel A). The existence of two such major SNP-based allotypes creates a simplified approach for understanding both the population structure and spread of *C. hominis* isolates worldwide, allowing genotyping approaches to focus on only three classes of GP60 alleles, namely "A", "B", and the less frequent "Other" types (Figure 3.3.6). Using this approach, intra-family GP60 alleles were categorised geographically, revealing clear distribution disparities between the various allele types (Appendices XIII-XVII, panel B). Allotype A was more than twice as prevalent as the other two classes for all subtype families except If. This indicates the existence of a single, dominant genotype within subtypes Ia-Ie that accounts for a substantial proportion of human cryptosporidiosis. From a geographic perspective, there were proportional disparities between and within the various allotypes; Ia type A was, for example, overwhelmingly present in Central Asian and Sub-Saharan African samples (>75.0% of type A), while South American and North American samples comprised more than two-thirds of Ia type B alleles. In contrast, Ib type A (which corresponds to the well-known IbA10G2 *C. hominis* subtype), was most common in European samples (>50.0%), while Ib type B was slightly more prevalent in isolates that had Central Asian origins. Subtypes Id and Ie had relatively even proportions across all geographic regions for allotype A, but for type B showed a predominance in Central Asian and East Asian samples, respectively. Proportions of allotypes were entirely dissimilar across the geographic categories and subtype families analysed. GP60 subtypes Ia and Ib were the most heterogeneous in terms of global allelic distribution, with 62.5% and 75.0% of geographic regions harbouring both A and B allotypes, respectively, and all of the remaining geographic regions containing either allotype A or B together with another less common ("other") allotype (Appendices XIII-XVII, panel C). For subtype Id, on the other hand, only half of the geographic regions contained a mixture of types A and B, and two

regions (North Africa and South America) revealed an apparently homogeneous allelic population structure comprised of allotype A only. Subtype Ie was similarly homogeneous in European samples (100.0% allotype A), and in subtype If homogeneity outweighed heterogeneity for analysed GP60 sequences (62.5% of geographic regions comprised entirely of type A, type B, or "other"), although the bias introduced by the significantly smaller sample size in this instance makes If allelic diversity more difficult to assess reliably. The overall pattern across the entirety of comparisons points towards a relatively heterogeneous global GP60 allelic population structure, with less than one-fifth of the geographic regions analysed (7 of 38 comparisons, 18.4%) exhibiting an entirely homogeneous sample. This study revealed an epidemiological feature (i.e. geographic distribution and composition of alleles) that appears as diverse at the intra-subtype level as it is at the inter-subtype level, indicating a ubiquitous spread of *C. hominis* allotypes and the absence of any defined associations between geographic boundaries and population structure for global GP60 alleles.

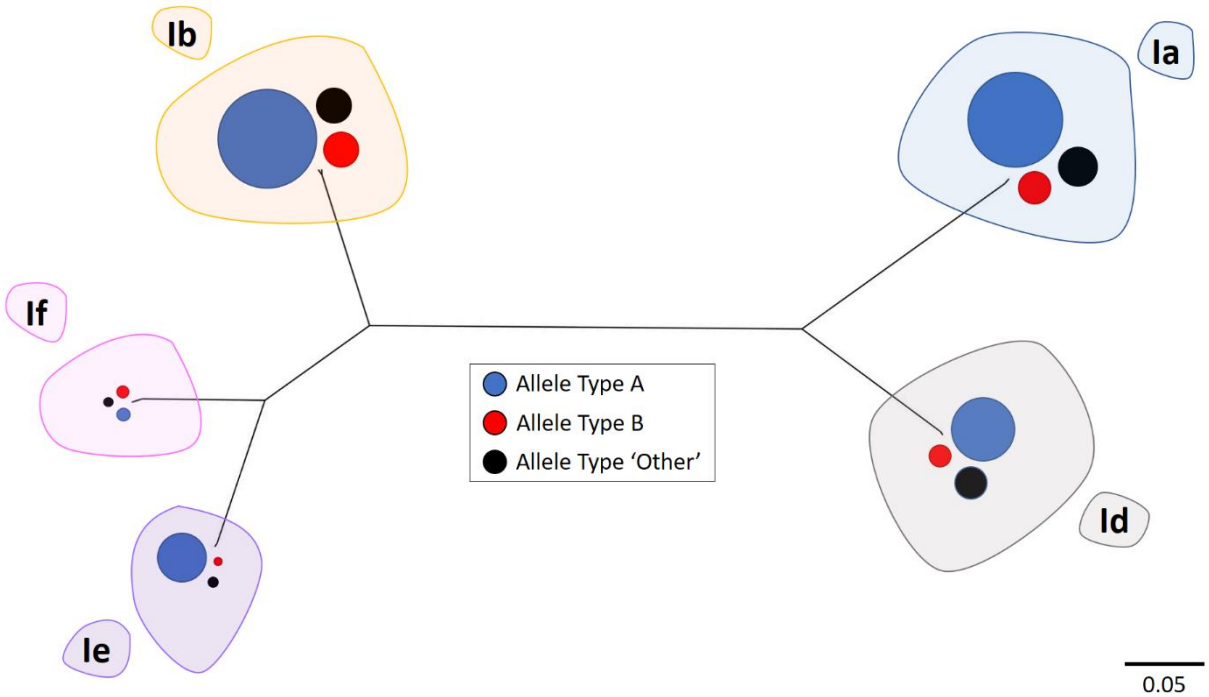


Figure 3.3.6. GP60 sequences form binary dominant allele types across *C. hominis* subtype families

Relative proportions of dominant allele types A and B, along with composite proportions of rare allele types 'other', are here illustrated for *C. hominis* GP60 subtype families Ia-If. Allele types A, B, and 'other' were based on the following no. of sequences per subtype family: Ia (73, 25, 30), Ib (76, 27, 27), Id (49, 17, 25), Ie (38, 6, 8), and If (10, 9, 7). Allele type frequencies were calculated using DnaSP v 5.10.1 (Librado & Rozas, 2009) for aligned GP60 sequences excluding gaps, and are presented here in combination with the relative phylogenetic distribution of subtype families. The radial view phylogeny was generated as a maximum likelihood tree according to the Tamura-Nei model and Maximum Composite Likelihood approach, with scale bars describing branch lengths according to the number of substitutions per site (MEGA v 7.0; Kumar *et al*. 2016).

(v) F_{st} reveals contrasting population structures between regions and GP60 subtype families

Population structure of subtype families was further evaluated through F_{st} analyses, to evaluate whether regionally-divergent populations are well-defined and genetically isolated (F_{st} = 1) or freely interbreeding (F_{st} = 0). Geographically-defined GP60 isolates showed a wide spectrum of genetic differentiation, from apparent panmixia to strongly defined population structure across all of the comparisons (Table 3.3.3). Values ranged from 0 to 1, with subtype Ib showing both the highest mean F_{st} values (mean F_{st} = 0.423) compared to the other four subtypes Ia, Id, Ie, and If (mean F_{st} = 0.208, 0.075, 0.088, and 0.353, respectively), and the highest proportion of significant ($P<0.05$) F_{st} values as well (53.6% of intra-Ib geographic comparisons). This indicates that the genetic variation observed in Ib isolates is largely accounted for by strong population structure amongst global sub-populations, and that a relatively low degree of gene flow is expected between geographically-diverse populations. Conversely, subtypes Id and Ie showed mostly weak to no genetic differentiation, with only one-tenth of regional comparisons exhibiting a significant F_{st}. For subtype Id, significant population structuring was only observed between isolates from Central Asian versus East Asian, North African and South American isolates, whereas for subtype Ie, the same F_{st} trend was found between isolates from Eastern Asia and European, North African and Sub-Saharan African samples. A reasonable proportion of Ia sequences demonstrated strong genetic differentiation (35.7%), with almost all geographic regions (barre Australia) exhibiting significant F_{st} values for at least one of the cross-regional comparisons. Subtype If contained too few samples for any comprehensive genetic differentiation to be assessed across geographic regions, although it was interesting to observe moderately strong and very strong F_{st} signals for Central Asian isolates versus Sub-Saharan African and European isolates, respectively (F_{st} = 0.248 and 1.000).

Central Asian and East Asian isolates separately showed the greatest proportion of significantly positive F_{st} values across all intra-subtype comparisons (most notably within subtype Ib), with nearly half of Central Asian and East Asian isolates exhibiting strong genetic differentiation compared to other geographic regions (46.7% each). This indicates significant population structuring and absence of genetic admixture of isolates between

these regions and other geographic localities. It is also an interesting contrast to the Australasian and European samples, which show a substantially lower overall degree of genetic differentiation (2/35 and 7/35 significant F_{st} values, respectively). This could point to a greater incidence of interbreeding and/or a more recent introduction of *C. hominis* isolates within these regions, based on genetic differentiation at the GP60 locus.

Table 3.3.3. F_{st} (bottom left) and D_{xy} (top right) values for inter-regional comparisons of *C. hominis* GP60 alleles

Ia	AU	CA	EA	EU	NAF	NAM	SAM	SSA
AU		0.00878	0.00598	0.00993	0.01656	0.00663	0.00654	0.00915
CA	0.12073		0.00609	0.00673	0.01248	0.00535	0.00651	0.00516
EA	-0.06250	0.40678*		0.00702	0.01382	0.00316	0.00268	0.00665
EU	0.04511	-0.00297	0.23404*		0.01423	0.00652	0.00738	0.00695
NAF	0.19079	0.14561	0.32770*	0.1269		0.0131	0.01438	0.01252
NAM	-0.07151	0.18382	0.05258	0.05907	0.23342*		0.00364	0.00578
SAM	-0.03922	0.37728*	-0.00274	0.21104*	0.32308*	0.05585		0.00709
SSA	0.13043	-0.01114	0.42125*	-0.00585	0.12973	0.20437*	0.39431*	
Ib	AU	CA	EA	EU	NAF	NAM	SAM	SSA
AU		0.00837	0.00982	0.00535	0.00619	0.00667	0.00529	0.00690
CA	0.43120*		0.00670	0.00696	0.00542	0.00355	0.00684	0.00552
EA	0.47477*	0.82353*		0.00710	0.00713	0.00665	0.00702	0.00722
EU	0.06402	0.85144*	0.79893*		0.00294	0.00395	0.00121	0.00376
NAF	-0.03065	0.55537*	0.60662*	0.09595		0.00397	0.00287	0.00437
NAM	0.04905	0.33333*	0.58553*	0.33906*	-0.00301		0.00387	0.00442
SAM	0.05275	0.84817*	0.79592*	-0.05797	0.07477	0.32407*		0.00369
SSA	-0.01925	0.44643*	0.52273*	0.12153	-0.07040	-0.04703	0.10437	
Id	AU	CA	EA	EU	NAF	NAM	SAM	SSA
AU		0.01363	0.01038	0.01267	0.00997	0.01029	0.00997	0.01300
CA	0.11026		0.00535	0.00745	0.00499	0.00461	0.00499	0.00817
EA	0.14141	0.24735*		0.00326	0.00041	0.00136	0.00041	0.00364
EU	0.10427	0.13166	0		0.00285	0.00365	0.00285	0.00595
NAF	0.14719	0.27368*	0	0		0.00095	0	0.00324
NAM	0.08112	0.00760	0	-0.04000	0		0.00095	0.00419
SAM	0.14719	0.27368*	0	0	0	0		0.00324
SSA	0.09804	0.16216	0.00338	-0.02138	0.00381	0.00295	0.00381	
Ie	AU	CA	EA	EU	NAF	NAM	SAM	SSA
AU		0.00416	0.00263	0.00287	0.00359	0.00268	0.00326	0.00323
CA	0.14362		0.00259	0.00155	0.00227	0.00236	0.00243	0.00191
EA	-0.07813	0.08537		0.00123	0.00195	0.00137	0.00178	0.00159
EU	0.30000*	0	0.33333*		0.00072	0.00096	0.00096	0.00036
NAF	0.24000*	0	0.21053*	0		0.00168	0.00168	0.00108
NAM	-0.10714	-0.06250	-0.30000	0	0		0.00160	0.00132
SAM	0.08824	-0.03030	0	0	0	-0.20000		0.00132
SSA	0.26667*	0	0.25806*	0	0	0	0	
If	AU	CA	EA	EU	NAF	NAM	SAM	SSA
AU		ND	ND	ND	ND	ND	ND	ND
CA	0.00722		ND	ND	ND	ND	ND	ND
EA	ND	ND		ND	ND	ND	ND	ND
EU	-0.22330	1.00000*	ND		ND	ND	ND	ND
NAF	ND	ND	ND	ND		ND	ND	ND
NAM	ND	ND	ND	ND	ND		ND	ND
SAM	ND	ND	ND	ND	ND	ND		ND
SSA	-0.03484	0.24788*	0	ND	ND	ND	ND	

*Sig. P-value < 0.05
ND: No Data

(vi) Human migration as a possible determinant of GP60 population structure for *C. hominis* subtypes Ia and Ib

In order to investigate the potential link between human movement (emigration and immigration) and GP60 subtype allele population structures (i.e. F_{st} values), standard regression analyses between fixation indices and net migration values were performed (Figure 3.3.7). Estimates of the net migration rate were taken from the last global census in 2015 (The World Bank, 2015), which represent the difference between numbers leaving versus numbers entering a country for a given year. This revealed an overall inverse association between regional composite (mean) F_{st} values and net migration. The regression model was only statistically significant for subtypes Ia and Ib (Regression: R-squared: 0.954, $F = 41.79$, $p < 0.05$ and R-squared: 0.936, $F = 29.96$, $p < 0.05$, respectively), while an absence of any statistically significant correlation was determined for subtypes Id and Ie (Regression: R-squared: 0.004, $F = 0.01$, $p > 0.05$ and R-squared: 0.441, $F = 1.58$, $p > 0.05$, respectively). The phenomenon of human movement (and thereby intra-regional population expansion or contraction) and the effect this may have on *C. hominis* population changes hence only applies to the Ia and Ib alleles within this dataset.

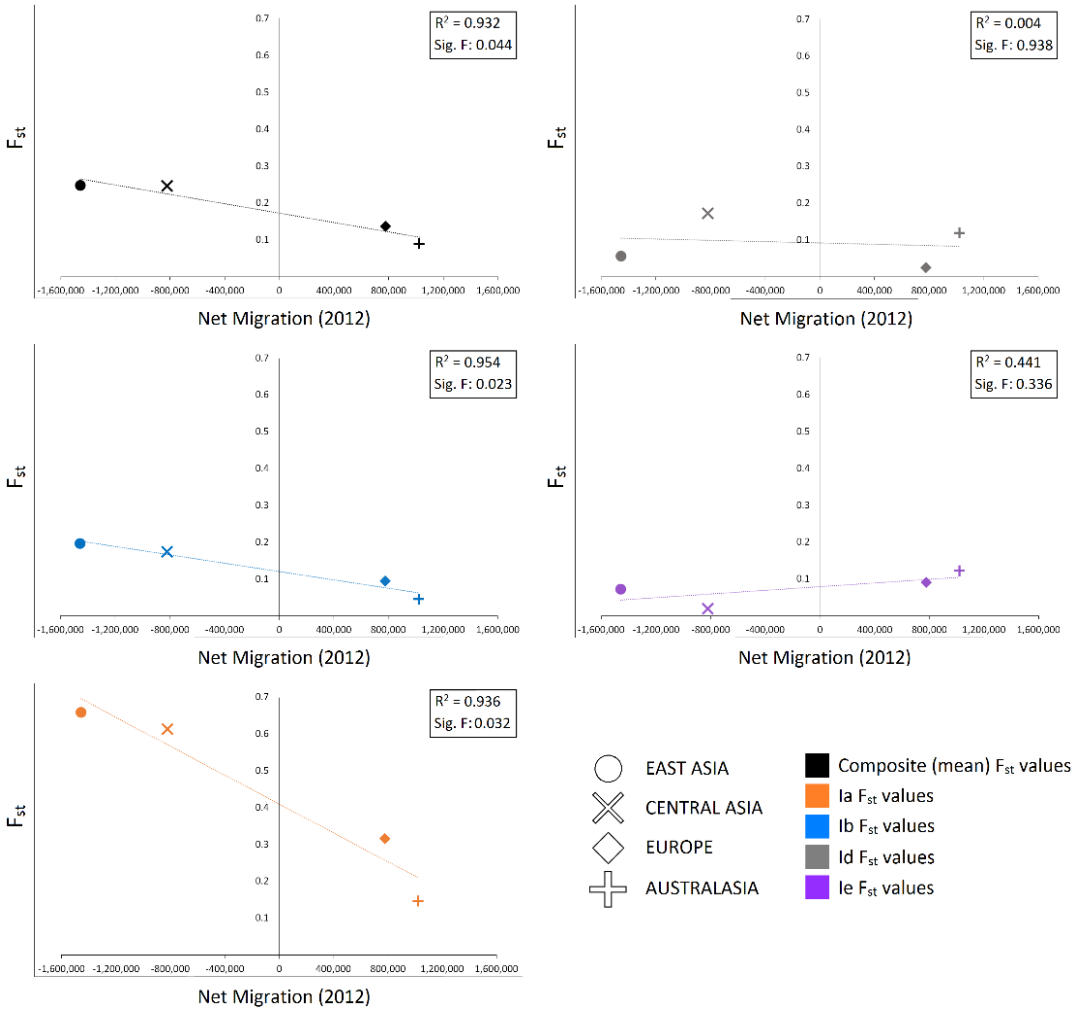


Figure 3.3.7. Regressions of F_{st} values for *C. hominis* GP60 alleles versus regional patterns of human migration

F_{st} values within geographically-defined *C. hominis* GP60 allele groups (Table 3.3.3) were coupled with net migration values defined for the same geographic regions, and statistically significant correlations were assessed using standard regression curve analyses (R-squared and F-test P-values, Sig. < 0.05). Net migration estimates from the last global census in 2015 indicate regional values for immigration (positive) and emigration (negative), and provides a generalised notion of the degree of diversification amongst the described geographic regions (The World Bank, 2015).

(vii) Intra-subtype phylogeography reveals potential geographic overlaps of closely-related allele types for GP60 subtypes Ia and Ib

C. hominis intra-subtype phylogeography of GP60 alleles was assessed across the five subtype families Ia, Ib, Id, Ie, and If. In this analysis, the differentiation of allele frequencies (G_{st}) was compared to nucleotide similarities of the alleles (N_{st}), taking into account all intra-subtype geographic comparisons (Appendix XVIII). This analysis revealed a number of differential associations between genealogy and geographic distribution. An initial evaluation of the spread of G_{st}-N_{st} values within subtype families revealed that closely related Ia, Ie, and particularly Ib GP60 alleles have the tendency to overlap geographically. This would suggest that the rate of gene flow is lower than the mutation rate for these subtypes. In contrast, subtype Id alleles were more likely to occupy separate defined geographic regions (Figure 3.3.8). Subtype If showed a relatively even distribution around the mean and contained relatively equal degrees of phylogeographic overlap versus differentiation. This analysis was performed more formally to identify geographic comparisons with particularly significant phylogeographic associations ($\mu \pm 2SD$). Hence, normal distribution curves for G_{st}-N_{st} values were analysed (Appendix XII), which showed an absence of statistical significance for a vast majority of the geographic intra-subtype comparisons (94.1% non-significant). However, significant phylogeographic correlations were found across the various subtype families for a limited number of inter-regional comparisons. For subtype Ia alleles, isolates from North Africa and South America appear to show geographic overlap of closely-related alleles ($N_{st}>G_{st}$). For subtype Ib alleles, the only statistically significant phylogeographic association was identified between Central Asian and European isolates ($N_{st}>G_{st}$), while subtype Id showed a similarly significant geographic overlap of closely-related alleles from Australasia and Central Asia, and an absence of geographic overlap ($N_{st}<G_{st}$) for Australasian and South American alleles. Subtype Ie alleles contained two comparisons with statistically significant $N_{st}>G_{st}$, namely between East Asian and North African isolates, and between East Asian and Sub-Saharan African isolates, while subtype If contained no statistically significant phylogeographic comparisons. Significant associations were more likely to follow a pattern of overlap between phylogeny and distribution ($N_{st}>G_{st}$; 71.4% of significant comparisons), which correlates with the frequently homogenous population structure of worldwide isolates, as indicated by the predominantly insignificant F_{st} values

(69.5% of regional comparisons across subtypes). Both results suggest that genetic migration is a less important evolutionary force than the mutation rate for the allelic and nucleotide diversity of GP60.

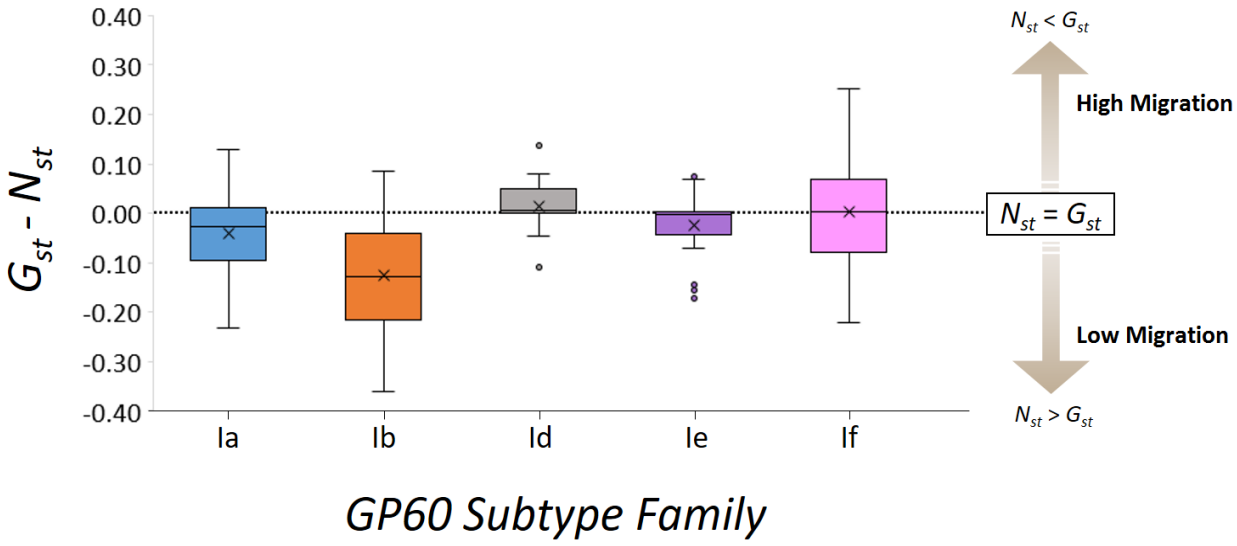


Figure 3.3.8. Intra-subtype phylogeography of global *C. hominis* GP60 alleles

Differentiation of allele frequencies (Gst) was quantitatively compared to allelic similarities (Nst) of all pairwise geographic comparisons (Australasia, Central Asia, East Asia, Europe, North Africa, North America, South America, and Sub-Saharan Africa), The figure illustrates the associations between phylogeny and geographic grouping of allele types (= phylogeography). The greater Nst > Gst, the more likely closely related alleles are to co-occupy two geographic regions under comparison; alternatively, Nst < Gst indicates that closely related alleles are less likely to coincide geographically. Nst = Gst demonstrates an absence of phylogeographic associations between intra-subtype allele-types (Lowe *et al*. 2004).

3.4 DISCUSSION

The present study describes the first comprehensive analysis of population genetics for *C. hominis* GP60 alleles from a global perspective, providing new insight into the genetic drivers of strain diversification for this species. One of the most significant findings from the divergence analyses was a substantial difference in signals of selection and population diversities between the various subtype families and geographic regions. Isolates belonging to subtypes Ia and Id, specifically those originating in Central Asia and Sub-Saharan Africa, revealed the highest degrees of nucleotide diversity and selective pressure, and were associated with the most significantly negative Tajima's D values. This points towards a recent population expansion event for these isolates, possibly following a bottleneck event. The ecological basis for this apparent selective sweep amongst certain geographically-defined isolates is difficult to infer in this instance, as the potentially relevant epidemiological and clinical data is too scarce and sporadic to establish any significantly associated drivers of adaptation. One possible explanation is that these subtypes have been introduced to certain regions more recently than others, and that heightened selective pressure may be reflective of an ongoing adaptation event between the parasite and its newfound host population. From another speculative perspective, part of the answer may derive from largescale historical immunological changes in certain human populations, and the capacity of certain allelic variants to rapidly adapt to said changes with a new heightened degree of pathogenicity and survival. Immunology and specifically a compromised versus competent immunity have long been associated with a differential propensity for *Cryptosporidium* to cause disease in humans (O'Connor *et al*. 2011), and hence a rise in immunodeficiency amongst a substantial proportion of a population might induce a previously avirulent allele type to rapidly adapt and expand. The significantly higher intra-allelic divergences and positive selective pressures that were identified for subtypes Ia and Id in Central Asia and Sub-Saharan Africa may therefore be reflective of a comparatively recent and novel divergence event, and possibly one that correlates with changes in human immunity (Rocha *et al*. 2006; Weber *et al*. 2014). In reviewing recent worldwide epidemics with disproportionately higher immunological impacts on resource-poor regions such as Central Asia and Sub-Saharan Africa, the outbreak and spread of HIV/AIDS since the 1970s stands out. Sub-Saharan Africa and Central Asia have consistently harboured the first and second highest proportions of

global cases, together accounting for 74.8%, 87.8%, and 83.4% of people living with HIV in 1990, 2000, and 2015, respectively (UNAIDS Global Report, 2000 and 2016). If the *Cryptosporidium* parasite experienced an opportunity to improve its evolutionary fitness through the enhanced ability of Ia and Id alleles to cause severe disease and infection in this new and rapidly expanding population of immunodeficient individuals, then this would explain the strong signals for a selective sweep and recent population expansion identified at the GP60 locus. GP60 is further predicted to be GPI-linked, a posttranslational modification that allows for the anchored extension of a protein in the outer cell membrane. This reinforces the accessibility of factors involved in parasite-host interactions, and therefore often links proteins to important roles within virulence and host immunity (Paulick & Bertozzi, 2008).

Fixation indices were used to assess genetic differentiation of subtypes across geographic regions, and together with allelic diversity provided a preliminary view of GP60 population structure amongst global *C. hominis* isolates. Signatures of divergent selection were inferred from significantly elevated F_{st} values amongst regional comparisons, and highlighted a particularly strong signal for divergent selection amongst Central Asian and East Asian isolates belonging to subtype Ib (F_{st} range: 0.33-0.85). This describes genetically-isolated population structures with limited gene flow, and is mirrored by the absence of the widespread and most frequent Ib GP60 allele type "A" in either of these geographic localities (Appendix XIV). Moderate genetic differentiation ($F_{st} \approx 0.3$) amongst Ib alleles was further identified between European and North American, and North American and South American, isolates, but for all of the remaining comparisons low to no genetic structuring could be identified ($F_{st} < 0.2$), which speaks to a largely ubiquitous distribution of Ib allele types worldwide, and complements the relatively low degree of heterogeneity identified across all geographic regions (Appendix VIII). The remaining subtypes (Ia, Id, Ie, and If) showed variable degrees of differentiation between geographic regions, with no single region exhibiting unilaterally high F_{st} values, which could be reflective of the high rates of inter-allelic divergence. This may be particularly true for subtypes Ia and Id, where a higher rate of rare allele types may have blurred any potential outlines of a defined population structure. The differential degree of genetic differentiation across and between the geographic comparisons did give rise, however, to one particular trend: the association between net human migration rates and average F_{st}.

Although the link between human behaviour and the parasite's genetic structure is entirely theoretical in this instance, the idea that higher immigration (or indeed foreign travel) rates would give rise to more parasites entering a region, and hence a greater potential diversity of allele types, is entirely plausible. Subtype Ib exhibited the most significant difference in mean F_{st} values between Central/East Asian samples, and Australasian/European samples (Central Asian mean = 0.643 [95% CI: 0.460 – 0.826]; East Asian mean = 0.689 [95% CI: 0.583 – 0.794]; Australasian mean = 0.023 [95% CI: -0.016 – 0.062]; European mean = 0.113 [95% CI: -0.014 – 0.239]; $p < 0.005$). Net migration may indeed have played a role, but another equally plausible explanation is that subtype Ib is in fact native to the two regions in which it is also the most prevalent, i.e. Australasia and Europe. A relatively recent introduction of type Ib *C. hominis* isolates to Central and East Asia could explain the rather homogenous population structure of isolates, because a short divergence timeframe could have limited its ability to significantly differentiate, or because the opportunity to diverge has been limited as there are other more efficiently-adapted (and more prevalent) allele types already present in the population, such as subtype Ia.

Allelic diversity was used to further characterise the population structure and distribution of *C. hominis* GP60 alleles in this study. Of the 142 allele types identified (defined in this context as intra-subtype alleles that differed by at least one SNP, disregarding indel diversities), Ia alleles showed the highest mean allelic diversity (mean AD = 0.712, range: 0.372-1.000), and If alleles showed the lowest (mean AD = 0.232, range: 0.000 – 0.889), although If sample size was not adjusted for and this almost certainly affected the observed absence of allelic diversity for most geographic regions. Australian isolates exhibited a significantly greater allelic diversity across all subtypes (mean AD = 0.909, range: 0.862 – 1.000; $p < 0.05$), while South American isolates showed the lowest mean allelic diversity (mean AD = 0.309, range: 0.000 – 0.719), but this was not statistically significant compared to the remaining six geographic regions ($p > 0.05$). The reason for this apparent excess of allelic diversity in Australian *C. hominis* isolates is difficult to assess in the absence of associated epidemiological or clinical data, and could not in this study be attributed to sample size (following the assumption of "more sequences = more potential for diversity"). However, it is interesting to note that the geographic region with the highest diversity simultaneously was the region with the highest net immigration

rate. In the most recent global census, a net immigration rate of 10.3 migrants per 1,000 inhabitants was estimated for Australasia, a figure which is more than double that of North America (4.76 immigrants/1,000 population), which had the second highest net migration rate (The World Bank, 2015). South America, on the other hand, had a net emigration rate of -2 migrants per 1,000 inhabitants, which comprised the highest regional emigration rate worldwide. The phenomenon of immigrants, refugees, and travellers importing intestinal parasitic infections has repeatedly been demonstrated, with prevalences of up to 66% identified in newly-arrived immigrants from East Africa and Southeast Asia to Australia (Caruana *et al*. 2006), and a 3-30 times higher notification rate for intestinal protozoa in immigrants and refugees to Europe (Ekdahl & Andersson, 2005). In one *Cryptosporidium*-specific example, a prospective study in Sweden found that more than 83% of positive samples were from travellers, versus only 17% in non-travellers. This provides some level of support that human movement may impact global patterns in GP60 diversity, although the significant limitations of sample size inherent in this study means that a much greater and diverse sample set is required in order to truly trace the genetic origins and spread of *Cryptosporidium* parasites across the world, both past and present.

Another defining feature of *C. hominis* GP60 subtype families was the existence of two predominant inter-allelic subgroups, designated allele types A and B, which revealed themselves through clusters on the radial phylogenies (Appendices XIII-XVII). This study sought to investigate the global relationship between geography and said allele-types, but instead found that there were few consistent associations to be made. For subtypes Ia and Ib, both predominant allele types were present (albeit at variable proportions) in most of the geographic regions identified (Appendices XIII & XIV), although it important to note that the most common *C. hominis* allele type worldwide, Ib type A (commonly referred to as IbA10G2), was entirely absent from Central and East Asian samples. For subtypes Id and Ie, allele type A was ubiquitous, but for allele type B a more stringent distribution pattern seems to emerge. Id type B was only found in Australasian, Central Asian, European, and North American samples, and Ie type B was found only in Australasian, East Asian and North American samples. Subtype If showed completely divergent patterns across regions for which sequences were available, but sample size was again such a limiting factor for this (and indeed all) GP60 subtype families that the

reliability and interpretability of these results remain constrained. But should these preliminary trends hold true for more expansive sequence datasets, the question remains as to whether there is a biological basis (such as host differences) that drive intra-subtype allele population formations.

Exhaustive characterisation of intra-subtype allele types also served to evaluate how effectively current subtyping approaches, which are largely based on GP60 repeat elements, describe true phylogenetic distances between global *C. hominis* isolates. For many of the notorious (i.e. most frequently-reported) *C. hominis* GP60 subtypes described in humans, such as **IbA9G3** (Gatei *et al*. 2007; Jex, *et al*. 2008; Geurden *et al*. 2009; Hijawi *et al*. 2010; Waldron *et al*. 2011; Flecha *et al*. 2015), **IdA15G1** (Gatei *et al*. 2007; Areeshi *et al*. 2008; O'Brien *et al*. 2008; Ng *et al*. 2010; Hira *et al*. 2011; Koehler *et al*. 2013; Mbae *et al*. 2015), **IeA11G3T3** (Waldron *et al*. 2009; Ajjampur *et al*. 2010; Ye *et al*. 2012; Valenzuela *et al*. 2014; Eibach *et al*. 2015; Mbae *et al*. 2015; Peralta *et al*. 2016), and **IfA12G1** (Hadfield *et al*. 2007; Jex & Gasser, 2008; O'Brien *et al*. 2008; Ng *et al*. 2010; Abu Samra *et al*. 2013; Mbae *et al*. 2015), conventionally-subtyped isolate groups produced repeat-based phylogenies that were in agreement with those resulting from the 360bp consensus sequences. This indicates a strong homogeneity within some of the major described GP60 subtype groups, and supports the use of traditional subtyping approaches for understanding allelic diversity within global *C. hominis* populations. One very important and significant exception to this, however, was identified for the most prevalent and well-known *C. hominis* GP60 subtype, **IbA10G2**. This study found that more than 15% of sequences labelled online as IbA10G2 did not follow the expected SNP pattern for the predominant allele type 'A', and that certain alleles were more divergent from type 'A' than was allele type 'B', or IbA9G3 (Figure 3.4.1). Although these sequences did in fact possess the 'IbA10G2' nucleotide repeat profile, and the designation as such is therefore not flawed, the potential for misinterpretations and confusion outside of the *Cryptosporidium* community is a valid concern. Individuals who are unfamiliar with the approach to GP60 subtype classification may be misled by the excess reporting of subtype IbA10G2 worldwide, and incorrectly assume that this shared subtype designation indicates allelic homogeneity amongst sampled populations. Even more troubling is the closer proximity of certain IbA10G2 subtypes to non-IbA10G2 subtypes. The need for new and improved genotyping markers for *Cryptosporidium* isolates is not unique to *C.*

hominis, but applies to all human-infective species and subspecies, and is a need which has already been highlighted by past research (Widmer & Lee, 2010). The release of new whole genome sequences in higher quantities will undoubtedly help fill this need, as they will provide a much-anticipated insight into the true allelic diversity at this and other loci, and help us understand how allelic divergence correlates with pathogenicity and virulence on a larger scale.

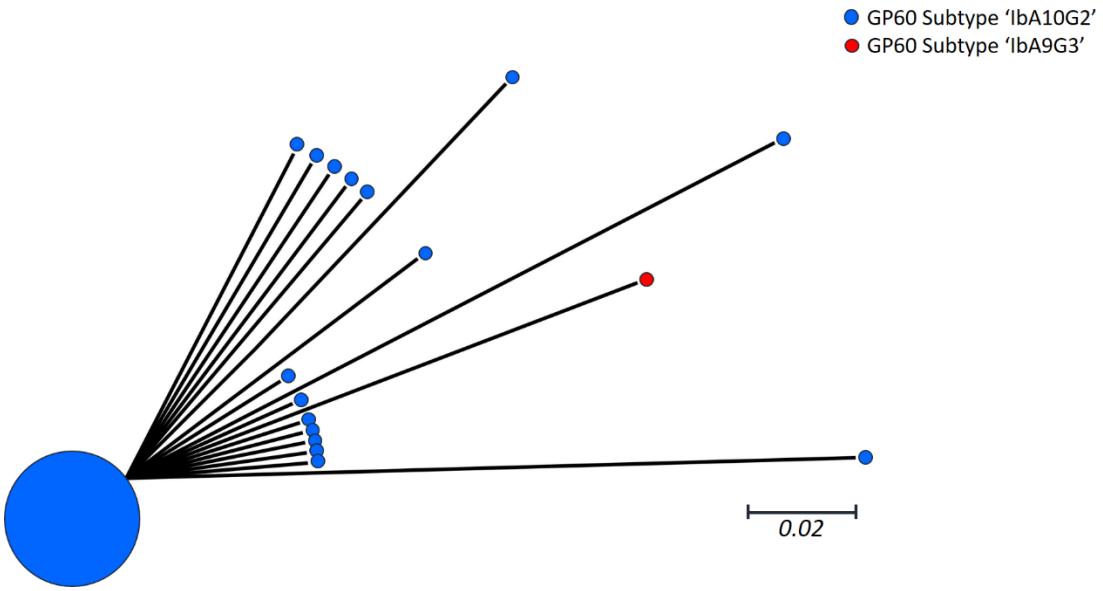


Figure 3.4.1. Maximum Likelihood Phylogeny of 'IbA10G2' sequences reveals lack of genetic identity between GenBank submissions
All previously-submitted *C. hominis* GP60 sequences labelled as "IbA10G2" on GenBank (http://www.ncbi.nlm.gov/genbank/) were retrieved and edited to achieve a 303bp consensus sequence, and aligned through ClustalW using MEGA v 7.0 (Kumar *et al*. 2016). The phylogenetic relationship between sequences was subsequently analysed using the Tamura-Nei model to construct a Maximum Likelihood tree, estimating pairwise distances using the Maximum Composite Likelihood (MLC) method.

A final point to discuss in this chapter is the improved epidemiological classification of *C. hominis* GP60 subtypes that was achieved. For the first time, systematic review of all available GP60 sequences has been used to characterise the geographic distributions and host types of subtype families on a global scale. Subtype Ib was reaffirmed as the most frequently-isolated GP60 subtype in European and Australasian samples (Jex & Gasser, 2008; Waldron *et al*. 2009), and is the overall most prevalent subtype characterized to date, accounting for more than half (59%) of the 2,972 published *C. hominis* GP60 sequences reviewed. This reaffirms subtype Ib isolates as the leading cause of sporadic human cryptosporidiosis cases due to *Cryptosporidium hominis* worldwide, as has been previously reported (Xiao *et al*. 2010; Waldron *et al*. 2011). The overlay of epidemiological and population genetics data becomes somewhat telling here, as the most common

subtype Ib largely revealed lower levels of nucleotide diversity and selective pressure than other subtypes, particularly in samples isolated outside of European and Australasian regions (Appendix VIII). This phenomenon may be explained, in a counterintuitive way, by the vast predominance itself; by hypothesizing that the most widely distributed and hence potentially the most pathogenic GP60 subtype isolates would achieve said success due to an already optimized host adaptation, and hence has little evolutionary incentive to diverge. The identified link between prevalence and divergence is further not unique to *Cryptosporidium hominis*, or in fact protozoans, as the most widespread and prevalent HTLV (Human T-cell Lymphotropic Virus) subtype "A" similarly exhibits the lowest overall divergence amongst all global subtypes at the hypervariable GP21 locus (Angus & Weiss, 1999). Geographic data also reaffirmed subtype Ia as being particularly present in developing regions, with the largest proportion of samples (~70%) originating from Sub-Saharan Africa, South America, and Central Asia (Figure 3.3.2). Subtype Ie closely mirrored this trend, with more than three-fourths (>75.0%) of samples originating in Sub-Saharan Africa, Central Asia, and South America. In addition to distribution data, host range patterns have also been expanded and refined through this epidemiological review. *C. hominis* subtypes Ii and Ik demonstrated predominantly zoonotic host ranges, having most frequently been isolated from equine hosts in the Eastern Asian region (specifically China). This may reflect a recent host switch from humans that is unique to this region. Evaluating the genomic changes between these zoonotic and other anthroponotic *C. hominis* isolates may provide insight into the nature and chronology surrounding this jump from humans to equines, and could be an important priority for future work. Subtype Ib demonstrated the broadest host range of the classically anthroponotic subtypes, having been isolated from ten different host species to date. A number of reports exist that describe subtype Ib as more virulent than other subtypes (Ng *et al*. 2010; Li *et al*. 2013; Ng-Hublin *et al*. 2017), and this heightened virulence may thus account for the broader range of hosts this subtype seems to infect. Host range data for subtype Id seems to complement this trend, as this similarly 'more virulent' subtype (Cama *et al*. 2007) also exhibited a broader host range compared to the remaining anthroponotic subtypes. In direct contrast to this, subtype Ia which frequently is classed as less or 'avirulent' (Mbae *et al*. 2015) had a much narrower host range, isolated from humans only with the exception of three cases in rhesus macaques. This clear distinction in terms of distribution, host range, virulence and allelic diversity

between certain subtype families, such as Ia and Ib, means it is important to understand the types and degrees of genomic changes that may account for the observed phenotypic variations. Comparative genomics studies are therefore key to understanding the genetic basis of phenotypic traits that define and divide *C. hominis* subtypes and/or subspecies.

Genomic comparison of *C. hominis* subtypes Ia and Ib reveals the major modalities of lineage-specific evolution

4.1 INTRODUCTION

The aim of this study was to perform an exhaustive comparative genomics review of two *C. hominis* WGS from different GP60 subtype families. Leading on from the previous chapter, this study intended to explore GP60 variability within a genomic context, but also to assess how this and other hypervariable loci (or those under significantly heightened selective pressure) may contribute to phenotypic disparities within *C. hominis*. This whole genome comparison specifically focused on the GP60 subtype families Ia and Ib. Inclusion of these specific subtypes was based on a number of criteria, such as the worldwide predominance of subtype Ib and the fact that the original *C. hominis* TU502 reference genome belonged to subtype Ia, but the most compelling criteria was the seemingly different host ranges and global distributions described in chapter 3 (Figure 4.1.1). Although both are predominantly anthroponotic, Ia has infrequently been isolated from rhesus monkeys, while Ib has been found in rhesus monkeys as well as baboon, cattle, chimpanzee, deer, flying fox, kangaroo, sheep, and wallaby hosts. This associates subtype Ib with a tentatively broader host range, and warranted investigating whether any genotypic disparities could explain the increased (albeit limited) propensity for zoonotic infection. The predominance of subtype Ia in resource-poor settings (Central Asia, South America, and Sub-Saharan Africa), compared to the more developed origins of Ib isolates (Europe and Australasia) was also of interest, as genomic changes such as elevated selective pressure could be footprints left behind from recent population-defined host adaptation events. As such, this study aimed to first broadly describe divergence in terms of the degree and genomic location of significantly elevated SNP rates and positive selective pressure. Divergence was subsequently dissected to explore specific features such as protein function, telomeric proximity, and drivers of variation (SNPs versus gaps), as well as identify subtype-specific coding regions.

By illustrating and describing these features, this study provides the first comprehensive overview of the genomic differences that separate *C. hominis* subtypes Ia and Ib, and explores how divergence may influence and drive some of the variation in phenotype and epidemiology previously observed and described.

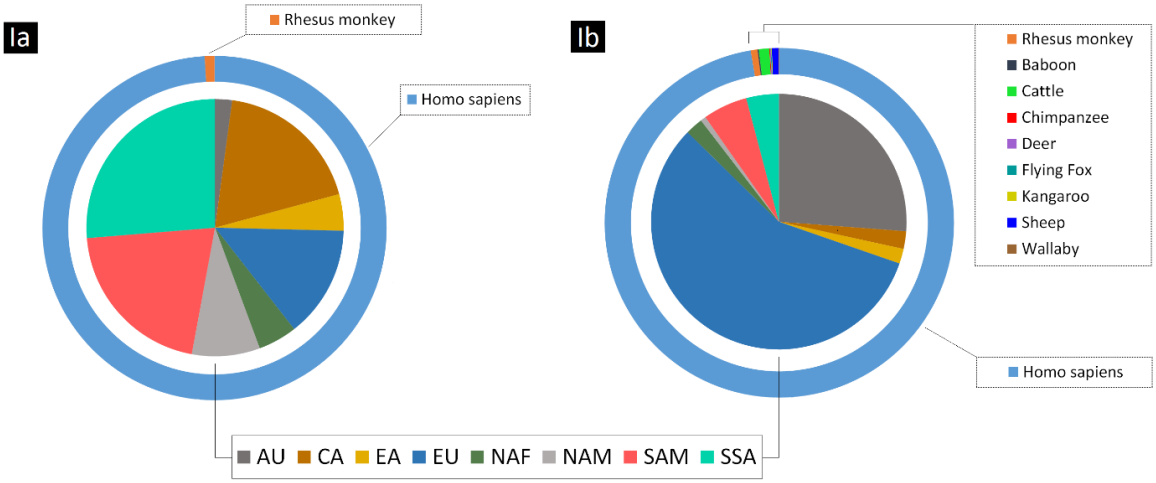


Figure 4.1.1. Host range and geographic distribution of Ia and Ib-type GP60 alleles reveals phenotypic and epidemiological differences

Host types and geographic origins of *C. hominis* GP60 sequences belonging to subtype families Ia and Ib are shown, adapted from figure 3.3.1.

A number of studies have reported variable virulence and infectivity between certain *C. hominis* GP60 subtype families, as discussed in the previous chapter, but the potential genomic features driving these and other phenotypic differences have yet to be explored. This limits the data that currently exists to describe the degree and nature of divergence between *Cryptosporidium hominis* isolates. Molecular characterization has overwhelmingly focused on two genetic targets, the highly-conserved 18S rRNA gene and the hypervariable GP60 gene. The former has primarily been used to differentiate *C. hominis* from other *Cryptosporidium spp*., rather than facilitate any phenotype-associated molecular discrimination within *C. hominis*, while the latter describes and explores genetic heterogeneity in diverse *C. hominis* populations, and designates isolates into distinct subtype groups (Figure 4.1.2). Although these targets are used for entirely contrasting purposes, together they provide an insight into the co-existence of conservation and variability within *C. hominis* populations. This provides an important incentive for exploring genome-wide features between divergent *C. hominis* GP60 subtype families. Other frequent targets for molecular characterization of *C. hominis*

include COWP (*Cryptosporidium* oocyst wall protein), heat shock proteins (primarily HSP70), microsatellite regions, CP47 (glycoprotein), and β-tubulin (Table 4.1.1), but a majority of the associated studies provided molecular discrimination of single targets rather than using a multilocus approach. Furthermore, studies focused on *C. hominis* tend to be geographically confined, and only rarely have associations between clinical and genetic features been performed. The predominant use of GP60 as a typing tool has also confounded our understanding of *C. hominis* genome diversity, due to its inability to provide an accurate portrayal of genomic proximity between certain strains. Describing the specific genomic changes between varying GP60 subtype WGS is therefore key to understanding how distinct *C. hominis* lineages have arisen through time. The genomic comparison will also identify potential drivers of phenotypic differences, such as unique virulence patterns and host niches, and will hopefully reveal novel markers for molecular typing of strains in a manner that correlates site-specific polymorphisms with whole genome divergence.

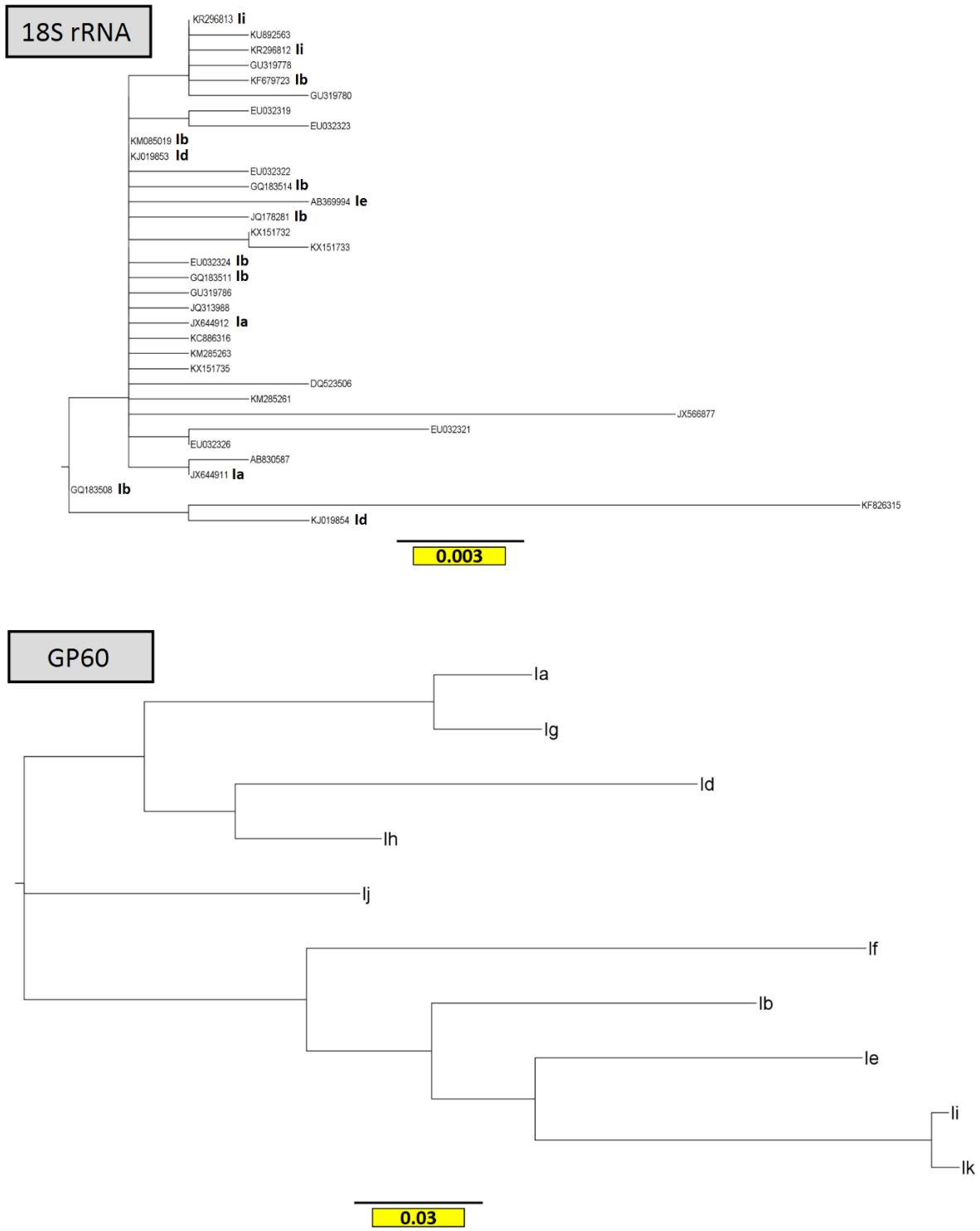


Figure 4.1.2. Maximum likelihood trees for 18S rRNA and GP60 sequences reveal dissimilar phylogenetic grouping and distances between *C. hominis* isolates

Maximum likelihood trees for *C. hominis* 18S rRNA (top tree) and GP60 (bottom tree) sequences are shown. All 18S rRNA sequences for *C. hominis* deposited to GenBank were retrieved (https://www.ncbi.nlm.nih.gov/genbank/), and single examples of genotypes were edited to achieve a 734bp consensus (N=34) and aligned using the alignment editor in MEGA v 7.0 (Kumar *et al*. 2016). A single representative sequence for each *C. hominis* GP60 subtype family (N=10) was also retrieved and a 976bp consensus sequence was similarly aligned. Where available, descriptions of GP60 subtype families of 18S rRNA-sequenced isolates are included. Maximum likelihood phylogenies were created to obtain evolutionary distances for the respective trees, according to the Tamura-Nei model and the Nearest-Neighbour-Interchange heuristic method (MEGA 7.0; Kumar *et al*. 2016). Branch lengths are illustrated in units of number of substitutions per site between consensus sequences, as indicated by the scalebars. Final tree figures were produced using the FigTree graphical phylogeny editor (http://tree.bio.ed.ac.uk/software/figtree).

Table 4.1.1. Molecular typing and sequence submission of *C. hominis* isolates has predominantly relied on 10 genetic targets

Genomic DNA/RNA GenBank submissions for *Cryptosporidium hominis* (N = 1,393) were extracated and reviewed, to quantitatively evaluate the ten most frequent molecular targets used (https://www.ncbi.nlm.nih.gov/genbank/).

Molecular Target	Number of GenBank Nucleotide Submissions
GP60	648
18S rRNA	238
COWP (Oocyst wall protein)	97
Heat shock Protein	92
Microsatellite	87
CP47 (glycoprotein)	52
Beta Tubulin	44
Serine repeat antigen	39
Hydroxyproline-rich glycoprotein precursor (DZHRGP)	18
Actin	14

In addition to the largely GP60-biased approach to molecular exploration of intra-*C. hominis* heterogeneity (Table 4.1.1), only two whole genome comparative studies between *C. hominis* strains exists (Isaza *et al*. 2015; Guo et al. 2015). The primary focus in Isaza *et al.* was the release of a new *C. hominis* WGS genome, along with an improved annotation of the already-published *C. p. parvum* Iowa II reference, and as such, the whole genome comparisons between *C. hominis* strains was a minor component. The strains included in this comparative analysis were the IbA10G2 isolates "UKH1" and "TU502 New" (both available on CryptoDB), and a novel WGS belonging to subtype IeA11G3T3 ("UdeA01"; BioProject ID: PRJEB10000). The results from this comparison were limited to the delineation of the genome-wide proportion of non-synonymous versus synonymous SNPs, and identifying polymorphic loci based on regions that exceeded a pre-determined threshold for SNP density (2.1 SNPs/kb), calculated based on the sum of two standard deviations of the median for SNPs/kb across protein coding sequences. Although loci with higher-than-average SNP rates were further annotated based on gene ontologies, Isaza *et al*. (2015) did not elaborate on the chromosomal locations (such as telomeric proximity) or evaluate selective pressures and potential recombination signals. There was little attempt to highlight epidemiological differences between the included *C. hominis* GP60 subtype families or to associate phenotypic differences with specific divergence traits. The potential biological implication of genes containing "higher-than-average" SNP rates was also only briefly discussed. Guo *et al.*

(2015) provided more detail in their comparative genomics approach, discussing hypervariable virulence factors from a functional perspective and annotating possible recombination events between *C. hominis* subtypes Ia and Ib, but selective pressure was not explored and genomic differences were only discussed in a phenotypic context to a limited degree. This means there still exists a significant need to perform enhanced, comprehensive comparative genomics studies, so that the molecular determinants of lineage-specific traits between and within *C. hominis* subtype families can be described.

4.2 METHODS

(i) *C. hominis* WGS – empirical data

The genomes of two *C. hominis* isolates belonging to GP60 subtype families Ia and Ib were selected for this comparative analysis. These isolates, known commonly as UKH4 and UKH1, belong to subtypes IaA14R3 and IbA10G2, respectively, and were both derived from human clinical samples in the United Kingdom. Methodology describing the DNA extraction, purification, and next generation sequencing of UKH4 has previously been described (Hadfield *et al*. 2015), but for UKH1 there is no protocol available as the WGS was submitted without an associated publication. The whole genome sequences for subtypes Ia (UKH4) and Ib (UKH1) are both publicly available, and were retrieved in FASTA format from their respective BioProject records on NCBI (Accession: PRJNA253838 and PRJNA222837). Sequence data in the form of contigs from each WGS were manually reassembled and rearranged against the *C. parvum* Iowa II reference sequence (http://www.cryptoDB.org; Abrahamsen *et al*. 2004) into eight complete chromosomes, and the two genomes were subsequently aligned using the ClustalW algorithm, all performed in MEGA v 7.0.

(ii) Computational approaches for comparative genomics

Intergenomic analyses were all performed using DnaSP v 5.10.1 (Librado & Rozas, 2009). Nucleotide (π) and indel (πi) diversities were both assessed by the 'DNA polymorphism' test, and were computed using a sliding window approach for a 1000bp window and 1000bp step size. Microsoft Excel (2013) was used to create databases describing SNP frequencies and chromosome locations, to determine associations between divergence and telomeric proximity. In order to determine genome-wide Ka/ks trends, as well as compare rates of SNPs and indels in coding versus non-coding regions, concatenated sequences of predicted open reading frames (ORFs) were prepared. A minimum ORF size of 300bp (100 amino acids) was selected, as this arbitrary cut-off level has previously been implemented for extracting coding sequences from whole genome coding datasets (McDougall, 1990; Oliver *et al*. 1992; Swanson, 2001). All putative coding sequences above this 300bp threshold were automatically extracted from WGS in FASTA format,

excluding stop codons (EMBOSS GetOrf; http://www.bioinformatics.nl/cgi-bin/emboss/getorf), and manually concatenated chromosome-by-chromosome in the 5' -> 3' direction. Species-specific ORFs were manually identified and removed, and final concatenations were aligned according to the ClustalW codon alignment function in MEGA v 7.0. Ka/ks ratios were calculated using a sliding window approach (150bp window, 75bp step) in the 'Polymorphism and Divergence' function of DnaSP v 5.10.1. Phylogenetic analysis of species-specific coding sequences was conducted in MEGA v 7.0, using the Maximum Likelihood test according to the Tamura-Nei model, with the highest log likelihood topology selected to construct the final tree.

4.3 RESULTS

(i) Whole genome comparison between *C. hominis* subtypes Ia and Ib identifies putative drivers of host specificity and virulence

In order to characterise putative genomic drivers of host-parasite interactions, a whole genome comparison between two phenotypically-diverse *C. hominis* subtypes (Ia and Ib) was performed. Divergence analyses between *C. hominis* subtypes Ia and Ib whole genome sequences revealed only a limited number of loci that showed significant degrees of nucleotide diversity and selective pressure (figure 4.3.1). Only seven of the 1000bp π windows exhibited significantly higher rates of genomic nucleotide diversity (>3SD of mean π), all of which encompassed putative coding sequences previously annotated on CryptoDB (chromosome 1: Chro.10427; chromosome 2: Chro.20050, Chro.20051, and Chro.20052; chromosome 6: Chro.60016, Chro.60138, and Chro.60606; chromosome 7: Chro.70055, and Chro.70447; Table 4.3.1). The window spanning GP60 (Chro.60138) exhibited a four-fold higher diversity compared to the next highest pi-value (0.084 versus 0.021), indicating an incomparable degree of hyperpolymorphism at this locus compared to the rest of the genome. Of the 9 highly-divergent coding sequences, six were annotated as putative mucins or glycoproteins (66.7%), with the remaining three each annotated as a putative cell surface anchor, a protein ligase, and as hypothetical. The sliding window ka/ks analysis revealed only five significant signals of selective pressure (ka/ks>1) across the whole genome alignment. These 150bp windows indicating positive selective pressure were all nested within five of the 9 highly-divergent coding sequences (Chro.10427, Chro.20050, Chro.20051, Chro.20052, and Chro.60606), the most significant of which was identified for Chro.20050 (ka/ks: 5.534).

Table 4.3.1. Summary characteristics for genetic loci exhibiting hyperpolymorphism between *C. hominis* subtypes Ia and Ib

Coding sequences exhibiting >3 SD from the mean SNP rates are characterized here by gene ID (Chromosome no., and CD no. relative to 5'-end), corresponding CryptoDB ID, *C. parvum* ortholog, location relative to the 5'-telomeric repeat, size, and putative protein function (BLAST hits E-value < 10^{-5}; http://www.uniprot.org).

Gene ID	CryptoDB ID	*C. parvum* ortholog	Chromosome	Start (bp)	End (bp)	CD size (bp)	Protein size (AA)	Putative Protein Function
C1_481	Chro.10427	cgd1_3810	1	853504	856230	2727	909	Cell surface anchor
C2_5	Chro. 20050	cgd2_430	2	7711	7115	585	195	Mucin
C2_4	Chro. 20051	cgd2_440	2	6670	6266	405	135	Mucin
C2_3	Chro. 20052	cgd2_450	2	5863	5267	561	187	Mucin
C6_6	Chro. 60016	cgd6_60	6	17239	21948	4710	1570	Glycoprotein
C6_133	Chro. 60138	cgd6_1080	6	264544	263513	1008	336	Glycoprotein
C6_675	Chro. 60606	cgd6_5270	6	1245813	1246295	483	161	Uncharacterized
C7_69	Chro. 70055	cgd7_420	7	130093	124031	5958	1986	Protein ligase
C7_506	Chro. 70447	cgd7_4020	7	936147	930805	5334	1778	Mucin

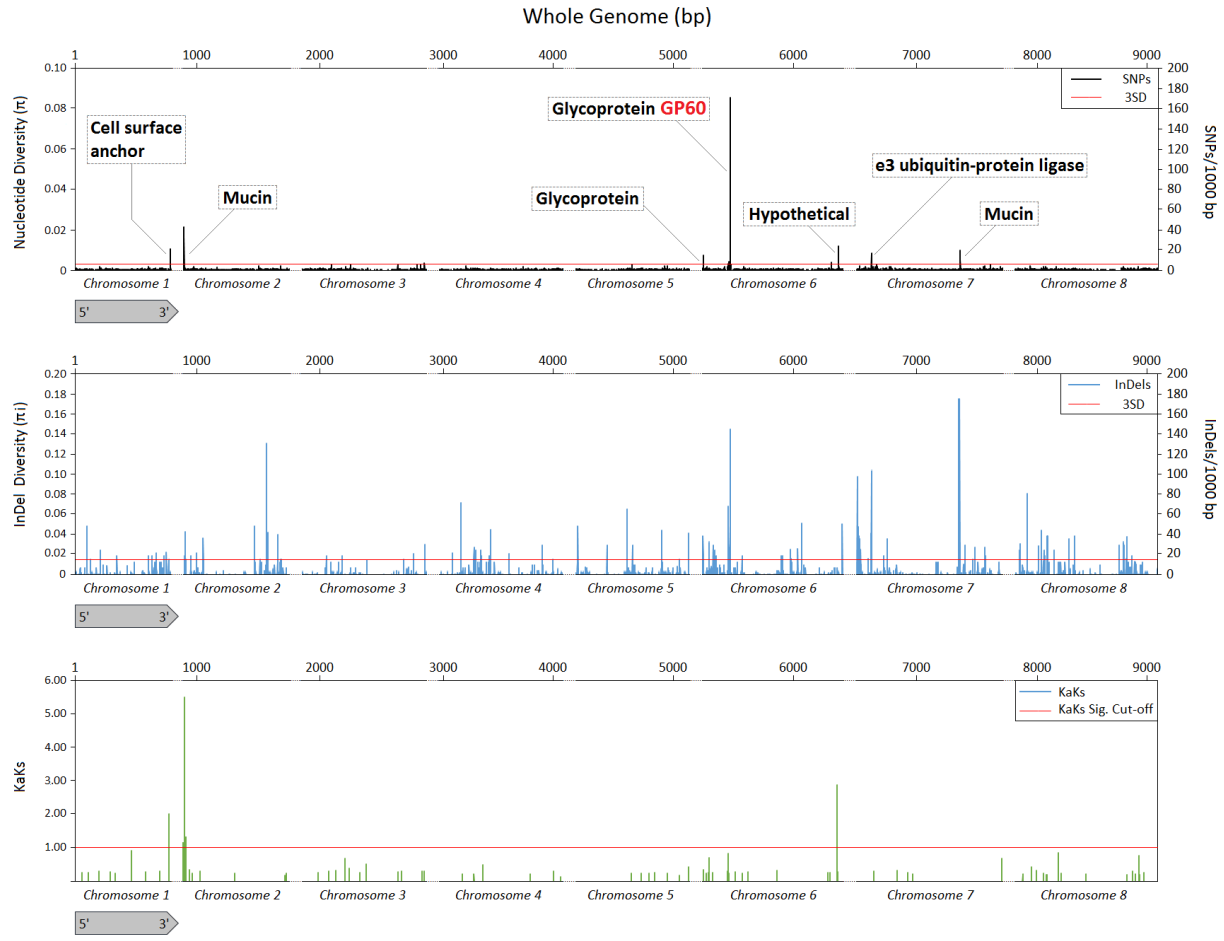


Figure 4.3.1. Sliding window divergence between *C. hominis* subtypes Ia (UKH4) and Ib (UKH1) reveals lineage-specific allelic variation

Sliding window diversity analyses (1000bp window, 1000bp step) were performed on concatenated whole genome alignments (*C. hominis* WGS UKH1 and UKH4) to determine genome-wide nucleotide and indel diversity (top and middle panels, respectively). Sliding window ka/ks (150bp window, 75bp step) was additionally performed on concatenated coding sequence alignments (bottom panel), constructed by extracting all potential CDs based on a 150bp minimum ORF size (EMBOSS GetORF; Rice *et al.* 2000). Significance cut-off values were set at 3 standard deviations from the mean for indel and nucleotide diversity, and at 1.00 for ka/ks. All alignments were created using the ClustalW algorithm in MEGA (Tamura-Nei Model, MEGA v 7.0), and all sliding window analyses were run using DnaSP (DnaSP v 5.10.1; Librado & Rozas, 2009). Putative protein functions of highly-divergent loci were annotated by BLASTing amino acid sequences in UniProt and selecting hits with E-values <10^{-5} (http://www.uniprot.org; UniProt Consortium, 2015).

(ii) Lozalization of genomic divergence between *C. hominis* subtypes Ia and Ib reveals subtelomeric bias

The relative distribution of nucleotide polymorphisms (SNPs) across the genomes was additionally assessed, to determine whether there existed a bias towards higher rates of mutations closer to telomere ends versus the centromeres (Figure 4.3.2). Subspeciation events in other species often exhibit some form of preferential localization of divergence (Carneiro *et al*. 2009; Bishop *et al*. 2000), so it was important to investigate whether any such significant polymorphism trends accompanied the split of *C. hominis* into subtypes Ia and Ib. An assessment of genome-wide SNP distribution revealed a significant, albeit moderate linear regression (R^2 = 0.5884, F. sig < 0.01), indicating an association between decreasing SNP rates and increasing distance from the telomeres, and the proportion of SNPs located within 10% of either telomere end (5' or 3') was found the be statistically significantly higher than the internal regions (75.0 ± 29.6 vs 39.5 ± 7.38, P = 0.0259). Individual chromosomes did not conform to this trend, however, with chromosomes 4 and 5 exhibiting no association between SNP rates and telomere proximity (F. sig > 0.5), and chromosome 6 showing only a weak, non-significant association (R^2 = 0.1107, F. sig > 0.1). There was a statistically significant difference between mean Ka/Ks values within 10% genomic distance of the telomeres compared to the rest of the genome (0.65 ± 0.23 vs 0.34 ± 0.03, P = 0.013), indicating an enhanced positive selective pressure in regions proximal to the telomere ends.

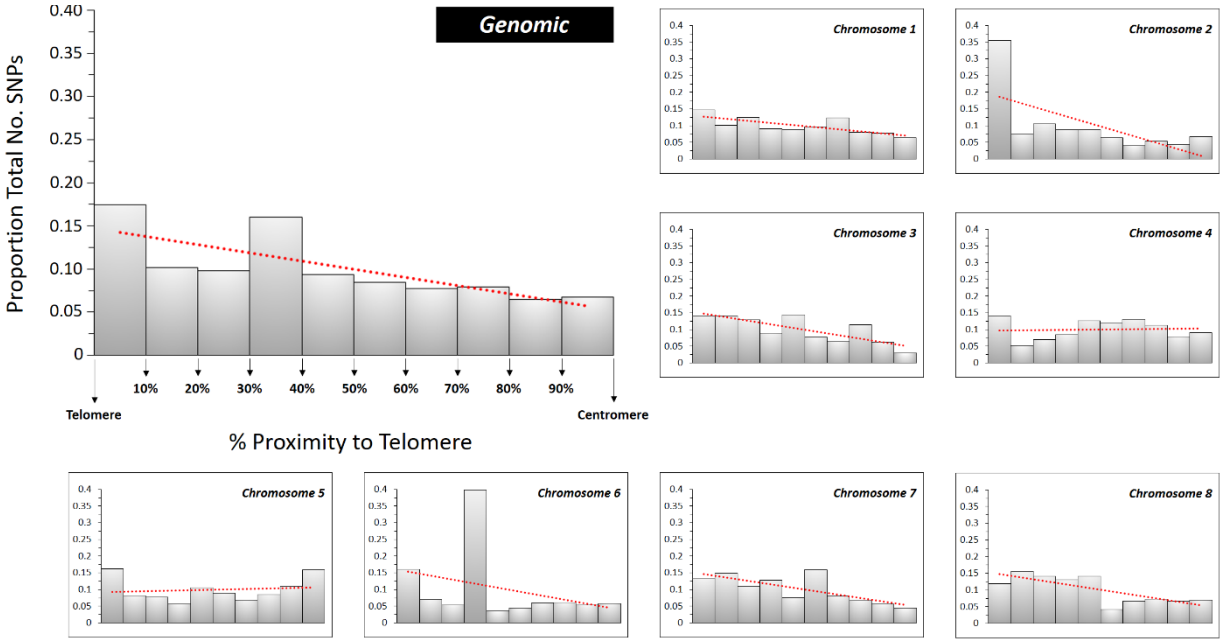


Figure 4.3.2. Distribution of single nucleotide polymorphisms (SNPs) relative to telomere ends show evidence of increased telomeric and subtelomeric divergence

SNPs were categorised into groups based on their proximity to the telomere ends. This was done by creating blocks of % proximity to describe relative position to either the telomere or centromere on the chromosome. SNPs that were located along sequence positions within 10% of the total distance from either telomere end (3' and 5') comprise block 1, SNPs located between 10% and 20% of the total distance comprise block 2, and so on; this continues until block ten, which represents the centromere, or middlemost, portion of the chromosome. Polymorphism data was collected using DnaSP (DnaSP v 5.10.1; Librado & Rozas, 2009), and manually organised into blocks. Regression fit (R^2) and strength (F. sig) were assessed genome-wide and for each consecutive chromosome; genomic (R^2 = 0.5884, F. sig < 0.01), chromosome 1 (R^2 = 0.5607, F. sig < 0.05), chromosome 2 (R^2 = 0.4115, F. sig < 0.05), chromosome 3 (R^2 = 0.6516, F. sig < 0.01), chromosome 4 (R^2 = 0.004, F. sig > 0.5), chromosome 5 (R^2 = 0.0132, F. sig > 0.5), chromosome 6 (R^2 = 0.1107, F. sig > 0.1), chromosome 7 (R^2 = 0.5934, F. sig < 0.01), and chromosome 8 (R^2 = 0.5758, F. sig < 0.05).

(iii) Indel diversity exceeds nucleotide diversity in driving lineage-specific evolution between *C. hominis* subtypes Ia and Ib

Predicted coding sequences made up 76.74% of the genome, and this composition did not differ significantly between the eight chromosomes (mean = 76.66%, 95% C.I.: 75.29-78.04; Appendices XIX-XXVI). Genetic changes were categorised based on nucleotide diversity (π) and indel diversity (πi), to assess patterns of divergence across the genome, and found that single indels occurred at 1.5 times the frequency of single nucleotide polymorphisms (SNPs) for both coding and non-coding regions across the genomes (Figure 4.3.3). This divergence pattern was consistent across all of the chromosomes, with the exception of chromosomes 2, 3, and 6. Two of these (chromosomes 2 and 6) exhibited patterns that reflect the unusually elevated rate of polymorphisms and significant selective pressure previously outlined in figure 4.3.1.

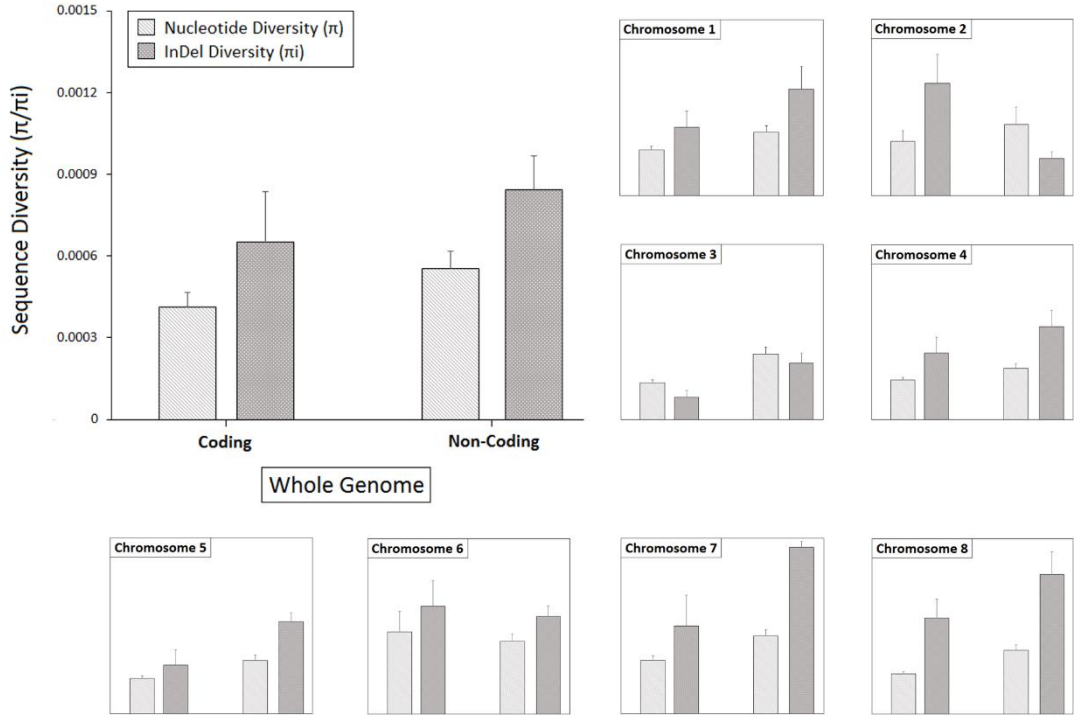


Figure 4.3.3. Distribution of nucleotide (π) and indel (πi) diversity between *C. hominis* subtype Ia and Ib whole genome sequences reveals nature of divergence driving lineage-specific development

The average number of differences per site between *C. hominis* Ia and Ib WGS were defined in terms of nucleotide (π) and indel (πi) diversity, for both coding and non-coding regions. This analysis revealed notably higher diversity rates for non-coding versus coding sequences across most chromosomes, and indels were frequently found to be the cause of more sequence diversity than SNPs. All analysis was performed using DnaSP v 5.10.1 (Librado & Rozas, 2009).

A detailed characterization of indel divergence between coding and non-coding sequences was additionally performed, identifying a strikingly differential nature in the manner indels appear to manifest themselves throughout the genomes (Figure 4.3.4). There was a statistically significant difference for indel event diversity between coding and non-coding sequences ($\pi ie = 0.00005 \pm 0.00001$ vs 0.00035 ± 0.00006, $p < 0.0001$) as well as for mean indel length between coding and non-coding sequences ($\pi ie = 12.33 \pm 2.38$ vs 2.23 ± 0.43, $p < 0.0001$). This reflects a significantly lower incidence of indel events in coding regions, but simultaneously shows that these less common events result in significantly larger gaps and hence substantial changes in ORF sizes. More than 90.0% of gaps in coding sequences greater than 6 bp in size affected microsatellite regions in *C. hominis* Ia and Ib WGS, and hence indels appear to have constituted an important mechanism for regulating the size of amino acid repeats and hence intra-species divergence in protein coding sequences.

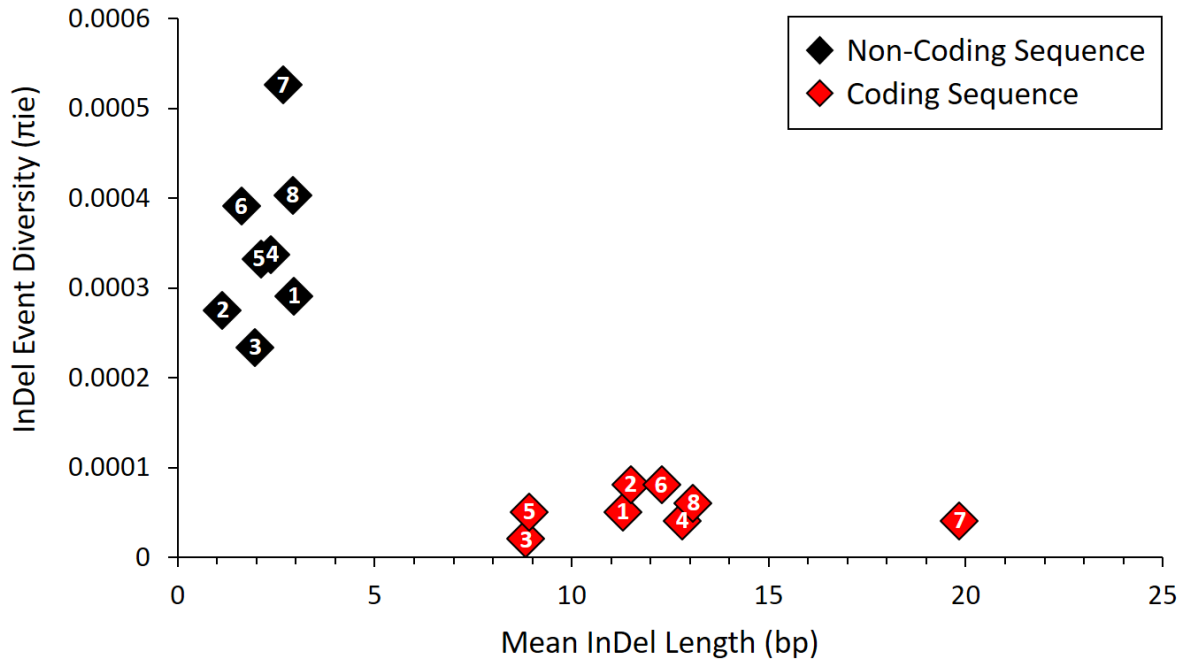


Figure 4.3.4. Indel event diversity (πie) versus mean indel event length (bp) reveal fewer but longer indel events for coding versus non-coding sequences between *C. hominis* Ia and Ib WGS

The nature of indel manifestation between *C. hominis* Ia and Ib WGS was analysed for coding versus non-coding chromosome sequences. Linked sequence gaps were characterised as single indel events and genomic diversity due to these events was subsequently calculated (πie). Mean indel size representing the average base pair length of indel events per chromosome was also determined. Results for coding sequences are illustrated in red, while non-coding sequences are illustrated in black, and data points are linked to corresponding chromosomes (1-8) numerically. All analysis was performed using DnaSP v 5.10.1 (Librado & Rozas, 2009).

(iv) Missing subtelomeres and subtype-specific genes: possible drivers of phenotypic differences between *C. hominis* subtypes Ia and Ib

A total of 5,645 bp of subtype-specific subtelomeric sequence data was identified in *C. hominis* Ia and Ib subtype whole genome sequences, of which 5319 bp constituted deletions at the 5' and 3'-ends of chromosome 2 in *C. hominis* Ia, and 326 bp constituted a deletion at the 3'-end of chromosome 6 in *C. hominis* Ib (Figure 4.3.5). The absence and/or presence of subtype-specific sequence data was confirmed for subtype Ia in the reference *C. hominis* IaA25R3 TU502 whole genome sequence data (http://cryptoDB.org), and for subtype Ib in two novel IbA10G2 whole genome sequences (UKH3 and UKH5) available to this study (source: Cryptosporidium Reference Unit; Singleton Hospital, Swansea, Wales).

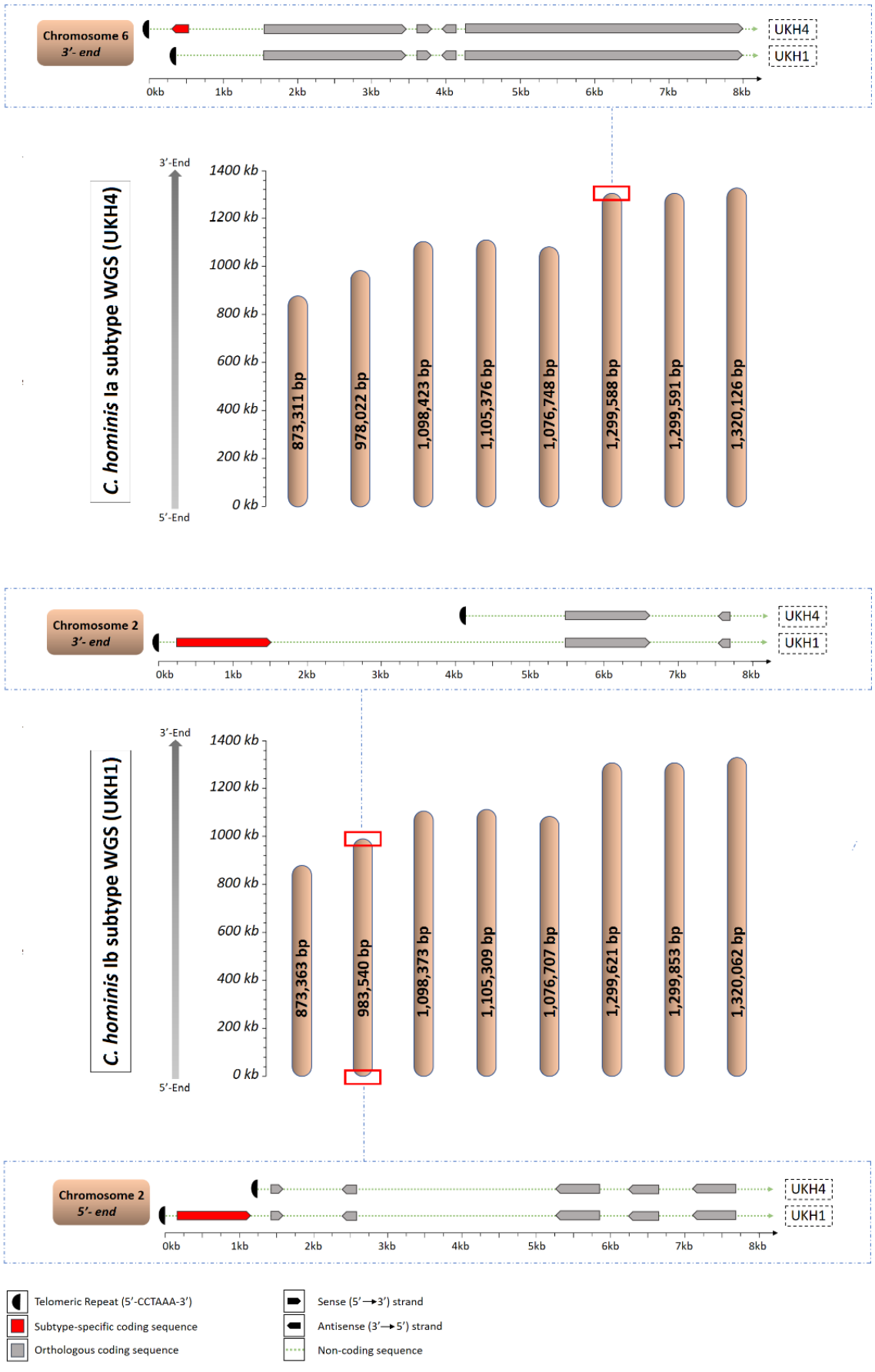


Figure 4.3.5. Missing subtelomeric sequence data encodes subtype-specific coding sequences in *C. hominis* Ia (UKH4) and Ib (UKH1) subtype WGS

Figure 4.3.5. Missing subtelomeric sequence data encodes subtype-specific coding sequences in *C. hominis* Ia (UKH4) and Ib (UKH1) subtype WGS

Whole chromosome alignments of *C. hominis* Ia and Ib genomic sequence data, stretching from 5'- to 3'-end telomeric repeats, were constructed by the ClustalW algorithm and analysed in an alignment viewer to identify missing sequence data >250 bp (MEGA v 7.0; Kumar *et al*. 2016). Subtype-specific sequence data was subsequently extracted from the WGS, and putative open reading frames were annotated using EMBOSS GetORF (Rice *et al.* 2000). Three subtype-specific stretches of subtelomeric sequence data were identified, two of which are specific to *C. hominis* Ib subtype "UKH1" and are located at either end of chromosome 2, and one *C. hominis* Ia (UKH4)-specific sequence at the 3'-end of chromosome 6. Single open reading frames encoding putative genes were identified within each subtype-specific sequence, with variable sizes of 132 bp (UKH4, chromosome 6, 3'-end), 990 bp (UKH1, chromosome 2, 5'-end) and 1263 bp (UKH1, chromosome 2, 3'-end).

Single open reading frames comprising >150 bp were identified within each of the three subtype-specific subtelomeric sequences, with the two *C. hominis* Ib subtype-specific genes encoding significantly larger putative proteins of 330 and 421 amino acids in length. The larger of the two, located at the 3'-end of chromosome 2, was formerly annotated as a *C. parvum*-specific gene (CryptoDB ID: cgd2_4380) due to its absence in the original *C. hominis* Ia subtype TU502 reference genome (Bouzid *et al*. 2013). However, this comparative study revealed that the absence or presence of this gene is variable across *C. hominis*, and attributable to the loss of subtelomeric sequence data in distinct *C. hominis* subtypes. The second Ib-specific gene at the 3'-end of chromosome 2 is a duplicate gene, homologous to Chro.00007, which is located at the telomeric 5'-end of chromosome 5. This Ib-specific homolog is similarly present in a variety of *C. parvum* subtypes (confirmed in *C. parvum* IIa (UKP4), *C. parvum* IId (UKP8), *C. parvum* IIcA5G3a (UKP15), and *C. parvum* IIcA5G3j (UKP16).

In addition to these major subtelomeric deletion events, minor genomic changes resulting in frameshifts similarly resulted in a small number of subtype-specific genes (Table 4.3.2). In total, six subtype-specific coding sequences were identified, five of which (83.3%) resulted in the absence or significant alteration of putative coding sequences (ORFs > 300bp) in UKH4. Putative protein functions of these subtype-specific coding sequences were additionally explored, revealing two possible adhesins, a proteophosphoglycan, and a protein kinase, as well as two uncharacterised (hypothetical) proteins. Subtelomeric location was a predominant feature, with four out of six subtype-specific coding sequences (66.7%) located within the first 15,000 bp extending from the closest telomere end. Phylogenetic relationships were explored for subtype-specific coding sequences in *C. hominis* I and Ib WGS and their orthologous counterparts in a number of zoonotic (*C.*

parvum IIa (UKP4), *C. parvum* IId (UKP8), and *C. meleagridis*) and anthroponotic (*C. parvum IIc* (UKP15)) *Cryptosporidium spp.* (Figure 4.3.6). This illustrated a closer evolutionary proximity of *C. hominis* Ib-specific genes to *C. parvum* isolates than to *C. hominis* Ia for all but one of the phylogenies (UKH4C4_553).

Table 4.3.2. Subtype-specific coding sequences in *C. hominis* Ia (UKH4) and Ib (UKH1) whole genome sequences

Subtype-specific coding sequences are characterized by gene ID (WGS, chromosome no., and CD no. relative to 5'-end), corresponding CryptoDB ID, location relative to the 5'-telomeric repeat, length, underlying molecular cause, and putative protein function (BLAST hits E-value < 10^{-5}; http://www.uniprot.org).

Gene ID	CryptoDB ID	*C. parvum* ortholog	Chromosome	Start (bp)	End (bp)	CD Size (bp)	Protein size (AA)	Genomic impact	Molecular cause	Putative protein function
UKH1C2_1	Unannotated	Unannotated	2	145	1131	990	330	No CD sequence in UKH4	Subtelomeric deletion	Uncharacterised
UKH1C2_224	Chro.20185	cgd2_1710	2	351677	355180	3507	1169	Split into 2 CDs in UKH4	Four InDels	Adhesin
UKH1C2_599	Unannotated	cgd2_4380	2	981996	983255	1263	421	No CD sequence in UKH4	Subtelomeric deletion	Proteophosphoglycan
UKH4C4_553	Chro.40480	cgd4_4210	4	1028823	1031171	2352	784	Split into 2 CDs in UKH1	SNP (C -> T)	Adhesin
UKH1C6_6	Chro.60015	cgd6_50	6	14958	13837	1125	375	Split into 2 CDs in UKH4	SNP (C -> G)	Protein kinase
UKH1C8_1	Chro.80010	cgd8_10	8	96	962	870	290	CD 336bp shorter in UKH4	Single InDel	Uncharacterized

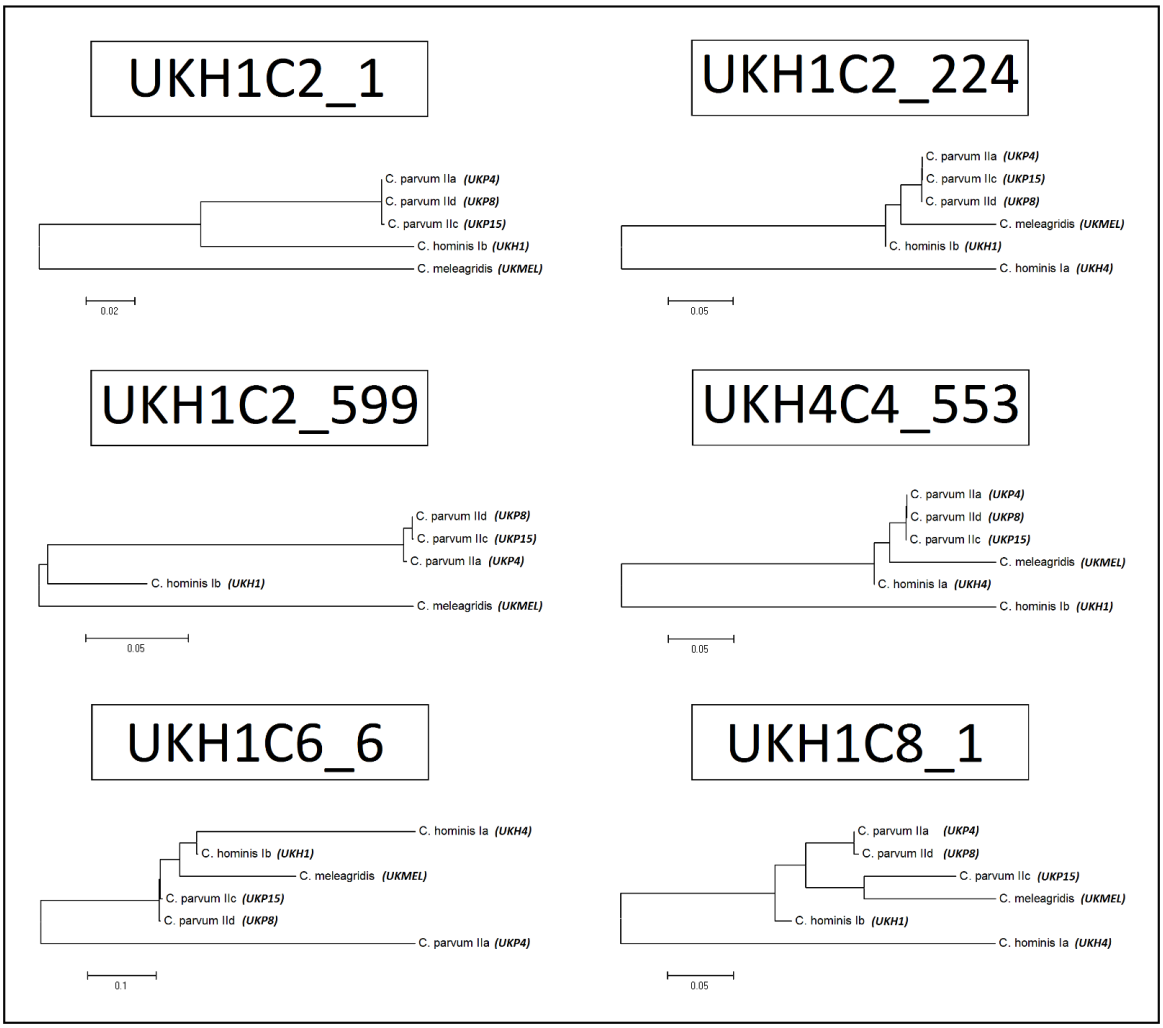


Figure 4.3.6. Maximum likelihood phylogenies of subtype-specific coding sequences in *C. hominis* Ia and Ib WGS

Protein coding sequences for subtype-specific genes in *C. hominis* Ia (UKH4) and Ib (UKH1) were extracted from whole genome datasets, along with corresponding orthologous protein coding sequences in *C. parvum* IIa (UKP4), *C. parvum* IId (UKP8), and *C. parvum* IIc (UKP15), with *C. meleagridis* (UKMEL) included as the outgroup. Amino acid sequences were aligned using the ClustalW algorithm, and maximum likelihood trees were subsequently constructed according to the Tamura-Nei model to illustrate phylogenetic relationships and distances (MEGA v 7.0; Kumar *et al*. 2016).

4.4 DISCUSSION

The current study provides a detailed illustration of the degree, nature, and potential biological impact of divergence between two closely-related *C. hominis* WGS belonging to GP60 subtypes Ia and Ib. Localized, limited genomic divergence appears to driven the evolution and co-adaptation of these separate lineages, and the discovery of subtype-specific coding sequences in either acts as a preliminary explanation for the minor phenotypic differences known to date. A majority of hyperpolymorphic loci were identified within close proximity (<20,000 bp) of either telomere end, which re-iterates the potential importance of subtelomeric location for proteins associated with lineage-specific evolution and adaptation. Selective pressure was also found to act strongest on subtelomeric loci, with three significant ka/ks signals (>1) affecting loci directly adjacent to telomeric repeats (Figure 4.3.1). These high subtelomeric evolutionary rates have been a common feature of many studies on protozoan divergence and adaptation (Kuo & Kissinger, 2008), particularly due to the frequent clustering of important virulence factors and host-parasite interactors in multi-copy form near the telomeres (Freitas-Junior *et al*. 2000; Barry *et al*. 2003; Dreesen *et al*. 2007). Surface antigens in protozoan parasites such as *Plasmodium spp*. and *Trypanosoma spp*. are particularly significant causes of heightened subtelomeric divergence, and this phenomenon has potential relevance for one of the major findings of this study – namely, the discovery that the highest rates for selective pressure were identified for three consecutive mucins at the 5'-end of chromosome 2. These mucins lie amongst an array of seven, all of which are annotated as putative mucins or surface glycoproteins, and the selective manifestation of hyperpolymorphism in some but not all of these genes could imply variable expression that is linked to host immune responses. The discovery that several other glycoproteins and a cell surface anchor were among the hypervariable loci further reveals potential key host-parasite interactors that evolved as the separate *C. hominis* lineages emerged, possibly as a result of adaptation to novel or altered host immune environments. Divergent evolution of surface glycoproteins has previously been described for closely-related lineages of other infectious organisms, and associations between prolonged immune response exposure and increased heterozygosity have been described (Holmes *et al*. 1992). The refinement of interactions between parasite glycoproteins and host

responses may thus have driven lineage-specific divergence as observed between *C. hominis* subtypes Ia and Ib.

Another key feature of lineage diversification was the presence of subtype-specific coding sequences, several of which appear to have resulted from major insertion or deletion events at the subtelomeres (Table 4.3.2). Gene loss and gain can have potent effects on an organism's physiological processes, and are therefore commonly identified in lineage- or species-specific adaptations. The biological purpose and/or consequence of these can differ significantly, however, and it is important to note that in silico findings cannot determine whether a subtype-specific coding sequence arose from loss versus gain through this type of single pairwise comparison alone. Broader gene repertoires can be associated with generalist versus specialist host ranges (Agrawal *et al*. 2016), increased virulence (Strouts *et al*. 2012; Linz *et al*. 2016), or an enhanced ability to withstand external stressors or regulate resistance (Xiao *et al*. 2009; Donoghue *et al*. 2011). The Ib-type WGS (UKH1) was found to contain five subtype-specific genes that were not present in the Ia-type WGS (UKH4), but that were conserved in *C. parvum* WGS belonging to subtypes IIa, IIc, and IId. A number of the encoded proteins were annotated as having adhesin or proteophosphoglycan-type properties, which could indicate potential roles in attachment and invasion or pathogenicity. The slightly extended host range in subtype Ib (Figure 4.1.1) and higher global human prevalence (chapter 3) could thus indicate that this lineage retained some generalist properties through specific gene conservation as it evolved from the common *C. hominis/C. parvum* ancestral strain, or that *C. parvum* and *C. hominis* subtype Ib separately gained certain advantageous genes through convergent evolution. On the other hand, gene loss during evolution can have equally beneficial effects. Deletion of genes during lineage-specific evolution and adaptation has been found to improve fitness, drive host specialization processes and even increase resistance to drugs for a range of pathogenic organisms (Gu *et al*. 2016). The loss or pseudo-functionalization of a handful of genes in Ia-type *C. hominis* lineages may therefore reflect evolutionary changes that occurred when Ia split from Ib and adapted towards a narrower anthroponotic lifestyle. The closer phylogenetic proximity of UKH1 to *C. parvum* versus UKH4 to *C. parvum* at a majority of these loci further lends support to this theory (Figure 4.3.6), and illustrates how variable conservation and adaptation of certain proteins may have shaped the phenotypic and epidemiological differences as they

currently stand (Figure 4.1.1). Although the presence of frameshift-causing InDels and subtype-specific sequence data was confirmed in other WGS of similar GP60 subtypes, it is important to note that the raw sequence reads from which the genomes were assembled were not examined to determine the reliability of these sequenced positions. This means that some uncertainty remains regarding the confidence of these findings. Performing an added level of quality control by referring to these reads in future may thus add or retract credibility surrounding the presence or absence of subtype-specificity at certain loci of these genomes, and is an important priority for intra-*C. hominis* comparative genomics moving forward.

GP60 variability quite literally towered over any other nucleotide divergence that could be detected across the genome, and was the only locus to produce a phylogeny that revealed potential evidence of recombination between *C. hominis* and *C. parvum* (Figure 4.4.1). The phylogenetic grouping of subtype Ib within the *C. parvum* clade, specifically its proximity to the anthroponotic *C. parvum* subtype IIc, provides some insight into the potential importance of this hypervariable locus in determining host-parasite interactions, especially in light of the comparable host ranges (predominantly anthroponotic with incidental zoonotic transmission) between subtypes IIc and Ib. The Ka/Ks ratio was insignificant and notably lower at this recombinant locus (~0.605) than the remaining divergent loci identified (mean Ka/Ks > 1.0), which is a typical occurrence for genomic fragments which have been imported from more distant lineages (Castillo-Ramírez *et al*. 2011). The retention of these highly-diverged recombinant GP60 alleles and absence of inter-subtype selective pressure could indicate a successful and rapid niche specialisation event. Intra-subtype divergences further imply a more recent timescale for Ia-subtype emergence than for Ib, as heightened levels of non-synonymous mutations often suggest a more recent diversification. Through this chronological hypothesis, the proximity of subtype Ib to zoonotic strains at GP60, and differential phenotypic traits we gain insight into the evolution of host specialisation in this species, which may imply that truly anthroponotic strains such as subtype Ia have only recently emerged within the *Cryptosporidium* genus, and that recombination at the GP60 locus was one of the key events involved.

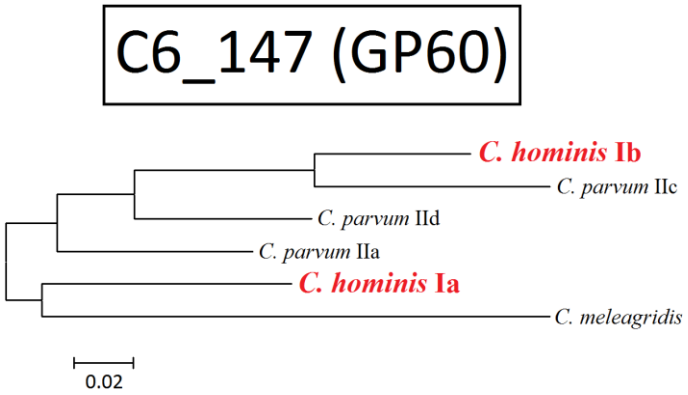


Figure 4.4.1. GP60 protein sequence phylogeny reveals potential phenotype-driven recombination between zoonotic and anthroponotic *Cryptosporidium spp.*

Whole gene GP60 sequences were extracted from WGS, translated into protein sequences, aligned (ClustalW), and evaluated phylogenetically using the Maximum Likelihood (ML) method, according to the Tamura-Nei model (MEGA v 7.0; Kumar *et al*. 2016), for *C. hominis* Ia (UKH4), *C. hominis* Ib (UKH1), *C. parvum* IIa (UKP4), *C. parvum* IId (UKP8), *C. parvum* IIc (34044), and *C. meleagridis* (UKMEL).

In addition to GP60, a further eight highly-divergent coding sequences were identified across the genome, along with six species-specific genes resulting from subtelomeric deletions and frameshifts. This finding limits the list of putative phenotype drivers to fifteen candidates, many of which have predicted protein functions in line with known host-parasite interactors, such as mucins (Hicks *et al*. 2000), glycoproteins (Ravidà *et al*. 2016), and adhesins (Boulanger *et al*. 2010). The three variable mucins directly adjacent to the 5'-telomere on chromosome 2 (Chro. 20052, 20051, and 20050) emerged as the most positively selected for according to Ka/Ks, reaching values above 5.0 (Figure 4.3.1), raising the question as to whether these proteins are selectively activated or modified in a manner that resembles antigenic variation in other parasitic pathogens (Andreas & Gojobori, 2004; Baranasic *et al*. 2014). These three mucins are the first in an array of seven, and the differential variability of certain members in this mucin array in zoonotic versus anthroponotic strains are a possible sign of antigenic variation. In zoonotic *C. parvum* subtypes IIa and IId, for example, the three mucins identified as divergent within *C. hominis* are entirely conserved (100% nucleotide identities), while the subsequent two in the array (Chro.20049 and Chro.20048) are divergent within *C. parvum* but entirely conserved between *C. hominis* subtypes Ia and Ib. In this chapter I also found a significant association between telomeric proximity and elevated SNP rates, which mirrors the genomic changes that seem to have accompanied speciation events separating *C. hominis* and *C. parvum*, as discussed in chapter 1.2. The *C. hominis*/*C. hominis* and *C. hominis*/*C. parvum* comparisons also shared the feature of significant sequence deletion or insertion

events at the telomeres, and an excess of frameshift-causing mutations across the genome, resulting in a number of species-specific genes. If the rule of 'more genes = more hosts' (or more virulence) applies, these species-specific genes may explain the phenotypic variations observed between *C. hominis* subtype Ia and Ib whole genome sequences. Thus, the level and nature of divergence between these two isolates, in conjunction with the unique host ranges and distributions they exhibit, may be enough to describe them as subspecies rather than subtypes, in a similar manner to which *C. parvum* IIc was recently re-classified as *C. parvum anthroponosum*. The only limit to this designation is that the epidemiological and clinical data seem too scarce at present to definitively describe true phenotypic differences, and that further WGS would be needed to determine whether genome-wide divergences are preserved across all Ia and Ib-type isolates.

Although in silico analyses on their own can only generate hypothetical scenarios of how protein changes affect biology, the results from this study provide the first comprehensive genomic hypothesis on how host specificity and virulence may be modulated in *Cryptosporidium hominis* at the molecular level. It also provides a new *C. hominis* species- and subtype-specific list of putative targets for genotyping, and has thereby expanded the range of molecular options for inter- and intra-molecular discrimination (Table 4.3.2). The next step forward inevitably lies in translation of in silico findings into an applied method. This involves validating whether the putative markers of virulence and host specificity identified in this study are more efficient at phylogenetic discrimination of strains (in a way that associates genetic features with phenotype) than the classic GP60-based approach. Investigating the molecular epidemiology of the hypervariable loci identified in this study across a broad range of *C. hominis* parasites is an important component of this, as extended associations between divergence and host data (such as clinical manifestations, immune responses, even geographic location) could reveal more about the functional consequence of gene diversification. Further exploration of the potential cellular role these proteins may play in host-parasite interactions is also needed, both to better understand the infection biology behind cryptosporidiosis but also to identify potential cellular pathways that could be targeted and interrupted through novel vaccination and/or treatment approaches. Preliminary work on recombination between *C. hominis* isolates has already begun to identify a number of recombinant fragments between divergent GP60 subtype families, which

means that much more remains to be learned about the nature of divergence between and within *Cryptosporidium spp*. This study provides a basis upon which such future studies can build, highlighting the types and consequences of genetic changes to look for. As epidemiological and clinical knowledge on *C. hominis* subtypes worldwide continues to grow, the results from this study may help to understand some of the genetic changes that impact phenotypic differences and accompany lineage-specific evolution. *C. hominis* lineages Ia and Ib appear to have arisen from a combination of gene loss/gain, hyperpolymorphism in a limited number of putative host-parasite interactors and significant positive selective pressure acting on telomerically-encoded surface antigens. Our understanding on how *Cryptosporidium* species and subtypes evolve and adapt is thereby improved, and may create a useful reference for future whole genome comparisons. The missing subtelomeric sequence data also provides a particularly useful opportunity to develop new molecular diagnostic methods, such as those that require subtype-specific sequence targets to achieve reliable florescence-based detection and discrimination. The data from this study has thereby introduced novel data on *Cryptosporidium* genomics that molecular techniques may benefit from. Through this, we get one step closer to understanding more about the genetic diversity and virulence that characterizes human infections for this largely neglected tropical disease.

Novel In Situ PCR/Fluorescence In Situ Hybridization (IS-PCR/FISH) technique for sensitive and specific detection of *C. p. parvum* oocysts

5.1 INTRODUCTION

Throughout most of the twentieth century, *Cryptosporidium* diagnosis was largely reliant on a trio of interdependent methods: concentration, staining, and microscopy. Concentration and staining have long been pre-requisites for the identification of the infectious stages of the parasite, the former due to the persistent presence of oocysts in low numbers (particularly in asymptomatic infections), and the latter due to the invisibility of oocysts through direct smears or less potent stains such as iodine. Many staining methods were successfully tailored to cryptosporidiosis very early on in the era of non-molecular detection, and few non-molecular diagnostic revolutions (i.e. largescale popularizations of new or improved methods) have taken place since this parasite entered into the limelight only decades ago. The advent of AIDS-related cryptosporidiosis and the consequential need for more sensitive and reliable DNA-based technologies in the 80s and 90s brought *Cryptosporidium* detection to an entirely new level, and simultaneously revolutionized our understanding of the genetic diversity and phylogenetic structuring of species and subtypes within this genus. This short review aims to discuss how *Cryptosporidium* detection methods have evolved through time, describing both the advancements that have been made through the introduction of DNA-based molecular technologies, as well as evaluating areas where the type and range of molecular information retrieved through diagnostics can still be improved.

(i) A history of diagnoses in cryptosporidiosis

For the first half of the twentieth century only a handful of studies were published describing *Cryptosporidium* infections, and these were limited to mice and chickens in the United Kingdom (Tyzzer 1907; Tyzzer 1910; Tyzzer 1912; Tyzzer 1929; Elton *et al*. 1931) and snakes in Uganda (Hoare 1933). Through these studies the first two species were described, *Cryptosporidium muris* and nondescript species in chicken simply referred to as *Cryptosporidium* sp. (later renamed *Cryptosporidium tyzzeri*; Levine, 1961), but no reliable host range patterns or epidemiological characterisations were available. In all of the pre-1950 studies that were published, histological sections of the intestine were used in order to detect and describe oocysts and other life cycle stages such as schizonts, merozoites, gametocytes and gametes. Most reported the use of a haematoxylin- and eosin-based (HE) staining approach (Tyzzer 1907-1929; Hoare 1933), with a single exception where Gram's stain had been used (Elton *et al*. 1931). The accepted description of *Cryptosporidium* as a parasite of limited and entirely zoonotic concern at the time explains the scarce and sporadic interest demonstrated by the scientific community between 1900 and 1950; interest during this era was instead focused on other coccidian parasites, such as *Eimeria*, which was published on more than 100 times before 1950 (database search 'Eimeria': PubMed, https://www.ncbi.nlm.nih.gov/pubmed and Wiley Online Library, http://onlinelibrary.wiley.com/). This trend was mirrored in academic textbooks and public health manuals of the time, where *Cryptosporidium* consistently received only brief mentions or single paragraphs, compared to entire chapters dedicated to other protozoa (Figure 5.1.1; Minchin, 1922; Wenyon, 1926; Manter, 1938; Zeliff, 1947). Reviewing the literature that would have been available to academics and lay persons alike during this time period therefore shows a striking lack of information and understanding, and drastically contrasts with the >4000 peer-reviewed articles that have been published on *Cryptosporidium* since 1950.

sporoblasts by division; each secreting about it a **sporocyst wall**, thus becoming a **spore** (sporocyst). Frequently some residual cytoplasm is left here also and in some types under this family the spore development does not occur until external environment acts as the stimulus. The oocysts pass out with the feces and the cycle is completed. Study a section of liver and draw the stages in natural order.

E. perforans, E. magna, E. irresidua and **E. media** are species which occur in the intestinal epithelium of the rabbit. See the reference for species in rabbits.

E. bovis, E. smithi, E. ellipsoidalis, E. zurnii are the most pathogenic to cattle in the U. S. (N. Am. Vet., March 1945) among the ten species described. **E. deblecki** of pigs, **E. faurei** of sheep and goats, **E. canis** of dogs, **E. felina** of cats, **E. miyairii (E. nieschulzi)** of brown rats, **E. falciformis** of mice, **E. ranarum** of frogs, **E. neglecta** of frog tadpoles, **E. sardinae** of sardines and **E. cleupearum (E. wenyoni)** of herring and mackerel are some other species, exclusive of those of birds; and several erroneously reported in man. At least seven species occur in sheep (Vet. Med., Sept., 1943).

Eimeria of Man

Eimeria gubleri, E. snijdersi, E. oxyspora and **E. wenyoni** have all been reported from man but the three latter species are known to be normally in fish and probably the first one was also accidentally ingested by eating fish, or is identical with **E. stiedae** of rabbits, and acquired as the others.

Eimeria of Birds

Eimeria tenella, E. mitis, E. acervulina, E. maxima, E. praecox and **E. necatrix** occur in chickens. Tyzzer (1929) described the first four species, and Tyzzer, Jones and Theiler (1932) confirmed the last two, which were described by Johnson (1930). The first species lives in the caeca and is very pathogenic, as is also **E. necatrix. E. acervulina** produces a chronic coccidiosis. Hemorrhage and lysis of the intestinal epithelium occur. **E. hagani** and **E. brunetti** also pathogenic, have recently been described. Study and draw various stages.

E. melegrimitis and **E. meleagridis** occur in domestic turkeys. **E. phasiani** and **E. dispersa** occur in pheasants.

Genus: **Cryptosporidium**

Cryptosporidium parvum, a very minute coccidian, also occurs in the intestinal epithelium of chickens.

Suborder: Adeleina
Family: Adeleidae
Genus: **Klossiella**

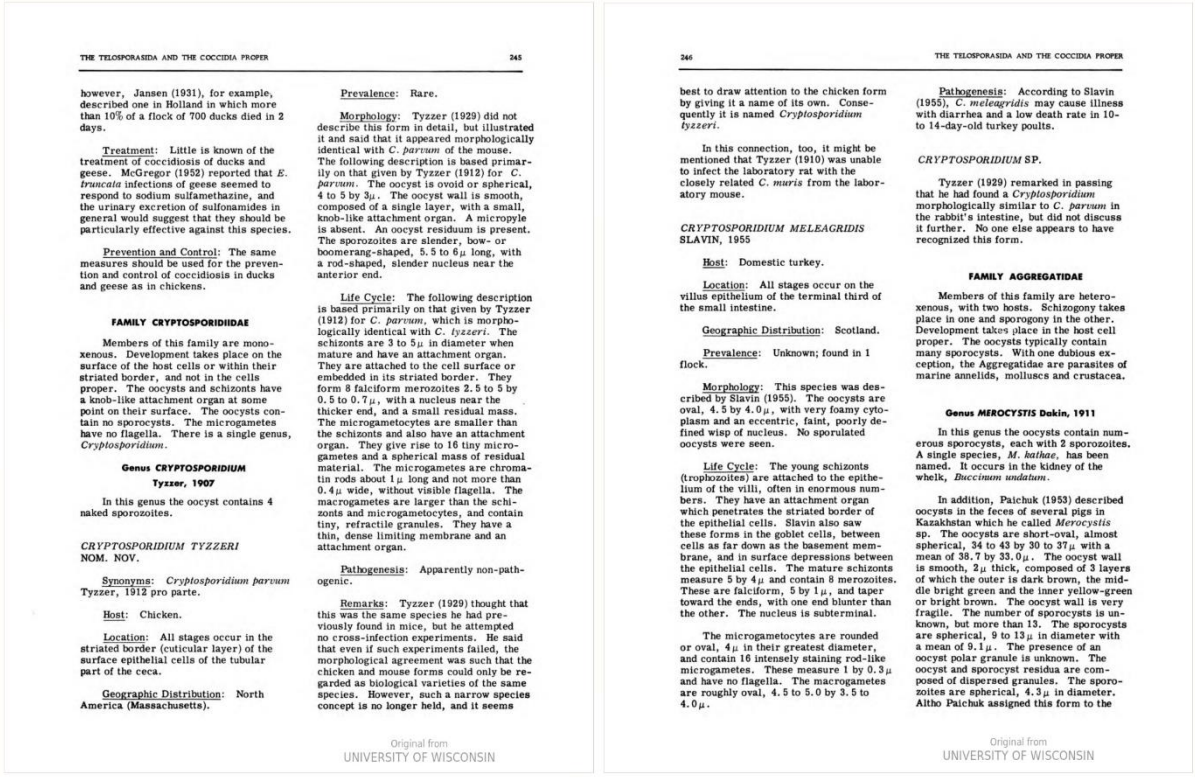
Klossiella muris. - This occurs in the kidney of mice and other species occur in other mammals. Schizogony may occur in other organs. Thirty or more sporozoites are formed.

Suborder: Haemogregarinina
Family: Hemogregarinidae
Genera: **Haemogregarina, Hepatozoon** and **Karyolysus**

Haemogregarina stepanowi. - Gametocytes occur as crescentlike forms in the erythrocytes of turtles; completion of the life cycle occurring in **Placobdella,** a leech. Uncertified species have been reported in man. Many occur in frogs and fish. No sporocyst (spore) is formed.

Figure 5.1.1. Scanned excerpt showing the only included description for *Cryptosporidium* in an early parasitological manual, from 'Manual of medical parasitology with techniques for laboratory diagnosis and notes on related animal parasites' (Page 34; Zeliff, 1947)

In the 1950's and 1960's a few more published articles started to trickle through, extending the range of hosts to include an Australian dingo (Bearup, 1954), Indian cat (Dubey & Pande, 1963), a North American guinea pig (Jervis *et al*. 1966), and Scottish turkeys (Slavin, 1955). One of the first comprehensive textbook descriptions of *Cryptosporidium* was also published in 1961 (Levine, 1961), providing a summary of what limited preliminary data existed regarding geographic distribution, pathogenesis, prevalence, morphology and life cycle (Figure 5.1.2). Reviewing the diagnostic approaches in this time period proved difficult, as two of the studies had no full text available (Dubey & Pande, 1963; Jervis *et al*. 1966) and one of the studies did not include a methodology section (Bearup, 1954). The fourth study reported use of the MacNeal staining procedure (Slavin, 1955), which similar to pre-1950 *Cryptosporidium* staining methods used eosin but also incorporated methylene blue as a counterstain (Adwan, 1952). Two more species emerged during this period, *Cryptosporidium meleagridis* (Slavin, 1955) and *Cryptosporidium tyzzeri* (Levine, 1961), bringing the total accepted species-types to four. This started to increase public understanding and thereby interest in cryptosporidiosis, and in the decades that followed the availability of data (and by association the range of diagnostic techniques) started to increase, eventually reaching exponential proportions during the 1980's AIDS era and beyond.



THE TELOSPORASIDA AND THE COCCIDIA PROPER 245

however, Jansen (1931), for example, described one in Holland in which more than 10% of a flock of 700 ducks died in 2 days.

Treatment: Little is known of the treatment of coccidiosis of ducks and geese. McGregor (1952) reported that *E. truncata* infections of geese seemed to respond to sodium sulfamethazine, and the urinary excretion of sulfonamides in general would suggest that they should be particularly effective against this species.

Prevention and Control: The same measures should be used for the prevention and control of coccidiosis in ducks and geese as in chickens.

FAMILY CRYPTOSPORIDIIDAE

Members of this family are monoxenous. Development takes place on the surface of the host cells or within their striated border, and not in the cells proper. The oocysts and schizonts have a knob-like attachment organ at some point on their surface. The oocysts contain no sporocysts. The microgametes have no flagella. There is a single genus, *Cryptosporidium*.

Genus CRYPTOSPORIDIUM Tyzzer, 1907

In this genus the oocyst contains 4 naked sporozoites.

CRYPTOSPORIDIUM TYZZERI NOM. NOV.

Synonyms: *Cryptosporidium parvum* Tyzzer, 1912 pro parte.

Host: Chicken.

Location: All stages occur in the striated border (cuticular layer) of the surface epithelial cells of the tubular part of the ceca.

Geographic Distribution: North America (Massachusetts).

Prevalence: Rare.

Morphology: Tyzzer (1929) did not describe this form in detail, but illustrated it and said that it appeared morphologically identical with *C. parvum* of the mouse. The following description is based primarily on that given by Tyzzer (1912) for *C. parvum*. The oocyst is ovoid or spherical, 4 to 5 by 3 μ. The oocyst wall is smooth, composed of a single layer, with a small, knob-like attachment organ. A micropyle is absent. An oocyst residuum is present. The sporozoites are slender, bow- or boomerang-shaped, 5.5 to 6 μ long, with a rod-shaped, slender nucleus near the anterior end.

Life Cycle: The following description is based primarily on that given by Tyzzer (1912) for *C. parvum*, which is morphologically identical with *C. tyzzeri*. The schizonts are 3 to 5 μ in diameter when mature and have an attachment organ. They are attached to the cell surface or embedded in its striated border. They form 8 falciform merozoites 2.5 to 5 by 0.5 to 0.7 μ, with a nucleus near the thicker end, and a small residual mass. The microgametocytes are smaller than the schizonts and also have an attachment organ. They give rise to 16 tiny microgametes and a spherical mass of residual material. The microgametes are chromatin rods about 1 μ long and not more than 0.4 μ wide, without visible flagella. The macrogametes are larger than the schizonts and microgametocytes, and contain tiny, refractile granules. They have a thin, dense limiting membrane and an attachment organ.

Pathogenesis: Apparently non-pathogenic.

Remarks: Tyzzer (1929) thought that this was the same species he had previously found in mice, but he attempted no cross-infection experiments. He said that even if such experiments failed, the morphological agreement was such that the chicken and mouse forms could only be regarded as biological varieties of the same species. However, such a narrow species concept is no longer held, and it seems

246 THE TELOSPORASIDA AND THE COCCIDIA PROPER

best to draw attention to the chicken form by giving it a name of its own. Consequently it is named *Cryptosporidium tyzzeri*.

In this connection, too, it might be mentioned that Tyzzer (1910) was unable to infect the laboratory rat with the closely related *C. muris* from the laboratory mouse.

CRYPTOSPORIDIUM MELEAGRIDIS SLAVIN, 1955

Host: Domestic turkey.

Location: All stages occur on the villus epithelium of the terminal third of the small intestine.

Geographic Distribution: Scotland.

Prevalence: Unknown; found in 1 flock.

Morphology: This species was described by Slavin (1955). The oocysts are oval, 4.5 by 4.0 μ, with very foamy cytoplasm and an eccentric, faint, poorly defined wisp of nucleus. No sporulated oocysts were seen.

Life Cycle: The young schizonts (trophozoites) are attached to the epithelium of the villi, often in enormous numbers. They have an attachment organ which penetrates the striated border of the epithelial cells. Slavin also saw these forms in the goblet cells, between cells as far down as the basement membrane, and in surface depressions between the epithelial cells. The mature schizonts measure 5 by 4 μ and contain 8 merozoites. These are falciform, 5 by 1 μ, and taper toward the ends, with one end blunter than the other. The nucleus is subterminal.

The microgametocytes are rounded or oval, 4 μ in their greatest diameter, and contain 16 intensely staining rod-like microgametes. These measure 1 by 0.3 μ and have no flagella. The macrogametes are roughly oval, 4.5 to 5.0 by 3.5 to 4.0 μ.

Pathogenesis: According to Slavin (1955), *C. meleagridis* may cause illness with diarrhea and a low death rate in 10- to 14-day-old turkey poults.

CRYPTOSPORIDIUM SP.

Tyzzer (1929) remarked in passing that he had found a *Cryptosporidium* morphologically similar to *C. parvum* in the rabbit's intestine, but did not discuss it further. No one else appears to have recognized this form.

FAMILY AGGREGATIDAE

Members of this family are heteroxenous, with two hosts. Schizogony takes place in one and sporogony in the other. Development takes place in the host cell proper. The oocysts typically contain many sporocysts. With one dubious exception, the Aggregatidae are parasites of marine annelids, molluscs and crustacea.

Genus *MEROCYSTIS* Dakin, 1911

In this genus the oocysts contain numerous sporocysts, each with 2 sporozoites. A single species, *M. kathae*, has been named. It occurs in the kidney of the whelk, *Buccinum undatum*.

In addition, Paichuk (1953) described oocysts in the feces of several pigs in Kazakhstan which he called *Merocystis* sp. The oocysts are short-oval, almost spherical, 34 to 43 by 30 to 37 μ with a mean of 38.7 by 33.0 μ. The oocyst wall is smooth, 2 μ thick, composed of 3 layers of which the outer is dark brown, the middle bright green and the inner yellow-green or bright brown. The oocyst wall is very fragile. The number of sporocysts is unknown, but more than 13. The sporocysts are spherical, 9 to 13 μ in diameter with a mean of 9.1 μ. The presence of an oocyst polar granule is unknown. The oocyst and sporocyst residua are composed of dispersed granules. The sporozoites are spherical, 4.3 μ in diameter. Altho Paichuk assigned this form to the

Figure 5.1.2. Scanned excerpt showing one of the first comprehensive textbook entries for *Cryptosporidium*, from 'Protozoan parasites of domestic animals and of man' (Pages 245-246; Levine, 1961)

The 1970's marked a significant shift in our understanding of *Cryptosporidium* and its pathogenic impact, with the first human case documented in a 3-year-old child suffering from acute, self-limiting gastroenteritis in 1976 (Nime *et al*. 1976). By the end of the 70's, two further cases in humans had been reported, both in immunocompromised adults (Meisel *et al*. 1976; Lasser *et al*. 1979), providing a first insight into what would become a defining feature of cryptosporidiosis in humans. The spectrum of animal hosts also continued to broaden, with *Cryptosporidium* now being identified in rhesus monkeys (Kovatch & White, 1972; Cockrell *et al*. 1974), a goose (Proctor & Kemp, 1974), lambs (Berg *et al*. 1979), calves (Barker & Carbonell, 1974; Schmitz & Smith, 1975; Pohlenz *et al*. 1978), pigs (Bergeland, 1977; Kennedy *et al*. 1977), foals (Snyder *et al*. 1978), a number of snake species (Brownstein *et al*. 1977; McKenzie *et al*. 1978), parrots (Doster *et al*. 1979), and rabbits (Rehg *et al*. 1979; Inman & Takeuchi, 1979). First reports of respiratory cryptosporidiosis in birds had also begun to appear (Hoerr *et al*. 1978). Diagnostic approaches seemed surprisingly standardized during this decade, with available studies universally reporting use of an HE-based staining technique, and a few describing slight modifications by adding Giemsa and/or periodic acid-Schiff (PAS) staining (Figure 5.1.3; Vetterling *et al*. 1971; Nime *et al*. 1976; Meisel *et al*. 1976; Brownstein *et al*. 1977; Snyder *et al*. 1978). This decade also features one of the first reports of a stool concentration technique to improve oocyst recovery (Vetterling *et al*. 1971), through Ritchie's formaldehyde and ether-based sedimentation technique (Ritchie, 1948). This would eventually be incorporated into routine clinical and epidemiological diagnosis of *Cryptosporidium* worldwide, and many studies today still use Ritchie's technique (or modified versions thereof) as an indispensable part of oocyst detection.

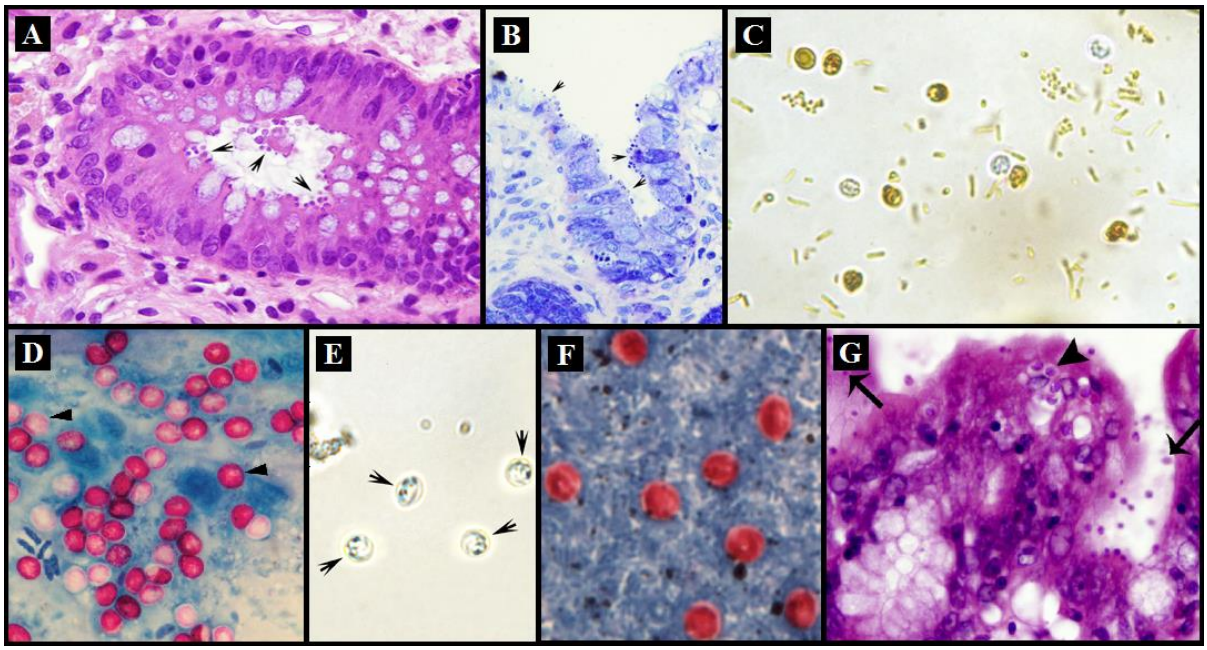


Figure 5.1.3. Microscopy images of various non-molecular *Cryptosporidium* staining techniques
(A) Haematoxylin and Eosin, (B) Giemsa, (C) Iodine, (D) Kinyoun's, (E) Direct, (F) Ziehl-Neelsen, (G) Periodic-acid Schiff (Sources: http://www.atlas-protozoa.com/gallery.php?SOT_CAP=CRIPTO#0, https://www.cdc.gov/dpdx/cryptosporidiosis/, Nakamura & Meireles 2015, Barugahare *et al*. 2011). Images reproduced with written permission from the publishing authors.

The birth and spread of the global HIV/AIDS epidemic in the early 1980's projected *Cryptosporidium* into an entirely new category of public health concern. In the preceding years, scientists had uncovered a tentative link between immune status and the risk of cryptosporidiosis, but it is difficult to imagine anyone having predicted the scale of devastation this parasite was about to inflict upon a new global population of immunodeficient hosts. This was no longer the rare zoonotic parasite of limited human importance, but one of the primary causes of substantial morbidity and mortality in the ever-increasing number of HIV-positive individuals worldwide. Hundreds of reports and published studies on cryptosporidiosis emerged throughout the 80's, more than 10-fold the number in the preceding 70 years, which appears to correlate with the continuously increasing number of individuals living with HIV (Figure 5.1.4). With this sharp rise in importance and interest came the need for improved diagnostic methods that would effectively gather as much information on the epidemiology and pathogenicity of this parasite as possible. One of the major diagnostic introductions throughout this decade was the adaptation of Ziehl-Neelsen staining for *Cryptosporidium* detection. The Ziehl-Neelsen (ZN) acid-fast staining technique, which had been developed more than a century before (Koch, 1882; Ehrlich, 1882; Ziehl, 1882; Rindfleisch, 1882; Neelsen, 1883; Bishop & Neumann, 1970), was first introduced for *Cryptosporidium* diagnosis in 1983 when Garcia

et al. found that the *Mycobacterium tuberculosis* stain was more sensitive at detecting *Cryptosporidium* oocysts than any alternative diagnostic technique available at the time (Garcia *et al*. 1983). This study also popularized a number of other methods such as the Kinyoun acid-fast method, variations of which still remain in use today (Neres-Norberg *et al*. 2014; Ghaffari *et al*. 2014; Uysal *et al*. 2016). ZN microscopy was a method that proved so successful at staining and picking up oocysts, that it became the most frequently used technique for decades to follow, and is still often used in place of molecular approaches for largescale prevalence studies (Maurya *et al*. 2016; Kiani *et al*. 2016; Rzymski *et al*. 2017). Preservation of stools in 10% formalin became routine practice during the 80's, as did stool concentration by means such as Sheather's sugar flotation method, formalin-ether or formalin-ethyl acetate sedimentation (Current *et al*. 1983). The use of membrane filters (Ongerth & Stibbs, 1987) and cartridge filters (DoE/SCA, 1989) for diagnostic processing of very large environmental samples was also introduced.

Immunofluorescence and antigen-based serological tests were increasingly developed and implemented for rapid and simple detection of *Cryptosporidium* antibodies, but the presence of a serological response long after an infection had passed meant this type of method had to tendency to overinflate prevalence estimates, and would not become the preferred approach for a majority of largescale prevalence and incidence surveys of cryptosporidiosis (Tzipori & Campbell, 1981; Campbell & Current, 1983; Casemore, 1987; Garcia *et al*. 1988). Direct fluorescence was also integrated into the diagnostic repertoire during the 80's, which simplified initial screening procedures by producing brightly fluorescing and easily-identifiable oocysts, commonly through Auramine (Ratnam *et al*. 1985; Stibbs & Ongerth, 1986; Ley *et al*. 1988) or acridine orange (Garcia *et al*. 1983) staining. The acceleration of detection methods experienced in this decade reflects the urgency that was felt across the public health community, and demonstrates the rapid attempt that was made to meet the diagnostic needs of this parasite of increasing public health importance. While none of the methods developed during this pre-molecular era were developed with the capacity of more advanced species and subtype discrimination, they facilitated the integration of routine diagnosis of *Cryptosporidium* in a rapid, sensitive, and efficient way, and further laid a solid platform of techniques and methodology upon which the molecular era would build.

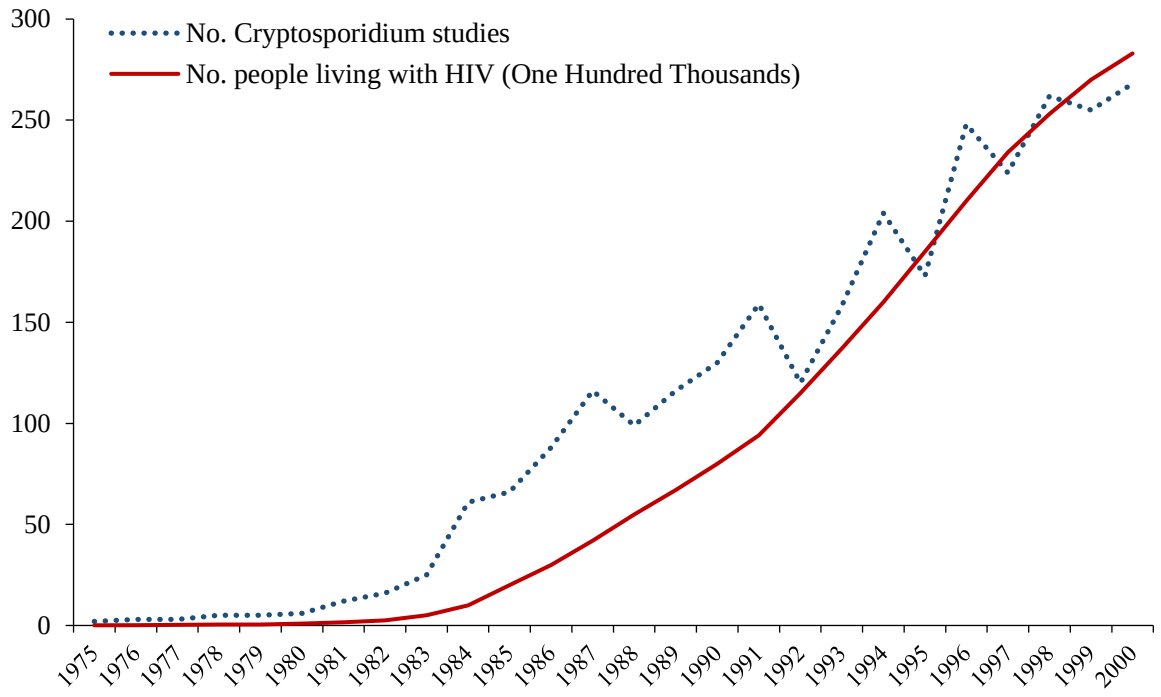


Figure 5.1.4. Total number of published *Cryptosporidium* studies (PubMed: https://www.ncbi.nlm.nih.gov/pubmed) and global number of HIV-positive individuals (WHO/UNAIDS/UN, 2009), 1975-2000

Throughout the 1980's, heightened interest in *Cryptosporidium* led to improved knowledge about its infectivity and epidemiology, and by the end of the decade, comprehensive descriptions on topics such as geographic distribution, life cycle, pathogenesis, transmission, and diagnosis were beginning to form (Garcia & Current, 1989), and 19 individual species had already been described (Current, 1986). Taxonomic relationships based on biological traits such as morphology and reproduction placed *Cryptosporidium spp*. somewhere between *Isospora belli* and *Sarcocystis spp*. on the protozoan tree of life, but it was only during the molecular era that followed that the true phylogenetic relationships between and within *Cryptosporidium* species began to emerge.

(ii) FISH in *Cryptosporidium*

The first adaptations of the polymerase chain reaction (PCR) for *Cryptosporidium spp.* started to appear in the early 1990s
(Goozé *et al*. 1991; Laxer *et al*. 1992; Cai *et al*. 1992). Initially only capable of discriminating *Cryptosporidium* from other protozoa (Johnson *et al*. 1995), this method would evolve within a few years to achieve inter-species discrimination (Leng *et al*. 1996; Morgan *et al*. 1997), and shortly after the new millennium begin to describe intra-species

and even intra-subtype phylogenies and divergences as well (Zhou *et al*. 2003; Alves *et al*. 2006). Previous chapters in this thesis have described some of the popular targets of PCR-based diagnostics through the years, and discussed the degree, consistency, and accuracy of discrimination, demonstrating the significant advancements yet room for improvement that exists within molecular diagnosis of *Cryptosporidium*. One molecular technique that has shown significant promise in recent years is the use of fluorescence in situ hybridization (FISH) for simple, sensitive, and informative discrimination between phenotypically-unique species. Even before reference genomes first became available for *Cryptosporidium spp*. in 2004, this method was able to successfully identify and discriminate *C. parvum* oocysts in environmental and clinical samples (Vesey *et al*. 1998, Graczyk *et al*. 2003; Smith *et al*. 2004). Studies in *Cryptosporidium* have predominantly used the popular DNA target '18S rRNA' (Lemos *et al*. 2005; Taguchi *et al*. 2006; Bednarska *et al*. 2007; Sunderland *et al*. 2007; Graczyk *et al*. 2008; Lucy *et al*. 2008; Alagappan *et al*. 2008; Örmeci & Linden, 2008; Alagappan *et al*. 2009; Zintl *et al*. 2010; Bougiouklis *et al*. 2013), as it is the only currently-known target that exists in enough copies to facilitate a strong enough fluorescence signal for detection. However, the largely conserved nature of 18S rRNA across *Cryptosporidium spp*. means there are limitations in terms of discriminatory power within subspecies and subtypes. Studies in other apicomplexan parasites such as *Neospora caninum* have attempted to overcome this by combining FISH with a preceding In Situ PCR (IS-PCR) step (Löschenberger *et al*. 2004). This method uses viable parasites in place of DNA extract and performs PCR within the cell (i.e. *In Situ*), and can theoretically target and amplify any genetic fragment to which primers can successfully bind, thereby expanding the list of suitable multi-copy FISH targets to a seemingly endless number of candidates. No such combined method has previously been validated and/or published on in *Cryptosporidium*, but with the true genetic diversity of human-infective species beginning to emerge, the value of having access to this type of diagnostic method is increasingly evident.

The simple and rapid identification and numeration of oocysts through fluorescence microscopy (Deere *et al*. 1998), as well as its ability to directly differentiate between viable and non-viable oocysts, makes this method an appealing alternative for *Cryptosporidium* diagnostics compared to some of the more laborious gold standard molecular techniques such as PCR. Despite being able to differentiate between species

types (Alagappan *et al*. 2009), current FISH methods in *Cryptosporidium* do not include multi-copy targets that exhibit enough variation between strains to facilitate discrimination of oocysts based on subtype and subspecies. This study saw an opportunity to fill this diagnostic gap, by advancing and broadening the specificity of fluorescence detection through a combined In Situ PCR/Fluorescence In Situ Hybridization (IS-PCR/FISH) approach. The initial aim was to demonstrate whether this method was sensitive enough to consistently produce visible fluorescence signals for routine diagnosis. The secondary aim involved comparative application of the technique on phenotypically-diverse human-infective specie – to demonstrate that species-specific targeting of the technique could in fact be achieved.

5.2 METHODS

(i) Oocysts

Two sets of oocysts were used for validation of the *C. p. parvum* IS-PCR/FISH detection method - one belonging to the *C. parvum* species group and another belonging to *C. hominis*. Purified *C. p. parvum* oocysts belonging to GP60 subtype IIaA19G2R1 (stock concentration 10^8/ml) were generously provided by the Moredun Research Institute (Dr. Frank Katzer; Midlothian, Scotland, United Kingdom). Purified *C. hominis* oocysts belonging to GP60 subtype family Ia (stock concentration 10^6/ml) were generously provided by the *Cryptosporidium* Reference Unit (Guy Robinson; Singleton University Hospital, Swansea, Wales, United Kingdom). Both oocyst suspensions were provided in a 1X concentrated solution of phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4; pH 7.4) and stored under sterile conditions at 4°C for the duration of this study.

(ii) Primers and Probe

PCR primers for In Situ PCR were designed based on optimal feature specifications regarding size, melting temperature (T_m), GC content (%), and self-complementarity (Chavali *et al*. 2005). A ~8,200 bp *C. parvum*-specific fragment at the 3'-end of chromosome 6 was blasted using NCBI's Primer-BLAST software tool (Ye *et al*. 2012) to identify effective forward and reverse primers for a pre-specified amplicon length of 100-300 bp. Primer sequence hits were then manually checked for cross-reactivity against *C. p. anthroponosum*, *C. hominis*, and a range of zoonotic species for which WGS were available through the present study (*C. meleagridis*, *C. baileyi*, *C. muris*, *C. cuniculus*, and *C. ubiquitum*). The selected primer pair, titled CPP6T3 (CPP = *C. p. parvum*, 6 = chromosome 6, T3 = 3'-telomere) was subsequently checked to ensure 100% conservation across the *C. p. parvum* subspecies. A 208 bp target sequence is amplified by the primers chosen for this study (Figure 5.2.1).

Table 5.2.1. Primer ID's and specifications for *C. p. parvum*-specific In Situ PCR protocol

ID	Sequence (5' -> 3')	Length	Tm	GC content
CPP6T3F (forward primer)	5'-AAAACCGCCAGCATACAAGA-3'	20 bp	58.1 °C	45.0%
CPP6T3R (reverse primer)	5'-TCGGAAGCCATAATGATCCCG-3'	21 bp	60.0 °C	52.4%

The 6-FAM (6-carboxyfluorescein)-labelled oligonucleotide probe sequence was similarly designed based on pre-determined specifications for optimised binding efficiency. Hybridization probe design was accomplished using the ARB software package (Ludwig *et al*. 2004), which identified the optimal 20 bp CPP6T3 probe (5'-TCCTTATTAGTAGACCAGGA-3') used in this study. The CPP6T3 DNA probe was designed to include a 5'-6-FAM modification, for visualization under a FITC fluorescence microscope channel upon successful hybridization to the target sequence. All primers and probe used in this study were ordered through Life Technologies (Thermo Fisher Scientific, United Kingdom), at a desalted purification and a synthesis scale of 25 nmol.

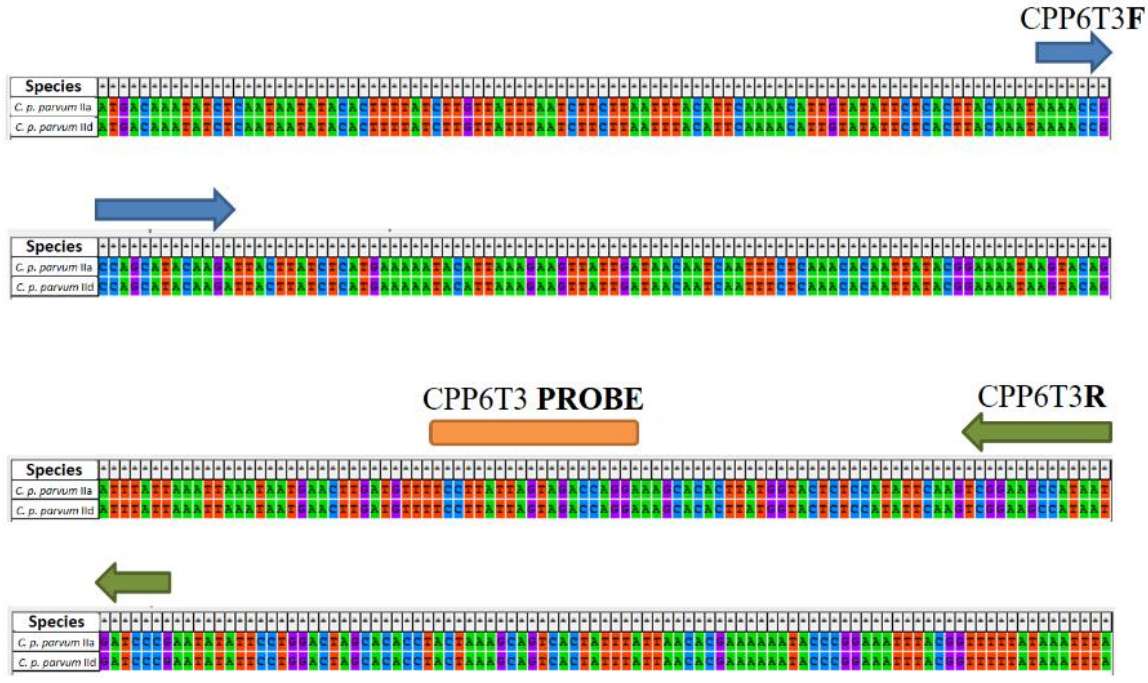


Figure 5.2.1. Primer and probe positions along 208 bp IS-PCR/FISH target sequence

The primer and probe target sequences along the 208 bp IS-PCR amplicon are here shown. Two *C. p. parvum* GP60 subtype (IIa and IId) target sequences are illustrated, to demonstrate successful binding capacity of the designed primers and probe across a range of *C. p. parvum* subtypes. The forward primer CPP6T3F, reverse primer CPP6T3R, and CPP6T3 probe are shown in blue, green, and orange, respectively.

(iii) Optimization of oocyst permeabilization/fixation and the IS-PCR/FISH protocol

The initial approach for permeabilization and fixation of *Cryptosporidium* oocysts was based on a 10 to 20-minute incubation of oocysts in 1:1 PBS/ethanol absolute solution at 80°C (Deere *et al*. 1998; Dorsch & Veal, 2001; Smith *et al*. 2004; Chu *et al*. 2011). Through numerous trial-and-error approaches to investigate the strength and sensitivity of the IS-PCR/FISH method, this was subsequently adapted to a 60-minute incubation period in 1:1 PBS/ethanol absolute at 4°C. An additional reagent step was later added, largely to ensure sufficient permeabilization of the thick-walled *Cryptosporidium* oocysts, through addition of a 15-minute incubation (room temperature) in 0.5% Triton X-100, as this detergent has proved effective at permeabilizing cell walls in *Cryptosporidium* and a number of other protozoa (Elliott & Clark, 2000; Paluszynski *et al*. 2014; Bandini *et al*. 2016; Currà *et al*. 2016).

The PCR reagent set-up and cycling conditions were developed through a range of troubleshooting approaches, and were initially based on the reagents and concentrations generally used for standard PCR (Butler, 2005), and the manufacturer's thermocycling recommendations outlined for the DNA polymerase kit used (Qiagen Non-Hot Start Taq DNA polymerase; 3 min 94°C initial denaturation, followed by 25-35 cycles of: 94°C denaturation for 0.5–1 min, 50–68°C annealing for 0.5-1min (approximately 5°C below T_m of primers), and 1 min 72°C extension, followed by a final extension for 10 min at 72°C). Protocols using a broad spectrum of PCR reagent volumes and combinations were designed and tested (1mM-5mM $MgCl_2$; 12mM-20mM Tris HCl; 30mM-50mM KCl; 12µM-20µM dNTPs (each), 0.1µM -1.0µM primer (each), and 0.5-5 units Non-Hot Start Taq polymerase), adjusting ingredients and temperatures based on the final outcome of each IS-PCR/FISH experiment. An unconventional change was further made through the addition of 0.1% Triton X-100 in the PCR master mix, based on promising results from an IS-PCR/FISH protocol in *Neospora caninum* (Löschenberger *et al*. 2004). The process of conducting all steps before, during, and directly after IS-PCR directly in a single PCR tube was also incorporated, as it improved the oocyst recovery rate compared to shifting between various Eppendorf and PCR tubes during the many permeabilization, fixation, centrifugation and amplification steps

Oocyst fixation and dehydration post IS-PCR in preparation for probe hybridization was adapted from previous protocols recommending air drying slides through 37°C incubation and ethanol dehydration series ranging from 70.0-100.0% (Erlandsen *et al*. 2005; Gookin *et al*. 2006; Timur & Wang, 2006; Kernohan & Bérubé, 2014). The recipes for hybridization buffer and wash buffer were retrieved through personal communication with FISH specialists at the University of Vienna (Dr. Sílvia Cervero-Aragó, Dr. Alexander Kirschner; Institute for Hygiene and Applied Immunology, Medical University of Vienna), and later adapted according to similar published protocols (Seviour & Nielsen, 2010; Kumar *et al*. 2013). Hybridization times varied significantly in previously-published studies, ranging from a few hours (Kumar *et al*. 2013) to several days (Hasegawa *et al*. 1991; Acloque *et al*. 2008), although the most commonly reported duration seemed to lie somewhere between 14 and 18 hours (Sirinarumitr *et al*. 2000; Fuyita *et al*. 2012; Matthiesen & Hansen, 2012; Li *et al*. 2015; Eyada *et al*. 2016). As a result, experiments were consistently run in quadruplicates at four separate hybridization durations (2 hours, 8 hours, 16 hours, and 24 hours), until the optimal time for hybridization of probe to target was achieved.

(iv) IS-PCR/FISH protocol for identification of purified *Cryptosporidium parvum* oocysts

- <u>Fixation and permeabilization</u>

Briefly, the purified oocyst suspension (concentration 10^4/ml to 10^6/ml) was centrifuged at 13,000 rpm (18,928 rcf) and the supernatant was removed. 1ml of 1:1 PBS/Ethanol was added and the mixture was incubated for 60 minutes at 4°C, followed by two washes at 13,000 rpm (18,928 rcf); supernatant was removed. Subsequently, 500 µl of 0.5% Triton X-100 in PBS was added and the solution was incubated for 15 minutes at room temperature, followed by two washes at 5,000 rpm (2,800 rcf); supernatant was removed except ~5 µl.

- In Situ PCR

A 50 µl amplification mixture was prepared containing 18 µl H_2O, 5 µl 10X buffer (500 mM KCl, 100 mM Tris-HCl (pH 9.0 at 25°C)), 6 µl $MgCl_2$ (50 mM), 2 µl dNTP mix (10 mM), 6 µl Forward primer (10 µM), 6 µl Reverse primer (10 µM), 1 µl Non-Hot Start Polymerase (5 U/µl; Qiagen, Hilden, Germany), and 1 µl 0.1% Triton X-100; the PCR mastermix was added to the permeabilized and fixated 5 µl oocyst suspension. PCR was conducted using a Veriti™ Thermal Cycler (Applied Biosystems, California, United States). An initial denaturation step at 94°C for 3 minutes was followed by 35 cycles of 94°C for 30 seconds, 56°C for 30 seconds, and 72°C for 50 seconds, followed by a final extension at 72°C for 10 minutes and 10°C for 10 minutes. PCR product was subsequently washed twice in 1X PBS by centrifugation at 2,500 rpm (700 rcf) for 30 seconds.

- Hybridization and Washing

10 µl of IS-PCR product was inserted onto a multiwell Teflon-coated microscope slide (Tekdon® Inc., Florida, United States), and incubated at 37°C for 30 minutes. The fixed slides subsequently underwent ethanol dehydration by insertion into 50%, 80%, and 96% ethanol solutions for 3 minutes each. Slides were subsequently air-dried under a fume hood. 8 µl of hybridization buffer (20% formamide, see table 5.2.2) plus 1 µl of FISH probe (0.3 mM working concentration) were mixed using a pipette and added to the well. A humid hybridization chamber was prepared, and the slide was carefully inserted and incubated at 46°C for 10-16 hours.

Following hybridization, 50ml of washing buffer (10% EDTA, see table 5.2.3) was pre-heated in a 48°C water bath, and the slide was added directly from the hybridization chamber. The slide was subsequently incubated in a dark water bath at 48°C for 15 minutes. Following hybridization washing, the slide was briefly dipped into ice-cold MiliQ water and air-dried under dark conditions under a fume hood. In preparation for microscopy visualisation, 5 µl of DAPI mixture (1 µg/ml DAPI) was applied to the well and mounted using a coverslip.

Hybridization Buffer

180 µl 5M NaCl
20 µl 1M Tris/HCl, pH 8
Xy MQ
Xy Formamide
1 µl 10% SDS

Table 5.2.2. Reagent volumes for FISH hybridization buffer according to varying formamide concentrations

FA%	0%	5%	10%	20%	25%	30%	35%	40%	45%	50%	55%	70%
NaCl	180 µl	180 µl	180 µl	180 µl	180 µl	180 µl	180 µl	180 µl	180 µl	180 µl	180 µl	180 µl
Tris	20 µl	20 µl	20 µl	20 µl	20 µl	20 µl	20 µl	20 µl	20 µl	20 µl	20 µl	20 µl
MQ	800 µl	750 µl	700 µl	600 µl	550 µl	500 µl	450 µl	400 µl	350 µl	300 µl	250 µl	100 µl
FA	0 µl	50 µl	100 µl	200 µl	250 µl	300 µl	350 µl	400 µl	450 µl	500 µl	550 µl	700 µl
SDS	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl

Washing Buffer

1 ml 1M Tris HCl, pH 8
Xy 5M NaCl
Xy 0.5M EDTA pH 8
Top up with ddH2O

Table 5.2.3. Reagent volumes for FISH washing buffer according to varying formamide concentrations

FA%	0%	5%	10%	20%	25%	30%	35%	40%	45%	50%	55%	70%
NaCl	9 ml	6.3 ml	4.5 ml	2150 µl	1490	1020	700	460	300	180	100	0
Tris	1 ml	1 ml	1 ml	1 ml	1 ml	1 ml	1 ml	1 ml	1 ml	1 ml	1 ml	1 ml
EDTA	0	0	0	500 ml	500 ml	500 ml	500 ml	500 ml	500 ml	500 ml	500 ml	500 ml
MQ	Add to 50 ml	Add to 50 ml	Add to 50 ml	Add to 50 ml	Add to 50 ml	Add to 50 ml	Add to 50 ml	Add to 50 ml	Add to 50 ml	Add to 50 ml	Add to 50 ml	Add to 50 ml

(v) Microscopy

A mounting medium was used to mount microscopy slides post IS-PCR/FISH, which contained mounting and antifade properties plus a generic nuclear counterstaining agent (1X phosphate-buffered saline, VECTASHIELD®, Fluoromount™, and DAPI (1.0 µg/ml); source: Medical University of Vienna). After application of the medium and a coverslip to the IS-PCR/FISH slide, visualization and photographic documentation of fluorescence was achieved using a Leica TCS SP2 confocal laser scanning microscope (Leica Microsystems GmbH, Wetzlar, Germany). DAPI staining was visualised through an ultraviolet/violet excitation (340-380 nm) band pass fluorescence filter, and 6-FAM green fluorescence was visualised through a blue excitation (450-490 nm) filter.

5.3 RESULTS

(i) Selection of IS-PCR/FISH primer and probe targets from whole genome comparisons

The selection of a *C. parvum*-specific DNA target for IS-PCR amplification largely relied on results from the cross-species genomic comparative analysis in chapter 1, which revealed a number of significant (>500 bp) stretches of sequence data that were present in *C. parvum* but absent from *C. hominis* (Figure 1.3.6). The *C. parvum*-specific target selected for this study comprises a ~8,200 bp sequence that extends from the 3'-end telomeric repeat (5'-CCTAAA-3') on chromosome 6 (Figure 5.3.1). This large subtelomeric fragment and its encoded gene sequences appear to be associated with a broader spectrum of zoonosis, as it has been conserved in *C. meleagridis*, a species which similar to *C. parvum* infects a number of animal hosts, and is furthermore the third most frequently isolated human-infective cause, but was absent from *C. baileyi*, which like *C. hominis* has evolved to become a more specialist *Cryptosporidium spp*. with a narrow host range.

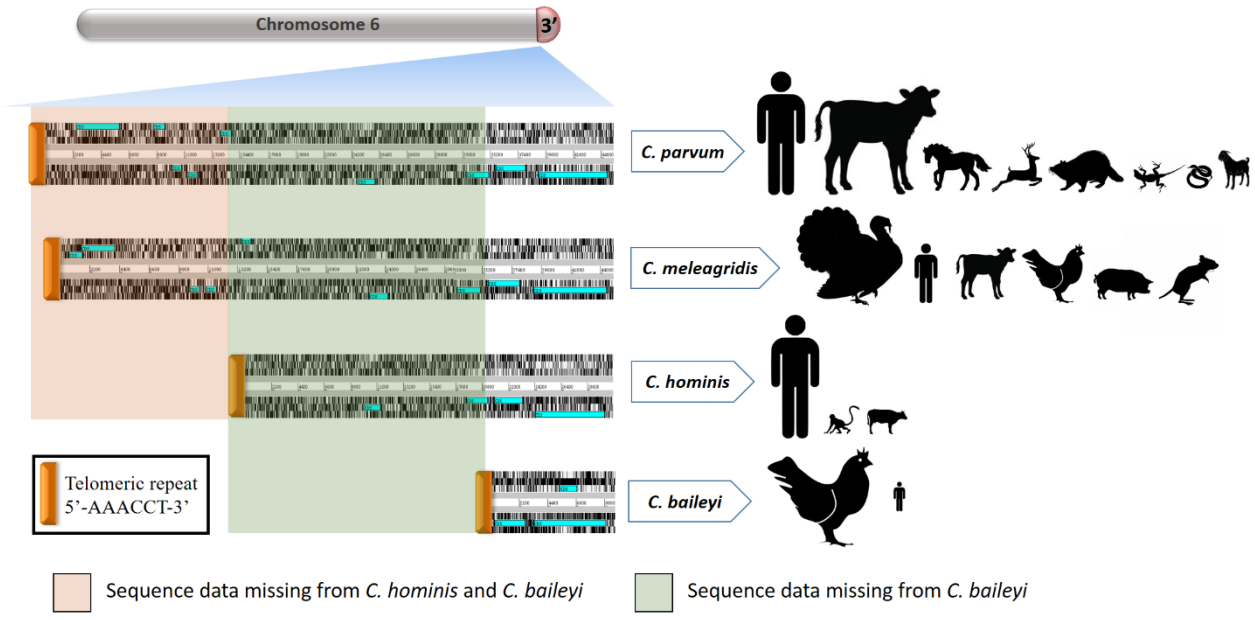


Figure 5.3.1. Cross-species ACT alignment at the 3'-telomere of chromosome 6 reveals sequence specificity to broad-range zoonotic *Cryptosporidium spp*.

Annotated views of putative protein coding sequences (ORFs > 150 bp) extending from the 3'-telomere of chromosome 6 were generated using the Artemis Comparison Tool WebACT (Abbott *et al.* 2007). FASTA files containing the first 40,000 bp of sequence data adjacent to the 3'-end of chromosome 6 were extracted and uploaded to WebACT from WGS sequence data for *C. parvum* subtype IIaA15G2R1 (UKP6, source: this study), *C. hominis* subtype IaA14R3 (UKH4, source: this study), *C. meleagridis* (UKMEL, source: http://www.cryptoDB.org), and *C. baileyi* (TAMU-09Q1, source: http://www.cryptoDB.org).

A 208 bp fragment within this species-specific target was chosen for the design of the In Situ PCR primers and FISH probe. The aim was to identify a target that was not only specific to the overall *C. parvum* species group, but in fact unique to the zoonotic *C. p. parvum* subspecies, which meant that the selected target further had to demonstrate sufficient intra-species polymorphism to ensure selective amplification. The 208 bp fragment used in this study was therefore chosen from a sufficiently polymorphic region of *C. parvum*/*C. meleagridis*-specific telomeric sequence data, which exhibited 100% conservation within *C. p. parvum* but a number of mismatched nucleotides between *C. p. parvum* and *C. p. anthroponosum*/*C. meleagridis* (Figure 5.3.2; Appendix XXVII).

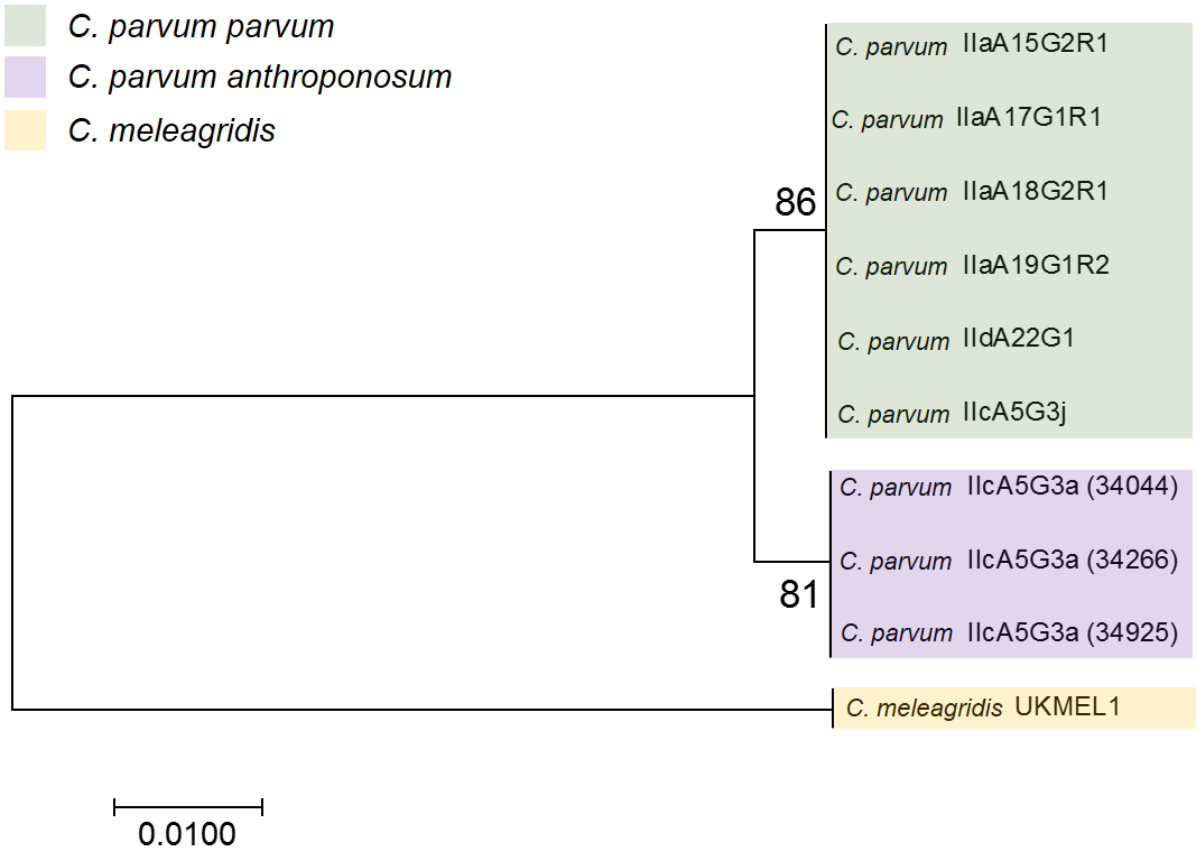


Figure 5.3.2. Maximum likelihood phylogeny of IS-PCR target amplicon reveals discriminatory specificity to the *C. p. parvum* subspecies

The evolutionary history between 9 unique *C. parvum* GP60 subtype WGS (source: this study) and *C. meleagridis* UKMEL1 (source: http://www.cryptoDB.org) was inferred for the 208 bp IS-PCR target sequence, using the Tamura-Nei substitution model in MEGA v 7.0 (Kumar *et al*. 2016). Six of the *C. parvum* WGS belong to the *C. p. parvum* subspecies (UKP2, UKP3, UKP6, UKP7, UKP8, UKP16) and three belong to the *C. p. anthroponosum* subspecies (UKP13, UKP14, UKP15). Phylogeny was calculated using the Nearest-Neighbour-Interchange (NNI) heuristic method, and 2,000 bootstrap replications were implemented to infer confidence values on the phylogenetic tree.

(ii) Molecular validation of new *C. p. parvum*-targeted PCR assay shows high species and subtype level specificity and sensitivity

The approximate sensitivity of this *C. p. parvum*-specific IS-PCR assay was tested through a limit of detection (LOD) test using purified *C. p. parvum* Iowa II genomic DNA. Using a ten-fold serial dilution, starting at a stock concentration of 100 ng/µl of purified *C. p. parvum* Iowa II reference DNA and ending at a highly-diluted DNA concentration of 10^{-8} ng/µl, the sensitivity of this assay (and hence the potential minimum detection capacity for *Cryptosporidium* oocysts) was evaluated. The dilution series showed that the PCR assay should be able to detect concentrations of *C. p. parvum* DNA as low as 10^{-4} ng/µl (Figure 5.3.3). At an approximate genomic weight of 3.9 x 10^{-5} ng DNA per oocyst (Figure 5.3.4), the results therefore demonstrate a theoretical LOD of 3 oocysts per µl of input.

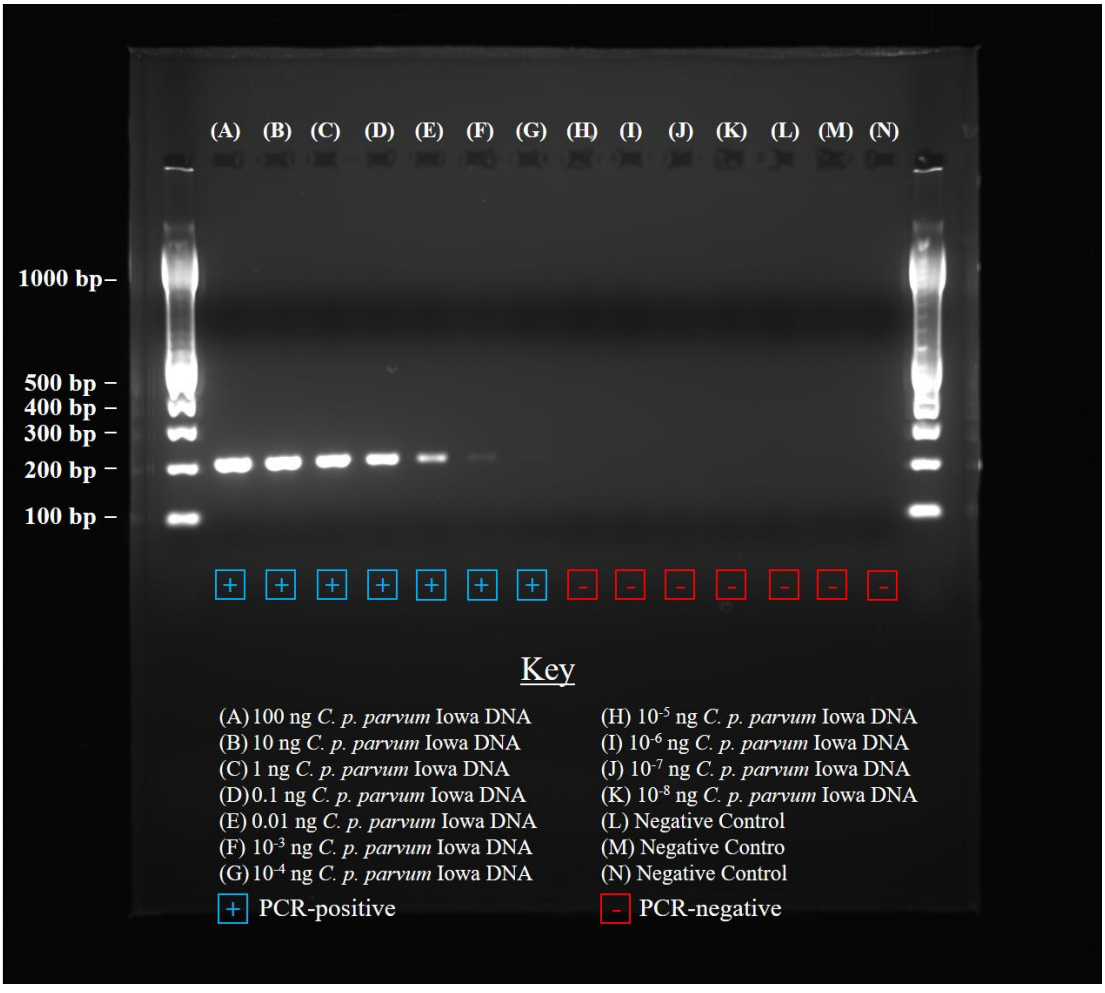


Figure 5.3.3. Limit of Detection of CP6T3 PCR protocol for *C. parvum* Iowa II genomic DNA

A ten-fold serial DNA dilution from a *C. parvum* Iowa II stock concentration of 100 ng/µl was performed (10^{0} to 10^{-8}), to determine the approximate minimum no. of oocysts the IS-PCR protocol would be capable of detecting. A triplicate of negative controls containing no DNA were additionally included. 5µl PCR product was run at 120V for 45 minutes on a 1.5% agarose gel.

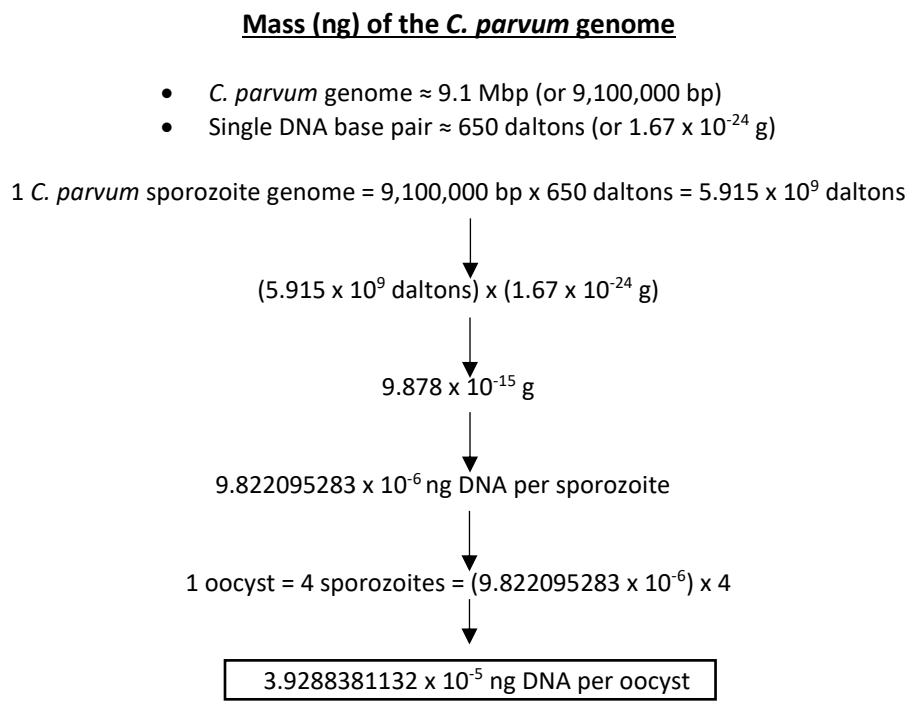


Figure 5.3.4. Flow diagram illustrating ng weight calculation of *C. p. parvum* haploid oocyst genomic DNA content
In order to calculate the approximate weight of a single *C. p. parvum* oocyst, the reported genome size for the reference *Cryptosporidium parvum parvum* Iowa II WGS (~9.1 Mbp; Abrahamsen et al. 2004) and 650 dalton weight for a single base pair (http://www.epibio.com/tech-support/useful-calculations) were used. A single *Cryptosporidium parvum parvum* sporozoite is thereby estimated to contain 9.822095283 x 10^{-6} ng DNA, while an oocyst (containing 4 sporozoites) is estimated to contain 3.9288381132 x 10^{-5} ng DNA.

The sensitivity of the assay was further explored through molecular validation, by testing the *C. parvum parvum*-specific primers on a panel of *C. p. parvum* (IIaA15G2R1, IIdAG2R1), *C. p. anthroponosum* (IIcA5G3a) and *C. hominis* (IaA14R3, IbA10G2) isolates, in addition to zoonotic species *C. meleagridis* and *C. cuniculus* (Figure 5.3.5). Results from triplicate PCR tests showed strong and 100% specific amplification of *C. p. parvum* isolates only, with no cross-reactivity demonstrated for any of the remaining *Cryptosporidium spp.* tested. Importantly, no cross-reactivity was seen for the closely related *C. p anthroponosum* isolate IIcA5G3a, which illustrates that this assay can selectively target strains within the *C. p. parvum* species-group and therefore efficiently discriminate between zoonotic and anthroponotic human-infective oocysts.

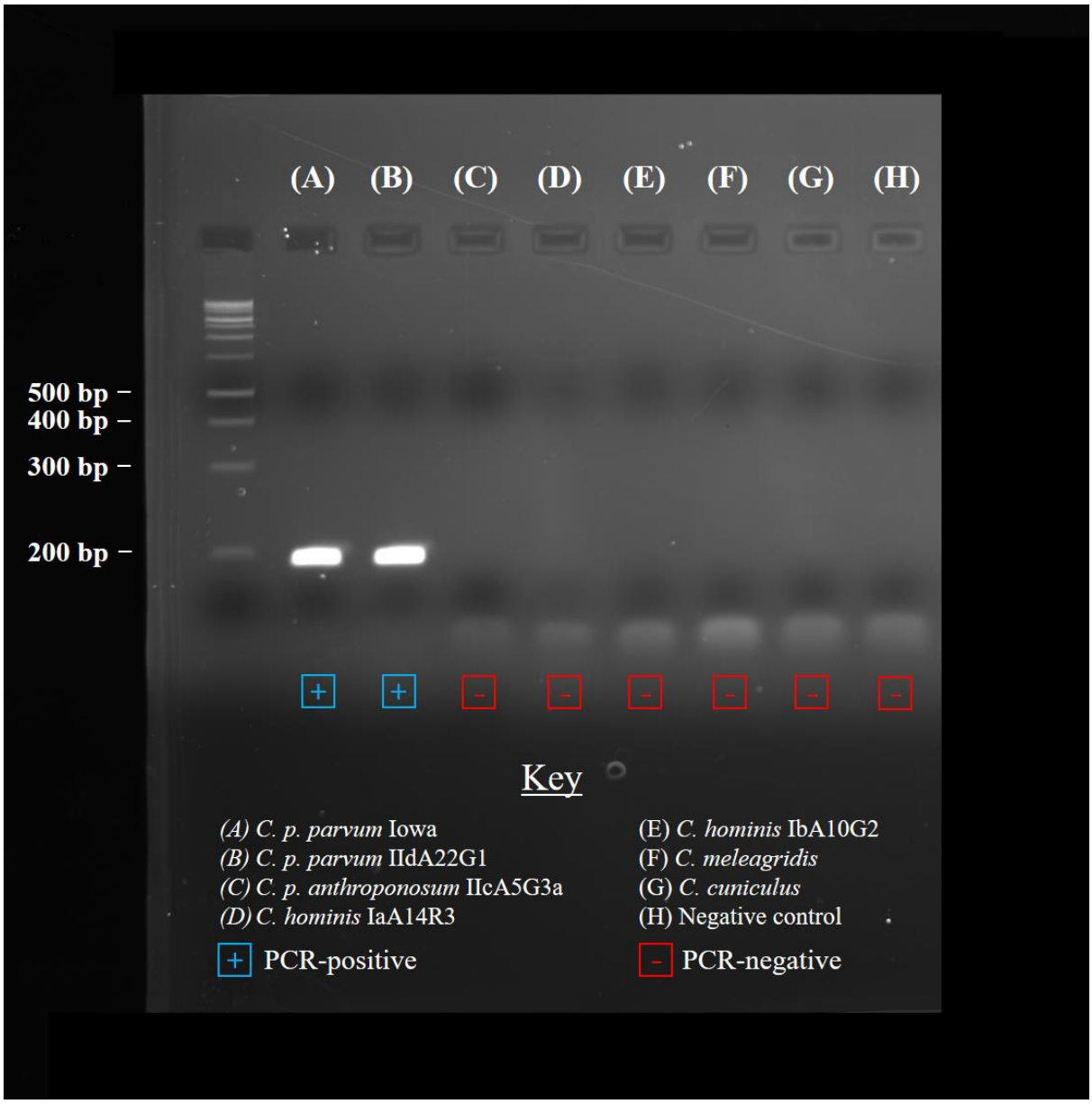


Figure 5.3.5. Specificity of CP6T3 *C. p. parvum*-specific primers across a range of zoonotic and anthroponotic *Cryptosporidium spp*. The species-specificity of the PCR assay was determined by positive and negative amplification of DNA from a variety of *Cryptosporidium spp*. Standardized DNA concentrations of 1 ng/µl input from *C. p. parvum* Iowa II (subtype IIaA15G2R1), *C. p. parvum* IIdA22G1, *C. p. anthroponosum* IIcA5G3a, *C. hominis* IaA14R3, *C. hominis* IbA10G2, *C. meleagridis* and *C. cuniculus* were tested in triplicate, exhibiting specificity (100%) to the zoonotic *C. p. parvum* human-infective species group. 5µl PCR product was run at 120V for 45 minutes on a 1.5% agarose gel.

(iii) Novel IS-PCR/FISH technique shows promising discrimination between *C. p. parvum* and *C. hominis* oocysts

A total of 127 trials were performed towards achieving optimization and validation of the proposed IS-PCR/FISH method, testing a variety of reagents at a number of concentrations and combinations based on previously-published recipes for IS-PCR and FISH in protozoa and bacteria. Out of 127 trials, 54 resulted in at least half (>50.0%) of observed oocysts exhibiting detectable fluorescence signals. These trials revealed a range of fluorescence, ranging from strong to no signal (Figure 5.3.6), but a majority of successful trials consistently showed significantly higher rates ($p<0.001$) of positive versus negative fluorescence for experimental versus control conditions (Figure 5.3.7). This trend was further validated through spectrophotometer absorbance readings at the optimal excitation wavelength for the FAM-labelled probe (495 nm), which repeatedly showed significantly ($p<0.01$) higher absorbance rates in the experimental compared to the control (Figure 5.3.8).

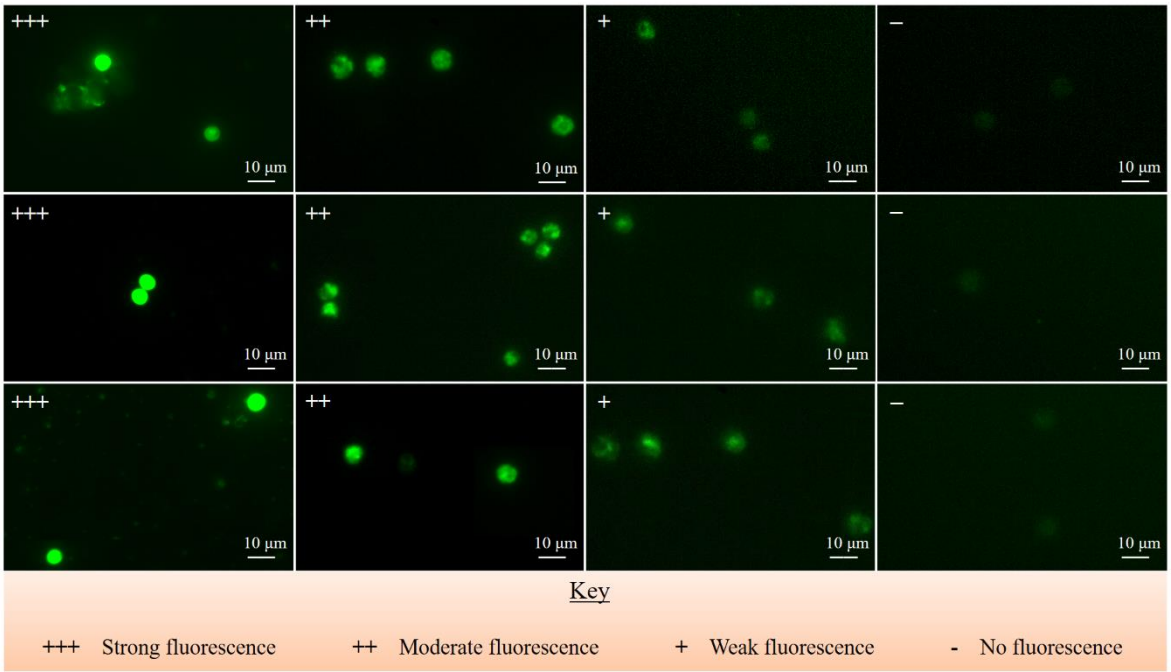


Figure 5.3.6. Fluorescence microscope images of *C. p. parvum* oocysts reveals variable fluorescence strengths post IS-PCR/FISH
Fluorescence microscopy images (60x magnification) of oocysts under the FITC channel demonstrates variable fluorescence stregths observed post IS-PCR/FISH trials. Oocysts are spherical/ovoid in shape and approximately 6 µm in size, and appear green with a black non-fluorescing background. Visualization strengths ranging from strong (+++), to moderate (++), to weak (+) and finally to negative (-) fluorescence are shown.

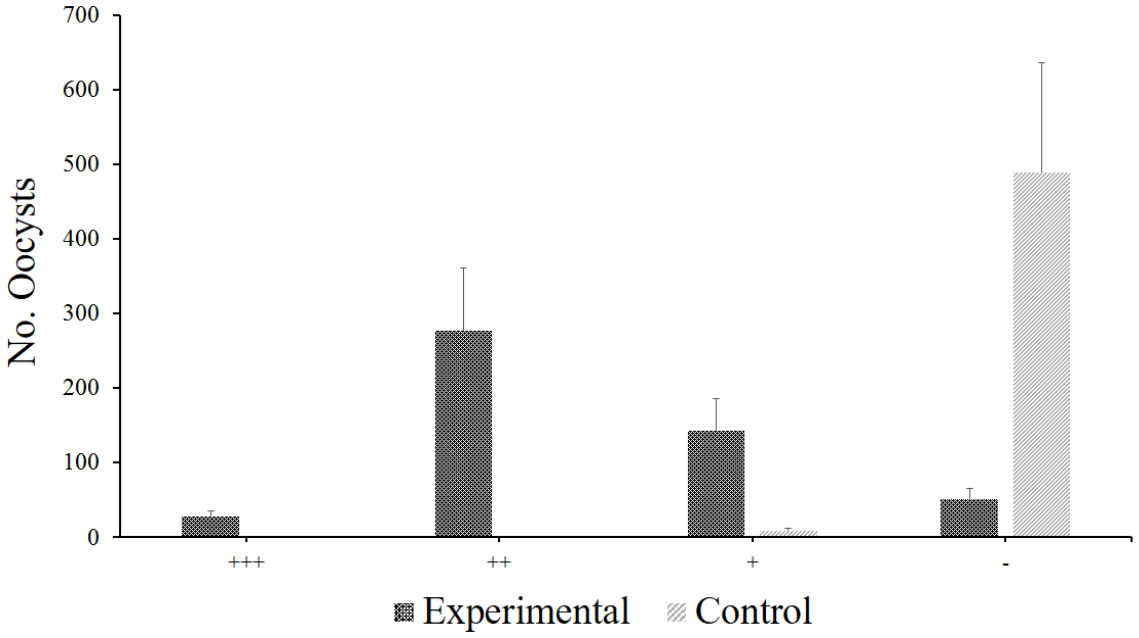


Figure 5.3.7. Average oocyst counts per 500 oocysts (with SEM error bars) detected for experimental versus control IS-PCR/FISH conditions, stratified based on strength of fluorescence (strong, moderate, weak, and no fluorescence)

The total number of oocysts detected in successful IS-PCR/FISH trials (i.e. at least one fluorescing oocyst under experimental conditions) were tabulated based on strong (+++), moderate (++), weak (+), or no (-) fluorescence, and are presented here in terms of proportion of oocysts per 500 oocysts detected. Control conditions refer to IS-PCR/FISH trials when Taq polymerase and/or primer was excluded from the IS-PCR. A statistically significant difference at $P<0.001$, as determined through a one-tailed student's T-test, was observed between the experimental and the control conditions for all fluorescence strengths.

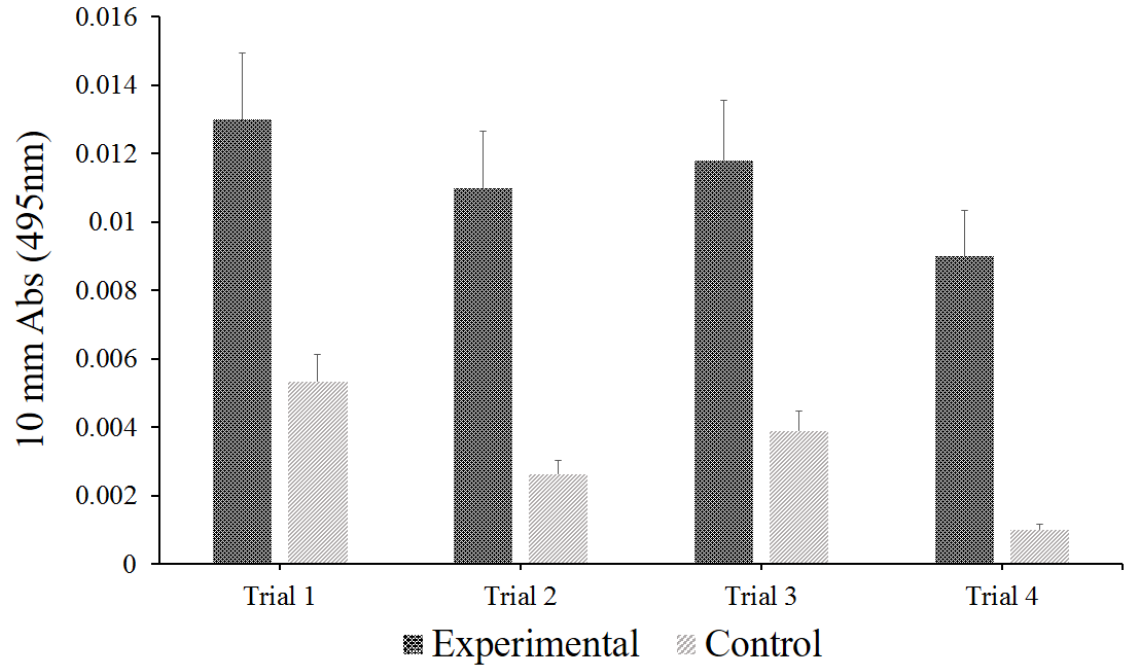


Figure 5.3.8. Absorbance rates of oocyst suspensions (with SEM error bars) following IS-PCR/FISH trials for experimental versus control conditions

Absorbance measurements (10 mm) of oocyst suspensions were taken following the final hybridization wash step in the IS-PCR/FISH protocol, measured at a 495nm (FITC absorbance peak) wavelength using a NanoDrop 1000 spectrophotometer (NanoDrop® ND-1000 Spectrophotometer V3.3; Thermo Fisher Scientific Inc). Control conditions refer to IS-PCR/FISH trials when Taq polymerase and/or primer was excluded from the IS-PCR. This experiment was performed in quadruplicate. A statistically significant difference at $P<0.01$, as determined through a one-tailed student's T-test, was observed for all oocyst counts between experimental and the control trials.

Composite fluorescence results for this study reveal a reliable molecular sensitivity and specificity of detection for this novel IS-PCR/FISH method. Oocysts were consistently visible at a range of magnifications (40x, 63x, and 100x), and often revealed a fluorescence signal that exceeded that demonstrated by the generic blue-fluorescent DNA stain DAPI (Figure 5.3.9). A comparative overlay of the two staining techniques did reveal a higher oocyst detection rate for DAPI, but showed that the IS-PCR/FISH approach frequently picked up a high proportion of DAPI-positive oocysts. This detection rate averaged at 78.6% across the entirety of successful trials conducted in this study (range: 57.1% - 100.0%, SEM± 0.07).

A final important validation component of this study was to test the potential cross-reactivity of this technique with non-*C. p. parvum* oocysts. A majority of trials therefore included *C. hominis* oocysts, which acted as both a negative control and species-specificity indicator for the IS-PCR/FISH technique. Fluorescence images of FITC versus DAPI staining revealed that *C. hominis* oocysts were consistently "invisible" post IS-PCR/FISH, a result which itself highlights the incomparably intense fluorescence frequently observed for the targeted *C. p. parvum* oocysts (Figure 5.3.10). More than 95% of trials revealed an absence of any fluorescence signal for runs involving *C. hominis* oocysts, with the remaining 5% infrequently showing oocysts with week to barely visible fluorescence (Figure 5.3.7). This demonstration of a reliable species-specificity of the designed primers and probe to the intended target, together with the high oocyst detection rate and sensitivity (both strength of fluorescence and the capacity to pick up oocysts at a range of magnifications), therefore speak to the overall success of this IS-PCR/FISH method, and the potential expansion of this technique to target and discriminate between a greater range of *Cryptosporidium* species and subtypes in future.

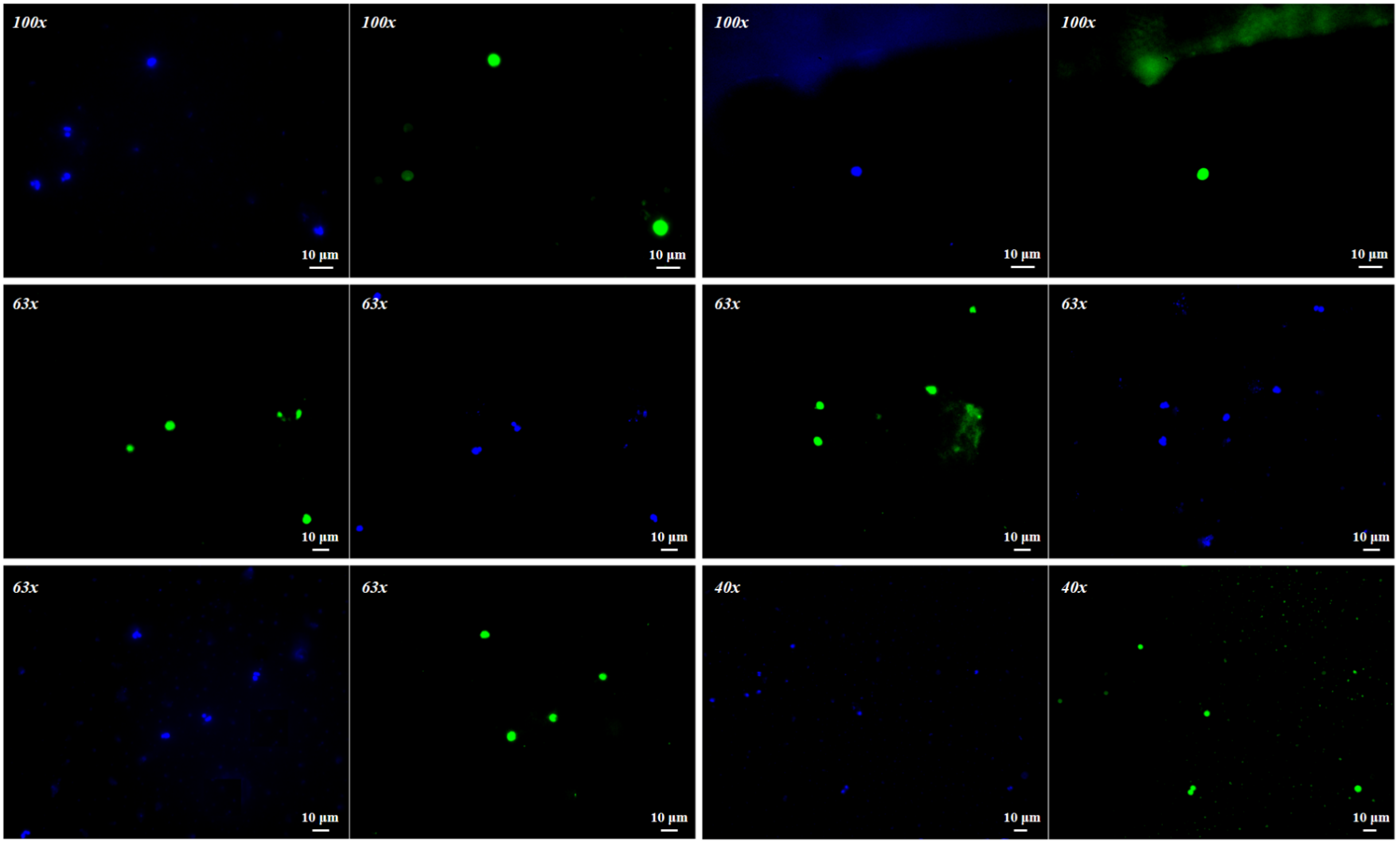


Figure 5.3.9. Fluorescence microscopy images of *C. p. parvum* IIa-subtype oocysts demonstrate sensitive and strong signals of detection post IS-PCR/FISH

Parallel fluorescence images of FITC (bright green) and DAPI (bright blue) staining of oocysts are shown, visualised through a blue (450-490 nm) and ultraviolet/violet (340-380 nm) excitation band pass filter, respectively. Oocysts appear as spherical objects and measure 4-6 µm across, with distinct sporozoites measuring 1-2 µm additionally visible by DAPI staining. Microscopy was performed using a Leica TCS SP2 confocal laser scanning microscope, at 40x, 63x and 100x magnifications.

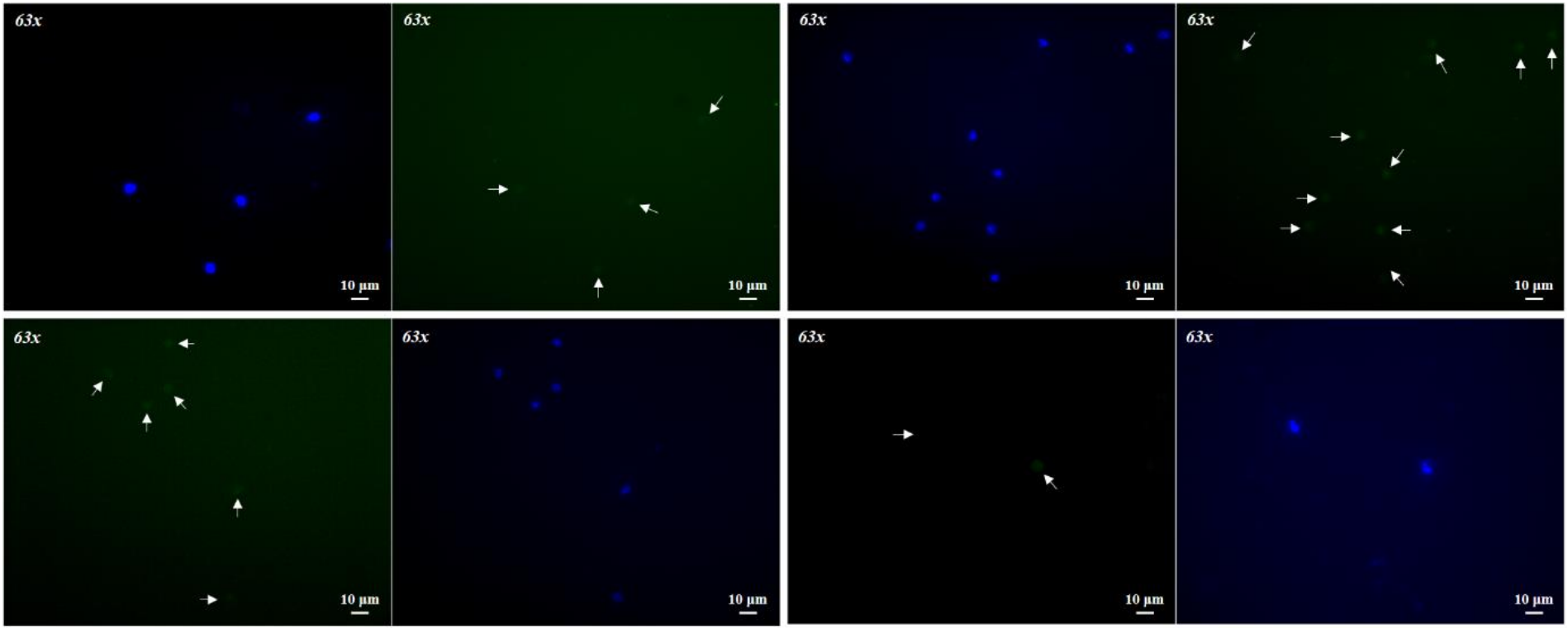


Figure 5.3.10. Fluorescence microscopy images of *C. hominis* Ia-subtype oocysts post IS-PCR/FISH demonstrate *C. p. parvum*-specificity of the 6-FAM-conjugated CPP6T3 probe

Parallel fluorescence images of FITC (bright green) and DAPI (bright blue) staining of oocysts are shown, visualised through a blue (450-490 nm) and ultraviolet/violet (340-380 nm) excitation band pass filter, respectively. White arrows point to oocysts that are positively fluorescing post DAPI staining but non-fluorescing under the FITC channel. Oocysts appear as spherical objects and measure 4-6 µm across, with distinct sporozoites measuring 1-2 µm additionally visible by DAPI staining. Microscopy was performed using a Leica TCS SP2 confocal laser scanning microscope, at a 63x magnification.

5.4 DISCUSSION

The current study presents results from one of the first validated IS-PCR/FISH protocols for molecular detection of *Cryptosporidium parvum parvum* oocysts. This fills a potentially important gap within *Cryptosporidium* diagnostics, as it provides the opportunity to approach species and subtype discrimination with an enhanced sensitivity and specificity control. Very few published protocols for IS-PCR in *Cryptosporidium* currently exist, which made methodological troubleshooting the most strenuous and time-consuming component of this study. The most comprehensive IS-PCR/FISH method that could be found through exhaustive review of the literature was presented as part of a "*Cryptosporidium* detection method" patent application to the United States government by Rochelle & Leon in 1997, but this protocol did not provide specific primer or probe sequences and only provided preliminary validation results (Rochelle *et al*. 1997), and the patent was deemed to be withdrawn in 2007. Another early study by Wagner-Wiening *et al*. (1998) similarly declined to include primer sequences and is only available in German, in addition to providing only limited agarose electrophoresis validation of the success of IS-PCR and without the incorporation of a FISH-based step. Only two more recent studies on IS-PCR in *Cryptosporidium* could be identified (Kimura, 2001; Chmura, 2010), both of which reported successful IS-PCR trials, but for neither study full text articles were available, and after repeated failed attempts to contact the authors the methodology remained inaccessible. The detailed methodology and extensive trial-and-error approach to optimization provided by this study is therefore somewhat novel within the field, and provides the opportunity for future adaptation, expansion or implementation of this technique for epidemiological and clinical purposes.

The sensitivity and specificity of the IS-PCR technique were validated through a number of approaches, and focused around the inclusion of appropriate negative and positive controls. This methodological validation included testing of cross-reactivity of primers with non-*C. p. parvum* extracted DNA (Figure 5.3.5) and whole *C. hominis* oocysts (Figure 5.3.10), both of which revealed significant specificity of the method by *C. p. parvum*. The inclusion of these negative controls and significant absence of cross-reactivity or false-positive results demonstrated the ability of the IS-PCR/FISH to sensitively and specifically detect strong fluorescence signals in *C. p. parvum* oocysts. The theoretical LOD of the

method, as described through PCR-based validation of a dilution series of *C. p. parvum* DNA, also revealed that the IS-PCR/FISH may be sensitive enough to pick up as few as 3 oocysts per µl of input, making the designed targets potentially valuable for largescale evaluation of oocyst contamination in environmental sources. The results therefore demonstrate a potentially comparable efficiency of detection to that traditionally demonstrated by the more gold standard molecular approaches, such as standard and real-time PCR, although this was not evaluated in the present study. In addition, the described method can theoretically be adapted to any 100-300 bp DNA fragment across the *Cryptosporidium spp*. genomes, which means a seemingly endless trove of targets exists that can be utilised for diagnostic purposes. This potentially expands the current toolbox of fluorescence-based detection of cryptosporidiosis to include countless potential targets for specific subtype and even strain discrimination.

Despite the initial success achieved through the novel IS-PCR/FISH technique, it is important to express and discuss a number of factors that imposed limitations on the present study. One important aspect to discuss is the choice of marker itself. The region of genomic sequence data that was chosen is present in *C. parvum* but absent from *C. hominis* (and divergent between *C. p. parvum* and *C. p. anthroponosum*), making it a good target for discrimination between the main human-infective strains. However, the sequence data is present in *C. meleagridis*, a more distantly related species, which means that the presence/absence of this stretch of DNA may not be evolutionarily stable. Since this sequence produces unexpected patterns of phylogenetic proximity between isolates compared to genome-wide topologies, it is possible that the presence and polymorphism of this subtelomeric stretch is variable at the intra-species and even intra-subtype level. Increased access to WGS across *C. hominis* and *C. parvum* subtypes to evaluate this genetic target is therefore an important next step in the validation of this method, and may eventually highlight an important potential limitation of the effective applicability of this technique for routine diagnosis. Sensitivity of detection was one aspect that showed significant promise overall, but the repeated presence of false-negative oocysts and predominance of moderate versus strong fluorescence in experimental trials indicate that extended troubleshooting may be needed to further enhance the ratio of positive versus negative fluorescence (Figure 5.3.7). However, FISH-based approaches for *Cryptosporidium spp*. in other studies similarly reported less than optimal (100%) levels of

fluorescence strength and rates of detection (Deere *et al*. 1998; Alagappan *et al*. 2008; Alagappan *et al*. 2009), which means that while there may be room for improvement, the coverage achieved by this preliminary IS-PCR/FISH method is possibly in line with acceptable scientific standards. The validation of assay specificity to *C. p. parvum* oocysts also experienced a number of limitations, most notably as a result of a lack of access to a wide panel of GP60 subtype oocysts. The unavailability of an inexpensive and simple in vitro approach to *Cryptosporidium* propagation means that only a very limited number of institutions around the world are capable of providing live oocysts, and the availability of oocysts from these institutions in turn depends on the success and frequency of oocyst harvesting from animal models. This means that only a few human-infective GP60 subtypes (most often zoonotic IIa, which readily propagates in animal systems) are commonly available, and meant that the specificity of the IS-PCR/FISH method could only be validated in what happened to be available at the time of this study. Use of stored historical samples (available from large reference units such as the CRU in the United Kingdom) was also considered but subsequently excluded from the present study, as the extended storage of oocysts frequently damages oocyst viability and structural integrity (Deere *et al*. 1998; Bednarska *et al*. 2007). This meant that only fresh oocysts were applicable for validating the species-specificity of the IS-PCR/FISH method. The consequence of this limitation meant that only two oocyst subtypes could be compared, *C. p. parvum* (IIa-subtype family) and *C. hominis* (Ia-subtype family), and that the true strength and breadth of species and subspecies-specificity remains to be investigated. However, tests on DNA extracts from a number of zoonotic and anthroponotic *Cryptosporidium spp*. (Figure 5.3.5) showed that this method has a potentially promising discriminatory capacity. It also indicates that the extension of this technique into a multiplex IS-PCR/FISH, with differentially-labelled probes capable of simultaneous detection and discrimination of different subtypes in mixed samples, is within the realm of possibilities.

The final topic of discussion worth having is the feasibility and applicability of implementation. This involves an honest evaluation of the benefits and disadvantages presented by the IS-PCR/FISH approach to *Cryptosporidium* diagnosis. The primary benefits of the current study can, in my mind, easily be defined: firstly, the specificity control one gains from being able to detect the species and potentially subtype of an

infectious organism through any single-copy genetic target; and secondly, the simple and rapid nature of the interpretation of the results. Many of the time-consuming and expensive preparatory steps of conventional PCR are additionally avoided, such as the not-so-easy process of DNA extraction and purification from oocysts, and the messy process of pouring and loading gels for electrophoresis. But what IS-PCR/FISH gains in terms of benefits and improvements, it more than makes up for in losses and impracticalities. The entire protocol described in this study takes more than twenty hours to complete, and requires an impressive list of more than 30 reagents and pieces of equipment, including a pricey thermocycler and fluorescence microscope. It also requires a range of storage and incubation conditions, spanning from sterile fume hoods to -20°C freezers to 4°C fridges to 37°C incubators to 48°C water baths. This means that there is a signifcant financial and logistical pre-requisite for being able to successfully implement an IS-PCR/FISH based strategy, which instantaneously segregates the accessibility of the technique between resource-poor and resource-rich settings. Hence, regions where cryptosporidiosis is associated with the greatest incidence rates, morbidity and mortality, such as poverty-stricken Sub-Saharan Africa, will likely lack the basic resources needed to apply this method routinely. Even in places where funding and resources may be available, such as the well-equipped reference laboratories of Western Europe, staff may consider the process too complex to incorporate it into routine diagnostic practices. The immediate integration of the IS-PCR/FISH method into widescale diagnostic use therefore does not seem practical or feasible; however, the potential relevance of this study for the development of new and advanced detection methods should nevertheless not be underestimated. One area of particular importance could be the adaptation of IS-PCR/FISH probes into microfluidic DNA chip-based technologies for rapid diagnostic testing of cryptosporidiosis (Reinholt *et al*. 2014). In recent years, significant interest and investment in these “lab-on-a-chip” (LOC) pathogen detection devices that are fast, accurate, low maintenance and portable has increased (Yager *et al*. 2006; Mairhofer *et al*. 2009; Lee *et al*. 2010; Foudeh *et al*. 2012), and a number of these have already been rolled out for use in field settings for diseases such as HIV (Sia *et al*. 2004; Lee *et al*. 2008), Malaria (Gascoyne *et al*. 2004; Antia *et al*. 2007), and Tuberculosis (Ke *et al*. 2004). A number of comparable microfluidic devices have been developed for detection of *Cryptosporidium* oocysts (Esch *et al*. 2001; Taguchi *et al*. 2007; Angus *et al*. 2012; Zhang *et al*. 2015), but most of these are only suitable for water samples and not evaluated or

adapted for clinical use. In addition, they are unilaterally developed for *Cryptosporidium parvum* detection and so cannot provide any discriminatory species data or even target non-*C. parvum* oocysts. With improved availability of WGS for *Cryptosporidium spp*., enhanced reports of the genetic diversity of human infections, and increasing knowledge of the true impact and consequences of cryptosporidiosis, there is an ever-growing need for enhanced diagnostic approaches that combine efficiency-of-use with accuracy-of-detection. A rapid diagnostic tool that could discriminate between the many subtypes and species responsible for human cryptosporidiosis would be valuable for both clinical and epidemiological purposes. The evaluation of DNA probes for species-specific detection through colour-change technologies, as was one of the primary aims of this study, is therefore potentially crucial for bridging the gaps between in silico analyses, molecular adaptation, and clinical implementation in future.

Cryptosporidiosis is a disease that remained largely unresearched for the first three-quarters of the 20th century. Subsequent advancements in understanding the severity of burden imposed by this disease worldwide, particularly in at-risk populations such as children under 5 years and immunocompromised individuals, led to the inclusion of this disease on the WHO Neglected Diseases Initiative in 2004 (Savioli et al. 2006), and since then the public health profile and scientific understanding of this disease continues to grow. To date, at least 26 species have been described, ubiquitous in nature and specialised to infect a broad range of hosts that includes humans, cattle, marsupials, amphibians, reptiles, felines, fish, rodents, pigs, sheep and rabbits (Rollinson & Stothard, 2015). The genetic divergence that separates host-adapted isolates into species according to common molecular markers such as 18S rRNA is well-characterised, and GP60-based subtyping has provided insight into the heterogeneity and recombination that further separates human species into strains and subtypes. A number of comparative genomics studies have previously described how features such as gene loss/duplication events and hypervariability have driven species evolution and strain emergence in *C. hominis* and *C. parvum*, but the diversity of subtypes explored to date remains narrow, the recombination that has been characterised between phenotypically-diverse subtypes is limited, and the extent to which single-marker variability correlates with genomic divergence remains largely unexplored. This three-year PhD project was built around trying to fill some of these gaps, aiming to provide greater descrimination of evolutionary divergence between and within human-infective species, while simultaneously characterising the genetic drivers of host range and virulence and expanding on the global picture of subtype distribution, population genetics and host type epidemiology (Figure 6.1). The concatenated phylogeny and whole genome divergence presented in chapter 1 provided novel genetic validation that IIcA5G3a/k subtype *Cryptosporidium* isolates form a separate subspecies within the *C. parvum* species group, and that reclassification of zoonotic and anthroponotic subtypes as *C. parvum parvum* and *C. parvum anthroponosum* respectively is appropriate. The observation of a IIc-type GP60 locus in a predominantly IIa-type genome for subtype IIcA5G3j warranted the extensive recombination analyses presented in chapter 2, which provided a detailed

characterisation of genome-wide recombination events across subtypes IIaA15G2R1, IIcA5G5a, IIcA5G3j, and IbA10G2, and generated estimated divergence dates for some of the important speciation and subspeciation events in human-infective *Cryptosporidium spp*. Variability at the GP60 locus was reaffirmed as one of the predominant examples of recombination-driven divergence in both *C. hominis* and *C. parvum*, which motivated the population genetics piece presented in chapter 3. Defining population genetics and differentiation of GP60 in *C. hominis* isolates served to increase our understanding of allelic diversity at this locus, and assess how allelic distribution and population structure differ between the various subtype families. Results showed that subtype emergence and evolution are subject to differential degrees of selective pressure and divergence between subtype families, that certain subtype populations in defined geographic regions (such as subtypes Ia and Id in Central Asia and Sub-Saharan Africa) show evidence of recent population expansion events, and that population structures vary from well-defined to poorly-defined for a majority of subtype Ib versus Ia/Id populations, respectively. Significant epidemiological differences between 'virulent' Ib-type isolates and 'avirulent' Ia-type isolates also emerged through the GP60 systematic review, inspiring the exhaustive comparative genomics study presented in chapter 4. New data describing genomic restructuring (missing telomere ends), subtype-specific coding sequences and patterns of selective pressure supplement previous descriptions of the genome-wide drivers of lineage-specific divergence between subtypes Ia and Ib (Guo *et al*. 2015). The nature of divergence was further dissected to describe a subtelomeric localization bias of hypervariability and positive selective pressure, and the relative contribution of indels versus SNPs was additionally described, identifying indel diversity as the more prominent cause of genomic divergence. The final chapter of this thesis aimed to demonstrate the translation of in silico results into an applied method, using species-specific subtelomeric sequence data to develop a novel combined IS-PCR/FISH method for sensitive and specific detection of *C. p. parvum* oocysts. Extensive optimization and molecular validation of the primers and probe on a range of *Cryptosporidium spp*. DNA and *C. hominis* oocysts indicate efficient detection and discrimination of this technique. These results show promise for the further development of this technique to achieve strain discrimination, a technique that would be useful for tracking strain emergence and spread during outbreaks, as well as the adaptation of this method to microfluidics DNA chip-based rapid diagnostic tests. The advancements that

this thesis has achieved within *Cryptosporidium* genomics, epidemiology and molecular diagnostics are hereby summarized, and the improved understanding of how human-infective strains, subspecies and species emerge, evolve, and adapt within this genus will hopefully provide a strong basis of knowledge on which future studies can build. The data presented in this thesis has thereby provided novel insight into the way genetic diversity manifests itself and influences phenotypic diversity in *C. hominis* and *C. parvum*, moving us one step closer towards accumulating vital in silico data to support the prevention, treatment, and eventual elimination of this neglected cause of diarrhoeal disease worldwide.

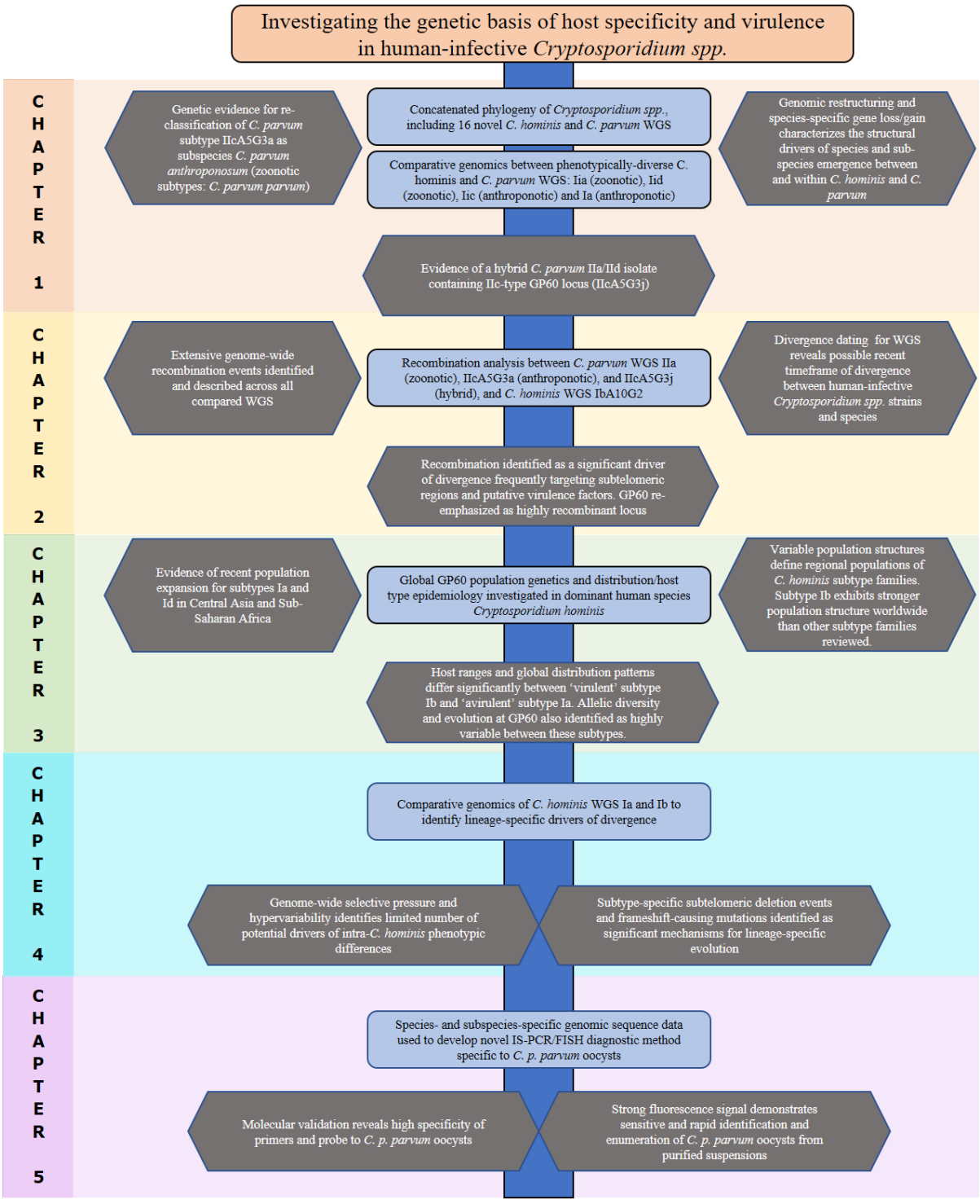


Figure 6. Summary of thesis study aims and results

The aims and main results of individual chapters of this PhD thesis are described and summarized. Study aims are presented in blue, while results and conclusions are illustrated in dark grey.

Appendix I

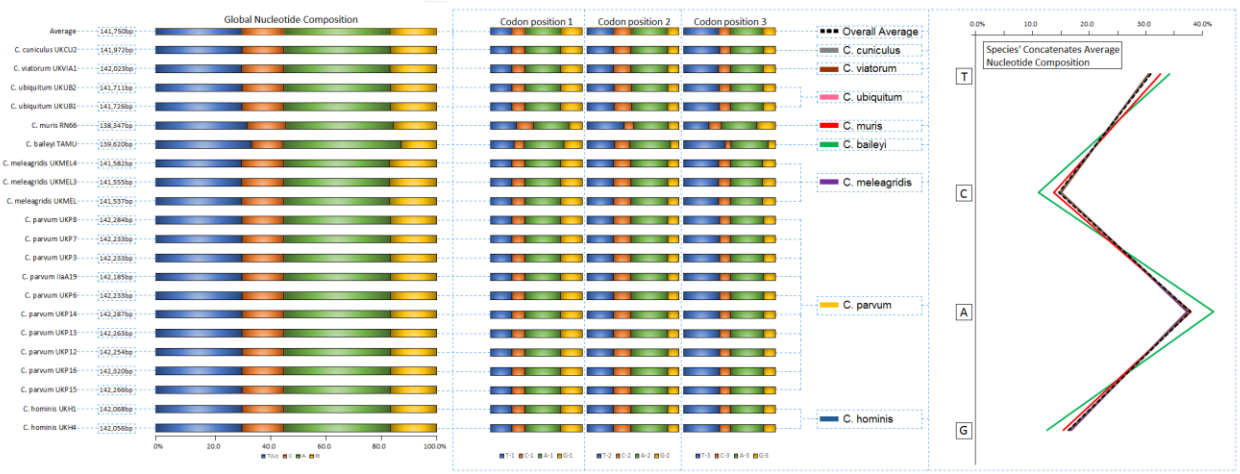
Whole Genome Sequence (WGS) Assembly Summary Statistics

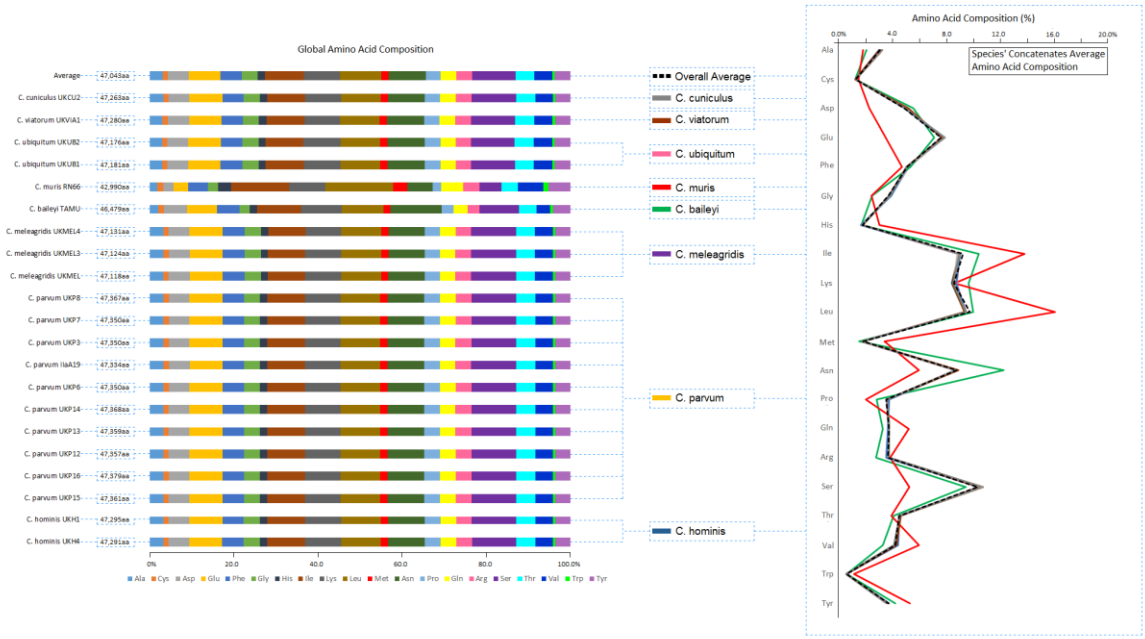
Assembly summary statistics are reported for the 20 WGS presented in Table 1.2.1. These include the length of the WGS (bp), the number of scaffolds, mean scaffold length, longest scaffold length, scaffold N50, and the total number of non-specific (N) nucleotides.

WGS ID	Length (bp)	No. Scaffolds	Mean Scaffold Length	Longest Scaffold Length	N50	Total Length of (N) Characters
RN66	9245251	97	95311.9	1182920	520347	6471
TAMU-09Q1	8502994	145	58641.3	145259	203018	0
UKCU2	7853083	7881	996.5	12218	1646	1
UKH1	9141398	156	58598.7	245088	179408	0
UKH4	9502496	2375	4001.1	283681	29100	22892
UKMEL1	8973224	57	157425	732862	322908	0
UKMEL3	12181424	15386	791.7	34869	4134	258412
UKMEL4	9663721	4645	2080.5	54682	9598	55969
UKP2	9759982	3084	3164.7	418246	105241	6400
UKP3	9172817	919	9981.3	247529	57514	22631
UKP6	9112937	18	506274.3	1296567	1013555	0
UKP7	9035974	1158	7803.1	254551	30196	18798
UKP8	9377202	1792	5232.8	302930	41951	20766
UKP12	8038639	7357	1092.7	16726	1842	12
UKP13	8416777	11108	757.8	16088	1315	7
UKP14	8594935	4847	1773.2	36377	3661	36
UKP15	10134261	6085	1665.4	102582	11563	4
UKP16	9496690	3527	2692.6	127893	15768	2
UKUB1	9060260	873	10378.3	223141	38238	14303
UKUB2	9069162	680	13337	261712	46750	10231

Appendix II

Nucleotide Composition (A, T, G, C) and Amino Acid Composition of Concatenated Sequences
Sequence compositions of 21 *Cryptosporidium spp*. concatenations were generated in tabulated format using MEGA v5.2 (Molecular Evolutionary Genetics Analysis version 5.2; Tamura *et al.* 2011). High sequence similarity (>99.0% AA IDs) between WGS isolates belonging to *C. hominis*, *C. parvum*, and *C. meleagridis* species types resulted in average values being implemented for graphical representation of comparative trends.





Appendix III

Amino acid composition, putative protein function and predicted protein localization of genes included in the concatenated phylogeny

Putative protein function of 61 genes included in the concatenated phylogeny were ascertained through annotations on CryptoDB (http://cryptoDB.org) and through Blasting amino acid sequences against the UniProt Database Resource (Cut-off E-value 10^{-10}; http://www.uniprot.org/). Predicted protein localizations were determined using WoLF PSORT (Protein Subcellular Localization Prediction Tool; https://wolfpsort.hgc.jp/). Localizations were categorized as follows: nucleus (nucl), cytoplasm (cyto), mitochondria (mito), cytoskeleton (cysk), plasma membrane (plas), golgi apparatus (golg), extracellular (extr), and peroxisome (pero). Microsoft Excel (2013) was used to determine the composition of amino acids for each protein, according to the twenty standard amino acids: alanine (A), arginine (R), asparagine (N), aspartic acid (D), cysteine (C), glutamine (Q), glutamic acid (E), glycine (G), histidine (H), isoleucine (I), leucine (L), lysine (K), methionine (M), phenylalanine (F), proline (P), serine (S), threonine (T), tryptophan (W), tyrosine (Y), and valine (V).

CryptoDB ID	Putative Protein Function	Predicted Protein Localization	Protein Length (AA)	Amino Acid Composition (%)																		
				A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
Chro.10151	Glycosyltransferase family 41	**Plas**	1049	3.2%	2.9%	9.1%	5.5%	3.4%	2.6%	6.1%	3.7%	2.2%	12.2%	11.4%	7.0%	1.4%	5.7%	3.2%	8.3%	3.2%	0.7%	4.9%	3.2%
Chro.10166	Uncharacterized	**Cyto**	290	4.5%	7.3%	7.0%	3.1%	1.0%	3.1%	7.3%	3.5%	2.1%	11.2%	10.5%	10.1%	1.0%	3.8%	2.1%	10.8%	2.8%	0.3%	2.8%	5.2%
Chro.10167	Uncharacterized	**Extr**	797	2.4%	2.8%	9.1%	4.5%	1.1%	4.5%	7.4%	1.4%	1.5%	8.9%	13.1%	11.7%	1.8%	4.3%	1.9%	10.7%	4.0%	0.6%	3.1%	5.1%
Chro.10199	Haloacid dehalogenase family	**Extr**	420	4.8%	3.1%	8.5%	4.1%	1.2%	1.4%	8.0%	6.0%	1.2%	8.2%	9.9%	8.2%	1.9%	5.8%	2.9%	10.4%	3.6%	0.0%	3.4%	7.2%
Chro.10229	U3 small nucleolar RNA-associated	**Nucl**	708	2.9%	5.0%	11.0%	4.9%	0.9%	5.3%	9.3%	4.0%	1.7%	8.8%	8.3%	11.0%	2.4%	2.6%	2.6%	9.2%	3.6%	1.1%	2.4%	2.9%
Chro.10411	Uncharacterized	**Extr**	356	4.3%	4.6%	6.0%	3.7%	1.7%	3.1%	5.7%	1.1%	1.4%	10.3%	12.3%	13.1%	1.1%	5.4%	2.8%	12.8%	3.1%	0.9%	3.1%	3.4%
Chro.10424	Transcription regulator	**E.R.**	826	4.8%	4.1%	4.3%	5.4%	0.2%	3.8%	12.4%	3.2%	1.5%	5.4%	9.2%	9.7%	1.4%	3.7%	4.4%	15.6%	3.6%	0.1%	2.7%	4.3%
Chro.10425	Leucine-rich repeat protein	**Plas**	1265	4.7%	3.3%	8.1%	6.1%	0.6%	3.5%	9.2%	3.9%	1.8%	8.1%	8.0%	9.7%	1.3%	5.3%	4.7%	9.1%	4.5%	0.5%	2.9%	4.6%
Chro.20024	Uncharacterized	**Cyto**	200	1.5%	3.0%	9.5%	5.5%	1.0%	5.5%	13.5%	5.0%	1.0%	9.5%	11.0%	14.0%	2.0%	1.5%	0.5%	6.0%	3.0%	0.5%	3.0%	3.5%
Chro.20105	Uncharacterized	**Plas**	1079	3.3%	2.5%	9.0%	5.4%	1.2%	4.5%	5.3%	4.3%	2.1%	7.5%	7.1%	11.1%	1.3%	3.2%	4.3%	11.3%	5.6%	0.6%	2.7%	7.5%
Chro.20144	Chaperone Protein	**Cyto**	285	1.4%	2.8%	8.1%	4.6%	3.5%	1.1%	3.5%	9.5%	1.4%	11.2%	8.8%	11.9%	2.1%	4.6%	4.6%	8.8%	4.6%	0.4%	2.8%	4.6%
Chro.20223	Transcription Factor	**Plas**	709	3.0%	4.1%	8.3%	3.9%	1.4%	4.7%	6.6%	3.7%	1.1%	11.1%	12.0%	7.9%	1.3%	4.5%	2.8%	11.3%	4.2%	1.0%	3.9%	3.1%
Chro.20326	Uncharacterized	**Plas**	1034	6.9%	5.1%	4.8%	3.7%	1.5%	3.1%	4.4%	3.9%	1.5%	7.1%	9.5%	3.8%	2.3%	4.3%	6.6%	12.0%	8.1%	0.6%	4.4%	6.7%
Chro.20388	Peptidylprolyl isomerase	**Nucl**	494	2.4%	6.7%	7.7%	5.7%	1.6%	5.5%	7.3%	8.3%	1.4%	4.0%	4.0%	11.9%	1.6%	3.8%	4.7%	10.9%	5.7%	0.6%	3.4%	2.6%
Chro.20406	Interacting factor-like phosphatase	**Nucl/Cyto**	473	1.9%	2.7%	8.9%	4.9%	0.6%	2.5%	6.8%	2.5%	1.9%	8.9%	7.6%	9.9%	0.8%	4.4%	4.4%	17.8%	5.3%	0.6%	4.2%	3.2%
Chro.30055	Uncharacterized	**Nucl**	1079	4.7%	3.3%	11.8%	3.8%	0.8%	3.4%	6.1%	9.0%	3.3%	6.4%	6.3%	6.9%	1.8%	2.8%	6.1%	10.9%	5.6%	0.4%	2.7%	3.8%
Chro.30132	Uncharacterized	**Cyto**	251	2.0%	4.8%	7.2%	3.6%	1.6%	4.0%	7.6%	3.6%	0.8%	10.0%	7.2%	8.0%	1.2%	8.8%	1.2%	13.9%	2.8%	1.2%	6.0%	4.8%
Chro.30206	Uncharacterized	**Extr**	274	4.4%	2.2%	5.2%	1.9%	3.7%	4.1%	7.0%	2.6%	0.7%	5.2%	11.9%	11.5%	2.6%	7.4%	1.9%	12.2%	3.7%	0.4%	6.7%	4.8%
Chro.30299	Uncharacterized	**Plas**	1201	2.9%	2.4%	10.2%	3.3%	1.8%	2.8%	7.5%	1.9%	1.5%	14.9%	11.5%	7.7%	1.5%	7.2%	1.9%	8.5%	4.9%	0.7%	4.2%	2.6%
Chro.30349	WD-40 repeat-containing protein	**Cyto**	595	2.5%	2.0%	10.8%	3.7%	2.0%	2.9%	8.2%	3.5%	1.2%	12.8%	10.3%	8.9%	0.8%	5.7%	0.8%	10.9%	4.5%	0.7%	3.5%	4.2%
Chro.30377	Uncharacterized	**Cyto**	721	2.1%	3.9%	9.7%	7.0%	1.1%	4.4%	15.9%	3.5%	1.3%	7.5%	9.2%	9.2%	1.1%	3.5%	1.0%	9.7%	3.5%	0.4%	3.2%	2.7%
Chro.30413	Centrin like EF hand protein	**Cyto**	472	0.6%	2.1%	10.0%	5.1%	0.0%	3.4%	12.1%	3.0%	1.1%	11.0%	12.7%	12.5%	1.3%	5.3%	0.8%	7.6%	3.6%	0.4%	4.2%	3.2%
Chro.30476	Kinesin-5	**Plas**	558	1.3%	6.1%	10.4%	3.4%	0.7%	3.0%	6.5%	3.6%	1.8%	9.5%	12.5%	8.6%	2.5%	6.6%	2.0%	8.8%	2.9%	0.9%	5.2%	3.8%
Chro.40248	Dumpy-19 like protein	**Plas**	1133	2.7%	3.6%	7.6%	3.5%	1.5%	2.8%	4.8%	2.4%	1.7%	12.0%	10.9%	7.6%	2.1%	9.1%	2.3%	11.0%	4.1%	1.2%	5.6%	3.4%
Chro.40252	Adaptin	**Cyto**	899	3.8%	3.0%	9.8%	4.2%	2.0%	3.7%	7.1%	2.3%	1.6%	11.6%	11.5%	8.1%	1.8%	3.3%	2.7%	10.2%	4.6%	0.4%	4.3%	4.0%
Chro.40294	Uncharacterized	**Nucl/Cyto**	330	3.1%	6.5%	9.6%	7.4%	0.9%	2.5%	9.6%	2.5%	1.2%	9.3%	8.6%	9.9%	1.5%	7.1%	0.6%	10.2%	2.5%	0.0%	3.7%	3.4%
Chro.40317	ZincFinger	**Cyto**	336	1.8%	2.7%	12.4%	6.1%	0.9%	3.6%	10.3%	1.5%	1.2%	7.6%	7.3%	15.2%	1.2%	4.8%	2.7%	10.3%	5.2%	0.3%	2.4%	2.4%
Chro.40433	Rrp8p like methyltransferase	**Cyto**	375	1.4%	3.5%	10.0%	3.8%	1.9%	3.3%	4.3%	5.1%	2.4%	9.2%	9.2%	13.0%	1.4%	5.4%	2.2%	10.3%	4.9%	0.8%	4.9%	3.0%
Chro.40495	Ring finger domain containing protein	**Nucl**	658	2.1%	4.9%	11.7%	5.3%	2.4%	3.8%	8.7%	3.6%	1.7%	7.4%	8.1%	9.4%	1.1%	2.6%	2.6%	12.9%	5.9%	0.3%	2.4%	3.0%
Chro.40503	ATP-binding cassettetransportera bc1	**Plas**	1344	2.9%	3.4%	8.0%	3.9%	1.8%	3.0%	6.1%	3.3%	1.6%	11.8%	12.2%	7.6%	1.8%	7.0%	1.9%	8.8%	4.6%	0.7%	5.3%	4.3%

CryptoDB ID	Putative Protein Function	Predicted Protein Localization	Protein Length (AA)	Amino Acid Composition (%)																		
				A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
Chro.50012	Uncharacterized	**Plas**	457	2.4%	4.6%	5.9%	4.6%	1.3%	1.5%	8.8%	2.0%	0.9%	9.6%	11.8%	9.0%	2.4%	9.4%	2.2%	10.5%	5.7%	0.4%	3.1%	3.9%
Chro.50084	Uncharacterized	**Extr**	351	1.4%	3.2%	11.6%	4.3%	2.6%	2.3%	4.0%	4.0%	2.3%	12.1%	8.4%	7.2%	1.4%	9.8%	2.9%	10.1%	3.2%	0.9%	3.8%	4.3%
Chro.50103	Caldesmon1	**Cyto**	454	1.8%	7.5%	7.3%	4.8%	1.3%	4.6%	12.1%	4.8%	2.0%	8.1%	10.8%	12.1%	1.8%	2.0%	2.9%	5.9%	2.2%	0.2%	4.0%	3.7%
Chro.50107	Dynein beta chain	**Extr**	337	3.9%	3.3%	8.1%	2.7%	2.1%	2.7%	7.5%	3.3%	1.2%	8.4%	14.5%	11.4%	2.4%	5.7%	1.8%	6.3%	4.5%	0.3%	5.4%	4.2%
Chro.50155	Uncharacterized	**Nucl**	1424	0.8%	1.7%	12.6%	6.9%	0.5%	3.9%	13.8%	2.2%	2.6%	7.6%	7.7%	11.7%	1.3%	3.0%	2.5%	9.7%	4.1%	0.1%	3.7%	3.6%
Chro.50195	Las1-like protein	**Cyto**	741	2.4%	3.0%	9.6%	4.3%	1.6%	3.0%	8.0%	2.3%	2.0%	12.1%	11.1%	9.9%	3.0%	4.6%	2.0%	8.4%	3.8%	1.6%	3.9%	3.5%
Chro.50420	Uncharacterized	**Cyto**	886	2.3%	2.1%	15.3%	5.4%	0.5%	3.7%	10.2%	2.6%	1.9%	8.2%	8.5%	8.4%	2.7%	3.3%	2.4%	8.1%	3.7%	0.7%	5.3%	4.7%
Chro.60245	Uncharacterized	**Plas**	2659	2.3%	3.7%	9.9%	3.8%	1.5%	2.4%	5.6%	2.3%	1.5%	13.1%	11.8%	8.8%	1.8%	6.8%	2.3%	10.0%	3.6%	1.2%	4.1%	3.3%
Chro.60295	Spm1 protein	**Nucl**	1133	3.9%	2.6%	10.3%	5.2%	0.4%	4.3%	6.2%	4.9%	1.2%	5.0%	7.1%	6.8%	0.9%	4.4%	6.6%	16.6%	7.6%	0.4%	1.7%	3.7%
Chro.60314	Uncharacterized	**Plas**	426	4.5%	4.3%	7.6%	3.6%	2.6%	4.8%	5.2%	5.2%	1.9%	8.3%	5.2%	7.6%	2.1%	3.1%	10.7%	8.3%	5.0%	0.0%	2.6%	7.1%
Chro.60470	Uncharacterized	**Cyto**	343	1.5%	2.9%	11.4%	5.2%	0.9%	2.9%	10.5%	3.8%	2.9%	9.3%	10.2%	7.0%	0.9%	3.5%	2.9%	12.2%	3.8%	0.3%	4.4%	3.5%
Chro.60490	Uncharacterized	**Extr**	462	1.8%	3.3%	9.9%	6.6%	2.6%	3.1%	5.9%	2.9%	2.9%	7.0%	9.9%	9.0%	0.9%	6.8%	4.2%	9.2%	5.1%	0.4%	3.5%	5.1%
Chro.60610	Uncharacterized	**Cyto**	426	4.7%	2.6%	5.6%	4.7%	1.2%	5.9%	7.0%	5.2%	1.9%	8.9%	7.0%	9.9%	2.1%	3.3%	4.9%	9.2%	6.8%	0.5%	2.6%	6.1%
Chro.60619	Uncharacterized	**Nucl**	350	0.6%	2.3%	11.9%	5.8%	1.4%	5.8%	10.1%	2.0%	2.3%	7.8%	7.0%	6.7%	1.2%	6.7%	3.8%	12.8%	5.8%	0.3%	3.2%	2.6%
Chro.70047	Golgi reassembly stacking protein	**Cyto**	745	4.0%	3.8%	8.5%	6.1%	1.2%	5.6%	10.8%	5.5%	2.9%	7.4%	7.5%	6.1%	1.2%	3.5%	5.3%	9.3%	4.6%	0.3%	2.3%	4.1%
Chro.70111	Uncharacterized	**Nucl**	749	2.1%	4.1%	9.3%	4.1%	1.1%	12.6%	3.9%	4.4%	3.5%	5.2%	7.1%	4.5%	2.5%	2.3%	5.2%	17.2%	3.1%	0.7%	3.7%	3.3%
Chro.70152	Uncharacterized	**Plas**	4235	4.7%	4.5%	6.2%	5.7%	0.7%	3.4%	6.2%	4.4%	1.8%	8.5%	10.9%	5.6%	2.0%	7.0%	4.2%	10.8%	4.7%	0.7%	3.7%	4.2%
Chro.70160	Proteinkinase	**Nucl**	816	3.8%	3.4%	7.6%	3.7%	1.0%	4.4%	9.4%	5.9%	1.2%	7.4%	8.0%	11.8%	1.1%	3.9%	2.9%	11.8%	4.9%	0.7%	2.9%	4.2%
Chro.70211	Beta-mannosyl transferase	**Nucl**	680	3.1%	4.1%	6.8%	5.4%	1.6%	3.4%	6.2%	4.4%	1.9%	8.8%	12.8%	8.7%	1.6%	4.9%	2.9%	10.4%	4.9%	0.7%	2.6%	4.7%
Chro.70267	Uncharacterized	**Plas**	2223	3.0%	4.6%	8.8%	5.2%	0.9%	3.1%	7.8%	3.1%	1.1%	9.4%	9.3%	9.0%	2.7%	5.8%	4.0%	8.1%	3.9%	0.9%	5.1%	4.1%
Chro.70296	Pyruvyl transferase1	**Plas**	1471	2.4%	3.0%	10.0%	5.6%	2.0%	4.0%	6.4%	4.4%	2.0%	10.1%	8.9%	7.2%	1.4%	6.1%	3.7%	12.4%	4.6%	0.5%	2.2%	3.1%
Chro.70395	LTV1 Protein	**Mito**	490	2.5%	4.1%	10.8%	9.1%	1.0%	3.3%	8.9%	3.7%	1.5%	6.6%	8.9%	10.0%	1.2%	4.4%	2.9%	9.8%	3.5%	0.4%	2.7%	4.6%
Chro.80024	Uncharacterized	**Nucl**	1061	4.2%	4.5%	9.0%	4.4%	2.0%	4.3%	6.2%	8.1%	1.7%	6.2%	5.4%	3.7%	0.8%	4.1%	5.4%	11.7%	7.4%	0.5%	5.1%	5.4%
Chro.80102	Uncharacterized	**Cyto**	267	2.6%	4.1%	8.3%	5.6%	0.0%	3.8%	13.5%	4.5%	0.4%	6.0%	7.9%	15.0%	1.1%	3.4%	3.0%	9.8%	3.0%	0.0%	3.8%	4.1%
Chro.80229	Uncharacterized	**Nucl/Cyto**	696	2.4%	3.0%	9.2%	7.2%	0.9%	4.5%	15.3%	2.3%	1.2%	9.2%	8.6%	9.6%	1.9%	4.0%	2.0%	8.9%	3.9%	0.1%	1.9%	3.9%
Chro.80245	Extensin-like protein	**Nucl/Cyto**	605	3.8%	1.8%	8.4%	2.6%	0.8%	7.9%	5.1%	4.6%	3.8%	7.4%	5.8%	5.1%	6.0%	2.6%	16.2%	8.8%	3.3%	0.0%	1.7%	4.1%
Chro.80353	Uncharacterized	**Extr**	235	3.0%	7.7%	6.8%	6.4%	0.0%	1.3%	9.4%	2.1%	0.0%	6.0%	5.6%	16.2%	2.6%	4.7%	4.7%	9.0%	5.1%	0.9%	3.8%	4.7%
Chro.80409	Uncharacterized	**Plas**	504	4.0%	4.4%	5.4%	3.4%	1.8%	4.4%	6.8%	3.4%	2.0%	8.7%	8.2%	7.0%	3.6%	7.8%	4.4%	12.5%	5.2%	0.4%	3.8%	3.2%
Chro.80605	Mature-parasite-infected erythrocyte surface antigen	**Plas**	1041	3.6%	5.6%	6.6%	6.7%	0.2%	3.7%	11.8%	4.8%	0.6%	6.7%	6.8%	11.4%	1.5%	3.8%	5.6%	10.4%	4.5%	0.0%	1.4%	4.2%

Appendix IV

RDP4 recombination detection results for IIaA15G2R1, IIcA5G3a, IIcA5G3j, and IbA10G2 whole genome comparisons

Summary of individual recombination events detected between *C. parvum* GP60 subtypes IIaA15G2R1 (IIaA), IIcA5G3j (IIcAj), and IIcA5G3a (IIcAa), and *C. hominis* GP60 subtype IbA10G2 (IbA) by RDP4 (Recombination Detection Program v. 4.66; Martin *et al*. 2015). Recombination events are characterised by start/end recombinant fragment breakpoints, major versus minor parental strains, E-values of recombination signals, and CDs/intergenic regions affected. Highly significant recombination events (E-value < E-40) are additionally highlighted. Recombinant sequences covering or within 20,000bp of putative virulence factors (telomeric, glycoprotein-like, low-complexity, and/or hypervariable between species) are additionally annotated with a red star (✸).

Breakpoint position in recombinant sequence (bp from 5' telomere)		Recombinant	Major parent	Minor Parent	Detection Method	CDs encoded within	
Beginning	***End***				***RDP***		
CHROMOSOME 1							
82251	104422	IIaA	IIcAj	Unknown	8.28E-240	cgd1_370 - cgd1_490	✸
82251	93181	IbA	IIcA/IIcAj	Unknown	4.46E-08	cgd1_370 - cgd1_430	✸
100170	100278	IIaA	IIcA/IIcAj	Unknown	3.07E-03	cgd1_470	✸
100631	100831	IbA	Unknown	IIaA	5.25E-04		
109846	110180	IbA	Unknown	IIcAa	8.34E-13	Intergenic cgd1_510 - cgd1_520	✸
111232	111726	IIcAa	IIaA/IIcAj	IbA	1.26E-13	cgd1_530	✸
115061	116161	IIcAj/IIaA	IIcAa	IbA	2.78E-07	cgd1_550	✸
127173	136648	IIaA	IIcAj	IIcAa	2.93E-72	cgd1_580 - cgd1_590	✸
136649	140781	IIcAj	IIaA	IIcAa	1.18E-15	cgd1_590 - cgd1_600	✸
142478	150610	IIaA	IIcAj	IIcAa	8.96E-16	cgd1_610 - cgd1_640	✸
376602	386949	IIcAa	Unknown	IIcAj/IIaA	2.92E-02	cgd1_1580 - cgd1_1640	
734690	744935	IIaA	IIcAj	IIcAa	4.18E-04	cgd1_3290 - cgd1_3340	✸
CHROMOSOME 2							
53785	55454	IIcAa	IIaA/IIcAj	IbA	1.38E-14	cgd2_160	
57056	57358	IIcAa	IIaA/IIcAj	IbA	1.54E-06	cgd2_140	
58483	58812	IIcAa	IIaA/IIcAj	Unknown	3.16E-08	cgd2_120	
61997	62206	IIcAa	IIaA/IIcAj	IbA	1.78E-09	Intergenic cgd2_110 - cgd2_100	
64582	65341	IIcAa	IIaA/IIcAj	IbA	8.90E-11	cgd2_90	
67242	67933	IIcAa	IIaA/IIcAj	IbA	2.23E-10		
71503	72990	IIcA/IIaA	IIcAj	IbA	2.82E-19	cgd2_80	
75512	76343	IIcAa	IIaA/IIcAj	IbA	5.18E-08	cgd2_70	
79931	80238	IIcAa	IIaA/IIcAj	IbA	4.95E-10	Intergenic cgd2_70 - cgd2_60	
294024	294928	IIaA	IIcAj	Unknown	4.32E-10	cgd2_1370	
341750	405045	IIcAa	Unknown	IIcAj	4.28E-06	cgd2_1690 - cgd2_2040	
432700	506795	IIcAj	IIaA15	Unknown	6.55E-04	cgd2_2170 - cgd2_2560	✸
625528	632428	IIaA	IIcAj	Unknown	1.93E-06	cgd2_3080 - cgd2_3110	✸
CHROMOSOME 3							
220866	220932	IIcAa	IIcAj/IIaA	Unknown	4.15E-08	cgd3_720	✸
272798	279815	IIcAj	IIaA	IIcAa	5.11E-04	cgd3_920 - cgd3_960	✸
319189	319570	IIcAa	IIcAj/IIaA	Unknown	9.41E-05	cgd3_1150	✸
321883	322660	IIcAj	IIaA	IIcAa	1.75E-15	cgd3_1160	✸
797968	799943	IIcAa	IIcAj/IIaA	IbA	1.43E-77	cgd3_3370	
995078	1030425	IIcAj	IIaA	IIcAa	2.18E-17	cgd3_4190 - cgd3_4280	✸
CHROMOSOME 4							
3370	5132	IIaA/IIcAj	IIcAa	IbA	1.70E-24	cgd4_20	
5137	5788	IIcAa	IIaA/IIcAj	IbA	5.46E-41		✸
848234	849840	IIcAj	IIaA	IIcAa	2.06E-13	cgd4_3630	✸
865724	865737	IIcAa	IIaA/IIcAj	IbA	1.89E-06	cgd4_3690	✸
1054213	1054636	IIcAa	IIcAj/IIaA	IbA	4.80E-14	cgd4_4480	✸
1057053	1058582	IIcAj/IIaA	IIcAa	IbA	1.27E-54	cgd4_4490	✸
1058583	1058932	IIcAa	IIaA/IIcAj	IbA	2.10E-13	Intergenic cgd4_4490 - cgd4_4500	✸
1059044	1059293	IIcAj/IIaA	IIcAa	IbA	5.69E-12		
1059418	1060146	IIcAa	IIaA/IIcAj	Unknown	3.67E-47	cgd4_4500	✸
1060336	1060469	IIaA/IIcAj	IIcAa	IbA	3.22E-07		

1060678	1060737	IIcAa	IIaA/IIcAj	IbA	1.45E-06	
1061059	1061153	IIcAa	IIaA/IIcAj	IbA	2.89E-04	
1061156	1061888	IIcAj/IIaA	IIcAa	IbA	1.65E-77	
1061941	1062512	IIcAa	IIcAj/IIaA	Unknown	5.83E-27	Intergenic cgd4_4500 - 3' telomere
1062847	1063606	IIcAj/IIaA	IIcAa	IbA	1.97E-61	

CHROMOSOME 5

3694	6176	IIaA	IIcAj	IIcAa	1.03E-23	Chro.50010
585260	586337	IIcAj	IIcA/IIaA	IbA	3.76E-28	cgd5_2180
649071	649362	IbA	Unknown	IIaA/IIcAj	8.09E-51	cgd5_1940
1031972	1033136	IIcAj/IIaA	IIcAa	IbA	2.15E-45	cgd5_40

CHROMOSOME 6

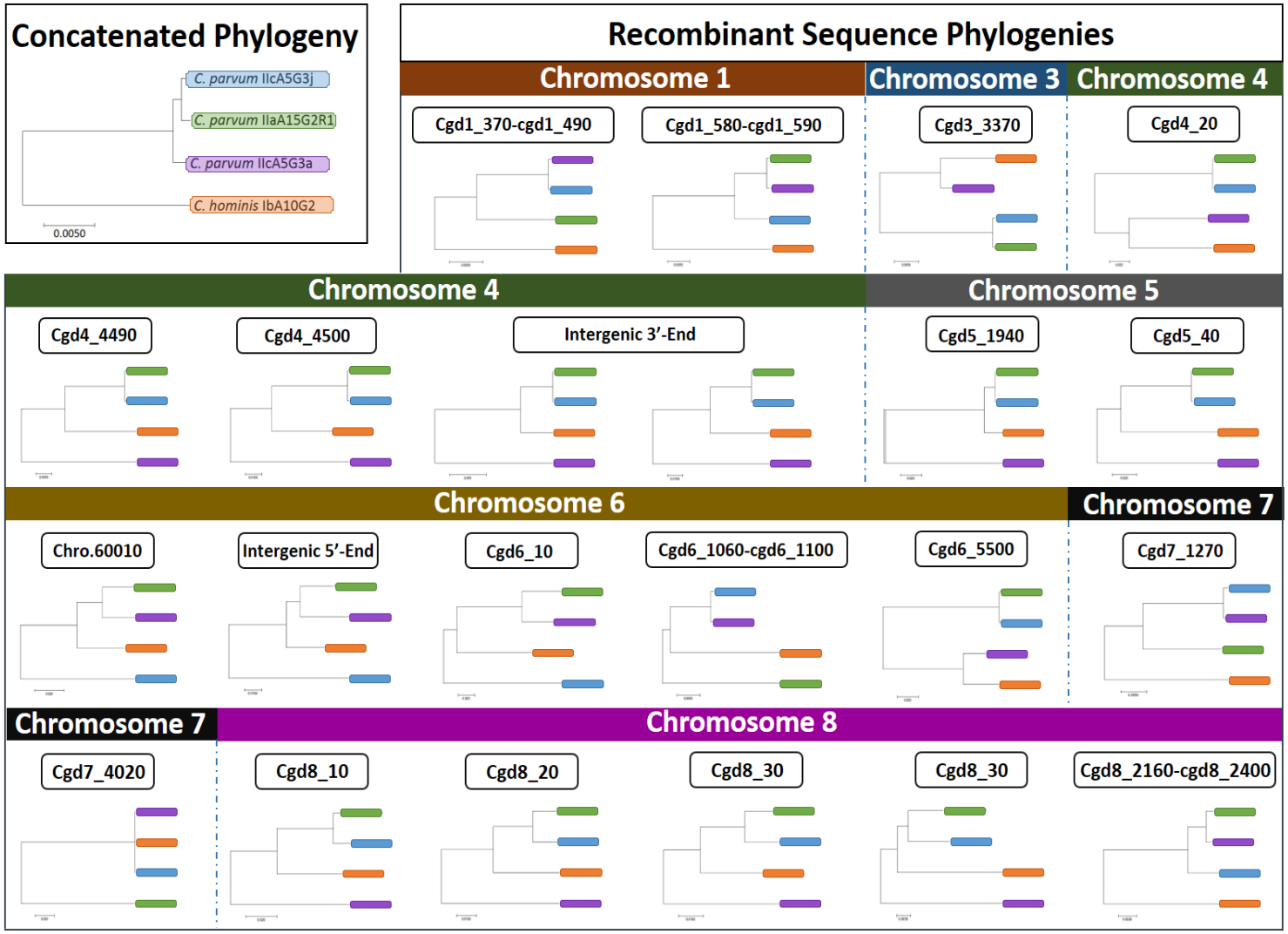
49	140	IIcAj	IIaA	Unknown	3.56E-05	Chro.60010
146	1792	IIcAj	IIaA	Unknown	2.59E-164	
1793	1905	IIcAj	IIaA	IbA	1.43E-12	
1986	2351	IIcAj	IIaA	Unknown	7.34E-43	
2352	2537	IIcAj	IIaA	IbA	8.61E-23	Intergenic Chro.60010 - cgd6_10
2538	2963	IIcAj	IIaA	Unknown	3.91E-09	
3510	3670	IIcAa	IIaA	IbA	3.10E-02	
4026	6334	IIaA	IIcAj	IIcAa	6.78E-144	cgd6_10
7166	7713	IIcAj/IIaA	IIcAa	IbA	5.31E-18	Intergenic cgd6_10 - cgd6_20
7784	7896	IIaA/IIcAj	IIcAa	IbA	2.21E-04	
8033	8972	IbA	Unknown	IIcAa	2.03E-14	cgd6_20
9758	9992	IIcAa	IIaA	IbA	1.12E-03	
10386	12685	IIcAj	IIaA	IIcAa	3.16E-32	cgd6_30 - cgd6_40
13148	13482	IIcAa	IIcAj/IIaA	Unknown	7.84E-06	Intergenic cgd6_40 - cgd6_50
14883	18178	IIaA	IIcAj	IIcAa	2.09E-24	cgd6_50
20061	20401	IIcAa	IIaA/IIcAj	Unknown	3.04E-19	cgd6_60
20936	21391	IIcAj/IIaA	IIcAa	IbA	7.46E-14	
186255	187077	IIaA/IIcAj	IIcAa	IbA	1.44E-06	cgd6_800
240190	240717	IIaA/IIcAj	IIcAa	IbA	3.02E-06	cgd6_1020
245902	247871	IIcAa	Unknown	IIaA	2.52E-31	cgd6_1060
247872	256568	IIaA	IIcAj	Unknown	2.04E-228	cgd6_1060 - cgd6_1100
1225101	1225478	IIcAa	IIaA/IIcAj	IbA	1.49E-04	cgd6_5260
1226191	1226342	IIcAa	IIcAj/IIaA	Unknown	1.28E-15	cgd6_5260 - cgd6_5270
1226343	1226614	IIcAa	IIaA/IIcAj	IbA	8.40E-13	cgd6_5270
1276817	1278061	IIcAa	IIaA	IbA	7.88E-28	cgd6_5450
1278062	1278345	IIcAa	IIaA	Unknown	5.11E-07	Intergenic cgd6_5450 - cgd6_5500
1278346	1280578	IIcAa	IIaA	IbA	4.13E-90	cgd6_5500

CHROMOSOME 7

285016	292270	IIcAj	IIaA	Unknown	3.04E-06	cgd7_1150 - cgd7_1170
317243	319588	IIcAj	IIaA	IIcAa	3.97E-46	cgd7_1270
878265	897621	IIcAj	IIaA	IIcAa	2.69E-08	cgd7_3910 - cgd7_4020
897622	898690	IIcAj	IIaA	IIcAa	7.11E-24	
897728	898242	IIaA	IIcAj	Unknown	1.07E-62	cgd7_4020
898691	899005	IbA	Unknown	IIcAj/IIaA	2. 16E-14	
899011	935740	IIcAj	IIaA	IIcAa	5.23E-25	cgd7_4020 - cgd7_4220
1055570	1063864	IIcAj	IIaA	IIcAa	1.45E-03	cgd7_4710 - cgd7_4750

CHROMOSOME 8

80	1150	IIcAa	IIaA/IIcAj	Unknown	9.69E-75	cgd8_10
1334	1408	IIcAa	IIaA/IIcAj	Unknown	2.72E-08	Intergenic cgd8_10 - cgd8_20
1409	1526	IIcAa	IIaA/IIcAj	IbA	6.92E-07	
3201	3369	IIcAa	IIaA/IIcAj	Unknown	1.18E-08	
3623	5676	IIcAj/IIaA	IIcAa	IbA	1.29E-110	cgd8_20
5677	5972	IIcAa	IIcAj/IIaA	IbA	7.91E-06	
6026	7033	IIcAj/IIaA	IIcAa	IbA	6.26E-40	
7274	9938	IIcA/IIcAj	IIaA	Unknown	2.95E-78	cgd8_30
10005	11970	IIcAj	IIaA	IIcAa	6.41E-07	cgd8_40
12805	14933	IIcAj	IIaA	IIcAa	9.67E-17	cgd8_40 - cgd8_50
15040	26389	IIaA	IIcAj	IIcAa	4.29E-19	cgd8_50 - cgd8_100
42714	48676	IIaA	IIcAj	IIcAa	6.45E-14	cgd8_170 - cgd8_180
75004	84938	IIcAj	IIaA	Unknown	4.04E-06	cgd8_300 - cgd8_350
547697	563658	IIaA	IIcAj	IIcAa	5.21E-33	cgd8_2090 - cgd8_2150
563659	564762	IIcAj	IIaA	IIcAa	2.16E-26	cgd8_2160
564902	618348	IIaA	IIcAj	IIcAa	1. 76E-115	cgd8_2160 - cgd8_2400
584382	584669	IbA	Unknown	IIcAj	5.59E-04	cgd8_2260
618349	628553	IIaA	IIcAj	Unknown	2.73E-08	cgd8_2400 - cgd8_2440
1085940	1086106	IIcAa	IIaA/IIcAj	Unknown	9.40E-32	cgd8_4480



Appendix V

Phylogenetic reconstruction of highly significant recombination events between *Cryptosporidium spp*. subtypes IIaA15G2R1, IIcA5G3a, IIcA5G3j, and IbA10G2

Highly significant recombination events (E-value < E-40) were identified from the genome-wide RDP4 recombination analysis (Appendix IV) of *C. parvum parvum* IIaA15G2R1 (UKP6), *C. parvum anthroponosum* IIcA5G3a (UKP13), *C. parvum parvum* IIcA5G3j (UKP16), and *C. hominis* IbA10G2 (UKH1) whole genome sequences. Sequence data within breakpoint positions were individually extracted from WGS and re-aligned using the ClustalW alignment algorithm imbedded in MEGA v 7.0. Phylogenetic trees were subsequently developed based on the Tamura-Nei model and the Maximum Likelihood method. Trees are drawn to scale, and branch lengths are represented in terms of number of substitutions per site. The concatenated phylogeny (section 1.1) for the included strains is additionally shown for comparative purposes. All analyses were conducted in MEGA version 7.0 (Kumar et al. 2015).

Appendix VI

Summary Table: Worldwide distribution, host origin and GP60 subtype family characterisation of *C. hominis* isolates

Publication	Source of Isolate	Host	Subtype Family	No. of Isolates	GenBank Submission	Submission year
Insulander M *et al*. 2013	Afghanistan	Human	Id	1	N/A	NS
	Albania	Human	Ib	1	N/A	NS
Laatamna AE *et al*. 2015	Algeria	Horse	Ik	1	KJ941148.1	2014
Peralta RH *et al*. 2016	Argentina	Human	Ia	2	KT381976.1 - KT381977.1	2015
	Argentina	Human	Ib	3	N/A	NS
	Argentina	Human	Ie	1	N/A	NS
Ebner J *et al*. 2015	Australia	Human	Id	1	KT123175.1	2015
Jex AR *et al*. 2007	Australia	Human	Ib	73	EU164805.1, EU164807.1	2007
	Australia	Human	Id	1	EU164806.1	2007
Jex AR *et al*. 2008	Australia	Human	Ib	34	N/A	NS
	Australia	Human	If	1	N/A	NS
	Australia	Human	Ig	3	EU379544.1	2008
Koehler AV *et al*. 2013	Australia	Human	Ib	53	KC632533.1	2013
	Australia	Human	Id	12	KC632532.1	2013
	Australia	Human	If	1	KC632534.1	2013
	Australia	Human	Ig	5	KC632529.1 - KC632531.1	2013
Koehler AV *et al*. 2016	Australia	Wallaby	Ib	1	KU531699	2015
Ng J *et al*. 2008	Australia	Human	Id	1	N/A	NS
Ng J *et al*. 2008	Australia	Human	Ib	2	N/A	NS
Ng J *et al*. 2010	Australia	Human	Ib	39	GU933438.1	2010
	Australia	Human	Id	88	GU933439.1 - GU933443.1	2010
	Australia	Human	Ie	4	GU933444.1 - GU933446.1	2010
	Australia	Human	If	3	GU933447.1 - GU933448.1	2010
	Australia	Human	Ig	60	GU933449.1 - GU933452.1	2010
Nolan MJ *et al* 2013	Australia	Deer	Ib	1	KC282993.1	2012
O'Brien E *et al*. 2008	Australia	Human	Ia	1	N/A	NS
	Australia	Human	Ib	24	EF025578.1 - EF025579.1	2006
	Australia	Human	Id	13	N/A	NS
	Australia	Human	If	3	EF025580.1	NS
Pangasa A *et al*. 2010	Australia	Human	Ig	1	GU214344.1	2009
	Australia	Human	If	1	GU214355.1	2009
Schiller SE *et al*. 2016	Australia	Flying Fox	Ib	1	N/A	NS
Sulaiman IM *et al*. 2001	Australia	Human	Id	1	AF402291.1	2001
Unpublished (Monis *et al*)	Australia	Kangaroo	Id	1	KX375347.1	2016
	Australia	Kangaroo	Ib	1	KX375348.1	2016
	Australia	Cattle	Ib	1	KX375349.1	2016
Waldron LS *et al*. 2009	Australia	Human	Ia	2	FJ839878.1, FJ839883.1	2009
	Australia	Human	Ib	28	FJ839881.1, FJ839882.1; FJ861209.1 - FJ861217.1; FJ861220.1 - FJ861224.1; FJ861227.1 - FJ861238.1	2009
	Australia	Human	Id	5	FJ839875.1, FJ861218.1 - FJ861219.1, FJ861225.1 - FJ861226.1	2009
	Australia	Human	Ie	1	FJ839874.1	2009
	Australia	Human	If	1	FJ839873.1	2009
Waldron LS *et al*. 2011	Australia	Cattle	Ib	3	JF727780.1 - JF727782.1	2011

	Australia	Human	Ia	4	JF727784.1	2011
	Australia	Human	Ib	156	JF727751.1 - JF727752.1, JF727785.1 - JF727788.1	2011
	Australia	Human	Id	1	JF727753.1	2011
	Australia	Human	Ie	1	JF727754.1	2011
	Australia	Human	If	4	JF727789.1 - JF727791.1	2011
	Australia	Human	U	1	JF727750.1	2011
	Australia	Human	U	1	JF727783.1	2011
Hira KG *et al*. 2011	Bangladesh	Human	Ia	8	AY700384.1, AY700390.1, AY700400.1; JF927177.1, JF927186.1, JF927190.1, JF927194.1, JF927196.1	2004 & 2011
	Bangladesh	Human	Ib	8	AY700387.1, JF927172.1, JF927178.1, JF927180.1, JF927181.1, JF927192.1, JF927195.1, JF927198.1	2004 & 2011
	Bangladesh	Human	Id	6	AY700388.1, AY700392.1, AY700398.1, JF927171.1, JF927179.1, JF927182.1	2004 & 2011
	Bangladesh	Human	Ie	10	AY700383.1, AY700389.1, AY700399.1, JF927169.1, JF927184.1, JF927189.1, JF927193.1, JF927197.1, JF927199.1, JF927200.1	2004 & 2011
	Bangladesh	Human	If	11	AY700391.1, AY700393.1, AY700397.1, JF927170.1, JF927173.1 - JF927176.1, JF927183.1, JF927185.1, JF927188.1	2004 & 2011
	Bangladesh	Human	U	1	AY700386.1	2004
Insulander M *et al*. 2013	Bangladesh	Human	Ia	1	N/A	NS
Geurden T *et al*. 2009	Belgium	Human	Ib	13	N/A	NS
Insulander M *et al*. 2013	Brazil	Human	Ib	1	N/A	NS
Peralta RH *et al*. 2016	Brazil	Human	Ib	13	N/A	NS
Unpublished (Inacio SV *et al*)	Brazil	Horse	Ik	1	KT948748.1	2015
Moore CE *et al*. 2016	Cambodia	Human	Ia	4	N/A	NS
	Cambodia	Human	Id	1	N/A	NS
	Cambodia	Human	If	1	N/A	NS
Thivierge K *et al.* 2016	Canada	Human	Id	14	N/A	NS
Trotz-Williams LA *et al.* 2006	Canada	Human	Ie	1	DQ192509.1	2005
	Canada	Human	Ia	2	DQ192510.1 - DQ192511.1	2005
	Canada	Human	Id	1	DQ192512.1	2005
Jex AR & Gasser RB 2008	Caribbean Islands	Human	Ib	1	N/A	NS
Gu Y *et al.* 2016	China	Rhesus monkey	Ib	4	KU945279.1 - KU945282.1	2016
	China	Chimpanzee	Ib	1	KU945283.1	2016
	China	Gibbon	Id	1	KU945284.1	2016
Insulander M *et al*. 2013	China	Human	Ib	1	N/A	NS
Jian F *et al*. 2016	China	Horse	Ik	1	N/A	NS
	China	Donkey	Ik	60	N/A	NS
Jiang Y *et al*. 2014	China	Human	Id	1	KF851244.1	2013
	China	Human	Ib	1	KF851245.1	2013
Karim MR *et al.* 2014	China	Cynomolgus monkey	Ii	1	KF679724.1	2013
	China	Rhesus monkey	Ib	7	KF679725.1	2013
Liu X *et al.* 2015	China	Squirrel monkey	U	1	KP314263.1	2014
Peng MM *et al.* 2001	China	Human	Ie	1	AF402286.1	2001
	China	Human	Id	1	AF403169.1	2001
	China	Human	Ib	1	AF403175.1	2001
Unpublished (Wang L *et al*. 2013)	China	Human	Ia	2	KC734574.1, KC734577.1	2013

	China	Human	Ib	1	KC734579.1	2013
	China	Human	Id	2	KC734575.1 - KC734576.1	2013
	China	Human	Ig	1	KC734578.1	2013
Unpublished (Wang RJ *et al*.)	China	Human	Ib	3	FJ707312.1 - FJ707314.1	2009
	China	Human	Ia	1	FJ707315.1	2009
	China	Human	Id	1	FJ707316.1	2009
Unpublished (Zhang H *et al*.)	China	Human	Ie	1	JQ796095.1	2012
Unpublished (Zhang L & Jian F)	China	Horse	Ii	1	KU200958.1	2015
	China	Donkey	Ii	2	KU200962.1 - KU200963.1	2015
Unpublished (Zhu HL *et al*.)	China	Human	Id	2	EF591785.1 - EF591786.1	2007
Wang L *et al*. 2013	China	Human	Ib	2	JF691560.1	2011
	China	Human	Ie	1	JQ029724.1	2011
Ye J *et al*. 2012	China	Rhesus monkey	Ia	3	JX000568.1 - JX000569.1	2012
	China	Rhesus monkey	Id	1	N/A	NS
	China	Rhesus monkey	Ie	1	N/A	NS
	China	Rhesus monkey	If	1	JX000570.1	2012
Insulander M *et al*. 2013	Congo	Human	Ib	1	N/A	NS
Gherasim A *et al*. 2012	Cuba	Human	Id	1	JQ028868.1	2011
Pelayo L *et al*. 2008	Cuba	Human	Ia	1	N/A	NS
	Cuba	Human	Ib	3	N/A	NS
	Cuba	Human	Id	2	N/A	NS
Chalmers RM *et al*. 2008	Cyprus	Human	Ib	3	N/A	NS
Jex AR & Gasser RB 2008	Cyprus	Human	Ib	4	EU877240.1	2008
	Denmark	Human	Ib	1	N/A	2008
	Dominica	Human	Ib	1	N/A	2008
	Ecuador	Human	Ie	1	EU877248.1	2008
Insulander M *et al*. 2013	Egypt	Human	Ia	1	N/A	NS
Jex AR & Gasser RB 2008	Egypt	Human	Ia	1	EU877235.1	2008
Unpublished (Abd El-Kader NM *et al*.)	Egypt	Human	Ia	1	HQ389257.1	2010
Unpublished (Beser J *et al*.)	Egypt	Human	Ib	1	KU852722.1	2016
Unpublished (Ghallab MMI *et al*.)	Egypt	Human	Ib	6	KX397528.1, KX397533.1, KX397542.1, KX397549.1, KX397554.1, KX397567.1	2016
	Egypt	Human	Id	23	KX397527.1, KX397529.1, KX397531.1, KX397532.1, KX397534.1, KX397536.1, KX397537.1, KX397541.1, KX397543.1 - KX397546.1, KX397548.1, KX397551.1 - KX397553.1, KX397555.1, KX397558.1, KX397560.1, KX397562.1, KX397564.1, KX397565.1, KX397568.1	2016
	Egypt	Human	Ie	9	KX397530.1, KX397535.1, KX397538.1, KX397547.1, KX397550.1, KX397557.1, KX397559.1, KX397561.1, KX397569.1	2016
Blanco MA *et al*. 2014	Equatorial Guinea	Human	Ia	3	N/A	NS
	Equatorial Guinea	Human	Ib	1	N/A	NS
	Equatorial Guinea	Human	Id	3	N/A	NS
Insulander M *et al*. 2013	Eritrea	Human	Ia	1	N/A	NS
Adamu H *et al*. 2010	Ethiopia	Human	Ib	1	N/A	NS
Flecha MJ *et al*. 2015	Ethiopia	Human	Ib	1	KP026302.1	2014
Insulander M *et al*. 2013	Ethiopia	Human	Id	2	N/A	NS

	Ethiopia	Human	Ib	1	N/A	NS
Chalmers RM *et al*. 2008	France	Human	Ib	4	N/A	NS
Jex AR & Gasser RB 2008	France	Human	Ib	2	N/A	NS
Insulander M *et al*. 2013	France	Human	Ib	1	N/A	NS
	Gambia	Human	If	2	N/A	NS
Eibach D *et al*. 2015	Ghana	Human	Ia	25	N/A	NS
	Ghana	Human	Ib	17	N/A	NS
	Ghana	Human	Id	1	N/A	NS
	Ghana	Human	Ie	8	N/A	NS
Chalmers RM *et al*. 2008	Greece	Human	Ib	2	N/A	NS
Jex AR & Gasser RB 2008	Greece	Human	Ib	5	N/A	NS
Strong WB *et al*. 2000	Guatemala	Human	Id	1	AF178695.1	1999
	Guatemala	Human	Ie	1	AF178697.1	1999
Sulaiman IM *et al*. 2001	Guatemala	Human	Ie	1	AF402290.1	2001
Ajjampur SS *et al*. 2010	India	Human	Ia	16	GQ384438.1, GQ384443.1	2009
	India	Human	Ib	11	N/A	NS
	India	Human	Id	7	N/A	NS
	India	Human	Ie	16	GQ384440.1, GQ384444.1	2009
	India	Human	If	5	GQ384439.1, GQ384441.1, GQ384442.1	2009
Feng Y *et al*. 2007	India	Cattle	Id	1	DQ871347.1	2006
Gatei W *et al*. 2007	India	Human	Ia	6	DQ665690.1, DQ665691.1, DQ665693.1, DQ665694.1	2006
	India	Human	Ib	13	DQ665688.1	2006
	India	Human	Id	12	DQ665692.1, DQ665695.1	2006
	India	Human	Ie	4	DQ665689.1	2006
Insulander M *et al*. 2013	India	Human	Ib	1	N/A	NS
	India	Human	Id	1	N/A	NS
Peng MM *et al*. 2001	India	Human	Ia	3	AF403171.1, AF403172.1, AF403785.1	2001
	India	Human	Ib	1	AF403173.1	2001
	India	Human	Id	1	AF403165.2	2001
Unpublished (Beser J *et al*.)	India	Human	Ia	1	KU852723.1	2016
Unpublished (Chamundeeswari S *et al*.)	India	Human	Ia	2	JQ517310.1, JQ517311.1	2012
Unpublished (Khalil S *et al*.)	India	Human	Ia	3	KU169226.1, KU169227.1, KU169235.1	2015
	India	Human	Ib	3	KU169228.1 - KU169230.1	2015
	India	Human	Id	2	KU169231.1, KU169234.1	2015
	India	Human	Ie	2	KU169232.1, KU169233.1	2015
Unpublished (Sharma P *et al*.)	India	Human	Ia	12	HQ241930.1, JF268622.1 - JF268627.1, JF495136.1 - JF495138.1, JF495154.1, JF495155.1	2010 & 2011
	India	Human	Ib	1	JF268628.1	2010 & 2011
	India	Human	Id	8	HQ241929.1, JF268629.1 - JF268633.1, JF495139.1, JF495156.1	2010 & 2011
	India	Human	Ie	13	HQ241927.1, JF268634.1 - JF268640.1, JF495140.1 - JF495143.1, JF495157.1	2010 & 2011
	India	Human	If	5	HQ241932.1, JF268641.1 - JF268643.1, JF495158.1	2010 & 2011
Unpublished (Sharma P. & Khurana S)	India	Human	Ia	3	KC813480.1, KC813486.1, KC813487.1	2013
	India	Human	Id	2	KC813481.1, KC813482.1	2013
	India	Human	Ie	3	KC813483.1, KC813488.1, KC813490.1	2013
Nazemalhosseini-Mojarad E *et al*. 2011	Iran	Human	Id	1	AB560749.1	2010
	Iran	Human	If	2	AB560750.1	2010

Zintl A *et al*. 2009	Ireland	Human	Ib	25	N/A	NS
Del Chierico F *et al*. 2011	Italy	Human	Ia	1	HM371371.1	2010
Wu Z *et al*. 2003	Italy	Human	Ib	2	AY167596.1	2002
Gatei W *et al*. 2008	Jamaica	Human	Ie	1	EU141719.1	2007
	Jamaica	Human	Ib	1	EU141720.1	2007
Jex AR & Gasser RB 2008	Jamaica	Human	Ib	1	N/A	NS
Abe N *et al*. 2006	Japan	Human	Ia	1	AB237130.1	2005
	Japan	Human	Ib	1	AB237135.1	2005
	Japan	Human	Ie	1	AB237138.1	2005
Wu Z *et al*. 2003	Japan	Human	Ie	1	AY167593.1	2002
	Japan	Human	Ia	1	AY167594.1	2002
Hijjawi N *et al*. 2010	Jordan	Human	Ib	8	N/A	NS
	Jordan	Human	Id	7	N/A	NS
Hijjawi N *et al*. 2016	Jordan	Human	Ib	2	N/A	NS
Chalmers RM *et al*. 2008	Kenya	Human	Ia	1	EU161651.1	2007
Jex AR & Gasser RB 2008	Kenya	Human	Ia	1	EU877238.1	2008
	Kenya	Human	Id	1	N/A	NS
Li W *et al*. 2011	Kenya	Baboon	Ib	2	JF681173.1	2011
	Kenya	Baboon	If	2	JF681172.1	2011
	Kenya	Baboon	Ij	1	JF681174.1	2011
Mbae C *et al*. 2015	Kenya	Human	Ia	5	N/A	NS
	Kenya	Human	Ib	20	KF974523.1, KF974525.1, KF974527.1	2013
	Kenya	Human	Id	32	KF984192.1	2013
	Kenya	Human	Ie	23	KF974521.1, KF974522.1	2013
	Kenya	Human	If	3	N/A	NS
Peng MM *et al*. 2001	Kenya	Human	Ie	2	AF402287.1, AF402288.1	2001
	Kenya	Human	Ia	1	AF403170.1	2001
Sulaiman IM *et al*. 2001	Kenya	Human	Ia	1	AF402293.1	2001
Unpublished (Gatei W)	Kenya	Human	Ia	5	EU146126.1	2007
	Kenya	Human	Id	7	EU146132.1	2007
	Kenya	Human	Ie	1	EU146131.1	2007
Sulaiman IM *et al*. 2005	Kuwait	Human	Ib	2	AY738187.1, AY738196.1	2004
	Kuwait	Human	Id	1	AY738192.1	2004
	Kuwait	Human	Ie	1	AY738184.1	2004
Osman M *et al*. 2015	Lebanon	Human	Id	8	KM215759.1 - KM215766.1	2014
Osman M *et al*. 2016	Lebanon	Human	Ia	4	KU311887.1	2015
	Lebanon	Human	Ib	16	KU311884.1, KU311885.1, KU311888.1, KU311890.1, KU311892.1, KU311900.1, KU311908.1, KU311909.1	2015
Areeshi M *et al*. 2008	Madagascar	Human	Ia	3	N/A	NS
	Madagascar	Human	Id	4	N/A	NS
	Madagascar	Human	Ie	1	N/A	NS
Ghaffari S & Kalantari N	Malawi	Human	Id	1	N/A	NS
Peng MM *et al*. 2003	Malawi	Human	Ia	4	AY382671.1	2003
	Malawi	Human	Ib	5	AY382672.1	2003
	Malawi	Human	Id	11	AY382669.1, AY382673.1	2003
	Malawi	Human	Ie	6	AY382668.1	2003
Unpublished (Ghaffari S *et al*.)	Malawi	Human	Id	1	KF577772.1	2013
Lim YA *et al*. 2011	Malaysia	Human	Ia	2	HQ631406.1, HQ631408.1	2010
	Malaysia	Human	Ib	2	HQ631410.1, HQ836661.1	2010

	Malaysia	Human	Id	1	HQ836664.1	2010
	Malaysia	Human	Ie	2	HQ631409.1, HQ836663.1	2010
	Malaysia	Human	If	1	HQ836662.1	2010
	Malaysia	Human	U	1	HQ631407.1	2010
Insulander M *et al*. 2013	Mexico	Human	Ib	2	N/A	NS
Jex AR & Gasser RB 2008	Mexico	Human	Ia	1	EU877234.1	2008
	Mexico	Human	Ib	2	EU877241.1	2008
Valenzuela O *et al*. 2014	Mexico	Human	Ia	5	KJ460366.1 - KJ460369.1, KJ460373.1	2014
	Mexico	Human	Ib	1	KJ460363.1	2014
	Mexico	Human	Id	1	KJ460362.1	2014
	Mexico	Human	Ie	3	KJ460364.1, KJ460365.1, KJ460371.1	2014
Wu Z *et al*. 2003	Nepal	Human	Ia		AY167595.1	2002
	Nepal	Human	Ia	1	N/A	NS
Abeywardena H *et al*. 2012	New Zealand	Cattle	Ib	12	JQ837920.1, JQ837922.1, JQ837926.1 - JQ837929.1	2012
Chalmers RM *et al*. 2008	New Zealand	Human	Ig	1	EF214734.1	2006
Akinbo FO *et al*. 2013	Nigeria	Human	Id	2	JX524490.1	2012
Ayinmode AB *et al*. 2014	Nigeria	Human	Ia	3	JX644914.1, JX644916.1, JX644917.1	2012
	Nigeria	Human	Ib	1	JX644913.1	2012
	Nigeria	Human	Ie	1	JX644915.1	2012
Jex AR & Gasser RB 2008	Nigeria	Human	Ib	2	EU877242.1	2008
Maikai BV *et al*. 2012	Nigeria	Human	Ia	2	JQ798143.1, JQ798144.1	2012
Molloy SF *et al*. 2010	Nigeria	Human	Ia	10	N/A	NS
	Nigeria	Human	Ib	10	N/A	NS
	Nigeria	Human	Id	4	N/A	NS
	Nigeria	Human	Ie	3	N/A	NS
	Nigeria	Human	Ih	1	FJ971716.1	2009
Chalmers RM *et al*. 2008	Pakistan	Human	Ia	3	EU161648.1, EU161649.1, EU161652.1	2007
Jex AR & Gasser RB 2008	Pakistan	Human	Ia	1	EU877237.1	2008
	Pakistan	Human	Ib	1	N/A	NS
	Pakistan	Human	Id	3	EU877243.1, EU877244.1, EU877247.1	2008
	Pakistan	Human	Ie	4	N/A	NS
	Pakistan	Human	If	1	EU877249.1	2008
Koinari M *et al*. 2013	Papua New Guinea	Fish	Id	1	N/A	NS
Koinari M *et al*. 2014	Papua New Guinea	Goat	Id	1	N/A	NS
Cama VA *et al*. 2007	Peru	Human	Ia	35	N/A	NS
	Peru	Human	Ib	39	EF035554.1	2006
	Peru	Human	Id	40	N/A	NS
	Peru	Human	Ie	13	N/A	NS
Cama VA *et al*. 2008	Peru	Human	Ia	20	EU095258.1 - EU095263.1	2007
	Peru	Human	Ib	23	N/A	NS
	Peru	Human	Id	12	EU095264.1, EU095265.1	2007
	Peru	Human	Ie	19	N/A	NS
Li N *et al*. 2013	Peru	Human	Ia	2	JX088407.1, JX088408.1	2012
	Peru	Human	Ib	1	JX088406.1	2012
	Peru	Human	Id	2	JX088403.1, JX088404.1	2012
	Peru	Human	Ie	1	JX088405.1	2012
Sturbaum GD *et al*. 2003	Peru	Human	Ia	1	AF528755.1	2002
	Peru	Human	Ib	1	AF528761.1	2002

	Peru	Human	Id	2	AF528756.1, AF528757.1	2002
	Peru	Human	Ie	1	AF528759.1	2002
Sturbaum GD *et al*. 2008	Peru	Human	Ib	1	EU052233.1	2007
	Peru	Human	Ia	1	EU052234.1	2007
Insulander M *et al*. 2013	Philippines	Human	Ia	1	N/A	NS
Alves M *et al*. 2003	Portugal	Human	Ib	1	AY166807.1	2002
	Portugal	Human	Ie	1	AY166808.1	2002
	Portugal	Human	If	1	AY166810.1	2002
Alves M *et al*. 2006	Portugal	Human	Ia	1	DQ280499.1	2005
	Portugal	Human	Ib	10	N/A	NS
	Portugal	Human	Id	1	DQ280498.1	2005
	Portugal	Human	Ie	2	N/A	NS
	Portugal	Human	If	1	N/A	NS
Insulander M *et al*. 2013	Portugal	Human	Ib	3	N/A	NS
Jex AR & Gasser RB 2008	Portugal	Human	Ib	1	N/A	NS
Peng MM *et al.* 2001	Portugal	Human	Ib	1	AF403174.1	2001
Neagoe IM *et al.* 2014	Romania	Human	Ib	1	HG423391.1	2010
Unpublished (Neagoe IM *et al*)	Romania	Human	Ia	1	LT556067.1	2016
	Romania	Human	Ib	2	LT556061.1, LT556068.1	2016
	Romania	Human	Id	1	LT556069.1	2016
Lobo ML *et al*. 2014	São Tomé	Human	Ia	7	N/A	NS
	São Tomé	Human	Ie	7	N/A	NS
Jex AR & Gasser RB 2008	Saudi Arabia	Human	Ib	1	EU877240l	2008
Unpublished (Areeshi MY *et al*.)	Saudi Arabia	Human	Ie	1	AB369996.1	2007
Unpublished (Al-Braikan FAA)	Saudi Arabia	Human	Ib	1	AJ973150.1	2005
Soba B & Logar J 2008	Slovenia	Human	Ia	1	AM937018.1	2008
	Slovenia	Human	Ib	1	AM988862.1	2008
Unpublished (Beser J *et al*.)	Somalia	Human	Ia	1	KU852724.1	2016
Insulander M *et al*. 2013	Somalia	Human	Ib	1	N/A	NS
Abu Samra N *et al* 2013	South Africa	Human	Ia	3	N/A	NS
	South Africa	Human	Ib	2	N/A	NS
	South Africa	Human	Id	5	N/A	NS
	South Africa	Human	Ie	4	N/A	NS
	South Africa	Human	If	5	N/A	NS
Abu Samra N *et al*. 2016	South Africa	Human	Ib	2	KU168244.1, KU168245.1	2015
	South Africa	Human	Ie	1	KU168246.1	2015
Insulander M *et al*. 2013	South Africa	Human	If	1	JN867334.1	2011
Leav BA *et al*. 2002	South Africa	Human	Ia	1	AF440634.1	2001
	South Africa	Human	Ib	4	AF440626.1, AF440628.1, AF440637.1, AF440640.1	2001
	South Africa	Human	Id	5	AF440623.1 - AF440625.1, AF440632.1, AF440635.1	2001
	South Africa	Human	If	5	AF440629.1, AF440630.1, AF440633.1, AF440638.1, AF440639.1	2001
Insulander M *et al*. 2013	South Africa	Human	Id	2	N/A	NS
Abal-Fabeiro JL *et al* 2013	Spain	Human	Ib	9	JQ349234.1 - JQ349242.1	2012
	Spain	Human	Id	1	JQ349243.1	2012
Abal-Fabeiro JL *et al* 2014	Spain	Human	Ia	3	N/A	NS
	Spain	Human	Ib	336	N/A	NS
	Spain	Human	Id	27	N/A	NS
	Spain	Human	Ie	4	N/A	NS

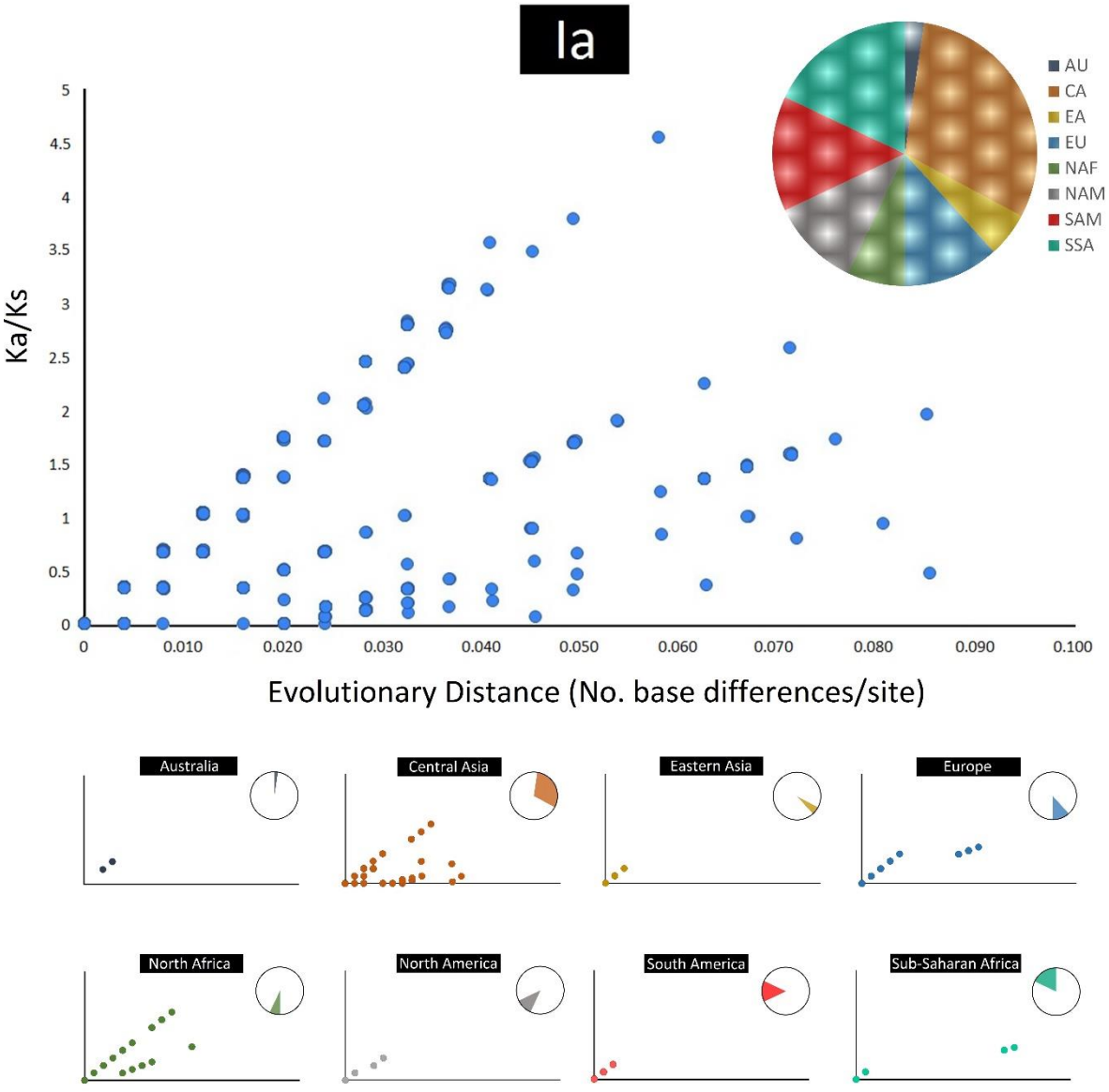
Chalmers RM *et al*. 2008	Spain	Human	Ib	14	N/A	NS
de Lucio A *et al*. 2016	Spain	Human	Ib	58	KT764962.1	2015
	Spain	Human	Ie	1	KT764963.1	2015
Gherasim A *et al*. 2012	Spain	Human	Ib	1	N/A	NS
Insulander M *et al*. 2013	Spain	Human	Ib	3	N/A	NS
	Spain	Human	Id	3	N/A	NS
Jex AR & Gasser RB 2008	Spain	Human	Ia	1	N/A	NS
	Spain	Human	Ib	41	N/A	NS
Ramo A *et al* 2015	Spain	Human	Ia	6	KM222527.1	2014
	Spain	Human	Ib	35	KM222525.1, KM222526.1	2014
Insulander M *et al*. 2013	Sweden	Human	Ia	1	N/A	NS
	Sweden	Human	Ib	12	N/A	NS
Unpublished (Beser J *et al*.)	Sweden	Human	Ik	1	KU727290.1	2016
	Sweden	Human	U	1	KU852719.1	2016
O'Brien E *et al*. 2008	Switzerland	Human	Ib	2	N/A	NS
	Switzerland	Human	Id	1	N/A	NS
Cheng YJ *et al*. 2014	Taiwan	Human	If	1	N/A	NS
Unpublished (Beser J *et al*.)	Tanzania	Human	Ia	1	KU727289.1	2016
	Tanzania	Human	Id	1	KU852721.1	2016
Koehler AV *et al*. 2014	Tasmania	Human	Ib	30	KJ506838.1, KJ506839.1	2014
	Tasmania	Human	Id	29	KJ506844.1, KJ506846.1	2014
	Tasmania	Human	Ie	1	KJ506842.1	2014
	Tasmania	Human	If	1	KJ506843.1	2014
Insulander M *et al*. 2013	Thailand	Human	Ib	9	N/A	NS
	Thailand	Human	Id	1	N/A	NS
Unpublished (Beser J *et al*. 2015)	Thailand	Human	Ii	2	KR296810.1, KR296811.1	2015
Jex AR & Gasser RB 2008	The Netherlands	Human	Ib	1	N/A	NS
Roelfsema JH *et al*. 2016	The Netherlands	Human	Ia	23	N/A	NS
	The Netherlands	Human	Ib	218	N/A	NS
	The Netherlands	Human	Ie	3	N/A	NS
	The Netherlands	Human	If	1	N/A	NS
Wielinga PR *et al*. 2008	The Netherlands	Human	Ib	63	EF576982.1 - EF577044.1	2007
	The Netherlands	Human	Id	2	EF576980.1, EF576981.1	2007
Chalmers RM *et al*. 2008	Tunisia	Human	Ib	2	N/A	NS
Unpublished (Essid R)	Tunisia	Human	Ia	9	KU057075.1 - KU057083.1	2015
Insulander M *et al*. 2013	Turkey	Human	Ib	2	N/A	NS
Chalmers RM *et al*. 2008	Turkey	Human	Ib	1	N/A	NS
Jex AR & Gasser RB 2008	Turkey	Human	Ib	1	N/A	NS
	Turks and Caicos	Human	Id	1	EU877245.1	2008
Akiyoshi DE *et al*. 2006	Uganda	Human	Ia	4	AY873783.1	2005
	Uganda	Human	Ib	1	N/A	NS
	Uganda	Human	Id	2	AY873782.1, AY873784.1	2005
	Uganda	Human	Ie	7	N/A	NS
	Uganda	Human	U	1	AY873781.1	2016
Insulander M *et al*. 2013	Uganda	Human	Ie	1	N/A	NS
Jex AR & Gasser RB 2008	Uganda	Human	Ie	1	N/A	NS
Widmer & Lee 2010	Uganda	Human	Ia	3	HM365230.1, HM365231.1, HM365235.1	2010
	Uganda	Human	Ie	3	HM365232.1 - HM365234.1	2010 - 2012
Xu P *et al*. 2004	Uganda	Human	Ia	1	XM_663000.1	2004

Bouzid M *et al*. 2010	United Kingdom	Human	Ib	4	GQ983356.1, GQ983358.1, GQ983360.1, GQ983362.1	2009
Chalmers RM *et al*. 2008	United Kingdom	Human	Ia	1	EU161650.1	2007
	United Kingdom	Human	Ib	66	EU161653.1, EU161654.1	2007
	United Kingdom	Human	If	1	EU161655.1	2007
	United Kingdom	Human	Ig	1	EF214735.1	2006
Chalmers RM *et al*. 2013	United Kingdom	Human	Ia	1	JX043493.1	2012
	United Kingdom	Human	Ib	2	JX294571.1, JX294572.1	2012
Connelly L *et al*. 2013	United Kingdom	Sheep	Ib	10	KC679684.1 - KC679693.1	2013
	United Kingdom	Sheep	U	1	KC679695.1	2013
Davies AP *et al*. 2009	United Kingdom	Human	Ib	1	FJ031237.1	2008
Ghaffari S & Kalantari N, 2014	United Kingdom	Human	Ib	13	N/A	NS
Ghaffari S *et al*. 2014	United Kingdom	Cattle	Ib	1	KF537685.1	2013
Giles M *et al*. 2009	United Kingdom	Sheep	Ib	1	EU186152.1	2007
Hadfield SJ *et al*. 2011	United Kingdom	Human	Ia	1	HQ149032.1	2010
	United Kingdom	Human	Ib	2	HQ149031.1, HQ149033.1	2010
	United Kingdom	Human	Id	1	HQ149034.1	2010
	United Kingdom	Human	Ie	1	HQ149035.1	2010
	United Kingdom	Human	If	1	HQ149036.1	2010
Hadfield SJ *et al*. 2015	United Kingdom	Human	Ia	1	KM012047.1	2014
	United Kingdom	Human	Ib	2	KM012048.1, KM085027.1	2014
Jex AR & Gasser RB 2008	United Kingdom	Human	Ia	1	EU877236.1	2008
	United Kingdom	Human	Ib	2	EU877239.1	2008
	United Kingdom	Human	Id	1	EU877246.1	2008
O'Brien E *et al*. 2008	United Kingdom	Human	Ib	3	N/A	NS
Pangasa A *et al*. 2010	United Kingdom	Human	Ia	3	GU214343.1, GU214345.1, GU214346.1	2009
	United Kingdom	Human	Ib	3	GU214347.1 - GU214349.1	2009
	United Kingdom	Human	Id	4	GU214350.1 - GU214353.1	2009
	United Kingdom	Human	Ie	1	GU214354.1	2009
Unpublished (Elwin K *et al*)	United Kingdom	Human	Ib	3	JQ413446.1, KF660596.1, KF660597.1	2012 & 2013
	United Kingdom	Human	Id	1	KF539855.1	2013
Unpublished (Ghaffari S *et al*)	United Kingdom	Human	Ib	1	KF577771	2013
Unpublished (Leoni F *et al*)	United Kingdom	Human	Ie	1	EF495194.1	2007
Widmer & Lee 2010	United Kingdom	Human	Ib	5	HM365224.1 - HM365228.1	2010
	United Kingdom	Human	Ie	1	HM365222.1	2010
	United Kingdom	Human	If	1	HM365223.1	2010
Feltus DC *et al*. 2006	United States	Human	Ib	1	N/A	NS
Feng Y *et al*. 2011	United States	Rhesus monkey	Ii	1	HM234173.1	2010
Feng Y *et al*. 2014	United States	Human	Ia	18	KF682372.1, KF682373.1	2013
	United States	Human	Id	8	KF682371.1	2013
	United States	Human	Ig	1	KF682374.1	2013
Jex AR & Gasser RB 2008	United States	Human	Ib	2	N/A	NS
Strong WB *et al*. 2000	United States	Human	Ia	5	AF164502.1 - AF164505.1, AF224462.1	1999
	United States	Human	Ib	6	AF164498.1 - AF164500.1, AF178690.1, AF224463.1, AF224464.1	1999
	United States	Human	Id	1	AF164497.1	1999
Sturbaum GD *et al*. 2003	United States	Human	Ib	1	AF528762.1	2002
Sulaiman IM *et al*. 2001	United States	Human	Ie	1	AF402289.1	2001
	United States	Human	Ia	2	AF403161.1, AF403162.1	2001

Unpublished (Beser J *et al*.)	United States	Human	Ia	1	KU852725.1	2016
Unpublished (Fan X *et al*)	United States	Human	Ib	1	JN571736.1	2011
Widmer & Lee 2010	United States	Human	Ib	1	HM370430.1	2010
	United States	Human	Ie	1	HM370431.1	2010
Insulander M *et al*. 2013	Venezuela	Human	Ib	5	N/A	NS
Alyousefi NA *et al*. 2013	Yemen	Human	Ie	1	JX032693.1	2012

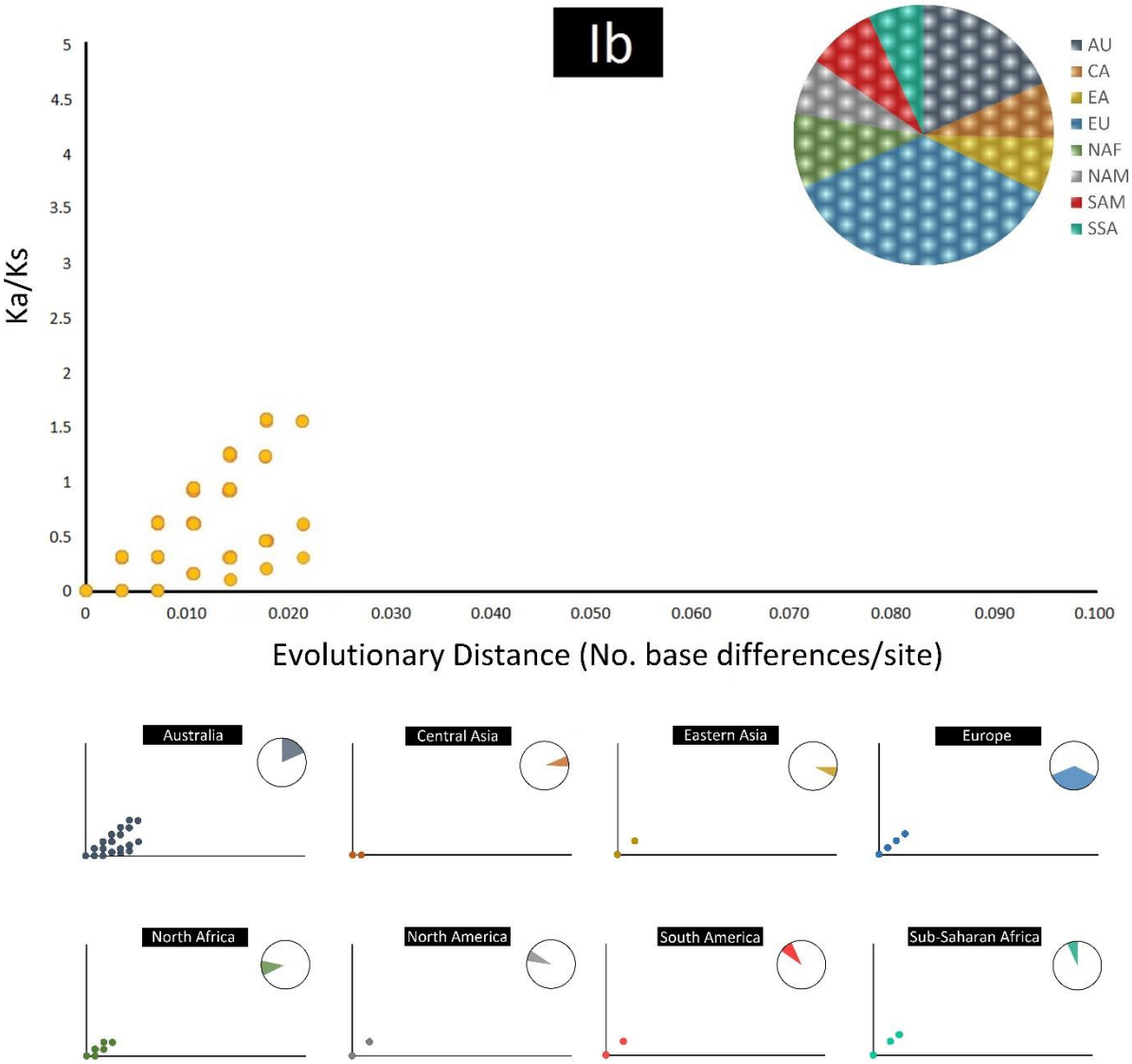
Appendix VII

Geographical disparities of *C. hominis* Ia-subtype GP60 allelic diversity



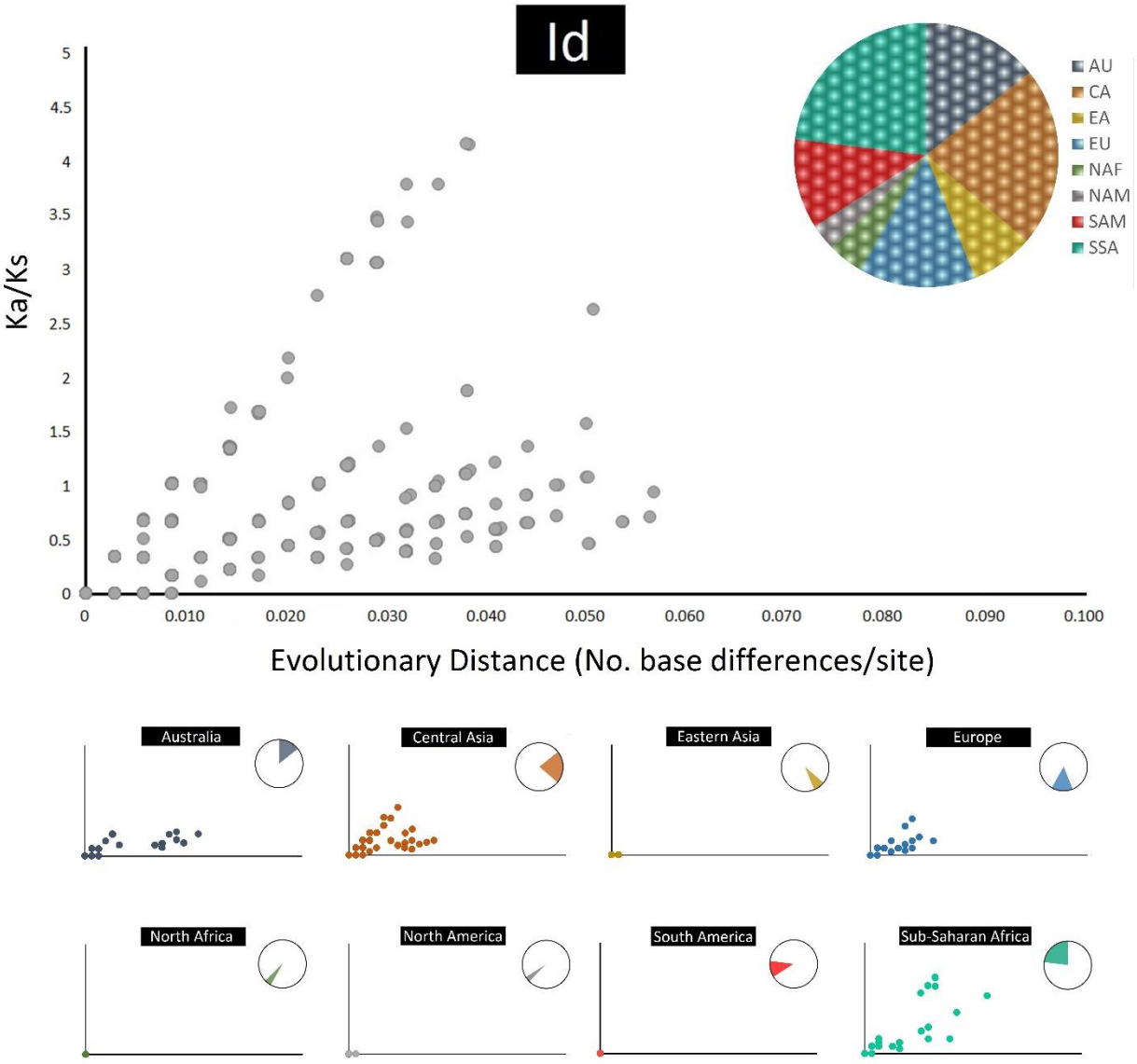
Appendix VIII

Geographical disparities of *C. hominis* Ib-subtype GP60 allelic diversity



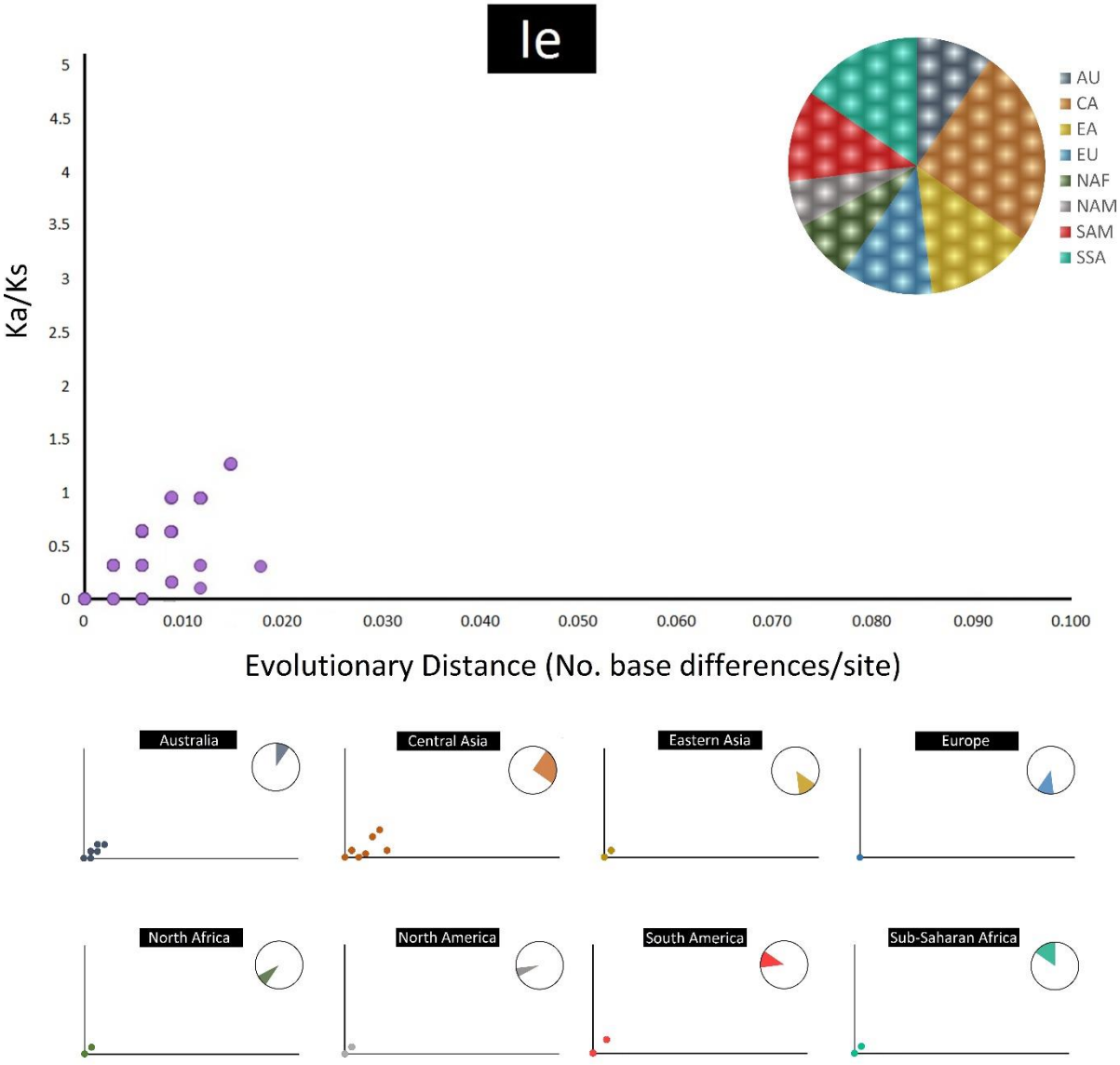
Appendix IX

Geographical disparities of *C. hominis* Id-subtype GP60 allelic diversity



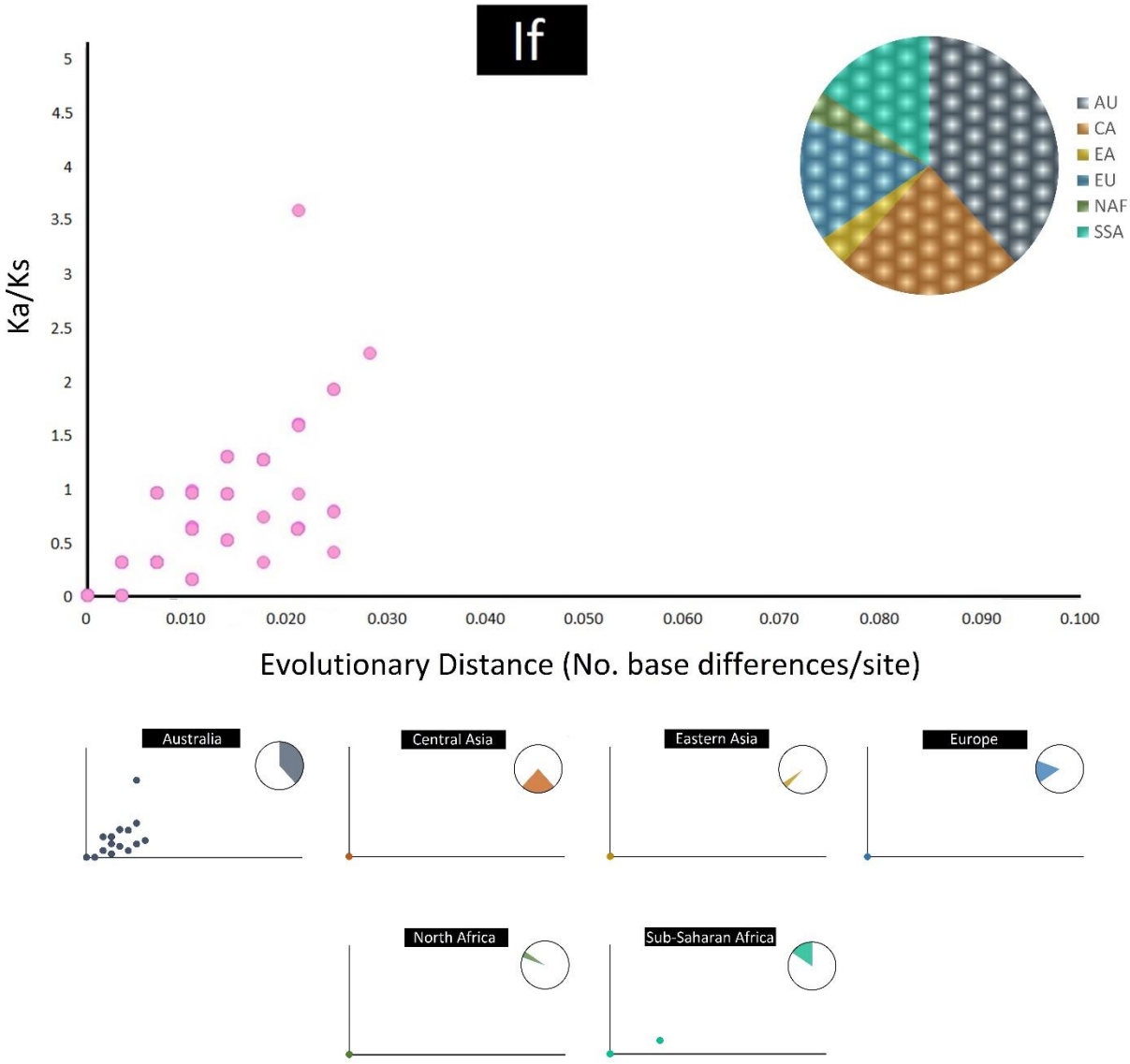
Appendix X

Geographical disparities of *C. hominis* Ie-subtype GP60 allelic diversity



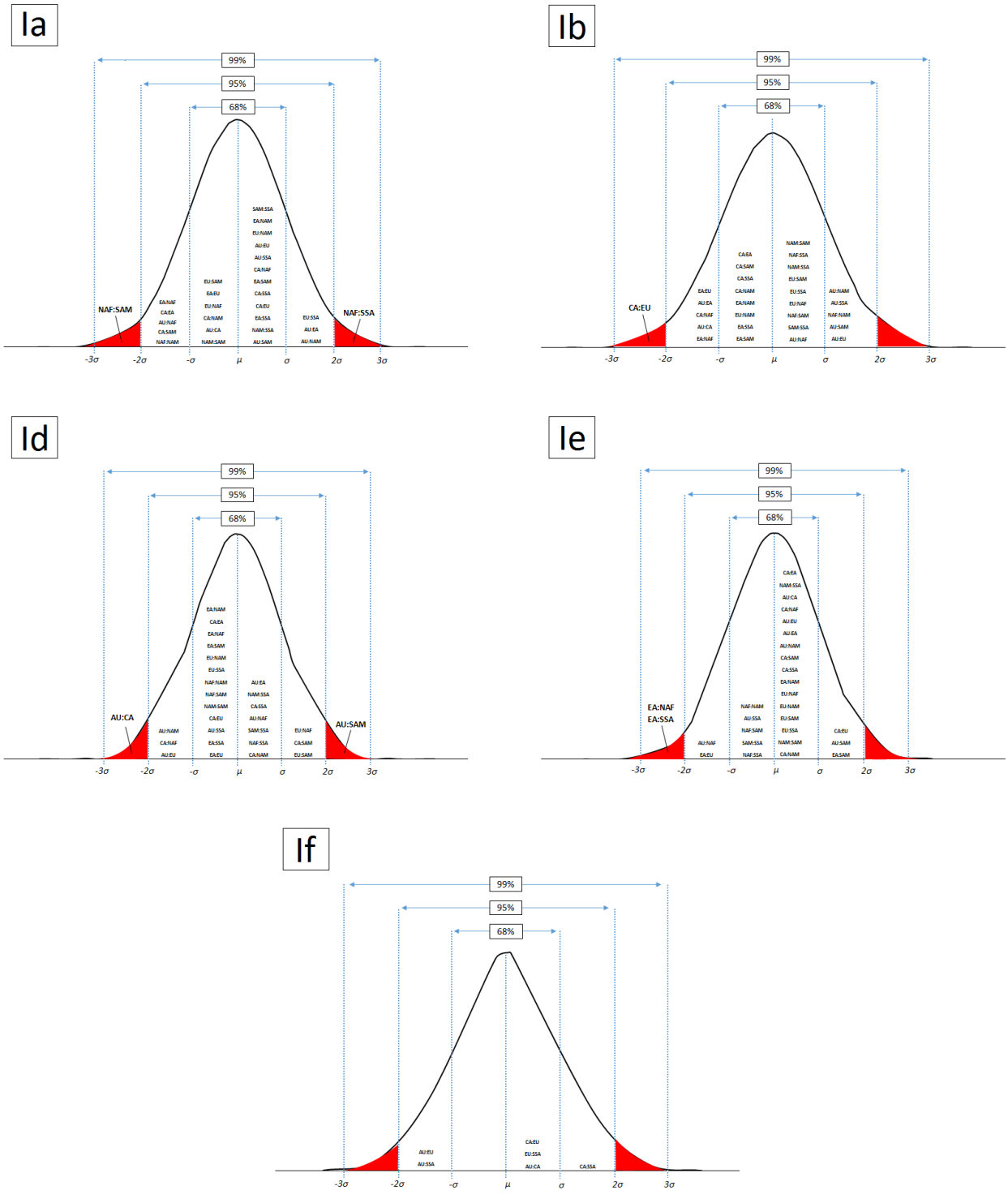
Appendix XI

Geographical disparities of *C. hominis* If-subtype GP60 allelic diversity



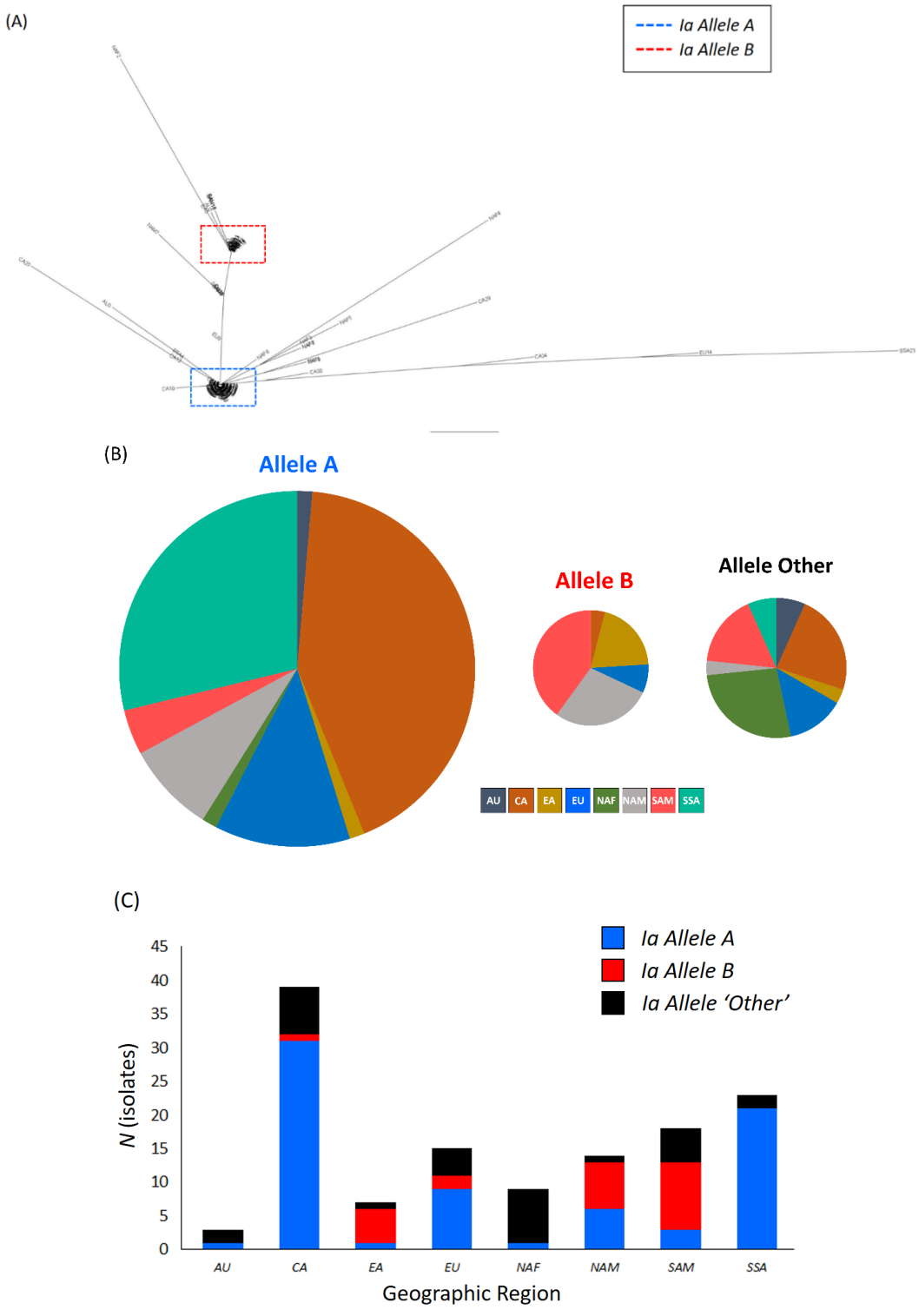
Appendix XII

Normal curves showing percentiles and standard deviations (±1-3 SD) for phylogeographic estimations (G_{st}-N_{st}) of geographically-defined *C. hominis* GP60 allele groups



Appendix XIII

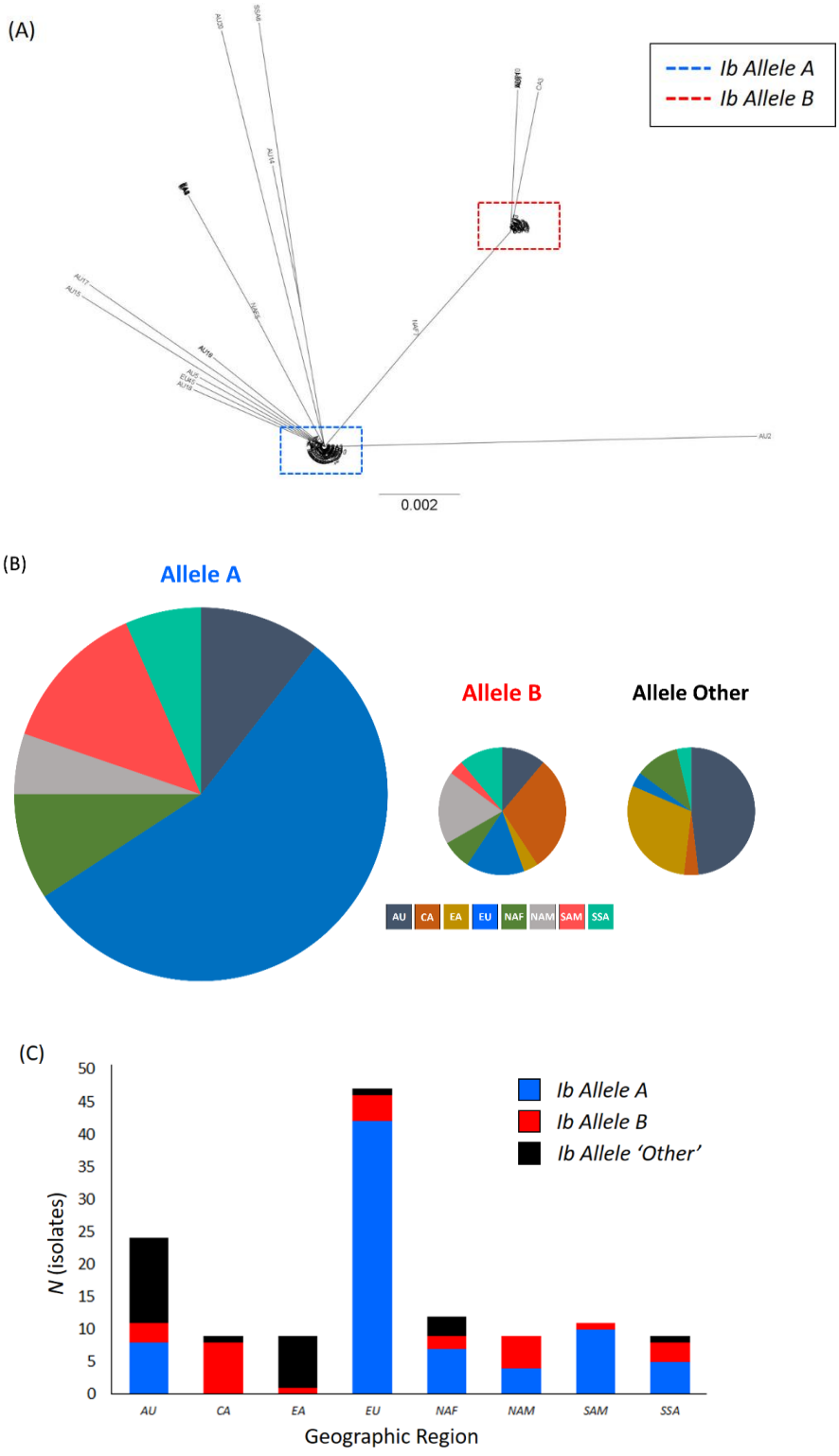
Allelic diversity of Ia-type *C. hominis* GP60 alleles



Appendix Figure XI. Phylogenetic clustering and geographic distribution of Ia-type GP60 alleles
25 unique alleles were identified from SNP-based comparisons of Ia GP60 consensus sequence alignments (DnaSP v 5.10.1; Librado & Rozas, 2009). Maximum likelihood phylogenetic reconstruction of ClustalW aligned gapless consensus sequences (A) revealed distinct clustering of isolates into predominantly two alleles, type A (57.0%) and type B (19.5%), with the remaining isolates comprising 'other' alleles at substantially lower total frequencies (0.8-3.1%) (Tamura-Nei Model, MEGA v 7.0). Geographic composition of these three subtype classes revealed inter-allelic differences, with allele A predominating in Central Asia and Sub-Saharan Africa, and allele B predominating in North and South America, and less conventional alleles dominating in North African samples (B). Relative sizes of individual pie figures are accurate proportional representations of total isolates analysed. Regional differences of allelic composition were additionally assessed, revealing coexistence of both alleles A and B in most defined geographic regions (62.5%), and differential proportions for intra-regional compositions of all allele types (C).

Appendix XIV

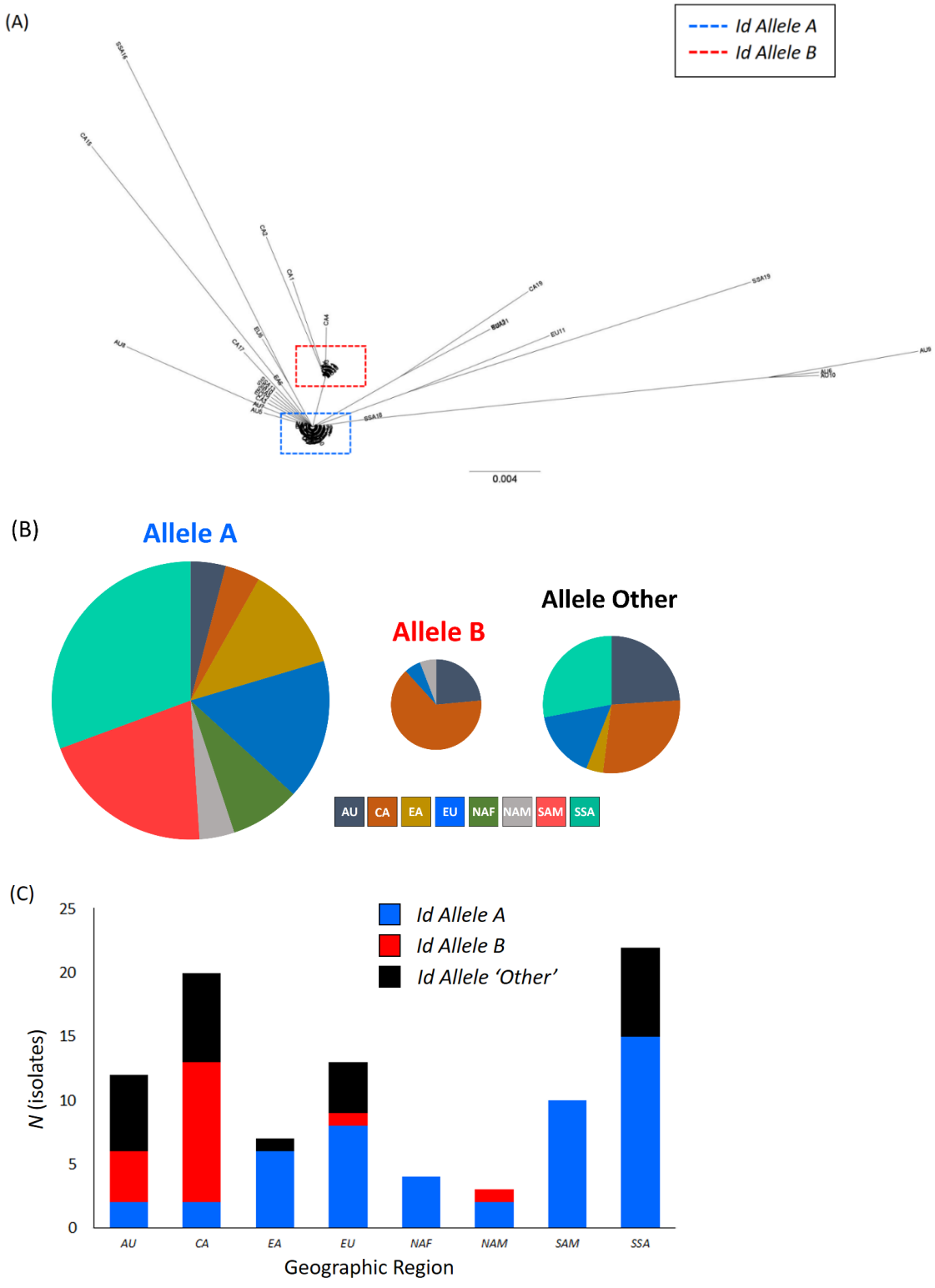
Allelic diversity of Ib-type *C. hominis* GP60 alleles



Appendix Figure XII. Phylogenetic clustering and geographic distribution of Ib-type GP60 alleles
17 unique alleles were identified from SNP-based comparisons of Ib GP60 consensus sequence alignments (DnaSP v 5.10.1; Librado & Rozas, 2009). Maximum likelihood phylogenetic reconstruction of ClustalW aligned gapless consensus sequences (A) revealed distinct clustering of isolates into predominantly two alleles, type A (58.5%) and type B (20.8%), with the remaining isolates comprising 'other' alleles at substantially lower total frequencies (0.8-6.2%) (Tamura-Nei Model, MEGA v 7.0). Geographic composition of these three subtype classes revealed inter-allelic differences, with allele A predominating in Europe, and allele B predominating in Central Asia and North America, and less conventional alleles dominating in Australasian and East Asian samples (B). Relative sizes of individual pie figures are accurate proportional representations of total isolates analysed. Regional differences of allelic composition were additionally assessed, revealing coexistence of both alleles A and B in all but two of the defined geographic regions (75.0%), and differential proportions for intra-regional compositions of all allele types (C).

Appendix XV

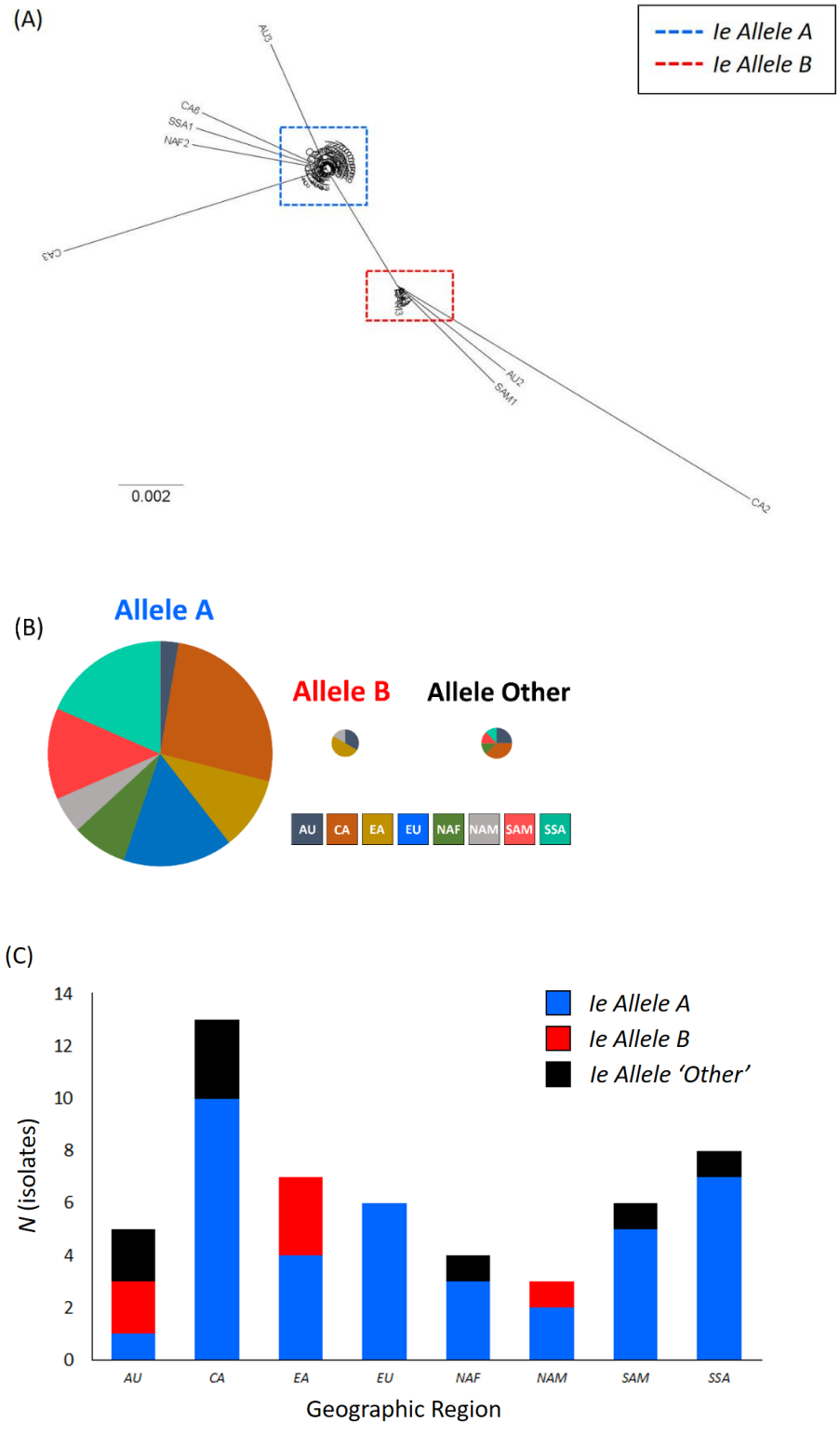
Allelic diversity of Id-type *C. hominis* GP60 alleles



Appendix Figure XIII. Phylogenetic clustering and geographic distribution of Id-type GP60 alleles
26 unique alleles were identified from SNP-based comparisons of Id GP60 consensus sequence alignments (DnaSP v 5.10.1; Librado & Rozas, 2009). Maximum likelihood phylogenetic reconstruction of ClustalW aligned gapless consensus sequences (A) revealed distinct clustering of isolates into predominantly two alleles, type A (53.8%) and type B (18.7%), with the remaining isolates comprising 'other' alleles at substantially lower total frequencies (1.1-2.2%) (Tamura-Nei Model, MEGA v 7.0). Geographic composition of these three subtype classes revealed inter-allelic differences, with allele A predominating in Sub-Saharan Africa and South America, and allele B predominating in Central Asia, and less conventional alleles dominating in Australasian, Central Asian and Sub-Saharan African samples (B). Relative sizes of individual pie figures are accurate proportional representations of total isolates analysed. Regional differences of allelic composition were additionally assessed, revealing coexistence of both alleles A and B in only half of the geographic regions compared (50.0%), and homogenous B-type allele populations in North Africa and South America (C).

Appendix XVI

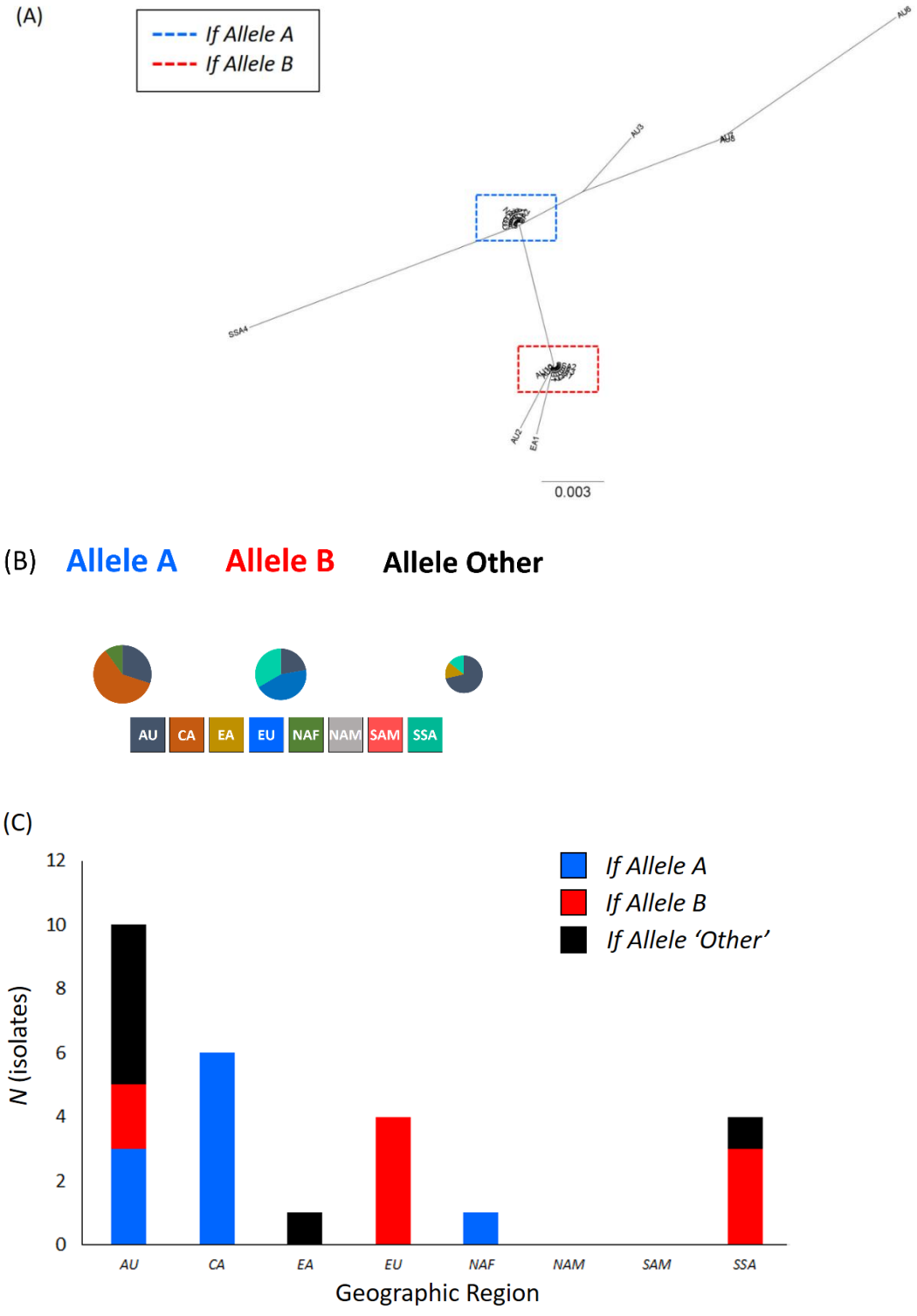
Allelic diversity of Ie-type *C. hominis* GP60 alleles



Appendix Figure XIV. Phylogenetic clustering and geographic distribution of Ie-type GP60 alleles
10 unique alleles were identified from SNP-based comparisons of Ie GP60 consensus sequence alignments (DnaSP v 5.10.1; Librado & Rozas, 2009). Maximum likelihood phylogenetic reconstruction of ClustalW aligned gapless consensus sequences (A) revealed distinct clustering of isolates into predominantly two alleles, type A (73.1%) and type B (11.5%), with the remaining isolates comprising 'other' alleles at substantially lower frequencies (1.9% each) (Tamura-Nei Model, MEGA v 7.0). Geographic composition of these three subtype classes revealed inter-allelic differences, with allele A predominating in Central Asia and Sub-Saharan Africa, and allele B predominating in Eastern Asia, and less conventional alleles dominating in Australasian and Central Asian samples (B). Relative sizes of individual pie figures are accurate proportional representations of total isolates analysed. Regional differences of allelic composition were additionally assessed, revealing coexistence of both alleles A and B in less than half of the geographic regions compared (37.5%), and homogenous A-type allele populations in Europe (C).

Appendix XVII

Allelic diversity of If-type *C. hominis* GP60 alleles



Appendix Figure XV. Phylogenetic clustering and geographic distribution of If-type GP60 alleles
8 unique alleles were identified from SNP-based comparisons of If GP60 consensus sequence alignments (DnaSP v 5.10.1; Librado & Rozas, 2009). Maximum likelihood phylogenetic reconstruction of ClustalW aligned gapless consensus sequences (A) revealed distinct clustering of isolates into predominantly two alleles, type A (38.5%) and type B (34.6%), with the remaining isolates comprising 'other' alleles at substantially lower frequencies (3.8-7.7%) (Tamura-Nei Model, MEGA v 7.0). Geographic composition of these three subtype classes revealed inter-allelic differences, with allele A predominating in Central Asia, and allele B predominating in Europe and Sub-Saharan Africa, and less conventional alleles dominating in Australasian samples (B). Relative sizes of individual pie figures are accurate proportional representations of total isolates analysed. Regional differences of allelic composition were additionally assessed, revealing coexistence of both alleles A and B in only a single geographic region (16.7%), and homogenous A-type allele populations in Central Asia and North Africa, versus homogenous B-type allele populations in Europe (C).

Appendix XVIII

N_{st} and G_{st} values for pairwise geographic comparisons of GP60 allele populations

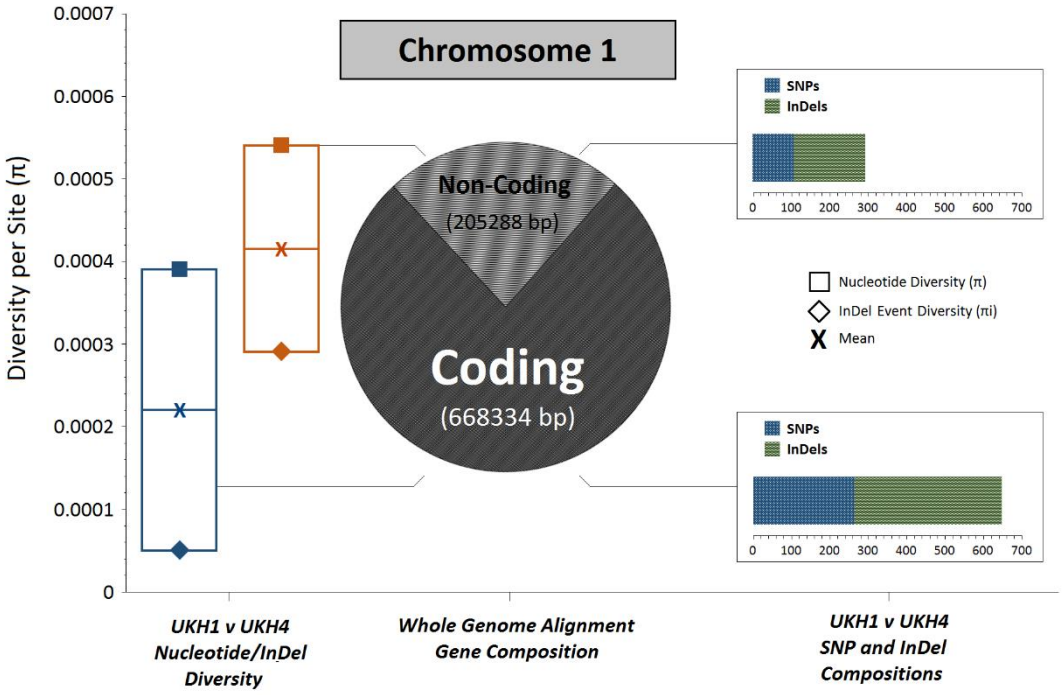
Subtype Ia				
Population 1	**Population 2**	**G_{st}**	**N_{st}**	**G_{st}-N_{st}**
AU	*CA*	0.07532	0.12196	-0.04664
AU	*EA*	0.12207	0.06417	0.05790
AU	*EU*	0.03389	0.04706	-0.01317
AU	*NAF*	0.01856	0.19236	-0.17380
AU	*NAM*	0.06196	0.00719	0.05477
AU	*SAM*	0.08184	0.04026	0.04158
AU	*SSA*	0.12339	0.13125	-0.00786
CA	*EA*	0.22744	0.40590	-0.17846
CA	*EU*	0.01149	0.00314	0.00835
CA	*NAF*	0.14070	0.14400	-0.00330
CA	*NAM*	0.12823	0.18388	-0.05565
CA	*SAM*	0.23920	0.37688	-0.13768
CA	*SSA*	0.00588	0.00110	0.00478
EA	*EU*	0.14776	0.23310	-0.08534
EA	*NAF*	0.13099	0.32626	-0.19527
EA	*NAM*	0.01907	0.05246	-0.03339
EA	*SAM*	0.00143	0.00026	0.00117
EA	*SSA*	0.42980	0.41649	0.01331
EU	*NAF*	0.06843	0.12636	-0.05793
EU	*NAM*	0.03849	0.05921	-0.02072
EU	*SAM*	0.11111	0.21070	-0.09959
EU	*SSA*	0.05575	0.00696	0.04879
NAF	*NAM*	0.09633	0.23226	-0.13593
NAF	*SAM*	0.09155	0.32223	-0.23068
NAF	*SSA*	0.25488	0.12681	0.12807
NAM	*SAM*	0.01463	0.05576	-0.04113
NAM	*SSA*	0.23860	0.20225	0.03635
SAM	*SSA*	0.35643	0.39054	-0.03411
Subtype Ib				
Population 1	**Population 2**	**G_{st}**	**N_{st}**	**G_{st}-N_{st}**
AU	*CA*	0.17206	0.43119	-0.25913
AU	*EA*	0.20880	0.47464	-0.26584
AU	*EU*	0.14715	0.06233	0.08482
AU	*NAF*	0.00696	0.00312	0.00384
AU	*NAM*	0.04336	0.04879	-0.00543
AU	*SAM*	0.10175	0.05100	0.05075
AU	*SSA*	0.01869	0.00194	0.01675
CA	*EA*	0.60440	0.82373	-0.21933
CA	*EU*	0.49233	0.85166	-0.35933
CA	*NAF*	0.29579	0.55550	-0.25971
CA	*NAM*	0.13103	0.33374	-0.20271
CA	*SAM*	0.63971	0.84836	-0.20865
CA	*SSA*	0.24082	0.44655	-0.20573
EA	*EU*	0.51415	0.79896	-0.28481
EA	*NAF*	0.35624	0.60677	-0.25053

EA	*NAM*	0.41395	0.58553	-0.17158
EA	*SAM*	0.66108	0.79592	-0.13484
EA	*SSA*	0.38206	0.52252	-0.14046
EU	*NAF*	0.04288	0.09613	-0.05325
EU	*NAM*	0.16948	0.33912	-0.16964
EU	*SAM*	0.00500	0.05797	-0.05297
EU	*SSA*	0.06733	0.12151	-0.05418
NAF	*NAM*	0.02941	0.00324	0.02617
NAF	*SAM*	0.03369	0.07489	-0.04120
NAF	*SSA*	0.02380	0.00703	0.01677
NAM	*SAM*	0.20032	0.32407	-0.12375
NAM	*SSA*	0.02517	0.00470	0.02047
SAM	*SSA*	0.06530	0.10429	-0.03899
Subtype Id				
Population 1	**Population 2**	**G_{st}**	**N_{st}**	**G_{st}-N_{st}**
AU	*CA*	0.00232	0.11206	-0.10974
AU	*EA*	0.15940	0.13943	0.01997
AU	*EU*	0.06399	0.10500	-0.04101
AU	*NAF*	0.19155	0.14456	0.04699
AU	*NAM*	0.03526	0.08057	-0.04531
AU	*SAM*	0.28227	0.14456	0.13771
AU	*SSA*	0.10448	0.09820	0.00628
CA	*EA*	0.21927	0.24485	-0.02558
CA	*EU*	0.13647	0.13111	0.00536
CA	*NAF*	0.22603	0.27063	-0.04460
CA	*NAM*	0.06812	0.00597	0.06215
CA	*SAM*	0.34279	0.27063	0.07216
CA	*SSA*	0.19745	0.16135	0.03610
EA	*EU*	0.01284	0.00084	0.01200
EA	*NAF*	0.00725	0.00000	0.00725
EA	*NAM*	-0.03194	0.00038	-0.03232
EA	*SAM*	0.00189	0.00000	0.00189
EA	*SSA*	0.01244	0.00242	0.01002
EU	*NAF*	0.06539	0.00151	0.06388
EU	*NAM*	-0.00228	-0.04007	0.03779
EU	*SAM*	0.08064	0.00151	0.07913
EU	*SSA*	0.01122	0.00212	0.00910
NAF	*NAM*	0.00297	0.00000	0.00297
NAF	*SSA*	0.05753	0.00000	0.05753
NAM	*SAM*	0.03045	0.00217	0.02828
NAM	*SSA*	0.03086	0.00000	0.03086
SAM	*SSA*	0.05410	0.00268	0.05142
Subtype Ie				
Population 1	**Population 2**	**G_{st}**	**N_{st}**	**G_{st}-N_{st}**
AU	*CA*	0.12817	0.14325	-0.01508
AU	*EA*	-0.01248	-0.07848	0.06600
AU	*EU*	0.29825	0.29938	-0.00113
AU	*NAF*	0.09416	0.24013	-0.14597
AU	*NAM*	-0.02094	-0.10747	0.08653
AU	*SAM*	0.15439	0.08768	0.06671

AU	*SSA*	0.19421	0.26645	-0.07224
CA	*EA*	0.06145	0.08478	-0.02333
CA	*EU*	0.03566	-0.00065	0.03631
CA	*NAF*	-0.00854	0.00022	-0.00876
CA	*NAM*	0.00256	0.00063	0.00193
CA	*SAM*	-0.01295	-0.03032	0.01737
CA	*SSA*	-0.00993	0.00013	-0.01006
EA	*EU*	0.1878	0.33333	-0.14553
EA	*NAF*	0.03625	0.21101	-0.17476
EA	*NAM*	-0.08808	-0.30000	0.21192
EA	*SAM*	0.07182	0.00044	0.07138
EA	*SSA*	0.10024	0.25834	-0.15810
EU	*NAF*	0.00415	0.00000	0.00415
EU	*NAM*	0.01370	0.00000	0.01370
EU	*SAM*	0.00000	0.00000	0.00000
EU	*SSA*	0.00149	0.00000	0.00149
NAF	*NAM*	-0.07348	0.00055	-0.07403
NAF	*SAM*	-0.04803	0.00055	-0.04858
NAF	*SSA*	-0.03226	0.00032	-0.03258
NAM	*SAM*	-0.04348	-0.20023	0.15675
NAM	*SSA*	-0.02227	0.00035	-0.02262
SAM	*SSA*	-0.03544	0.00035	-0.03579
Subtype If				
Population 1	**Population 2**	**G_{st}**	**N_{st}**	**G_{st}-N_{st}**
AU	*CA*	0.18019	0.17297	0.00722
AU	*EU*	0.19656	0.41986	-0.22330
AU	*SSA*	0.08901	0.12385	-0.03484
CA	*EU*	1.00000	1.00000	0.00000
CA	*SSA*	0.64339	0.39551	0.24788
EU	*SSA*	0.00000	0.00000	0.00000

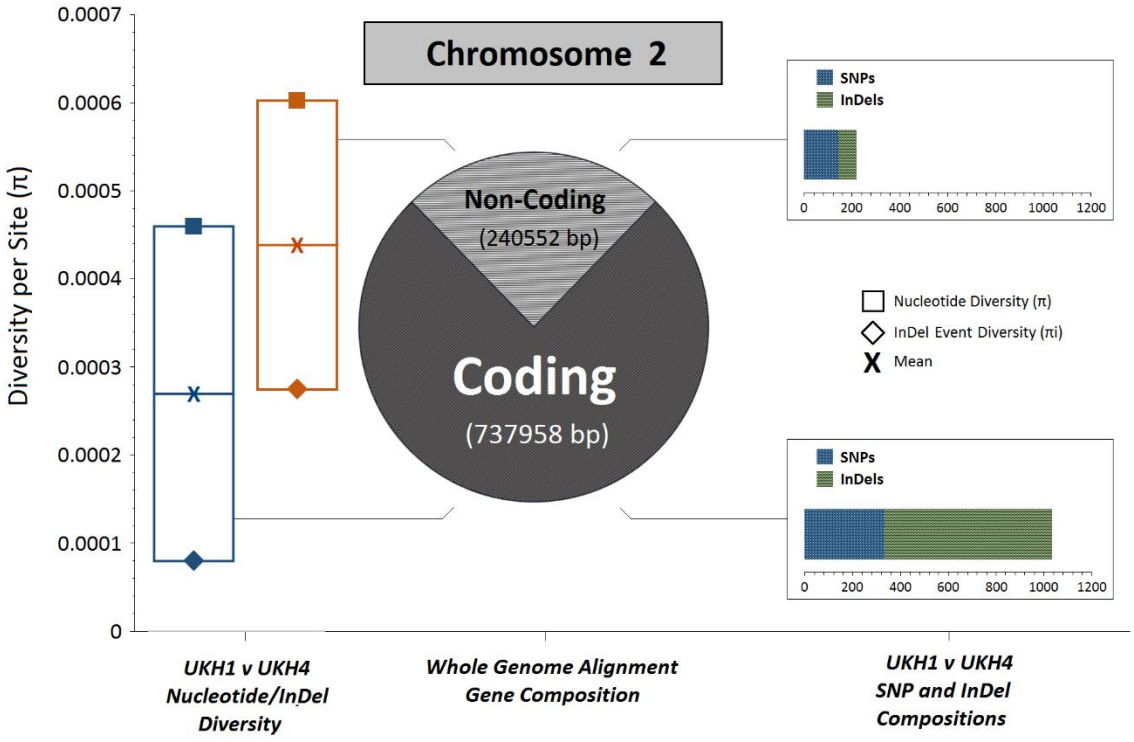
Appendix XIX

Characterization of genomic differences between *C. hominis* Ia and Ib whole genome sequences: Chromosome 1



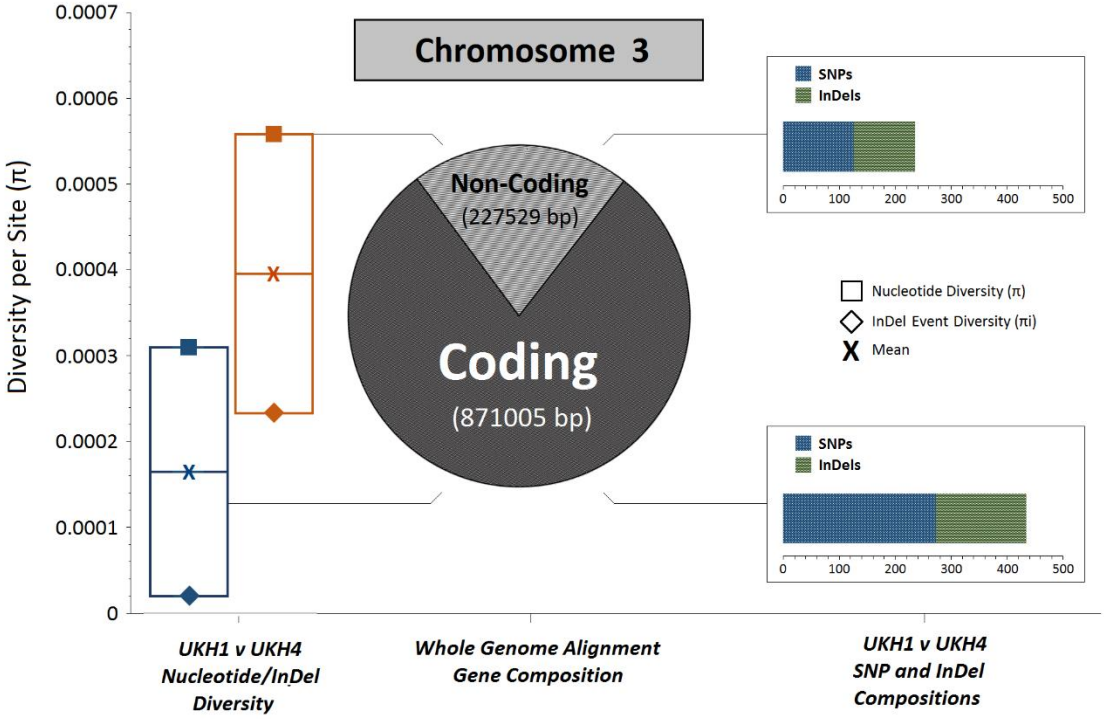
Appendix XX

Characterization of genomic differences between *C. hominis* Ia and Ib whole genome sequences: Chromosome 2



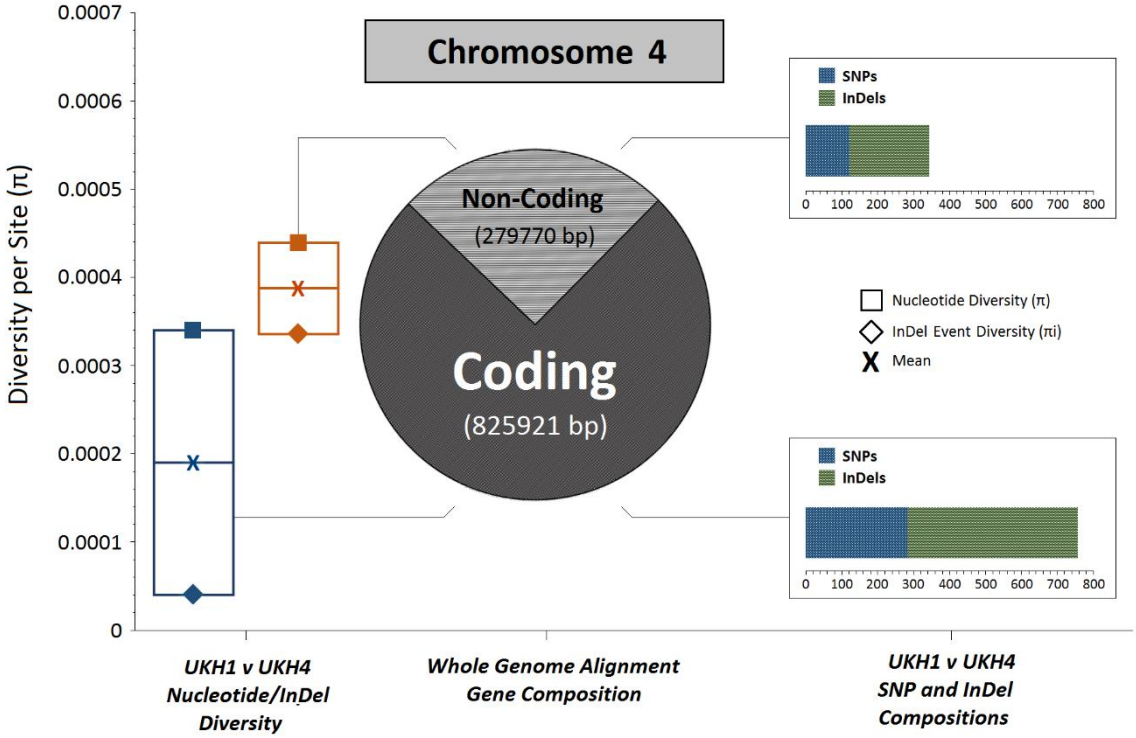
Appendix XXI

Characterization of genomic differences between *C. hominis* Ia and Ib whole genome sequences: Chromosome 3



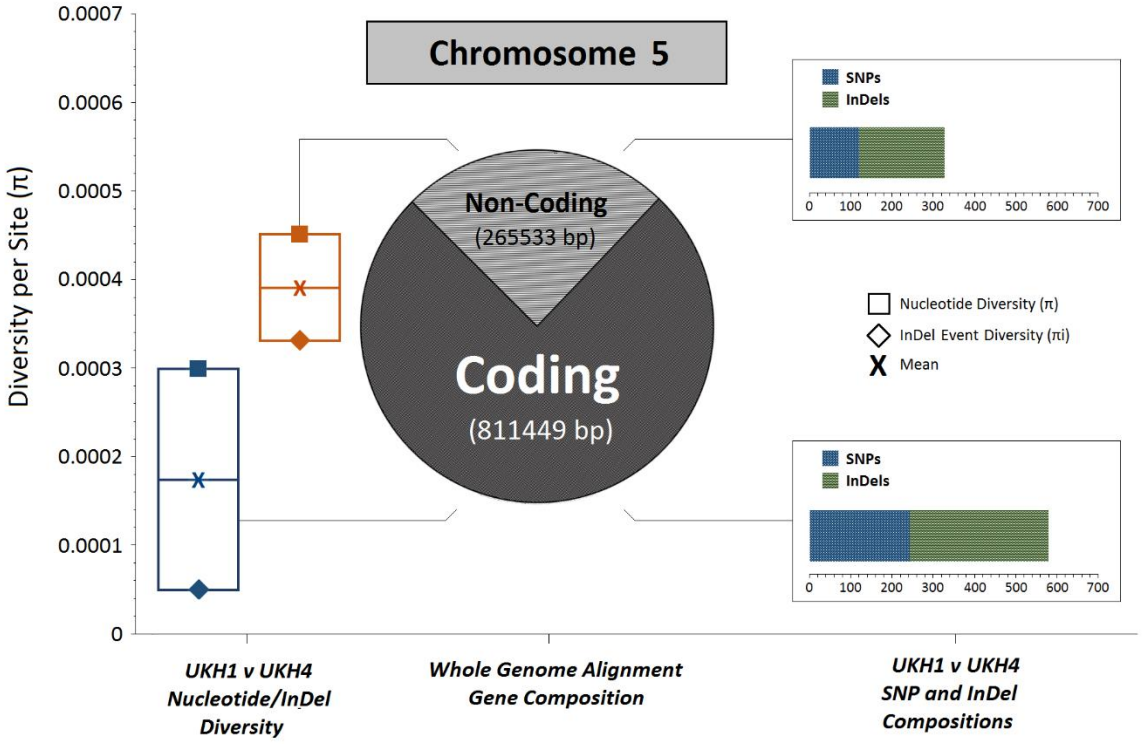
Appendix XXII

Characterization of genomic differences between *C. hominis* Ia and Ib whole genome sequences: Chromosome 4



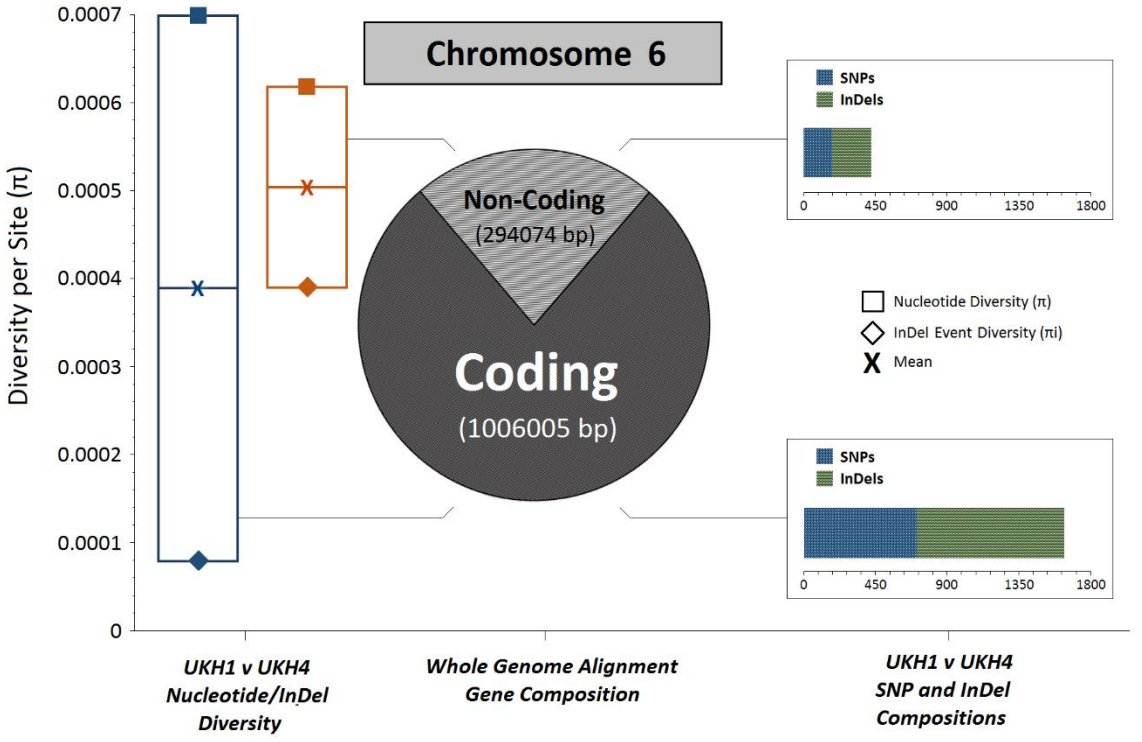
Appendix XXIII

Characterization of genomic differences between *C. hominis* Ia and Ib whole genome sequences: Chromosome 5



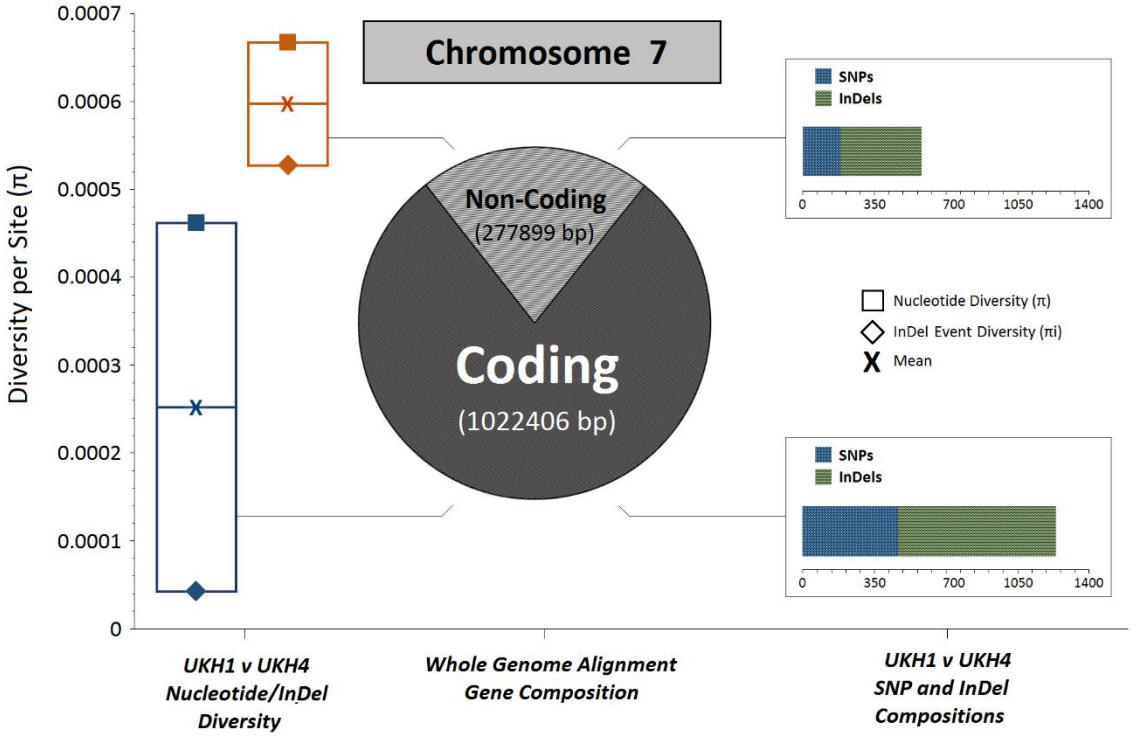
Appendix XXIV

Characterization of genomic differences between *C. hominis* Ia and Ib whole genome sequences: Chromosome 6



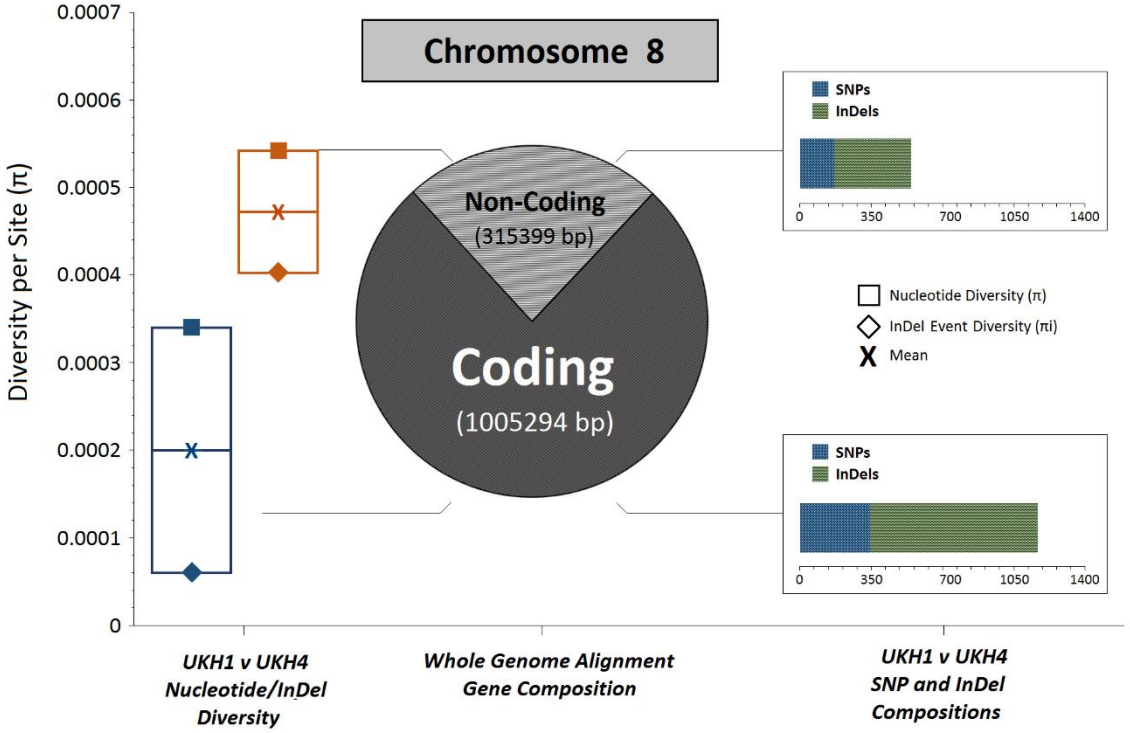
Appendix XXV

Characterization of genomic differences between *C. hominis* Ia and Ib whole genome sequences: Chromosome 7



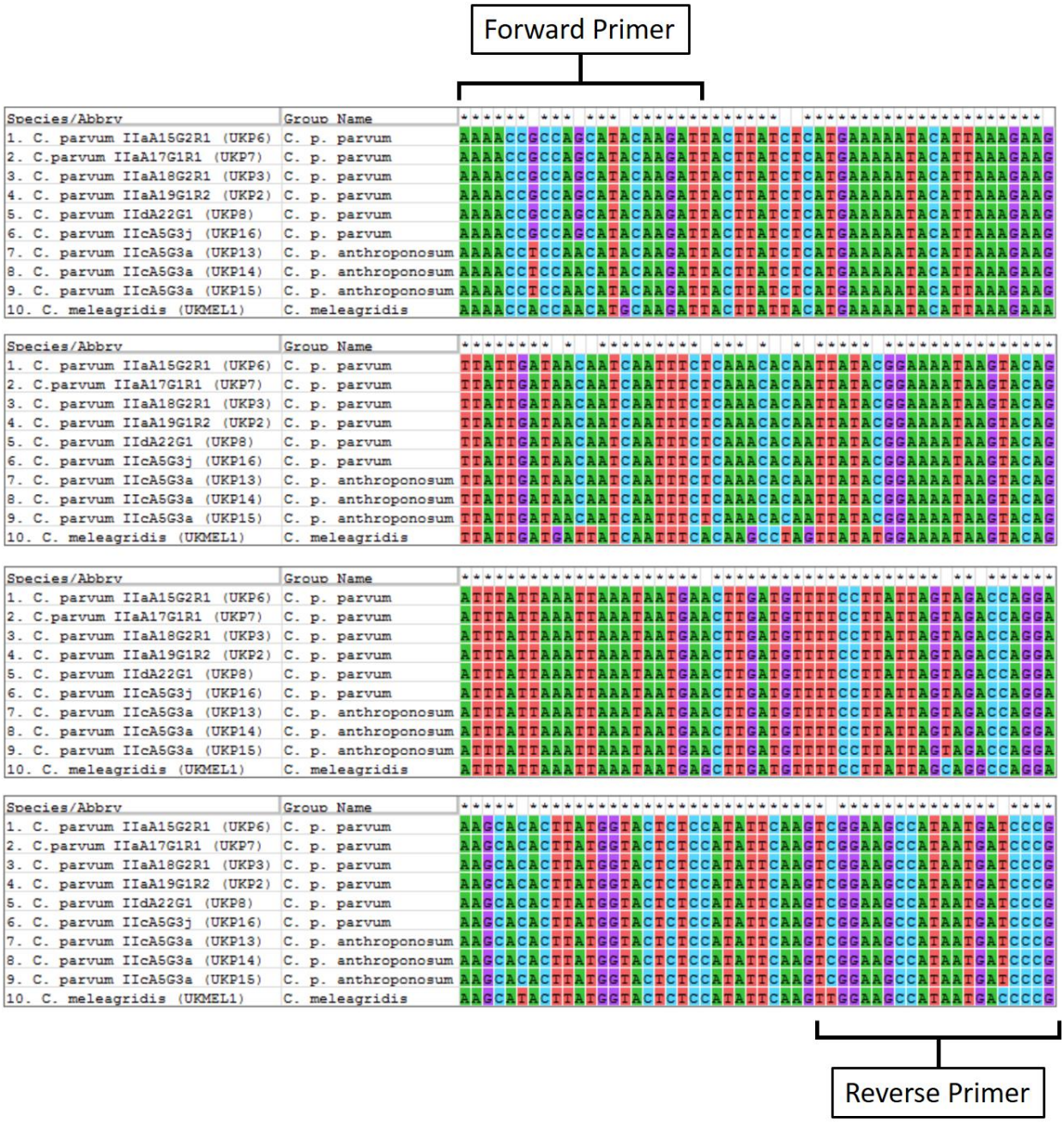
Appendix XXVI

Characterization of genomic differences between *C. hominis* Ia and Ib whole genome sequences: Chromosome 8



Appendix XXVII

Sequence alignment of 208 bp In Situ PCR target across variable *C. parvum* GP60 subtypes and *C. meleagridis* (Screenshots MEGA v 7.0, Kumar *et al*. 2016)



Species/Abbrv	Group Name
1. C. parvum IIaA15G2R1 (UKP6)	C. p. parvum
2. C.parvum IIaA17G1R1 (UKP7)	C. p. parvum
3. C. parvum IIaA18G2R1 (UKP3)	C. p. parvum
4. C. parvum IIaA19G1R2 (UKP2)	C. p. parvum
5. C. parvum IIdA22G1 (UKP8)	C. p. parvum
6. C. parvum IIcA5G3j (UKP16)	C. p. parvum
7. C. parvum IIcA5G3a (UKP13)	C. p. anthroponosum
8. C. parvum IIcA5G3a (UKP14)	C. p. anthroponosum
9. C. parvum IIcA5G3a (UKP15)	C. p. anthroponosum
10. C. meleagridis (UKMEL1)	C. meleagridis

REFERENCES

Abal-Fabeiro JL, Maside X, Bello X, Llovo J, Bartolomé C. Multilocus patterns of genetic variation across *Cryptosporidium* species suggest balancing selection at the gp60 locus. Mol Ecol. 2013 Sep;22(18):4723-32. doi: 10.1111/mec.12425. Epub 2013 Aug 5.

Abal-Fabeiro JL, Maside X, Llovo J, Bello X, Torres M, Treviño M, Moldes L, Muñoz A, Carracedo A, Bartolomé C. High-throughput genotyping assay for the large-scale genetic characterization of *Cryptosporidium* parasites from human and bovine samples. Parasitology. 2014 Apr;141(4):491-500. doi: 10.1017/S0031182013001807. Epub 2013 Nov 15.

Abal-Fabeiro JL, Maside X, Llovo J, Bartolomé C. Aetiology and epidemiology of human cryptosporidiosis cases in Galicia (NW Spain), 2000-2008. Epidemiol Infect. 2015 Feb 20:1-15. [Epub ahead of print]

Abbott JC, Aanensen DM, Bentley SD. WebACT: an online genome comparison suite. Methods Mol Biol. 2007; 395:57-74.

Abd El Kader NM, Blanco MA, Ali-Tammam M, Abd El Ghaffar Ael R, Osman A, El Sheikh N, Rubio JM, de Fuentes I. Detection of *Cryptosporidium parvum* and *Cryptosporidium hominis* in human patients in Cairo, Egypt. Parasitol Res. 2012 Jan;110(1):161-6. doi: 10.1007/s00436-011-2465-6. Epub 2011 May 24.

Abdel-Hafeez EH, Ahmad AK, Ali BA, Moslam FA. Opportunistic parasites among immunosuppressed children in Minia District, Egypt. Korean J Parasitol. 2012 Mar;50(1):57-62. doi: 10.3347/kjp.2012.50.1.57. Epub 2012 Mar 6.

Abdel-Hafeez EH, Belal US, Abdellatif MZ, Naoi K, Norose K. Breast-feeding protects infantile diarrhea caused by intestinal protozoan infections. Korean J Parasitol. 2013 Oct;51(5):519-24. doi: 10.3347/kjp.2013.51.5.519. Epub 2013 Oct 31.

Abdel-Messih IA, Wierzba TF, Abu-Elyazeed R, Ibrahim AF, Ahmed SF, Kamal K, Sanders J, Frenck R. Diarrhea associated with *Cryptosporidium parvum* among young children of the Nile River Delta in Egypt. J Trop Pediatr. 2005 Jun;51(3):154-9. Epub 2005 Apr 14.

Abdul-Sada KM. Molecular and Epidemiological Study of *Cryptosporidium spp*. in Mid-Euphrates Area. KUFA JOURNAL FOR NURSING SCIENCES. 2015 Jan; 5(1).

Abe N, Matsubayashi M, Kimata I, Iseki M. Subgenotype analysis of *Cryptosporidium parvum* isolates from humans and animals in Japan using the 60-kDa glycoprotein gene sequences. Parasitol Res. 2006 Aug;99(3):303-5. Epub 2006 Mar 25.

Abeywardena H, Jex AR, Nolan MJ, Haydon SR, Stevens MA, McAnulty RW, Gasser RB. Genetic characterisation of *Cryptosporidium* and Giardia from dairy calves: discovery of species/genotypes consistent with those found in humans. Infect Genet Evol. 2012 Dec;12(8):1984-93. doi: 10.1016/j.meegid.2012.08.004. Epub 2012 Sep 5.

Abrahamsen MS, Templeton TJ, Enomoto S, Abrahante JE, Zhu G, Lancto CA, Deng M, Liu C, Widmer G, Tzipori S, Buck GA, Xu P, Bankier AT, Dear PH, Konfortov BA, Spriggs HF, Iyer L, Anantharaman V, Aravind L, Kapur V. Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. Science. 2004 Apr 16;304(5669):441-5. Epub 2004 Mar 25.

Abreu-Acosta N, Quispe MA, Foronda-Rodríguez P, Alcoba-Florez J, Lorenzo-Morales J, Ortega-Rivas A, Valladares B. *Cryptosporidium* in patients with diarrhoea, on Tenerife, Canary Islands, Spain. Ann Trop Med Parasitol. 2007 Sep;101(6):539-45.

Abu Samra N, Thompson PN, Jori F, Frean J, Poonsamy B, du Plessis D, Mogoye B, Xiao L. Genetic characterization of *Cryptosporidium spp*. in diarrhoeic children from four provinces in South Africa. Zoonoses Public Health. 2013 Mar;60(2):154-9. doi: 10.1111/j.1863-2378.2012.01507.x. Epub 2012 Jun 20.

Abu Samra N, Jori F, Cacciò SM, Frean J, Poonsamy B, Thompson PN. *Cryptosporidium* genotypes in children and calves living at the wildlife or livestock interface of the Kruger National Park, South Africa. Onderstepoort J Vet Res. 2016 May 20;83(1):a1024. doi: 10.4102/ojvr.v83i1.1024.

Abu-Alrub SM, Abusada GM, Farraj MA, Essawi TA. Prevalence of *Cryptosporidium spp*. in children with diarrhoea in the West Bank, Palestine. J Infect Dev Ctries. 2008 Feb 1;2(1):59-62.

Acloque H, Wilkinson DG, Nieto MA. In situ hybridization analysis of chick embryos in whole-mount and tissue sections. Methods Cell Biol. 2008;87:169-85. doi: 10.1016/S0091-679X(08)00209-4.

Adam AA, Hassan HS, Shears P, Elshibly E. *Cryptosporidium* in Khartoum, Sudan. East Afr Med J. 1994 Nov;71(11):745-6.

Adamu H, Petros B, Hailu A, Petry F. Molecular characterization of *Cryptosporidium* isolates from humans in Ethiopia. Acta Trop. 2010 Jul-Aug;115(1-2):77-83. doi: 10.1016/j.actatropica.2010.02.003. Epub 2010 Mar 4.

Adamu H, Wegayehu T, Petros B. High prevalence of diarrhoegenic intestinal parasite infections among non-ART HIV patients in Fitche Hospital, Ethiopia. PLoS One. 2013 Aug 26;8(8):e72634. doi: 10.1371/journal.pone.0072634. eCollection 2013.

Adamu H, Petros B, Zhang G, Kassa H, Amer S, Ye J, Feng Y, Xiao L. Distribution and clinical manifestations of *Cryptosporidium* species and subtypes in HIV/AIDS patients in Ethiopia.PLoS Negl Trop Dis. 2014 Apr 17;8(4):e2831. doi: 10.1371/journal.pntd.0002831. eCollection 2014.

Addy PA, Antepim G, Frimpong EH. Prevalence of pathogenic Escherichia coli and parasites in infants with diarrhoea in Kumasi, Ghana. East Afr Med J. 2004 Jul;81(7):353-7.

Adegbola RA, Demba E, De Veer G, Todd J. *Cryptosporidium* infection in Gambian children less than 5 years of age. J Trop Med Hyg. 1994 Apr;97(2):103-7.

Adjei A, Lartey M, Adiku TK, Rodrigues O, Renner L, Sifah E, Mensah JD, Akanmori B, Otchere J, Bentum BK, Bosompem KM. *Cryptosporidium* oocysts in Ghanaian AIDS patients with diarrhoea. East Afr Med J. 2003 Jul;80(7):369-72.

Adjei AA, Armah H, Rodrigues O, Renner L, Borketey P, Ayeh-Kumi P, Adiku T, Sifah E, Lartey M. *Cryptosporidium spp*., a frequent cause of diarrhea among children at the Korle-Bu Teaching Hospital, Accra, Ghana. Jpn J Infect Dis. 2004 Oct;57(5):216-9.

Adwan A. MacNeal's stain for blood films. Am J Med Technol. 1952 Nov-Dec;18(6):311-2.

Aghamolaie S, Rostami A, Fallahi Sh, Tahvildar Biderouni F, Haghighi A, Salehi N. Evaluation of modified Ziehl-Neelsen, direct fluorescent-antibody and PCR assay for detection of *Cryptosporidium spp*. in children faecal specimens. J Parasit Dis. 2016 Sep;40(3):958-63. doi: 10.1007/s12639-014-0614-4. Epub 2014 Nov 26.

Agholi M, Hatam GR, Motazedian MH. HIV/AIDS-associated opportunistic protozoal diarrhea. AIDS Res Hum Retroviruses. 2013 Jan;29(1):35-41. doi: 10.1089/AID.2012.0119. Epub 2012 Sep 11.

Agholi M, Hatam GR, Motazedian MH. Microsporidia and coccidia as causes of persistence diarrhea among liver transplant children: incidence rate and species/genotypes. Pediatr Infect Dis J. 2013 Feb;32(2):185-7. doi: 10.1097/INF.0b013e318273d95f.

Agholi M, Shahabadi SN, Motazedian MH, Hatam GR. Prevalence of Enteric Protozoan Oocysts with Special Reference to Sarcocystis cruzi among Fecal Samples of Diarrheic Immunodeficient Patients in Iran. Korean J Parasitol. 2016 Jun;54(3):339-44. doi: 10.3347/kjp.2016.54.3.339. Epub 2016 Jun 30.

Agrawal Y, Narwani T, Subramanian S. Genome sequence and comparative analysis of clavicipitaceous insect-pathogenic fungus Aschersonia badia with Metarhizium spp. BMC Genomics. 2016 May 17;17:367. doi: 10.1186/s12864-016-2710-6.

Ahmed NH, Chowdhary A. Pattern of co-infection by enteric pathogenic parasites among HIV sero-positive individuals in a Tertiary Care Hospital, Mumbai, India. Indian J Sex Transm Dis. 2015 Jan-Jun;36(1):40-7. doi: 10.4103/0253-7184.156707.

Ajjampur SS, Gladstone BP, Selvapandian D, Muliyil JP, Ward H, Kang G. Molecular and spatial epidemiology of cryptosporidiosis in children in a semiurban community in South India. J Clin Microbiol. 2007 Mar;45(3):915-20. Epub 2007 Jan 24.

Ajjampur SS, Liakath FB, Kannan A, Rajendran P, Sarkar R, Moses PD, Simon A, Agarwal I, Mathew A, O'Connor R, Ward H, Kang G. Multisite study of cryptosporidiosis in children with diarrhea in India. J Clin Microbiol. 2010 Jun;48(6):2075-81. doi: 10.1128/JCM.02509-09. Epub 2010 Apr 14.

Akinbo FO, Okaka CE, Omoregie R. Prevalence of intestinal parasitic infections among HIV patients in Benin City, Nigeria. Libyan J Med. 2010 Oct 29;5. doi: 10.3402/ljm.v5i0.5506.

Akinbo FO, Okaka CE, Omoregie R. Seasonal Variation of Intestinal Parasitic Infections Among HIV-Positive Patients in Benin City, Nigeria. Ethiop J Health Sci. 2011 Nov;21(3):191-4.

Akinbo FO, Okaka CE, Omoregie R, Adamu H, Xiao L. Unusual Enterocytozoon bieneusi genotypes and *Cryptosporidium hominis* subtypes in HIV-infected patients on highly active antiretroviral therapy. Am J Trop Med Hyg. 2013 Jul;89(1):157-61. doi: 10.4269/ajtmh.12-0635. Epub 2013 Apr 29.

Akiyoshi DE, Tumwine JK, Bakeera-Kitaka S, Tzipori S. Subtype analysis of *Cryptosporidium* isolates from children in Uganda. J Parasitol. 2006 Oct;92(5):1097-100.

Aksoy U, Tuncay S. [Short communication: investigation of intestinal coccidia in patients with diarrhea]. Mikrobiyol Bul. 2007 Jan;41(1):127-31.

Alagappan A, Tujula NA, Power M, Ferguson CM, Bergquist PL, Ferrari BC. Development of fluorescent in situ hybridization for *Cryptosporidium* detection reveals zoonotic and anthroponotic transmission of sporadic cryptosporidiosis in Sydney. J Microbiol Methods. 2008 Dec;75(3):535-9. doi: 10.1016/j.mimet.2008.08.007. Epub 2008 Aug 27.

Alagappan A, Bergquist PL, Ferrari BC. Development of a two-color fluorescence in situ hybridization technique for species-level identification of human-infectious *Cryptosporidium spp*. Appl Environ Microbiol. 2009 Sep;75(18):5996-8. doi: 10.1128/AEM.00643-09. Epub 2009 Jul 24.

Albert MJ, Rotimi VO, Iqbal J, Chehadeh W. Evaluation of the xTAG Gastrointestinal Pathogen Panel Assay for the Detection of Enteric Pathogens in Kuwait. Med Princ Pract. 2016;25(5):472-6. doi: 10.1159/000447698. Epub 2016 Jun 20.

Al-Braiken FA, Amin A, Beeching NJ, Hommel M, Hart CA. Detection of *Cryptosporidium* amongst diarrhoeic and asymptomatic children in Jeddah, Saudi Arabia. Ann Trop Med Parasitol. 2003 Jul;97(5):505-10.

Al-Braiken FA. Is intestinal parasitic infection still a public health concern among Saudi children? Saudi Med J. 2008 Nov;29(11):1630-5.

Al-Brikan FA, Salem HS, Beeching N, Hilal N. Multilocus genetic analysis of *Cryptosporidium* isolates from Saudi Arabia. J Egypt Soc Parasitol. 2008 Aug;38(2):645-58.

Al-Delaimy AK, Al-Mekhlafi HM, Nasr NA, Sady H, Atroosh WM, Nashiry M, Anuar TS, Moktar N, Lim YA, Mahmud R. Epidemiology of intestinal polyparasitism among Orang Asli school children in rural Malaysia. PLoS Negl Trop Dis. 2014 Aug 21;8(8):e3074. doi: 10.1371/journal.pntd.0003074. eCollection 2014.

Alemu A, Shiferaw Y, Getnet G, Yalew A, Addis Z. Opportunistic and other intestinal parasites among HIV/AIDS patients attending Gambi higher clinic in Bahir Dar city, North West Ethiopia. Asian Pac J Trop Med. 2011 Aug;4(8):661-5. doi: 10.1016/S1995-7645(11)60168-5.

Al-Harthi SA, Jamjoom MB. Preliminary study of the prevalence of intestinal parasites among diarrheic inhabitants in Makkah Al-Mukarramah. J Egypt Soc Parasitol. 2007 Aug;37(2):671-80.

Al-Hindi AI, Elmanama AA, Elnabris KJA. Cryptosporidiosis among children attending Al-Nasser Pediatric Hospital, Gaza, Palestine. Turk J Med Sci. 2007; 37: 367-72.

Ali MB, Ghenghesh KS, Aissa RB, Abuhelfaia A, Dufani M. Etiology of childhood diarrhea in Zliten, Libya. Saudi Med J. 2005 Nov;26(11):1759-65.

Ali S, Mumar S, Kalam K, Raja K, Baqi S. Prevalence, clinical presentation and treatment outcome of cryptosporidiosis in immunocompetent adult patients presenting with acute diarrhoea. J Pak Med Assoc. 2014 Jun;64(6):613-8.

Al-Megrin WA. Intestinal parasites infection among immunocompromised patients in Riyadh, Saudi Arabia. Pak J Biol Sci. 2010 Apr 15;13(8):390-4.

Almeida AA, Delgado ML, Soares SC, Castro AO, Moreira MJ, Mendonça CM, Canada NB, Da Costa JM, Coelho HG. Genetic characterization of *Cryptosporidium* isolates from humans in northern Portugal. J Eukaryot Microbiol. 2006;53 Suppl 1:S26-7.

Al-Mekhlafi HM, Mahdy MA, 'Azlin MY, Fatmah MS, Norhayati M. Childhood *Cryptosporidium* infection among aboriginal communities in Peninsular Malaysia. Ann Trop Med Parasitol. 2011 Mar;105(2):135-43. doi: 10.1179/136485911X12899838683368.

Al-Qobati SA, Al-Maktari MT, Al-Zoa AM, Derhim M. Intestinal parasitosis among Yemeni patients with cancer, Sana'a, Yemen. J Egypt Soc Parasitol. 2012 Dec;42(3):727-34.

Al-Shamiri A, Al-Zubairy A, Al-Mamari R. The Prevalence of *Cryptosporidium spp*. in Children, Taiz District, Yemen. Iran J Parasitol. 2010 Jun;5(2):26-32.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic Local Alignment Search Tool. J. Mol. Biol. 1990. 215: 403-410.

Alver O, Töre O. [The prevalence and distribution of intestinal parasites detected by the Uludag University Medical School]. Turkiye Parazitol Derg. 2006;30(4):296-301.

Alver O, Oral B, Töre O. [The distribution of intestinal parasites detected in the Uludag University Medical School Hospital between 2005 and 2008]. Turkiye Parazitol Derg. 2011;35(4):194-8. doi: 10.5152/tpd.2011.46.

Alves M, Matos O, Spano F, Antunes F. PCR-RFLP analysis of *Cryptosporidium parvum* isolates from HIV-infected patients in Lisbon, Portugal. Ann Trop Med Parasitol. 2000 Apr;94(3):291-7.

Alves M, Matos O, Pereira Da Fonseca I, Delgado E, Lourenço AM, Antunes F. Multilocus genotyping of *Cryptosporidium* isolates from human HIV-infected and animal hosts. J Eukaryot Microbiol. 2001;Suppl:17S-18S.

Alves M, Xiao L, Sulaiman I, Lal AA, Matos O, Antunes F. Subgenotype analysis of *Cryptosporidium* isolates from humans, cattle, and zoo ruminants in Portugal. J Clin Microbiol. 2003 Jun;41(6):2744-7.

Alves M, Xiao L, Antunes F, Matos O. Distribution of *Cryptosporidium* subtypes in humans and domestic and wild ruminants in Portugal. Parasitol Res. 2006 Aug;99(3):287-92. Epub 2006 Mar 22.

Alyousefi NA, Mahdy MA, Mahmud R, Lim YA. Factors associated with high prevalence of intestinal protozoan infections among patients in Sana'a City, Yemen. PLoS One. 2011;6(7):e22044. doi: 10.1371/journal.pone.0022044. Epub 2011 Jul 18.

Alyousefi NA, Mahdy MA, Lim YA, Xiao L, Mahmud R. First molecular characterization of *Cryptosporidium* in Yemen. Parasitology. 2013 May;140(6):729-34. doi: 10.1017/S0031182012001953. Epub 2013 Feb 1.

Amadi B, Kelly P, Mwiya M, Mulwazi E, Sianongo S, Changwe F, Thomson M, Hachungula J, Watuka A, Walker-Smith J, Chintu C. Intestinal and systemic infection, HIV, and mortality in Zambian children with persistent diarrhea and malnutrition. J Pediatr Gastroenterol Nutr. 2001 May;32(5):550-4.

Amar CF, East CL, Gray J, Iturriza-Gomara M, Maclure EA, McLauchlin J. Detection by PCR of eight groups of enteric pathogens in ,627 faecal samples: re-examination of the English case-control Infectious Intestinal Disease Study (1993-1996). Eur J Clin Microbiol Infect Dis. 2007 May;26(5):311-23.

Amatya R, Poudyal N, Gurung R, Khanal B. Prevalence of *Cryptosporidium* species in paediatric patients in Eastern Nepal. Trop Doct. 2011 Jan;41(1):36-7. doi: 10.1258/td.2010.100110. Epub 2010 Nov 25.

Amin OM. Seasonal prevalence of intestinal parasites in the United States during 2000. Am J Trop Med Hyg. 2002 Jun;66(6):799-803.

Anderson MZ, Wigen LJ, Burrack LS, Berman J. Real-Time Evolution of a Subtelomeric Gene Family in Candida albicans. Genetics. 2015 Jul;200(3):907-19. doi: 10.1534/genetics.115.177451. Epub 2015 May 8.

Andrews TD, Gojobori T. Strong positive selection and recombination drive the antigenic variation of the PilE protein of the human pathogen Neisseria meningitidis. Genetics. 2004 Jan;166(1):25-32.

Angal L, Mahmud R, Samin S, Yap NJ, Ngui R, Amir A, Ithoi I, Kamarulzaman A, Lim YA. Determining intestinal parasitic infections (IPIs) in inmates from Kajang Prison, Selangor, Malaysia for improved prison management. BMC Infect Dis. 2015 Oct 29;15:467. doi: 10.1186/s12879-015-1178-3.

Angus SV, Kwon HJ, Yoon JY. Field-deployable and near-real-time optical microfluidic biosensors for single-oocyst-level detection of *Cryptosporidium parvum* from field water samples. J Environ Monit. 2012 Dec;14(12):3295-304. doi: 10.1039/c2em30700f. Epub 2012 Nov 14.

Aniesona AT, Bamaiyi PH. Retrospective study of cryptosporidiosis among diarrhoeic children in the arid region of north-eastern Nigeria. Zoonoses Public Health. 2014 Sep;61(6):420-6. doi: 10.1111/zph.12088. Epub 2013 Nov 18.

ANOFEL *Cryptosporidium* National Network. Laboratory-based surveillance for *Cryptosporidium* in France, 2006-2009. Euro Surveill. 2010 Aug 19;15(33):19642.

Anon. The development of a national collection for oocysts of *Cryptosporidium*. Foundation for Water Research, Marlow, Bucks, UK (http://www.fwr.org/); 2002.

Antia M, Herricks T, Rathod PK. Microfluidic modeling of cell-cell interactions in malaria pathogenesis. PLoS Pathog. 2007;3:e99

Araújo AJ, Kanamura HY, Almeida ME, Gomes AH, Pinto TH, Da Silva AJ. Genotypic identification of *Cryptosporidium spp*. isolated from HIV-infected patients and immunocompetent children of São Paulo, Brazil. Rev Inst Med Trop Sao Paulo. 2008 May-Jun;50(3):139-43. Epub 2008 May 14.

Areeshi M, Dove W, Papaventsis D, Gatei W, Combe P, Grosjean P, Leatherbarrow H, Hart CA. *Cryptosporidium* species causing acute diarrhoea in children in Antananarivo, Madagascar. Ann Trop Med Parasitol. 2008 Jun;102(4):309-15. doi: 10.1179/136485908X278793.

Arenas-Pinto A, Certad G, Ferrara G, Castro J, Bello MA, Núñez LT. Association between parasitic intestinal infections and acute or chronic diarrhoea in HIV-infected patients in Caracas, Venezuela. Int J STD AIDS. 2003 Jul;14(7):487-92.

Arslan H, Inci EK, Azap OK, Karakayali H, Torgay A, Haberal M. Etiologic agents of diarrhea in solid organ recipients. Transpl Infect Dis. 2007 Dec;9(4):270-5. Epub 2007 May 19.

Asma I, Johari S, Sim BL, Lim YA. How common is intestinal parasitism in HIV-infected patients in Malaysia? Trop Biomed. 2011 Aug;28(2):400-10.

Asma I, Sim BL, Brent RD, Johari S, Yvonne Lim AL. Molecular epidemiology of *Cryptosporidium* in HIV/AIDS patients in Malaysia. Trop Biomed. 2015 Jun;32(2):310-22.

Aspöck H, Hassl A. Parasitic infections in HIV patients in Austria: first results of a long-term study. Zentralbl Bakteriol. 1990 Apr;272(4):540-6.

Assefa S, Erko B, Medhin G, Assefa Z, Shimelis T. Intestinal parasitic infections in relation to HIV/AIDS status, diarrhea and CD4 T-cell count. BMC Infect Dis. 2009 Sep 18;9:155. doi: 10.1186/1471-2334-9-155.

Assis DC, Resende DV, Cabrine-Santos M, Correia D, Oliveira-Silva MB. Prevalence and genetic characterization of *Cryptosporidium spp*. and Cystoisospora belli in HIV-infected patients. Rev Inst Med Trop Sao Paulo. 2013;55(3). pii: S0036-46652013000300149. doi: 10.1590/S0036-46652013000300002.

Assoumou A, Kone M, Penali LK, Coulibaly M, N'Draman AA. [Cryptosporidiosis and HIV in Abidjan (Ivory Coast)]. Bull Soc Pathol Exot. 1993;86(2):85-6.

Atayde VD, Aslan H, Townsend S, Hassani K, Kamhawi S, Olivier M. Exosome Secretion by the Parasitic Protozoan Leishmania within the Sand Fly Midgut. Cell Rep. 2015 Nov 3;13(5):957-67. doi: 10.1016/j.celrep.2015.09.058. Epub 2015 Oct 22.

Athman A, Tanz SK, Conn VM, Jordans C, Mayo GM, Ng WW, Burton RA, Conn SJ, Gilliham M. Protocol: a fast and simple in situ PCR method for localising gene expression in plant tissue. Plant Methods. 2014 Sep 18;10:29. doi: 10.1186/1746-4811-10-29. eCollection 2014.

Atterholm I, Castor B, Norlin K. Cryptosporidiosis in southern Sweden. Scand J Infect Dis. 1987;19(2):231-4.

Australian Society for Parasitology. *Cryptosporidium*. Para-Site. The Australian Society for Parasitology Inc., 16 June 2010. Web. 17 Nov. 2016.

Ayalew D, Boelee E, Endeshaw T, Petros B. *Cryptosporidium* and Giardia infection and drinking water sources among children in Lege Dini, Ethiopia. Trop Med Int Health. 2008 Apr;13(4):472-5. doi: 10.1111/j.1365-3156.2008.02024.x. Epub 2008 Feb 14.

Aye T, Moe K, Nyein MM, Swe T. Cryptosporidiosis in Myanmar infants with acute diarrhea. Southeast Asian J Trop Med Public Health. 1994 Dec;25(4):654-6.

Aygun G, Yilmaz M, Yasar H, Aslan M, Polat E, Midilli K, Ozturk R, Altas K. Parasites in nosocomial diarrhoea: are they underestimated? J Hosp Infect. 2005 Jul;60(3):283-5.

Ayinmode AB, Fagbemi BO, Xiao L. Molecular characterization of *Cryptosporidium* in children in Oyo State, Nigeria: implications for infection sources. Parasitol Res. 2012 Jan;110(1):479-81. doi: 10.1007/s00436-011-2531-0. Epub 2011 Jul 9.

Ayinmode AB, Zhang H, Dada-Adegbola HO, Xiao L. *Cryptosporidium hominis* subtypes and Enterocytozoon bieneusi genotypes in HIV-infected persons in Ibadan, Nigeria. Zoonoses Public Health. 2014 Jun;61(4):297-303. doi: 10.1111/zph.12072. Epub 2013 Jul 22.

Azami M, Moghaddam DD, Salehi R, Salehi M. [The identification of *Cryptosporidium* species (protozoa) in Ifsahan, Iran by PCR-RFLP analysis of the 18S rRNA gene]. Mol Biol (Mosk). 2007 Sep-Oct;41(5):934-9.

Bachur TP, Vale JM, Coêlho IC, Queiroz TR, Chaves Cde S. Enteric parasitic infections in HIV/AIDS patients before and after the highly active antiretroviral therapy. Braz J Infect Dis. 2008 Apr;12(2):115-22.

Bacon V, Jellum E, Patton W, Pereira W, Halpern B. Peptide sequencing by low resolution mass spectrometry. I. The use of acetylacetonyl derivatives to identify N-terminal residues. Biochem Biophys Res Commun. 1969 Dec 4;37(6):878-82.

Badugu SB, Nabi SA, Vaidyam P, Laskar S, Bhattacharyya S, Bhattacharyya MK. Identification of Plasmodium falciparum DNA Repair Protein Mre11 with an Evolutionarily Conserved Nuclease Function. PLoS One. 2015 May 4;10(5):e0125358. doi: 10.1371/journal.pone.0125358. eCollection 2015.

Bajer A, Bednarska M, Cacciò SM, Wolska-Kuśnierz B, Heropolitanska-Pliszka E, Bernatowska E, Wielopolska M, Paziewska A, Welc-faleciak R, Siński E. Genotyping of *Cryptosporidium* isolates from human clinical cases in Poland. Parasitol Res. 2008 Jun;103(1):37-42. doi: 10.1007/s00436-008-0924-5.

Ballal M, Shivananda PG. Rotavirus and enteric pathogens in infantile diarrhoea in Manipal, South India. Indian J Pediatr. 2002 May;69(5):393-6.

Bandin F, Kwon T, Linas MD, Guigonis V, Valentin A, Cassaing S, Carol A, Garnier A, Baudouin V, Decramer S. Cryptosporidiosis in paediatric renal transplantation. Pediatr Nephrol. 2009 Nov;24(11):2245-55. doi: 10.1007/s00467-009-1274-y.

Bandini G, Haserick JR, Motari E, Ouologuem DT, Lourido S, Roos DS, Costello CE, Robbins PW, Samuelson J. O-fucosylated glycoproteins form assemblies in close proximity to the nuclear pore complexes of Toxoplasma gondii. Proc Natl Acad Sci U S A. 2016 Oct 11;113(41):11567-11572. Epub 2016 Sep 23.

Banerjee D, Deb R, Dar L, Mirdha BR, Pati SK, Thareja S, Falodia S, Ahuja V. High frequency of parasitic and viral stool pathogens in patients with active ulcerative colitis: report from a tropical country. Scand J Gastroenterol. 2009;44(3):325-31. doi: 10.1080/00365520802556809.

Banisch DM, El-Badry A, Klinnert JV, Ignatius R, El-Dib N. Simultaneous detection of Entamoeba histolytica/dispar, Giardia duodenalis and cryptosporidia by immunochromatographic assay in stool samples from patients living in the Greater Cairo Region, Egypt. World J Microbiol Biotechnol. 2015 Aug;31(8):1251-8. doi: 10.1007/s11274-015-1875-5. Epub 2015 May 28.

Banwat EB, Egah DZ, Onile BA, Angyo IA, Audu ES. Prevalence of *Cryptosporidium* infection among undernourished children in Jos, Central Nigeria. Niger Postgrad Med J. 2003 Jun;10(2):84-7.

Baptista Rde P, D'Ávila DA, Segatto M, do Valle ÍF, Franco GR, Valadares HM, Gontijo ED4, Galvão LM, Pena SD, Chiari E, Machado CR, Macedo AM. Evidence of substantial recombination among Trypanosoma cruzi II strains from Minas Gerais. Infect Genet Evol. 2014 Mar;22:183-91. doi: 10.1016/j.meegid.2013.11.021. Epub 2013 Dec 1.

Baqai R, Anwar S, Kazmi SU. Detection of *Cryptosporidium* in immunosuppressed patients. J Ayub Med Coll Abbottabad. 2005 Jul-Sep;17(3):38-40.

Baranasic D, Oppermann T, Cheaib M, Cullum J, Schmidt H, Simon M. Genomic characterization of variable surface antigens reveals a telomere position effect as a prerequisite for RNA interference-mediated silencing in Paramecium tetraurelia. MBio. 2014 Nov 11;5(6):e01328. doi: 10.1128/mBio.01328-14.

Barboni G, Candi M, Inés Villacé M, Leonardelli A, Balbaryski J, Gaddi E. [Intestinal cryptosporidiosis in HIV infected children]. Medicina (B Aires). 2008;68(3):213-8.

Barker IK, Carbonell PL. *Cryptosporidium* agni sp.n. from lambs, and *Cryptosporidium* bovis sp.n. from a calf, with observations on the oocyst. Z Parasitenkd. 1974;44(4):289-98.

Barry JD Ginger ML, Burton P, McCulloch R. Why are parasite contingency genes often associated with telomeres? Int J Parasitol. 2003 Jan;33(1):29-45.

Barugahare R, Becker JA, Landosb M, Šlapeta J, Dennis MM. Gastric cryptosporidiosis in farmed Australian Murray cod, Maccullochella peelii peelii. Aquaculture. 2011; 314 (1-4), 1-6.

Baumgartner A, Marder HP, Munzinger J, Siegrist HH. Frequency of *Cryptosporidium spp*. as cause of human gastrointestinal disease in Switzerland and possible sources of infection. Schweiz Med Wochenschr. 2000 Sep 9;130(36):1252-8.

Beal SG, Couturier MR, Gander RM, Doern CD. Diagnostic Algorithm for the Diagnosis of Pediatric Parasitic Gastroenteritis. J Clin Lab Anal. 2016 Mar;30(2):155-60. doi: 10.1002/jcla.21829. Epub 2014 Dec 26.

Bearup AJ. The coccidia of carnivores in Sydney. Austral. Vet. J. 1954; 30, 185-6.

Beckmann C, Heininger U, Marti H, Hirsch HH. Gastrointestinal pathogens detected by multiplex nucleic acid amplification testing in stools of pediatric patients and patients returning from the tropics. Infection. 2014 Dec;42(6):961-70.

Bednarska M, Bajer A, Sinski E, Girouard AS, Tamang L, Graczyk TK. Fluorescent in situ hybridization as a tool to retrospectively identify *Cryptosporidium parvum* and Giardia lamblia in samples from terrestrial mammalian wildlife. Parasitol Res. 2007 Feb;100(3):455-60. Epub 2006 Sep 27.

Ben Abda I, Essid R, Mellouli F, Aoun K, Bejaoui M, Bouratbine A. [*Cryptosporidium* infection in patients with major histocompatibility complex class II deficiency syndrome in Tunisia: description of five cases]. Arch Pediatr. 2011 Sep;18(9):939-44. doi: 10.1016/j.arcped.2011.06.015. Epub 2011 Aug 3.

Ben Rashed MB, Abulhassan M, Tabit A, Hawas A. Demographic features of intestinal parasitic infection among Libyan children. Jamahiriya Med J. 2006;6:138–40.

Ben Salem-Ben Nejma I, Hassine Zaafrane M, Hassine F, Sdiri-Loulizi K, Ben Said M, Aouni M, Mzoughi R. Etiology of Acute Diarrhea in Tunisian Children with Emphasis on Diarrheagenic Escherichia coli: Prevalence and Identification of E. coli Virulence Markers. Iran J Public Health. 2014 Jul;43(7):947-60.

Bentley C, Laubach H, Spalter J, Ginter E, Jensen L. Relationship of cryptosporidiosis to abdominal pain and diarrhea in Mayan Indians. Rev Inst Med Trop Sao Paulo. 2004 Jul-Aug;46(4):235-7. Epub 2004 Sep 3.

Bera P, Das S, Saha R, Ramachandran VG, Shah D. *Cryptosporidium* in children with diarrhea: a hospital-based Study. Indian Pediatr. 2014 Nov;51(11):906-8.

Berg IE, Peterson AC, Freeman TP. Ovine Cryptosporidiosis. Journal of American Veterinary Medical Association. 1979; 173: 1586-158

Bergeland ME. Necrotic enteritis in nursing piglets. Proc Am Assoc Vet Lab Diagn. 1977; 20:151–158.

Berger, S. Infectious Diseases of Austria: 2016 Edition. Los Angeles: GIDEON Informatics, 2016. Print.

Berger S. Infectious diseases of the united states. 2017. California, USA: Gideon Informatics Inc. Print.

Berkman DS, Lescano AG, Gilman RH, Lopez SL, Black MM. Effects of stunting, diarrhoeal disease, and parasitic infection during infancy on cognition in late childhood: a follow-up study. Lancet. 2002 Feb 16;359(9306):564-71.

Bern C, Hernandez B, Lopez MB, Arrowood MJ, De Merida AM, Klein RE. The contrasting epidemiology of Cyclospora and *Cryptosporidium* among outpatients in Guatemala. Am J Trop Med Hyg. 2000 Nov-Dec;63(5-6):231-5.

Bern C, Ortega Y, Checkley W, Roberts JM, Lescano AG, Cabrera L, Verastegui M, Black RE, Sterling C, Gilman RH. Epidemiologic differences between cyclosporiasis and cryptosporidiosis in Peruvian children. Emerg Infect Dis. 2002 Jun;8(6):581-5.

Berrilli F, D'Alfonso R, Giangaspero A, Marangi M, Brandonisio O, Kaboré Y, Glé C, Cianfanelli C, Lauro R, Di Cave D. Giardia duodenalis genotypes and *Cryptosporidium* species in humans and domestic animals in Côte d'Ivoire: occurrence and evidence for environmental contamination. Trans R Soc Trop Med Hyg. 2012 Mar;106(3):191-5. doi: 10.1016/j.trstmh.2011.12.005. Epub 2012 Jan 20.

Bhadauria D, Goel A, Kaul A, Sharma RK, Gupta A, Ruhela V, Gupta A, Vardhan H, Prasad N. *Cryptosporidium* infection after renal transplantation in an endemic area. Transpl Infect Dis. 2015 Feb;17(1):48-55. doi: 10.1111/tid.12336. Epub 2015 Jan 12.

Bhalchandra S, Ludington J, Coppens I, Ward HD. Identification and characterization of *Cryptosporidium parvum* Clec, a novel C-type lectin domain-containingmucin-like glycoprotein. Infect Immun. 2013 Sep;81(9):3356-65. doi: 10.1128/IAI.00436-13. Epub 2013 Jul 1.

Bhatia V, Kapoor A, Sibal A. Prevalence of *Cryptosporidium* in immunocompetent Indian children with recurrent abdominal pain. Trop Gastroenterol. 2014 Oct-Dec;35(4):277.

Bhattacharyya MK, Kumar N. Identification and molecular characterisation of DNA damaging agent induced expression of Plasmodium falciparum recombination protein PfRad51. Int J Parasitol. 2003 Oct;33(12):1385-92.

Bhattacharyya MK, Norris DE, Kumar N. Molecular players of homologous recombination in protozoan parasites: implications for generating antigenic variation. Infect Genet Evol. 2004 Jun;4(2):91-8.

Bilenko N, Levy A, Dagan R, Deckelbaum RJ, El-On Y, Fraser D. Does co-infection with Giardia lamblia modulate the clinical characteristics of enteric infections in young children? Eur J Epidemiol. 2004;19(9):877-83.

Bilenko N, Ghosh R, Levy A, Deckelbaum RJ, Fraser D. Partial breastfeeding protects Bedouin infants from infection and morbidity: prospective cohort study. Asia Pac J Clin Nutr. 2008;17(2):243-9.

Binka E, Vermund SH, Armah GE. Rotavirus diarrhea among children less than 5 years of age in urban Ghana. Pediatr Infect Dis J. 2011 Aug;30(8):716-8. doi: 10.1097/INF.0b013e318223bd85.

Bird RG, Smith MD. Cryptosporidiosis in man: parasite life cycle and fine structural pathology. J Pathol. 1980 Nov;132(3):217-33.

Bishop PJ, Neumann G. The history of the Ziehl-Neelsen stain. Tubercle. 1970 Jun;51(2):196-206.

Bishop R, Gobright E, Nene V, Morzaria S, Musoke A, Sohanpal B. Polymorphic open reading frames encoding secretory proteins are located less than 3 kilobases from Theileria parva telomeres. Mol Biochem Parasitol. 2000 Oct;110(2):359-71.

Blanco MA, Iborra A, Vargas A, Nsie E, Mbá L, Fuentes I. Molecular characterization of *Cryptosporidium* isolates from humans in Equatorial Guinea. Trans R Soc Trop Med Hyg. 2009 Dec;103(12):1282-4. doi: 10.1016/j.trstmh.2009.02.016. Epub 2009 Apr 10.

Blanco MA, Montoya A, Iborra A, Fuentes I. Identification of *Cryptosporidium* subtype isolates from HIV-seropositive patients in Equatorial Guinea. Trans R Soc Trop Med Hyg. 2014 Sep;108(9):594-6. doi: 10.1093/trstmh/tru108. Epub 2014 Jul 4.

Boaitey YA, Nkrumah B, Idriss A, Tay SC. Gastrointestinal and urinary tract pathogenic infections among HIV seropositive patients at the Komfo Anokye Teaching Hospital in Ghana. BMC Res Notes. 2012 Aug 21;5:454. doi: 10.1186/1756-0500-5-454.

Bodhidatta L, McDaniel P, Sornsakrin S, Srijan A, Serichantalergs O, Mason CJ. Case-control study of diarrheal disease etiology in a remote rural area in Western Thailand. Am J Trop Med Hyg. 2010 Nov;83(5):1106-9. doi: 10.4269/ajtmh.2010.10-0367.

Bodhidatta L, Wongstitwilairoong B, Khantapura P, Shrestha SK, Wangchuk S, Raj Pant A, Ram Shrestha B, Regmi S, Ratna Dhakwa J, Tsheten, Mason1 CJ, and Swierczewski BE. Cryptosporidiosis: Prevalence in Children in Nepal and Bhutan. J Vet Med Res. 2016; 3(4): 1059

Bonnin A, Fourmaux MN, Dubremetz JF, Nelson RG, Gobet P, Harly G, Buisson M, Puygauthier-Toubas D, Gabriel-Pospisil G, Naciri M, Camerlynck P. Genotyping human and bovine isolates of *Cryptosporidium parvum* by polymerase chain reaction-restriction fragment length polymorphism analysis of a repetitive DNA sequence. FEMS Microbiol Lett. 1996 Apr 1;137(2-3):207-11.

Börekçi G, Uzel A. [The determination of intestinal parasites, physical growth and hygiene behaviors of children in the Mersin City Social Service Child Care Centre]. Turkiye Parazitol Derg. 2009;33(1):63-72.

Borges JD, Alarcón RS, Neto VA, Gakiya E. [Intestinal parasitosis in Indians of the Mapuera community (Oriximiná, State of Pará, Brazil): high prevalence of Blastocystis hominis and finding of *Cryptosporidium* sp and Cyclospora cayetanensis]. Rev Soc Bras Med Trop. 2009 May-Jun;42(3):348-50.

Borodina ON, Zhukova EV, Kravets ZF. [The detection of the causative agent of cryptosporidiosis in man and animals in Turkmenistan]. Med Parazitol (Mosk). 1994 Jan-Mar;(1):8-11.

Botero JH, Castaño A, Montoya MN, Ocampo NE, Hurtado MI, Lopera MM. A preliminary study of the prevalence of intestinal parasites in immunocompromised patients with and without gastrointestinal manifestations. Rev Inst Med Trop Sao Paulo. 2003 Jul-Aug;45(4):197-200. Epub 2003 Sep 17.

Bouckenooghe AR, Jiang ZD, De La Cabada FJ, Ericsson CD, DuPont HL. Enterotoxigenic Escherichia coli as cause of diarrhea among Mexican adults and US travelers in Mexico. J Travel Med. 2002 May-Jun;9(3):137-40.

Bougiouklis PA, Weissenböck H, Wells A, Miller WA, Palmieri C, Shivaprasad HL. Otitis media associated with *Cryptosporidium baileyi* in a Saker falcon (Falco cherrug). J Comp Pathol. 2013 May;148(4):419-23. doi: 10.1016/j.jcpa.2012.09.005. Epub 2012 Nov 2.

Boulanger MJ, Tonkin ML, Crawford J. Apicomplexan parasite adhesins: novel strategies for targeting host cell carbohydrates. Curr Opin Struct Biol. 2010 Oct;20(5):551-9. doi: 10.1016/j.sbi.2010.08.003. Epub 2010 Sep 16.

Bouyou-Akotet MK, Owono-Medang M, Moussavou-Boussougou MN, Mamfoumbi MM, Mintsa-Nguema R, Mawili-Mboumba DP, Kombila M. Low sensitivity of the ImmunocardSTAT® Crypto/Giardia Rapid Assay test for the detection of Giardia and *Cryptosporidium* in fecal samples from children living in Libreville, Central Africa. J Parasit Dis. 2016 Dec;40(4):1179-1183. Epub 2015 Feb 14.

Bouzid M, Tyler KM, Christen R, Chalmers RM, Elwin K, Hunter PR. Multi-locus analysis of human infective *Cryptosporidium* species and subtypes using ten novel genetic loci. BMC Microbiol. 2010 Aug 9;10:213. doi: 10.1186/1471-2180-10-213.

Bouzid M, Hunter PR, Chalmers RM, Tyler KM. *Cryptosporidium* pathogenicity and virulence. Clin Microbiol Rev. 2013 Jan;26(1):115-34. doi: 10.1128/CMR.00076-12.

Bouzid M, Hunter PR, McDonald V, Elwin K, Chalmers RM, Tyler KM. A new heterogeneous family of telomerically encoded *Cryptosporidium* proteins. Evol Appl. 2013 Feb;6(2):207-17. doi: 10.1111/j.1752-4571.2012.00277.x. Epub 2012 Jun 14.

Branco N, Leal DA, Franco RM. A parasitological survey of natural water springs and inhabitants of a tourist city in southeastern Brazil. Vector Borne Zoonotic Dis. 2012 May;12(5):410-7. doi: 10.1089/vbz.2011.0679. Epub 2012 Jan 4.

Brannan DK, Greenfield RA, Owen WL, Welch DF, Kuhls TL. Protozoal colonization of the intestinal tract in institutionalized Romanian children. Clin Infect Dis. 1996 Mar;22(3):456-61.

Brantley RK, Williams KR, Silva TM, Sistrom M, Thielman NM, Ward H, Lima AA, Guerrant RL. AIDS-associated diarrhea and wasting in Northeast Brazil is associated with subtherapeutic plasma levels of antiretroviral medications and with both bovine and human subtypes of *Cryptosporidium parvum*. Braz J Infect Dis. 2003 Feb;7(1):16-22. Epub 2003 Dec 2.

Breurec S, Vanel N, Bata P, Chartier L, Farra A, Favennec L, Franck T, Giles-Vernick T, Gody JC, Luong Nguyen LB, Onambélé M, Rafaï C, Razakandrainibe R, Tondeur L, Tricou V, Sansonetti P, Vray M. Etiology and Epidemiology of Diarrhea in Hospitalized Children from Low Income Country: A Matched Case-Control Study in Central African Republic. PLoS Negl Trop Dis. 2016 Jan 5;10(1):e0004283. doi: 10.1371/journal.pntd.0004283. eCollection 2016.

Brink AK, Mahé C, Watera C, Lugada E, Gilks C, Whitworth J, French N. Diarrhea, CD4 counts and enteric infections in a community-based cohort of HIV-infected adults in Uganda. J Infect. 2002 Aug;45(2):99-106.

Brownstein DG, Strandberg JD, Montali RJ, Bush M, Fortner J. *Cryptosporidium* in snakes with hypertrophic gastritis. Vet Pathol. 1977 Nov;14(6):606-17.

Bruijnesteijn van Coppenraet LE, Wallinga JA, Ruijs GJ, Bruins MJ, Verweij JJ. Parasitological diagnosis combining an internally controlled real-time PCR assay for the detection of four protozoa in stool samples with a testing algorithm for microscopy. Clin Microbiol Infect. 2009 Sep;15(9):869-74. doi: 10.1111/j.1469-0691.2009.02894.x. Epub 2009 Jul 14.

Bruijnesteijn van Coppenraet LE, Dullaert-de Boer M, Ruijs GJ, van der Reijden WA, van der Zanden AG, Weel JF, Schuurs TA. Case-control comparison of bacterial and protozoan microorganisms associated with gastroenteritis: application of molecular detection. Clin Microbiol Infect. 2015 Jun;21(6):592.e9-19. doi: 10.1016/j.cmi.2015.02.007. Epub 2015 Feb 18.

Budu-Amoako E, Greenwood SJ, Dixon BR, Sweet L, Ang L, Barkema HW, McClure JT. Molecular epidemiology of *Cryptosporidium* and Giardia in humans on Prince Edward Island, Canada: evidence of zoonotic transmission from cattle. Zoonoses Public Health. 2012 Sep;59(6):424-33. doi: 10.1111/j.1863-2378.2012.01474.x. Epub 2012 Mar 6.

Bushen OY, Kohli A, Pinkerton RC, Dupnik K, Newman RD, Sears CL, Fayer R, Lima AA, Guerrant RL. Heavy cryptosporidial infections in children in northeast Brazil: comparison of *Cryptosporidium hominis* and *Cryptosporidium parvum*. Trans R Soc Trop Med Hyg. 2007 Apr;101(4):378-84. Epub 2006 Aug 24.

Butler, J. Forensic DNA typing biology, technology, and genetics of STR markers. 2nd ed. England: ACADEMIC Press INC. 2005. Print.

Büttner D, Bonas U. Regulation and secretion of Xanthomonas virulence factors. FEMS Microbiol Rev. 2010 Mar;34(2):107-33. doi: 10.1111/j.1574-6976.2009.00192.x. Epub 2009 Oct 13.

Büyükbaba Boral O, Uysal H, Alan S, Nazlican O. [Investigation of intestinal parasites in AIDS patients]. Mikrobiyol Bul. 2004 Jan-Apr;38(1-2):121-8.

Cacciò SM, Homan W, Camilli R, Traldi G, Kortbeek T, Pozio E. A microsatellite marker reveals population heterogeneity within human and animal genotypes of *Cryptosporidium parvum*. Parasitology. 2000 Mar;120 (Pt 3):237-44.

Cacciò SM, Thompson RC, McLauchlin J, Smith HV. Unravelling *Cryptosporidium* and Giardia epidemiology. Trends Parasitol. 2005 Sep;21(9):430-7.

Cacciò SM, Sannella AR, Mariano V, Valentini S, Berti F, Tosini F, Pozio E. A rare *Cryptosporidium parvum* genotype associated with infection of lambs and zoonotic transmission in Italy. Vet Parasitol. 2013 Jan 16;191(1-2):128-31. doi: 10.1016/j.vetpar.2012.08.010. Epub 2012 Aug 20.

Cacciò SM, Chalmers RM. Human cryptosporidiosis in Europe. Clin Microbiol Infect. 2016 Jun;22(6):471-80. doi: 10.1016/j.cmi.2016.04.021. Epub 2016 May 10.

Cai J, Collins MD, McDonald V, Thompson DE. PCR cloning and nucleotide sequence determination of the 18S rRNA genes and internal transcribed spacer 1 of the protozoan parasites *Cryptosporidium parvum* and *Cryptosporidium muris*. Biochim Biophys Acta. 1992 Jul 15;1131(3):317-20.

Cai X, Lancto CA, Abrahamsen MS, Zhu G. Intron-containing beta-tubulin transcripts in *Cryptosporidium parvum* cultured in vitro. Microbiology. 2004 May;150(Pt 5):1191-5.

Calderaro A, Montecchini S, Gorrini C, Dettori G, Chezzi C. Similar diagnostic performances of antigen detection and nucleic acid detection of *Cryptosporidium spp*. in a low-prevalence setting. Diagn Microbiol Infect Dis. 2011 May;70(1):72-7. doi: 10.1016/j.diagmicrobio.2010.11.017.

Cama VA, Bern C, Sulaiman IM, Gilman RH, Ticona E, Vivar A, Kawai V, Vargas D, Zhou L, Xiao L. *Cryptosporidium* species and genotypes in HIV-positive patients in Lima, Peru. J Eukaryot Microbiol. 2003;50 Suppl:531-3.

Cama VA, Ross JM, Crawford S, Kawai V, Chavez-Valdez R, Vargas D, Vivar A, Ticona E, Navincopa M, Williamson J, Ortega Y, Gilman RH, Bern C, Xiao L. Differences in clinical manifestations among *Cryptosporidium* species and subtypes in HIV-infected persons. J Infect Dis. 2007 Sep 1;196(5):684-91. Epub 2007 Jul 13.

Cama VA, Bern C, Roberts J, Cabrera L, Sterling CR, Ortega Y, Gilman RH, Xiao L. *Cryptosporidium* species and subtypes and clinical manifestations in children, Peru. Emerg Infect Dis [serial on the Internet]. 2008 Oct [21/08/205]. Available from http://wwwnc.cdc.gov/eid/article/14/10/07-1273

Campbell PN, Current WL. Demonstration of serum antibodies to *Cryptosporidium* sp. in normal and immunodeficient humans with confirmed infections. J Clin Microbiol. 1983 Jul;18(1):165-9.

Cancrini G, Bartoloni A, Nuñez L, Paradisi F. Intestinal parasites in the Camiri, Gutierrez and Boyuibe areas, Santa Cruz Department, Bolivia. Parassitologia. 1988 May-Dec;30(2-3):263-9.

Cárcamo C, Hooton T, Wener MH, Weiss NS, Gilman R, Arevalo J, Carrasco J, Seas C, Caballero M, Holmes KK. Etiologies and manifestations of persistent diarrhea in adults with HIV-1 infection: a case-control study in Lima, Peru. J Infect Dis. 2005 Jan 1;191(1):11-9. Epub 2004 Dec 1.

Cardenas VM, Mena KD, Ortiz M, Karri S, Variyam E, Behravesh CB, Snowden KF, Flisser A, Bristol JR, Mayberry LF, Ortega YR, Fukuda Y, Campos A, Graham DY. Hyperendemic H. pylori and tapeworm infections in a U.S.-Mexico border population. Public Health Rep. 2010 May-Jun;125(3):441-7.

Cardona GA, Carabin H, Goñi P, Arriola L, Robinson G, Fernández-Crespo JC, Clavel A, Chalmers RM, Carmena D. Identification and molecular characterization of *Cryptosporidium* and Giardia in children and cattle populations from the province of Álava, North of Spain. Sci Total Environ. 2011 Dec 15;412-413:101-8. doi: 10.1016/j.scitotenv.2011.09.076. Epub 2011 Oct 26.

Cardoso LV, Galisteu KJ, Schiesari Júnior A, Chahla LA, Canille RM, Belloto MV, Franco C, Maia IL, Rossit AR, Machado RL. Enteric parasites in HIV-1/AIDS-infected patients from a Northwestern São Paulo reference unit in the highly active antiretroviral therapy era. Rev Soc Bras Med Trop. 2011 Nov-Dec;44(6):665-9. Epub 2011 Nov 21.

Carme B, M'Pele P, Mbitsi A, Kissila AM, Aya GM, Mouanga-Yidika G, Mboussa J, Itoua-Ngaporo A. [Opportunistic parasitic diseases and mycoses in AIDS. Their frequencies in Brazzaville (Congo)]. Bull Soc Pathol Exot Filiales. 1988;81(3):311-6.

Carneiro M, Ferrand N, Nachman MW. Recombination and Speciation: Loci Near Centromeres Are More Differentiated Than Loci Near Telomeres Between Subspecies of the European Rabbit (Oryctolagus cuniculus) Genetics. 2009 Feb;181(2):593-606. doi: 10.1534/genetics.108.096826. Epub 2008 Nov 17.

Caruana SR, Kelly HA, Ngeow JY, Ryan NJ, Bennett CM, Chea L, Nuon S, Bak N, Skull SA, Biggs BA. Undiagnosed and potentially lethal parasite infections among immigrants and refugees in Australia. J Travel Med. 2006 Jul-Aug;13(4):233-9.

Carvalho-Costa FA, Gonçalves AQ, Lassance SL, de Albuquerque CP, Leite JP, Bóia MN. Detection of *Cryptosporidium spp* and other intestinal parasites in children with acute diarrhea and severe dehydration in Rio de Janeiro. Rev Soc Bras Med Trop. 2007 May-Jun;40(3):346-8.

Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. 2005 June 23. 21 (16): 3422-3423.doi: 10.1093/bioinformatics/bti553

Casemore, DP, Sands, RL, Curry, A. *Cryptosporidium* species a "new" human pathogen. J Clin Pathol. 1985 Dec;38(12):1321-36.

Casemore DP. The antibody response to *Cryptosporidium*: development of a serological test and its use in a study of immunologically normal persons. J Infect. 1987 Mar;14(2):125-34.

Cazorla D, Acosta ME, Acosta ME, Morales P. [Clinical and epidemiological study of intestinal coccidioses in a rural population of a semiarid region from Falcon state, Venezuela]. Invest Clin. 2012 Sep;53(3):273-88.

Centers for Disease Control and Prevention (CDC). Cryptosporidiosis outbreak at a summer camp--North Carolina, 2009. MMWR Morb Mortal Wkly Rep. 2011 Jul 15;60(27):918-22.

Certad G, Arenas-Pinto A, Pocaterra L, Ferrara G, Castro J, Bello A, Núñez L. Cryptosporidiosis in HIV-infected Venezuelan adults is strongly associated with acute or chronic diarrhea. Am J Trop Med Hyg. 2005 Jul;73(1):54-7.

Certad G, Ngouanesavanh T, Hernan A, Rojas E, Contreras R, Pocaterra L, Nuñez L, Dei-Cas E, Guyot K. First molecular data on cryptosporidiosis in Venezuela. J Eukaryot Microbiol. 2006;53 Suppl 1:S30-2.

Cevallos AM, Bhat N, Verdon R, Hamer DH, Stein B, Tzipori S, Pereira ME, Keusch GT, Ward HD. Mediation of *Cryptosporidium parvum* infection in vitro by mucin-like glycoproteins defined by a neutralizing monoclonal antibody. Infect Immun. 2000 Sep;68(9):5167-75.

Chacín-Bonilla L, Sánchez-Chávez Y. Intestinal parasitic infections, with a special emphasis on cryptosporidiosis, in Amerindians from western Venezuela. Am J Trop Med Hyg. 2000 Mar;62(3):347-52.

Chacín-Bonilla L, Barrios F, Sanchez Y. Environmental risk factors for *Cryptosporidium* infection in an island from Western Venezuela. Mem Inst Oswaldo Cruz. 2008 Feb;103(1):45-9. Epub 2008 Jan 31.

Chai JY, Kim NY, Guk SM, Park YK, Seo M, Han ET, Lee SH. High prevalence and seasonality of cryptosporidiosis in a small rural village occupied predominantly by aged people in the Republic of Korea. Am J Trop Med Hyg. 2001 Nov;65(5):518-22.

Chain PS, Carniel E, Larimer FW, Lamerdin J, Stoutland PO, Regala WM, Georgescu AM, Vergez LM, Land ML, Motin VL, Brubaker RR, Fowler J, Hinnebusch J, Marceau M, Medigue C, Simonet M, Chenal-Francisque V, Souza B, Dacheux D, Elliott JM, Derbise A, Hauser LJ, Garcia E. Insights into the evolution of Yersinia pestis through whole-genome comparison with Yersinia pseudotuberculosis. Proc Natl Acad Sci U S A. 2004 Sep 21;101(38):13826-31. Epub 2004 Sep 9.

Chalmers RM, Ferguson C, Cacciò S, Gasser RB, Abs EL-Osta YG, Heijnen L, Xiao L, Elwin K, Hadfield S, Sinclair M, Stevens M. Direct comparison of selected methods for genetic categorisation of *Cryptosporidium parvum* and *Cryptosporidium hominis* species. Int J Parasitol. 2005 Apr 1;35(4):397-410.

Chalmers RM, Hadfield SJ, Jackson CJ, Elwin K, Xiao L, Hunter P. Geographic linkage and variation in *Cryptosporidium hominis*. Emerg Infect Dis. 2008 Mar;14(3):496-8. doi: 10.3201/eid1403.071320.

Chalmers RM, Elwin K, Thomas AL, Guy EC, Mason B. Long-term *Cryptosporidium* typing reveals the aetiology and species-specific epidemiology of human cryptosporidiosis in England and Wales, 2000 to 2003. Euro Surveill. 2009 Jan 15;14(2). pii: 19086.

Chalmers RM, Elwin K, Hadfield SJ, Robinson G. Sporadic human cryptosporidiosis caused by *Cryptosporidium cuniculus*, United Kingdom, 2007-2008. Emerg Infect Dis. 2011 Mar;17(3):536-8. doi: 10.3201/eid1703.100410.

Chalmers RM, Smith R, Elwin K, Clifton-Hadley FA, Giles M. Epidemiology of anthroponotic and zoonotic human cryptosporidiosis in England and Wales, 2004-2006. Epidemiol Infect. 2011 May;139(5):700-12. doi: 10.1017/S0950268810001688. Epub 2010 Jul 12.

Chalmers RM, Elwin K, Cheesbrough J, Hadfield SJ, Beeching NJ. Detection of IgG antibodies in sera from patients with *Cryptosporidium parvum* and *Cryptosporidium hominis*. J Infect. 2013 Sep;67(3):231-7. doi: 10.1016/j.jinf.2013.04.019. Epub 2013 May 2.

Chamberlain, C. J. A Phylogenetic Study of Cycads. Proc Natl Acad Sci U S A. 1915 Feb;1(2):86-90.

Chappell CL, Okhuysen PC, Langer-Curry RC, Akiyoshi DE, Widmer G, Tzipori S. Cryptosporidium meleagridis: infectivity in healthy adult volunteers. Am J Trop Med Hyg. 2011 Aug;85(2):238-42. doi: 10.4269/ajtmh.2011.10-0664.

Charles M, Delva GG, Boutin J, Severe K, Peck M, Mabou MM, Wright PF, Pape JW. Importance of cholera and other etiologies of acute diarrhea in post-earthquake Port-au-Prince, Haiti. Am J Trop Med Hyg. 2014 Mar;90(3):511-7. doi: 10.4269/ajtmh.13-0514. Epub 2014 Jan 20.

Chau ML, Hartantyo SH, Yap M, Kang JS, Aung KT, Gutiérrez RA, Ng LC, Tam CC, Barkham T. Diarrheagenic pathogens in adults attending a hospital in Singapore. BMC Infect Dis. 2016 Jan 28;16:32. doi: 10.1186/s12879-016-1354-0.

Chauve CM, Reynaud MC. & Gounel JM. Description d'Eimeria mulardi N. sp. chez le canard mulard. Etude de la phase endogene de son cycle evolutif avec mise en evidence du developpement intranucleaire. Parasite. 1994; 1, 15-22.

Chavali S, Mahajan A, Tabassum R, Maiti S, Bharadwaj D. Oligonucleotide properties determination and primer designing: a critical examination of predictions. Bioinformatics. 2005 Oct 15;21(20):3918-25. Epub 2005 Aug 16.

Chen XM, Keithly JS, Paya CV, LaRusso NF. Cryptosporidiosis. N Engl J Med. 2002 May 30;346(22):1723-31.

Cheng YJ, Lin WC, Liu CE, Ji DD. First microbiologically confirmed case of cryptosporidiosis due to *Cryptosporidium hominis* gp60-If sub-genotype in a patient with HIV-1 infection in Taiwan. J Formos Med Assoc. 2014 Apr;113(4):264-5. doi: 10.1016/j.jfma.2012.02.003. Epub 2012 Aug 4.

Cheun HI, Choi TK, Chung GT, Cho SH, Lee YH, Kimata I, Kim TS. Genotypic characterization of *Cryptosporidium* oocysts isolated from healthy people in three different counties of Korea. J Vet Med Sci. 2007 Oct;69(10):1099-101.

Cheun HI, Cho SH, Lim YY, Lee BC, Kim JY, Ju JW, Na BK, Kimata I, Yu JR, Kim TS. *Cryptosporidium parvum* in Korea: prevalence in individuals residing in three major river valleys and genetic characteristics of the isolates. J Vet Med Sci. 2010 Feb;72(2):167-72. Epub 2009 Nov 27.

Cheun HI, Cho SH, Lee JH, Lim YY, Jeon JH, Yu JR, Kim TS, Lee WJ, Cho SH, Lee DY, Park MS, Jeong HS, Chen DS, Ji YM, Kwon MH. Infection status of hospitalized diarrheal patients with gastrointestinal protozoa, bacteria, and viruses in the Republic of Korea. Korean J Parasitol. 2010 Jun;48(2):113-20. doi: 10.3347/kjp.2010.48.2.113. Epub 2010 Jun 17.

Cheun HI, Kim K, Yoon S, Lee WJ, Park WY, Sim S, Yu JR. *Cryptosporidium hominis* infection diagnosed by real-time PCR-RFLP. Korean J Parasitol. 2013 Jun;51(3):353-5. doi: 10.3347/kjp.2013.51.3.353. Epub 2013 Jun 30.

Chhin S, Harwell JI, Bell JD, Rozycki G, Ellman T, Barnett JM, Ward H, Reinert SE, Pugatch D. Etiology of chronic diarrhea in antiretroviral-naive patients with HIV infection admitted to Norodom Sihanouk Hospital, Phnom Penh, Cambodia. Clin Infect Dis. 2006 Oct 1;43(7):925-32. Epub 2006 Aug 24.

China B, Goffaux F. Secretion of virulence factors by Escherichia coli. Vet Res. 1999 Mar-Jun;30(2-3):181-202.

Chmelík V, Ditrich O, Trnovcová R, Gutvirth J. Clinical features of diarrhoea in children caused by *Cryptosporidium parvum*. Folia Parasitol (Praha). 1998;45(2):170-2.

Chmura, SE. Development of an in Situ PCR Assay to Detect *Cryptosporidium* Oocysts in Environmental Samples. 2010. MSc Dissertation. University at Albany, Department of Biomedical Science.

Chokephaibulkit K, Wanachiwanawin D, Tosasuk K, Pavitpok J, Vanprapar N, Chearskul S. Intestinal parasitic infections among human immunodeficiency virus-infected and -uninfected children hospitalized with diarrhea in Bangkok, Thailand. Southeast Asian J Trop Med Public Health. 2001 Dec;32(4):770-5.

Chopra RD, Dworkin MS. Descriptive epidemiology of enteric disease in Chennai, India. Epidemiol Infect. 2013 May;141(5):953-7. doi: 10.1017/S0950268812001409. Epub 2012 Jul 17.

Chu XN, Hu JY, Guo HL, Tan XL. Rapid method for viable *Cryptosporidium parvum* oocysts detection. Water Science and Technology. December 2011; 11 (6) 737-744; DOI: 10.2166/ws.2011.109

Çiçek M, Körkoca H, Gül A. [Investigation of *Cryptosporidium* sp. in workers of the Van municipality slaughterhouse and in slaughtered animals]. Turkiye Parazitol Derg. 2008;32(1):8-11.

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Research. 2016;44(Database issue):D67-D72. doi:10.1093/nar/gkv1276.

Cockrell BY, Valerio MG, Garner FM. Cryptosporidiosis in the intestine of Rhesus monkeys (Macaca mulatta). Laboratory Animal Science. 1974; 24: 881-887.

Cohen JD, Ruhlig L, Jayich SA, Tong MJ, Lechago J, Snape WJ Jr. *Cryptosporidium* in acquired immunodeficiency syndrome. Dig Dis Sci. 1984 Aug;29(8):773-7.

Соколова О.И., Демьянов А.В., Дедов В.А., Соколова Ю.Я., Яковлев А.А. КРИПТОСПОРИДИОЗ У ВИЧ-ИНФИЦИРОВАННЫХ ПАЦИЕНТОВ: ДИАГНОСТИКА И КЛИНИКО-ЛАБОРАТОРНАЯ ХАРАКТЕРИСТИКА. Журнал инфектологии. 2012;4(2):26-32.

Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honoré N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG. Massive gene decay in the leprosy bacillus. Nature. 2001 Feb 22;409(6823):1007-11.

Comparative genomic analysis reveals occurrence of genetic recombination in virulent *Cryptosporidium hominis* subtypes and telomeric gene duplications in *Cryptosporidium parvum*. BMC Genomics. 2015 Apr 18;16:320. doi: 10.1186/s12864-015-1517-1.

Connelly L, Craig BH, Jones B, Alexander CL. Genetic diversity of *Cryptosporidium spp*. within a remote population of Soay Sheep on St. Kilda Islands, Scotland. Appl Environ Microbiol. 2013 Apr;79(7):2240-6. doi: 10.1128/AEM.02823-12. Epub 2013 Jan 25.

Conway C, Proudfoot C, Burton P, Barry JD, McCulloch R. Two pathways of homologous recombination in Trypanosoma brucei. Mol Microbiol. 2002 Sep;45(6):1687-700.

Cooke ML, Goddard EA, Brown RA. Endoscopy findings in HIV-infected children from sub-Saharan Africa. J Trop Pediatr. 2009 Aug;55(4):238-43. doi: 10.1093/tropej/fmn114. Epub 2009 Jan 15.

Cooper MA, Adam RD, Worobey M, Sterling CR. Population genetics provides evidence for recombination in Giardia. Curr Biol. 2007 Nov 20;17(22):1984-8. Epub 2007 Nov 1.

Cordova Paz Soldan O, Vargas Vásquez F, Gonzalez Varas A, Peréz Cordón G, Velasco Soto JR, Sánchez-Moreno M, Rodríguez Gonzalez I, Rosales Lombardo MJ. Intestinal parasitism in Peruvian children and molecular characterization of *Cryptosporidium* species. Parasitol Res. 2006 May;98(6):576-81. Epub 2006 Jan 24.

Costa LB, JohnBull EA, Reeves JT, Sevilleja JE, Freire RS, Hoffman PS, Lima AA, Oriá RB, Roche JK, Guerrant RL, Warren CA. *Cryptosporidium*-malnutrition interactions: mucosal disruption, cytokines, and TLR signaling in a weaned murine model. J Parasitol. 2011 Dec;97(6):1113-20. doi: 10.1645/GE-2848.1. Epub 2011 Jun 28.

Coupe S, Sarfati C, Hamane S, Derouin F. Detection of *Cryptosporidium* and identification to the species level by nested PCR and restriction fragment length polymorphism. J Clin Microbiol. 2005 Mar;43(3):1017-23.

Cranendonk RJ, Kodde CJ, Chipeta D, Zijlstra EE, Sluiters JF. *Cryptosporidium parvum* and Isospora belli infections among patients with and without diarrhoea. East Afr Med J. 2003 Aug;80(8):398-401.

Creek TL, Kim A, Lu L, Bowen A, Masunge J, Arvelo W, Smit M, Mach O, Legwaila K, Motswere C, Zaks L, Finkbeiner T, Povinelli L, Maruping M, Ngwaru G, Tebele G, Bopp C, Puhr N, Johnston SP, Dasilva AJ, Bern C, Beard RS, Davis MK. Hospitalization and mortality among primarily nonbreastfed children during a large outbreak of diarrhea and malnutrition in Botswana, 2006.J Acquir Immune Defic Syndr. 2010 Jan;53(1):14-9. doi: 10.1097/QAI.0b013e3181bdf676.

Cuomo CA, Güldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, Decaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes

HW, Mitterbauer R, Muehlbauer G, Münsterkötter M, Nelson D, O'donnell K, Ouellet T, Qi W, Quesneville H, Roncero MI, Seong KY, Tetko IV, Urban M, Waalwijk C, Ward TJ, Yao J, Birren BW, Kistler HC. The Fusarium graminearum genome reveals a link between localized polymorphism and pathogen specialization. Science. 2007 Sep 7;317(5843):1400-2.

Currà C, Gessmann R, Pace T, Picci L, Peruzzi G, Varamogianni-Mamatsi V, Spanos L, Garcia CR, Spaccapelo R, Ponzi M, Siden-Kiamos I. Release of Plasmodium sporozoites requires proteins with histone-fold dimerization domains. Nat Commun. 2016 Dec 16;7:13846. doi: 10.1038/ncomms13846.

Current WL, Reese NC, Ernst JV, Bailey WS, Heyman MB, Weinstein WM. Human cryptosporidiosis in immunocompetent and immunodeficient persons. Studies of an outbreak and experimental transmission. N Engl J Med. 1983 May 26;308(21):1252-7.

Current WL, Haynes TB. Complete development of *Cryptosporidium* in cell culture. Science. 1984 May 11;224(4649):603-5.

Current WL. *Cryptosporidium*: its biology and potential for environmental transmission, Crzf. Rev. Environ. Control. 1986; 17, 21.

Dalgleish, Angus, and Robin Weiss. HIV and the new viruses. 2nd ed. San Diego, CA: Academic Press, 1999. Print.

Daniels ME, Shrivastava A, Smith WA, Sahu P, Odagiri M, Misra PR, Panigrahi P, Suar M, Clasen T, Jenkins MW. *Cryptosporidium* and Giardia in Humans, Domestic Animals, and Village Water Sources in Rural India.Am J Trop Med Hyg. 2015 Sep;93(3):596-600. doi: 10.4269/ajtmh.15-0111. Epub 2015 Jun 29.

Daryani A, Sharif M, Meigouni M, Mahmoudi FB, Rafiei A, Gholami Sh, Khalilian A, Gohardehi Sh, Mirabi AM. Prevalence of intestinal parasites and profile of CD4+ counts in HIV+/AIDS people in north of Iran, 2007-2008. Pak J Biol Sci. 2009 Sep 15;12(18):1277-81.

Dash M, Padhi S, Panda P, Parida B. Intestinal protozoans in adults with diarrhea. N Am J Med Sci. 2013 Dec;5(12):707-12. doi: 10.4103/1947-2714.123261.

Davies AP, Campbell B, Evans MR, Bone A, Roche A, Chalmers RM. Asymptomatic carriage of protozoan parasites in children in day care centers in the United Kingdom. Pediatr Infect Dis J. 2009 Sep;28(9):838-40. doi: 10.1097/INF.0b013e31819d646d.

Davies AP, Slatter M, Gennery AR, Robinson G, Crouch N, Elwin K, Hadfield SJ, Cant AJ, Davies EG, Chalmers RM. Prevalence of *Cryptosporidium* Carriage and Disease in Children with Primary Immune-Deficiencies Undergoing Hematopoietic Stem Cell Transplant in Northern Europe. Pediatr Infect Dis J. 2016 Dec 27. doi: 10.1097/INF.0000000000001517. [Epub ahead of print]

de Buck E, Anné J, Lammertyn E. The role of protein secretion systems in the virulence of the intracellular pathogen Legionella pneumophila. Microbiology. 2007 Dec;153(Pt 12):3948-53.

de Carvalho TB, de Carvalho LR, Mascarini LM. Occurrence of enteroparasites in day care centers in Botucatu (São Paulo State, Brazil) with emphasis on *Cryptosporidium* sp., Giardia duodenalis and Enterobius vermicularis. Rev Inst Med Trop Sao Paulo. 2006 Sep-Oct;48(5):269-73.

de Koning-Ward TF, Gilson PR, Crabb BS. Advances in molecular genetic systems in malaria. Nat Rev Microbiol. 2015 Jun;13(6):373-87. doi: 10.1038/nrmicro3450.

de Lucio A, Merino FJ, Martínez-Ruiz R, Bailo B, Aguilera M, Fuentes I, Carmena D. Molecular genotyping and sub-genotyping of *Cryptosporidium spp*. isolates from symptomatic individuals attending two major public hospitals in Madrid, Spain. Infect Genet Evol. 2016 Jan;37:49-56. doi: 10.1016/j.meegid.2015.10.026. Epub 2015 Oct 28.

de Lucio A, Amor-Aramendía A, Bailo B, Saugar JM, Anegagrie M, Arroyo A, López-Quintana B, Zewdie D, Ayehubizu Z, Yizengaw E, Abera B, Yimer M, Mulu W, Hailu T, Herrador Z, Fuentes I, Carmena D. Prevalence and Genetic Diversity of Giardia duodenalis and *Cryptosporidium spp*. among School Children in a Rural Area of the Amhara Region, North-West Ethiopia. PLoS One. 2016 Jul 28;11(7):e0159992. doi: 10.1371/journal.pone.0159992. eCollection 2016.

de Mol P, Mukashema S, Bogaerts J, Hemelhof W, Butzler JP. *Cryptosporidium* related to measles diarrhoea in Rwanda. Lancet. 1984 Jul 7;2(8393):42-3.

de Oliveira-Silva MB, de Oliveira LR, Resende JC, Peghini BC, Ramirez LE, Lages-Silva E, Correia D. Seasonal profile and level of CD4+ lymphocytes in the occurrence of cryptosporidiosis and cystoisosporidiosis in HIV/AIDS patients in the Triângulo Mineiro region, Brazil. Rev Soc Bras Med Trop. 2007 Sep-Oct;40(5):512-5.

de Paz VC, Bríinguez MB, Viamonte BV, Suárez CL, Rodríguez AM, Martińex ZA. [Diagnosis of coccidia and microspores in specimens of diarrheal feces from Cuban HIV seropositive patients: first report of microspores in Cuba]. Rev Cubana Med Trop. 2003 Jan-Apr;55(1):14-8.

de Silva NR, de Silva HJ, Jayapani VP. Intestinal parasitoses in the Kandy area, Sri Lanka. Southeast Asian J Trop Med Public Health. 1994 Sep;25(3):469-73.

de Waele V, Van den Broeck F, Huyse T, McGrath G, Higgins I, Speybroeck N, Berzano M, Raleigh P, Mulcahy GM, Murphy TM. Panmictic structure of the *Cryptosporidium parvum* population in Irish calves: influence of prevalence and host movement. Appl Environ Microbiol. 2013 Apr;79(8):2534-41. doi: 10.1128/AEM.03613-12. Epub 2013 Feb 8.

de Wit MA, Koopmans MP, Kortbeek LM, Wannet WJ, Vinjé J, van Leusden F, Bartelds AI, van Duynhoven YT. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. Am J Epidemiol. 2001 Oct 1;154(7):666-74.

Deere D, Vesey G, Milner M, Williams K, Ashbolt N, Veal D. Rapid method for fluorescent in situ ribosomal RNA labelling of *Cryptosporidium parvum*. J Appl Microbiol. 1998 Nov;85(5):807-18.

Dehkordy AB, Rafiei A, Alavi S, Latifi S. Prevalence of *Cryptosporidium* infection in immunocompromised patients, in South-west of iran, 2009-10. Iran J Parasitol. 2010 Dec;5(4):42-7.

Del Chierico F, Onori M, Di Bella S, Bordi E, Petrosillo N, Menichella D, Cacciò SM, Callea F, Putignani L. Cases of cryptosporidiosis co-infections in AIDS patients: a correlation between clinical presentation and GP60 subgenotype lineages from aged formalin-fixed stool samples. Ann Trop Med Parasitol. 2011 Jul;105(5):339-49. doi: 10.1179/1364859411Y.0000000025.

den Hartog J, Rosenbaum L, Wood Z, Burt D, Petri WA Jr. Diagnosis of multiple enteric protozoan infections by enzyme-linked immunosorbent assay in the Guatemalan highlands. Am J Trop Med Hyg. 2013 Jan;88(1):167-71. doi: 10.4269/ajtmh.2012.12-0142. Epub 2012 Nov 26.

Deshpande A, Alexander CL, Coyne M, Brownlie S, Smith-Palmer A, Jones B. Molecular diversity of Scottish *Cryptosporidium hominis* isolates. Epidemiol Infect. 2015 Apr;143(6):1219-24. doi: 10.1017/S0950268814002337. Epub 2014 Sep 4.

Devera R, Blanco Y, Cabello E. [High prevalence of Cyclospora cayetanensis among indigenous people in Bolivar State, Venezuela]. Cad Saude Publica. 2005 Nov-Dec;21(6):1778-84. Epub 2006 Jan 9.

Devulder G, Pérouse de Montclos M, Flandrois JP. A multigene approach to phylogenetic analysis using the genus Mycobacterium as a model. Int J Syst Evol Microbiol. 2005 Jan;55(Pt 1):293-302.

Dey A, Ghoshal U, Agarwal V, Ghoshal UC. Genotyping of *Cryptosporidium* Species and Their Clinical Manifestations in Patients with Renal Transplantation and Human Immunodeficiency Virus Infection. J Pathog. 2016;2016:2623602. doi: 10.1155/2016/2623602. Epub 2016 Feb 14.

Dias S, Longacre S, Escalante AA, Udagama-Randeniya PV. Genetic diversity and recombination at the C-terminal fragment of the merozoite surface protein-1 of Plasmodium vivax (PvMSP-1) in Sri Lanka. Infect Genet Evol. 2011 Jan;11(1):145-56. doi: 10.1016/j.meegid.2010.09.007. Epub 2010 Oct 8.

Diaz E, Mondragon J, Ramirez E, Bernal R. Epidemiology and control of intestinal parasites with nitazoxanide in children in Mexico. Am J Trop Med Hyg. 2003 Apr;68(4):384-5.

Dieng T, Ndir O, Diallo S, Coll-Seck AM, Dieng Y. [Prevalence of *Cryptosporidium* sp and Isospora belli in patients with acquired immunodeficiency syndrome (AIDS) in Dakar (Senegal)]. Dakar Med. 1994;39(2):121-4.

Dillingham RA, Pinkerton R, Leger P, Severe P, Guerrant RL, Pape JW, Fitzgerald DW. High early mortality in patients with chronic acquired immunodeficiency syndrome diarrhea initiating antiretroviral therapy in Haiti: a case-control study. Am J Trop Med Hyg. 2009 Jun;80(6):1060-4.

Ditrich O, Palkovic L, Stěrba J, Prokopic J, Loudová J, Giboda M. The first finding of *Cryptosporidium baileyi* in man. Parasitol Res. 1991;77(1):44-7.

Djeddi D, Wallois C, Agnamey P, Kongolo G, Vanrenterghem A, Léké A, Mounard J, Raccurt F. [Incidence of *Cryptosporidium* sp in infantile acute gastroenteritis]. Arch Pediatr. 2010 Mar;17(3):301-2. doi: 10.1016/j.arcped.2009.11.021. Epub 2010 Feb 4.

Dlamini MS, Nkambule SJ, Grimason AM. First report of cryptosporidiosis in paediatric patients in Swaziland. Int J Environ Health Res. 2005 Oct;15(5):393-6.

DoE/SCA. Isolation and identification of Giardia cysts, *Cryptosporidium* oocysts, and free living pathogenic amoeba in water etc., 1989. (Methods for the examination of waters and associated materials). Department of the environment standing committee of analysts. 1989. London: HMSO.

Doğan N, Oz Y, Koçman NU, Nursal AF. [Comparison of individual differences in the direct microscopic examination in the diagnosis of intestinal parasites]. Turkiye Parazitol Derg. 2012;36(4):211-4. doi: 10.5152/tpd.2012.51.

Doganci T, Araz E, Ensari A, Tanyuksel M, Doganci L. Detection of *Cryptosporidium parvum* infection in childhood using various techniques. Med Sci Monit. 2002 Dec;8(12):MT223-6.

Donoghue MT, Keshavaiah C, Swamidatta SH, Spillane C. Evolutionary origins of Brassicaceae specific genes in Arabidopsis thaliana. BMC Evol Biol. 2011 Feb 18;11:47. doi: 10.1186/1471-2148-11-47.

Dorsch MR, Veal DA. Oligonucleotide probes for specific detection of Giardia lamblia cysts by fluorescent in situ hybridization. J Appl Microbiol. 2001 May;90(5):836-42.

Doster AR, Mahaffey EA, McClearen JR. Cryptosporidia in the cloacal coprodeum of red-lored parrots (Amazona autumnalis). Avian Diseases. 1979; 23: 654-661.

Doumbo O, Rossignol JF, Pichard E, Traore HA, Dembele TM, Diakite M, Traore F, Diallo DA. Nitazoxanide in the treatment of cryptosporidial diarrhea and other intestinal parasitic infections associated with acquired immunodeficiency syndrome in tropical Africa. Am J Trop Med Hyg. 1997 Jun;56(6):637-9.

Dowle M, Hill NJ, Power ML. *Cryptosporidium* from a free-ranging marsupial host: bandicoots in urban Australia. Vet Parasitol. 2013 Nov 15;198(1-2):197-200. doi: 10.1016/j.vetpar.2013.08.017. Epub 2013 Aug 31.

Dreesen O, Li B, Cross GA. Telomere structure and function in trypanosomes: a proposal. Nat Rev Microbiol. 2007 Jan;5(1):70-5. Epub 2006 Dec 11.

Drinkard LN, Halbritter A, Nguyen GT, Sertich PL, King M, Bowman S, Huxta R, Guagenti M. Notes from the Field: Outbreak of Cryptosporidiosis Among Veterinary Medicine Students--Philadelphia, Pennsylvania, February 2015. MMWR Morb Mortal Wkly Rep. 2015 Jul 24;64(28):773.

Drumo R, Widmer G, Morrison LJ, Tait A, Grelloni V, D'Avino N, Pozio E, Cacciò SM. Evidence of host-associated populations of *Cryptosporidium parvum* in Italy. Appl Environ Microbiol. 2012 May;78(10):3523-9. doi: 10.1128/AEM.07686-11. Epub 2012 Mar 2.

Dubey TP, Pande BP. Observations on the coccidian oocysts from Indian jungle cat (Felis chaus). Ind J Microbiol. 1963; 3, 103-8.

Duedu KO, Karikari YA, Attah SK, Ayeh-Kumi PF. Prevalence of intestinal parasites among patients of a Ghanaian psychiatry hospital. BMC Res Notes. 2015 Nov 5;8:651. doi: 10.1186/s13104-015-1634-6.

Duong TH, Kombila M, Dufillot D, Richard-Lenoble D, Owono Medang M, Martz M, Gendrel D, Engohan E, Moreno JL. [Role of cryptosporidiosis in infants in Gabon. Results of two prospective studies]. Bull Soc Pathol Exot. 1991;84(5 Pt 5):635-44.

Duong TH, Dufillot D, Koko J, Nze-Eyo'o R, Thuilliez V, Richard-Lenoble D, Kombila M. [Digestive cryptosporidiosis in young children in an urban area in Gabon]. Sante. 1995 May-Jun;5(3):185-8.

Dusaynski DW. Two new coccidia (Protozoa: Eimeriidae) from Costa Rican Lizards with a review of the Eimerio from Lizards. J. Protozool. 1969; 16, 581-5.

Dwivedi KK, Prasad G, Saini S, Mahajan S, Lal S, Baveja UK. Enteric opportunistic parasites among HIV infected individuals: associated risk factors and immune status. Jpn J Infect Dis. 2007 May;60(2-3):76-81.

Eassa SM, El-Wahab EW, Lotfi SE, El Masry SA, Shatat HZ, Kotkat AM. Risk Factors Associated with Parasitic Infection Among Municipality Solid-Waste Workers in an Egyptian Community. J Parasitol. 2016 Apr;102(2):214-21. doi: 10.1645/15-782. Epub 2015 Nov 25.

Ebner J, Koehler AV, Robertson G, Bradbury RS, Jex AR, Haydon SR, Stevens MA, Norton R, Joachim A, Gasser RB. Genetic analysis of Giardia and *Cryptosporidium* from people in Northern Australia using PCR-based tools. Infect Genet Evol. 2015 Dec;36:389-95. doi: 10.1016/j.meegid.2015.08.034. Epub 2015 Aug 29.

Efunshile MA, Ngwu BA, Kurtzhals JA, Sahar S, König B, Stensvold CR. Molecular Detection of the Carriage Rate of Four Intestinal Protozoa with Real-Time Polymerase Chain Reaction: Possible Overdiagnosis of Entamoeba histolytica in Nigeria. Am J Trop Med Hyg. 2015 Aug;93(2):257-62. doi: 10.4269/ajtmh.14-0781. Epub 2015 Jun 22.

Egorov A, Paulauskis J, Petrova L, Tereschenko A, Drizhd N, Ford T. Contamination of water supplies with *Cryptosporidium parvum* and Giardia lamblia and diarrheal illness in selected Russian cities. Int J Hyg Environ Health. 2002 May;205(4):281-9.

Ehrlich, P. V. Aus dem Verein fur innere Medizin zu Berlin. Sitzung vom 1. Mai 1882. Deutsche Medizinische Wochenschrift. 1882; 8, 269.

Ehsan AM, Geurden T, Casaert S, Parvin SM, Islam TM, Ahmed UM, Levecke B, Vercruysse J, Claerebout E. Assessment of zoonotic transmission of Giardia and *Cryptosporidium* between cattle and humans in rural villages in Bangladesh. PLoS One. 2015 Feb 19;10(2):e0118239. doi: 10.1371/journal.pone.0118239. eCollection 2015.

Eibach D, Krumkamp R, Al-Emran HM, Sarpong N, Hagen RM, Adu-Sarkodie Y, Tannich E, May J. Molecular characterization of *Cryptosporidium spp*. among children in rural Ghana. PLoS Negl Trop Dis. 2015 Mar 6;9(3):e0003551. doi: 10.1371/journal.pntd.0003551. eCollection 2015.

Ekdahl K, Andersson Y. Imported giardiasis: impact of international travel, immigration, and adoption. Am J Trop Med Hyg. 2005 Jun;72(6):825-30.

Elaine J, Sanford C. The Travel and Tropical Medicine Manual E-Book. 4th ed. N.p.: Elsevier Inc., 2008. Print.

El-Badry AA, Al-Antably AS, Hassan MA, Hanafy NA, Abu-Sarea EY. Molecular seasonal, age and gender distributions of *Cryptosporidium* in diarrhoeic Egyptians: distinct endemicity. Eur J Clin Microbiol Infect Dis. 2015 Dec;34(12):2447-53. doi: 10.1007/s10096-015-2502-y. Epub 2015 Oct 6.

ElBakri A, Mogane L, Ezzedine S, AbuOdeh R, Bessong P, Potgeiter N, Samie A. Real Time PCR detection and prevalence determination of *Cryptosporidium* species in stool specimens from the United Arab Emirates. 16th ICID Abstracts / International Journal of Infectious Diseases 21S (2014) 1–460

Elfving K, Andersson M, Msellem MI, Welinder-Olsson C, Petzold M, Björkman A, Trollfors B, Mårtensson A, Lindh M. Real-time PCR threshold cycle cutoffs help to identify agents causing acute childhood diarrhea in Zanzibar. J Clin Microbiol. 2014 Mar;52(3):916-23. doi: 10.1128/JCM.02697-13. Epub 2014 Jan 8.

Elgun G, Koltas IS. Investigation of *Cryptosporidium spp*. antigen by ELISA method in stool specimens obtained from patients with diarrhea. Parasitol Res. 2011 Feb;108(2):395-7. doi: 10.1007/s00436-010-2079-4. Epub 2010 Oct 12.

Ellegren, H. Microsatellites: simple sequences with complex evolution. Nat Rev Genet. 2004 Jun;5(6):435-45.

Elliott DA, Clark DP. *Cryptosporidium parvum* induces host cell actin accumulation at the host-parasite interface. Infect Immun. 2000 Apr;68(4):2315-22.

El-Mahallawy HA, El-Din NH, Salah F, El-Arousy M, El-Naga SA. Epidemiologic profile of symptomatic gastroenteritis in pediatric oncology patients receiving chemotherapy. Pediatr Blood Cancer. 2004 Apr;42(4):338-42.

El-Mohamady H, Abdel-Messih IA, Youssef FG, Said M, Farag H, Shaheen HI, Rockabrand DM, Luby SB, Hajjeh R, Sanders JW, Monteville MR, Klena JD, Frenck RW. Enteric pathogens associated with diarrhea in children in Fayoum, Egypt. Diagn Microbiol Infect Dis. 2006 Sep;56(1):1-5. Epub 2006 May 3.

Elnadi NA, Hassanien HA, Ahmad AM, Abd Ellah AK. Intestinal parasites in diabetic patients in Sohag University Hospitals, Egypt. J Egypt Soc Parasitol. 2015 Aug;45(2):443-9.

El-Refaey AM, Abdelbasset A, Atia G, Matar M, Awad SI, Yahya RS. Duodenal microbiasis in children on regular hemodialysis. Saudi J Kidney Dis Transpl. 2012 Nov;23(6):1278-80. doi: 10.4103/1319-2442.103576.

Elsafi SH, Al-Maqati TN, Hussein MI, Adam AA, Hassan MM, Al Zahrani EM. Comparison of microscopy, rapid immunoassay, and molecular techniques for the detection of Giardia lamblia and *Cryptosporidium parvum*. Parasitol Res. 2013 Apr;112(4):1641-6. doi: 10.1007/s00436-013-3319-1. Epub 2013 Feb 15.

El-Shabrawi M, Salem M, Abou-Zekri M, El-Naghi S, Hassanin F, El-Adly T, El-Shamy A. The burden of different pathogens in acute diarrhoeal episodes among a cohort of Egyptian children less than five years old. Prz Gastroenterol. 2015;10(3):173-80. doi: 10.5114/pg.2015.51186. Epub 2015 Apr 28.

El-Shazly LB, El-Faramawy AA, El-Sayed NM, Ismail KA, Fouad SM. Intestinal parasitic infection among Egyptian children with chronic liver diseases. J Parasit Dis. 2015 Mar;39(1):7-12. doi: 10.1007/s12639-013-0346-x. Epub 2013 Sep 3.

Elswaifi SF, Palmieri JR, El-Tantawy N, El-Hussiny M, Besheer T, Abohashem E. Comparison of microscopic and immunoassay examination in the diagnosis of intestinal protozoa of humans in Mansoura, Egypt. J Parasit Dis. 2016 Sep;40(3):580-5. doi: 10.1007/s12639-014-0542-3. Epub 2014 Aug 31.

Elton C, Ford EB, Baker JR, Gardner AD. The Health and Parasites of a Wild Mouse Population. Proceedings of the Zoological Society of London. September 1931; Volume 101, Issue 3, Pages: 657–721.

Elwin K, Hadfield SJ, Robinson G, Chalmers RM. The epidemiology of sporadic human infections with unusual cryptosporidia detected during routine typing in England and Wales, 2000-2008. Epidemiol Infect. 2012 Apr;140(4):673-83. doi: 10.1017/S0950268811000860. Epub 2011 Jun 7.

Elyan D, Wasfy M, El Mohammady H, Hassan K, Monestersky J, Noormal B, Oyofo B. Non-bacterial etiologies of diarrheal diseases in Afghanistan. Trans R Soc Trop Med Hyg. 2014 Aug;108(8):461-5. doi: 10.1093/trstmh/tru096. Epub 2014 Jun 18.

Emmett KJ, Lee A, Khiabanian H, Rabadan R. High-resolution Genomic Surveillance of 2014 Ebolavirus Using Shared Subclonal Variants. PLoS Curr. 2015 Feb 9;7. pii: ecurrents.outbreaks.c7fd7946ba606c982668a96bcba43c90. doi: 10.1371/currents.outbreaks.c7fd7946ba606c982668a96bcba43c90.

Enemark HL, Ahrens P, Juel CD, Petersen E, Petersen RF, Andersen JS, Lind P, Thamsborg SM. Molecular characterization of Danish *Cryptosporidium parvum* isolates. Parasitology. 2002 Oct;125(Pt 4):331-41.

Enserink R, van den Wijngaard C, Bruijning-Verhagen P, van Asten L, Mughini-Gras L, Duizer E, Kortbeek T, Scholts R, Nagelkerke N, Smit HA, Kooistra-Smid M, van Pelt W. Gastroenteritis attributable to 16 enteropathogens in children attending day care: significant effects of rotavirus, norovirus, astrovirus, *Cryptosporidium* and Giardia. Pediatr Infect Dis J. 2015 Jan;34(1):5-10. doi: 10.1097/INF.0000000000000472.

Eraky MA, El-Hamshary AM, Hamadto HH, Abdallah KF, Abdel-Hafed WM, Abdel-Had S. Predominance of *Cryptosporidium parvum* genotype among diarrheic children from Egypt as an indicator for zoonotic transmission. Acta Parasitol. 2014 Mar 1;60(1):26-34. doi: 10.1515/ap-2015-0004.

Erhabor O, Obunge O, Awah I. Cryptosporidiosis among HIV-infected persons in the Niger Delta of Nigeria. Niger J Med. 2011 Jul-Sep;20(3):372-5.

Erlandsen SL, Jarroll E, Wallis P, van Keulen H. Development of species-specific rDNA probes for Giardia by multiple fluorescent in situ hybridization combined with immunocytochemical identification of cyst wall antigens. J Histochem Cytochem. 2005 Aug;53(8):917-27. Epub 2005 May 6.

Esch MB, Locascio LE, Tarlov MJ, Durst RA. Detection of viable *Cryptosporidium parvum* using DNA-modified liposomes in a microfluidic chip. Anal Chem. 2001 Jul 1;73(13):2952-8.

Essers B, Burnens AP, Lanfranchini FM, Somaruga SG, von Vigier RO, Schaad UB, Aebi C, Bianchetti MG. Acute community-acquired diarrhea requiring hospital admission in Swiss children. Clin Infect Dis. 2000 Jul;31(1):192-6.

Essid R, Mousli M, Aoun K, Abdelmalek R, Mellouli F, Kanoun F, Derouin F, Bouratbine A. Identification of *Cryptosporidium* species infecting humans in Tunisia. Am J Trop Med Hyg. 2008 Nov;79(5):702-5.

Esteban JG, Flores A, Aguirre C, Strauss W, Angles R, Mas-Coma S. Presence of very high prevalence and intensity of infection with Fasciola hepatica among Aymara children from the Northern Bolivian Altiplano. Acta Trop. 1997 Jun 24;66(1):1-14.

Esteban JG, Aguirre C, Flores A, Strauss W, Angles R, Mas-Coma S. High *Cryptosporidium* prevalences in healthy Aymara children from the northern Bolivian Altiplano. Am J Trop Med Hyg. 1998 Jan;58(1):50-5.

Ethelberg S, Lisby M, Vestergaard LS, Enemark HL, Molbak K. Cryptosporidiosis outbreak associated with eating in a canteen, Denmark, August 2005. Euro Surveill. 2005 Oct 27;10(10):E051027.4.

Evans MR, Gardner D. Cryptosporidiosis outbreak associated with an educational farm holiday. Commun Dis Rep CDR Rev. 1996 Mar 1;6(3):R50-1.

Eyada TM, Botros SK, El Noshokaty EH, Nabil G. Prevalence of Chromosome 8 Aneuploidy in Egyptian Acute Myeloid Leukemia Patients using Fluorescence in situ Hybridization. J Clin Exp Pathol. 2016; 6:297. doi: 10.4172/2161-0681.1000297

FAO. The State of food and agriculture. Italy: Food & Agriculture Organization of the United Nations, 2013. Print.

Fathy MM, Abdelrazek NM, Hassan FA, El-Badry AA. Molecular copro-prevalence of *Cryptosporidium* in Egyptian children and evaluation of three diagnostic methods. Indian Pediatr. 2014 Sep;51(9):727-9.

Faye B, Dieng T, Tine RC, Diouf L, Sylla K, Ndiaye M, Sow D, Ndiaye JL, Ndiaye D, Ndiaye M, Badiane AS, Seck MC, Dieng Y, Faye O, Ndir O, Gaye O. [Cryptosporidiosis in Senegalese children: prevalence study and use of ELISA serologic diagnosis]. Bull Soc Pathol Exot. 2013 Oct;106(4):258-63. doi: 10.1007/s13149-013-0316-7. Epub 2013 Oct 16.

Feltus DC, Giddings CW, Schneck BL, Monson T, Warshauer D, McEvoy JM. Evidence supporting zoonotic transmission of *Cryptosporidium spp*. in Wisconsin. J Clin Microbiol. 2006 Dec;44(12):4303-8. Epub 2006 Sep 27.

Feng X, Rich SM, Akiyoshi D, Tumwine JK, Kekitiinwa A, Nabukeera N, Tzipori S, Widmer G. Extensive polymorphism in *Cryptosporidium parvum* identified by multilocus microsatellite analysis. Appl Environ Microbiol. 2000 Aug;66(8):3344-9.

Feng X, Rich SM, Tzipori S, Widmer G. Experimental evidence for genetic recombination in the opportunistic pathogen *Cryptosporidium parvum*. Mol Biochem Parasitol. 2002 Jan;119(1):55-62.

Feng Y, Ortega Y, He G, Das P, Xu M, Zhang X, Fayer R, Gatei W, Cama V, Xiao L. Feng Y1, Ortega Y, He G, Das P, Xu M, Zhang X, Fayer R, Gatei W, Cama V, Xiao L. Vet Parasitol. 2007 Mar 15;144(1-2):1-9. Epub 2006 Nov 13.

Feng Y, Dearen T, Cama V, Xiao L. 90-kilodalton heat shock protein, Hsp90, as a target for genotyping *Cryptosporidium spp*. known to infect humans. Eukaryot Cell. 2009 Apr;8(4):478-82. doi: 10.1128/EC.00294-08. Epub 2009 Jan 23.

Feng Y, Lal AA, Li N, Xiao L. Subtypes of *Cryptosporidium spp*. in mice and other small mammals. Exp Parasitol. 2011 Jan;127(1):238-42. doi: 10.1016/j.exppara.2010.08.002. Epub 2010 Aug 6.

Feng Y, Wang L, Duan L, Gomez-Puerta LA, Zhang L, Zhao X, Hu J, Zhang N, Xiao L. Extended outbreak of cryptosporidiosis in a pediatric hospital, China. Emerg Infect Dis. 2012 Feb;18(2):312-4. doi: 10.3201/eid1802.110666.

Feng Y, Torres E, Li N, Wang L, Bowman D, Xiao L. Population genetic characterisation of dominant *Cryptosporidium parvum* subtype IIaA15G2R1. Int J Parasitol. 2013 Dec;43(14):1141-7. doi: 10.1016/j.ijpara.2013.09.002. Epub 2013 Oct 12.

Feng Y, Tiao N, Li N, Hlavsa M, Xiao L. Multilocus sequence typing of an emerging *Cryptosporidium hominis* subtype in the United States. J Clin Microbiol. 2014 Feb;52(2):524-30. doi: 10.1128/JCM.02973-13. Epub 2013 Dec 4.

Ferreira MU, da Silva Nunes M, Wunderlich G. Antigenic diversity and immune evasion by malaria parasites. Clin Diagn Lab Immunol. 2004 Nov;11(6):987-95.

Firdu T, Abunna F, Girma M. Intestinal Protozoal Parasites in Diarrheal Children and Associated Risk Factors at Yirgalem Hospital, Ethiopia: A Case-Control Study. Int Sch Res Notices. 2014 Oct 29;2014:357126. doi: 10.1155/2014/357126. eCollection 2014.

Flecha MJ, Benavides CM, Tissiano G, Tesfamariam A, Cuadros J, de Lucio A, Bailo B, Cano L, Fuentes I, Carmena D. Detection and molecular characterisation of Giardia duodenalis, *Cryptosporidium spp*. and Entamoeba spp. among patients with gastrointestinal symptoms in Gambo Hospital, Oromia Region, southern Ethiopia. Trop Med Int Health. 2015 Sep;20(9):1213-1222. doi: 10.1111/tmi.12535. Epub 2015 May 19.

Fleming, R. "*Cryptosporidium*: Could It Be in Your Water?" Factsheet. Ontario Ministry of Agriculture, Food, and Rural Affairs, Oct. 2015. Web. 15 Oct. 2016

Florén CH, Chinenye S, Elfstrand L, Hagman C, Ihse I. ColoPlus, a new product based on bovine colostrum, alleviates HIV-associated diarrhoea. Scand J Gastroenterol. 2006 Jun;41(6):682-6.

Flórez AC, García DA, Moncada L, Beltrán M. [Prevalence of microsporidia and other intestinal parasites in patients with HIV infection, Bogota, 2001]. Biomedica. 2003 Sep;23(3):274-82.

Forsell J, Granlund M, Samuelsson L, Koskiniemi S, Edebro H, Evengård B. High occurrence of Blastocystis sp. subtypes 1-3 and Giardia intestinalis assemblage B among patients in Zanzibar, Tanzania. Parasit Vectors. 2016 Jun 29;9(1):370. doi: 10.1186/s13071-016-1637-8.

Foudeh AM, Fatanat Didar T, Veres T, Tabrizian M. Microfluidic designs and techniques using lab-on-a-chip devices for pathogen detection for point-of-care diagnostics. Lab Chip. 2012 Sep 21;12(18):3249-66. doi: 10.1039/c2lc40630f. Epub 2012 Aug 2.

Fournet N, Deege MP, Urbanus AT, Nichols G, Rosner BM, Chalmers RM, Gorton R, Pollock KG, van der Giessen JW, Wever PC, Dorigo-Zetsma JW, Mulder B, Mank TG, Overdevest I, Kusters JG, van Pelt W, Kortbeek LM. Simultaneous increase of *Cryptosporidium* infections in the Netherlands, the United Kingdom and Germany in late summer season, 2012. Euro Surveill. 2013 Jan 10;18(2). pii: 20348.

Fregonesi BM, Suzuki MN, Machado CS, Tonani KA, Fernandes AP, Monroe AA, Cervi MC, Segura-Muñoz S. Emergent and re-emergent parasites in HIV-infected children: immunological and socio-environmental conditions that are involved in the transmission of Giardia spp. and *Cryptosporidium spp*. Rev Soc Bras Med Trop. 2015 Nov-Dec;48(6):753-8. doi: 10.1590/0037-8682-0119-2015.

Freitas-Junior LH, Bottius E, Pirrit LA, Deitsch KW, Scheidig C, Guinet F, Nehrbass U, Wellems TE, Scherf A. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature. 2000 Oct 26;407(6807):1018-22.

Freites A, Colmenares D, Pérez M, García M, Díaz de Suárez O. [*Cryptosporidium* sp infections and other intestinal parasites in food handlers from Zulia state, Venezuela]. Invest Clin. 2009 Mar;50(1):13-21.

Fretz R, Svoboda P, Ryan UM, Thompson RC, Tanners M, Baumgartner A. Genotyping of *Cryptosporidium spp*. isolated from human stool samples in Switzerland. Epidemiol Infect. 2003 Aug;131(1):663-7.

Frickmann H, Schwarz NG, Rakotozandrindrainy R, May J, Hagen RM. PCR for enteric pathogens in high-prevalence settings. What does a positive signal tell us? Infect Dis (Lond). 2015 Jul;47(7):491-8. doi: 10.3109/23744235.2015.1022212. Epub 2015 Mar 12.

Fry K, Poon R, Whitcome P, Idriss J, Salser W, Mazrimas J, Hatch F. Nucleotide sequence of HS-beta satellite DNA from kangaroo rat Dipodomys ordii. Proc Natl Acad Sci U S A. 1973 Sep;70(9):2642-6.

Fu YX, Li WH. Statistical tests of neutrality of mutations. Genetics. 1993 Mar;133(3):693-709.

Fujita Y, Suda K, Kimura H, Matsumoto K, Arao T, Nagai T, Saijo N, Yatabe Y, Mitsudomi T, Nishio K. Highly sensitive detection of EGFR T790M mutation using colony hybridization predicts favorable prognosis of patients with lung cancer harboring activating EGFR mutation. J Thorac Oncol. 2012 Nov;7(11):1640-4. doi: 10.1097/JTO.0b013e3182653d7f.

Gadagkar SR, Rosenberg MS, Kumar S. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. J Exp Zool B Mol Dev Evol. 2005 Jan 15;304(1):64-74.

Gait R, Soutar RH, Hanson M, Fraser C, Chalmers R. Outbreak of cryptosporidiosis among veterinary students. Vet Rec. 2008 Jun 28;162(26):843-5.

García C, Rodríguez E, Do N, López de Castilla D, Terashima A, Gotuzzo E. [Intestinal parasitosis in patients with HIV-AIDS]. Rev Gastroenterol Peru. 2006 Jan-Mar;26(1):21-4.

Garcia LS, Bruckner DA, Brewer TC, Shimizu RY. Techniques for the recovery and identification of *Cryptosporidium* oocysts from stool specimens. J Clin Microbiol. 1983 Jul;18(1):185-90.

Garcia LS, Brewer TC, Bruckner DA. Incidence of *Cryptosporidium* in all patients submitting stool specimens for ova and parasite examination: monoclonal antibody IFA method. Diagn Microbiol Infect Dis. 1988 Sep;11(1):25-7.

Garcia LS, Current WL. Cryptosporidiosis: clinical features and diagnosis. Crit Rev Clin Lab Sci. 1989;27(6):439-60.

Garcia-R JC, Hayman DTS. Evolutionary processes in populations of *Cryptosporidium* inferred from gp60 sequence data. Parasitol Res. 2017 May 13. doi: 10.1007/s00436-017-5459-1. [Epub ahead of print]

Garzón M, Pereira-da-Silva L, Seixas J, Papoila AL, Alves M, Ferreira F, Reis A. Association of enteric parasitic infections with intestinal inflammation and permeability in asymptomatic infants of São Tomé Island. Pathog Glob Health. 2017 Mar 10:1-12. doi: 10.1080/20477724.2017.1299831. [Epub ahead of print]

Gascoyne P, Satayavivad J, Ruchirawat M. Microfluidic approaches to malaria detection. Acta Trop. 2004 Feb;89(3):357-69.

Gasparinho C, Mirante MC, Centeno-Lima S, Istrate C, Mayer AC, Tavira L, Nery SV, Brito M. Etiology of Diarrhea in Children Younger Than 5 Years Attending the Bengo General Hospital in Angola. Pediatr Infect Dis J. 2016 Feb;35(2):e28-34. doi: 10.1097/INF.0000000000000957.

Gassama A, Sow PS, Fall F, Camara P, Guèye-N'diaye A, Seng R, Samb B, M'Boup S, Aïdara-Kane A. Ordinary and opportunistic enteropathogens associated with diarrhea in Senegalese adults in relation to human immunodeficiency virus serostatus. Int J Infect Dis. 2001;5(4):192-8.

Gatei W, Suputtamongkol Y, Waywa D, Ashford RW, Bailey JW, Greensill J, Beeching NJ, Hart CA. Zoonotic species of *Cryptosporidium* are as prevalent as the anthroponotic in HIV-infected patients in Thailand. Ann Trop Med Parasitol. 2002 Dec;96(8):797-802.

Gatei W, Greensill J, Ashford RW, Cuevas LE, Parry CM, Cunliffe NA, Beeching NJ, Hart CA. Molecular analysis of the 18S rRNA gene of *Cryptosporidium* parasites from patients with or without human immunodeficiency virus infections living in Kenya, Malawi, Brazil, the United Kingdom, and Vietnam. J Clin Microbiol. 2003 Apr;41(4):1458-62.

Gatei W, Wamae CN, Mbae C, Waruru A, Mulinge E, Waithera T, Gatika SM, Kamwati SK, Revathi G, Hart CA. Cryptosporidiosis: prevalence, genotype analysis, and symptoms associated with infections in children in Kenya. Am J Trop Med Hyg. 2006 Jul;75(1):78-82.

Gatei W, Das P, Dutta P, Sen A, Cama V, Lal AA, Xiao L. Multilocus sequence typing and genetic structure of *Cryptosporidium hominis* from children in Kolkata, India. Infect Genet Evol. 2007 Mar;7(2):197-205. Epub 2006 Sep 28.

Gatei W, Barrett D, Lindo JF, Eldemire-Shearer D, Cama V, Xiao L. Unique *Cryptosporidium* population in HIV-infected persons, Jamaica. Emerg Infect Dis. 2008 May;14(5):841-3. doi: 10.3201/eid1405.071277.

Gaunt MW, Yeo M, Frame IA, Stothard JR, Carrasco HJ, Taylor MC, Mena SS, Veazey P, Miles GA, Acosta N, de Arias AR, Miles MA. Mechanism of genetic exchange in American trypanosomes. Nature. 2003 Feb 27;421(6926):936-9.

Gautam H, Bhalla P, Saini S, Uppal B, Kaur R, Baveja CP, Dewan R. Epidemiology of opportunistic infections and its correlation with CD4 T-lymphocyte counts and plasma viral load among HIV-positive patients at a tertiary care hospital in India. J Int Assoc Physicians AIDS Care (Chic). 2009 Nov-Dec;8(6):333-7. doi: 10.1177/1545109709346881. Epub 2009 Sep 15.

Gay-Andrieu E, Adehossi E, Illa H, Garba Ben A, Kourna H, Boureima H. [Prevalence of cryptosporidiosis in pediatric hospital patients in Niamey, Niger]. Bull Soc Pathol Exot. 2007 Aug;100(3):193-6.

Gbakima AA, Konteh R, Kallon M, Mansaray H, Sahr F, Bah ZJ, Spencer A, Luckay A. Intestinal protozoa and intestinal helminthic infections in displacement camps in Sierra Leone. Afr J Med Med Sci. 2007 Mar;36(1):1-9.

GBD HIV Collaborators. Estimates of global, regional, and national incidence, prevalence, and mortality of HIV, 1980-2015: the Global Burden of Disease Study 2015. Lancet HIV. 2016 Aug;3(8):e361-87. doi: 10.1016/S2352-3018(16)30087-X. Epub 2016 Jul 19.

Genois MM, Plourde M, Éthier C, Roy G, Poirier GG, Ouellette M, Masson JY. Roles of Rad51 paralogs for promoting homologous recombination in Leishmania infantum. Nucleic Acids Res. 2015 Mar 11;43(5):2701-15. doi: 10.1093/nar/gkv118. Epub 2015 Feb 24.

Getaneh A, Medhin G, Shimelis T. *Cryptosporidium* and Strongyloides stercoralis infections among people with and without HIV infection and efficiency of diagnostic methods for Strongyloides in Yirgalem Hospital, southern Ethiopia. BMC Res Notes. 2010 Apr 1;3:90. doi: 10.1186/1756-0500-3-90.

Geurden T, Levecke B, Cacció SM, Visser A, De Groote G, Casaert S, Vercruysse J, Claerebout E. Multilocus genotyping of *Cryptosporidium* and Giardia in non-outbreak related cases of diarrhoea in human patients in Belgium. Parasitology. 2009 Sep;136(10):1161-8. doi: 10.1017/S0031182009990436. Epub 2009 Jul 27.

Ghaffari S, Kalantari N, A Hart C. A Multi-Locus Study for Detection of *Cryptosporidium* Species Isolated from Calves Population, Liverpool; UK. Int J Mol Cell Med. 2014 Winter;3(1):35-42.

Ghaffari S, Kalantari N. A multi-locus study of *Cryptosporidium* parasites isolated from patients living in iran, Malawi, Nigeria, the United kingdom, and Vietnam. Iran J Parasitol. 2014 Mar;9(1):79-89.

Ghallab MM, Aziz IZ, Shoeib EY, El-Badry AA. Laboratory utility of coproscopy, copro immunoassays and copro nPCR assay targeting Hsp90 gene for detection of *Cryptosporidium* in children, Cairo, Egypt. J Parasit Dis. 2016 Sep;40(3):901-5. doi: 10.1007/s12639-014-0601-9. Epub 2014 Oct 30.

Gherasim A, Lebbad M, Insulander M, Decraene V, Kling A, Hjertqvist M, Wallensten A. Two geographically separated food-borne outbreaks in Sweden linked by an unusual *Cryptosporidium parvum* subtype, October 2010. Euro Surveill. 2012 Nov 15;17(46). pii: 20318.

Gholami R, Gholami S, Emadi-Kouchak H, Abdollahi A, Shahriari M. Clinical Characteristic of the HIV/AIDS Patients with Cryptosporidiosis Referring to Behavioral Diseases Consultation Center, Imam Khomeini Hospital, Tehran in 2013. Iran J Pathol. 2016 Winter;11(1):27-34.

Gholami S, Khanmohammadi M, Ahmadpour E, Pagheh AS, Khadem Nakhjiri S, Ramazannipour H, Shahbazi A. *Cryptosporidium* infection in patients with gastroenteritis in sari, iran. Iran J Parasitol. 2014 Apr-Jun;9(2):226-32.

Ghoshal U, Dey A, Ranjan P, Khanduja S, Agarwal V, Ghoshal UC. Identification of opportunistic enteric parasites among immunocompetent patients with diarrhoea from Northern India and genetic characterisation of *Cryptosporidium* and Microsporidia. Indian J Med Microbiol. 2016 Jan-Mar;34(1):60-6. doi: 10.4103/0255-0857.174114.

Giacometti A, Cirioni O, Fortuna M, Drenaggi D, Veccia S, D'Errico MM, Scalise G. [Giardiasis: a parasitic disease of continued topicality. Study of prevalence among a selected adult population]. Infez Med. 2000;8(2):82-86.

Gibney KB, O'Toole J, Sinclair M, Leder K. Disease burden of selected gastrointestinal pathogens in Australia, 2010. Int J Infect Dis. 2014 Nov;28:176-85. doi: 10.1016/j.ijid.2014.08.006. Epub 2014 Oct 2.

Gibson W. Liaisons dangereuses: sexual recombination among pathogenic trypanosomes. Res Microbiol. 2015 Jul-Aug;166(6):459-66. doi: 10.1016/j.resmic.2015.05.005. Epub 2015 May 28.

Gil FF, Barros MJ, Macedo NA, Júnior CG, Redoan R, Busatti H, Gomes MA, Santos JF. Prevalence of intestinal parasitism and associated symptomatology among hemodialysis patients. Rev Inst Med Trop Sao Paulo. 2013 Mar-Apr;55(2):69-74.

Giles M, Chalmers R, Pritchard G, Elwin K, Mueller-Doblies D, Clifton-Hadley F. *Cryptosporidium hominis* in a goat and a sheep in the UK. Vet Rec. 2009 Jan 3;164(1):24-5.

Girma M, Teshome W, Petros B, Endeshaw T. Cryptosporidiosis and Isosporiasis among HIV-positive individuals in south Ethiopia: a cross sectional study. BMC Infect Dis. 2014 Feb 22;14:100. doi: 10.1186/1471-2334-14-100.

Girotto KG, Grama DF, da Cunha MJ, Faria ES, Limongi JE, Pinto Rde M, Cury MC. Prevalence and risk factors for intestinal protozoa infection in elderly residents at Long Term Residency Institutions in Southeastern Brazil. Rev Inst Med Trop Sao Paulo. 2013 Jan-Feb;55(1):19-24.

Glaeser C, Grimm F, Mathis A, Weber R, Nadal D, Deplazes P. Detection and molecular characterization of *Cryptosporidium spp*. isolated from diarrheic children in Switzerland. Pediatr Infect Dis J. 2004 Apr;23(4):359-61.

Glover L, Jun J, Horn D. Microhomology-mediated deletion and gene conversion in African trypanosomes. Nucleic Acids Res. 2011 Mar;39(4):1372-80. doi: 10.1093/nar/gkq981. Epub 2010 Oct 21.

Gnida A, Kunda K, Ziembińska A, Luczkiewicz A, Felis E, Surmacz-Górska J. Detection of sulfonamide resistance genes via in situ PCR-FISH. Pol J Microbiol. 2014;63(2):167-73.

Gobet, P, Toze, S. Sensitive genotyping of *Cryptosporidium parvum* by PCR-RFLP analysis of the 70-kilodalton heat shock protein (HSP70) gene. FEMS Microbiol Lett. 2001 Jun 12;200(1):37-41.

Goh S, Reacher M, Casemore DP, Verlander NQ, Chalmers R, Knowles M, Williams J, Osborn K, Richards S. Sporadic cryptosporidiosis, North Cumbria, England, 1996-2000. Emerg Infect Dis. 2004 Jun;10(6):1007-15.

Gołąb E, Waloch M, Rozej W, Wernik T, Piotrowska M, Wąsik M, Dzbeński TH. Evaluation of prevalence of *Cryptosporidium spp*. in children with diarrhoea. 2007. Wiad Parazytol 53(suppl):103

Goldfarb DM, Dixon B, Moldovan I, Barrowman N, Mattison K, Zentner C, Baikie M, Bidawid S, Chan F, Slinger R. Nanolitre real-time PCR detection of bacterial, parasitic, and viral agents from patients with diarrhoea in Nunavut, Canada. Int J Circumpolar Health. 2013;72:19903. doi: 10.3402/ijch.v72i0.19903. Epub 2013 Apr 8.

González-Moreno O, Domingo L, Teixidor J, Gracenea M. Prevalence and associated factors of intestinal parasitisation: a cross-sectional study among outpatients with gastrointestinal symptoms in Catalonia, Spain. Parasitol Res. 2011 Jan;108(1):87-93. doi: 10.1007/s00436-010-2044-2. Epub 2010 Sep 23.

Gookin JL, Nordone SK, Argenzio RA. Host responses to *Cryptosporidium* infection. J Vet Intern Med. 2002 Jan-Feb;16(1):12-21.

Gookin JL, Chiang S, Allen J, Armstrong MU, Stauffer SH, Finnegan C, Murtaugh MP. NF-kappaB-mediated expression of iNOS promotes epithelial defense against infection by *Cryptosporidium parvum* in neonatal piglets. Am J Physiol Gastrointest Liver Physiol. 2006 Jan;290(1):G164-74. Epub 2005 Aug 25.

Goozé L, Kim K, Petersen C, Gut J, Nelson RG. Amplification of a *Cryptosporidium parvum* gene fragment encoding thymidylate synthase. J Protozool. 1991 Nov-Dec;38(6):56S-58S.

Gormley FJ, Little CL, Chalmers RM, Rawal N, Adak GK. Zoonotic cryptosporidiosis from petting farms, England and Wales, 1992-2009. Emerg Infect Dis. 2011 Jan;17(1):151-2. doi: 10.3201/eid1701.100902.

Gosselin KB, Aboud S, McDonald CM, Moyo S, Khavari N, Manji K, Kisenge R, Fawzi W, Kellogg M, Tran HQ, Kibiki G, Gratz J, Liu J, Gewirtz A, Houpt E, Duggan C. Etiology of Diarrhea, Nutritional Outcomes, and Novel Intestinal Biomarkers in Tanzanian Infants. J Pediatr Gastroenterol Nutr. 2017 Jan;64(1):104-108. doi: 10.1097/MPG.0000000000001323.

Graczyk TK, Grimes BH, Knight R, Da Silva AJ, Pieniazek NJ, Veal DA. Detection of *Cryptosporidium parvum* and Giardia lamblia carried by synanthropic flies by combined fluorescent in situ hybridization and a monoclonal antibody. Am J Trop Med Hyg. 2003 Feb;68(2):228-32.

Graczyk TK, Shiff CK, Tamang L, Munsaka F, Beitin AM, Moss WJ. The association of Blastocystis hominis and Endolimax nana with diarrheal stools in Zambian school-age children. Parasitol Res. 2005 Dec;98(1):38-43. Epub 2005 Oct 25.

Graczyk TK, Kacprzak M, Neczaj E, Tamang L, Graczyk H, Lucy FE, Girouard AS. Occurrence of *Cryptosporidium* and Giardia in sewage sludge and solid waste landfill leachate and quantitative comparative analysis of sanitization treatments on pathogen inactivation. Environ Res. 2008 Jan;106(1):27-33. Epub 2007 Jun 21.

Grigg ME, Suzuki Y. Sexual recombination and clonal evolution of virulence in Toxoplasma. Microbes Infect. 2003 Jun;5(7):685-90.

Gu Y, Xing S, He C. Genome-Wide Analysis Indicates Lineage-Specific Gene Loss during Papilionoideae Evolution. Genome Biol Evol. 2016 Feb 11;8(3):635-48. doi: 10.1093/gbe/evw021.

Gu Y, Wang X, Zhou C, Li P, Xu Q, Zhao C, Liu W, Xu W. Investigation on *Cryptosporidium* infections in wild animals in a zoo in Anhui Province. J Zoo Wildl Med. 2016 Sep;47(3):846-854.

Guarino A, Canani RB, Casola A, Pozio E, Russo R, Bruzzese E, Fontana M, Rubino A. Human intestinal cryptosporidiosis: secretory diarrhea and enterotoxic activity in Caco-2 cells. J Infect Dis. 1995 Apr;171(4):976-83.

Guk SM, Seo M, Park YK, Oh MD, Choe KW, Kim JL, Choi MH, Hong ST, Chai JY. Parasitic infections in HIV-infected patients who visited Seoul National University Hospital during the period 1995-2003. Korean J Parasitol. 2005 Mar;43(1):1-5.

Gumbo T, Sarbah S, Gangaidzo IT, Ortega Y, Sterling CR, Carville A, Tzipori S, Wiest PM. Intestinal parasites in patients with diarrhea and human immunodeficiency virus infection in Zimbabwe. AIDS. 1999 May 7;13(7):819-21.

Guo Y, Tang K, Rowe LA, Li N, Roellig DM, Knipe K, Frace M, Yang C, Feng Y, Xiao L. Comparative genomic analysis reveals occurrence of genetic recombination in virulent *Cryptosporidium hominis* subtypes and telomeric gene duplications in *Cryptosporidium parvum*. BMC Genomics. 2015 Apr 18;16:320. doi: 10.1186/s12864-015-1517-1.

Gupta K, Bala M, Deb M, Muralidhar S, Sharma DK. Prevalence of intestinal parasitic infections in HIV-infected individuals and their relationship with immune status. Indian J Med Microbiol. 2013 Apr-Jun;31(2):161-5. doi: 10.4103/0255-0857.115247.

Gupta S, Narang S, Nunavath V, Singh S. Chronic diarrhoea in HIV patients: prevalence of coccidian parasites. Indian J Med Microbiol. 2008 Apr-Jun;26(2):172-5.

Guyot K, Follet-Dumoulin A, Lelièvre E, Sarfati C, Rabodonirina M, Nevez G, Cailliez JC, Camus D, Dei-Cas E. Molecular characterization of *Cryptosporidium* isolates obtained from humans in France. J Clin Microbiol. 2001 Oct;39(10):3472-80.

Hadfield SJ, Robinson G, Elwin K, Chalmers RM. Detection and differentiation of *Cryptosporidium spp*. in human clinical samples by use of real-time PCR. J Clin Microbiol. 2011 Mar;49(3):918-24. doi: 10.1128/JCM.01733-10. Epub 2010 Dec 22.

Hadfield SJ, Pachebat JA, Swain MT, Robinson G, Cameron SJ, Alexander J, Hegarty MJ, Elwin K, Chalmers RM. Generation of whole genome sequences of new *Cryptosporidium hominis* and *Cryptosporidium parvum* isolates directly from stool samples. BMC Genomics. 2015 Aug 29;16:650. doi: 10.1186/s12864-015-1805-9.

Haider SS, Baqai R, Qureshi FM, Boorom K. Blastocystis spp., *Cryptosporidium spp*., and Entamoeba histolytica exhibit similar symptomatic and epidemiological patterns in healthcare-seeking patients in Karachi. Parasitol Res. 2012 Sep;111(3):1357-68. doi: 10.1007/s00436-012-2972-0. Epub 2012 Jul 5.

Hailu AW, G/Selassie S, Merid Y, Gebru AA, Ayene YY, Asefa MK. The case control studies of HIV and Intestinal parasitic infections rate in active pulmonary tuberculosis patients in Woldia General Hospital and Health Center in North Wollo, Amhara Region, Ethiopia. Int J Pharma Sci. 2015 May-Jun;5(3):1092-1099. Epub 2015 Jun 10.

Hajdu A, Vold L, Østmo TA, Helleve A, Helgebostad SR, Krogh T, Robertson L, de Jong B, Nygård K. Investigation of Swedish cases reveals an outbreak of cryptosporidiosis at a Norwegian hotel with possible links to in-house water systems. BMC Infect Dis. 2008 Nov 1;8:152. doi: 10.1186/1471-2334-8-152.

Hajdusek O, Ditrich O, Slapeta J. Molecular identification of *Cryptosporidium spp*. in animal and human hosts from the Czech Republic. Vet Parasitol. 2004 Jul 14;122(3):183-92.

Hakim SL, Gan CC, Malkit K, Azian MN, Chong CK, Shaari N, Zainuddin W, Chin CN, Sara Y, Lye MS. Parasitic infections among Orang Asli (aborigine) in the Cameron Highlands, Malaysia. Southeast Asian J Trop Med Public Health. 2007 May;38(3):415-9.

Hall, T. A. BioEdit: A user-friendly biological sequence alignment program for Windows 95/98/NT. 1999. Nucleic Acids Symposium Series 01/1999; 41:95-98.

Hamedi Y, Safa O, Haidari M. *Cryptosporidium* infection in diarrheic children in southeastern Iran. Pediatr Infect Dis J. 2005 Jan;24(1):86-8.

Hancock-Allen J, Alden NB, Cronquist AB. Cryptosporidiosis outbreak at an academic animal research laboratory-Colorado, 2014. Am J Ind Med. 2017 Feb;60(2):208-214. doi: 10.1002/ajim.22630.

Haque R, Mondal D, Kirkpatrick BD, Akther S, Farr BM, Sack RB, Petri WA Jr. Epidemiologic and clinical characteristics of acute diarrhea with emphasis on Entamoeba histolytica infections in preschool children in an urban slum of Dhaka, Bangladesh. Am J Trop Med Hyg. 2003 Oct;69(4):398-405.

Haque R, Mondal D, Karim A, Molla IH, Rahim A, Faruque AS, Ahmad N, Kirkpatrick BD, Houpt E, Snider C, Petri WA Jr. Prospective case-control study of the association between common enteric protozoal parasites and diarrhea in Bangladesh. Clin Infect Dis. 2009 May 1;48(9):1191-7. doi: 10.1086/597580.

Hartley CL, McCulloch R. Trypanosoma brucei BRCA2 acts in antigenic variation and has undergone a recent expansion in BRC repeat number that is important during homologous recombination. Mol Microbiol. 2008 Jun;68(5):1237-51. doi: 10.1111/j.1365-2958.2008.06230.x. Epub 2008 Apr 22.

Hasajová A, Valenčáková A, Malčeková B, Danišová O, Halán M, Goldová M, Sak B, Květoňová D, Kváč M, Halánová M. Significantly higher occurrence of *Cryptosporidium* infection in Roma children compared with non-Roma children in Slovakia. Eur J Clin Microbiol Infect Dis. 2014 Aug;33(8):1401-6. doi: 10.1007/s10096-014-2082-2. Epub 2014 Mar 12.

Hasegawa T, Matsumoto Y, Goitsuka R, Tsujimoto H, Ono K, Hasegawa A. In situ hybridization for the detection of feline interleukin 1 alpha mRNA on the paraffin-embedded section using biotin-labeled probes. J Vet Med Sci. 1991 Jun;53(3):451-6.

Hassan KM, Ali FE, Bella MA. Cryptosporidiosis among children attending the Maternity and Paediatrics Teaching Hospital, Wad Medani, Central Region, Sudan. J Egypt Soc Parasitol. 1991 Apr;21(1):213-8.

Hassen Amer O, Ashankyty IM, Haouas NA. Prevalence of intestinal parasite infections among patients in local public hospitals of Hail, Northwestern Saudi Arabia. Asian Pac J Trop Med. 2016 Jan;9(1):44-8. doi: 10.1016/j.apjtm.2015.12.009. Epub 2015 Dec 18.

Hawash Y. DNA extraction from protozoan oocysts/cysts in feces for diagnostic PCR. Korean J Parasitol. 2014 Jun;52(3):263-71. doi: 10.3347/kjp.2014.52.3.263. Epub 2014 Jun 26.

Hawash Y, Dorgham LSh, Al-Hazmi AS, Al-Ghamdi MS. Prevalence of *Cryptosporidium*-associated diarrhea in a high altitude-community of Saudi Arabia detected by conventional and molecular methods. Korean J Parasitol. 2014 Oct;52(5):479-85. doi: 10.3347/kjp.2014.52.5.479. Epub 2014 Oct 22.

Hawash Y, Ghonaim MM, Al-Hazmi AS. Internal Amplification Control for a *Cryptosporidium* Diagnostic PCR: Construction and Clinical Evaluation. Korean J Parasitol. 2015 Apr;53(2):147-54. doi: 10.3347/kjp.2015.53.2.147. Epub 2015 Apr 22.

Hawash YA, Dorgham LSh, Amir el-AM, Sharaf OF. Prevalence of Intestinal Protozoa among Saudi Patients with Chronic Renal Failure: A Case-Control Study. J Trop Med. 2015;2015:563478. doi: 10.1155/2015/563478. Epub 2015 Sep 28.

Hayashida K, Hara Y, Abe T, Yamasaki C, Toyoda A, Kosuge T, Suzuki Y, Sato Y, Kawashima S, Katayama T, Wakaguri H, Inoue N, Homma K, Tada-Umezaki M, Yagi Y, Fujii Y, Habara T, Kanehisa M, Watanabe H, Ito K, Gojobori T, Sugawara H, Imanishi T, Weir W, Gardner M, Pain A, Shiels B, Hattori M, Nene V, Sugimoto C. Comparative genome analysis of three eukaryotic parasites with differing abilities to transform leukocytes reveals key mediators of Theileria-induced leukocyte transformation. MBio. 2012 Sep 4;3(5):e00204-12. doi: 10.1128/mBio.00204-12. Print 2012.

Heidarnegadi S, Mohebali M, Maraghi Sh, Babaei Z, Farnia Sh, Bairami A, Rezaeian M. *Cryptosporidium spp*. Infection in human and domestic animals. Iran J Parasitol. 2012;7(1):53-8.

Hellard ME, Sinclair MI, Hogg GG, Fairley CK. Prevalence of enteric pathogens among community based asymptomatic individuals. J Gastroenterol Hepatol. 2000 Mar;15(3):290-3.

Hellard ME, Sinclair MI, Forbes AB, Fairley CK. A randomized, blinded, controlled trial investigating the gastrointestinal health effects of drinking water quality. Environ Health Perspect. 2001 Aug;109(8):773-8.

Helmy YA, Krücken J, Nöckler K, von Samson-Himmelstjerna G, Zessin KH. Molecular epidemiology of *Cryptosporidium* in livestock animals and humans in the Ismailia province of Egypt. Vet Parasitol. 2013 Mar 31;193(1-3):15-24. doi: 10.1016/j.vetpar.2012.12.015. Epub 2012 Dec 20.

Henry MC, Alary M, Desmet P, Gerniers M, Muteteke D, Nku I, Mutombo L, Piot P. Community survey of diarrhoea in children under 5 years in Kinshasa, Zaire. Ann Soc Belg Med Trop. 1995 Jun;75(2):105-14.

Herges GR, Widmer G, Clark ME, Khan E, Giddings CW, Brewer M, McEvoy JM. Evidence that *Cryptosporidium parvum* populations are panmictic and unstructured in the Upper Midwest of the United States. Appl Environ Microbiol. 2012 Nov;78(22):8096-101. doi: 10.1128/AEM.02105-12. Epub 2012 Sep 14.

Herwaldt BL, de Arroyave KR, Wahlquist SP, de Merida AM, Lopez AS, Juranek DD. Multiyear prospective study of intestinal parasitism in a cohort of Peace Corps volunteers in Guatemala. J Clin Microbiol. 2001 Jan;39(1):34-42.

Heyer WD. Regulation of recombination and genomic maintenance. Cold Spring Harb Perspect Biol. 2015 Aug 3;7(8):a016501. doi: 10.1101/cshperspect.a016501.

Hicks SJ, Theodoropoulos G, Carrington SD, Corfield AP. The role of mucins in host-parasite interactions. Part I-protozoan parasites. Parasitol Today. 2000 Nov;16(11):476-81.

Hijjawi NS, Meloni BP, Ng'anzo M, Ryan UM, Olson ME, Cox PT, Monis PT, Thompson RC. Complete development of *Cryptosporidium parvum* in host cell-free culture. Int J Parasitol. 2004 Jun;34(7):769-77.

Hijjawi N, Ng J, Yang R, Atoum MF, Ryan U. Identification of rare and novel *Cryptosporidium* GP60 subtypes in human isolates from Jordan. Exp Parasitol. 2010 Jun;125(2):161-4. doi: 10.1016/j.exppara.2010.01.011. Epub 2010 Jan 28.

Hijjawi N, Mukbel R, Yang R, Ryan U. Genetic characterization of *Cryptosporidium* in animal and human isolates from Jordan. Vet Parasitol. 2016 Sep 15;228:116-120. doi: 10.1016/j.vetpar.2016.08.015. Epub 2016 Aug 21.

Hilmarsdóttir I, Baldvinsdóttir GE, Harðardóttir H, Briem H, Sigurðsson SI. Enteropathogens in acute diarrhea: a general practice-based study in a Nordic country. Eur J Clin Microbiol Infect Dis. 2012 Jul;31(7):1501-9. doi: 10.1007/s10096-011-1470-0. Epub 2011 Nov 5.

Hira KG, Mackay MR, Hempstead AD, Ahmed S, Karim MM, O'Connor RM, Hibberd PL, Calderwood SB, Ryan ET, Khan WA, Ward HD. Genetic diversity of *Cryptosporidium spp*. from Bangladeshi children. J Clin Microbiol. 2011 Jun;49(6):2307-10. doi: 10.1128/JCM.00164-11. Epub 2011 Apr 6.

Hoare CA. Studies on some new ophidian and avian coccidia from Uganda, with a revision of the classification of the Eimeriidea. Parasitology. 1933; 25: 359-388.

Hoek MR, Oliver I, Barlow M, Heard L, Chalmers R, Paynter S. Outbreak of *Cryptosporidium parvum* among children after a school excursion to an adventure farm, south west England. J Water Health. 2008 Sep;6(3):333-8.

Hoerr FJ, Ranck FM, Hastings TF. Respiratory cryptosporidiosis in turkeys. Journal of American Veterinary Medical Association. 1978; 173: 1591-159

Holmes EC, Zhang LQ, Simmonds P, Ludlam CA, Brown AJ. Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient. Proc Natl Acad Sci U S A. 1992 Jun 1;89(11):4835-9.

Homer MJ, Aguilar-Delfin I, Telford SR 3rd, Krause PJ, Persing DH. Babesiosis. Clin Microbiol Rev. 2000 Jul;13(3):451-69.

Hong SH, Anu D, Jeong YI, Abmed D, Cho SH, Lee WJ, Lee SE. Molecular characterization of Giardia duodenalis and *Cryptosporidium parvum* in fecal samples of individuals in Mongolia. Am J Trop Med Hyg. 2014 Jan;90(1):43-7. doi: 10.4269/ajtmh.13-0271. Epub 2013 Nov 18.

Hormoz S. Amino acid composition of proteins reduces deleterious impact of mutations. Sci Rep. 2013 Oct 10;3:2919. doi: 10.1038/srep02919.

Horn D, McCulloch R. Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. Curr Opin Microbiol. 2010 Dec;13(6):700-5. doi: 10.1016/j.mib.2010.08.009. Epub 2010 Sep 29.

Houpt ER, Bushen OY, Sam NE, Kohli A, Asgharpour A, Ng CT, Calfee DP, Guerrant RL, Maro V, Ole-Nguyaine S, Shao JF. Short report: asymptomatic *Cryptosporidium hominis* infection among human immunodeficiency virus-infected patients in Tanzania. Am J Trop Med Hyg. 2005 Sep;73(3):520-2.

Howard P, Alexander ND, Atkinson A, Clegg AO, Gerega G, Javati A, Kajoi M, Lupiwa S, Lupiwa T, Mens M, Saleu G, Sanders RC, West B, Alpers MP. Bacterial, viral and parasitic aetiology of paediatric diarrhoea in the highlands of Papua New Guinea. J Trop Pediatr. 2000 Feb;46(1):10-4.

Hu X, Xiao G, Zheng P, Shang Y, Su Y, Zhang X, Liu X, Zhan S, St Leger RJ, Wang C. Trajectory and genomic determinants of fungal-pathogen speciation and host adaptation. Proc Natl Acad Sci U S A. 2014 Nov 25;111(47):16796-801. doi: 10.1073/pnas.1412662111. Epub 2014 Nov 3.

Huang X, Hou L, Xu X, Chen H, Ji H, Zhu S. One-PCR-tube approach for in situ DNA isolation and detection. Analyst. 2011 Oct 21;136(20):4254-9. doi: 10.1039/c1an15116a. Epub 2011 Aug 30.

Hudson RR, Slatkin M, Maddison WP. Estimation of Levels of Gene Flow from DNA Sequence Data. Genetics. 1992;132(2):583-589.

Huh JW, Moon SG, Lim YH. A survey of intestinal protozoan infections among gastroenteritis patients during a 3-year period (2004-2006) in Gyeonggi-do (province), South Korea. Korean J Parasitol. 2009 Sep;47(3):303-5. doi: 10.3347/kjp.2009.47.3.303. Epub 2009 Aug 28.

Huh S, Yu JR, Kim JI, Gotov C, Janchiv R, Seo JS. Intestinal protozoan infections and echinococcosis in the inhabitants of Dornod and Selenge, Mongolia (2003). Korean J Parasitol. 2006 Jun;44(2):171-4.

Huhulescu S, Kiss R, Brettlecker M, Cerny RJ, Hess C, Wewalka G, Allerberger F. Etiology of acute gastroenteritis in three sentinel general practices, Austria 2007. Infection. 2009 Apr;37(2):103-8. doi: 10.1007/s15010-008-8106-z. Epub 2008 Dec 3.

Hung CC, Tsaihong JC, Lee YT, Deng HY, Hsiao WH, Chang SY, Chang SC, Su KE. Prevalence of intestinal infection due to *Cryptosporidium* species among Taiwanese patients with human immunodeficiency virus infection. J Formos Med Assoc. 2007 Jan;106(1):31-5.

Hunter PR, Hughes S, Woodhouse S, Syed Q, Verlander NQ, Chalmers RM, Morgan K, Nichols G, Beeching N, Osborn K. Sporadic cryptosporidiosis case-control study with genotyping. Emerg Infect Dis. 2004 Jul;10(7):1241-9.

Hurst, LD. The Ka/Ks ratio: diagnosing the form of sequence evolution. Trends Genet. 2002 Sep;18(9):486.

Huruy K, Kassu A, Mulu A, Worku N, Fetene T, Gebretsadik S, Biadglegne F, Belyhun Y, Muche A, Gelaw A, Anagaw B, Yifru S, Wondie Y, Bekele A, Tiruneh M, Reissig D, Moges F. Intestinal parasitosis and shigellosis among diarrheal patients in Gondar teaching hospital, northwest Ethiopia. BMC Res Notes. 2011 Oct 31;4:472. doi: 10.1186/1756-0500-4-472.

Hussain G, Roychoudhury S, Singha B, Paul J. Incidence of *Cryptosporidium* andersoni in diarrheal patients from southern Assam, India: a molecular approach. Eur J Clin Microbiol Infect Dis. 2017 Jan 19. doi: 10.1007/s10096-016-2887-2. [Epub ahead of print]

Højlyng N, Mølbak K, Jepsen S. *Cryptosporidium spp*., a frequent cause of diarrhea in Liberian children. J Clin Microbiol. 1986 Jun;23(6):1109-13.

IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

Ibrahim MA, Abdel-Ghany AE, Abdel-Latef GK, Abdel-Aziz SA, Aboelhadid SM. Epidemiology and public health significance of *Cryptosporidium* isolated from cattle, buffaloes, and humans in Egypt. Parasitol Res. 2016 Jun;115(6):2439-48. doi: 10.1007/s00436-016-4996-3. Epub 2016 Apr 5.

Idris MA, Al-Jabri AM. Usefulness of Kato-Katz and trichrome staining as diagnostic methods for parasitic infections in clinical laboratories. J Sci Res Med Sci. 2001 Oct;3(2):65-8.

Idris NS, Dwipoerwantoro PG, Kurniawan A, Said M. Intestinal parasitic infection of immunocompromised children with diarrhoea: clinical profile and therapeutic response. J Infect Dev Ctries. 2010 Jun 3;4(5):309-17.

Ignatius R, Klemm T, Zander S, Gahutu J, Kimmig P, Mockenhaupt FP, Regnath T. Highly specific detection of *Cryptosporidium spp*. oocysts in human stool samples by undemanding and inexpensive phase contrast microscopy. Parasitol Res. 2016 Mar;115(3):1229-34. doi: 10.1007/s00436-015-4859-3. Epub 2015 Dec 8.

Inman LR, Takeuchi A. Spontaneous cryptosporidiosis in an adult female rabbit Veterinary Pathology. 1979; 16: 89-95.

Inpankaew T, Traub R, Thompson RC, Sukthana Y. Canine parasitic zoonoses in Bangkok temples. Southeast Asian J Trop Med Public Health. 2007 Mar;38(2):247-55.

Insulander M, Silverlås C, Lebbad M, Karlsson L, Mattsson JG, Svenungsson B. Molecular epidemiology and clinical manifestations of human cryptosporidiosis in Sweden. Epidemiol Infect. 2013 May;141(5):1009-20. doi: 10.1017/S0950268812001665. Epub 2012 Aug 9.

Inungu JN, Morse AA, Gordon C. Risk factors, seasonality, and trends of cryptosporidiosis among patients infected with human immunodeficiency virus. Am J Trop Med Hyg. 2000 Mar;62(3):384-7.

Iqbal A, Lim YA, Surin J, Sim BL. High diversity of *Cryptosporidium* subgenotypes identified in Malaysian HIV/AIDS individuals targeting gp60 gene. PLoS One. 2012;7(2):e31139. doi: 10.1371/journal.pone.0031139. Epub 2012 Feb 8.

Iqbal A, Goldfarb DM, Slinger R, Dixon BR. Prevalence and molecular characterization of *Cryptosporidium spp*. and Giardia duodenalis in diarrhoeic patients in the Qikiqtani Region, Nunavut, Canada. Int J Circumpolar Health. 2015 Jun 19;74:27713. doi: 10.3402/ijch.v74.27713. eCollection 2015.

Iqbal J, Hira PR, Al-Ali F, Philip R. Cryptosporidiosis in Kuwaiti children: seasonality and endemicity. Clin Microbiol Infect. 2001 May;7(5):261-6.

Iqbal J, Hira PR, Al-Ali F, Khalid N. Cyclospora cayetanensis: first report of imported and autochthonous infections in Kuwait. J Infect Dev Ctries. 2011 May 28;5(5):383-90.

Iqbal J, Khalid N, Hira PR. Cryptosporidiosis in Kuwaiti children: association of clinical characteristics with *Cryptosporidium* species and subtypes. J Med Microbiol. 2011 May;60(Pt 5):647-52. doi: 10.1099/jmm.0.028001-0. Epub 2011 Jan 13.

Irisarri-Gutiérrez MJ, Mingo MH, de Lucio A, Gil H, Morales L, Seguí R, Nacarapa E, Muñoz-Antoli C, Bornay-Llinares FJ, Esteban JG, Carmena D. Association between enteric protozoan parasites and gastrointestinal illness among HIV- and tuberculosis-infected individuals in the Chowke district, southern Mozambique. Acta Trop. 2017 Mar 13. pii: S0001-706X(16)30785-9. doi: 10.1016/j.actatropica.2017.03.010. [Epub ahead of print]

Ishaque M, Rashid R, Mubarak M. Gastrointestinal complications in renal transplant recipients detected by endoscopic biopsies in a developing country. Indian J Gastroenterol. 2015 Jan;34(1):51-7. doi: 10.1007/s12664-015-0537-8. Epub 2015 Mar 12.

Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bauser C, Beck A, Beverley SM, Bianchettin G, Borzym K, Bothe G, Bruschi CV, Collins M, Cadag E, Ciarloni L, Clayton C, Coulson RM, Cronin A, Cruz AK, Davies RM, De Gaudenzi J, Dobson DE, Duesterhoeft A, Fazelina G, Fosker N, Frasch AC, Fraser A, Fuchs M, Gabel C, Goble A, Goffeau A, Harris D, Hertz-Fowler C, Hilbert H, Horn D, Huang Y, Klages S, Knights A, Kube M, Larke N, Litvin L, Lord A, Louie T, Marra M, Masuy D, Matthews K, Michaeli S, Mottram JC, Müller-Auer S, Munden H, Nelson S, Norbertczak H, Oliver K, O'neil S, Pentony M, Pohl TM, Price C, Purnelle B, Quail MA, Rabbinowitsch E, Reinhardt R, Rieger M, Rinta J, Robben J, Robertson L, Ruiz JC, Rutter S, Saunders D, Schäfer M, Schein J, Schwartz DC, Seeger K, Seyler A, Sharp S, Shin H, Sivam D, Squares R, Squares S, Tosato V, Vogt C, Volckaert G, Wambutt R, Warren T, Wedler H, Woodward J, Zhou S, Zimmermann W, Smith DF, Blackwell JM, Stuart KD, Barrell B, Myler PJ. The genome of the kinetoplastid parasite, Leishmania major. Science. 2005 Jul 15;309(5733):436-42.

Iyer RN, Rao J, Venkatalakshmi A, Nahdi FB. Clinical and Microbiology Profile and Outcome of Diarrhea by Coccidian Parasites in Immunocompetent Children. Pediatr Infect Dis J. 2015 Sep;34(9):937-9. doi: 10.1097/INF.0000000000000770.

Izadi M, Jonaidi-Jafari N, Saburi A, Eyni H, Rezaiemanesh MR, Ranjbar R. Prevalence, molecular characteristics and risk factors for cryptosporidiosis among Iranian immunocompromised patients. Microbiol Immunol. 2012 Dec;56(12):836-42. doi: 10.1111/j.1348-0421.2012.00513.x.

Izadi M, Jonaidi-Jafari N, Saburi A, Eyni H, Rezaiemanesh MR, Ranjbar R. Cryptosporidiosis in Iranian Farm Workers and Their Household Members: A Hypothesis about Possible Zoonotic Transmission. J Trop Med. 2014;2014:405875. doi: 10.1155/2014/405875. Epub 2014 Apr 27.

Jacobsen KH, Ribeiro PS, Quist BK, Rydbeck BV. Prevalence of intestinal parasites in young Quichua children in the highlands of rural Ecuador. J Health Popul Nutr. 2007 Dec;25(4):399-405.

Jafari R, Maghsood AH, Fallah M. Prevalence of *Cryptosporidium* infection among livestock and humans in contact with livestock in Hamadan district, Iran, 2012. J Res Health Sci. 2012 May 29;13(1):86-9.

Jafari R, Maghsood AH, Safari M, Latifi M, Fallah M. Comparison of fecal antigen detection using enzyme linked immunosorbent assay with the auramine phenol staining method for diagnosis of human cryptosporidiosis. Jundishapur J Microbiol. 2015 Feb 2;8(2):e16470. doi: 10.5812/jjm.16470. eCollection 2015.

Jehi SE, Li X, Sandhu R, Ye F, Benmerzouga I, Zhang M, Zhao Y, Li B. Suppression of subtelomeric VSG switching by Trypanosoma brucei TRF requires its TTAGGG repeat-binding activity. Nucleic Acids Res. 2014 Nov 10;42(20):12899-911. doi: 10.1093/nar/gku942. Epub 2014 Oct 13.

Jervis HR, Merrill TG, Sprinz, H. Coccidiosis in the guinea pig small intestine due to a *Cryptosporidium*. Am. J. Vet. Res. 1966; 27, 408-14.

Jex AR, Whipp M, Campbell BE, Cacciò SM, Stevens M, Hogg G, Gasser RB. A practical and cost-effective mutation scanning-based approach for investigating genetic variation in *Cryptosporidium*. Electrophoresis. 2007 Nov;28(21):3875-83.

Jex AR, Pangasa A, Campbell BE, Whipp M, Hogg G, Sinclair MI, Stevens M, Gasser RB. Classification of *Cryptosporidium* species from patients with sporadic cryptosporidiosis by use of sequence-based multilocus analysis following mutation scanning. J Clin Microbiol. 2008 Jul;46(7):2252-62. doi: 10.1128/JCM.00116-08. Epub 2008 Apr 30.

Jex AR, Gasser RB. Analysis of the genetic diversity within *Cryptosporidium hominis* and *Cryptosporidium parvum* from imported and autochtonous cases of human cryptosporidiosis by mutation scanning. Electrophoresis. 2008 Nov;29(20):4119-29. doi: 10.1002/elps.200800422.

Jian F, Liu A, Wang R, Zhang S, Qi M, Zhao W, Shi Y, Wang J, Wei J, Zhang L, Xiao L. Common occurrence of *Cryptosporidium hominis* in horses and donkeys. Infect Genet Evol. 2016 Sep;43:261-6. doi: 10.1016/j.meegid.2016.06.004. Epub 2016 Jun 3.

Jiang H, Li N, Gopalan V, Zilversmit MM, Varma S, Nagarajan V, Li J, Mu J, Hayton K, Henschen B, Yi M, Stephens R, McVean G, Awadalla P, Wellems TE, Su XZ. High recombination rates and hotspots in a Plasmodium falciparum genetic cross. Genome Biol. 2011;12(4):R33. doi: 10.1186/gb-2011-12-4-r33. Epub 2011 Apr 4.

Jiang Y, Ren J, Yuan Z, Liu A, Zhao H, Liu H, Chu L, Pan W, Cao J, Lin Y, Shen Y. *Cryptosporidium* andersoni as a novel predominant *Cryptosporidium* species in outpatients with diarrhea in Jiangsu Province, China. BMC Infect Dis. 2014 Oct 25;14:555. doi: 10.1186/s12879-014-0555-7.

Jiménez Gutiérrez E, Pineda V, Calzada JE, Guerrant RL, Lima Neto JB, Pinkerton RC, Saldaña A. Enteric parasites and enteroaggregative Escherichia coli in children from Cañazas County, Veraguas Province, Panama. Am J Trop Med Hyg. 2014 Aug;91(2):267-72. doi: 10.4269/ajtmh.13-0438. Epub 2014 Jun 30.

Jiménez-Cardoso, E, Eligio-García L, Cano-Estrada A, Cortés-Campos A, Medina-Sansón A, Molina-Martínez D. Frequency of Emerging Parasites in HIV/AIDS and Oncological Patients Stool by Coprological and Molecular Analysis. Advances in Infectious Diseases; 201, , 162-171

Johnson DW, Pieniazek NJ, Griffin DW, Misener L, Rose JB. Development of a PCR protocol for sensitive detection of *Cryptosporidium* oocysts in water samples. Appl Environ Microbiol. 1995 Nov;61(11):3849-55.

Jovelin, R, Cutter, AD. Fine-scale signatures of molecular evolution reconcile models of indel-associated mutation. Genome Biol Evol. 2013;5(5):978-86. doi: 10.1093/gbe/evt051.

Kabayiza JC, Andersson ME, Nilsson S, Baribwira C, Muhirwa G, Bergström T, Lindh M. Diarrhoeagenic microbes by real-time PCR in Rwandan children under 5 years of age with acute gastroenteritis. Clin Microbiol Infect. 2014 Dec;20(12):O1128-35. doi: 10.1111/1469-0691.12698. Epub 2014 Aug 11.

Kadende P, Nkurunziza T, Floch JJ, Mpfizi B, Laroche R, Ndabaneze E, Aubry P. [Infectious diarrhea in African acquired immunodepression syndrome (AIDS). Apropos of 100 patients studied in Bujumbura (Burundi)]. Med Trop (Mars). 1989 Apr-Jun;49(2):129-33.

Kaminsky RG. Parasitism and diarrhoea in children from two rural communities and marginal barrio in Honduras. Trans R Soc Trop Med Hyg. 1991 Jan-Feb;85(1):70-3.

Kange'the E, McDermott B, Grace D, Mbae C, Mulinge E, Monda J, Nyongesa C, Ambia J, Njehu A. Prevalence of cryptosporidiosis in dairy cattle, cattle-keeping families, their non-cattle-keeping neighbours and HIV-positive individuals in Dagoretti Division, Nairobi, Kenya. Trop Anim Health Prod. 2012 Sep;44 Suppl 1:S11-6. doi: 10.1007/s11250-012-0201-6. Epub 2012 Aug 10.

Kaniyarakkal V, Mundangalam N, Moorkoth AP, Mathew S. Intestinal Parasite Profile in the Stool of HIV Positive Patients in relation to Immune Status and Comparison of Various Diagnostic Techniques with Special Reference to *Cryptosporidium* at a Tertiary Care Hospital in South India. Adv Med. 2016;2016:3564359. doi: 10.1155/2016/3564359. Epub 2016 Jul 17.

Kanta RC, Ravindranath G. Enteric protozoan parasitic infection in HIV seropositive individuals with diarrhoea. J Commun Dis. 2008 Mar;40(1):65-70.

Karadag G, Tamer GS, Dervisoglu E. Investigation of intestinal parasites in dialysis patients. Saudi Med J. 2013 Jul;34(7):714-8.

Karim MR, Zhang S, Jian F, Li J, Zhou C, Zhang L, Sun M, Yang G, Zou F, Dong H, Li J, Rume FI, Qi M, Wang R, Ning C, Xiao L. Multilocus typing of *Cryptosporidium spp*. and Giardia duodenalis from non-human primates in China. Int J Parasitol. 2014 Nov;44(13):1039-47. doi: 10.1016/j.ijpara.2014.07.006. Epub 2014 Aug 18.

Kashyap B, Sinha S, Das S, Rustagi N, Jhamb R. Efficiency of diagnostic methods for correlation between prevalence of enteric protozoan parasites and HIV/AIDS status--an experience of a tertiary care hospital in East Delhi. J Parasit Dis. 2010 Oct;34(2):63-7. doi: 10.1007/s12639-011-0020-0. Epub 2011 Feb 8.

Kassi RR, Kouassi RA, Yavo W, Barro-Kiki CP, Bamba A, Menan HI, Kone M. [Cryptosporidiosis and isosporiasis in children suffering from diarrhoea in Abidjan]. Bull Soc Pathol Exot. 2004 Nov;97(4):280-2.

Kassouha M, Soukkarieh C, Alkhaled A. First genotyping of *Cryptosporidium spp*. in pre-weaned calves, broiler chickens and children in Syria by PCR-RFLP analysis. Vet Parasitol. 2016 Jul 30;225:86-90. doi: 10.1016/j.vetpar.2016.06.009. Epub 2016 Jun 6.

Katsumata T, Hosea D, Wasito EB, Kohno S, Hara K, Soeparto P, Ranuh IG. Cryptosporidiosis in Indonesia: a hospital-based study and a community-based survey. Am J Trop Med Hyg. 1998 Oct;59(4):628-32.

Kattula D, Jeyavelu N, Prabhakaran AD, Premkumar PS, Velusamy V, Venugopal S, Geetha JC, Lazarus RP, Das P, Nithyanandhan K, Gunasekaran C, Muliyil J, Sarkar R, Wanke C, Ajjampur SS, Babji S, Naumova EN, Ward HD, Kang G. Natural History of Cryptosporidiosis in a Birth Cohort in Southern India. Clin Infect Dis. 2017 Feb 1;64(3):347-354. doi: 10.1093/cid/ciw730. Epub 2016 Nov 9.

Kaur R, Rawat D, Kakkar M, Uppal B, Sharma VK. Intestinal parasites in children with diarrhea in Delhi, India. Southeast Asian J Trop Med Public Health. 2002 Dec;33(4):725-9.

Ke C, Berney H, Mathewson A, Sheehan MM. Rapid amplification for the detection of Mycobacterium tuberculosis using a non-contact heating method in a silicon microreactor based thermal cycler. Sensors and Actuators B. 2004;102:308–314.

Kennedy GA, Kreitner GL, Strafuss AC. Cryptosporidiosis in three pigs. J Am Vet Med Assoc. 1977 Feb 1;170(3):348-50.

Kernohan KD, Bérubé NG. Three dimensional dual labelled DNA fluorescent in situ hybridization analysis in fixed tissue sections. MethodsX. 2014 May 9;1:30-5. doi: 10.1016/j.mex.2014.04.001. eCollection 2014.

Khalil S, Mirdha BR, Sinha S, Panda A, Singh Y, Joseph A, Deb M. Intestinal Parasitosis in Relation to Anti-Retroviral Therapy, CD4(+) T-cell Count and Diarrhea in HIV Patients. Korean J Parasitol. 2015 Dec;53(6):705-12. doi: 10.3347/kjp.2015.53.6.705. Epub 2015 Dec 31.

Khalil S, Mirdha BR, Paul J, Panda A, Makharia G, Chaudhry R, Bhatnagar S. Development and evaluation of molecular methods for detection of *Cryptosporidium spp*. in human clinical samples. Exp Parasitol. 2016 Nov;170:207-213. doi: 10.1016/j.exppara.2016.10.001. Epub 2016 Oct 4.

Khan SM, Debnath C, Pramanik AK, Xiao L, Nozaki T, Ganguly S. Molecular characterization and assessment of zoonotic transmission of *Cryptosporidium* from dairy cattle in West Bengal, India. Vet Parasitol. 2010 Jul 15;171(1-2):41-7. doi: 10.1016/j.vetpar.2010.03.008. Epub 2010 Mar 11.

Khan WA, Rogers KA, Karim MM, Ahmed S, Hibberd PL, Calderwood SB, Ryan ET, Ward HD. Cryptosporidiosis among Bangladeshi children with diarrhea: a prospective, matched, case-control study of clinical features, epidemiology and systemic antibody responses. Am J Trop Med Hyg. 2004 Oct;71(4):412-9.

Kiang KM, Scheftel JM, Leano FT, Taylor CM, Belle-Isle PA, Cebelinski EA, Danila R, Smith KE. Recurrent outbreaks of cryptosporidiosis associated with calves among students at an educational farm programme, Minnesota, 2003. Epidemiol Infect. 2006 Aug;134(4):878-86. Epub 2006 May 3.

Kiani H, Haghighi A, Rostami A, Azargashb E, Tabaei SJ, Solgi A, Zebardast N. Prevalence, risk factors, and symptoms associated to intestinal parasite infections among patients with gastrointestinal disorders in Nahavand, Western Iran. Rev Inst Med Trop Sao Paulo. 2016;58:42. doi: 10.1590/S1678-9946201658042. Epub 2016 May 24.

Kiani H, Haghighi A, Salehi R, Azargashb E. Distribution and risk factors associated with intestinal parasite infections among children with gastrointestinal disorders. Gastroenterol Hepatol Bed Bench. 2016 Dec;9(Suppl1):S80-S87.

Kilani RT, Wenman WM. Geographical variation in 18S rRNA gene sequence of *Cryptosporidium parvum*. Int J Parasitol. 1994 Apr;24(2):303-6.

Kimura A. Analysis of species and types of *Cryptosporidium* from water using molecular biological techniques. 2000-2001. Final research report summary. Abstract Only. URL: https://kaken.nii.ac.jp/en/grant/KAKENHI-PROJECT-12670385/

Kimura A, Edagawa A, Okada K, Takimoto A, Yonesho S, Karanis P. Detection and genotyping of *Cryptosporidium* from brown rats (Rattus norvegicus) captured in an urban area of Japan. Parasitol Res. 2007 May;100(6):1417-20. Epub 2007 Feb 21.

Kinross P, Beser J, Troell K, Axén C, Björkman C, Lebbad M, Winiecka-Krusnell J, Lindh J, Löfdahl M. *Cryptosporidium parvum* infections in a cohort of veterinary students in Sweden. Epidemiol Infect. 2015 Oct;143(13):2748-56. doi: 10.1017/S0950268814003318. Epub 2015 Jan 30.

Kirkman LA, Deitsch KW. Recombination and Diversification of the Variant Antigen Encoding Genes in the Malaria Parasite Plasmodium falciparum. Microbiol Spectr. 2014 Dec;2(6). doi: 10.1128/microbiolspec.MDNA3-0022-2014.

Kirkpatrick BD, Haque R, Duggal P, Mondal D, Larsson C, Peterson K, Akter J, Lockhart L, Khan S, Petri WA Jr. Association between *Cryptosporidium* infection and human leukocyte antigen class I and class II alleles. J Infect Dis. 2008 Feb 1;197(3):474-8. doi: 10.1086/525284.

Kiros H, Nibret E, Munshea A, Kerisew B, Adal M. Prevalence of intestinal protozoan infections among individuals living with HIV/AIDS at Felegehiwot Referral Hospital, Bahir Dar, Ethiopia. Int J Infect Dis. 2015 Jun;35:80-6. doi: 10.1016/j.ijid.2015.04.012. Epub 2015 Apr 23.

Kissinger JC, DeBarry J. Genome cartography: charting the apicomplexan genome. Trends Parasitol. 2011 Aug;27(8):345-54. doi: 10.1016/j.pt.2011.03.006. Epub 2011 Jul 19.

Koch R. Die Aetiolozie der Tuberkulose. Berliner klinische Wochenschrift. 1882; 19. 221.

Koehler AV, Bradbury RS, Stevens MA, Haydon SR, Jex AR, Gasser RB. Genetic characterization of selected parasites from people with histories of gastrointestinal disorders using a mutation scanning-coupled approach. Electrophoresis. 2013 Jun;34(12):1720-8. doi: 10.1002/elps.201300100. Epub 2013 May 26.

Koehler AV, Whipp M, Hogg G, Haydon SR, Stevens MA, Jex AR, Gasser RB. First genetic analysis of *Cryptosporidium* from humans from Tasmania, and identification of a new genotype from a traveller to Bali. Electrophoresis. 2014 Sep;35(18):2600-7. doi: 10.1002/elps.201400225. Epub 2014 Jul 28.

Koehler AV, Haydon SR, Jex AR, Gasser RB. *Cryptosporidium* and Giardia taxa in faecal samples from animals in catchments supplying the city of Melbourne with drinking water (2011 to 2015). Parasit Vectors. 2016 Jun 1;9(1):315. doi: 10.1186/s13071-016-1607-1.

Koinari M, Karl S, Ng-Hublin J, Lymbery AJ, Ryan UM. Identification of novel and zoonotic *Cryptosporidium* species in fish from Papua New Guinea. Vet Parasitol. 2013 Nov 15;198(1-2):1-9. doi: 10.1016/j.vetpar.2013.08.031. Epub 2013 Sep 7.

Koinari M, Lymbery AJ2, Ryan UM3. *Cryptosporidium* species in sheep and goats from Papua New Guinea. Exp Parasitol. 2014 Jun;141:134-7. doi: 10.1016/j.exppara.2014.03.021. Epub 2014 Apr 3.

Kołodziejczyk L, Kuźna-Grygiel W, Sulżyc-Bielicka V, Kładny J, Bielicki D, Stępień-Korzonek M, Balicka-Ramisz A. Cryptosporidiosis in patients with colon cancer. 2003. Wiad Parazytol 49:99

Konate A, Minta D, Diarra M, Dolo A, Dembele M, Diarra B, Maiga MY, Traore HA, Doumbo O. [Intestinal parasitosis during AIDS diarrhoea]. Bull Soc Pathol Exot. 2005 Apr;98(1):33-5.

Kong BN, Harwell JI, Suos P, Lynen L, Mohiuddin S, Reinert S, Pugatch D. Opportunistic infections and HIV clinical disease stage among patients presenting for care in Phnom Penh, Cambodia. Southeast Asian J Trop Med Public Health. 2007 Jan;38(1):62-8.

Körkoca H, Göz Y, Ataş AD, Kurtoğlu MG, Ekici K, Berktaş M. Prevalence of *Cryptosporidium spp*. in asymptomatic food workers. Turkiye Parazitol Derg. 2013;37(4):241-4. doi: 10.5152/tpd.2013.2981.

Korpe PS, Haque R, Gilchrist C, Valencia C, Niu F, Lu M, Ma JZ, Petri SE, Reichman D, Kabir M, Duggal P, Petri WA Jr. Natural History of Cryptosporidiosis in a Longitudinal Study of Slum-Dwelling Bangladeshi Children: Association with Severe Malnutrition. PLoS Negl Trop Dis. 2016 May 4;10(5):e0004564. doi: 10.1371/journal.pntd.0004564. eCollection 2016.

Kosek M, Alcantara C, Lima AA, Guerrant RL. Cryptosporidiosis: an update. Lancet Infect Dis. 2001 Nov;1(4):262-9.

Kosuwin R, Putaporntip C, Pattanawong U, Jongwutiwes S. Clonal diversity in Giardia duodenalis isolates from Thailand: evidences for intragenic recombination and purifying selection at the beta giardin locus. Gene. 2010 Jan 1;449(1-2):1-8. doi: 10.1016/j.gene.2009.09.010. Epub 2009 Sep 29.

Kovatch RM, White JD. Cryptosporidiosis in two juvenile rhesus monkeys. Vet Pathol. 1972; 9: 426–440.

Krause I, Amir J, Cleper R, Dagan A, Behor J, Samra Z, Davidovits M. Cryptosporidiosis in children following solid organ transplantation. Pediatr Infect Dis J. 2012 Nov;31(11):1135-8. doi: 10.1097/INF.0b013e31826780f7.

Krawczyk AI, van Leeuwen AD, Jacobs-Reitsma W, Wijnands LM, Bouw E, Jahfari S, van Hoek AH, van der Giessen JW, Roelfsema JH, Kroes M, Kleve J, Dullemont Y, Sprong H, de Bruin A. Presence of zoonotic agents in engorged ticks and hedgehog faeces from Erinaceus europaeus in (sub) urban areas. Parasit Vectors. 2015 Apr 9;8:210. doi: 10.1186/s13071-015-0814-5.

Krogh BO, Symington LS. Recombination proteins in yeast. Annu Rev Genet. 2004;38:233-71.

Kryazhimskiy S, Plotkin JB. The population genetics of dN/dS. PLoS Genet. 2008 Dec;4(12):e1000304. doi: 10.1371/journal.pgen.1000304. Epub 2008 Dec 12.

Kubatko LS, Degnan JH. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst Biol. 2007 Feb;56(1):17-24.

Kucerova Z, Sokolova OI, Demyanov AV, Kvac M, Sak B, Kvetonova D, Secor WE. Microsporidiosis and Cryptosporidiosis in HIV/AIDS Patients in St. Petersburg, Russia: Serological identification of microsporidia and *Cryptosporidium parvum* in sera samples from HIV/AIDS patients. AIDS Res Hum Retroviruses. 2011 Jan;27(1):13-5. doi: 10.1089/aid.2010.0132. Epub 2010 Oct 7.

Kuhls K, Cupolillo E, Silva SO, Schweynoch C, Boité MC, Mello MN, Mauricio I, Miles M, Wirth T, Schönian G. Population structure and evidence for both clonality and recombination among Brazilian strains of the subgenus Leishmania (Viannia). PLoS Negl Trop Dis. 2013 Oct 31;7(10):e2490. doi: 10.1371/journal.pntd.0002490. eCollection 2013.

Kukuruzovic R, Robins-Browne RM, Anstey NM, Brewster DR. Enteric pathogens, intestinal permeability and nitric oxide production in acute gastroenteritis. Pediatr Infect Dis J. 2002 Aug;21(8):730-9.

Kulik RA, Falavigna DL, Nishi L, Araujo SM. Blastocystis sp. and other intestinal parasites in hemodialysis patients. Braz J Infect Dis. 2008 Aug;12(4):338-41.

Kumar A, Chery L, Biswas C, Dubhashi N, Dutta P, Dua VK, Kacchap M, Kakati S, Khandeparkar A, Kour D, Mahajan SN, Maji A, Majumder P, Mohanta J, Mohapatra PK, Narayanasamy K, Roy K, Shastri J, Valecha N, Vikash R, Wani R, White J, Rathod PK. Malaria in South Asia: prevalence and control. Acta Trop. 2012 Mar;121(3):246-55. doi: 10.1016/j.actatropica.2012.01.004. Epub 2012 Jan 8.

Kumar GV, Tuohy MG, Ayyachamy M, O'Donovan A, Turner KM. Laboratory protocols in fungal biology: current methods in fungal biology. New York, NY: Springer. 2013. Print.

Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016 Jul;33(7):1870-4. doi: 10.1093/molbev/msw054. Epub 2016 Mar 22.

Kumar SS, Ananthan S, Lakshmi P. Intestinal parasitic infection in HIV infected patients with diarrhoea in Chennai. Indian J Med Microbiol. 2002 Apr-Jun;20(2):88-91.

Kuo CH, Kissinger JC. Consistent and contrasting properties of lineage-specific genes in the apicomplexan parasites Plasmodium and Theileria. BMC Evol Biol. 2008 Apr 11;8:108. doi: 10.1186/1471-2148-8-108.

Kurniawan A, Karyadi T, Dwintasari SW, Sari IP, Yunihastuti E, Djauzi S, Smith HV. Intestinal parasitic infections in HIV/AIDS patients presenting with diarrhoea in Jakarta, Indonesia. Trans R Soc Trop Med Hyg. 2009 Sep;103(9):892-8. doi: 10.1016/j.trstmh.2009.02.017. Epub 2009 Mar 26.

Kurniawan A, Dwintasari SW, Connelly L, Nichols RA, Yunihastuti E, Karyadi T, Djauzi S. *Cryptosporidium* species from human immunodeficiency-infected patients with chronic diarrhea in Jakarta, Indonesia. Ann Epidemiol. 2013 Nov;23(11):720-3. doi: 10.1016/j.annepidem.2013.07.019. Epub 2013 Aug 23.

Kvác M, Kvetonová D, Sak B, Ditrich O. *Cryptosporidium* pig genotype II in immunocompetent man. Emerg Infect Dis. 2009 Jun;15(6):982-3. doi: 10.3201/eid1506.071621.

Kyes SA, Kraemer SM, Smith JD. Antigenic variation in Plasmodium falciparum: gene organization and regulation of the var multigene family. Eukaryot Cell. 2007 Sep;6(9):1511-20. Epub 2007 Jul 20.

Laatamna AE, Wagnerová P, Sak B, Květoňová D, Xiao L, Rost M, McEvoy J, Saadi AR, Aissi M, Kváč M. Microsporidia and *Cryptosporidium* in horses and donkeys in Algeria: detection of a novel *Cryptosporidium hominis* subtype family (Ik) in a horse. Vet Parasitol. 2015 Mar 15;208(3-4):135-42. doi: 10.1016/j.vetpar.2015.01.007. Epub 2015 Jan 19.

Lalonde LF, Reyes J, Gajadhar AA. Application of a qPCR assay with melting curve analysis for detection and differentiation of protozoan oocysts in human fecal samples from Dominican Republic. Am J Trop Med Hyg. 2013 Nov;89(5):892-8. doi: 10.4269/ajtmh.13-0106. Epub 2013 Sep 9.

Lander RL, Lander AG, Houghton L, Williams SM, Costa-Ribeiro H, Barreto DL, Mattos AP, Gibson RS. Factors influencing growth and intestinal parasitic infections in preschoolers attending philanthropic daycare centers in Salvador, Northeast Region of Brazil. Cad Saude Publica. 2012 Nov;28(11):2177-88.

Lange H, Johansen OH, Vold L, Robertson LJ, Anthonisen IL, Nygard K. Second outbreak of infection with a rare *Cryptosporidium parvum* genotype in schoolchildren associated with contact with lambs/goat kids at a holiday farm in Norway. Epidemiol Infect. 2014 Oct;142(10):2105-13. doi: 10.1017/S0950268813003002. Epub 2013 Dec 5.

Lanternier F, Amazzough K, Favennec L, Mamzer-Bruneel MF, Abdoul H, Touret J, Decramer S, Zuber J, Scemla A, Legendre C, Lortholary O, Bougnoux ME; ANOFEL *Cryptosporidium* National Network; Transplant *Cryptosporidium* Study Group. *Cryptosporidium spp*. Infection in Solid Organ Transplantation: The Nationwide "TRANSCRYPTO" Study. Transplantation. 2016 Sep 27. [Epub ahead of print]

Larrosa-Haro A, Ruiz-Pérez M, Aguilar-Benavides S. [Utility of studying feces for the diagnosis and management of infants and preschool children with acute diarrhea]. Salud Publica Mex. 2002 Jul-Aug;44(4):328-34.

Larrosa-Haro A, Macias-Rosales R, Sánchez-Ramírez CA, Cortés-López MC, Aguilar-Benavides S. Seasonal variation of enteropathogens in infants and preschoolers with acute diarrhea in western Mexico. J Pediatr Gastroenterol Nutr. 2010 Oct;51(4):534-6. doi: 10.1097/MPG.0b013e3181df5b66.

Lasser KH, Lewin KJ, Ryning FW. Cryptosporidial enteritis in a patient with congenital hypogammaglobulinemia. Hum Pathol. 1979 Mar;10(2):234-40.

Laubach HE, Bentley CZ, Ginter EL, Spalter JS, Jensen LA. A study of risk factors associated with the prevalence of *Cryptosporidium* in villages around Lake Atitlan, Guatemala. Braz J Infect Dis. 2004 Aug;8(4):319-23. Epub 2004 Nov 19.

Lavdovskaia MV, Lysenko AIa, Basova EN, Lozovaia GA, Baleva LS, Rybalkina TN. [The host-opportunistic protozoa system. The effect of ionizing radiation on the prevalence of cryptosporidiosis and pneumocystosis]. Parazitologiia. 1996 Mar-Apr;30(2):153-8.

Lavdovskaia MV, Lysenko AIa, Gorbunova IuP, Derbina ZP, Pivovarova GP, Mikhalkina OP, Zhelezniak LG. [The characteristics of cryptosporidiosis epidemiology in the European part of Russia]. Med Parazitol (Mosk). 1996 Jul-Sep;(3):8-11.

Laxer MA, D'Nicuola ME, Patel RJ. Detection of *Cryptosporidium parvum* DNA in fixed, paraffin-embedded tissue by the polymerase chain reaction. Am J Trop Med Hyg. 1992 Oct;47(4):450-5.

Learmonth JJ, Ionas G, Ebbett KA, Kwan ES. Genetic characterization and transmission cycles of *Cryptosporidium* species isolated from humans in New Zealand. Appl Environ Microbiol. 2004 Jul;70(7):3973-8.

Leav BA, Mackay MR, Anyanwu A, O' Connor RM, Cevallos AM, Kindra G, Rollins NC, Bennish ML, Nelson RG, Ward HD. Analysis of sequence diversity at the highly polymorphic Cpgp40/15 locus among *Cryptosporidium* isolates from human immunodeficiency virus-infected children in South Africa. Infect Immun. 2002 Jul;70(7):3881-90.

Lebbad M, Norrgren H, Nauclér A, Dias F, Andersson S, Linder E. Intestinal parasites in HIV-2 associated AIDS cases with chronic diarrhoea in Guinea-Bissau. Acta Trop. 2001 Sep 1;80(1):45-9.

Lederer W, Allerberger F, Dierich MP. [Cryptosporidia associated gastroenteritis: epidemiologic study of children in Tyrol and Vorarlberg]. Wien Klin Wochenschr. 1996;108(20):649-53.

Lee AH, Symington LS, Fidock DA. DNA repair mechanisms and their biological roles in the malaria parasite Plasmodium falciparum. Microbiol Mol Biol Rev. 2014 Sep;78(3):469-86. doi: 10.1128/MMBR.00059-13.

Lee JD, Chiu ML, Chung LY, Yen CM. Prevalence of *Cryptosporidium* for foreign workers in Taiwan. Trop Doct. 2004 Jul;34(3):185-6.

Lee JK, Song HJ, Yu JR. Prevalence of diarrhea caused by *Cryptosporidium parvum* in non-HIV patients in Jeollanam-do, Korea. Korean J Parasitol. 2005 Sep;43(3):111-4.

Lee SC, Ngui R, Tan TK, Muhammad Aidil R, Lim YA. Neglected tropical diseases among two indigenous subtribes in peninsular Malaysia: highlighting differences and co-infection of helminthiasis and sarcocystosis. PLoS One. 2014 Sep 23;9(9):e107980. doi: 10.1371/journal.pone.0107980. eCollection 2014.

Lee SH, Kim SW, Kang JY, Ahn CH. A polymer lab-on-a-chip for reverse transcription (RT)-PCR based point-of-care clinical diagnostics. Lab Chip. 2008;8:2121–2127.

Lee WG, Kim YG, Chung BG, Demirci U, Khademhosseini A. Nano/Microfluidics for diagnosis of infectious diseases in developing countries. Adv Drug Deliv Rev. 2010 Mar 18;62(4-5):449-57. doi: 10.1016/j.addr.2009.11.016. Epub 2009 Nov 30.

Leelayoova S, Vithayasai N, Watanaveeradej V, Chotpitayasunondh T, Therapong V, Naaglor T, Mungthin M. Intestinal microsporidiosis in HIV-infected children with acute and chronic diarrhea. Southeast Asian J Trop Med Public Health. 2001 Mar;32(1):33-7.

Lehman LG, Kangam L, Mbenoun ML, Zemo Nguepi E, Essomba N, Tonga C, Bilong Bilong CF. Intestinal parasitic and candida infection associated with HIV infection in Cameroon. J Infect Dev Ctries. 2013 Feb 15;7(2):137-43. doi: 10.3855/jidc.2757.

Leitch GJ, He Q. Cryptosporidiosis-an overview. J Biomed Res. 2012 Jan;25(1):1-16. Epub 2012 Feb 21.

Lemieux J, Tran A, Freimark L, Goethert H, Bazner S, McGrath G, Rosenberg E, Telford S, Bailey J, Sabeti P. Whole-Genome Sequencing of Babesia microti From Patients Reveals Recent Origin, Extensive Population Structure, and Recent Expansion of Zoonotic Strains. Open Forum Infect Dis (Fall 2015) 2 (suppl 1): doi: 10.1093/ofid/ofv131.34

Lemos V, Graczyk TK, Alves M, Lobo ML, Sousa MC, Antunes F, Matos O. Identification and determination of the viability of Giardia lamblia cysts and *Cryptosporidium parvum* and *Cryptosporidium hominis* oocysts in human fecal and water supply samples by fluorescent in situ hybridization (FISH) and monoclonal antibodies. Parasitol Res. 2005 Dec;98(1):48-53. Epub 2005 Nov 1.

Leng X, Mosier DA, Oberst RD. Differentiation of *Cryptosporidium parvum*, *C. muris*, and *C. baileyi* by PCR-RFLP analysis of the 18S rRNA gene. Vet Parasitol. 1996 Mar;62(1-2):1-7.

Leoni F, Amar C, Nichols G, Pedraza-Díaz S, McLauchlin J. Genetic analysis of *Cryptosporidium* from 2414 humans with diarrhoea in England between 1985 and 2000. J Med Microbiol. 2006 Jun;55(Pt 6):703-7.

Levallois P, Chevalier P, Gingras S, Déry P, Payment P, Michel P, Rodriguez M. Risk of infectious gastroenteritis in young children living in Québec rural areas with intensive animal farming: results of a case-control study (2004-2007). Zoonoses Public Health. 2014 Feb;61(1):28-38. doi: 10.1111/zph.12039. Epub 2013 Feb 14.

Levine ND. Protozoan Parasites of Domestic Animals and of Man. Minneapolis, USA: Burgess Pub. Co. 1961. Print.

Ley DH, Levy MG, Hunter L, Corbett W, Barnes HJ. Cryptosporidia-positive rates of avian necropsy accessions determined by examination of auramine O-stained fecal smears. Avian Dis. 1988 Jan-Mar;32(1):108-13.

Li J, Feng J, Liu Y, Ho TV, Grimes W, Ho HA, Park S, Wang S, Chai Y. BMP-SHH signaling network controls epithelial stem cell fate via regulation of its niche in the developing tooth. Dev Cell. 2015 Apr 20;33(2):125-35. doi: 10.1016/j.devcel.2015.02.021. Epub 2015 Apr 9.

Li N, Xiao L, Cama VA, Ortega Y, Gilman RH, Guo M, Feng Y. Genetic recombination and *Cryptosporidium hominis* virulent subtype IbA10G2. Emerg Infect Dis. 2013 Oct;19(10):1573-82. doi: 10.3201/eid1910.121361.

Li W, Kiulia NM, Mwenda JM, Nyachieo A, Taylor MB, Zhang X, Xiao L. Cyclospora papionis, *Cryptosporidium hominis*, and human-pathogenic Enterocytozoon bieneusi in captive baboons in Kenya. J Clin Microbiol. 2011 Dec;49(12):4326-9. doi: 10.1128/JCM.05051-11. Epub 2011 Sep 28.

Librado, P. and Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009. 25: 1451-1452.

Lim YA, Rohela M, Sim BL, Jamaiah I, Nurbayah M. Prevalence of cryptosporidiosis in HIV-infected patients in Kajang Hospital, Selangor. Southeast Asian J Trop Med Public Health. 2005;36 Suppl 4:30-3.

Lim YA, Iqbal A, Surin J, Sim BL, Jex AR, Nolan MJ, Smith HV, Gasser RB. First genetic classification of *Cryptosporidium* and Giardia from HIV/AIDS patients in Malaysia. Infect Genet Evol. 2011 Jul;11(5):968-74. doi: 10.1016/j.meegid.2011.03.007. Epub 2011 Mar 23.

Lindo JF, Dubon JM, Ager AL, de Gourville EM, Solo-Gabriele H, Klaskala WI, Baum MK, Palmer CJ. Intestinal parasitic infections in human immunodeficiency virus (HIV)-positive and HIV-negative individuals in San Pedro Sula, Honduras. Am J Trop Med Hyg. 1998 Apr;58(4):431-5.

Lindo JF, Levy VA, Baum MK, Palmer CJ. Epidemiology of giardiasis and cryptosporidiosis in Jamaica. Am J Trop Med Hyg. 1998 Nov;59(5):717-21.

Linz B, Ivanov YV, Preston A, Brinkac L, Parkhill J, Kim M, Harris SR, Goodfield LL, Fry NK, Gorringe AR, Nicholson TL, Register KB, Losada L, Harvill ET.

Liò, P., Goldman, N. Models of Molecular Evolution and Phylogeny. Genome Res. 1998. 8: 1233-1244.

Liu C, Abrahamsen, MS. Identification of a putative telomeric repeat DNA binding factor of *Cryptosporidium parvum*. J Eukaryot Microbiol. 1999 Sep-Oct;46(5):50S-51S.

Liu C, Wang L, Lancto CA, Abrahamsen MS. Characterization of a *Cryptosporidium parvum* protein that binds single-stranded G-strand telomeric DNA. Mol Biochem Parasitol. 2009 Jun;165(2):132-41. doi: 10.1016/j.molbiopara.2009.01.013. Epub 2009 Feb 4.

Liu H, Shen Y, Yin J, Yuan Z, Jiang Y, Xu Y, Pan W, Hu Y, Cao J. Prevalence and genetic characterization of *Cryptosporidium*, Enterocytozoon, Giardia and Cyclospora in diarrheal outpatients in China. BMC Infect Dis. 2014 Jan 13;14:25. doi: 10.1186/1471-2334-14-25.

Liu R, Ma X, Liu A, Zhang L, Cai J, Wang M. Identification and characterization of a cathepsin-L-like peptidase in Eimeria tenella. Parasitol Res. 2014 Dec;113(12):4335-48. doi: 10.1007/s00436-014-4107-2. Epub 2014 Sep 25.

Liu X, Xie N, Li W, Zhou Z, Zhong Z, Shen L, Cao S, Yu X, Hu Y, Chen W, Peng G. Emergence of *Cryptosporidium hominis* Monkey Genotype II and Novel Subtype Family Ik in the Squirrel Monkey (Saimiri sciureus) in China. PLoS One. 2015 Oct 28;10(10):e0141450. doi: 10.1371/journal.pone.0141450. eCollection 2015.

Llewellyn S, Inpankaew T, Nery SV, Gray DJ, Verweij JJ, Clements AC, Gomes SJ, Traub R, McCarthy JS. Application of a Multiplex Quantitative PCR to Assess Prevalence and Intensity Of Intestinal Parasite Infections in a Controlled Clinical Trial. PLoS Negl Trop Dis. 2016 Jan 28;10(1):e0004380. doi: 10.1371/journal.pntd.0004380. eCollection 2016.

Llorente MT, Clavel A, Goñi MP, Varea M, Seral C, Becerril R, Suarez L, Gómez-Lus R. Genetic characterization of *Cryptosporidium* species from humans in Spain. Parasitol Int. 2007 Sep;56(3):201-5. Epub 2007 Feb 22.

Lobo ML, Augusto J, Antunes F, Ceita J, Xiao L, Codices V, Matos O. *Cryptosporidium spp*., Giardia duodenalis, Enterocytozoon bieneusi and other intestinal parasites in young children in Lobata province, Democratic Republic of São Tomé and Principe. PLoS One. 2014 May 20;9(5):e97708. doi: 10.1371/journal.pone.0097708. eCollection 2014.

Logar J, Poljsak-Prijatelj M, Andlovic A. Incidence of *Cryptosporidium parvum* in patients with diarrhea. J Eukaryot Microbiol. 1996 Sep-Oct;43(5):67S.

López L, Castillo FJ, Fernández MA, Clavel A, Rubio MC, Gómez-Lus R, Cutillas B. Astrovirus infection among children with gastroenteritis in the city of Zaragoza, Spain. Eur J Clin Microbiol Infect Dis. 2000 Jul;19(7):545-7.

Lowe, Andrew, Stephen Harris, and Paul Ashton. Ecological Genetics: Design, Analysis, and Application. Malden, MA, USA: Blackwell Pub., 2004. Print.

Lowery CJ, Millar BC, Moore JE, Xu J, Xiao L, Rooney PJ, Crothers L, Dooley JS. Molecular genotyping of human cryptosporidiosis in Northern Ireland: epidemiological aspects and review. Ir J Med Sci. 2001 Oct-Dec;170(4):246-50.

Lu J, Li CP. [The survey of *Cryptosporidium* infection among young children in kindergartens in Anhui Province]. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 2004 Dec;22(6):331-3.

Lucy FE, Graczyk TK, Tamang L, Miraflor A, Minchin D. Biomonitoring of surface and coastal water for *Cryptosporidium*, Giardia, and human-virulent microsporidia using molluscan shellfish. Parasitol Res. 2008 Nov;103(6):1369-75. doi: 10.1007/s00436-008-1143-9. Epub 2008 Aug 14.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, Förster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, König A, Liss T, Lüssmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer KH. ARB: a software environment for sequence data. Nucleic Acids Res. 2004 Feb 25;32(4):1363-71. Print 2004.

Lule JR, Mermin J, Awor A, Hughes P, Kigozi A, Wafula W, Nakanjako D, Kaharuza F, Downing R, Quick R. Aetiology of diarrhoea among persons with HIV and their family members in rural Uganda: a community-based study. East Afr Med J. 2009 Sep;86(9):422-9.

Löschenberger K, Szölgyényi W, Peschke R, Prosl H. Detection of the protozoan Neospora caninum using in situ polymerase chain reaction. Biotech Histochem. 2004 Apr;79(2):101-5.

Ma, P, Soave, R. Three-step stool examination for cryptosporidiosis in 10 homosexual men with protracted watery diarrhea. J Infect Dis. 1983 May;147(5):824-8.

Macfarlane DE, Horner-Bryce J. Cryptosporidiosis in well-nourished and malnourished children. Acta Paediatr Scand. 1987 May;76(3):474-7.

Madison-Antenucci S, Relich RF, Doyle L, Espina N, Fuller D, Karchmer T, Lainesse A, Mortensen JE, Pancholi P, Veros W, Harrington SM. Multicenter Evaluation of BD Max Enteric Parasite Real-Time PCR Assay for Detection of Giardia duodenalis, *Cryptosporidium hominis*, *Cryptosporidium parvum*, and Entamoeba histolytica. J Clin Microbiol. 2016 Nov;54(11):2681-2688. Epub 2016 Aug 17.

Magi B, Canocchi V, Tordini G, Cellesi C, Barberi A. *Cryptosporidium* infection: diagnostic techniques. Parasitol Res. 2006 Jan;98(2):150-2. Epub 2005 Nov 18.

Mahdavi Poor B, Rashedi J, Asgharzadeh M, Fallah E, Hatam-Nahavandi K, Dalimi A. Molecular Characterization of *Cryptosporidium* Species in Children with Diarrhea in North West of Iran. Int J Mol Cell Med. 2015 Fall;4(4):235-9.

Mahdi NK, Ali NH. Cryptosporidiosis among animal handlers and their livestock in Basrah, Iraq. East Afr Med J. 2002 Oct;79(10):550-3.

Mahdi NK, Ali NH. Cryptosporidiosis and other intestinal parasitic infections in patients with chronic diarrhea. Saudi Med J. 2004 Sep;25(9):1204-7.

Mahdi NK, Ali NH. Intestinal parasites, including *Cryptosporidium* species, in Iraqi patients with sickle-cell anaemia. East Mediterr Health J. 2002 Mar-May;8(2-3):345-9.

Mahgoub ES, Almahbashi A, Abdulatif B. Cryptosporidiosis in children in a north Jordanian paediatric hospital. East Mediterr Health J. 2004 Jul-Sep;10(4-5):494-501.

Maiga MY, Dembele MY, Traoré HA, Kouyate M, Traoré AK, Maiga II, Bougoudogo F, Doumbo O, Guindo A. [Gastrointestinal manifestations of AIDS in adults in Mali]. Bull Soc Pathol Exot. 2002 Nov;95(4):253-6.

Maikai BV, Umoh JU, Lawal IA, Kudi AC, Ejembi CL, Xiao L. Molecular characterizations of *Cryptosporidium*, Giardia, and Enterocytozoon in humans in Kaduna State, Nigeria. Exp Parasitol. 2012 Aug;131(4):452-6. doi: 10.1016/j.exppara.2012.05.011. Epub 2012 Jun 1.

Mairhofer J, Roppert K, Ertl P. Microfluidic systems for pathogen sensing: a review. Sensors (Basel). 2009;9(6):4804-23. doi: 10.3390/s90604804. Epub 2009 Jun 17.

Maizels, N. M. The nucleotide sequence of the lactose messenger ribonucleic acid transcribed from the UV5 promoter mutant of Escherichia coli. Proc Natl Acad Sci U S A. 1973 Dec;70(12):3585-9.

Makni F, Cheikrouhou F, Ayadi A. [Parasitoses and immunodepression]. Arch Inst Pasteur Tunis. 2000;77(1-4):51-4.

Malik SB, Pightling AW, Stefaniak LM, Schurko AM, Logsdon JM Jr. An expanded inventory of conserved meiotic genes provides evidence for sex in Trichomonas vaginalis. PLoS One. 2007 Aug 6;3(8):e2879. doi: 10.1371/journal.pone.0002879.

Mallon M, MacLeod A, Wastling J, Smith H, Reilly B, Tait A. Population structures and the role of genetic exchange in the zoonotic pathogen *Cryptosporidium parvum*. J Mol Evol. 2003 Apr;56(4):407-17.

Manocha H, Dua S, Chander Y, Tailang M. Cryptosporidiosis, whether it is more prevalent in Southern India. Trop Parasitol. 2014 Jul;4(2):125-7. doi: 10.4103/2229-5070.138543.

Manter, HW. A laboratory manual in animal parasitology: with special reference to the animal parasites of man. Minneapolis, USA: Burgess Pub. Co. 1938. Print.

Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol. 2015 May 26;1(1):vev003. eCollection 2015.

Martín-Ampudia M, Mariscal A, Lopez-Gigosos RM, Mora L, Fernandez-Crehuet J. Under-notification of cryptosporidiosis by routine clinical and laboratory practices among non-hospitalised children with acute diarrhoea in Southern Spain. Infection. 2012 Apr;40(2):113-9. doi: 10.1007/s15010-011-0188-3. Epub 2011 Sep 6.

Masarat S, Ahmad F, Chisti M, Hamid S, Sofi BA. Prevalence of *Cryptosporidium* species among HIV positive asymptomatic and symptomatic immigrantpopulation in Kashmir, India. Iran J Microbiol. 2012 Mar;4(1):35-9.

Mascarini LM, Donalísio MR. [Giardiasis and cryptosporidiosis in children institutionalized at daycare centers in the state of São Paulo]. Rev Soc Bras Med Trop. 2006 Nov-Dec;39(6):577-9.

Mata L, Bolaños H, Pizarro D, Vives M. Cryptosporidiosis in children from some highland Costa Rican rural and urban areas. Am J Trop Med Hyg. 1984 Jan;33(1):24-9.

Mata L, Bolaños H, Pizarro D, Vives M. [Cryptosporidiosis in children of Costa Rica: a cross section and longitudinal study]. Rev Biol Trop. 1984 Jun;32(1):129-35.

Mateo M, Mateo M, Montoya A, Bailo B, Saugar JM, Aguilera M, Fuentes I, Carmena D. Detection and molecular characterization of Giardia duodenalis in children attending day care centers in Majadahonda, Madrid, Central Spain. Medicine (Baltimore). 2014 Oct;93(15):e75. doi: 10.1097/MD.0000000000000075.

Mathur MK, Verma AK, Makwana GE, Sinha M. Study of opportunistic intestinal parasitic infections in human immunodeficiency virus/acquired immunodeficiency syndrome patients. J Glob Infect Dis. 2013 Oct;5(4):164-7. doi: 10.4103/0974-777X.122012.

Matos O, Alves M, Xiao L, Cama V, Antunes F. *Cryptosporidium* felis and *C. meleagridis* in persons with HIV, Portugal. Emerg Infect Dis. 2004 Dec;10(12):2256-7.

Matthiesen SH, Hansen CM. Fast and non-toxic in situ hybridization without blocking of repetitive sequences. PLoS One. 2012;7(7):e40675. doi: 10.1371/journal.pone.0040675. Epub 2012 Jul 24.

Maurya PS, Sahu S, Sudhakar NR, Jaiswal V, Prashant DG, Rawat S, Verma H. Cryptosporidiosis in a buffalo calf at Meerut, Uttar Pradesh and its successful therapeutic management. J Parasit Dis. 2016 Dec;40(4):1583-1585. Epub 2016 Jan 13.

Mazurie, AJ, Alves, JM, Ozaki, LS, Zhou, S, Schwartz, DC, Buck, GA. Comparative Genomics of *Cryptosporidium*. International Journal of Genomics. 2013; 832756. doi:10.1155/2013/832756

Mbae CK, Nokes DJ, Mulinge E, Nyambura J, Waruru A, Kariuki S. Intestinal parasitic infections in children presenting with diarrhoea in outpatient and inpatient settings in an informal settlement of Nairobi, Kenya. BMC Infect Dis. 2013 May 27;13:243. doi: 10.1186/1471-2334-13-243.

Mbae CK, Mulinge E, Waruru A, Ngugi B, Wainaina J, Kariuki S. Genetic Diversity of *Cryptosporidium* in Children in an Urban Informal Settlement of Nairobi, Kenya. PLoS One. 2015 Dec 21;10(12):e0142055. doi: 10.1371/journal.pone.0142055. eCollection 2015.

McAuliffe GN, Anderson TP, Stevens M, Adams J, Coleman R, Mahagamasekera P, Young S, Henderson T, Hofmann M, Jennings LC, Murdoch DR. Systematic application of multiplex PCR enhances the detection of bacteria, parasites, and viruses in stool samples. J Infect. 2013 Aug;67(2):122-9. doi: 10.1016/j.jinf.2013.04.009. Epub 2013 Apr 18.

McCulloch R, Barry JD. A role for RAD51 and homologous recombination in Trypanosoma brucei antigenic variation. Genes Dev. 1999 Nov 1;13(21):2875-88.

McCulloch R, Morrison LJ, Hall JP. DNA Recombination Strategies During Antigenic Variation in the African Trypanosome. Microbiol Spectr. 2015 Apr;3(2):MDNA3-0016-2014. doi: 10.1128/microbiolspec.MDNA3-0016-2014.

McDougall, J. K. Cytomegaloviruses (1st ed.). 1990. Berlin: Springer-Verlag.

McGuigan C. *Cryptosporidium* outbreak after a visit to a wildlife centre in northeast Scotland: 62 confirmed cases. Euro Surveill. 2005 Apr 28;10(4):E050428.2.

McGuigan CC, Steven K, Pollock KG. Cryptosporidiosis associated with wildlife center, Scotland. Emerg Infect Dis. 2010 May;16(5):895-6. doi: 10.3201/eid1605.091468.

McKenzie RA, Green PE, Hartley WJ, Pollitt CC. *Cryptosporidium* in a red-bellied snake (Pseudechis porphyriacus). Aust Vet J. 1978 Jul;54(7):365-6.

McLauchlin J, Amar C, Pedraza-Díaz S, Nichols GL. Molecular epidemiological analysis of *Cryptosporidium spp*. in the United Kingdom: results of genotyping *Cryptosporidium spp*. in ,705 fecal samples from humans and 105 fecal samples from livestock animals. J Clin Microbiol. 2000 Nov;38(11):3984-90.

McMullan M, Gardiner A, Bailey K, Kemen E, Ward BJ, Cevik V, Robert-Seilaniantz A, Schultz-Larsen T, Balmuth A, Holub E, van Oosterhout C, Jones JD. Evidence for suppression of immunity as a driver for genomic introgressions and host range expansion in races of Albugo candida, a generalist parasite. Elife. 2015 Feb 27;4. doi: 10.7554/eLife.04550.

Meamar AR, Guyot K, Certad G, Dei-Cas E, Mohraz M, Mohebali M, Mohammad K, Mehbod AA, Rezaie S, Rezaian M. Molecular characterization of *Cryptosporidium* isolates from humans and animals in Iran. Appl Environ Microbiol. 2007 Feb;73(3):1033-5. Epub 2006 Dec 1.

Medeiros MI, Neme SN, da Silva P, Capuano DM, Errera MC, Fernandes SA, do Valle GR, de Avila FA. Etiology of acute diarrhea among children in Ribeirão Preto-SP, Brazil. Rev Inst Med Trop Sao Paulo. 2001 Jan-Feb;43(1):21-4.

Mehta KD, Vacchani A, Mistry MM, Kavathia GU, Goswami YS. To Study the Prevalence of Various Enteric Parasitic Infections Among HIV Infected Individuals in the P.D.U. Medical College and Hospital, Rajkot, Gujarat, India. J Clin Diagn Res. 2013 Jan;7(1):58-60. doi: 10.7860/JCDR/2012/4864.2670. Epub 2012 Nov 8.

Meisel JL, Perera DR, Meligro C, Rubin CE. Overwhelming watery diarrhea associated with a *Cryptosporidium* in an immunosuppressed patient. Gastroenterology. 1976 Jun;70(6):1156-60.

Mejia R, Vicuña Y, Broncano N, Sandoval C, Vaca M, Chico M, Cooper PJ, Nutman TB. A novel, multi-parallel, real-time polymerase chain reaction approach for eight gastrointestinal parasites provides improved diagnostic capabilities to resource-limited at-risk populations. Am J Trop Med Hyg. 2013 Jun;88(6):1041-7. doi: 10.4269/ajtmh.12-0726. Epub 2013 Mar 18.

Mengelle C, Mansuy JM, Prere MF, Grouteau E, Claudet I, Kamar N, Huynh A, Plat G, Benard M, Marty N, Valentin A, Berry A, Izopet J. Simultaneous detection of gastrointestinal pathogens with a multiplex Luminex-based molecular assay in stool samples from diarrhoeic patients. Clin Microbiol Infect. 2013 Oct;19(10):E458-65. doi: 10.1111/1469-0691.12255. Epub 2013 May 28.

Mengist HM, Taye B, Tsegaye A. Intestinal parasitosis in relation to CD4+T cells levels and anemia among HAART initiated and HAART naive pediatric HIV patients in a Model ART center in Addis Ababa, Ethiopia. PLoS One. 2015 Feb 6;10(2):e0117715. doi: 10.1371/journal.pone.0117715. eCollection 2015.

Menon BS, Abdullah S, Mahamud F, Morgan UM, Malik AS, Choo KE, Singh B. Low prevalence of *Cryptosporidium parvum* in hospitalized children in Kota Bharu, Malaysia. Southeast Asian J Trop Med Public Health. 2001 Jun;32(2):319-22.

Mergani MH, Mohammed MA, Khan N, Bano M, Khan AH. Detection of intestinal protozoa by using different methods. Dent Med Res [serial online] 2014 [cited 2017 Mar 6];2:28-32

Miller SA, Rosario CL, Rojas E, Scorza JV. Intestinal parasitic infection and associated symptoms in children attending day care centres in Trujillo, Venezuela. Trop Med Int Health. 2003 Apr;8(4):342-7.

Minchin, EA. An introduction to the study of the protozoa, with special reference to the parasitic forms. London: Edward Arnold. 1922. Print.

Minot S, Melo MB, Li F, Lu D, Niedelman W, Levine SS, Saeij JP. Admixture and recombination among Toxoplasma gondii lineages explain global genome diversity. Proc Natl Acad Sci U S A. 2012 Aug 14;109(33):13458-63. doi: 10.1073/pnas.1117047109. Epub 2012 Jul 30.

Minta DK, Dembélé M, Dolo A, Sidibé AT, Diarra AS, Konaté A, Diarra M, Diakité A, Sidibé AF, Traoré AK, Maiga MY, Pichard E, Traoré HA, Doumbo O. [Digestive parasitic diseases to HIV/AIDS infected patients of internal medicine and infectious diseases wards of the hôpital du Point "G" Bamako - Mali]. Mali Med. 2007;22(1):33-6.

Mirzaei M. Prevalence of *Cryptosporidium* sp. infection in diarrheic and non-diarrheic humans in Iran. Korean J Parasitol. 2007 Jun;45(2):133-7.

Missaye A, Dagnew M, Alemu A, Alemu A. Prevalence of intestinal parasites and associated risk factors among HIV/AIDS patients with pre-ART and on-ART attending dessie hospital ART clinic, Northeast Ethiopia. AIDS Res Ther. 2013 Feb 25;10(1):7. doi: 10.1186/1742-6405-10-7.

Mitra S, Mukherjee A, Khanra D, Bhowmik A, Roy K, Talukdar A. Enteric Parasitic Infection Among Antiretroviral Therapy Naïve HIV-Seropositive People: Infection Begets Infection-Experience from Eastern India. J Glob Infect Dis. 2016 Apr-Jun;8(2):82-6. doi: 10.4103/0974-777X.182124.

Mitsui-Sekinaka K, Imai K, Sato H, Tomizawa D, Kajiwara M, Nagasawa M, Morio T, Nonoyama S. Clinical features and hematopoietic stem cell transplantations for CD40 ligand deficiency in Japan. J Allergy Clin Immunol. 2015 Oct;136(4):1018-24. doi: 10.1016/j.jaci.2015.02.020. Epub 2015 Mar 31.

Mladenova Z, Steyer A, Steyer AF, Ganesh B, Petrov P, Tchervenjakova T, Iturriza-Gomara M. Aetiology of acute paediatric gastroenteritis in Bulgaria during summer months: prevalence of viral infections. J Med Microbiol. 2015 Mar;64(Pt 3):272-82. doi: 10.1099/jmm.0.000018. Epub 2015 Jan 16.

Moghaddam AA. Symptomatic and asymptomatic cryptosporidiosis in young children in Iran. Pak J Biol Sci. 2007 Apr 1;10(7):1108-12.

Mohamed A, Mohammed A, Ali A, Abrika O, Masoud A, Khan A. Prevalence of Intestinal Parasites among HIV/AIDS Patients in South Western, Libya. Int. J. Appl. Med. Biol. Res. Vol 2 (1), 2017: p 16-21

Mohammadpour I, Bozorg-Ghalati F, Motazedian MH1. Molecular characterization and phylogenetic analysis of Microsporidia and *Cryptosporidium spp*. in patients with multiple bowel biopsies from Fars Province, Iran Ann Parasitol. 2016;62(4):321-330. doi: 10.17420/ap6204.68.

Mohandas, Sehgal R, Sud A, Malla N. Prevalence of intestinal parasitic pathogens in HIV-seropositive individuals in Northern India. Jpn J Infect Dis. 2002 Jun;55(3):83-4.

Mohanty I, Panda P, Sahu S, Dash M, Narasimham MV, Padhi S, Parida B. Prevalence of isosporiasis in relation to CD4 cell counts among HIV-infected patients with diarrhea in Odisha, India. Adv Biomed Res. 2013 Jul 30;2:61. doi: 10.4103/2277-9175.115810. eCollection 2013.

Molloy SF, Smith HV, Kirwan P, Nichols RA, Asaolu SO, Connelly L, Holland CV. Identification of a high diversity of *Cryptosporidium* species genotypes and subtypes in a pediatric population in Nigeria. Am J Trop Med Hyg. 2010 Apr;82(4):608-13. doi: 10.4269/ajtmh.2010.09-0624.

Molloy SF, Tanner CJ, Kirwan P, Asaolu SO, Smith HV, Nichols RA, Connelly L, Holland CV. Sporadic *Cryptosporidium* infection in Nigerian children: risk factors with species identification. Epidemiol Infect. 2011 Jun;139(6):946-54. doi: 10.1017/S0950268810001998. Epub 2010 Aug 27.

Mondal D, Petri WA Jr, Sack RB, Kirkpatrick BD, Haque R. Entamoeba histolytica-associated diarrheal illness is negatively associated with the growth of preschool children: evidence from a prospective study. Trans R Soc Trop Med Hyg. 2006 Nov;100(11):1032-8. Epub 2006 May 30.

Mondal D, Minak J, Alam M, Liu Y, Dai J, Korpe P, Liu L, Haque R, Petri WA Jr. Contribution of enteric infection, altered intestinal barrier function, and maternal malnutrition to infant malnutrition in Bangladesh. Clin Infect Dis. 2012 Jan 15;54(2):185-92. doi: 10.1093/cid/cir807. Epub 2011 Nov 21.

Monerawela C, James TC, Wolfe KH, Bond U. Loss of lager specific genes and subtelomeric regions define two different Saccharomyces cerevisiae lineages for Saccharomyces pastorianus Group I and II strains. FEMS Yeast Res. 2015 Mar;15(2). pii: fou008. doi: 10.1093/femsyr/fou008.

Moore CE, Elwin K, Phot N, Seng C, Mao S, Suy K, Kumar V, Nader J, Bousfield R, Perera S, Bailey JW, Beeching NJ, Day NP, Parry CM, Chalmers RM. Molecular Characterization of *Cryptosporidium* Species and Giardia duodenalis from Symptomatic Cambodian Children. PLoS Negl Trop Dis. 2016 Jul 7;10(7):e0004822. doi: 10.1371/journal.pntd.0004822. eCollection 2016.

Moore JE, Crothers L, Millar BC, Crothers E, Rooney PJ, Xiao L, Dooley JS, Lowery CJ. Low incidence of concurrent enteric infection associated with sporadic and outbreak-related human cryptosporidiosis in Northern Ireland. J Clin Microbiol. 2002 Aug;40(8):3107-8.

Mor SM, Tumwine JK, Ndeezi G, Srinivasan MG, Kaddu-Mulindwa DH, Tzipori S, Griffiths JK. Respiratory cryptosporidiosis in HIV-seronegative children in Uganda: potential for respiratory transmission. Clin Infect Dis. 2010 May 15;50(10):1366-72. doi: 10.1086/652140.

Moran P, Caras IW. Requirements for glycosylphosphatidylinositol attachment are similar but not identical in mammalian cells and parasitic protozoa. J Cell Biol. 1994 Apr;125(2):333-43.

Morgan UM, Constantine CC, Forbes DA, Thompson RC. Differentiation between human and animal isolates of *Cryptosporidium parvum* using rDNA sequencing and direct PCR analysis. J Parasitol. 1997 Oct;83(5):825-30.

Morgan UM, Weber R, Xiao L, Sulaiman I, Thompson RC, Ndiritu W, Lal A, Moore A, Deplazes P. Molecular characterization of *Cryptosporidium* isolates obtained from human immunodeficiency virus-infected individuals living in Switzerland, Kenya, and the United States. J Clin Microbiol. 2000 Mar;38(3):1180-3.

Morgan UM, Monis PT, Xiao L, Limor J, Sulaiman I, Raidal S, O'Donoghue P, Gasser R, Murray A, Fayer R, Blagburn BL, Lal AA, Thompson RC. Molecular and phylogenetic characterisation of *Cryptosporidium* from birds. Int J Parasitol. 2001 Mar;31(3):289-96.

Morgan-Ryan UM, Fall A, Ward LA, Hijjawi N, Sulaiman I, Fayer R, Thompson RC, Olson M, Lal A, Xiao L. *Cryptosporidium hominis* n. sp. (Apicomplexa: Cryptosporidiidae) from Homo sapiens. J Eukaryot Microbiol. 2002 Nov-Dec;49(6):433-40.

Morrison LJ, Mallon ME, Smith HV, MacLeod A, Xiao L, Tait A. The population structure of the *Cryptosporidium parvum* population in Scotland: a complex picture. Infect Genet Evol. 2008 Mar;8(2):121-9. Epub 2007 Nov 9.

Morrison LJ, Marcello L, McCulloch R. Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. Cell Microbiol. 2009 Dec;11(12):1724-34. doi: 10.1111/j.1462-5822.2009.01383.x. Epub 2009 Sep 14.

Morse TD, Nichols RA, Grimason AM, Campbell BM, Tembo KC, Smith HV. Incidence of cryptosporidiosis species in paediatric patients in Malawi. Epidemiol Infect. 2007 Nov;135(8):1307-15. Epub 2007 Jan 15.

Mosli M, Gregor J, Chande N, Lannigan R. Nonutility of routine testing of stool for ova and parasites in a tertiary care Canadian centre. Can J Microbiol. 2012 May;58(5):653-9. doi: 10.1139/w2012-039. Epub 2012 Apr 27.

Mousa N, Abdel-Razik A, El-Nahas H, El-Shazly A, Abdelaziz M, Nabih M, Hamed M, Eissa M, Effat N, Eldars W. Cryptosporidiosis in patients with diarrhea and chronic liver diseases. J Infect Dev Ctries. 2014 Dec 15;8(12):1584-90. doi: 10.3855/jidc.5166.

Moyo SJ, Gro N, Matee MI, Kitundu J, Myrmel H, Mylvaganam H, Maselle SY, Langeland N. Age specific aetiological agents of diarrhoea in hospitalized children aged less than five years in Dar es Salaam, Tanzania. BMC Pediatr. 2011 Feb 23;11:19. doi: 10.1186/1471-2431-11-19.

Mukherjee AK, Chowdhury P, Bhattacharya MK, Ghosh M, Rajendran K, Ganguly S. Hospital-based surveillance of enteric parasites in Kolkata. BMC Res Notes. 2009 Jun 19;2:110. doi: 10.1186/1756-0500-2-110.

Mukhopadhyay C, Wilson G, Pradhan D, Shivananda PG. Intestinal protozoal infestation profile in persistent diarrhea in children below age 5 years in western Nepal. Southeast Asian J Trop Med Public Health. 2007 Jan;38(1):13-9.

Muñoz-Antoli C, Pavón A, Marcilla A, Toledo R, Esteban JG. Prevalence and molecular characterization of *Cryptosporidium* in schoolchildren from department of Rio San Juan (Nicaragua). Trop Biomed. 2011 Apr;28(1):40-7.

Musa HA, Salim GA, Ismael AA, Elfadil A, Kafi S. Diarrhea due to *Cryptosporidium parvum* in immunocompromised and immunocompetent patients in Khartoum State. Volume , Number 1 (March 2011) , 39-42.

Muthusamy D, Rao SS, Ramani S, Monica B, Banerjee I, Abraham OC, Mathai DC, Primrose B, Muliyil J, Wanke CA, Ward HD, Kang G. Multilocus genotyping of *Cryptosporidium* sp. isolates from human immunodeficiency virus-infected individuals in South India. J Clin Microbiol. 2006 Feb;44(2):632-4.

Nacro B, Bonkoungou P, Nagalo K, Tall FR, Curtis V. [Clinical profile of cryptosporidiosis in a pediatric hospital environment in Burkina Faso]. Med Trop (Mars). 1998;58(1):47-50.

Nadjm B, Behrens RH. Malaria: an update for physicians. Infect Dis Clin North Am. 2012 Jun;26(2):243-59. doi: 10.1016/j.idc.2012.03.010.

Nagamani K, Pavuluri PR, Gyaneshwari M, Prasanthi K, Rao MI, Saxena NK. Molecular characterisation of *Cryptosporidium*: an emerging parasite. Indian J Med Microbiol. 2007 Apr;25(2):133-6.

Nahrevanian H, Assmar M, Samin MG. Cryptosporidiosis among immunocompetent patients with gastroenteritis in Iran: a comparison with other enteropathogenic parasites. J Microbiol Immunol Infect. 2007 Apr;40(2):154-6.

Nair P, Mohamed JA, DuPont HL, Figueroa JF, Carlin LG, Jiang ZD, Belkind-Gerson J, Martinez-Sandoval FG, Okhuysen PC. Epidemiology of cryptosporidiosis in North American travelers to Mexico. Am J Trop Med Hyg. 2008 Aug;79(2):210-4.

Nakamura AA, Meireles MV. *Cryptosporidium* infections in birds--a review. Rev Bras Parasitol Vet. 2015 Jul-Sep;24(3):253-67. doi: 10.1590/S1984-29612015063.

Nascimento WR, Cavalcanti IM, Irmão JI, Rocha FJ. [Presence of *Cryptosporidium spp* in children with acute diarrhea in a public daycare center in Recife, State of Pernambuco]. Rev Soc Bras Med Trop. 2009 Mar-Apr;42(2):175-8.

Nassar SA, Oyekale TO, Oluremi AS. Prevalence of *Cryptosporidium* infection and related risk factors in children in Awo and Iragberi, Nigeria. J Immunoassay Immunochem. 2017;38(1):2-9. doi: 10.1080/15321819.2016.1178652. Epub 2016 May 9.

Nath J, Hussain G, Singha B, Paul J, Ghosh SK. Burden of major diarrheagenic protozoan parasitic co-infection among amoebic dysentery cases from North East India: a case report. Parasitology. 2015 Sep;142(10):1318-25.

Natividad FF, Buerano CC, Lago CB, Mapua CA, de Guzman BB, Seraspe EB, Samentar LP, Endo T. Prevalence rates of Giardia and *Cryptosporidium* among diarrheic patients in the Philippines. Southeast Asian J Trop Med Public Health. 2008 Nov;39(6):991-9.

Navarro-i-Martinez L, da Silva AJ, Botero Garces JH, Montoya Palacio MN, del Aguila C, Bornay-Llinares FJ. Cryptosporidiosis in HIV-positive patients from Medellín, Colombia. J Eukaryot Microbiol. 2006;53 Suppl 1:S37-9.

Navarro-I-Martinez L, da Silva AJ, Llovo Taboada J, Del Águila C, Pieniazek NJ, Bornay-Llinares FJ. [Molecular characterization of *Cryptosporidium spp*. isolated in humans in two different locations in Spain]. Enferm Infecc Microbiol Clin. 2013 Oct;31(8):506-10. doi: 10.1016/j.eimc.2012.11.005. Epub 2012 Dec 27.

Nazeer JT, El Sayed Khalifa K, von Thien H, El-Sibaei MM, Abdel-Hamid MY, Tawfik RA, Tannich E. Use of multiplex real-time PCR for detection of common diarrhea causing protozoan parasites in Egypt. Parasitol Res. 2013 Feb;112(2):595-601. doi: 10.1007/s00436-012-3171-8. Epub 2012 Nov 1.

Nazemalhosseini-Mojarad E, Haghighi A, Taghipour N, Keshavarz A, Mohebi SR, Zali MR, Xiao L. Subtype analysis of *Cryptosporidium parvum* and *Cryptosporidium hominis* isolates from humans and cattle in Iran. Vet Parasitol. 2011 Jun 30;179(1-3):250-2. doi: 10.1016/j.vetpar.2011.01.051. Epub 2011 Feb 2.

Neagoe, IM, Miclos, S, Popescu, D, Savastru, D, Niculescu, VIR, Damian, M, Lazar, L, Dontu, S. and Tautan, M. Wavelet spectrogram - based DNA analysis for the assessment of *Cryptosporidium spp*. Gp60 subgenotypes variation. Optoelectronics and advanced materials-rapid communications. 2014 July; 8(7):814-819

Neelsen, F. Ein casuistischer Beitrag zur Lehre von der Tuberkulose. Centralblatt fiir die medicinischen Wissenschaften. 1883; No. 28, p. 497.

Nei, M. Molecular Evolutionary Genetics. 1987. Columbia University Press, New York, NY.

Nei M, Kumar S. Molecular Evolution and Phylogenetics. 2000. Oxford University Press, New York.

Neira OP, Muñoz SN, Wilson LG, Barthel M ME, Rosales L MJ, Henríquez RC. [*Cryptosporidium* species in immunodeficient and immunocompetent patients of Valparaíso: a descriptive study]. Rev Chilena Infectol. 2012 Feb;29(1):63-71. doi: 10.4067/S0716-10182012000100011. Epub 2012 Apr 10.

Neira-Otero P, Muñoz-Saldías N, Sanchez-Moreno M, Rosales-Lombardo MJ. Molecular characterization of *Cryptosporidium* species and genotypes in Chile. Parasitol Res. 2005 Aug;97(1):63-7. Epub 2005 Jun 11.

Nel ED, Rabie H, Goodway J, Cotton MF. A retrospective study of cryptosporidial diarrhea in a region with high HIV prevalence. J Trop Pediatr. 2011 Aug;57(4):289-92. doi: 10.1093/tropej/fmq094. Epub 2010 Oct 14.

Neres-Norberg A, Guerra-Sanches F, Blanco Moreira-Norberg PR, Madeira-Oliveira JT, Santa-Helena AA, Serra-Freire NM. [Intestinal Parasitism in Terena Indigenous People of the Province of Mato Grosso do Sul, Brazil]. Rev Salud Publica (Bogota). 2014 Nov-Dec;16(6):859-70.

Nétor Velásquez J, Marta E, Alicia di Risio C, Etchart C, Gancedo E, Victor Chertcoff A, Bruno Malandrini J, Germán Astudillo O, Carnevale S. Molecular identification of protozoa causing AIDS-associated cholangiopathy in Buenos Aires, Argentina. Acta Gastroenterol Latinoam. 2012 Dec;42(4):301-8.

Neves RF, Fernandes AC, Meyer-Fernandes JR, Souto-Padrón T. Trypanosoma cruzi-secreted vesicles have acid and alkaline phosphatase activities capable of increasing parasite adhesion and infection. Parasitol Res. 2014 Aug;113(8):2961-72. doi: 10.1007/s00436-014-3958-x. Epub 2014 Jun 7.

Ng J, Eastwood K, Durrheim D, Massey P, Walker B, Armson A, Ryan U. Evidence supporting zoonotic transmission of *Cryptosporidium* in rural New South Wales. Exp Parasitol. 2008 May;119(1):192-5. doi: 10.1016/j.exppara.2008.01.010. Epub 2008 Feb 2.

Ng J, MacKenzie B, Ryan U. Longitudinal multi-locus molecular characterisation of sporadic Australian human clinical cases of cryptosporidiosis from 2005 to 2008. Exp Parasitol. 2010 Aug;125(4):348-56. doi: 10.1016/j.exppara.2010.02.017. Epub 2010 Mar 4.

Ng JS, Eastwood K, Walker B, Durrheim DN, Massey PD, Porigneaux P, Kemp R, McKinnon B, Laurie K, Miller D, Bramley E, Ryan U. Evidence of *Cryptosporidium* transmission between cattle and humans in northern New South Wales. Exp Parasitol. 2012 Apr;130(4):437-41. doi: 10.1016/j.exppara.2012.01.014. Epub 2012 Feb 7.

Ngindu A, Kamar K, Choge A, Maritim A, Akoru C, Diero L, Smith J, Bartlett M. Survey of faecal parasites in patients from western Kenya. J Egypt Soc Parasitol. 2002 Apr;32(1):1-7.

Nhampossa T, Mandomando I, Acacio S, Quintó L, Vubil D, Ruiz J, Nhalungo D, Sacoor C, Nhabanga A, Nhacolo A, Aide P, Machevo S, Sigaúque B, Nhama A, Kotloff K, Farag T, Nasrin D, Bassat Q, Macete E, Levine MM, Alonso P. Diarrheal Disease in Rural Mozambique: Burden, Risk Factors and Etiology of Diarrheal Disease among Children Aged 0-59 Months Seeking Care at Health Facilities. PLoS One. 2015 May 14;10(5):e0119824. doi: 10.1371/journal.pone.0119824. eCollection 2015.

Niang M, Bei AK, Madnani KG, Pelly S, Dankwa S, Kanjee U, Gunalan K, Amaladoss A, Yeo KP, Bob NS, Malleret B, Duraisingh MT, Preiser PR. STEVOR is a Plasmodium falciparum erythrocyte binding protein that mediates merozoite invasion and rosetting. Cell Host Microbe. 2014 Jul 9;16(1):81-93. doi: 10.1016/j.chom.2014.06.004.

Nichols R. Gene trees and species trees are not the same. Trends Ecol Evol. 2001 Jul 1;16(7):358-364.

Nime FA, Burek JD, Page DL, Holscher MA, Yardley JH. Acute enterocolitis in a human being infected with the protozoan *Cryptosporidium*. Gastroenterology. 1976 Apr;70(4):592-8.

Nimri LF. Cyclospora cayetanensis and other intestinal parasites associated with diarrhea in a rural area of Jordan. Int Microbiol. 2003 Jun;6(2):131-5. Epub 2003 May 29.

Nimri LF, Meqdam M. Enteropathogens associated with cases of gastroenteritis in a rural population in Jordan. Clin Microbiol Infect. 2004 Jul;10(7):634-9.

Nizeyi JB, Sebunya D, Dasilva AJ, Cranfield MR, Pieniazek NJ, Graczyk TK. Cryptosporidiosis in people sharing habitats with free-ranging mountain gorillas (Gorilla gorilla beringei), Uganda. Am J Trop Med Hyg. 2002 Apr;66(4):442-4.

Nolan MJ, Jex AR, Koehler AV, Haydon SR, Stevens MA, Gasser RB. Molecular-based investigation of *Cryptosporidium* and Giardia from animals in water catchments in southeastern Australia. Water Res. 2013 Apr 1;47(5):1726-40. doi: 10.1016/j.watres.2012.12.027. Epub 2012 Dec 27.

Noor R, Saha SR, Rahman F, Munsi SK, Uddin MA, Rahman MM. Frequency of opportunistic and other intestinal parasitic infections in patients infected with human immunodeficiency virus in Bangladesh. Tzu Chi Med J. 2012;24:191–5.

Norhayati M, Fatmah MS, Yusof S, Edariah AB. Intestinal parasitic infections in man: a review. Med J Malaysia. 2003 Jun;58(2):296-305; quiz 306.

Nouri M, Moghadam A, Haghighatnia H. *Cryptosporidium* infection in human diarrhea patients in West Azerbaijan, Iran. Med J Islam Repub Iran. 1991;2:35–38.

Nsagha DS, Njunda AL, Assob NJ, Ayima CW, Tanue EA, Kibu OD, Kwenti TE. Intestinal parasitic infections in relation to CD4(+) T cell counts and diarrhea in HIV/AIDS patients with or without antiretroviral therapy in Cameroon. BMC Infect Dis. 2016 Jan 11;16:9. doi: 10.1186/s12879-016-1337-1.

Nuchjangreed C, Boonrod K, Ongerth J, Karanis P. Prevalence and molecular characterization of human and bovine *Cryptosporidium* isolates in Thailand. Parasitol Res. 2008 Nov;103(6):1347-53. doi: 10.1007/s00436-008-1139-5. Epub 2008 Aug 16.

Núñez FA, González OM, Bravo JR, Escobedo AA, Gonzaléz I. [Intestinal parasitosis in children admitted to the Pediatric Teaching Hospital of Cerro, Havana City, Cuba]. Rev Cubana Med Trop. 2003 Jan-Apr;55(1):19-26.

Núñez FA, González OM, González I, Escobedo AA, Cordoví RA. Intestinal coccidia in Cuban pediatric patients with diarrhea. Mem Inst Oswaldo Cruz. 2003 Jun;98(4):539-42. Epub 2003 Aug 18.

Nurminen N, Juuti R, Oikarinen S, Fan YM, Lehto KM, Mangani C, Maleta K, Ashorn P, Hyöty H. High-throughput multiplex quantitative polymerase chain reaction method for Giardia lamblia and *Cryptosporidium* species detection in stool samples. Am J Trop Med Hyg. 2015 Jun;92(6):1222-6. doi: 10.4269/ajtmh.15-0054. Epub 2015 Apr 27

Nwabuisi C. Childhood cryptosporidiosis and intestinal parasitosis in association with diarrhoea in Kwara State, Nigeria. West Afr J Med. 2001 Apr-Jun;20(2):165-8.

Nyamwange CI, Mkoji G, Mpoke S, Nyandieka HS. CRYPTOSPORIDIOSIS AND ITS GENOTYPES AMONG CHILDREN ATTENDING MOI TEACHING AND REFERRAL HOSPITAL IN ELDORET, KENYA. East Afr Med J. 2012 Jan;89(1):11-9.

Obana M, Sagara H, Aoki T, Kim R, Takizawa Y, Tsunoda T, Irimajiri S, Yamashita K. [The current status of infectious enteritis in Japan--reports of the "Research Group for Infectious Enteric Diseases, Japan" in the last 5 years (1996-2000)]. Kansenshogaku Zasshi. 2002 May;76(5):355-68.

Oberli A, Slater LM, Cutts E, Brand F, Mundwiler-Pachlatko E, Rusch S, Masik MF, Erat MC, Beck HP, Vakonakis I. A Plasmodium falciparum PHIST protein binds the virulence factor PfEMP1 and comigrates to knobs on the host cell surface. FASEB J. 2014 Oct;28(10):4420-33. doi: 10.1096/fj.14-256057. Epub 2014 Jun 30.

O'Brien E, McInnes L, Ryan U. *Cryptosporidium* GP60 genotypes from humans and domesticated animals in Australia, North America and Europe. Exp Parasitol. 2008 Jan;118(1):118-21. Epub 2007 Jun 12.

O'Connor RM, Burns PB, Ha-Ngoc T, Scarpato K, Khan W, Kang G, Ward H. Polymorphic mucin antigens CpMuc4 and CpMuc5 are integral to *Cryptosporidium parvum* infection in vitro. Eukaryot Cell. 2009 Apr;8(4):461-9. doi: 10.1128/EC.00305-08. Epub 2009 Jan 23.

O'Connor RM, Shaffie R, Kang G, Ward HD. Cryptosporidiosis in patients with HIV/AIDS. AIDS. 2011 Mar 13;25(5):549-60. doi: 10.1097/QAD.0b013e3283437e88.

Odiit M, Kansiime F, Enyaru JC. Duration of symptoms and case fatality of sleeping sickness caused by Trypanosoma brucei rhodesiense in Tororo, Uganda. East Afr Med J. 1997 Dec;74(12):792-5.

Oguntibeju OO. Prevalence of Intestinal Parasites In HIV-Positive/AIDS Patients. Malays J Med Sci. 2006 Jan;13(1):68-73.

Ojuromi OT, Izquierdo F, Fenoy S, Fagbenro-Beyioku A, Oyibo W, Akanmu A, Odunukwe N, Henriques-Gil N, del Aguila C. Identification and characterization of microsporidia from fecal samples of HIV-positive patients from Lagos, Nigeria. PLoS One. 2012;7(4):e35239. doi: 10.1371/journal.pone.0035239. Epub 2012 Apr 9.

Ojuromi OT, Duan L, Izquierdo F, Fenoy SM, Oyibo WA, Del Aguila C, Ashafa AO, Feng Y, Xiao L. Genotypes of *Cryptosporidium spp*. and Enterocytozoon bieneusi in Human Immunodeficiency Virus-Infected Patients in Lagos, Nigeria. J Eukaryot Microbiol. 2016 Jul;63(4):414-8. doi: 10.1111/jeu.12285. Epub 2015 Dec 28.

Ojurongbe O, Raji OA, Akindele AA, Kareem MI, Adefioye OA, Adeyeba AO. *Cryptosporidium* and other enteric parasitic infections in HIV-seropositive individuals with and without diarrhoea in Osogbo, Nigeria. Br J Biomed Sci. 2011;68(2):75-8.

Okhuysen PC, Rich SM, Chappell CL, Grimes KA, Widmer G, Feng X, Tzipori S. Infectivity of a Cryptosporidium parvum isolate of cervine origin for healthy adults and interferon-gamma knockout mice. J Infect Dis. 2002 May 1;185(9):1320-5. Epub 2002 Apr 16.

Olesen B, Neimann J, Böttiger B, Ethelberg S, Schiellerup P, Jensen C, Helms M, Scheutz F, Olsen KE, Krogfelt K, Petersen E, Mølbak K, Gerner-Smidt P. Etiology of diarrhea in young children in Denmark: a case-control study. J Clin Microbiol. 2005 Aug;43(8):3636-41.

Oliver SG, van der Aart QJ, Agostoni-Carbone ML, Aigle M, Alberghina L, Alexandraki D, Antoine G, Anwar R, Ballesta JP, Benit P, et al. The complete DNA sequence of yeast chromosome III. Nature. 1992 May 7;357(6373):38-46.

Omata Y, Taka A, Terada K, Koyama T, Kanda M, Saito A, Dubey JP. Isolation of coccidian enteroepithelial stages of Toxoplasma gondii from the intestinal mucosa of cats by Percoll density-gradient centrifugation. Parasitol Res. 1997;83(6):574-7.

Ormeci B, Linden KG. Development of a fluorescence in situ hybridization protocol for the identification of micro-organisms associated with wastewater particles and flocs. J Environ Sci Health A Tox Hazard Subst Environ Eng. 2008 Nov;43(13):1484-8. doi: 10.1080/10934520802293586.

Omrani VF, Fallahi Sh, Rostami A, Siyadatpanah A, Barzgarpour G, Mehravar S, Memari F, Hajialiani F, Joneidi Z. Prevalence of intestinal parasite infections and associated clinical symptoms among patients with end-stage renal disease undergoing hemodialysis. Infection. 2015 Oct;43(5):537-44. doi: 10.1007/s15010-015-0778-6. Epub 2015 Apr 14.

Ong CSL, Chow S, So PPL, Chen R, Xiao L, Sulaiman I et al. Identification of two different *Cryptosporidium hominis* subtypes from cases in the 2001 waterborne cryptosporidiosis outbreak in North Battleford, Saskatchewan. In: Robertson W, Brooks T, editors. Proceedings of the 11th Canadian National Conference and 2nd Policy Forum on Drinking Water; Calgary, Alberta; Apr 2004. Ottawa (Canada): Canadian Water and Wastewater Assoc., Health Canada & Alberta Environment; 2004. p. 628–39.

Ono K, Rai SK, Chikahira M, Fujimoto T, Shibata H, Wada Y, Tsuji H, Oda Y, Rai G, Shrestha CD, Masuda K, Shrestha HG, Matsumura T, Hotta H, Kawamura T, Uga S. Seasonal distribution of enteropathogens detected from diarrheal stool and water samples collected in Kathmandu, Nepal. Southeast Asian J Trop Med Public Health. 2001 Sep;32(3):520-6.

Opintan JA, Newman MJ, Ayeh-Kumi PF, Affrim R, Gepi-Attee R, Sevilleja JE, Roche JK, Nataro JP, Warren CA, Guerrant RL. Pediatric diarrhea in southern Ghana: etiology and association with intestinal inflammation and malnutrition. Am J Trop Med Hyg. 2010 Oct;83(4):936-43. doi: 10.4269/ajtmh.2010.09-0792.

Oram, PG Mr. Brian. "*Cryptosporidium parvum* Drinking Water Protozoan." 2014. *Cryptosporidium parvum* Drinking Water Protozoan. Water Research Centre, n.d. Web. 15 Oct. 2016.

Oshiro ET, Dorval ME, Nunes VL, Silva MA, Said LA. [Prevalence of *Cryptosporidium parvum* among children of less than 5 years of age in the urban zone of Campo Grande, Mato Grosso do Sul State, Brazil, 1996]. Rev Soc Bras Med Trop. 2000 May-Jun;33(3):277-80.

Osman M, El Safadi D, Benamrouz S, Guyot K, Dei-Cas E, Aliouat el M, Creusy C, Mallat H, Hamze M, Dabboussi F, Viscogliosi E, Certad G. Initial data on the molecular epidemiology of cryptosporidiosis in Lebanon. PLoS One. 2015 May 7;10(5):e0125129. doi: 10.1371/journal.pone.0125129. eCollection 2015.

Osman M, El Safadi D, Cian A, Benamrouz S, Nourrisson C, Poirier P, Pereira B, Razakandrainibe R, Pinon A, Lambert C, Wawrzyniak I, Dabboussi F, Delbac F, Favennec L, Hamze M, Viscogliosi E, Certad G. Prevalence and Risk Factors for Intestinal Protozoan Infections with *Cryptosporidium*, Giardia, Blastocystis and Dientamoeba among Schoolchildren in Tripoli, Lebanon. PLoS Negl Trop Dis. 2016 Mar 14;10(3):e0004496. doi: 10.1371/journal.pntd.0004496. eCollection 2016.

Otağ F, Aslan G, Emekdaş G, Aydin E, Ozkan AT, Ceber K. [Investigation of *Cryptosporidium* oocysts in elementary school students in Mersin]. Turkiye Parazitol Derg. 2007;31(1):17-9.

Oyedeji OA, Adejuyigbe E, Oninla SO, Akindele AA, Adedokun SA, Agelebe E. Intestinal Parasitoses in HIV Infected Children in a Nigerian Tertiary Hospital. J Clin Diagn Res. 2015 Nov;9(11):SC01-5. doi: 10.7860/JCDR/2015/12537.6736. Epub 2015 Nov 1.

Paboriboune P, Phoumindr N, Borel E, Sourinphoumy K, Phaxayaseng S, Luangkhot E, Sengphilom B, Vansilalom Y, Odermatt P, Delaporte E, Etard JF, Rabodonirina M. Intestinal parasitic infections in HIV-infected patients, Lao People's Democratic Republic. PLoS One. 2014 Mar 24;9(3):e91452. doi: 10.1371/journal.pone.0091452. eCollection 2014.

Paje-Villar E, Co BG, Caradang EH, Raymundo A, Lagamayo E, Lavadia E, Rigor E, Hart CA. Non-bacterial diarrhoea in children in the Philippines. Ann Trop Med Parasitol. 1994 Feb;88(1):53-8.

Palit A, Sur D, MitraDhar K, Saha MR. Asymptomatic cryptosporiosis in a periurban slum setting in Kolkata, India--a pilot study. Jpn J Infect Dis. 2005 Apr;58(2):110-1.

Palmer CJ, Xiao L, Terashima A, Guerra H, Gotuzzo E, Saldías G, Bonilla JA, Zhou L, Lindquist A, Upton SJ. *Cryptosporidium muris*, a rodent pathogen, recovered from a human in Perú. Emerg Infect Dis. 2003 Sep;9(9):1174-6.

Paluszynski J, Monahan Z, Williams M, Lai O, Morris C, Burns P, O'Connor R. Biochemical and functional characterization of CpMuc4, a *Cryptosporidium* surface antigen that binds to host epithelial cells. Mol Biochem Parasitol. 2014 Feb;193(2):114-21. doi: 10.1016/j.molbiopara.2014.03.005. Epub 2014 Mar 29.

Panaitescu D, Căpraru T, Bugarin V. Study of the incidence of intestinal and systemic parasitoses in a group of children with handicaps. Roum Arch Microbiol Immunol. 1995 Jan-Jun;54(1-2):65-74.

Pangasa A, Jex AR, Campbell BE, Bott NJ, Whipp M, Hogg G, Stevens MA, Gasser RB. High resolution melting-curve (HRM) analysis for the diagnosis of cryptosporidiosis in humans. Mol Cell Probes. 2009 Feb;23(1):10-5. doi: 10.1016/j.mcp.2008.10.003. Epub 2008 Oct 29.

Pangasa A, Jex AR, Nolan MJ, Campbell BE, Haydon SR, Stevens MA, Gasser RB. Highly sensitive non-isotopic restriction endonuclease fingerprinting of nucleotide variability in the gp60 gene within *Cryptosporidium* species, genotypes and subgenotypes infective to humans, and its implications. Electrophoresis. 2010 May;31(10):1637-47. doi: 10.1002/elps.200900706.

Paperna I, Finkelman S. Ultrastructural study of Sarcocystis muriviperae development in the intestine of its snake hosts. Folia Parasitol (Praha). 1996;43(1):13-9.

Parghi E, Dash L, Shastri J. Evaluation of different modifications of acid-fast staining techniques and stool enzyme-linked immunosorbent assay in detecting fecal *Cryptosporidium* in diarrheic HIV seropositive and seronegative patients. Trop Parasitol. 2014 Jul;4(2):99-104. doi: 10.4103/2229-5070.138537.

Park JH, Guk SM, Han ET, Shin EH, Kim JL, Chai JY. Genotype analysis of *Cryptosporidium spp*. prevalent in a rural village in Hwasun-gun, Republic of Korea. Korean J Parasitol. 2006 Mar;44(1):27-33.

Park JH, Kim HJ, Guk SM, Shin EH, Kim JL, Rim HJ, Lee SH, Chai JY. A survey of cryptosporidiosis among ,541 residents of 25 coastal islands in Jeollanam-Do (Province), Republic of Korea. Korean J Parasitol. 2006 Dec;44(4):367-72.

Parsons MB, Travis D, Lonsdorf EV, Lipende I, Roellig DM, Collins A, Kamenya S, Zhang H, Xiao L, Gillespie TR. Epidemiology and molecular characterization of *Cryptosporidium spp*. in humans, wild primates, and domesticated animals in the Greater Gombe Ecosystem, Tanzania. PLoS Negl Trop Dis. 2015 Feb 20;9(2):e0003529. doi: 10.1371/journal.pntd.0003529. eCollection 2015.

Patel SD, Kinariwala DM, Javadekar TB. Clinico-microbiological study of opportunistic infection in HIV seropositive patients. Indian J Sex Transm Dis. 2011 Jul;32(2):90-3. doi: 10.4103/0253-7184.85411.

Paulick MG, Bertozzi CR. The glycosylphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins. Biochemistry. 2008 Jul 8;47(27):6991-7000. doi: 10.1021/bi8006324. Epub 2008 Jun 17.

Pavie J, Menotti J, Porcher R, Donay JL, Gallien S, Sarfati C, Derouin F, Molina JM. Prevalence of opportunistic intestinal parasitic infections among HIV-infected patients with low CD4 cells counts in France in the combination antiretroviral therapy era. Int J Infect Dis. 2012 Sep;16(9):e677-9. doi: 10.1016/j.ijid.2012.05.1022. Epub 2012 Jul 6.

Pavlinac PB, John-Stewart GC, Naulikha JM, Onchiri FM, Denno DM, Odundo EA, Singa BO, Richardson BA, Walson JL. High-risk enteric pathogens associated with HIV infection and HIV exposure in Kenyan children with acute diarrhoea. AIDS. 2014 Sep 24;28(15):2287-96. doi: 10.1097/QAD.0000000000000396.

Pays E, Guyaux M, Aerts D, Van Meirvenne N, Steinert M. Telomeric reciprocal recombination as a possible mechanism for antigenic variation in trypanosomes. Nature. 1985 Aug 8-14;316(6028):562-4.

Pedersen SH, Wilkinson AL, Andreasen A, Warhurst DC, Kinung'hi SM, Urassa M, Mkwashapi DM, Todd J, Changalucha J, McDermid JM. *Cryptosporidium* prevalence and risk factors among mothers and infants 0 to 6 months in rural and semi-rural Northwest Tanzania: a prospective cohort study. PLoS Negl Trop Dis. 2014 Oct 2;8(10):e3072. doi: 10.1371/journal.pntd.0003072. eCollection 2014.

Pedraza-Díaz S, Amar C, Nichols GL, McLauchlin J. Nested polymerase chain reaction for amplification of the *Cryptosporidium* oocyst wall protein gene. Emerg Infect Dis. 2001 Jan-Feb;7(1):49-56.

Pelayo L, Nuñez FA, Rojas L, Wilke H, Furuseth Hansen E, Mulder B, Gjerde B, Robertson L. Molecular and epidemiological investigations of cryptosporidiosis in Cuban children. Ann Trop Med Parasitol. 2008 Dec;102(8):659-69. doi: 10.1179/136485908X355265.

Peng MM, Matos O, Gatei W, Das P, Stantic-Pavlinic M, Bern C, Sulaiman IM, Glaberman S, Lal AA, Xiao L. A comparison of *Cryptosporidium* subgenotypes from several geographic regions. J Eukaryot Microbiol. 2001;Suppl:28S-31S.

Peng MM, Meshnick SR, Cunliffe NA, Thindwa BD, Hart CA, Broadhead RL, Xiao L. Molecular epidemiology of cryptosporidiosis in children in Malawi. J Eukaryot Microbiol. 2003;50 Suppl:557-9.

Peng Y, Cai J, Wang W, Su B. Multiple inter-kingdom horizontal gene transfers in the evolution of the phosphoenolpyruvate carboxylase gene family. PLoS One. 2012;7(12):e51159. doi: 10.1371/journal.pone.0051159. Epub 2012 Dec 12.

Peralta RH, Velásquez JN, Cunha Fde S, Pantano ML, Sodré FC, Silva Sd, Astudillo OG, Peralta JM, Carnevale S. Genetic diversity of *Cryptosporidium* identified in clinical samples from cities in Brazil and Argentina. Mem Inst Oswaldo Cruz. 2016 Jan;111(1):30-6. doi: 10.1590/0074-02760150303.

Perch M, Sodemann M, Jakobsen MS, Valentiner-Branth P, Steinsland H, Fischer TK, Lopes DD, Aaby P, Mølbak K. Seven years' experience with *Cryptosporidium parvum* in Guinea-Bissau, West Africa. Ann Trop Paediatr. 2001 Dec;21(4):313-8.

Pereira MD, Atwill ER, Barbosa AP, Silva SA, García-Zapata MT. Intra-familial and extra-familial risk factors associated with *Cryptosporidium parvum* infection among children hospitalized for diarrhea in Goiânia, Goiás, Brazil. Am J Trop Med Hyg. 2002 Jun;66(6):787-93.

Perera J. *Cryptosporidium* associated childhood diarrhoea in Sri Lanka--a preliminary study. Ceylon Med J. 1988 Sep;33(3):101-4.

Peréz Cordón G, Cordova Paz Soldan O, Vargas Vásquez F, Velasco Soto JR, Sempere Bordes L, Sánchez Moreno M, Rosales MJ. Prevalence of enteroparasites and genotyping of Giardia lamblia in Peruvian children. Parasitol Res. 2008 Jul;103(2):459-65. doi: 10.1007/s00436-008-1007-3. Epub 2008 May 11.

Perin NM, Pires MM, Nassar SM. [Intestinal absorption of D-xylose in children infected with the human immunodeficiency virus]. Arq Gastroenterol. 2001 Oct-Dec;38(4):261-8.

Petrincová A, Valenčáková A, Luptáková L, Ondriska F, Kalinová J, Halánová M, Danišová O, Jarčuška P. Molecular characterization and first report of *Cryptosporidium* genotypes in human population in the Slovak Republic. Electrophoresis. 2015 Dec;36(23):2925-30. doi: 10.1002/elps.201500230. Epub 2015 Sep 18.

Pham-Duc P, Nguyen-Viet H, Hattendorf J, Zinsstag J, Phung-Dac C, Zurbrügg C, Odermatt P. Ascaris lumbricoides and Trichuris trichiura infections associated with wastewater and human excreta use in agriculture in Vietnam. Parasitol Int. 2013 Apr;62(2):172-80. doi: 10.1016/j.parint.2012.12.007. Epub 2013 Jan 8.

Pijnacker R, Mughini-Gras L, Vennema H, Enserink R, VAN DEN Wijngaard CC, Kortbeek T, VAN Pelt W. Characteristics of child daycare centres associated with clustering of major enteropathogens. Epidemiol Infect. 2016 Sep;144(12):2527-39. doi: 10.1017/S0950268816001011.

Pinlaor S, Mootsikapun P, Pinlaor P, Pipitgool V, Tuangnadee R. Detection of opportunistic and non-opportunistic intestinal parasites and liver flukes in HIV-positive and HIV-negative subjects. Southeast Asian J Trop Med Public Health. 2005 Jul;36(4):841-5.

Plotnikova AV, Andreev AI, Lysenko AIa. [The first cases of human cryptosporidiosis in Moscow]. Med Parazitol (Mosk). 1990 May-Jun;(3):58-60.

Pohjola S, Jokipii AM, Jokipii L. Sporadic cryptosporidiosis in a rural population is asymptomatic and associated with contact to cattle. Acta Vet Scand. 1986;27(1):91-102.

Pohlenz J, Bemrick WJ, Moon HW, Cheville NF. Bovine cryptosporidiosis: a transmission and scanning electron microscopic study of some stages in the life cycle and of the host-parasite relationship. Vet Pathol. 1978 May;15(3):417-27.

Pollock KG, Ternent HE, Mellor DJ, Chalmers RM, Smith HV, Ramsay CN, Innocent GT. Spatial and temporal epidemiology of sporadic human cryptosporidiosis in Scotland. Zoonoses Public Health. 2010 Dec;57(7-8):487-92. doi: 10.1111/j.1863-2378.2009.01247.x.

Prasad KN, Nag VL, Dhole TN, Ayyagari A. Identification of enteric pathogens in HIV-positive patients with diarrhoea in northern India. J Health Popul Nutr. 2000 Jun;18(1):23-6.

Predrag S, Branislava K, Miodrag S, Biljana MS, Suzana T, Natasa MT, Tatjana B. Clinical importance and representation of toxigenic and non-toxigenic Clostridium difficile cultivated from stool samples of hospitalized patients. Braz J Microbiol. 2012 Jan;43(1):215-23. doi: 10.1590/S1517-838220120001000023. Epub 2012 Jun 1.

Preiser G, Preiser L, Madeo L. An outbreak of cryptosporidiosis among veterinary science students who work with calves. J Am Coll Health. 2003 Mar;51(5):213-5.

Priest JW, Bern C, Xiao L, Roberts JM, Kwon JP, Lescano AG, Checkley W, Cabrera L, Moss DM, Arrowood MJ, Sterling CR, Gilman RH, Lammie PJ. Longitudinal analysis of *Cryptosporidium* species-specific immunoglobulin G antibody responses in Peruvian children. Clin Vaccine Immunol. 2006 Jan;13(1):123-31.

Proctor SJ, Kemp RL. *Cryptosporidium* anserinum sp. n. (Sporozoa) in a domestic goose Anser anser L., from Iowa. J Protozool. 1974 Nov;21(5):664-6.

Puiu D, Enomoto S, Buck GA, Abrahamsen MS, Kissinger JC. CryptoDB: the *Cryptosporidium* genome resource. Nucleic Acids Res. 2004 Jan 1;32(Database issue):D329-31.

Putignani L, Menichella D. Global distribution, public health and clinical impact of the protozoan pathogen *Cryptosporidium*. Interdiscip Perspect Infect Dis. 2010;2010. pii: 753512. doi: 10.1155/2010/753512. Epub 2010 Jul 14.

Qi M, Cai J, Wang R, Li J, Jian F, Huang J, Zhou H, Zhang L. Molecular characterization of *Cryptosporidium spp*. and Giardia duodenalis from yaks in the central western region of China. BMC Microbiol. 2015 May 21;15:108. doi: 10.1186/s12866-015-0446-0.

Quihui-Cota L, Lugo-Flores CM, Ponce-Martínez JA, Morales-Figueroa GG. Cryptosporidiosis: a neglected infection and its association with nutritional status in schoolchildren in northwestern Mexico. J Infect Dev Ctries. 2015 Aug 29;9(8):878-83. doi: 10.3855/jidc.6751.

Raccurt CP, Brasseur P, Verdier RI, Li X, Eyma E, Stockman CP, Agnamey P, Guyot K, Totet A, Liautaud B, Nevez G, Dei-Cas E, Pape JW. [Human cryptosporidiosis and *Cryptosporidium spp*. in Haiti]. Trop Med Int Health. 2006 Jun;11(6):929-34.

Raccurt CP, Fouché B, Agnamey P, Menotti J, Chouaki T, Totet A, Pape JW. Presence of Enterocytozoon bieneusi associated with intestinal coccidia in patients with chronic diarrhea visiting an HIV center in Haiti. Am J Trop Med Hyg. 2008 Oct;79(4):579-80.

Rafiei A, Rashno Z, Samarbafzadeh A, Khademvatan S. Molecular Characterization of *Cryptosporidium spp*. Isolated From Immunocompromised Patients and Children. Jundishapur J Microbiol. 2014 Apr;7(4):e9183. doi: 10.5812/jjm.9183. Epub 2014 Apr 1.

Rahman MM, Hossain MA, Paul SK, Ahmed S, Islam A, Ehsan MA, Alam MM, Kabir MR, Sarkar SR. Detection of Giardia lamblia and *Cryptosporidium parvum* by direct immunofluorescence assay in stool specimen. Mymensingh Med J. 2014 Jul;23(3):426-9.

Rahmouni I, Essid R, Aoun K, Bouratbine A. Glycoprotein 60 diversity in *Cryptosporidium parvum* causing human and cattle cryptosporidiosis in the rural region of Northern Tunisia. Am J Trop Med Hyg. 2014 Feb;90(2):346-50. doi: 10.4269/ajtmh.13-0522. Epub 2013 Dec 16.

Rahouma A, Klena JD, Krema Z, Abobker AA, Treesh K, Franka E, Abusnena O, Shaheen HI, El Mohammady H, Abudher A, Ghenghesh KS. Enteric pathogens associated with childhood diarrhea in Tripoli-Libya. Am J Trop Med Hyg. 2011 Jun;84(6):886-91. doi: 10.4269/ajtmh.2011.11-0116.

Raja K, Abbas Z, Hassan SM, Luck NH, Aziz T, Mubarak M. Prevalence of cryptosporidiosis in renal transplant recipients presenting with acute diarrhea at a single center in Pakistan. J Nephropathol. 2014 Oct;3(4):127-31. doi: 10.12860/jnp.2014.25. Epub 2014 Oct 1.

Ramakrishnan K, Shenbagarathai R, Uma A, Kavitha K, Rajendran R, Thirumalaikolundusubramanian P. Prevalence of intestinal parasitic infestation in HIV/AIDS patients with diarrhea in Madurai City, South India. Jpn J Infect Dis. 2007 Jul;60(4):209-10.

Ramly NZ, Rouzheinikov SN, Sedelnikova SE, Baker PJ, Chow YP, Wan KL, Nathan S, Rice DW. Crystallization and preliminary crystallographic analysis of a surface antigen glycoprotein, SAG19, from Eimeriatenella. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2013 Dec;69(Pt 12):1380-3. doi: 10.1107/S1744309113029734. Epub 2013 Nov 29.

Ramo A, Quílez J, Vergara-Castiblanco C, Monteagudo L, Del Cacho E, Clavel A. Multilocus typing and population structure of *Cryptosporidium* from children in Zaragoza, Spain. Infect Genet Evol. 2015 Apr;31:190-7. doi: 10.1016/j.meegid.2015.01.023. Epub 2015 Feb 7.

Ranjbar R, Baghaei K, Nazemalhosseini Mojarad E. Genetic characterization of *Cryptosporidium spp*. among patients with gastrointestinal complaints. Gastroenterol Hepatol Bed Bench. 2016 Fall;9(4):301-307.

Ranjbar-Bahadori Sh, Sangsefidi H, Shemshadi B, Kashefinejad M. Cryptosporidiosis and its potential risk factors in children and calves in Babol, north of Iran. Trop Biomed. 2011 Apr;28(1):125-31.

Rao Ajjampur SS, Asirvatham JR, Muthusamy D, Gladstone BP, Abraham OC, Mathai D, Ward H, Wanke C, Kang G. Clinical features & risk factors associated with cryptosporidiosis in HIV infected adults in India. Indian J Med Res. 2007 Dec;126(6):553-7.

Rashmi KS, Ravi Kumar KL. Intestinal Cryptosporidiosis and the Profile of the CD4 Counts in a Cohort of HIV Infected Patients. J Clin Diagn Res. 2013 Jun;7(6):1016-20. doi: 10.7860/JCDR/2013/5339.3062. Epub 2013 Apr 15.

Rathore D, Hrstka SC, Sacci JB Jr, De la Vega P, Linhardt RJ, Kumar S, McCutchan TF. Molecular mechanism of host specificity in Plasmodium falciparum infection: role of circumsporozoite protein. J Biol Chem. 2003 Oct 17;278(42):40905-10. Epub 2003 Aug 6.

Ratnam S, Paddock J, McDonald E, Whitty D, Jong M, Cooper R. Occurrence of *Cryptosporidium* oocysts in fecal samples submitted for routine microbiological examination. J Clin Microbiol. 1985 Sep;22(3):402-4.

Raub CB, Lee CC, Shibata D, Taylor C, Kartalov E. HistoMosaic Detecting KRAS G12V Mutation Across Colorectal Cancer Tissue Slices through in Situ PCR. Anal Chem. 2016 Mar 1;88(5):2792-8. doi: 10.1021/acs.analchem.5b04460. Epub 2016 Feb 12.

Realacci M, Perrone GA, Sale P, Tafani M, Di Vito M, Perrone-Donnorso R, Zotti G, Russo MA. Detection of oncogenic HPV and identification of 72Arg polymorphic p53 by in situ PCR for clinical routine purposes. Anticancer Res. 2006 Jul-Aug;26(4B):3095-103.

Regnath T, Klemm T, Ignatius R. Rapid and accurate detection of Giardia lamblia and *Cryptosporidium spp*. antigens in human fecal specimens by new commercially available qualitative immunochromatographic assays. Eur J Clin Microbiol Infect Dis. 2006 Dec;25(12):807-9.

Rehg JE, Lawton GW, Pakes SP. *Cryptosporidium cuniculus* in the rabbit (Oryctolagus cuniculus). Laboratory Animals Science. 1979; 29: 656-660.

Reif JS, Wimmer L, Smith JA, Dargatz DA, Cheney JM. Human cryptosporidiosis associated with an epizootic in calves. Am J Public Health. 1989 Nov;79(11):1528-30.

Reinholt SJ, Behrent A, Greene C, Kalfe A, Baeumner AJ. Isolation and amplification of mRNA within a simple microfluidic lab on a chip. Anal Chem. 2014 Jan 7;86(1):849-56. doi: 10.1021/ac403417z. Epub 2013 Dec 13.

Reinthaler FF, Linck G, Klem G, Mascher F, Sixl W. Intestinal parasites in children with diarrhea in El Salvador. Geogr Med. 1988;18:175-80.

Rey, F, Acosta, S, Martínez, W, Lena, A, Tort, C, Acuña, A, Míguez, D, Ureta, A. Implementación de métodos moleculares y microscópicos para estudios clínicos y ambientales relativos a *Cryptosporidium* sp. en Uruguay. INNOTEC. 2016; No. 12 (34 - 47).

Ribeiro PC, Pile E, Queiroz MM, Norberg AN, Tenório JR. Cryptosporidiosis occurrence in HIV+ patients attended in a hospital, Brazil. Rev Saude Publica. 2004 Jun;38(3):469-70. Epub 2004 Jul 8.

Ribes JA, Seabolt JP, Overman SB. Point prevalence of *Cryptosporidium*, Cyclospora, and Isospora infections in patients being evaluated for diarrhea. Am J Clin Pathol. 2004 Jul;122(1):28-32.

Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 2000 Jun;16(6):276-7.

Ridley RK, Olsen RM. Rapid diagnosis of bovine cryptosporidiosis with a modified commercial acid-fast staining procedure. J Vet Diagn Invest. 1991 Apr;3(2):182-3.

Rimhanen-Finne R, Sakari Jokiranta T, Virtanen MJ, Kuusi M. Giardia and *Cryptosporidium* infection in Finland: a registry-based study of their demographic determinants. APMIS. 2011 Nov;119(11):735-40. doi: 10.1111/j.1600-0463.2011.02759.x. Epub 2011 Jul 22.

Rimšelienė G, Vold L, Robertson L, Nelke C, Søli K, Johansen ØH, Thrana FS, Nygård K. An outbreak of gastroenteritis among schoolchildren staying in a wildlife reserve: thorough investigation reveals Norway's largest cryptosporidiosis outbreak. Scand J Public Health. 2011 May;39(3):287-95. doi: 10.1177/1403494810396557. Epub 2011 Feb 14.

Rindfleisch, EG. The bacillus of tubercle. British Medical Journal. 1882; ii, 142.

Ritchie LS. An ether sedimentation technique for routine stool examinations. Bull U S Army Med Dep. 1948 Apr;8(4):326.

Rivera WL, Yason JA, Rivera PT. Serological detection of cryptosporidiosis among Filipino cancer patients. Parasitol Res. 2005 Dec;98(1):75-6. Epub 2005 Oct 20.

Rizk H, Soliman M. Coccidiosis among malnourished children in Mansoura, Dakahlia Governorate, Egypt. J Egypt Soc Parasitol. 2001 Dec;31(3):877-86.

Robertson L, Gjerde B, Forberg T, Haugejorden G, Kielland C. A small outbreak of human cryptosporidiosis associated with calves at a dairy farm in Norway. Scand J Infect Dis. 2006;38(9):810-3.

Robinson, G, Chalmers, RM. Assessment of polymorphic genetic markers for multi-locus typing of *Cryptosporidium parvum* and *Cryptosporidium hominis*. Exp Parasitol. 2012 Oct;132(2):200-15. doi: 10.1016/j.exppara.2012.06.016. Epub 2012 Jul 7.

Roch, S, Steel, M. Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. Theoretical Population Biology. 2015; 100: 56-62.

Rocha EP, Smith JM, Hurst LD, Holden MT, Cooper JE, Smith NH, Feil EJ. Comparisons of dN/dS are time dependent for closely related bacterial genomes. J Theor Biol. 2006 Mar 21;239(2):226-35. Epub 2005 Oct 18.

Rochelle PA, R De Leon, DM Ferguson, MH Stewart and RL Wolfe. Optimization of an infectivity assay, combining cell culture and PCR, for waterborne *Cryptosporidium parvum*. In: 1997 International Symposium on Waterborne *Cryptosporidium* Proceedings (Fricker CR, JL Clancy and PA Rochelle, eds), pp 31–40, AWWA, Denver, Colorado, USA.

Rodrigues F, Davies EG, Harrison P, McLauchlin J, Karani J, Portmann B, Jones A, Veys P, Mieli-Vergani G, Hadzić N. Liver disease in children with primary immunodeficiencies. J Pediatr. 2004 Sep;145(3):333-9.

Roelfsema JH, Sprong H, Cacciò SM, Takumi K, Kroes M, van Pelt W, Kortbeek LM, van der Giessen JW. Molecular characterization of human *Cryptosporidium spp*. isolates after an unusual increase in late summer 2012. Parasit Vectors. 2016 Mar 10;9:138. doi: 10.1186/s13071-016-1397-5.

Roka M, Goñi P, Rubio E, Clavel A. Prevalence of intestinal parasites in HIV-positive patients on the island of Bioko, Equatorial Guinea: its relation to sanitary conditions and socioeconomic factors. Sci Total Environ. 2012 Aug 15;432:404-11. doi: 10.1016/j.scitotenv.2012.06.023. Epub 2012 Jul 4.

Rolando RF, Silva Sd, Peralta RH, Silva AJ, Cunha Fde S, Bello AR, Peralta JM. Detection and differentiation of *Cryptosporidium* by real-time polymerase chain reaction in stool samples from patients in Rio de Janeiro, Brazil. Mem Inst Oswaldo Cruz. 2012 Jun;107(4):476-9.

Roldán WH, Espinoza YA, Huapaya PE, Huiza AF, Sevilla CR, Jiménez S. Frequency of human toxocariasis in a rural population from Cajamarca, Peru determined by DOT-ELISA test. Rev Inst Med Trop Sao Paulo. 2009 Mar-Apr;51(2):67-71.

Rosinski-Chupin I, Sauvage E, Mairey B, Mangenot S, Ma L, Da Cunha V, Rusniok C, Bouchier C, Barbe V, Glaser P. Reductive evolution in Streptococcus agalactiae and the emergence of a host adapted lineage. BMC Genomics. 2013 Apr 15;14:252. doi: 10.1186/1471-2164-14-252.

Rossignol JF, Ayoub A, Ayers MS. Treatment of diarrhea caused by *Cryptosporidium parvum*: a prospective randomized, double-blind, placebo-controlled study of Nitazoxanide. J Infect Dis. 2001 Jul 1;184(1):103-6. Epub 2001 May 29.

Rossit AR, de Almeida MT, Nogueira CA, da Costa Oliveira JG, Barbosa DM, Moscardini AC, Mascarenhas JD, Gabbay YB, Marques FR, Cardoso LV, Cavasini CE, Machado RL. Bacterial, yeast, parasitic, and viral enteropathogens in HIV-infected children from São Paulo State, Southeastern Brazil. Diagn Microbiol Infect Dis. 2007 Jan;57(1):59-66.

Rowe JS, Shah SS, Motlhagodi S, Bafana M, Tawanana E, Truong HT, Wood SM, Zetola NM, Steenhoff AP. An epidemiologic review of enteropathogens in Gaborone, Botswana: shifting patterns of resistance in an HIV endemic region. PLoS One. 2010 Jun 2;5(6):e10924. doi: 10.1371/journal.pone.0010924.

Roy N, Bhattacharyya S, Chakrabarty S, Laskar S, Babu SM, Bhattacharyya MK. Dominant negative mutant of Plasmodium Rad51 causes reduced parasite burden in host by abrogating DNA double-strand break repair. Mol Microbiol. 2014 Oct;94(2):353-66. doi: 10.1111/mmi.12762. Epub 2014 Sep 5.

Rozej W, Gołab E, Waloch M, Wasik M, Sadkowska-Todys M, Czerwiński M, Dzbeński TH. [The occurrence of *Cryptosporidium* in a group of children and adults with diarrhoea of undetermined earlier aetiology]. Przegl Epidemiol. 2010;64(1):35-9.

Ryan UM, Bath C, Robertson I, Read C, Elliot A, McInnes L, Traub R, Besier B. Sheep may not be an important zoonotic reservoir for *Cryptosporidium* and Giardia parasites. Appl Environ Microbiol. 2005 Sep;71(9):4992-7.

Ryuda N, Hashimoto T, Ueno D, Inoue K, Someya T. Visualization and direct counting of individual denitrifying bacterial cells in soil by nirK-targeted direct in situ PCR. Microbes Environ. 2011;26(1):74-80.

Rzymski P, Słodkowicz-Kowalska A, Klimaszyk P, Solarczyk P, Poniedziałek B. Screening of protozoan and microsporidian parasites in feces of great cormorant (Phalacrocorax carbo). Environ Sci Pollut Res Int. 2017 Mar 2. doi: 10.1007/s11356-017-8652-y. [Epub ahead of print]

Sadaka HA, Gaafar MR, Mady RF, Hezema NN. Evaluation of ImmunoCard STAT test and ELISA versus light microscopy in diagnosis of giardiasis and cryptosporidiosis. Parasitol Res. 2015 Aug;114(8):2853-63. doi: 10.1007/s00436-015-4486-z. Epub 2015 May 1.

Sadraei J, Rizvi MA, Baveja UK. Diarrhea, CD4+ cell counts and opportunistic protozoa in Indian HIV-infected patients. Parasitol Res. 2005 Oct;97(4):270-3. Epub 2005 Jul 7.

Sahimin N, Lim YA, Ariffin F, Behnke JM, Lewis JW, Mohd Zain SN. Migrant Workers in Malaysia: Current Implications of Sociodemographic and Environmental Characteristics in the Transmission of Intestinal Parasitic Infections. PLoS Negl Trop Dis. 2016 Nov 2;10(11):e0005110. doi: 10.1371/journal.pntd.0005110. eCollection 2016.

Saigal K, Sharma A, Sehgal R, Sharma P, Malla N, Khurana S. Intestinal microsporidiosis in India: a two year study. Parasitol Int. 2013 Feb;62(1):53-6. doi: 10.1016/j.parint.2012.09.005. Epub 2012 Sep 25.

Saksirisampant W, Eampokalap B, Rattanasrithong M, Likanonsakul S, Wiwanitkit V, Nasingkarn A, Denmasae N. A prevalence of *Cryptosporidium* infections among Thai HIV-infected patients. J Med Assoc Thai. 2002 Jun;85 Suppl 1:S424-8.

Saksirisampant W, Prownebon J, Saksirisampant P, Mungthin M, Siripatanapipong S, Leelayoova S. Intestinal parasitic infections: prevalences in HIV/AIDS patients in a Thai AIDS-care centre. Ann Trop Med Parasitol. 2009 Oct;103(7):573-81. doi: 10.1179/000349809X12502035776072.

Salehi Sangani G, Mirjalali H, Farnia S, Rezaeian M. Prevalence of Intestinal Coccidial Infections among Different Groups of Immunocompromised Patients. Iran J Parasitol. 2016 Jul-Sep;11(3):332-338.

Sallon S, Deckelbaum RJ, Schmid II, Harlap S, Baras M, Spira DT. *Cryptosporidium*, malnutrition, and chronic diarrhea in children. Am J Dis Child. 1988 Mar;142(3):312-5.

Sallon S, el-Shawwa R, Khalil M, Ginsburg G, el Tayib J, el-Eila J, Green V, Hart CA. Diarrhoeal disease in children in Gaza. Ann Trop Med Parasitol. 1994 Apr;88(2):175-82.

Salser W, Fry K, Brunk C, Poon R. Nucleotide sequencing of DNA: preliminary characterization of the products of specific cleavages at guanine, cytosine, or adenine residues (bacteriophage M13-ribosubstitution-DNA polymerase I-electrophoresis-two-dimensional fingerprinting). Proc Natl Acad Sci U S A. 1972 Jan;69(1):238-42.

Salyer SJ, Gillespie TR, Rwego IB, Chapman CA, Goldberg TL. Epidemiology and molecular relationships of *Cryptosporidium spp*. in people, primates, and livestock from Western Uganda. PLoS Negl Trop Dis. 2012;6(4):e1597. doi: 10.1371/journal.pntd.0001597. Epub 2012 Apr 10.

Samie A, Bessong PO, Obi CL, Sevilleja JE, Stroup S, Houpt E, Guerrant RL. *Cryptosporidium* species: preliminary descriptions of the prevalence and genotype distribution among school children and hospital patients in the Venda region, Limpopo Province, South Africa. Exp Parasitol. 2006 Dec;114(4):314-22. Epub 2006 Jun 27.

Samie A, Guerrant RL, Barrett L, Bessong PO, Igumbor EO, Obi CL. Prevalence of intestinal parasitic and bacterial pathogens in diarrhoeal and non-diarroeal human stools from Vhembe district, South Africa. J Health Popul Nutr. 2009 Dec;27(6):739-45.

Samie A, Makuwa S, Mtshali S, Potgieter N, Thekisoe O, Mbati P, Bessong PO. Parasitic infection among HIV/AIDS patients at Bela-Bela clinic, Limpopo province, South Africa with special reference to *Cryptosporidium*. Southeast Asian J Trop Med Public Health. 2014 Jul;45(4):783-95.

Sam-Yellowe, TY. A Plasmodium gene family encoding Maurer's cleft membrane proteins: structural properties and expression profiling. Genome Res. 2004 Jun;14(6):1052-9. Epub 2004 May 12.

Sánchez A, Munoz M, Gómez N, Tabares J, Segura L, Salazar Á, Restrepo C, Ruíz M, Reyes P, Qian Y, Xiao L, López MC, Ramírez JD. Molecular Epidemiology of Giardia, Blastocystis and *Cryptosporidium* among Indigenous Children from the Colombian Amazon Basin. Front Microbiol. 2017 Feb 21;8:248. doi: 10.3389/fmicb.2017.00248. eCollection 2017.

Sánchez-Vega JT, Tay-Zavala J, Aguilar-Chiu A, Ruiz-Sánchez D, Malagón F, Rodríguez-Covarrubias JA, Ordóñez-Martínez J, Calderón-Romero L. Cryptosporidiosis and other intestinal protozoan infections in children less than one year of age in Mexico City. Am J Trop Med Hyg. 2006 Dec;75(6):1095-8.

Sander AF, Lavstsen T, Rask TS, Lisby M, Salanti A, Fordyce SL, Jespersen JS, Carter R, Deitsch KW, Theander TG, Pedersen AG, Arnot DE. DNA secondary structures are associated with recombination in major Plasmodium falciparum variable surface antigen gene families. Nucleic Acids Res. 2014 Feb;42(4):2270-81. doi: 10.1093/nar/gkt1174. Epub 2013 Nov 18.

Sanders JW, Putnam SD, Gould P, Kolisnyk J, Merced N, Barthel V, Rozmajzl PJ, Shaheen H, Fouad S, Frenck RW. Diarrheal illness among deployed U.S. military personnel during Operation Bright Star 2001--Egypt. Diagn Microbiol Infect Dis. 2005 Jun;52(2):85-90.

Saneian H, Yaghini O, Yaghini A, Modarresi MR, Soroshnia M. Infection Rate of *Cryptosporidium parvum* among Diarrheic Children in Isfahan. Iran J Pediatr. 2010 Sep;20(3):343-7.

Sanford SE, Josephson GK. Bovine Cryptosporidiosis: Clinical and Pathological Findings in Forty-two Scouring Neonatal Calves. Can Vet J. 1982 Dec;23(12):343-7.

Sangaré I, Bamba S, Cissé M, Zida A, Bamogo R, Sirima C, Yaméogo BK, Sanou R, Drabo F, Dabiré RK, Guiguemdé RT. Prevalence of intestinal opportunistic parasites infections in the University hospital of Bobo-Dioulasso, Burkina Faso.Infect Dis Poverty. 2015 Jul 27;4:32. doi: 10.1186/s40249-015-0065-x. eCollection 2015.

Saredi N, Bava J. Cryptosporidiosis in pediatric patients. Rev Inst Med Trop Sao Paulo. 1998 May-Jun;40(3):197-200.

Sarfati C, Bourgeois A, Menotti J, Liegeois F, Moyou-Somo R, Delaporte E, Derouin F, Ngole EM, Molina JM. Prevalence of intestinal parasites including microsporidia in human immunodeficiency virus-infected adults in Cameroon: a cross-sectional study. Am J Trop Med Hyg. 2006 Jan;74(1):162-4.

Sarkar R, Ajjampur SS, Prabakaran AD, Geetha JC, Sowmyanarayanan TV, Kane A, Duara J, Muliyil J, Balraj V, Naumova EN, Ward H, Kang G. Cryptosporidiosis among children in an endemic semiurban community in southern India: does a protected drinking water source decrease infection? Clin Infect Dis. 2013 Aug;57(3):398-406. doi: 10.1093/cid/cit288. Epub 2013 May 24.

Sayers GM, Dillon MC, Connolly E, Thornton L, Hyland E, Loughman E, O'Mahony MA, Butler KM. Cryptosporidiosis in children who visited an open farm. Commun Dis Rep CDR Rev. 1996 Sep 13;6(10):R140-4.

Sazmand A, Rasooli A, Nouri M, Hamidinejat H, Hekmatimoghaddam S. Prevalence of *Cryptosporidium spp*. in Camels and Involved People in Yazd Province, Iran. Iran J Parasitol. 2012;7(1):80-4.

Schiller SE, Webster KN, Power M. Detection of *Cryptosporidium hominis* and novel *Cryptosporidium* bat genotypes in wild and captive Pteropus hosts in Australia. Infect Genet Evol. 2016 Oct;44:254-60. doi: 10.1016/j.meegid.2016.07.002. Epub 2016 Jul 4.

Schindler AR, Abs El-Osta YG, Stevens M, Sinclair MI, Gasser RB. Capillary electrophoretic analysis of fragment length polymorphism in ribosomal markers of *Cryptosporidium* from humans. Mol Cell Probes. 2005 Dec;19(6):394-9. Epub 2005 Sep 19.

Schmitz JA, Smith DH. *Cryptosporidium* infection in a calf. Journal of the American Veterinary Medical Association. 1975; 167(8), 731-732.

Schuh AJ, Ward MJ, Leigh Brown AJ, Barrett AD. Dynamics of the emergence and establishment of a newly dominant genotype of Japanese encephalitis virus throughout Asia. J Virol. 2014 Apr;88(8):4522-32. doi: 10.1128/JVI.02686-13. Epub 2014 Feb 5.

Segura R, Prim N, Montemayor M, Valls ME, Muñoz C. Predominant virulent IbA10G2 subtype of *Cryptosporidium hominis* in human isolates in Barcelona: a five-year study. PLoS One. 2015 Mar 27;10(3):e0121753. doi: 10.1371/journal.pone.0121753. eCollection 2015.

Sejdini A, Mahmud R, Lim YA, Mahdy M, Sejdini F, Gjoni V, Xhaferraj K, Kasmi G. Intestinal parasitic infections among children in central Albania. Ann Trop Med Parasitol. 2011 Apr;105(3):241-50. doi: 10.1179/136485911X12987676649584.

Seo M, Huh S, Chai JY, Yu JR. An epidemiological survey on *Cryptosporidium parvum* infection of inhabitants in Chorwon-gun, Kangwon-do. Korean J Parasitol. 2001 Jun;39(2):201-3.

Seviour R, Nielsen PH. Microbial ecology of activated sludge. London: IWA Publishing. 2010. Print.

Seyrafian S, Pestehchian N, Kerdegari M, Yousefi HA, Bastani B. Prevalence rate of *Cryptosporidium* infection in hemodialysis patients in Iran. Hemodial Int. 2006 Oct;10(4):375-9.

Shah S, Kongre V, Kumar V, Bharadwaj R. A Study of Parasitic and Bacterial Pathogens Associated with Diarrhea in HIV-Positive Patients. Cureus. 2016 Sep 27;8(9):e807.

Shahdoust S, Niyyati M, Haghighi A, Azargashb E, Khataminejad MR. Prevalence of intestinal parasites in referred individuals to the medical centers of Tonekabon city, Mazandaran province. Gastroenterol Hepatol Bed Bench. 2016 Dec;9(Suppl1):S75-S79.

Shalaby I, Gherbawy Y, Jamjoom M, Banaja A. Prevalence and genotyping of *Cryptosporidium* in stool samples collected from children in Taif City (Saudi Arabia). Trop Biomed. 2014 Jun;31(2):215-24.

Shalaby NM, Shalaby NM. *Cryptosporidium parvum* infection among Egyptian school children. J Egypt Soc Parasitol. 2015 Apr;45(1):125-31.

Sharbatkhori M, Nazemalhosseini Mojarad E, Taghipour N, Pagheh AS, Mesgarian F. Prevalence and Genetic Characterization of *Cryptosporidium spp*. In Diarrheic Children from Gonbad Kavoos City, Iran. Iran J Parasitol. 2015 Jul-Sep;10(3):441-7.

Sharif M, Daryani A, Kia E, Rezaei F, Nasiri M, Nasrolahei M. Prevalence of intestinal parasites among food handlers of Sari, Northern Iran. Rev Inst Med Trop Sao Paulo. 2015 Mar-Apr;57(2):139-44. doi: 10.1590/S0036-46652015000200007.

Shariff M, Deb M, Singh R, Singh K. *Cryptosporidium* infection in children with diarrhoea of acute onset. J Trop Pediatr. 2002 Jun;48(3):187-8.

Sharma P, Sharma A, Sehgal R, Malla N, Khurana S. Genetic diversity of *Cryptosporidium* isolates from patients in North India. Int J Infect Dis. 2013 Aug;17(8):e601-5. doi: 10.1016/j.ijid.2012.12.003. Epub 2013 Jan 16.

Sharma P, Khurana S, Sharma A, Sehgal R, Malla N. Presence of intracellular viruses in human *Cryptosporidium* isolates. Ann Parasitol. 2016;62(2):139-47. doi: 10.17420/ap6202.46.

Shehata AI, Hassanein F. Intestinal parasitic infections among mentally handicapped individuals in Alexandria, Egypt. Ann Parasitol. 2015;61(4):275-81. doi: 10.17420/ap6104.19.

Shenoy S, Baliga S, Kurnvilla T, Prashanth HV, Dominic RM. Opportunistic intestinal parasitic infections in human immunodeficiency virus infected patients in Mangalore, South India. Trop Doct. 2003 Oct;33(4):250.

Sheoran A, Wiffin A, Widmer G, Singh P, Tzipori S. Infection With *Cryptosporidium hominis* Provides Incomplete Protection of the Host Against *Cryptosporidium parvum*. J Infect Dis. 2012 Mar 15;205(6):1019-23. doi: 10.1093/infdis/jir874. Epub 2012 Jan 25.

Shimelis T, Tassachew Y, Lambiyo T. *Cryptosporidium* and other intestinal parasitic infections among HIV patients in southern Ethiopia: significance of improved HIV-related care. Parasit Vectors. 2016 May 10;9(1):270. doi: 10.1186/s13071-016-1554-x.

Shio, MT. PKC/ROS-Mediated NLRP3 Inflammasome Activation Is Attenuated by Leishmania Zinc-Metalloprotease during Infection. PLoS Negl Trop Dis. 2015 Jun 26;9(6):e0003868. doi: 10.1371/journal.pntd.0003868. eCollection 2015.

Shoaib S, Tauheed S, Hafiz A. Frequency of *Cryptosporidium* in childhood diarrhoea--importance of modified acid fast technique. J Ayub Med Coll Abbottabad. 2003 Jul-Sep;15(3):3-5.

Shoukry NM, Dawoud HA, Haridy FM. Studies on zoonotic cryptosporidiosis parvum in Ismailia Governorate, Egypt.J Egypt Soc Parasitol. 2009 Aug;39(2):479-88.

Sia SK, Linder V, Parviz BA, Siegel A, Whitesides GM. An integrated approach to a portable and low-cost immunoassay for resource-poor settings. Angew Chem Int Ed Engl. 2004 Jan 16;43(4):498-502.

Sibley LD, Qiu W, Fentress S, Taylor SJ, Khan A, Hui R. Forward genetics in Toxoplasma gondii reveals a family of rhoptry kinases that mediates pathogenesis. Eukaryot Cell. 2009 Aug;8(8):1085-93. doi: 10.1128/EC.00107-09. Epub 2009 May 22.

Silva CV, Ferreira MS, Borges AS, Costa-Cruz JM. Intestinal parasitic infections in HIV/AIDS patients: experience at a teaching hospital in central Brazil. Scand J Infect Dis. 2005;37(3):211-5.

Silva RK, Pacheco FT, Martins AS, Menezes JF, Costa-Ribeiro H Jr, Ribeiro TC, Mattos ÂP, Oliveira RR, Soares NM, Teixeira MC. Performance of microscopy and ELISA for diagnosing Giardia duodenalis infection in different pediatric groups. Parasitol Int. 2016 Dec;65(6 Pt A):635-640. doi: 10.1016/j.parint.2016.08.012. Epub 2016 Aug 30.

Sim S, Yu JR, Lee YH, Lee JS, Jeong HG, Mohamed AA, Hong ST. Prevalence of *Cryptosporidium* Infection among Inhabitants of 2 Rural Areas in White Nile State, Sudan. Korean J Parasitol. 2015 Dec;53(6):745-7. doi: 10.3347/kjp.2015.53.6.745. Epub 2015 Dec 31.

Simango C, Mutikani S. Cryptosporidiosis in Harare, Zimbabwe. Cent Afr J Med. 2004 May-Jun;50(5-6):52-4.

Sinclair, C. B., Dunn, D. B. Surface printing of plant leaves for phylogenetic studies. Surface printing of plant leaves for phylogenetic studies. Stain Technol. 1961 Sep;36:299-304.

Singh A, Bairy I, Shivananda PG. Spectrum of opportunistic infections in AIDS cases. Indian J Med Sci. 2003 Jan;57(1):16-21.

Sinniah B, Sabaridah I, Soe MM, Sabitha P, Awang IP, Ong GP, Hassan AK. Determining the prevalence of intestinal parasites in three Orang Asli (Aborigines) communities in Perak, Malaysia. Trop Biomed. 2012 Jun;29(2):200-6.

Sirinarumitr T, Morozov I, Nawagitgul P, Sorden SD, Harms PA, Paul PS. Utilization of a rate enhancement hybridization buffer system for rapid in situ hybridization for the detection of porcine circovirus in cell culture and in tissues of pigs with postweaning multisystemic wasting syndrome. J Vet Diagn Invest. 2000 Nov;12(6):562-5.

Sirisena UM, Iddawela WM, Noordeen F, Wickramasinghe S. Prevalence and identification of *Cryptosporidium* species in paediatric patients with diarrhoea. Ceylon Med J. 2014 Sep;59(3):75-8. doi: 10.4038/cmj.v59i3.7467.

Siwila J, Phiri IG, Vercruysse J, Goma F, Gabriel S, Claerebout E, Geurden T. Asymptomatic cryptosporidiosis in Zambian dairy farm workers and their household members. Trans R Soc Trop Med Hyg. 2007 Jul;101(7):733-4. Epub 2007 Mar 21.

Siwila J, Phiri IG, Enemark HL, Nchito M, Olsen A. Intestinal helminths and protozoa in children in pre-schools in Kafue district, Zambia. Trans R Soc Trop Med Hyg. 2010 Feb;104(2):122-8. doi: 10.1016/j.trstmh.2009.07.024. Epub 2009 Sep 22.

Siwila J, Phiri IG, Enemark HL, Nchito M, Olsen A. Seasonal prevalence and incidence of *Cryptosporidium spp*. and Giardia duodenalis and associated diarrhoea in children attending pre-school in Kafue, Zambia. Trans R Soc Trop Med Hyg. 2011 Feb;105(2):102-8. doi: 10.1016/j.trstmh.2010.10.004. Epub 2010 Nov 18.

Slavin D. *Cryptosporidium meleagridis* (sp. nov.). J. Comp. Path. 1955; 65, 262-6.

Smith JJ, Gunasekera TS, Barardi CRM, Veal D, Vesey G. Determination of *Cryptosporidium parvum* oocyst viability by fluorescence in situ hybridization using a ribosomal RNA-directed probe. Journal of Applied Microbiology. 2004; 96, 409–417. doi:10.1046/j.1365-2672.2004.02150.x

Smith KE, Stenzel SA, Bender JB, Wagstrom E, Soderlund D, Leano FT, Taylor CM, Belle-Isle PA, Danila R. Outbreaks of enteric infections caused by multiple pathogens associated with calves at a farm day camp. Pediatr Infect Dis J. 2004 Dec;23(12):1098-104.

Snyder SP, England JJ, McChesney AE. Cryptosporidiosis in immunodeficient Arabian foals. Vet Pathol. 1978 Jan;15(1):12-7.

Soba B, Petrovec M, Mioc V, Logar J. Molecular characterisation of *Cryptosporidium* isolates from humans in Slovenia. Clin Microbiol Infect. 2006 Sep;12(9):918-21.

Soba B, Logar J. Genetic classification of *Cryptosporidium* isolates from humans and calves in Slovenia. Parasitology. 2008 Sep;135(11):1263-70. doi: 10.1017/S0031182008004800. Epub 2008 Jul 29.

Solórzano-Santos F, Penagos-Paniagua M, Meneses-Esquivel R, Miranda-Novales MG, Leaños-Miranda B, Angulo-González D, Fajardo-Gutiérrez A. [*Cryptosporidium parvum* infection in malnourished and non malnourished children without diarrhea in a Mexican rural population]. Rev Invest Clin. 2000 Nov-Dec;52(6):625-31.

Sopwith W, Osborn K, Chalmers R, Regan M. The changing epidemiology of cryptosporidiosis in North West England. Epidemiol Infect. 2005 Oct;133(5):785-93.

Souza Wd. Secretory organelles of pathogenic protozoa. An Acad Bras Cienc. 2006 Jun;78(2):271-91. Epub 2006 May 11.

Sow SO, Muhsen K, Nasrin D, Blackwelder WC, Wu Y, Farag TH, Panchalingam S, Sur D, Zaidi AK, Faruque AS, Saha D, Adegbola R, Alonso PL, Breiman RF, Bassat Q, Tamboura B, Sanogo D, Onwuchekwa U, Manna B, Ramamurthy T, Kanungo S, Ahmed S, Qureshi S, Quadri F, Hossain A, Das SK, Antonio M, Hossain MJ, Mandomando I, Nhampossa T, Acácio S, Omore R, Oundo JO, Ochieng JB, Mintz ED, O'Reilly CE, Berkeley LY, Livio S, Tennant SM, Sommerfelt H, Nataro JP, Ziv-Baran T, Robins-Browne RM, Mishcherkin V, Zhang J, Liu J, Houpt ER, Kotloff KL, Levine MM. The Burden of *Cryptosporidium* Diarrheal Disease among Children < 24 Months of Age in Moderate/High Mortality Regions of Sub-Saharan Africa and South Asia, Utilizing Data from the Global Enteric Multicenter Study (GEMS). PLoS Negl Trop Dis. 2016 May 24;10(5):e0004729. doi: 10.1371/journal.pntd.0004729. eCollection 2016.

Sree, KS, Varma A. Biocontrol of lepidopteran pests: use of soil microbes and their metabolites. Cham: Springer, 2015. Print.

Srijan A, Wongstitwilairoong B, Pitarangsi C, Serichantalergs O, Fukuda CD, Bodhidatta L, Mason CJ. Re-evaluation of commercially available enzyme-linked immunosorbent assay for the detection of Giardia lamblia and *Cryptosporidium spp* from stool specimens. Southeast Asian J Trop Med Public Health. 2005;36 Suppl 4:26-9.

Srisuphanunt M, Saksirisampant W, Karanis P. Prevalence and genotyping of *Cryptosporidium* isolated from HIV/AIDS patients in urban areas of Thailand. Ann Trop Med Parasitol. 2011 Sep;105(6):463-8. doi: 10.1179/1364859411Y.0000000032.

Stark D, Fotedar R, van Hal S, Beebe N, Marriott D, Ellis JT, Harkness J. Prevalence of enteric protozoa in human immunodeficiency virus (HIV)-positive and HIV-negative men who have sex with men from Sydney, Australia. Am J Trop Med Hyg. 2007 Mar;76(3):549-52.

Stark D, Al-Qassab SE, Barratt JL, Stanley K, Roberts T, Marriott D, Harkness J, Ellis JT. Evaluation of multiplex tandem real-time PCR for detection of *Cryptosporidium spp*., Dientamoeba fragilis, Entamoeba histolytica, and Giardia intestinalis in clinical stool samples. J Clin Microbiol. 2011 Jan;49(1):257-62. doi: 10.1128/JCM.01796-10. Epub 2010 Nov 3.

Starkey SR, Johnson AL, Ziegler PE, Mohammed HO. An outbreak of cryptosporidiosis among alpaca crias and their human caregivers. J Am Vet Med Assoc. 2007 Nov 15;231(10):1562-7.

Stensvold CR, Nielsen SD, Badsberg JH, Engberg J, Friis-Møller N, Nielsen SS, Nielsen HV, Friis-Møller A. The prevalence and clinical significance of intestinal parasites in HIV-infected patients in Denmark. Scand J Infect Dis. 2011 Feb;43(2):129-35. doi: 10.3109/00365548.2010.524659. Epub 2010 Oct 11.

Stensvold CR, Beser J, Axén C, Lebbad M. High applicability of a novel method for gp60-based subtyping of *Cryptosporidium meleagridis*. J Clin Microbiol. 2014 Jul;52(7):2311-9. doi: 10.1128/JCM.00598-14. Epub 2014 Apr 16.

Stensvold CR, Ethelberg S, Hansen L, Sahar S, Voldstedlund M, Kemp M, Hartmeyer GN, Otte E, Engsbro AL, Nielsen HV, Mølbak K. *Cryptosporidium* infections in Denmark, 2010-2014. Dan Med J. 2015 May;62(5). pii: A5086.

Stephenson N, Foley J. Parallelisms and Contrasts in the Diverse Ecologies of the Anaplasma phagocytophilum and Borrelia burgdorferi Complexes of Bacteria in the Far Western United States. Vet. Sci. 2016, 3, 26; doi:10.3390/vetsci3040026

Stibbs HH, Ongerth JE. Immunofluorescence detection of *Cryptosporidium* oocysts in fecal smears. J Clin Microbiol. 1986 Oct;24(4):517-21.

Stich A, Abel PM, Krishna S. Human African trypanosomiasis. BMJ. 2002 Jul 27;325(7357):203-6.

Stoletzki N, Eyre-Walker A. The positive correlation between dN/dS and dS in mammals is due to runs of adjacent substitutions. Mol Biol Evol. 2011 Apr;28(4):1371-80. doi: 10.1093/molbev/msq320. Epub 2010 Nov 29.

Streenu, VB. Microsatellite polymorphism across the M. tuberculosis and M. bovis genomes: implications on genome evolution and plasticity. BMC Genomics. 2006 Apr 10;7:78.

Strong WB, Gut J, Nelson RG. Cloning and sequence analysis of a highly polymorphic *Cryptosporidium parvum* gene encoding a 60-kilodalton glycoprotein and characterization of its 15- and 45-kilodalton zoite surface antigen products. Infect Immun. 2000 Jul;68(7):4117-34.

Strouts FR, Power P, Croucher NJ, Corton N, van Tonder A, Quail MA, Langford PR, Hudson MJ, Parkhill J, Kroll JS, Bentley SD. Lineage-specific virulence determinants of Haemophilus influenzae biogroup aegyptius. Emerg Infect Dis. 2012 Mar;18(3):449-57. doi: 10.3201/eid1803.110728.

Sturbaum GD, Jost BH, Sterling CR. Nucleotide changes within three *Cryptosporidium parvum* surface protein encoding genes differentiate genotype I from genotype II isolates. Mol Biochem Parasitol. 2003 Apr 25;128(1):87-90.

Sturbaum GD, Schaefer DA, Jost BH, Sterling CR, Riggs MW. Antigenic differences within the *Cryptosporidium hominis* and *Cryptosporidium parvum* surface proteins P23 and GP900 defined by monoclonal antibody reactivity. Mol Biochem Parasitol. 2008 Jun;159(2):138-41. doi: 10.1016/j.molbiopara.2008.02.009. Epub 2008 Mar 4.

Su C, Evans D, Cole RH, Kissinger JC, Ajioka JW, Sibley LD. Recent expansion of Toxoplasma through enhanced oral transmission. Science. 2003 Jan 17;299(5605):414-6.

Sudhir Kumar, Glen Stecher, and Koichiro Tamura MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. Molecular Biology and Evolution (submitted). 2015

Sulaiman IM, Lal AA, Xiao L. A population genetic study of the *Cryptosporidium parvum* human genotype parasites. J Eukaryot Microbiol. 2001;Suppl:24S-27S.

Sulaiman IM, Hira PR, Zhou L, Al-Ali FM, Al-Shelahi FA, Shweiki HM, Iqbal J, Khalid N, Xiao L. Unique endemicity of cryptosporidiosis in children in Kuwait. J Clin Microbiol. 2005 Jun;43(6):2805-9.

Suler D, Mullins D, Rudge T, Ashurst J. *Cryptosporidium parvum* Infection Following Contact with Livestock. N Am J Med Sci. 2016 Jul;8(7):323-5. doi: 10.4103/1947-2714.187162.

Sulzyc-Bielicka V, Kuźna-Grygiel W, Kołodziejczyk L, Bielicki D, Kładny J, Stepień-Korzonek M, Telatyńska-Smieszek B. Cryptosporidiosis in patients with colorectal cancer. J Parasitol. 2007 Jun;93(3):722-4.

Sulżyc-Bielicka V, Kołodziejczyk L, Jaczewska S, Bielicki D, Kładny J, Safranow K. Prevalence of *Cryptosporidium* sp. in patients with colorectal cancer. Pol Przegl Chir. 2012 Jul;84(7):348-51. doi: 10.2478/v10035-012-0058-4.

Sunderland D, Graczyk TK, Tamang L, Breysse PN. Impact of bathers on levels of *Cryptosporidium parvum* oocysts and Giardia lamblia cysts in recreational beach waters. Water Res. 2007 Aug;41(15):3483-9. Epub 2007 May 13.

Sungur T, Kar S, Güven E, Aktaş M, Karaer Z, Vatansever Z. [Detection of *Cryptosporidium spp*. in feces with nested PCR and carbol fuchsin staining method]. Turkiye Parazitol Derg. 2008;32(4):305-8.

Supcharassaeng S, Suankratay C. Antibiotic prescription for adults with acute diarrhea at King Chulalongkorn Memorial Hospital, Thailand. J Med Assoc Thai. 2011 May;94(5):545-50.

Sutton PL, Torres LP, Branch OH. Sexual recombination is a signature of a persisting malaria epidemic in Peru. Malar J. 2011 Oct 31;10:329. doi: 10.1186/1475-2875-10-329.

Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 2006 Jul 1;34(Web Server issue):W609-12.

Svenungsson B, Lagergren A, Ekwall E, Evengård B, Hedlund KO, Kärnell A, Löfdahl S, Svensson L, Weintraub A. Enteropathogens in adult patients with diarrhea and healthy control subjects: a 1-year prospective study in a Swedish clinic for infectious diseases. Clin Infect Dis. 2000 May;30(5):770-8. Epub 2000 May 18.

Swanson RV. Genome of Aquifex aeolicus. Methods Enzymol. 2001;330:158-69.

Swierczewski B, Odundo E, Ndonye J, Kirera R, Odhiambo C, Oaks E. Comparison of the Triage Micro Parasite Panel and Microscopy for the Detection of Entamoeba histolytica/Entamoeba dispar, Giardia lamblia, and *Cryptosporidium parvum* in Stool Samples Collected in Kenya. J Trop Med. 2012;2012:564721. doi: 10.1155/2012/564721. Epub 2012 Jul 10.

Tacconi EM, Tarsounas M. How homologous recombination maintains telomere integrity. Chromosoma. 2015 Jun;124(2):119-30. doi: 10.1007/s00412-014-0497-2. Epub 2014 Nov 29.

Taghipour N, Nazemalhosseini-Mojarad E, Haghighi A, Rostami-Nejad M, Romani S, Keshavarz A, Alebouyeh M, Zali M. Molecular epidemiology of cryptosporidiosis in Iranian children, tehran, iran. Iran J Parasitol. 2011 Dec;6(4):41-5.

Taguchi T, Shinozaki Y, Takeyama H, Haraguchi S, Yoshino M, Kaneko M, Ishimori Y, Matsunaga T. Direct counting of *Cryptosporidium parvum* oocysts using fluorescence in situ hybridization on a membrane filter. J Microbiol Methods. 2006 Nov;67(2):373-80. Epub 2006 Jun 21.

Taguchi T, Arakaki A, Takeyama H, Haraguchi S, Yoshino M, Kaneko M, Ishimori Y, Matsunaga T. Detection of *Cryptosporidium parvum* oocysts using a microfluidic device equipped with the SUS micromesh and FITC-labeled antibody. Biotechnol Bioeng. 2007 Feb 1;96(2):272-80.

Taherkhani H, Fallah M, Jadidian K, Vaziri S. A Study on the Prevalence of *Cryptosporidium* in HIV Positive Patients. J Res Health Sci. 2007 Dec 28;7(2):20-4.

Tahvildar-Biderouni F, Salehi N. Detection of *Cryptosporidium* infection by modified ziehl-neelsen and PCR methods in children with diarrheal samples in pediatric hospitals in Tehran. Gastroenterol Hepatol Bed Bench. 2014 Spring;7(2):125-30.

Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989 Nov;123(3):585-95

Tamer GS, Balikçi E, Erbay A. [The prevalence of cryptosporidiosis in children who were diagnosed with leukemia and lymphoma]. Turkiye Parazitol Derg. 2008;32(3):192-7.

Tamer GS, Turk M, Dagci H, Pektas B, Guy EC, Guruz AY, Uner A. The prevalence of cryptosporidiosis in Turkish children, and genotyping of isolates by nested polymerase chain reaction-restriction fragment length polymorphism. Saudi Med J. 2007 Aug;28(8):1243-6.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution. 2011; 28: 2731-2739.

Tang Y, Li N, Song M, Roellig DM, Feng Y, Xiao L. Development of a multilocus sequence typing tool for high-resolution subtyping and genetic structure characterization of *Cryptosporidium ubiquitum*. Infect Genet Evol. 2016 Sep 12;45:256-261. doi: 10.1016/j.meegid.2016.09.011. [Epub ahead of print]

Tanriverdi S, Blain JC, Deng B, Ferdig MT, Widmer G. Genetic crosses in the apicomplexan parasite *Cryptosporidium parvum* define recombination parameters. Mol Microbiol. 2007 Mar;63(5):1432-9.

Tassara R, Alarcón T, Larrañaga C, Wu E, Alvarez AM. [Digestive pathology in children infected with the human immunodeficiency virus (HIV), in Santiago de Chile]. Rev Med Chil. 2003 Jan;131(1):19-24.

Tavarez LA, Pena F, Placencia F, Mendoza HR, Polanco D. [Prevalence of protozoans in children with acute diarrheal disease]. Arch Domin Pediatr. 1991 May-Aug;27(2):43-7.

Taye B, Desta K, Ejigu S, Dori GU. The magnitude and risk factors of intestinal parasitic infection in relation to Human Immunodeficiency Virus infection and immune status, at ALERT Hospital, Addis Ababa, Ethiopia. Parasitol Int. 2014 Jun;63(3):550-6. doi: 10.1016/j.parint.2014.02.002. Epub 2014 Mar 4.

Taylor DN, Echeverria P, Pitarangsi C, Seriwatana J, Sethabutr O, Bodhidatta L, Brown C, Herrmann JE, Blacklow NR. Application of DNA hybridization techniques in the assessment of diarrheal disease among refugees in Thailand. Am J Epidemiol. 1988 Jan;127(1):179-87.

Teklemariam Z, Abate D, Mitiku H, Dessie Y. Prevalence of Intestinal Parasitic Infection among HIV Positive Persons Who Are Naive and on Antiretroviral Treatment in Hiwot Fana Specialized University Hospital, Eastern Ethiopia. ISRN AIDS. 2013 Jun 11;2013:324329. doi: 10.1155/2013/324329. eCollection 2013.

Tellevik MG, Moyo SJ, Blomberg B, Hjøllo T, Maselle SY, Langeland N, Hanevik K. Prevalence of *Cryptosporidium parvum*/hominis, Entamoeba histolytica and Giardia lamblia among Young Children with and without Diarrhea in Dar es Salaam, Tanzania. PLoS Negl Trop Dis. 2015 Oct 9;9(10):e0004125. doi: 10.1371/journal.pntd.0004125. eCollection 2015.

ten Hove R, Schuurman T, Kooistra M, Möller L, van Lieshout L, Verweij JJ. Detection of diarrhoea-causing protozoa in general practice patients in The Netherlands by multiplex real-time PCR. Clin Microbiol Infect. 2007 Oct;13(10):1001-7. Epub 2007 Aug 21.

Terletskaia-Ladwig E, Eggers M, Enders M, Regnath T. [Epidemiological aspects of gastrointestinal infections]. Dtsch Med Wochenschr. 2011 Jan;136(3):69-75. doi: 10.1055/s-0030-1269441. Epub 2011 Jan 11.

Thivierge K, Iqbal A, Dixon B, Dion R, Levesque B, Cantin P, Cédilotte L, Ndao M, Proulx JF, Yansouni CP. *Cryptosporidium hominis* Is a Newly Recognized Pathogen in the Arctic Region of Nunavik, Canada: Molecular Characterization of an Outbreak. PLoS Negl Trop Dis. 2016 Apr 8;10(4):e0004534. doi: 10.1371/journal.pntd.0004534. eCollection 2016.

Thomas DK, Lone AG, Selinger LB, Taboada EN, Uwiera RR, Abbott DW, Inglis GD. Comparative variation within the genome of Campylobacter jejuni NCTC 11168 in human and murine hosts. PLoS One. 2014 Feb 7;9(2):e88229. doi: 10.1371/journal.pone.0088229. eCollection 2014.

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994 Nov 11;22(22):4673-80.

Thompson LH, Schild D. The contribution of homologous recombination in preserving genome integrity in mammalian cells. Biochimie. 1999 Jan-Feb;81(1-2):87-105.

Tian LG, Chen JX, Wang TP, Cheng GJ, Steinmann P, Wang FF, Cai YC, Yin XM, Guo J, Zhou L, Zhou XN. Co-infection of HIV and intestinal parasites in rural area of China. Parasit Vectors. 2012 Feb 13;5:36. doi: 10.1186/1756-3305-5-36.

Tian LG, Cheng GJ, Chen JX, Cai YC, Guo J, Tong XM, Liu Q, Zhou XN. [Survey on co-infection with HIV and intestinal parasites in high prevalence areas of HIV/AIDS, China]. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi. 2012 Apr;24(2):168-72.

Tian LG, Wang TP, Lv S, Wang FF, Guo J, Yin XM, Cai YC, Dickey MK, Steinmann P, Chen JX. HIV and intestinal parasite co-infections among a Chinese population: an immunological profile. Infect Dis Poverty. 2013 Aug 23;2(1):18. doi: 10.1186/2049-9957-2-18.

Tiangtip R, Jongwutiwes S. Molecular analysis of *Cryptosporidium* species isolated from HIV-infected patients in Thailand. Trop Med Int Health. 2002 Apr;7(4):357-64.

Tibayrenc M, Kjellberg F, Ayala FJ. A clonal theory of parasitic protozoa: the population structures of Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas, and Trypanosoma and their medical and taxonomical consequences. Proc Natl Acad Sci U S A. 1990 Apr;87(7):2414-8.

Tikasingh ES, Smith S, Poon-King T, Bratt D. *Cryptosporidium* in man in Trinidad and Tobago. Trans R Soc Trop Med Hyg. 1986;80(6):998-9.

Timur AA, Wang QK. Fluorescence in situ hybridization in cardiovascular disease. Methods Mol Med. 2006;128:11-21.

Tiwari BR, Ghimire P, Malla S, Sharma B, Karki S. Intestinal parasitic infection among the HIV-infected patients in Nepal. J Infect Dev Ctries. 2013 Jul 15;7(7):550-5. doi: 10.3855/jidc.2785.

Tomita T, Bzik DJ, Ma YF, Fox BA, Markillie LM, Taylor RC, Kim K, Weiss LM. The Toxoplasma gondii cyst wall protein CST1 is critical for cyst wall integrity and promotes bradyzoite persistence. PLoS Pathog. 2013;9(12):e1003823. doi: 10.1371/journal.ppat.1003823. Epub 2013 Dec 26.

Torres ME, Pírez MC, Schelotto F, Varela G, Parodi V, Allende F, Falconi E, Dell'Acqua L, Gaione P, Méndez MV, Ferrari AM, Montano A, Zanetta E, Acuña AM, Chiparelli H, Ingold E. Etiology of children's diarrhea in Montevideo, Uruguay: associated pathogens and unusual isolates. J Clin Microbiol. 2001 Jun;39(6):2134-9.

Trotz-Williams LA, Martin DS, Gatei W, Cama V, Peregrine AS, Martin SW, Nydam DV, Jamieson F, Xiao L. Genotype and subtype analyses of *Cryptosporidium* isolates from dairy calves and humans in Ontario. Parasitol Res. 2006 Sep;99(4):346-52. Epub 2006 Mar 25.

Tsaihong JC, Ma P. Comparison of an indirect fluorescent antibody test and stool examination for the diagnosis of cryptosporidiosis. Eur J Clin Microbiol Infect Dis. 1990 Oct;9(10):770-3.

Tsubouchi H, Ogawa H. A novel mre11 mutation impairs processing of double-strand breaks of DNA during both mitosis and meiosis. Mol Cell Biol. 1998 Jan;18(1):260-8.

Tuli L, Gulati AK, Sundar S, Mohapatra TM. Correlation between CD4 counts of HIV patients and enteric protozoan in different seasons - an experience of a tertiary care hospital in Varanasi (India). BMC Gastroenterol. 2008 Aug 20;8:36. doi: 10.1186/1471-230X-8-36.

Tumwine JK, Kekitiinwa A, Nabukeera N, Akiyoshi DE, Rich SM, Widmer G, Feng X, Tzipori S. *Cryptosporidium parvum* in children with diarrhea in Mulago Hospital, Kampala, Uganda. Am J Trop Med Hyg. 2003 Jun;68(6):710-5.

Tumwine JK, Kekitiinwa A, Bakeera-Kitaka S, Ndeezi G, Downing R, Feng X, Akiyoshi DE, Tzipori S. Cryptosporidiosis and microsporidiosis in ugandan children with persistent diarrhea with and without concurrent infection with the human immunodeficiency virus. Am J Trop Med Hyg. 2005 Nov;73(5):921-5.

Turgay N, Yolasigmaz A, Erdogan DD, Zeyrek FY, Uner A. Incidence of cyclosporiasis in patients with gastrointestinal symptoms in western Turkey. Med Sci Monit. 2007 Jan;13(1):CR34-9. Epub 2006 Dec 18.

Turgay N, Unver-Yolasığmaz A, Oyur T, Bardak-Özcem S, Töz S. [Monthly distribution of intestinal parasites detected in a part of western Turkey between May 2009-April 2010-results of acid fast and modified trichrome staining methods]. Turkiye Parazitol Derg. 2012;36(2):71-4. doi: 10.5152/tpd.2012.18.

Turkcapar N, Kutlay S, Nergizoglu G, Atli T, Duman N. Prevalence of *Cryptosporidium* infection in hemodialysis patients. Nephron. 2002 Mar;90(3):344-6.

Turkiewicz D, Gorczyńska E, Wesołowska M, Kałwa K, Dyla A, Owoc J, Słociak M, Wójcik D, Ussowicz M, Chybicka A. Cryptosporidiosis in children undergoing allogeneic haematopoietic progenitur cell transplantation. 2006. Bone Marrow Transplant 37:S2, P786

Tyzzer EE. An extracellular Coccidium, *Cryptosporidium muris* (Gen. Et Sp. Nov.), of the gastric Glands of the Common Mouse. J Med Res. 1910 Nov;23(3):487-510.3.

Tzipori S, Campbell I. Prevalence of *Cryptosporidium* antibodies in 10 animal species. J Clin Microbiol. 1981 Oct;14(4):455-6.

Tzipori S. Cryptosporidiosis in perspective. Adv Parasitol. 1988;27:63-129.

Udgiri N, Minz M, Kashyap R, Heer M, Gupta CS, Mohandas K, Minz RW, Malla N. Intestinal cryptosporidiasis in living related renal transplant recipients. Transplant Proc. 2004 Sep;36(7):2128-9.

Ukwah BN, Ezeonu IM, Ezeonu CT, Roellig D, Xiao L. *Cryptosporidium* species and subtypes in diarrheal children and HIV-infected persons in Ebonyi and Nsukka, Nigeria. J Infect Dev Ctries. 2017 Feb 28;11(2):173-179. doi: 10.3855/jidc.8034.

UNAIDS. Report on the global HIV/AIDS epidemic, 2000. UNAIDS. Joint United Nations Programme on HIV/AIDS , June 2000. Web. <http://data.unaids.org/pub/Report/2000/2000_gr_en.pdf>.

UNAIDS. Global update on HIV treatment, 2013: results, impact and opportunities. WHO report in partnership with UNICEF and UNAIDS. June, 2013. Print.

UNAIDS. Global HIV Statistics. UNAIDS Fact Sheet 2016. UNAIDS, Nov. 2016. Web. 05 Feb. 2016. <http://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf>.

UniProt Consortium. UniProt: a hub for protein information. Nucleic Acids Res. 2015 Jan;43(Database issue):D204-12. doi: 10.1093/nar/gku989. Epub 2014 Oct 27.

Uppal B, Kashyap B, Bhalla P. Enteric Pathogens in HIV/AIDS from a Tertiary Care Hospital. Indian J Community Med. 2009 Jul;34(3):237-42. doi: 10.4103/0970-0218.55291.

Upton, SJ. "Basic Biology of *Cryptosporidium*." Division of Biology, KSU. Kansas State University, 17 Feb. 2008. Web. 15 Oct. 2016.

Utsi L, Smith SJ, Chalmers RM, Padfield S. Cryptosporidiosis outbreak in visitors of a UK industry-compliant petting farm caused by a rare *Cryptosporidium parvum* subtype: a case-control study. Epidemiol Infect. 2016 Apr;144(5):1000-9. doi: 10.1017/S0950268815002319. Epub 2015 Oct 1.

Uysal S, Tunalı V, Akdur Öztürk E, Ardeniz Ö, Işıkgöz Taşbakan M, Pullukçu H, Özensoy Töz S, Turgay N, Arda B. Incidence of Parasitic Diarrhea in Patients with Common Variable Immune Deficiency. Turkiye Parazitol Derg. 2016 Jun;40(2):67-71. doi: 10.5152/tpd.2016.4687.

Vahedi M, Gohardehi S, Sharif M, Daryani A. Prevalence of parasites in patients with gastroenteritis at East of Mazandaran Province, Northern Iran. Trop Biomed. 2012 Dec;29(4):568-74.

Valentiner-Branth P, Perch M, Nielsen J, Steinsland H, Garly ML, Fischer TK, Sommerfelt H, Mølbak K, Aaby P. Community cohort study of *Cryptosporidium parvum* infections: sex-differential incidences associated with BCG and diphtheria-tetanus-pertussis vaccinations. Vaccine. 2007 Mar 30;25(14):2733-41. Epub 2006 Feb 2.

Valenzuela O, González-Díaz M, Garibay-Escobar A, Burgara-Estrella A, Cano M, Durazo M, Bernal RM, Hernandez J, Xiao L. Molecular characterization of *Cryptosporidium spp*. in children from Mexico. PLoS One. 2014 Apr 22;9(4):e96128. doi: 10.1371/journal.pone.0096128. eCollection 2014.

Van Cleave, H. J. Morphological and phylogenetic interpretations of the cement glands in the Acanthocephala. J Morphol. 1949 May;84(3):427-57.

van Kraaij MG, Dekker AW, Verdonck LF, van Loon AM, Vinjé J, Koopmans MP, Rozenberg-Arska M. Infectious gastro-enteritis: an uncommon cause of diarrhoea in adult allogeneic and autologous stem cell transplant recipients. Bone Marrow Transplant. 2000 Aug;26(3):299-303.

Van Lint P, Rossen JW, Vermeiren S, Ver Elst K, Weekx S, Van Schaeren J, Jeurissen A. Detection of Giardia lamblia, *Cryptosporidium spp*. and Entamoeba histolytica in clinical stool samples by using multiplex real-time PCR after automated DNA isolation. Acta Clin Belg. 2013 May-Jun;68(3):188-92.

Vasco K, Graham JP, Trueba G. Detection of Zoonotic Enteropathogens in Children and Domestic Animals in a Semirural Community in Ecuador. Appl Environ Microbiol. 2016 Jun 30;82(14):4218-24. doi: 10.1128/AEM.00795-16. Print 2016 Jul 15.

Velasquez DE, Arvelo W, Cama VA, López B, Reyes L, Roellig DM, Kahn GD, Lindblade KA. Short report: Molecular insights for Giardia, *Cryptosporidium*, and soil-transmitted helminths from a facility-based surveillance system in Guatemala. Am J Trop Med Hyg. 2011 Dec;85(6):1141-3. doi: 10.4269/ajtmh.2011.11-0437.

Vesey G, Ashbolt N, Fricker EJ, Deere D, Williams KL, Veal DA, Dorsch M. The use of a ribosomal RNA targeted oligonucleotide probe for fluorescent labelling of viable *Cryptosporidium parvum* oocysts. J Appl Microbiol. 1998 Sep;85(3):429-40.

Vetterling JM, Jervis HR, Merrill TG, Sprinz H. *Cryptosporidium* wrairi sp. n. from the guinea pig Cavia porcellus, with an emendation of the genus. J Protozool. 1971 May;18(2):243-7.

Vieira PM, Mederle N, Lobo ML, Imre K, Mederle O, Xiao L, Darabus G, Matos O. Molecular characterisation of *Cryptosporidium* (Apicomplexa) in children and cattle in Romania. Folia Parasitol (Praha). 2015 Jan 1;62. pii: 2015.002. doi: 10.14411/fp.2015.002.

Vignesh R, Balakrishnan P, Shankar EM, Murugavel KG, Hanas S, Cecelia AJ, Thyagarajan SP, Solomon S, Kumarasamy N. High proportion of isosporiasis among HIV-infected patients with diarrhea in southern India. Am J Trop Med Hyg. 2007 Nov;77(5):823-4.

Vink C, Rudenko G, Seifert HS. Microbial antigenic variation mediated by homologous DNA recombination. FEMS Microbiol Rev. 2012 Sep;36(5):917-48. doi: 10.1111/j.1574-6976.2011.00321.x. Epub 2012 Jan 17.

Viriyavejakul P, Nintasen R, Punsawad C, Chaisri U, Punpoowong B, Riganti M. High prevalence of Microsporidium infection in HIV-infected patients. Southeast Asian J Trop Med Public Health. 2009 Mar;40(2):223-8.

Volkman SK, Barry AE, Lyons EJ, Nielsen KM, Thomas SM, Choi M, Thakore SS, Day KP, Wirth DF, Hartl DL. Recent origin of Plasmodium falciparum from a single progenitor. Science. 2001 Jul 20;293(5529):482-4.

Voth DE, Broederdorf LJ, Graham JG. Bacterial Type IV secretion systems: versatile virulence machines. Future Microbiol. 2012 Feb;7(2):241-57. doi: 10.2217/fmb.11.150.

Vouking MZ, Enoka P, Tamo CV, Tadenfok CN. Prevalence of intestinal parasites among HIV patients at the Yaoundé Central Hospital, Cameroon. Pan Afr Med J. 2014 Jun 11;18:136. doi: 10.11604/pamj.2014.18.136.3052. eCollection 2014.

Vyas N, Pathan N, Aziz A. Enteric pathogens in HIV-positive patients with diarrhoea and their correlation with CD4+ T-lymphocyte counts. Trop Parasitol. 2012 Jan;2(1):29-34. doi: 10.4103/2229-5070.97236.

Vyas N, Sood S, Sharma B, Kumar M. The Prevalence of Intestinal Parasitic Infestation and the Related Profile of the CD4 (+) Counts in HIV/AIDS People with Diarrhoea in Jaipur City. J Clin Diagn Res. 2013 Mar;7(3):454-6. doi: 10.7860/JCDR/2013/4817.2797. Epub 2013 Mar 1.

Waldron LS, Ferrari BC, Gillings MR, Power ML. Terminal restriction fragment length polymorphism for identification of *Cryptosporidium* species in human feces. Appl Environ Microbiol. 2009 Jan;75(1):108-12. doi: 10.1128/AEM.01341-08. Epub 2008 Oct 31.

Waldron LS, Ferrari BC, Power ML. Glycoprotein 60 diversity in *C. hominis* and *C. parvum* causing human cryptosporidiosis in NSW, Australia. Exp Parasitol. 2009 Jun;122(2):124-7. doi: 10.1016/j.exppara.2009.02.006. Epub 2009 Feb 20.

Waldron LS, Dimeski B, Beggs PJ, Ferrari BC, Power ML. Molecular epidemiology, spatiotemporal analysis, and ecology of sporadic human cryptosporidiosis in Australia. Appl Environ Microbiol. 2011 Nov;77(21):7757-65. doi: 10.1128/AEM.00615-11. Epub 2011 Sep 9.

Wang KX, Li CP, Wang J, Pan BR. Epidemiological survey of cryptosporidiosis in Anhui Province China. World J Gastroenterol. 2002 Apr;8(2):371-4.

Wang L, Zhang H, Zhao X, Zhang L, Zhang G, Guo M, Liu L, Feng Y, Xiao L. Zoonotic *Cryptosporidium* species and Enterocytozoon bieneusi genotypes in HIV-positive patients on antiretroviral therapy. J Clin Microbiol. 2013 Feb;51(2):557-63. doi: 10.1128/JCM.02758-12. Epub 2012 Dec 5.

Wang QQ, Guo JD, Cao ZG, Wang QZ, Liu DH, Wang TP. [Investigation on human *Cryptosporidium* infection in local area of Anhui Province]. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi. 2015 Jun;27(3):263-, 272.

Wang Y, Tian RM, Gao ZM, Bougouffa S, Qian PY. Optimal eukaryotic 18S and universal 16S/18S ribosomal RNA primers and their application in a study of symbiosis. PLoS One. 2014 Mar 3;9(3):e90053. doi: 10.1371/journal.pone.0090053. eCollection 2014.

Wang Z, Martin J, Abubucker S, Yin Y, Gasser RB, Mitreva M. Systematic analysis of insertions and deletions specific to nematode proteins and their proposed functional and evolutionary relevance. BMC Evol Biol. 2009 Jan 28;9:23. doi: 10.1186/1471-2148-9-23.

Wanyiri JW, Kanyi H, Maina S, Wang DE, Ngugi P, O'Connor R, Kamau T, Waithera T, Kimani G, Wamae CN, Mwamburi M, Ward HD. Infectious diarrhoea in antiretroviral therapy-naive HIV/AIDS patients in Kenya. Trans R Soc Trop Med Hyg. 2013 Oct;107(10):631-8. doi: 10.1093/trstmh/trt078.

Wanyiri JW, Kanyi H, Maina S, Wang DE, Steen A, Ngugi P, Kamau T, Waithera T, O'Connor R, Gachuhi K, Wamae CN, Mwamburi M, Ward HD. Cryptosporidiosis in HIV/AIDS patients in Kenya: clinical features, epidemiology, molecular characterization and antibody responses. Am J Trop Med Hyg. 2014 Aug;91(2):319-28. doi: 10.4269/ajtmh.13-0254. Epub 2014 May 27.

Ward BJ, van Oosterhout C. HYBRIDCHECK: software for the rapid detection, visualization and dating of recombinant regions in genome sequence data. Mol Ecol Resour. 2016 Mar;16(2):534-9. doi: 10.1111/1755-0998.12469. Epub 2015 Oct 12.

Warnow, T. Concatenation Analyses in the Presence of Incomplete Lineage Sorting. Version 1. PLoS Curr. 2015 May 22; 7: ecurrents.currents.tol.8d41ac0f13d1abedf4c4a59f5d17b1f7.

Waywa D, Kongkriengdaj S, Chaidatch S, Tiengrim S, Kowadisaiburana B, Chaikachonpat S, Suwanagool S, Chaiprasert A, Curry A, Bailey W, Suputtamongkol Y, Beeching NJ. Protozoan enteric infection in AIDS related diarrhea in Thailand. Southeast Asian J Trop Med Public Health. 2001;32 Suppl 2:151-5.

Weatherly DB, Peng D, Tarleton RL. Recombination-driven generation of the largest pathogen repository of antigen variants in the protozoan Trypanosoma cruzi. BMC Genomics. 2016 Sep 13;17(1):729. doi: 10.1186/s12864-016-3037-z.

Webb LM, Tubach SA, Hunt DC; Centers for Disease Control and Prevention (CDC). Outbreak of cryptosporidiosis among responders to a rollover of a truck carrying calves - Kansas, April 2013. MMWR Morb Mortal Wkly Rep. 2014 Dec 19;63(50):1185-8.

Weber CC, Nabholz B, Romiguier J, Ellegren H. Kr/Kc but not dN/dS correlates positively with body mass in birds, raising implications for inferring lineage-specific selection. Genome Biol. 2014;15(12):542.

Weber, JL. Interspersed repetitive DNA from Plasmodium falciparum. Mol Biochem Parasitol. 1988 Jun;29(2-3):117-24.

Weedall GD, Hall N. Sexual reproduction and genetic exchange in parasitic protists. Parasitology. 2015 Feb;142 Suppl 1:S120-7. doi: 10.1017/S0031182014001693. Epub 2014 Dec 22.

Wegayehu T, Adamu H, Petros B. Prevalence of Giardia duodenalis and *Cryptosporidium* species infections among children and cattle in North Shewa Zone, Ethiopia. BMC Infect Dis. 2013 Sep 8;13:419. doi: 10.1186/1471-2334-13-419.

Weisberg AJ, Elmarakeby HA, Heath LS, Vinatzer BA. Similarity-based codes sequentially assigned to ebolavirus genomes are informative of species membership, associated outbreaks, and transmission chains. Open Forum Infect Dis. 2015 Mar 12;2(1):ofv024. doi: 10.1093/ofid/ofv024. eCollection 2015.

Weiss, L. M., & Kim, K. Toxoplasma gondii: The model apicomplexan: Perspectives and methods. 2007. London: Academic Press.

Wenyon, CM. Protozoology: a manual for medical men, veterinarians and zoologists. Vol. II. London: Baillière. 1926. Print.

Werneck-Silva AL, Prado IB. Gastroduodenal opportunistic infections and dyspepsia in HIV-infected patients in the era of Highly Active Antiretroviral Therapy. J Gastroenterol Hepatol. 2009 Jan;24(1):135-9. doi: 10.1111/j.1440-1746.2008.05700.x. Epub 2008 Nov 26.

Werner A, Majewska AC, Słodkowicz A. Prevalence of intestinal protozoan parasites in humans from the Poznań area. 2001. Wiad Parazyt 47(suppl 2):51

Werner A, Majewska AC, Słodkowicz A, Juszczyk J, Barałkiewicz G. Frequency of occurrence of intestinal parasites in inhabitants of Poznań and its neighbourhood. 2004. Wiad Parazytol 50(suppl):127

Wesołowska M, Ditrich O. *Cryptosporidium parvum* in HIV patients from Wrocław. 2003. Wiad Parazytol 49:401

Wesołowska M, Mowszet K, Wróbel G, Jankowski S. [Cryptosporidiosis in children with chronic diarrhea]. Wiad Parazytol. 2004;50(3):393-6.

West SA, Smith TG, Read AF. Sex allocation and population structure in apicomplexan (protozoa) parasites. Proc Biol Sci. 2000 Feb 7;267(1440):257-63.

Wetzel R. Ein neues Coccid (*Cryptosporidium* vulpis sp. nov.) aus dem Rotfuchs. Arch. Wiss. Prakt. Tierheilk. 1938; 71, 39-40.

Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Tatusova TA, Wagner L. Database resources of the National Center for Biotechnology. Nucleic Acids Res. 2003 Jan 1;31(1):28-33.

Wheeler, W. M. The Phylogenetic Development of Subapterous and Apterous Castes in the Formicidae. Proc Natl Acad Sci U S A. 1917 Feb;3(2):109-17.

WHO/UNAIDS/UN. The Millennium Development Goals Report/WHO HIV/AIDS global data and statistics. 2009; p.32. Accessed from: http://www.who.int/hiv/data/global_data/en/

Widmer G, Tchack L, Chappell CL, Tzipori S. Sequence polymorphism in the beta-tubulin gene reveals heterogeneous and variable population structures in *Cryptosporidium parvum*. Appl Environ Microbiol. 1998 Nov;64(11):4477-81.

Widmer G, Akiyoshi D, Buckholt MA, Feng X, Rich SM, Deary KM, Bowman CA, Xu P, Wang Y, Wang X, Buck GA, Tzipori S. Animal propagation and genomic survey of a genotype 1 isolate of *Cryptosporidium parvum*. Mol Biochem Parasitol. 2000 May;108(2):187-97.

Widmer G, Lee Y. Comparison of single- and multilocus genetic diversity in the protozoan parasites *Cryptosporidium parvum* and *C. hominis*. Appl Environ Microbiol. 2010 Oct;76(19):6639-44. doi: 10.1128/AEM.01268-10. Epub 2010 Aug 13.

Widmer, G, Sullivan, S. Genomics and population biology of *Cryptosporidium* species. Parasite Immunol. 2012 Feb-Mar;34(2-3):61-71. doi: 10.1111/j.1365-3024.2011.01301.x.

Widmer G, Lee Y, Hunt P, Martinelli A, Tolkoff M, Bodi K. Comparative genome analysis of two *Cryptosporidium parvum* isolates with different host range. Infect Genet Evol. 2012 Aug;12(6):1213-21. doi: 10.1016/j.meegid.2012.03.027. Epub 2012 Apr 12.

Wielinga PR, de Vries A, van der Goot TH, Mank T, Mars MH, Kortbeek LM, van der Giessen JW. Molecular epidemiology of *Cryptosporidium* in humans and cattle in The Netherlands. Int J Parasitol. 2008 Jun;38(7):809-17. Epub 2007 Nov 4.

Williams, LE, Wernegreen, JJ. Sequence context of indel mutations and their effect on protein evolution in a bacterial endosymbiont. Genome Biol Evol. 2013;5(3):599-605. doi: 10.1093/gbe/evt033.

Winkelstein JA, Marino MC, Ochs H, Fuleihan R, Scholl PR, Geha R, Stiehm ER, Conley ME. The X-linked hyper-IgM syndrome: clinical and immunologic features of 79 patients. Medicine (Baltimore). 2003 Nov;82(6):373-84.

Witmer K, Schmid CD, Brancucci NM, Luah YH, Preiser PR, Bozdech Z, Voss TS. Analysis of subtelomeric virulence gene families in Plasmodium falciparum by comparative transcriptional profiling. Mol Microbiol. 2012 Apr;84(2):243-59. doi: 10.1111/j.1365-2958.2012.08019.x. Epub 2012 Mar 22.

Wiwanitkit V. Intestinal parasitic infections in Thai HIV-infected patients with different immunity status. BMC Gastroenterol. 2001;1:3. Epub 2001 Jun 1.

Won EJ, Kim SH, Kee SJ, Shin JH, Suh SP, Chai JY, Ryang DW, Shin MG. Multiplex Real-Time PCR Assay Targeting Eight Parasites Customized to the Korean Population: Potential Use for Detection in Diarrheal Stool Samples from Gastroenteritis Patients. PLoS One. 2016 Nov 18;11(11):e0166957. doi: 10.1371/journal.pone.0166957. eCollection 2016.

Wongstitwilairoong B, Srijan A, Serichantalergs O, Fukuda CD, McDaniel P, Bodhidatta L, Mason CJ. Intestinal parasitic infections among pre-school children in Sangkhlaburi, Thailand. Am J Trop Med Hyg. 2007 Feb;76(2):345-50.

World Bank. Net migration. Net migration | Data. The World Bank, n.d. Web. 22 Dec. 2016. <http://data.worldbank.org/indicator/SM.POP.NETM>.

Wright S. Evolution and the Genetics of Populations. Variability within and among Natural Population. 1978. Vol. IV. Chicago, Illinois: University of Chicago Press.

Wu Z, Nagano I, Boonmars T, Nakada T, Takahashi Y. Intraspecies polymorphism of *Cryptosporidium parvum* revealed by PCR-restriction fragment length polymorphism (RFLP) and RFLP-single-strand conformational polymorphism analyses. Appl Environ Microbiol. 2003 Aug;69(8):4720-6.

Wumba R, Enache-Angoulvant A, Develoux M, Mulumba A, Mulumba PM, Hennequin C, Odio TW, Biligui S, Sala J, Thellier M. [Prevalence of opportunistic digestive parasitic infections in Kinshasa, Democratic Republic of Congo. Results of a preliminary study in 50 AIDS patients]. Med Trop (Mars). 2007 Apr;67(2):145-8.

Wumba R, Longo-Mbenza B, Mandina M, Odio WT, Biligui S, Sala J, Breton J, Thellier M. Intestinal parasites infections in hospitalized AIDS patients in Kinshasa, Democratic Republic of Congo. Parasite. 2010 Dec;17(4):321-8.

Wumba R, Longo-Mbenza B, Menotti J, Mandina M, Kintoki F, Situakibanza NH, Kakicha MK, Zanga J, Mbanzulu-Makola K, Nseka T, Mukendi JP, Kendjo E, Sala J, Thellier M. Epidemiology, clinical, immune, and molecular profiles of microsporidiosis and cryptosporidiosis among HIV/AIDS patients. Int J Gen Med. 2012;5:603-11. doi: 10.2147/IJGM.S32344. Epub 2012 Jul 19.

Wumba RD, Zanga J, Mbanzulu KM, Mandina MN, Kahindo AK, Aloni MN, Ekila MB. *Cryptosporidium* identification in HIV-Infected Humans. Experience from Kinshasa, the Democratic Republic of Congo. Acta Parasitol. 2015 Dec;60(4):638-44. doi: 10.1515/ap-2015-0090.

Xiao L, Escalante L, Yang C, Sulaiman I, Escalante AA, Montali RJ, Fayer R, Lal AA. Phylogenetic analysis of *Cryptosporidium* parasites based on the small-subunit rRNA gene locus. Appl Environ Microbiol. 1999 Apr;65(4):1578-83.

Xiao L, Bern C, Limor J, Sulaiman I, Roberts J, Checkley W, Cabrera L, Gilman RH, Lal AA. Identification of 5 types of *Cryptosporidium* parasites in children in Lima, Peru. J Infect Dis. 2001 Feb 1;183(3):492-7. Epub 2000 Dec 20.

Xiao L, Sulaiman IM, Ryan UM, Zhou L, Atwill ER, Tischler ML, Zhang X, Fayer R, Lal AA. Host adaptation and host-parasite co-evolution in *Cryptosporidium*: implications for taxonomy and public health. Int J Parasitol. 2002 Dec 19;32(14):1773-85.

Xiao L, Fayer R, Ryan U, Upton SJ. *Cryptosporidium* taxonomy: recent advances and implications for public health. Clin Microbiol Rev. 2004 Jan;17(1):72-97.

Xiao L, Hlavsa MC, Yoder J, Ewers C, Dearen T, Yang W, Nett R, Harris S, Brend SM, Harris M, Onischuk L, Valderrama AL, Cosgrove S, Xavier K, Hall N, Romero S, Young S, Johnston SP, Arrowood M, Roy S, Beach MJ. Subtype analysis of *Cryptosporidium* specimens from sporadic cases in Colorado, Idaho, New Mexico, and Iowa in 2007: widespread occurrence of one *Cryptosporidium hominis* subtype and case history of an infection with the *Cryptosporidium* horse genotype. J Clin Microbiol. 2009 Sep;47(9):3017-20. doi: 10.1128/JCM.00226-09. Epub 2009 Jul 8.

Xiao L. Molecular epidemiology of cryptosporidiosis: an update. Exp Parasitol. 2010 Jan;124(1):80-9. doi: 10.1016/j.exppara.2009.03.018. Epub 2009 Apr 7.

Xiao W, Liu H, Li Y, Li X, Xu C, Long M, Wang S. A rice gene of de novo origin negatively regulates pathogen-induced defense response. PLoS One. 2009;4(2):e4603. doi: 10.1371/journal.pone.0004603. Epub 2009 Feb 25.

Xu P, Widmer G, Wang Y, Ozaki LS, Alves JM, Serrano MG, Puiu D, Manque P, Akiyoshi D, Mackey AJ, Pearson WR, Dear PH, Bankier AT, Peterson DL, Abrahamsen MS, Kapur V, Tzipori S, Buck GA. The genome of *Cryptosporidium hominis*. Nature. 2004 Oct 28;431(7012):1107-12.

Yadav P, Khalil S, Mirdha BR. Molecular appraisal of intestinal parasitic infection in transplant recipients. Indian J Med Res. 2016 Aug;144(2):258-263. doi: 10.4103/0971-5916.195041.

Yager P, Edwards T, Fu E, Helton K, Nelson K, Tam MR, Weigl BH. Microfluidic diagnostic technologies for global public health. Nature. 2006 Jul 27;442(7101):412-8.

Yagita K, Izumiyama S, Tachibana H, Masuda G, Iseki M, Furuya K, Kameoka Y, Kuroki T, Itagaki T, Endo T. Molecular characterization of *Cryptosporidium* isolates obtained from human and bovine infections in Japan. Parasitol Res. 2001 Nov;87(11):950-5.

Yakoob J, Abbas Z, Beg MA, Naz S, Khan R, Islam M, Jafri W. Prevalences of Giardia lamblia and *Cryptosporidium parvum* infection in adults presenting with chronic diarrhoea. Ann Trop Med Parasitol. 2010 Sep;104(6):505-10. doi: 10.1179/136485910X12786389891209.

Yang Y, Zhou Y, Cheng W, Pan X, Xiao P, Shi Y, Gao J, Song X, Chen Y, Jiang Q. Prevalence and Determinants of *Cryptosporidium* Infection in an Underdeveloped Rural Region of Southwestern China. Am J Trop Med Hyg. 2017 Jan 16. pii: 16-0732. doi: 10.4269/ajtmh.16-0732. [Epub ahead of print]

Yang, Z., Bielawski, J. P. Statistical methods for detecting molecular adaptation. Trends Ecol Evol. 2000 Dec 1;15(12):496-503.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 2012 Jun 18;13:134. doi: 10.1186/1471-2105-13-134.

Ye J, Xiao L, Ma J, Guo M, Liu L, Feng Y. Anthroponotic enteric parasites in monkeys in public park, China. Emerg Infect Dis. 2012 Oct;18(10):1640-3. doi: 10.3201/eid1810.120653.

Yones DA, Galal LA, Abdallah AM, Zaghlol KS. Effect of enteric parasitic infection on serum trace elements and nutritional status in upper Egyptian children. Trop Parasitol. 2015 Jan-Jun;5(1):29-35. doi: 10.4103/2229-5070.145581.

Youssef M, Shurman A, Bougnoux M, Rawashdeh M, Bretagne S, Strockbine N. Bacterial, viral and parasitic enteric pathogens associated with acute diarrhea in hospitalized children from northern Jordan. FEMS Immunol Med Microbiol. 2000 Jul;28(3):257-63.

Yu JR, Lee JK, Seo M, Kim SI, Sohn WM, Huh S, Choi HY, Kim TS. Prevalence of cryptosporidiosis among the villagers and domestic animals in several rural areas of Korea. Korean J Parasitol. 2004 Mar;42(1):1-6.

Yu Z, Li F, Zeng Z, Huang Z, Fan Z, Jin Y, Luo W, Xiang X, Deng Q. Prevalence and clinical significance of *Cryptosporidium* infection in patients with hepatitis B virus-associated acute-on-chronic liver failure. Int J Infect Dis. 2011 Dec;15(12):e845-8. doi: 10.1016/j.ijid.2011.08.006. Epub 2011 Oct 12.

Zaglool DA, Mohamed A, Khodari YA, Farooq MU. Crypto-Giardia antigen rapid test versus conventional modified Ziehl-Neelsen acid fast staining method for diagnosis of cryptosporidiosis. Asian Pac J Trop Med. 2013 Mar;6(3):212-5. doi: 10.1016/S1995-7645(13)60025-5.

Zaidah AR, Chan YY, Asma HS, Abdullah S, Nurhaslindawati AR, Salleh M, Zeehaida M, Lalitha P, Mustafa M, Ravichandran M. Detection of *Cryptosporidium parvum* in HIV-infected patients in Malaysia using a molecular approach. Southeast Asian J Trop Med Public Health. 2008 May;39(3):511-6.

Zali MR, Mehr AJ, Rezaian M, Meamar AR, Vaziri S, Mohraz M. Prevalence of intestinal parasitic pathogens among HIV-positive individuals in Iran. Jpn J Infect Dis. 2004 Dec;57(6):268-70.

Zeliff CC. Manual of medical parasitology with techniques for laboratory diagnosis and notes on related animal parasites. Pennsylvania: State College. 1947. Print.

Zeynudin A, Hemalatha K, Kannan S. Prevalence of opportunistic intestinal parasitic infection among HIV infected patients who are taking antiretroviral treatment at Jimma Health Center, Jimma, Ethiopia. Eur Rev Med Pharmacol Sci. 2013 Feb;17(4):513-6.

Zhang L, Fu Y, Jing W, Xu Q, Zhao W, Feng M, Tachibana H, Sui G, Cheng X. Rapid microfluidic immunoassay for surveillance and diagnosis of *Cryptosporidium* infection in human immunodeficiency virus-infected patients. Biomicrofluidics. 2015 Apr 14;9(2):024114. doi: 10.1063/1.4916229. eCollection 2015.

Zhang SX, Zhou YM, Xu W, Tian LG, Chen JX, Chen SH, Dang ZS, Gu WP, Yin JW, Serrano E, Zhou XN. Impact of co-infections with enteric pathogens on children suffering from acute diarrhea in southwest China. Infect Dis Poverty. 2016 Jun 27;5(1):64. doi: 10.1186/s40249-016-0157-2.

Zhang SX, Yang CL, Gu WP, Ai L, Serrano E, Yang P, Zhou X, Li SZ, Lv S, Dang ZS, Chen JH, Hu W, Tian LG, Chen JX, Zhou XN. Case-control study of diarrheal disease etiology in individuals over 5 years in southwest China. Gut Pathog. 2016 Nov 16;8:58. eCollection 2016.

Zhao Y, Davis RE, Lee IM. Phylogenetic positions of 'Candidatus Phytoplasma asteris' and Spiroplasma kunkelii as inferred from multiple sets of concatenated core housekeeping proteins. Int J Syst Evol Microbiol. 2005 Sep;55(Pt 5):2131-41.

Zhong Y, Hu S, Xu C, Zhao Y, Xu D, Zhao Y, Zhao J, Li Z, Zhang X, Zhang H, Li J. A novel method for detection of HBVcccDNA in hepatocytes using rolling circle amplification combined with in situ PCR. BMC Infect Dis. 2014 Dec 3;14:608. doi: 10.1186/s12879-014-0608-y.

Zhou L, Singh A, Jiang J, Xiao L. Molecular surveillance of *Cryptosporidium spp*. in raw wastewater in Milwaukee: implications for understanding outbreak occurrence and transmission dynamics. J Clin Microbiol. 2003 Nov;41(11):5254-7.

Ziehl F. Zur Farbung des Tuberkelbaclilus. Deutsche Medizinische Wochenschrif. 1882; 8,451.

Zintl A, Mulcahy G, de Waal T, de Waele V, Byrne C, Clyne M, Holden N, Fanning S. An Irish perspective on *Cryptosporidium*. Part 1. Ir Vet J. 2006 Aug 1;59(8):442-7. doi: 10.1186/2046-0481-59-8-442.

Zintl A, Proctor AF, Read C, Dewaal T, Shanaghy N, Fanning S, Mulcahy G. The prevalence of *Cryptosporidium* species and subtypes in human faecal samples in Ireland. Epidemiol Infect. 2009 Feb;137(2):270-7. doi: 10.1017/S0950268808000769. Epub 2008 May 12.

Zintl A, Keogh B, Ezzaty-Mirhashemi M, De Waal T, Scholz D, Mulcahy G. Survival of *Cryptosporidium parvum* oocysts in the presence of hydrated lime. Vet Rec. 2010 Mar 6;166(10):297-300. doi: 10.1136/vr.c1157.