2022

Novel Molecular Approaches to Diagnostic and Epidemiological Analyses of Cryptosporidium spp. Implicated in Clinical Infection

Jennifer K. O'Leary
Department of Biological Sciences, Munster Technological University, Cork, Ireland, R00059070@mycit.ie

Follow this and additional works at: https://sword.cit.ie/allthe

Part of the Parasitology Commons

Recommended Citation
O'Leary, Jennifer K., "Novel Molecular Approaches to Diagnostic and Epidemiological Analyses of Cryptosporidium spp. Implicated in Clinical Infection" (2022). Theses [online]. Available at: https://sword.cit.ie/allthe/42

This Doctoral Thesis is brought to you for free and open access by the Dissertations and Theses at SWORD - South West Open Research Deposit. It has been accepted for inclusion in Theses by an authorized administrator of SWORD - South West Open Research Deposit. For more information, please contact sword@cit.ie.
Novel Molecular Approaches to Diagnostic and Epidemiological Analyses of Cryptosporidium spp. Implicated in Clinical Infection

A dissertation presented to the Munster Technological University for the degree of

Doctor of Philosophy

by

Jennifer O’Leary, B.Sc.

Department of Biological Sciences, Munster Technological University, Bishopstown Campus, Cork, Ireland.

Research supervisors: Prof. Roy D. Sleator & Dr Brigid Lucey
Submitted to Munster Technological University

November 2021
Declaration

I declare that this thesis, which I submit to Munster Technological University, is my own personal effort. Where any of the content presented is the result of input or data from a related collaborative research programme, this is duly acknowledged, such that it is possible to ascertain how much of the work is my own. I have not already obtained a degree in Munster Technological University or elsewhere on the basis of this work. Furthermore, I took reasonable care to ensure that the work is original, and, to the best of my knowledge, does not breach copyright law, and has not been taken from other sources except where such work has been cited and acknowledged within the text.

Student signature:  
Jennifer O’Leary

Student number: R00059070

Supervisor signature:  
Prof. Roy D. Sleator

Supervisor signature:  
Dr Brigid Lucey

Date: 18-November-2021
For my parents, Angela and Michael
## Table of Contents

Acknowledgements ............................................................................................................. 1  
List of Posters, Presentations, Publications & Scientific Communications ...................... 3  
List of Abbreviations ........................................................................................................... 5  
Thesis Abstract ...................................................................................................................... 7  

Chapter I: Literature review - Cryptosporidium epidemiology, diagnosis, and research in the 21st century ................................................................................................. 9  

1.1 Abstract .......................................................................................................................... 10  
1.2 Introduction ...................................................................................................................... 11  
1.3 Taxonomy ........................................................................................................................ 12  
1.4 Epidemiology ................................................................................................................... 14  
1.4.1 A Global Perspective .................................................................................................. 14  
1.4.2 Epidemiology within Europe ..................................................................................... 16  
1.5 Life Cycle ........................................................................................................................ 18  
1.6 Pathophysiology .............................................................................................................. 22  
1.7 Clinical Features and Patient Spectrum ......................................................................... 23  
1.8 Diagnosis ........................................................................................................................ 25  
1.8.1 Brightfield and Fluorescent Staining ........................................................................ 28  
1.8.2 Enzyme Immunoassays, ELISA and Immunochromatographic Methods ............ 30  
1.8.3 Molecular Approaches ............................................................................................. 31  
1.9 Future Trends in Molecular Platforms ......................................................................... 38  
1.10 Future Trends in Omics Research ................................................................................. 40  
1.11 Treatment Limitations, Propagation via Cell Culture and Implications for Future Advancements .................................................................................................................... 41  
1.12 Concluding Remarks ...................................................................................................... 44
Chapter II: Validation of the EntericBio GastroPanel II for use in routine clinical
diagnosis of cryptosporidiosis, and comparison with microscopic diagnostic methods. 45

2.1 Abstract.........................................................................................................................46
2.2 Introduction ..................................................................................................................47
2.3 Materials and Methods ..............................................................................................48
    2.3.1 Sample Selection ....................................................................................................48
    2.3.2 Microscopy ............................................................................................................48
    2.3.3 EntericBio GastroPanel II Molecular Testing .........................................................50
2.4 Results ........................................................................................................................51
    2.4.1 Real-time PCR Validation Study ............................................................................51
    2.4.2 Ova, Cyst and Parasite Microscopy Study .............................................................51
2.5 Discussion ....................................................................................................................52

Chapter III: A longitudinal, epidemiological study of Cryptosporidium species in
patients presenting with gastroenteritis in Southern Ireland ........................................56

3.1 Abstract ......................................................................................................................57
3.2 Introduction ..................................................................................................................58
3.3 Materials and Methods ..............................................................................................60
    3.3.1 Clinical Sample Acquisition ..................................................................................60
    3.3.2 Control Isolates .....................................................................................................60
    3.3.3 DNA Extraction .....................................................................................................61
    3.3.4 Primers and Probes ...............................................................................................61
    3.3.5 Molecular Analysis ...............................................................................................62
    3.3.6 Statistical Analysis ...............................................................................................63
3.4 Results ........................................................................................................................64
    3.4.1 18S rRNA Species Identification via Duplex Real-time PCR ...............................64
    3.4.2 Cryptosporidium Reference Unit EQA Participation .............................................65
    3.4.3 Epidemiological and Statistical Analyses ...............................................................65
3.5 Discussion ....................................................................................................................70
Chapter IV: Clinical and epidemiological significance of novel and established Cryptosporidium subtypes identified via gp60 sequencing among an Irish patient cohort

4.1 Abstract ..................................................................................................................79
4.2 Introduction .........................................................................................................80
4.3 Materials and Methods ..................................................................................83
  4.3.1 Clinical Sample Collection ........................................................................83
  4.3.2 Molecular Species Determination .................................................................83
  4.3.3 Subtyping Primer Design ..............................................................................83
  4.3.4 Real-time PCR Amplification and Sequencing of the gp60 Tandem Repeat Region ................................................................................................................84
  4.3.5 Statistical and Phylogenetic Analysis .............................................................85
4.4 Results ................................................................................................................86
  4.4.1 Epidemiological Trends ................................................................................86
  4.4.2 Allocation of Cryptosporidium to Genotypes and gp60-subtypes based on Phylogenetic Analysis .................................................................90
  4.4.3 Seasonal Distribution of gp60-subtypes .........................................................94
4.5 Discussion .........................................................................................................97

Chapter V: Development of a novel high resolution melting analysis-based genotyping method for Cryptosporidium parvum .................................................................101
5.1 Abstract ............................................................................................................102
5.2 Introduction .....................................................................................................103
5.3 Materials and Methods ................................................................................105
  5.3.1 Clinical Sample Acquisition ......................................................................105
  5.3.2 Genetic Marker Selection ............................................................................106
  5.3.3 Subtyping Primer Design ............................................................................107
  5.3.4 Nested Real-time PCR Amplification .........................................................109
  5.3.5 DNA Sequence Analysis of First Round Amplicons ................................110
Chapter VI: Development of a novel high resolution melting analysis-based genotyping method for Cryptosporidium hominis

6.1 Abstract ..................................................................................... 135
6.2 Introduction .................................................................................. 136
6.3 Materials and Methods ................................................................. 138
  6.3.1 Clinical Sample Acquisition ..................................................... 138
  6.3.2 Genetic Marker Selection ......................................................... 139
  6.3.3 Subtyping Primer Design ........................................................ 139
  6.3.4 Nested Real-time PCR Amplification ........................................ 141
  6.3.5 DNA Sequence Analysis of First Round Amplicons ................. 143
  6.3.6 HRM Curve Acquisition and Analysis ....................................... 143
  6.3.7 Evaluation of Assay Sensitivity and Specificity ......................... 143
  6.3.8 Reproducibility of HRM Analyses ............................................ 144
6.4 Results ......................................................................................... 145
  6.4.1 DNA Sequencing of Genetic Marker gp60-subtypes .................. 145
  6.4.2 Interpretation of HRM Curve Profiles and Categorisation of Categorisation of C. hominis gp60-subtypes ................................................. 149
6.4.3 Detection Sensitivity and Specificity of C. hominis Tandem Repeat Locus Targets ........................................................................................................157
6.4.4 Assessment of Intra- and Inter-assay Variabilities ............................ 158
6.5 Discussion ............................................................................................. 160

Chapter VII: Thesis Summary ................................................................. 164
7.1 Concluding Remarks ........................................................................... 165

Appendices ............................................................................................... 168
Appendix I – Chapter III Supplementary Data ........................................ 169
  A1.1 Real-time PCR Species Identification Results of Clinical Cryptosporidium DNA Samples ................................................................. 169
  A1.2 Amplification Curves and Electrophoresis Gels .............................. 174
Appendix II – Chapter V Supplementary Data ........................................... 181
  A2.1 Sensitivity Analysis .......................................................................... 181
  A2.2 Specificity Analysis ......................................................................... 188
Appendix III – Chapter VI Supplementary Data ........................................ 194
  A3.1 Sensitivity Analysis .......................................................................... 194
  A3.2 Specificity Analysis ......................................................................... 200
Appendix IV – Addendum to Chapter V .................................................... 206
  A4.1 Introduction ..................................................................................... 206
  A4.2 Materials and Methods .................................................................. 206
  A4.3 Results ............................................................................................ 207
  A4.4 Discussion ....................................................................................... 215
Appendix V – Addendum to Chapter VI .................................................... 217
  A5.1 Introduction ..................................................................................... 217
  A5.2 Materials and Methods .................................................................. 217
  A5.3 Results ............................................................................................ 218
  A5.4 Discussion ....................................................................................... 223

References .................................................................................................. 224
Acknowledgements

To my supervisors, Prof Roy D. Sleator and Dr Brigid Lucey, I wish to express my profound thanks for your unerring support, enthusiasm and expertise at every step of my research. Your boundless optimism and understanding were pivotal in the completion of my PhD.

I wish to acknowledge the staff of the Department of Medical Microbiology, Cork University Hospital, who provided a collaborative keystone over the course my PhD. I am particularly indebted to Liam Blake and to Dr Dan Corcoran. Thank you both for facilitating our collaborative research project and for your invaluable commentary and feedback on all of my manuscripts.

To Dr Rachel Chalmers and Dr Kristen Elwin of the Cryptosporidium Reference Unit, thank you for generously providing resources that proved invaluable to my research, and your expert feedback on a number of manuscripts.

I would also like to thank the staff of Serosep, particularly James O’ Leary and Dr Monika Lambert, who provided me with essential training and resources at the beginning of my PhD, and Rodney Govender, who contributed resources and advice at more advanced stages of my research.

To the lecturing and technical staff in the Department of Biological Sciences, MTU, thank you for your support and advice during my time as an undergraduate laboratory practical tutor.

I am also grateful for, and want to express special thanks to, the friends I have made among my postgraduate peers over the years. Thank you, Laura and Kate, for all of the supportive chats, stress venting opportunities and Fulda memories. I’m so happy that we coincidentally wound up in the same new job, Kate! Thank you, Colin, for the sage scientific advice and games nights! Thank you, Caoimhe, for your magical company, evening chip breaks, and for allowing me the great honour of cat-sitting Mary! Thank you too, Gosia, for all the delicious snacks, and support! I am thankful to have met so many interesting and lovely characters, past and present, during my time in CREATE.
To my parents, Angela and Michael, I don’t think I will ever adequately express how grateful I am for all you have done for me over the years, but I’ll do my best to continue to show you! Thank you also to my brother, Brian – I have finally finished school now! And to one of my closest friends, my little sister, Aisling – thank you so much for your conviction that I would actually get through my PhD, and for always supporting me and cheering me on. Special thanks to Ophelia for being a fluffy source of comfort and for putting my PhD-related troubles into perspective with feline indifference.

To Karl, my very best friend, thank you for your unwavering patience, support and understanding over the past several years. Thank you for taking care of me when I was distracted, low and stressed, and for all of the countless other things you have done to prop me up during my PhD. I don’t think I am capable of articulating all of the ways you have helped me, and I will never be able to express my gratitude to you or repay your kindness enough. I am so excited for post-PhD life with you!
List of Posters, Presentations, Publications & Scientific Communications

Oral Presentations


Poster Presentations


Publications


**Scientific Communications**

### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIDS</td>
<td>Acquired Immune Deficiency Syndrome</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>ATCC</td>
<td>American Type Culture Collection</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CIR</td>
<td>crude incidence rate</td>
</tr>
<tr>
<td>cp</td>
<td>crossing point</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered Regularly Interspaced Palindromic Repeats</td>
</tr>
<tr>
<td>CRU</td>
<td>Cryptosporidium Reference Unit</td>
</tr>
<tr>
<td>CUH</td>
<td>Cork University Hospital</td>
</tr>
<tr>
<td>DALY</td>
<td>disability-adjusted life year</td>
</tr>
<tr>
<td>ddPCR</td>
<td>digital drop polymerase chain reaction</td>
</tr>
<tr>
<td>df</td>
<td>degrees of freedom</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>ECDC</td>
<td>European Centre for Disease Prevention and Control</td>
</tr>
<tr>
<td>EEC</td>
<td>European Economic Community</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FDA</td>
<td>US Food and Drug Administration</td>
</tr>
<tr>
<td>g</td>
<td>Grams</td>
</tr>
<tr>
<td>gp60</td>
<td>60 kDa glycoprotein</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrogen chloride</td>
</tr>
<tr>
<td>HPSC</td>
<td>Health Protection Surveillance Centre</td>
</tr>
<tr>
<td>HRM</td>
<td>high resolution melting</td>
</tr>
<tr>
<td>HSE</td>
<td>Health Service Executive</td>
</tr>
<tr>
<td>HTS</td>
<td>high-throughput screening</td>
</tr>
<tr>
<td>ICZN</td>
<td>International Code of Zoonotic Nomenclature</td>
</tr>
<tr>
<td>IEC</td>
<td>intestinal epithelial cell</td>
</tr>
<tr>
<td>LOD</td>
<td>limit of detection</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ml</td>
<td>millilitre</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>MLG</td>
<td>multi-locus genotyping</td>
</tr>
<tr>
<td>MLST</td>
<td>multi-locus sequence typing</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>MMV</td>
<td>Medicines for Malaria Venture</td>
</tr>
<tr>
<td>mZN</td>
<td>modified Zhiel-Neelsen</td>
</tr>
<tr>
<td>nM</td>
<td>nanomolar</td>
</tr>
<tr>
<td>oocysts/ml</td>
<td>oocysts per millilitre</td>
</tr>
<tr>
<td>p</td>
<td>p-value</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>dPCR</td>
<td>digital polymerase chain reaction</td>
</tr>
<tr>
<td>qPCR</td>
<td>qualitative polymerase chain reaction</td>
</tr>
<tr>
<td>RFLP</td>
<td>restriction fragment length polymorphism</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcriptase polymerase chain reaction</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>TR</td>
<td>tandem repeat</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>VNTR</td>
<td>variable number tandem repeats</td>
</tr>
<tr>
<td>VTEC</td>
<td>Verotoxigenic <em>Escherichia coli</em></td>
</tr>
<tr>
<td>WGS</td>
<td>whole genome sequencing</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>μL</td>
<td>microlitre</td>
</tr>
<tr>
<td>μm</td>
<td>micrometre</td>
</tr>
<tr>
<td>μM</td>
<td>micromolar</td>
</tr>
<tr>
<td>v/v</td>
<td>volume per volume</td>
</tr>
<tr>
<td>w/v</td>
<td>weight per volume</td>
</tr>
<tr>
<td>%</td>
<td>per cent</td>
</tr>
</tbody>
</table>
Thesis Abstract

Novel Molecular Approaches to Diagnostic and Epidemiological Analyses of Cryptosporidium spp. Implicated in Clinical Infection – Jennifer O’ Leary

Cryptosporidium remains an underestimated, yet highly prevalent, parasitic agent of gastrointestinal illness worldwide. In clinical settings, diagnosis has previously relied on microscopic techniques. However, enteric parasite targeting molecular tools, such as real-time PCR panels, have superseded microscopy in recent years. This body of work examines the nature, implications and opportunities of this paradigm shift.

This thesis is divided into seven chapters that consist of a literature review chapter, five experimental chapters focusing on the molecular diagnosis and epidemiological analysis of Cryptosporidium spp., and a final summary chapter.

Chapter I provides the contextual narrative and supportive reasoning behind the research undertaken in this thesis through discussion of the published literature.

A clinical validation study, introducing a real-time PCR based method for the clinical diagnosis of cryptosporidiosis to the Medical Microbiology Department of Cork University Hospital, and an examination of the impact of reducing routine enteric microscopy on detection rates of other enteric parasites provided the initial impetus behind this research. These studies are outlined and discussed in Chapter II.

Chapter III details the longitudinal epidemiological study that was conducted in collaboration with Cork University Hospital on a bank of clinical Cryptosporidium samples amassed through routine clinical testing. Routine diagnostic testing in this medical microbiology laboratory was limited to genus level identification. This study identified the species implicated in clinical infection among the amassed isolates through fluorescent probed-based real-time PCR analysis of the 18S rRNA gene.

Further epidemiological and phylogenetic analyses were conducted on the amassed clinical Cryptosporidium sample bank to identify isolates to the subspecies/subtype level. This was achieved thorough sequencing of the 60-kDa glycoprotein gene. This analysis resulted in the discovery of Cryptosporidium parvum and Cryptosporidium
hominis gp60-subtypes hitherto not detected in Ireland. This research is detailed in Chapter IV.

The diversification of the Cryptosporidium population in Ireland and the general heterogeneity of Cryptosporidium populations reported globally provided the basis for the research conducted in Chapter V. This chapter describes the development of a novel methodology to differentiate between C. parvum gp60-subtypes, without the need for DNA sequencing, through the real-time PCR application, high resolution melting analysis.

Chapter VI describes the application of an analogous subtyping method to that developed in the preceding chapter to differentiate between C. hominis gp60-subtypes.

Chapter VII provides a summary of the research described herein and also details future avenues of research for the advancement of the developed HRM analysis-based Cryptosporidium spp. subtyping methods.
Chapter I: Literature review - Cryptosporidium epidemiology, diagnosis, and research in the 21st century

A manuscript based upon this chapter has been published in Food and Waterborne Parasitology.

1.1 Abstract

The protozoan parasite Cryptosporidium has emerged as a leading cause of diarrhoeal illness worldwide, posing a significant threat to young children and immunocompromised patients. While endemic in the vast majority of developing countries, Cryptosporidium also has the potential to cause waterborne epidemics and large scale outbreaks in both developing and developed nations. Anthroponontic and zoonotic transmission routes are well defined, with the ingestion of faecally contaminated food and water supplies a common source of infection. Microscopy, the current diagnostic mainstay, is considered by many to be suboptimal. This has prompted a shift towards alternative diagnostic techniques in the advent of the molecular era. Molecular methods, particularly PCR, are gaining traction in a diagnostic capacity over microscopy in the diagnosis of cryptosporidiosis, given the laborious and often tedious nature of the latter. Until now, developments in the field of Cryptosporidium detection and research have been somewhat hampered by the intractable nature of this parasite. However, recent advances in the field have taken the tentative first steps towards bringing Cryptosporidium research into the 21st century. A review of these advances is detailed herein.
1.2 Introduction

*Cryptosporidium* is an obligate enteric protozoan parasite and a well-established pathogen of the gastrointestinal tract (Tzipori and Widmer, 2008). Originally described from histological preparations of murine gastric mucosa in 1907 (Tyzzer, 1910), *Cryptosporidium* was not linked with human infection until 1976 (Nime et al., 1976). Sentinel alert of the clinical significance of cryptosporidial infection was established after reports of fatal cryptosporidiosis in AIDS patients during the 1980s (Soave et al., 1984), and a large waterborne outbreak affecting approximately 400,000 Milwaukee residents in 1993 (D’Antonio et al., 1985; MacKenzie et al., 1995; Soave et al., 1984). *Cryptosporidium* is environmentally and geographically ubiquitous, comprising many species spanning a wide host range. To date over 40 species have been described, with at least 20 species having been reported in human infection (Feng et al., 2018; Xiao and Feng, 2017). However, *C. hominis* and *C. parvum* have been reported to account for the vast majority of infections in humans (Feng et al., 2018).

*Cryptosporidium* has now emerged as a leading cause of diarrhoeal illness worldwide, posing a significant threat to young children and immunocompromised patients. It has been reported to be a leading cause of moderate-to-severe gastrointestinal morbidity in children younger than 5 years in developing countries (Kotloff et al., 2013). A recent study into the global burden of gastrointestinal disease found that *Cryptosporidium* spp. accounted for in excess of 1 million deaths, almost half a million of which were in children under the age of five, and over 71 million disability-adjusted life years (DALYs) between 2005 and 2015; the highest mortality rates were observed in developing countries, particularly those in sub-Saharan Africa (Troeger et al., 2017).

While endemic in the vast majority of developing countries, *Cryptosporidium* also has the potential to cause waterborne epidemics and large scale outbreaks in both developing and developed nations (Efstratiou et al., 2017). Anthroponontic and zoonotic transmission routes are well defined, with the ingestion of faecally contaminated food and water supplies a common source of infection (King et al., 2019; McKerr et al., 2019; Ryan et al., 2018).

Microscopy, the current diagnostic mainstay, is considered by many to be suboptimal. This has prompted a shift towards alternative diagnostic techniques in the advent of the molecular era. Molecular methods, particularly PCR, are gaining traction in a diagnostic
capacity over microscopy in the detection of Cryptosporidium spp., given the laborious and tedious nature of the latter. The supersession of microscopy by molecular techniques, which offer rapid diagnosis and improved sensitivity, may also be attributed to diminishing microscopy skills in modern clinical laboratories (Meurs et al., 2017). Until now, developments in the field of Cryptosporidium detection and research have been somewhat hampered by the intractable nature of this parasite. However, recent advances, particularly in the application of the CRISPR/Cas system to produce genetically modified, tractable Cryptosporidium oocysts (Vinayak et al., 2015), have allowed the field to take tentative steps towards bringing Cryptosporidium research into the 21st century.

1.3 Taxonomy

Taxonomically, Cryptosporidium spp. are assigned to family Cryptosporidiidae, order Cryptogregida, subclass Cryptogregarinaria, class Gregarinomorpha, phylum Apicomplexa (Cavalier-Smith, 2014; Tzipori and Ward, 2002). The taxonomy of Cryptosporidium has been a source of debate for several decades due to biological, reproductive and phylogenetic characteristics atypical to those seen within the traditional Coccidian class, to which it was previously assigned (Cavalier-Smith, 2014; Chalmers and Katzer, 2013; Cavalier-Smith, 2014; Tzipori and Ward, 2002).

It has long been speculated that Cryptosporidium may represent a ‘missing link’ between more primitive gregarine parasites and coccidians, with phylogenetic analysis revealing that Cryptosporidium were likely to have undergone early evolutionary divergence from other apicomplexan taxa, to form a separate clade with the gregarines, a monophyletic apicomplexan group, (Carreno et al., 1999). This theory was also extensively supported by decades-worth of biochemical, microscopic, molecular and studies (Aldeyarbi and Karanis, 2016a, 2016b; Barta and Thompson, 2006; Borowski et al., 2008; Bull et al., 1998; Gibbons et al., 2014; Hijjawi, 2010; Hijjawi et al., 2010a, 2004, 2002; Karanis et al., 2008; Karanis and Aldeyarbi, 2011; Pohlenz et al., 1978; Rosales et al., 2005; Templeton et al., 2010; Thompson et al., 2016; Valigurová et al., 2015).

A formal re-evaluation of Cryptosporidium taxonomic designations occurred in 2014, based on phylogenetic analyses that lead to a revision of gregarine classification (Cavalier-Smith, 2014). Cryptosporidium was transferred from the subclass Coccidia,
class Coccidiomorphea to the newly devised subclass Cryptogregaria, class Gregarinomorphea. The genus *Cryptosporidium* is currently the sole member of Cryptogregaria, which is described as comprising epicellular parasites of vertebrates possessing a gregarine-like feeder organelle but lacking an apicoplast (Cavalier-Smith, 2014). The re-classification of *Cryptosporidium* as a member of Cryptogregaria currently remains unchallenged. According to the International Code of Zoological Nomenclature (ICZN), once a species has been formally re-classified in a peer-review publically available journal, the re-classification stands in the absence of subsequent challenges in the literature (Ryan *et al.*, 2016).

Taxonomic classification of *Cryptosporidium* within the phylum Apicomplexa has been largely based upon microscopic identification of morphological features within endogenous and exogenous stages, location of endogenous stages, and host specificity (Morgan *et al.*, 1999). These phenotypically based criteria were employed for *Cryptosporidium* classification, and were subsequently found to be inadequate (Fayer, 2010). Although oocyst morphology has proven particularly useful for genus and species designation in other Apicomplexan classification schemes, microscopic study of *Cryptosporidium* morphology is somewhat ambiguous, due to *Cryptosporidium* oocysts being some of the smallest within the Apicomplexa phylum (Fayer *et al.*, 2000).

The implementation of single host specificity as a speciation criterion for *Cryptosporidium* also proved to be fundamentally flawed. Several species identified on this basis, such as *C. agni* and *C. ansernum*, which were thought to be specific to sheep and geese, respectively, were subsequently invalidated by cross-transmission studies that demonstrated the ability of *Cryptosporidium* spp to infect multiple host species beyond those within which they were initially identified (Barker and Carbonell, 1974; Proctor and Kemp, 1974). However, cross-transmission studies led to the erroneous classification of several biologically distinct mammalian species as *C. parvum*, which was thought to be species specific within mammals at the time (Fayer *et al.*, 2000).

The taxonomic conundrum of differentiating *Cryptosporidium* spp. has been further exacerbated by the fact that no standardised concept or definition of what constitutes such a species has been proposed. Thus, no specific criteria have been established in order to define, identify or characterise *Cryptosporidium* spp., though several recommendations have been published. A polyphasic approach incorporating
morphological, developmental biology and genetic evidence has been proposed by Egyed and colleagues (Egyed et al., 2003). A similar, though arguably more comprehensive series of potential criteria have been put forward by Xiao et al. (Xiao et al., 2004), who consider oocyst morphometrics, genetic characterisation, identification of naturally and experimentally infected hosts, and compliance with the ICZN to be of utmost importance when defining Cryptosporidium spp. These criteria were recently employed in a comprehensive review of valid and non-valid Cryptosporidium spp. (Ryan et al., 2014).

Cryptosporidium taxonomic discoveries are ongoing, with a detailed 2019 study demonstrating the existence of two C. parvum sub-species with distinct host-associations. The two discrete lineages were identified as the previously cryptic Cryptosporidium parvum anthroponsum (C. p. anthroponsum), a more recently emerged, phylogenetically distinct, anthroponotic sub-species; and Cryptosporidium parvum parvum (C. p. parvum), the extant, zoonotic sub-species (Nader et al., 2019). Analyses conducted on 21 whole genome sequences and 467 60kDa glycoprotein gene (gp60) sequences identified these lineages to be distinct to the nucleotide level and their population genetic structures to differ markedly. It is likely that whole genomic sequencing and phylogenetic analyses such as these may stand to dramatically advance the depth of understanding pertaining to Cryptosporidium species and sub-species in the coming years.

1.4 Epidemiology

1.4.1 A Global Perspective

Cryptosporidium is a globally ubiquitous pathogen, with documented cases of human infection reported in over 90 countries across all continents aside from Antarctica (Ryan et al., 2014). Current molecular characterisation studies lack standardisation, resulting in a fragmented picture of global distribution (Chalmers et al., 2017). The burden of disease from cryptosporidiosis varies substantially between and within countries and geographical areas. The prevalence of Cryptosporidium in humans is reported to be 2.6–21.3% in African countries, 3.2–31.5% in Central and South American countries, 1.3–13.1% in Asia countries, 0.1–14.1% in Europe, and 0.3–4.3% in North America (Fayer, 2004). In developing countries, 8–19% of diarrheal diseases are attributed to
*Cryptosporidium*, and cryptosporidiosis is reported to account for 20% of all cases of diarrhoea in children (Gatei *et al.*, 2006; Ryan *et al.*, 2014). In developed countries, *Cryptosporidium* is less common and accounts for about 9% of diarrheal episodes in children (Ryan *et al.*, 2016).

Given the diversity exhibited by *Cryptosporidium* species, geographical and socio-economic variables impact significantly on species distribution (Feng *et al.*, 2018). For example, the predominantly anthroponotic *C. hominis* accounts for the vast majority of human infection in industrialised countries such as North and South America, Australia, Japan, Africa and much of the developing world (Ajjampur *et al.*, 2007; Braima *et al.*, 2019; Cacciò *et al.*, 2005; Gatei *et al.*, 2003; Hira *et al.*, 2011; Ng-Hublin *et al.*, 2017; Pelayo *et al.*, 2008; Sharma *et al.*, 2013; Tiangtip and Jongwutiwes, 2002; Yagita *et al.*, 2001). However, *C. parvum* predominates in rural areas of New Zealand, much of Europe and the Middle-East (Areeshi *et al.*, 2007; Iqbal *et al.*, 2011; Learmonth *et al.*, 2001; Meamar *et al.*, 2007).

Geographic variation in species distribution also varies within countries, such as New Zealand and Ireland, where there exists a marked disparity between *C. parvum* and *C. hominis* infection rates (Chalmers *et al.*, 2011b; Learmonth *et al.*, 2004; Zintl *et al.*, 2011). *C. parvum* generally predominates in rural areas, due to widespread livestock farming. *C. hominis* on the other hand, being predominantly specific to humans, with few rare exceptions, is generally associated with spread in densely populated, urban areas (Akiyoshi *et al.*, 2002; Ebeid *et al.*, 2003; Lake *et al.*, 2007).

Temporal variation also occurs; both spring and autumn infection peaks occur in the UK, Ireland and Oceanic countries, such as New Zealand (Chalmers *et al.*, 2019; Learmonth *et al.*, 2004; Zintl *et al.*, 2011), while a single autumnal peak is observed in the United States (Lal *et al.*, 2012). Spring peaks generally correlate with the autochthonous environmental spread of *C. parvum* oocysts occurring as a result of lambing and calving seasons (McLauchlin *et al.*, 2000; Zintl *et al.*, 2009). *C. hominis* exhibits similar variation such that infection rates often peak in urban areas during the late summer and early autumn. This can be attributed, as least in part, to anthroponotic transmission which is largely associated with travellers and recreational swimming pool facilities (Chalmers *et al.*, 2009; Learmonth *et al.*, 2001; Snel *et al.*, 2009). Variations also occur in tropical climates, such as those in North-Eastern Brazil and West Africa,
with the infectious burden of Cryptosporidiosis peaking during the warm, rainy season (Newman et al., 1999; Perch et al., 2001).

One review of the literature, pertaining to all reported water-borne protozoan outbreaks on a global scale (325 in total) and spanning almost a hundred years, determined 23.7% (77) of cases to be attributable to Cryptosporidium contamination (Karanis et al., 2007). The vast majority of cases were reported in the United States and Europe, particularly the United Kingdom, with a very limited number of cases having been reported in Japan, Australia, New Zealand and other nations (Karanis et al., 2007). A similar study examined all published literature on outbreaks between 2004 and 2010 and found that the majority of Cryptosporidium outbreaks were reported in North America (30.6%) and Europe (16.5%), with only 3.5% attributable to Asian cases (0.5% of cases occurred in Japan, China, India, Malaysia, with 1.5% occurring in Turkey) (Baldursson and Karanis, 2011). Markedly disproportionate asymptomatic carriage rates have also been reported between developed and developing countries, ranging from <1% and 10-30%, respectively (Baldursson and Karanis, 2011).

Reported prevalence and carriage rates indicate that the majority of outbreaks are likely to occur in developed countries, but this may not necessarily be the case. Such a lack of correlation between prevalence and outbreak rates is likely due to the advanced logistical, diagnostic, and reporting systems available in developed countries that are often not available in resource-limited countries (Karanis et al., 2007). In light of this, Cryptosporidium has been included in the World Health Organisation’s Neglected Disease Initiative since 2004 (Savioli et al., 2006). This initiative is aimed at combating the global burden exerted by a heterogeneous group of parasitic, bacterial and viral diseases, particularly in developing countries where climate, poverty and lack of access to health services impacts upon infection outcomes.

1.4.2 Epidemiology within Europe

Though not currently notifiable in all European countries, the European Surveillance System database has collated data since 2005, with EU member states submitting data pertaining to Cryptosporidium outbreaks in a process governed by the European Commission Council Directive 2008/246/EC. As of the most recent ECDC report, only 21 of the 28 member nations of the EU/EEA were reporting data on cryptosporidiosis,
suggesting underreporting in a number of European countries where formal report frameworks are not enforced (European Centre for Disease Prevention and Control, 2014). However, Ireland, the United Kingdom, Norway and Sweden report the highest incidence rates per 100,000 population, with incidence highest among children under the age of four (European Centre for Disease Prevention and Control, 2019, 2014, 2013, 2012).

In a similar trend to non-EU countries, the 2017 ECDC epidemiological report highlighted a distinct bimodal seasonal pattern featuring a pronounced autumnal peak (European Centre for Disease Prevention and Control, 2019). The autumnal peak is again thought to be caused by exposure to contaminated water during various recreational water sport activities favoured during the summer holiday season (Chalmers et al., 2019). Reports of several outbreaks among British tourists upon return from Majorcan holiday resorts, and a study of cryptosporidiosis in German travellers upon return from Asia are suggestive of this being the case (European Centre for Disease Prevention and Control, 2014; Galmes et al., 2003; Smearon, 2000; Weitzel et al., 2006b)

Foodborne and waterborne outbreaks are often reported within the EU. A large outbreak of cryptosporidiosis due to contamination of the source of reticulated drinking water was reported to have occurred in Östersund, Sweden during 2010 and 2011 (European Centre for Disease Prevention and Control, 2012). Further outbreaks emanating from contaminated water supplies have occurred in England, Wales, Northern Ireland and the Republic of Ireland (Carnicer-Pont et al., 2011; Jennings and Rhatigan, 2002; Smyth, 2001; Stanwell-Smith, 1997). Several foodborne outbreaks have been described in recent times in Finland, Sweden and the UK. Vegetables, particularly salads, are reputed to be an important source in foodborne Cryptosporidium outbreaks, and were directly implicated or strongly suspected as the source of contamination in these outbreaks (Insulander et al., 2008; McKerr et al., 2015; Pönka et al., 2009).

Epidemiological information, particularly pan-European based information, on species identification and subtyping of Cryptosporidium isolates is limited. However, studies conducted among European nations such as the United Kingdom, France, Switzerland and Portugal, Slovenia and Poland have reported C. parvum and C. hominis to be the predominant species, with C. meleagridis and C. felis contributing to infection rates to a
much lesser extent, in the case where identification was performed to species level. (Alves et al., 2003; Areeshi et al., 2007; Bajer et al., 2008; Chalmers et al., 2009; Fretz et al., 2003; Guyot et al., 2001; Iqbal et al., 2011; Learmonth et al., 2001; Meamar et al., 2007; Ong et al., 2002).

Current subtyping methodologies are also inadequate to describe the level of genetic variation observed within and between Cryptosporidium populations. The current mainstay in Cryptosporidium spp. subtype characterisation, the 60kDa glycoprotein (gp60) gene based nomenclature system, was first described by Sulaiman et al. (2005) and further refined by Jex et al. (2007), based on research conducted by Strong et al., (2000), (Jex et al., 2007; Strong et al., 2000; Sulaiman et al., 2005). Within this nomenclatural system, species are assigned a species specific prefix (‘I’ denoting C. hominis, ‘II’ denoting C. parvum and ‘III’ denoting C. meleagridis etc.) and a gp60-subtype family designation (Ia, Ib, Id, Ie etc. for C. hominis; IIa, IIb, IIc, IId etc. for C. parvum, IIIa, IIIb, IIIc, IIId etc. for C. meleagridis). gp60-subtypes are further differentiated on the basis of the number of trinucleotide repeat sequences (TCA, TCG and TCT), which encode the poly-serine tract of the gp40 glycoprotein, within the gp60 microsatellite region (Sulaiman et al., 2005; Jex et al., 2007).

However, owing to the occurrence of genetic recombination between genetically dissimilar haplotypes during the sexual phase of the Cryptosporidium life-cycle, reliable and discriminatory subtyping methods require the analysis of multiple loci (Chalmers et al., 2018). Consequently, current Cryptosporidium subtyping analyses utilise much broader multi-locus panels by which to characterise intra-species variation and conduct detailed population studies (Beser et al., 2017; Chalmers et al., 2017; De Waele et al., 2013; Feng et al., 2014; Robinson and Chalmers, 2012; Wang et al., 2015). Although called for, there has been no international adoption of a standardised multi-locus genotyping (MLG) scheme as of yet (Chalmers et al., 2017).

1.5 Life Cycle

Cryptosporidium spp. exhibit a monoxenous life-cycle with oocyst ingestion propagating infection within a single host (O’ Donoghue, 1995). The small, circular oocysts of C. parvum, C. hominis and C. meleagridis measure approximately 3 to 6 μm in diameter (Bones et al., 2019). Sporozoites exhibit gliding motility and optimal in vitro attachment and invasion are induced at 37°C and a pH of 7.4 to 7.6 (Huang et al.,
Exposure to acidic pH (pH ~ 2), followed by incubation in bile salts, reducing agents and proteases mimics transit through the acidic stomach to the alkaline small intestine and enhances excystation, the process wherein the ‘suture’ of the robust oocyst wall rupture, in vitro; however, the fact that oocysts can excyst in vivo and cause disease in extraintestinal locations indicates that some of these host derived triggers are not essential (Smith et al., 2005). Excystation results in the release of four naked sporozoites, measuring $2 \times 0.8 \mu$m, into the small intestine, as depicted in Figure 1.1 (Fayer and Leek, 1984; King et al., 2012; Smith et al., 2005).

Following attachment to the apical side of the enterocyte, the sporozoite is enveloped by the host epithelial cell. *Cryptosporidium* is capable of inducing the rearrangement of cytoskeleton molecules at the invaded apical pole, particularly ezrin, villin and smaller amounts of F-actin, to produce an intracellular, yet extracytoplasmic, niche between the membrane and the cytoplasm of the cell, surrounding the luminal and lateral sides of the parasite - a parasitophorous vacuole (Bonnin et al., 1999). This parasitophorous vacuole remains in the brush border area, disrupting the distribution of microvilli (Smith et al., 2005). Once internalised, the sporozoite undergoes a transformative process to form a replicative trophozoite, requiring both host and parasite derived factors to do so. The formation of a ‘feeder organelle’, termed the ‘epimerite’, a tube that connects the parasite to the host cell cytoplasm, also occurs at this point (Beyer et al., 2000; Bonnin et al., 1999; Huang et al., 2004). This process has been hypothesised to be a mechanism of myzocytosis, as observed in other gregarines. The parasite remains sequestered within this vacuole, accessing host cell nutrients via a feeder organelle membrane, while remaining protected from the host immune response and the hostile environment of the gut lumen (Tzipori and Ward, 2002).

From this intracellular vacuole, the trophozoite is primed to undergo both asexual and sexual reproduction (detailed in Figure 1.1). Endopolygeny leads to the formation of daughter cells within the mother meront through internal budding. Two such cycles occur, with these cycles yielding eight Type I meronts and four Type II meronts. Type I meronts, containing six to eight type I merozoites, are released upon maturation and go on to infect surrounding epithelial cells through the release of these Type I merozoites, and also undergo another round of merogeny to form type II meronts. Type II meronts undergo cell division to produce four type II merozoites (O’ Donoghue, 1995). These merozoites are responsible for sexual reproduction as they give rise to both micro-and
macrogamonts. The fusion of the non-flagellated microgamont and the uninucleate macrogamont leads to a fertilised diploid zygote (O’Hara and Chen, 2012). A subsequent process, not dissimilar to meiosis, produces four haploid sporozoites within an oocyst. Oocysts may be thick walled, which are egested in faeces, or thin-walled (O’Donoghue, 1995). Thin walled oocysts are not egested, but remain within the host and may sporulate to cause autoinfection and persistent diarrhoea, a mechanism initially observed by Tyzzer (Tyzzer, 1910).
Figure 1.1  Life Cycle of Cryptosporidium - Following oocyst ingestion by a host, and excystation, the sporozoites are released and parasitize epithelial cells of the gastrointestinal tract. In these cells, the parasites undergo asexual multiplication (schizogony or merogony) and then sexual multiplication (gametogony), producing microgamonts and macrogamonts. Upon fertilization of the macrogamonts by the microgametes, oocysts develop and sporulate in the infected host. Two different types of oocysts are produced: the thick-walled, which is commonly excreted by the host, after sporogony, and the thin-walled oocyst, which is primarily involved in autoinfection.
1.6 Pathophysiology

The severity and duration of clinical symptoms are generally dependent upon parasite infectivity, intestinal distribution, and the host immune response (Bouzid et al., 2013). Different Cryptosporidium spp. can also cause disease of varying severity in human hosts. For example, C. hominis appears to cause more severe acute disease, more extra-intestinal disease, and have more recurrent manifestations after an infection than do other Cryptosporidium spp. (Leitch and He, 2011). Extensive study on the infectivity of Cryptosporidium spp. in humans is lacking, although a small number of studies have been published. Although variable, the infectious dose (ID₅₀) has been ascertained to be relatively low, estimated at 10 to 83 C. hominis oocysts, and 300 to 375 C. parvum oocysts, based on studies examining both clinical and microbiological definitions of infection, respectively (Chappell et al., 2006; Okhuysen et al., 2002). ID₅₀ also varies between unexposed and previously exposed individuals. One study found a 23–fold difference to exist between the clinically derived ID₅₀ of serologically negative individuals and anti-C. parvum IgG producing individuals. A 57–fold difference existed between the microbiologically derived ID₅₀ of antibody negative and antibody positive volunteers (Chappell et al., 1999). Interestingly, infectious dose has also been found to vary between C. parvum isolates (Okhuysen et al., 1999).

While the precise mechanisms have not yet been elucidated, it has been proposed that diarrhoeic symptoms are a consequence of impaired intestinal absorption and increased secretion (Farthing, 2000). Infection of the proximal small intestine has been shown to produce increased levels of diarrhoea and is associated with higher mortality rates, while infection limited to the large intestine generally produces intermittent bouts of diarrhoea or is asymptomatic (Clayton et al., 1994; Phillips et al., 1992).

Epithelial invasion is minimal, with zoite forms confined to the superficial layer of cells (Guerrant, 1997). However, invasion induces considerable alterations in the architecture of the luminal surface. Villous atrophy and crypt hyperplasia, often accompanied by inflammatory infiltration of the underlying lamina propria by lymphocytes, macrophages and neutrophils, are common histological findings (Farthing, 2000). Morphological deterioration has been found to be dependent upon the number of infective organisms (Laurent et al., 1999).
Villous atrophy encumbers Na$^+$ absorption while crypt cell hyperplasia causes increased Cl$^-$ secretion (Argenzio et al., 1990). Elevated production of prostaglandins, known regulators of intestinal fluid secretion, has been observed in Cryptosporidium-infected human intestinal epithelial cells (Laurent et al., 1998). Similarly elevations in prostaglandin production have been reported in a porcine model, resulting in reduced NaCl absorption and increased secretion. Reversal of secretion upon administration of cyclo-oxygenase inhibitor, indomethacin, indicates that this is a plausible mechanism that may potentially underpin the pathophysiology of cryptosporidiosis (Farthing, 2000).

It has also been proposed that pathological manifestations of Cryptosporidium infection may be the result of T-cell mediated immune damage, similar to that seen in autoimmune diseases such as Crohn’s disease and ulcerative colitis, to some extent (Chalmers and Davies, 2010). Localisation of inflammatory cells within the lamina propria may cause cytokine and neuropeptide mediated damage as a result of increased intercellular permeability (Chalmers and Davies, 2010). In addition, damage to the microvillus border also diminishes the absorptive capacity of the intestine, whilst also limiting membrane-bound digestive enzyme activities (Adams et al., 1994; Griffiths et al., 1994). Cytoskeletal modifications mediated by Cryptosporidium also result in transient alterations in cellular architecture and impact on overall intestinal barrier function, as tight junctions are also compromised (Buret et al., 2003; O’Donoghue, 1995).

1.7 Clinical Features and Patient Spectrum

Clinical manifestation of symptomatic cryptosporidiosis in immunocompetent individuals results in the sudden onset of large volumes of watery diarrhoea, usually accompanied by fever, malaise, nausea and vomiting, pyrexia and anorexia. With an incubation period of typically 2 to 10 days, infection is largely self-limiting, usually resolving spontaneously within one to two weeks of symptom onset in the majority of cases, owing to innate and adaptive immune responses within the host response (Ludington and Ward, 2015). However, diarrhoea may persist for up to two months in some individuals, although this is rare among the immunocompetent (Chalmers and Davies, 2010). The highest burden of infection within an otherwise healthy population is seen in children under the age of 5, the majority of whom are < 2 years of age (Stark et al., 2009). This may potentially be attributed to the development of immunity in older
children through antibody generation after resolution of previous infections (Priest et al., 2006).

*Cryptosporidium* is a leading cause of prolonged (1 to 7 days) and persistent (> 14 days) bouts of diarrhoea in children in developing countries, (Lima et al., 2000; Snelling et al., 2007). A recent Global Enteric Multicenter study found *Cryptosporidium* to be the second most prevalent cause of diarrheal induced morbidity and mortality in children < 5 years, surpassed only by rotavirus (Kotloff et al., 2013). Suboptimal nutritional status and poor sanitation practices, both of which are considerable problems in developing countries, impact significantly on patient outcome and mortality (Moore, 2011). In addition, several studies have reported a distinct correlation between multiple episodes of persistent diarrhoea during infancy and malnutrition, with subsequent long-term growth impairment, and physical and cognitive deficits (Agnew et al., 1998; Berkman et al., 2002; Guerrant et al., 1999; Moore et al., 2010). A number of robust studies have highlighted how asymptomatic infection in children also exerts detrimental effects on cognitive and physical development, which given the greater prevalence of such infections in endemic areas, is a considerable cause for concern and should not be underestimated (Berkman et al., 2002; Checkley et al., 1997).

The impact of *Cryptosporidium* infection on HIV/AIDS infected patients is well documented, as is the relationship between increasing symptoms severity and extent of immuno-compromisation (Wang et al., 2018). Transient, self-limiting disease is generally associated with CD4+ counts in excess of 180/mm³, while fulminant diarrhoea, biliary infection and death are generally confined to patients exhibiting complete immune collapse, with CD4+ counts of less than 50/mm³ (Blanshard et al., 1992; Flanigan et al., 1992).

Patients, particularly children, with congenital immunodeficiencies, such as severe combined immunodeficiency syndrome (SCID), X-linked hyper IgM syndrome, interferon-γ deficiency, and primary CD4 lymphopenia are also susceptible to *Cryptosporidium* infection (Conley et al., 1999; Gomez Morales et al., 1996; Kocoshis et al., 1984; Rahman et al., 2012; Wolska-Kusnierz et al., 2007). Additionally, patients with haematological malignancies, and immunosuppressed haemodialysis patients are also at risk from severe, persistent, *Cryptosporidium* induced diarrhoea (Helmy et al., 2014; Hong et al., 2007; Legrand et al., 2011; Seyrafian et al., 2006; Tandon and...
Gupta, 2014; Trad et al., 2003; Turkcapar et al., 2002). Disseminated infection in immunocompromised patients is known to occur, although cases in immunocompetent individuals appear not to have been reported to date (Kutukculer et al., 2003).

1.8 Diagnosis

Given the broad spectrum of susceptibility and the significant morbidity and mortality rates associated with Cryptosporidium in immunosuppressed patient populations, the development of efficient and effective screening criteria and robust testing algorithms is vital in clinical laboratories (Bruijnesteijn van Coppenraet et al., 2009; Garcia et al., 2003; Hawash, 2014). However, there remains no international Standard Methods for the diagnosis of cryptosporidiosis. In some countries Cryptosporidium testing is limited to known HIV/AIDS patients, however, testing of adult samples is usually reliant upon stipulating factors such as watery or persistent diarrhoea, and when clinically suspected (Chalmers, 2008). Current UK microbiological standards recommend that all symptomatic cases of acute diarrhoea are investigated based on guidelines for faecal screening for Cryptosporidium, published in 1993 (Casemore and Roberts, 1993). These guidelines are the culmination of a two year prospective survey conducted in 16 clinical laboratories on some 62,000 patients, which provided data pertaining to the absolute and relative frequency of Cryptosporidium infection across all age groups. This allowed for meaningful age-based selection criteria to be determined (Casemore and Roberts, 1993; Palmer and Biffin, 1990; Public Health England, 2014). Compliance with these guidelines is not absolute. Several UK-based studies determined that routine screening of all diagnostic stool specimens for Cryptosporidium in participating clinical laboratories ranged from 33 to 72.5% (Chalmers and Thomas, 2002; Chalmers et al., 2015). The more recent of these studies found that 27.5% of surveyed laboratories conducted for Cryptosporidium spp. testing on samples based on one or more selection criteria, including stool consistency (19% [16 of 85]), patient age (21% [18 of 85]), history or clinical details (47% [40 of 85]), duration of hospitalization (21% [18 of 85]), or clinician requests (29% [25 of 85]), prompting further revision of UK national standards to encourage testing of all submitted samples (Chalmers et al., 2015).

Owing to higher prevalence of Cryptosporidium in children, patient age is currently used as the primary selection criterion for Cryptosporidium testing within the UK and Ireland (Chalmers and Davies, 2010). It is widely acknowledged that age bias impacts on the reported age distribution rates of a variety of pathogens, including
Adding to this bias are parental and health care professional behaviours towards gastrointestinal symptoms and stool samples collection, respectively, leading to higher reporting of *Cryptosporidium* cases in children than in adults (Garvey and McKeown, 2009). There also remains a wide-scale lack of standardisation in clinical *Cryptosporidium* detection practices both within and between nations, and comparative information is particularly limited (Manser et al., 2014).

A 2004 report by the Health Protection Surveillance Centre in Ireland (HPSC) recommended the testing of all stool samples from patients exhibiting clinical symptoms associated with cryptosporidiosis (Health Protection Surveillance Centre, 2004). However, should resource and logistical constraints prohibit this, it is recommended that all patients under the age of 10 be screened as an alternative, with a comparable age threshold of 15 advised in a similar UK report (Crook et al., 2002). It is important to note, however, that should such an age threshold be employed, it should only be applied for the investigation of sporadic cases rather than in outbreak situations. A study of a 2007 outbreak in Galway, in the Republic of Ireland concluded that 40% of all infections occurred in patients over the age of 15. Thus, it appears likely that laboratories using such thresholds also fail to detect a large proportion of sporadic cases (Pelly et al., 2007).

In many countries, such as the United States and France, *Cryptosporidium* screening is not a routine component of standard “ova plus parasite” examinations carried out in clinical laboratories regardless of patient age, clinical and epidemiological evidence, unless specifically requested by a clinician or recommended by the laboratory directorate. Thus, further contributing to the underreporting of cases (ANOFEL Cryptosporidium National Network, 2010; Chen et al., 2002). Meanwhile, a recent European survey of 18 laboratories found noteworthy variety between detection methods with almost all laboratories relying upon microscopic methods either alone or in combination with other detection methods. Half of all laboratories surveyed used at least two methods of detection, only two of which employed molecular diagnostics (Manser et al., 2014).

Expertise in the field of stool microscopy is declining among the modern clinical laboratory workforce, particularly in areas of decreasing prevalence of faecal parasites (McHardy et al., 2014). While a number of studies have reported that molecular
methods are employed by only a minority of routine clinical microbiology laboratories in Europe and the US (Chalmers and Davies, 2010; Fournet et al., 2013; Jones et al., 2004; Manser et al., 2014; ten Hove et al., 2007), the growing acceptance of molecular methods and increasing throughput of samples has prompted the necessity for automated, walk-away technology in these laboratories.

Additionally, given the low prevalence of *Cryptosporidium* within the population in developed countries at least, certain laboratories may not receive adequate levels of positive samples to establish and maintain expertise in this area (McHardy et al., 2014). Consequently, although microscopy is currently regarded as the gold standard in the diagnosis of cryptosporidiosis, it seems likely that molecular methods will eventually replace microscopy altogether.

A summary of the advantages and disadvantages associated with the various diagnostic methods discussed herein are outlined in Table 1.1.

**Table 1.1** Advantages and disadvantages of microscopy, immunological and molecular diagnostic methods for *Cryptosporidium* spp.

<table>
<thead>
<tr>
<th>Diagnostic Test</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Microscopy</strong></td>
<td>- Relatively low cost</td>
<td>- Poor sensitivity</td>
</tr>
<tr>
<td></td>
<td>- Widely available</td>
<td>- Time consuming</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Skilled microscopist essential</td>
</tr>
<tr>
<td><strong>Imunoassay based methods</strong></td>
<td>- Good sensitivity</td>
<td>- Not widely available in developing countries due to cost constraints and limited detection spectrum of kits</td>
</tr>
<tr>
<td></td>
<td>- Wide variety of kits available</td>
<td>- False positives</td>
</tr>
<tr>
<td></td>
<td>- Convenient adjunct to microscopic analysis</td>
<td></td>
</tr>
<tr>
<td><strong>Molecular/ nucleic acid amplification methods</strong></td>
<td>- Exceptional sensitivity</td>
<td>- Expensive reagents and instrumentation required</td>
</tr>
<tr>
<td></td>
<td>- Capable of species and subtypes identification</td>
<td>- Requires skilled technician</td>
</tr>
<tr>
<td></td>
<td>- Option to multilplex detection of several enteric pathogens</td>
<td></td>
</tr>
</tbody>
</table>
1.8.1 Brightfield and Fluorescent Staining

Faecal investigation for the presence of shed oocysts or antigens is the diagnostic mainstay in *Cryptosporidium* detection (Manser *et al.*, 2014). Conventional clinical diagnosis has largely relied on microscopic examination of tinctorially or fluorescently stained faecal smears. The acid-fast properties of *Cryptosporidium* were demonstrated in 1981, with the development of a modified Ziehl-Neelson (mZN) stain for differential staining (Henricksen and Pohlenz, 1981). Prior to this *Cryptosporidium* was largely identified through Giemsa staining of histological preparation of intestinal biopsy samples, with iodine, trichrome and iron haematoxylin stained faecal specimens yielding poor results (Kissinger, 2008; McNabb *et al.*, 1985).

A variety of stains including the acid-fast Kinyoun’s stain and differential stains such as the hot safrinin-methylene blue stain have also been employed by clinical laboratories (Baxby *et al.*, 1984; Kageruka *et al.*, 1984). In addition, a variety of negative and fluorescent stains have been developed. Both staining techniques provide a rapid, inexpensive, sensitive alternative to the acid-fast techniques (Casemore *et al.*, 1985; Garcia *et al.*, 1983; Hanscheid *et al.*, 2008; Khurana *et al.*, 2012; Vohra *et al.*, 2012). Despite this, acid-fast staining, particularly the mZN technique, predominates in clinical laboratories (Manser *et al.*, 2014). However, although there is a marked contrast between the red stained oocyst against the green background counterstain of the mZN, yeasts, fungal and bacterial spores may be erroneously identified as oocysts (Casemore, 1991). Intermittent oocyst shedding is inherent to the *Cryptosporidium* life cycle. Thus, in order to improve diagnostic sensitivity, faecal specimens are often collected over three different days, adding to the laboratory workload (Goñi *et al.*, 2012; van Gool *et al.*, 2003). Overall, this time consuming and tedious staining technique demands an experienced microscopist, but exhibits poor sensitivity (37-100%) (Abou El-Naga and Gaafar, 2014; Chalmers *et al.*, 2011a; Kaushik *et al.*, 2008; Tuli *et al.*, 2010; Zaglool *et al.*, 2013).

Unsurprisingly, given the variable levels of sensitivity associated with brightfield and fluorescent staining techniques, a number of oocyst concentration methods have been developed in order to maximise oocyst yields from faecal samples (Garcia *et al.*, 1983; Weber *et al.*, 1991). These techniques are most useful when preserved stool specimens are received, in epidemiological cases, asymptomatic cases, and in
immunocompromised patients with a clinical history of unexplained diarrhoea, as this patient population is susceptible to recrudescence following periods of remission (Casemore, 1991; Omoruyi et al., 2014).

Although faecal staining methods remain the cornerstone of parasitological investigations in both American and European clinical laboratories, these methods were previously surpassed by a variety of immunological and most significantly, molecular identification methods (Jones et al., 2004; Manser et al., 2014). Cryptosporidium targeting, immunofluorescent monoclonal antibodies (MAb) were initially introduced almost three decades ago following the advent of hybridoma technology, which allowed for the generation of highly specific antibodies (Milstein, 1999; Sterling and Arrowood, 1986). Monoclonal oocyst wall antibodies conjugated with fluorescent labels such as fluorescein isothiocyanate (FITC) and biotin hydrazide, which imparts a distinct apple green-to-yellow fluorescence to the oocysts against a dark background, allow visualisation of intact parasites (Arrowood and Sterling, 1989; Garcia et al., 1987; Sterling and Arrowood, 1986). Comparative studies have found the sensitivity and specificity of immunofluorescent techniques to outweigh the sensitivity and specificity exhibited by conventional brightfield staining techniques (Alles et al., 1995; Arrowood and Sterling, 1989; Current and Garcia, 1991; Elsafi et al., 2014; Garcia et al., 1992; Kamal et al., 2008). Additionally, indirect immunofluorescent techniques, although requiring an additional incubation step, have been reported to possess similar levels of sensitivity and specificity to those of their direct counterparts (Rusnak et al., 1989; Stibbs and Ongerth, 1986).

While immunofluorescent detection of Cryptosporidium spp. necessitates the use of a fluorescent microscope, which has prevented widespread utilisation of this method and may preclude the use of this technique in developing countries, the marked distinction of oocysts from the non-fluorescent background conveniently reduces the amount of time required for microscopic investigation (Vohra et al., 2012). In addition, faecal concentration is not a prerequisite when faecal samples contain a paucity of oocysts, given the sensitivity of this method (Elsafi et al., 2014). Immunofluorescence enhances the ease with which less experienced microscopists can definitively identify the presence of oocysts (Alles et al., 1995; Garcia et al., 1987).
1.8.2 Enzyme Immunoassays, ELISA and Immunochromatographic Methods

Faecal-antigen diagnostic techniques have been developed in order to obviate the need for skilled microscopists, laborious methodologies and specialised equipment, such as fluorescent microscopes, while also accommodating batch testing requirements (Helmy et al., 2014; Rosenblatt and Sloan, 1993; Ungar, 1990). Indeed, the colorimetric principles underlying quantitative enzyme immunoassays (EIA) and enzyme linked immuno-sorbent assays (ELISA) eliminate the requirement for extensive microscopy training of lab personnel and subjectivity associated with conventional microscopy (Goñi et al., 2012; Newman et al., 1993). Comparative studies investigating the diagnostic utility of EIA and ELISA kits have found that they provide significantly improved sensitivity (94 - 100%) and specificity (93 -100%) over conventional acid-fast staining methods (Kehl et al., 1995; Parghi et al., 2014; Siddons et al., 1992). However, comparisons between fluorescent and immunofluorescent staining methods have indicated enzyme-based immunological detection of Cryptosporidium to be inferior, with reduced capabilities of detecting low oocyst densities (Johnston et al., 2003; Kehl et al., 1995; Newman et al., 1993; Weitzel et al., 2006). In addition, several cases detailing the generation of false positive results by the ProSpect Cryptosporidium immunoassay (Alexon, Inc., Mountain View, California) have been reported, perhaps owing to faecal antigen shedding often persisting after intact oocyst shedding has abated (Doing et al., 1999; Johnston et al., 2003; Miller and Mojica, 1999).

Immunochromatographic kits provide a detection system that surpasses enzyme-based methods in terms of rapidity by eliminating the need for additional reagent additions, washing steps and incubations (Current and Garcia, 1991; Garcia et al., 2003). Antigen migration via capillary action allows detection of Cryptosporidium antigens by a discrete, colloidal dye labelled antibody impregnated in a line assay, permitting objective antigen detection (Llorente et al., 2002). Sensitivities and specificities of these qualitative faecal-antigen kits vary considerably, with one study which investigated four kits reporting ranges of 47 - 71 % and 98 - 100%, respectively, when compared to microscopy (Agnamey et al., 2011). Meanwhile three independent studies reported sensitivities and specificities of 98% and 100%, 98% and 100%, 100% and 99%, respectively, in immunochromatographic kits. In each study the
immunochromatographic kits in question were compared to microscopy, EIA and ELISA, respectively (Chan et al., 2000; Garcia and Shimizu, 2000; Youn et al., 2009).

Like their enzyme-based counterparts, immunochromatographic kits have been found to generate false positive results, resulting in one case, for example, in a lot recall of the CoulorPAC™ Cryptosporidium/Giardia rapid assay kit (Haupst and Davis, 2002). Immunoassay based kits also offer a reduced diagnostic spectrum, as many are tailored solely for the detection of C. parvum and C. hominis. Therefore the clinical utility of such kits is limited in regions where alternative Cryptosporidium species are attributable to a significant number of cryptosporidiosis cases (Agnamey et al., 2011; Llorente et al., 2002). Consequently, despite the logistical and economical improvements in assay methodology over conventional staining methods, enzyme and non-enzyme based immunoassays are not deemed to be a suitable substitution for such techniques in the modern clinical laboratory, even though they may be used as a confirmatory adjunct to conventional methods in clinical laboratories with limited experience in Cryptosporidium detection, or for epidemiological studies (Checkley et al., 1997; Goñi et al., 2012; Hanson and Cartwright, 2001; Weitzel et al., 2006)

1.8.3 Molecular Approaches

Following its inception in 1983, the polymerase chain reaction (PCR) has vastly improved molecular diagnostic approaches in many fields, including clinical microbiology (Espy et al., 2006; Tong and Giffard, 2012). PCR detection of Cryptosporidium has been proven to be more sensitive than conventional microscopic and immunological methods, while also permitting batch testing, species and sub-species identification of detected organisms (Chalmers et al., 2011a; Elsafi et al., 2013; Aghamolaie et al., 2016; Uppal et al., 2014).

A sequence survey identifying >250 kb of the C. parvum genome heralded the beginning of the genomic era of Cryptosporidium research in the late 1990s (Liu et al., 1999). Given the clinical significance and the dearth of epidemiological and molecular Cryptosporidium data at the time, the National Institute of Allergy and Infectious Diseases (NIAID) subsequently allocated funding to a consortium of three American universities, which were tasked with sequencing both C. parvum and C. hominis genomes (Widmer and Sullivan, 2012). Two separate studies, undertaking whole
genome shotgun sequencing strategies, subsequently yielded the fully sequenced genomes of *C. parvum* IOWA and *C. hominis* TU502 isolates (Abrahamsen et al., 2004; Xu et al., 2004). This research ultimately led to the establishment of CryptoDB in 2003, an online database of known *Cryptosporidium* genomes (Heiges et al., 2006; Puia et al., 2004). This collaborative effort integrates all genomic and functional genomic data pertaining to *Cryptosporidium* spp. in a single online, bioinformatics resource. At present, CryptoDB houses 15 genome sequences encompassing nine different species, while 52 genome assemblies are available for the *Cryptosporidium* genus in the NCBI GenBank (Baptista et al., 2021). These genomic advances have paved the way for current epidemiological studies, functional analyses, protein and metabolic pathway predictions and genome annotation (Isaza et al., 2015; Mazurie et al., 2013).

In the initial advent of molecular techniques, characterisation of *Cryptosporidium* spp. was largely achieved through PCR-mediated amplification of specific genetic regions, followed by enzymatic cleavage or sequencing (Cheun et al., 2013; Ibrahim et al., 2021; Insulander et al., 2013; Sulaiman et al., 2005). Prior to the *C. hominis* and *C. parvum* whole genome sequencing projects and the widespread availability of DNA sequencing techniques, RFLP was the primary means by which to conduct inter-species genotyping on *Cryptosporidium* spp. (Awad-el-Kariem et al., 1994; Leng et al., 1996). This technique utilises a number of key molecular markers amenable to PCR amplification and restriction digestion to produce unique banding patterns that are visualised via gel electrophoresis (Roellig and Xiao, 2020). However, RFLP based genotyping is limited in that it cannot resolve differences between isolates of the same species, particularly *C. hominis* and *C. parvum*, necessitating an alternative sub-genotyping technique (Roellig and Xiao, 2020).

The *C. parvum* and *C. hominis* genome sequencing projects enabled the identification of a number of highly polymorphic micro- and minisatellite loci and conserved loci flanking sequences (Aiello et al., 1999; Cacciò et al., 2000; Feng et al., 2000). These sequences permitted the development of microsatellite and minisatellite locus-specific PCR assays for both species regions, thereby permitting genotyping superior to that of RFLP, and ultimately the development of a technique that is also capable of identifying the subtleties of intra-species differentiation (Ramo et al., 2015; Xiao and Feng, 2017).
Key molecular markers for species identification include the small sub-unit rRNA (SSU rRNA), *Cryptosporidium* outer wall protein (COWP), 70-kDa heat shock protein (HSP70), thrombospondin-related adhesive protein (TRAP-C2) and actin genes (Roellig and Xiao, 2020; Elwin *et al.*, 2013; Hadfield *et al.*, 2011). These regions contain large numbers of interspecific polymorphisms, which make them ideal for basic species identification. Owing to low levels of intraspecific variation, the SSU rRNA gene is the most widely used of these genetic targets in genotypic differentiation between an array of human and animal infecting species (de Lucio *et al.*, 2016; Roellig and Xiao, 2020). Highly variable regions, such as the tandem repeat containing 60-kDa glycoprotein (GP60) gene and the microsatellite loci, ML1 and ML2, are predominantly used for this purpose given the marked amount of intra-species sequence heterogeneity expressed in these regions (Robinson and Chalmers, 2012). These regions have been pivotal in the determination of the extensive number of *C. parvum* and *C. hominis* subspecies, or subtypes. Extensive panels including these markers, among many others, are commonly employed in multi-locus sequence typing (MLST) based epidemiological studies to identify population structures, and inter- and intra-species genetic diversity (De Waele *et al.*, 2013; Feng *et al.*, 2014; Garcia-R *et al.*, 2020; Wang *et al.*, 2015; Xiao, 2010).

Within the *Cryptosporidium* genus and more specifically among the predominant human-pathogenic species, *C. parvum* and *C. hominis*, asexual and sexual life cycle stages, genetic recombination and selective pressures, such as parasite-host coevolution, host adaptation and geographic segregation, have led to generation of new subtype families and diverse genetic populations (Abal-Fabeiro *et al.*, 2013; Feng *et al.*, 2002; Garcia-R and Hayman, 2017). gp60, which is firmly established as a key marker of genetic variation within *Cryptosporidium* spp. (Abal-Fabeiro *et al.*, 2013), is subject to selective pressure which has resulted in a lack of global sub-structuring, with the same gp60 alleles emerging in different locations globally (Abal-Fabeiro *et al.*, 2013; Widmer, 2009). Thus, gp60 is not a sufficient descriptor of population structure to enable single locus typing (Robinson and Chalmers, 2012). Multi-locus genotyping (MLG) is necessary to adequately assess genetic variation and population structures within *Cryptosporidium* spp.

Research is ongoing into the identification of novel genetic markers, and refining known markers into an internationally standardised MLG scheme (Chalmers *et al.*, 2013).
Population-level analyses (Nucleotide diversity (θπ) and Watterson’s theta (θW), Tajima’s D statistic etc.) have been used in several studies seeking to assess genetic diversity and evolutionary processes at multiple loci within Cryptosporidium spp. in order to aid understanding of host-parasite adaptation and evolution (Garcia-R et al., 2020; Garcia-R and Hayman, 2017). MLG, accomplished via DNA sequencing of PCR amplified amplicons from specific loci, or real-time PCR based high resolution melting analysis of loci amplicons, is also being evaluated for epidemiological surveillance and outbreak investigations in clinical laboratories (Chalmers et al., 2017).

Overall, several studies have recommended the incorporation of PCR techniques into routine clinical Cryptosporidium diagnosis methods (Rubio et al., 2014; Stensvold et al., 2011; Uppal et al., 2014). Indeed, PCR has been reported as having improved sensitivity over current detection methods, with reported limits of detection ranging from $1 \times 10^5$ to 1 oocyst/Gram of faeces (Costa et al., 2021). This is a significant improvement in detection when compared to the limit of detection associated with microscopy based detection, which requires between $1 \times 10^4 – 5 \times 10^4$ oocysts per mL of faeces (Khurana and Chaudhary, 2018). Although promising results have been reported, standard PCR techniques, particularly nested PCR, are not ideally suited to routine human diagnostics as they have been associated with considerable contamination risks, owing to multiple rounds of DNA amplification and concomitant DNA manipulation steps.

The development of real-time PCR (or quantitative PCR, qPCR) offers a convenient alternative to conventional techniques. Completed within an hour or less, qPCR reaction times are superior to those seen in conventional PCR methods (Espy et al., 2006). Fluorescent probes are used to detect DNA amplification, while the closed reaction vessel ensures that contamination is comparably negligible to that associated with conventional PCR (Espy et al., 2006; Minarovičová et al., 2009). In addition, sensitivity and specificity levels equal to, or surpassing, those observed in conventional PCR have been reported (Elsafi et al., 2013; Hadfield et al., 2011; Liu et al., 2013). Table 1.2 provides a comparison between reported sensitivities and specificities of common microscopic methods versus currently available commercial immunological and DNA-based diagnostic panels.
Multiplex PCR has become a popular means by which to investigate the presence of multiple enteric parasites in a single sample. Initial studies found multiplex qPCR assays to be both 100% sensitive and specific when compared to the monoplex qPCR assays, with superior sensitivity and specificity over conventional microscopic methods for each of the protozoan parasites tested (Stark et al., 2011). This technology continues to increase in popularity for use in clinical laboratories, with in excess of 10 enteric pathogen-targeting commercial multiplex qPCR kits currently available, including Allplex, Amplidia, BD Max, Biofire FilmArray, FTD Stool Parasites, Gastro Panel EntericBio Panel II, Gastroenteritis/Parasite Panel I, Luminex, NanoChip, PARAGenie, RIDAGENE, and QIAstat (as detailed in Table 1.2) (Autier et al., 2020; Boers et al., 2020; Morio et al., 2019; Hannet et al., 2019; Koziel et al., 2013; Ryan et al., 2017; Paulos et al., 2019). In addition, multiplex assays significantly reduce reagent and labour costs, and outperform the majority of alternative methods currently in use (Haque et al., 2007; Stark et al., 2011; ten Hove et al., 2007; Verweij and Van Lieshout, 2011). qPCR also readily accommodates genotyping via melting curve analysis, and more recently high resolution melting (HRM) curve analysis (Chelbi et al., 2018; Lalonde et al., 2013).

While appearing to offer a myriad of diagnostic and logistical advantages, conventional PCR and qPCR have not yet been widely incorporated into routine Cryptosporidium detection procedures. This is likely to be due, at least in part, to the requirement for significant investment in reagents and analysers, coupled with extensive personnel training (Burnet et al., 2013; Checkley et al., 2014). However, a future molecular shift in the field of diagnostic parasitology would appear to be inevitable in order to provide essential improvement to current diagnostic services. Thus, the potential benefits must be carefully weighed against the perceived disadvantages associated with molecular methods in the context of current diagnostic limitations.
### Table 1.2  Sensitivities and Specificities of Microscopic Methods, and Currently Available Immunological and DNA-based Diagnostic Tests for *Cryptosporidium* spp.

#### Microscopic Staining Techniques

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>mZiehl-Neelsen stain</td>
<td>37–79.1%</td>
<td>100%</td>
<td>(Chalmers <em>et al.</em>, 2011a; Kaushik <em>et al.</em>, 2008; Khurana <em>et al.</em>, 2012)</td>
</tr>
<tr>
<td>Fluorescent – Auramine phenol</td>
<td>92.1 –100%</td>
<td>99.6–100%</td>
<td>Chalmers <em>et al.</em>, 2011a; Khurana <em>et al.</em>, 2012</td>
</tr>
<tr>
<td>stain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kinyoun’s acid fast stain</td>
<td>66.7–91.6%</td>
<td>88.2–100%</td>
<td>(Abou El-Naga and Gaafar, 2014; Elsafi <em>et al.</em>, 2014)</td>
</tr>
</tbody>
</table>

#### Commercial Immunological Diagnostic Tests for *Cryptosporidium* spp.

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Additional Pathogens Detected</th>
<th>Supplier</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crypto-Strip</td>
<td>47.2%</td>
<td>100%</td>
<td>N/A</td>
<td>Coris BioConcept, Gembloux, Belgium</td>
<td>(Agnamney <em>et al.</em>, 2011)</td>
</tr>
<tr>
<td><em>Cryptosporidium</em> and <em>Giardia</em> Duo-Strip</td>
<td>91.7%</td>
<td>100%</td>
<td><em>G. duodanalis</em></td>
<td>Coris BioConcept, Gembloux, Belgium</td>
<td>(Van den Bossche <em>et al.</em>, 2015)</td>
</tr>
<tr>
<td><em>Cryptosporidium</em> EZ VUE lateral-flow test strips</td>
<td>89%</td>
<td>99%</td>
<td>N/A</td>
<td>TechLab Inc., Blacksburg, Virginia, United States</td>
<td>(Johansen <em>et al.</em>, 2021)</td>
</tr>
<tr>
<td><em>Cryptosporidium</em> II test</td>
<td>71.8%</td>
<td>94.3%</td>
<td>N/A</td>
<td>TechLab Inc., Blacksburg, Virginia, United States</td>
<td>(Kabir <em>et al.</em>, 2018)</td>
</tr>
<tr>
<td><em>Giardia/Cryptosporidium</em> Quik Chek</td>
<td>92.3–100%</td>
<td>97.1–100%</td>
<td><em>G. duodanalis</em></td>
<td>TechLab Inc., Blacksburg, Virginia, United States</td>
<td>(Chalmers <em>et al.</em>, 2011a; Van den Bossche <em>et al.</em>, 2015; Kabir <em>et al.</em>, 2018; Minak <em>et al.</em>, 2012)</td>
</tr>
<tr>
<td>ImmunocardSTAT® C/G</td>
<td>5.5–96%</td>
<td>96.6–100%</td>
<td><em>G. duodanalis</em></td>
<td>Meridian Bioscience Inc., Cincinnati, Ohio, United States</td>
<td>(Agnamney <em>et al.</em>, 2011; El-Moamly and El-Sweify, 2012; Bouyou-Akotet <em>et al.</em>, 2016)</td>
</tr>
<tr>
<td>ImmunoCard STAT!® CGE</td>
<td>100%</td>
<td>45.6–100%</td>
<td><em>G. duodanalis; E. histolytica</em></td>
<td>Meridian Bioscience Inc., Cincinnati, Ohio, United States</td>
<td>(Van den Bossche <em>et al.</em>, 2015)</td>
</tr>
</tbody>
</table>
### Commercial Molecular Diagnostic Tests for *Cryptosporidium* spp.

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Additional Pathogens Detected</th>
<th>Supplier</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allplex™ Gastrointestinal Panel-Parasite Assay (GIPPA)</td>
<td>100%</td>
<td>Up to 100% – further testing needed.</td>
<td><em>G. duodenalis</em>, <em>E. histolytica</em>, <em>D. fragilis</em>, <em>B. hominis</em>, and <em>C. cayetanensis</em></td>
<td>Seegene Inc, Seoul, Korea</td>
<td>(Autier et al., 2020; Paulos et al., 2019)</td>
</tr>
<tr>
<td>Amplidiag® Stool Parasites</td>
<td>10³ oocysts/g</td>
<td>Not specified</td>
<td><em>G. duodenalis</em>, <em>E. histolytica</em>, <em>D. fragilis</em></td>
<td>Mobidiag, Espoo, Finland</td>
<td>(Costa et al., 2021)</td>
</tr>
<tr>
<td>BD Max parasitic panel (EPP)</td>
<td>95.5%</td>
<td>99.6%</td>
<td><em>G. duodenalis</em>, <em>E. histolytica</em></td>
<td>BD Diagnostics, Sparks, Maryland, United States</td>
<td>(Madison-Antenucci et al., 2016; Mölling et al., 2016)</td>
</tr>
<tr>
<td>Biofire FilmArray™ Gastrointestinal Panel</td>
<td>100%</td>
<td>99.6–100%</td>
<td>14 bacterial and 5 viral targets. 3 further parasites: <em>G. duodenalis</em>, <em>E. histolytica</em>, and <em>C. cayetanensis</em>.</td>
<td>BioFire Diagnostics, Salt Lake City, Utah, United States</td>
<td>(Murphy et al., 2017; Binnicker, 2015; Khare et al., 2014; Zhang et al., 2015)</td>
</tr>
<tr>
<td>EasyScreen™ Enteric Parasite Detection Kit</td>
<td>100%</td>
<td>100%</td>
<td><em>G. duodenalis</em>, <em>Entamoeba</em> complex; <em>D. fragilis</em>, <em>Blastocystis</em> spp.</td>
<td>Genetic Signatures, Sydney, Australia</td>
<td>(Stark et al., 2014)</td>
</tr>
<tr>
<td>Test</td>
<td>Sensitivity</td>
<td>Specificity</td>
<td>Target pathogens</td>
<td>Manufacturer and Location</td>
<td>Publication</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>EntericBio GastroPanel II</td>
<td>100%</td>
<td>100%</td>
<td>4 bacterial targets; G. duodenalis</td>
<td>Serosep, Limerick, Ireland</td>
<td>(McAuliffe et al., 2017)</td>
</tr>
<tr>
<td>FTD Stool Parasites</td>
<td>53.1%</td>
<td></td>
<td>Giardia spp; E. histolytica</td>
<td>Fast Track Diagnostics, Esch-sur-Alzette, Luxembourg</td>
<td>(Paulos et al., 2019)</td>
</tr>
<tr>
<td>Gastroenteritis/Parasite Panel I</td>
<td>92–100%</td>
<td></td>
<td>G. duodenalis; E. histolytica</td>
<td>Diagnose, Seraing, Belgium</td>
<td>(Paulos et al., 2019)</td>
</tr>
<tr>
<td>Luminex xTAG® Gastrointestinal Pathogen Panel</td>
<td>95–100%</td>
<td>100%</td>
<td>12 bacterial and viral targets and 2 further parasites: G. duodenalis; E. histolytica</td>
<td>Luminex Corporation, Austin, Texas, United States</td>
<td>(Patel et al., 2014; Wessels et al., 2014; Claas et al., 2013; Navidad et al., 2013; Perry et al., 2014; Zhang et al., 2015)</td>
</tr>
<tr>
<td>NanoCHIP® Gastrointestinal Panel (GIP)</td>
<td>Detection limit of 5 x 10³ oocysts</td>
<td></td>
<td>3 bacterial targets and G. duodenalis, E. histolytica, E. dispar, D. fragilis, and Blastocystis hominis</td>
<td>Savyon Diagnostics, Ashdod, Israel</td>
<td>(Dror et al., 2016)</td>
</tr>
<tr>
<td>ParaGENIE Crypto-Micro PCR</td>
<td>91.7%</td>
<td>100%</td>
<td>Also differentiates between Enterocytozoon bieneusi and Encephalitozoon intestinalis</td>
<td>Ademtech, Pessac, France</td>
<td>(Morio et al., 2019)</td>
</tr>
<tr>
<td>RIDA®GENE Parasitic Stool Panel</td>
<td>87.5%</td>
<td></td>
<td>G. duodenalis; E. histolytica; D. fragilis</td>
<td>R-Biopharm Diagnostic, Darmstadt, Germany</td>
<td>(Paulos et al., 2019)</td>
</tr>
<tr>
<td>QIAStat Dx® GIP</td>
<td>Not specified</td>
<td>Not specified</td>
<td>14 bacterial targets, 6 viral targets and 3 further parasites: C. cayetanensis; E. histolytica; G. lamblia</td>
<td>Qiagen, Hilden, Germany</td>
<td>(Boers et al., 2020)</td>
</tr>
</tbody>
</table>

*Sensitivity and specificity for Cryptosporidium spp are given here. Variable sensitivities and specificities reported for other pathogens detection by these panels.*
1.9 Future Trends in Molecular Platforms

In recent years a third-generation implementation of conventional PCR that obfuscates the need for calibration curves in the quantification of nucleic acid targets, digital PCR (dPCR), has been gaining traction for its utility in pathogen detection. The method, which is based on the principle of amplifying a single DNA template from maximally diluted samples, therefore generating amplicons that are exclusively derived from one template, based on Poisson statistics, remains a relatively new concept in the field of medical parasitology (Pomari et al., 2019). Unlike qPCR, which produces an exponential signal, dPCR generates linear, digital signals, allowing quantitative analysis of the PCR product, detecting very rare mutations with unprecedented precision and sensitivity (Pohl and Shih, 2004). As of yet, application of dPCR to the detection and quantification of Cryptosporidium in human faecal samples is limited to a single study (Yang et al., 2014). However, despite this, the reported precision of dPCR is consistently superior to that of qPCR, and the quantitative detection less affected by the presence of inhibitors. Thus, the initial application of dPCR to the detection and quantification of Cryptosporidium oocysts by Yang et al. (2014) may herald its incorporation into mainstream clinical diagnostic methodologies over the coming years.

The emergence and increasing prevalence of next generation sequencing (NGS) technologies is also likely to shape the field of protozoan parasitology over the course of the coming decade, (DeMone et al., 2020). Evaluations of Cryptosporidium isolates with NGS techniques have revealed unprecedented within-isolate genetic diversity to a degree that is not possible to discern through current conventional PCR and Sanger sequencing-based subtyping methodologies, given their inability to resolve complex DNA mixtures and detect low-abundance intra-isolate alleles (Grinberg et al., 2013; Grinberg and Widmer, 2016; Zahedi et al., 2017). Consequently, NGS-based studies have already advanced current knowledge on the taxonomic distribution and transmission dynamics of Cryptosporidium spp., and may also play a future role in outbreak identification and surveillance of new and virulent subtypes (Zahedi et al., 2018; Zahedi et al., 2017).

Despite difficulties in extracting high quality pure DNA from clinical samples, issues with uneven depth of read coverage that leads to gaps in the assembled genome sequence, all of which impact cost and may preclude the widespread use of NGS
platforms in clinical laboratories and Public Health agencies; NGS remains likely to exert an important, indirect impact on clinical diagnostics through informing the development of much needed multi-locus sequence typing (MLST) schemes (Cacciò and Chalmers, 2016; Morris et al., 2019). Furthermore, NGS will also undoubtedly play an increasingly pivotal role in epidemiological analyses of Cryptosporidium spp., in addition to vaccine and drug development, over the coming decade (Morris et al., 2019; Zahedi et al., 2017).

1.10 Future Trends in Omics Research

It would be imprudent to ignore the role of the rapidly-developing fields of bioinformatics and proteomics in Cryptosporidium research, particularly in this post-genomic era. Efforts to integrate biochemical and genomic data have led to the development of predictive computational models, known as GEMs (genome scale models), for a number of microorganisms in recent years. These models utilise genomic and environmentally based parameters to predict phenotypic outcomes and growth based on biochemical mechanisms (Monk and Palsson, 2014). This is a concept that has already been put into practice for well characterised microorganisms, such as E. coli (Carrera et al., 2014), while attempts to develop a genome scale metabolic model of C. hominis have already been reported (Vanee et al., 2010). Given the fact that the genomes of both C. parvum and C. hominis has been fully sequenced, and in light of the studies carried out by Vanee and colleagues (2010), it is conceivable that GEMs may soon be employed to predict the phenotypic outcome environmental variations for Cryptosporidium species (Vinayak et al., 2015).

In silico approaches to drug discovery have also been undertaken and offer the advantage of predicting the structure and biological characteristics of proteins. For example, one such in silico study utilised the CryptoDB database (Heiges et al., 2006) to identify novel drug targets using the predicted proteome of Cryptosporidium spp. (Shrivastava et al., 2017). BLAST analysis of the proteins within this database revealed one unique hypothetical C. hominis protein as a potential candidate. However, the results of this study are limited given the ambiguity surrounding the unique nature of the proposed hypothetical protein target and its restriction to a single Cryptosporidium species (Shrivastava et al., 2017).
Panda and Mahapatra (2018) also employed an in silico approach, through the use of comparative genomics and immunoinformatics, in an attempt to identify novel therapeutic and vaccine candidates against the *C. parvum* proteome. Further in silico screening strategies like subcellular localization, drug target prioritization parameters and major histocompatibility complex (MHC) binding potential were employed in the screening of novel therapeutic candidates. This approach reported the essential cytoplasmic, signal peptide protein, cgd7_1830, as a novel drug target (Panda and Mahapatra, 2018).

In recent times, proteomic approaches have also spear-headed the development of novel diagnostic techniques. Initial efforts in developing a reproducible MALDI-TOF MS methodology and spectral marker “fingerprint” for *C. parvum* were first reported by Magnuson and colleagues, and subsequently refined by Glassmeyer et al. (Glassmeyer et al., 2007; Magnuson et al., 2000). Although MALDI-TOF MS based analysis of *Cryptosporidium* is still very much in its infancy, this new era of proteomic research is very likely to underpin the development of the next generation of diagnostic techniques. Further research in this area is ongoing, with a 2021 study employing a refined purification protocol to yield distinct, *Cryptosporidium* spp. specific spectra for *C. parvum* and *C. hominis* oocysts from clinical faecal samples (Gathercole et al. 2021).

With several mass spectrometry-based diagnostic tools already described for several blood-borne and enteric parasites, this avenue of study has the potential to yield an effective *Cryptosporidium* specific platform in the future.

### 1.11 Treatment Limitations, Propagation via Cell Culture and Implications for Future Advancements

To date, *Cryptosporidium* has remained a largely enigmatic pathogen, owing to its limited tractability and the difficultly encountered in successfully propagating the parasite in vitro in cell lines (Karanis and Aldeyarbi, 2011). In vitro culture efforts generally result in low yields of mature parasites, as current *Cryptosporidium* cell culture methods generally suffer from rapid host cell overgrowth and ageing, resulting in premature termination of the *Cryptosporidium* lifecycle (Hijjawi, 2010). Normal intestinal epithelial cell (IEC) models fail to adequately recapitulate human intestinal structure and function. Support of parasitic infection by these and immortalised adenocarcinoma derived human IEC models is generally limited to only a few days,
precluding parasitic life-cycle completion or continuous propagation (Bhalchandra et al., 2018).

Lack of suitably facile animal models, and molecular tools have also hampered progress in key areas such as developmental biology; the elucidation of host-parasite interactions; biochemical, immunological and molecular studies development; and, perhaps most significantly, evaluation and development of effective anti-cryptosporidial drug therapies (Di Cristina and Carruthers, 2018; Hijjawi et al., 2001). Nitazoxanide, a broad spectrum antimicrobial agent, is currently the only FDA approved treatment for cryptosporidiosis in patients 1 year and older (Sparks et al., 2015). Three double-blind placebo controlled studies suggest the efficacy of nitazoxanide in immunocompetent patients (Amadi et al., 2002a; Rossignol et al., 2006, 2001). Nitazoxanide, however, is not without considerable limitations in its utility among patient populations that are most vulnerable to Cryptosporidium infection. A study conducted on malnourished children found nitazoxanide to improve diarrhoea and morality rates, but the response rate was limited to 56% of studied patients (Amadi et al., 2002b). Nitazoxanide has also been found to be ineffective in AIDS patients (Amadi et al., 2009), while various other drugs such as paromomycin, azithromycin, rifamycin, and HIV protease inhibitors have also been unsuccessful in the treatment of cryptosporidiosis in AIDS patients (Checkley et al., 2014; Sparks et al., 2015).

The pursuit of an optimal in vitro culture system for Cryptosporidium has spanned four decades, with the first complete development of C. parvum reported in human, and porcine cell line models in 1984 (Current and Haynes, 1984). In recent years a variety of both cell-free and axenic culture-based systems have yielded promising results. Key milestones in the drive to develop a suitable cell culture-based in vitro model have also been published in the past decade (Alcantara Warren et al., 2008; Castellanos-Gonzalez et al., 2008). More recent advancements have sought to extend periods of culture survival (Castellanos-Gonzalez et al., 2013; Jossé et al., 2019; Karanis, 2018; Miller et al., 2018; Varughese et al., 2014). Hollow fibre technologies and stem-cell derived organoids, have also provided a basis for sustained in-vitro oocyst propagation (Morada et al., 2016). The development of a system employing stem-cell derived small intestinal, and lung organoids, capable of recapitulating the in vivo physiology of their original tissues, and to model Cryptosporidium infection has been described. These organoids support propagation and completion of the parasitic life-cycle, generating infectious
oocysts equivalent to those derived from animal models (Dutta et al., 2019; Heo et al., 2018).

Advancements in cell culturing methodologies have enabled several compound screening studies to be conducted. These screening studies have yielded promising potential anti-cryptosporidial compounds. A HCT-8 cell-based, high-throughput screen (HTS) of the Medicines for Malaria (MMV) Open Access Malaria box, a collection of 400 compounds selected from 19,000 structurally unique molecules that were shown to have activity against the *Plasmodium falciparum*, was conducted by Bessoff et al. (2014) and identified several scaffold structures with activity against *C. parvum*. Another study screened a bank of over 6,000 compounds known to exhibit anti-protozoan activity via a high-content imaging infection assay in HCT-8 cells resulting in the identification of the *Cryptosporidium* orthologue to *Plasmodium* lipid kinase PI(4)K, *Cryptosporidium* lipid kinase PI(4)K, as a potential target for pyrazolopyridines based therapeutic candidates (Manjunatha et al., 2017).

Most significantly, however, is the recent application of the CRISPR/Cas9 system genome editing technology to *Cryptosporidium*. In a major breakthrough for *Cryptosporidium* in vitro research, Vinayak et al (2015) harnessed the CRISPR/Cas9 system, to successfully transfect *C. parvum* sprozoites in tissue culture and subsequently isolate genetically modified *C. parvum* sporozoites. Such a demonstration of genetically modified *Cryptosporidium* sporozoites was the first of its kind, paving the way for the development of subsequent parasite survival assays that are furthering contemporary understanding of the basic biology and virulence of this pathogen, and tractable in vivo models (Pawlowic et al., 2019; Sateriale et al., 2019; Vinayak et al., 2015).

Similar to the advances made in NGS technologies, future applications of CRISPR/Cas systems may result in far reaching implications and advancements in diverse areas of *Cryptosporidium* research. For example, given the severe limitations associated with current treatment options, the recent advancements employing the aforementioned CRISPR/Cas9 genetic modification of *C. parvum* may prove critical in the development of vaccines and more effective drug therapies, particularly given the utility of gene ablation in identifying alternative treatment strategies (Doudna and Charpentier, 2014; Vinayak et al., 2015). Additionally, the recent incorporation of the CRISPR/Cas12a system within a fluorescent lateral flow strip biosensor tailored for on-site diagnosis of
Cryptosporidium parvum subtype family IId. demonstrates the clinical potential and diagnostic utility of CRISPR/Cas system in Cryptosporidium detection (Yu et al., 2021).

1.12 Concluding Remarks
The 2015 Nobel Prize in Physiology or Medicine is a pertinent reminder of the importance and necessity for continued research in the field of parasitology. The work of Tu Youyou and the combined efforts of William C. Campbell and Satoshi Ōmura were recognised for the discoveries made concerning the development of novel therapies for malaria, a parasite also of the apicomplexan phylum, and roundworm, respectively (Długońska, 2015). Such high profile recognition in the field of parasitology is arguably suggestive of a resurgence of interest in the field, which has stagnated somewhat in terms of clinical, therapeutic and molecular diagnostic advances.

The renaissance of Cryptosporidium research is already underway; the major breakthrough in producing CRISPR/Cas9 modified, tractable C. parvum oocysts signifies a reinvigoration of the field. This advancement is likely aid the advancement of knowledge in areas such as host-parasite interactions, and the biochemical and immunological pathways at play within Cryptosporidium spp. It will also facilitate the identification and validation of much needed drug targets.

It is also crucial to note that the causative agent remains undetected in 70% of gastroenteritis cases. Advances in diagnostic approaches for Cryptosporidium spp., and parasitic species in general, may play a significant role in further elucidating the nature of this unknown, and potentially diverse, pathogenic conglomerate (Freedman et al., 2015). This is particularly relevant in light of the human and economic toll exerted by gastrointestinal infection on a global scale, which according to a recent WHO study affected 2 billion people, caused 1 billion deaths and 78.7 million disability adjusted life years in 2010 (Kirk et al., 2015). As a consequence, Cryptosporidium research has perhaps never been more pertinent than it is today.
Chapter II: Validation of the EntericBio GastroPanel II for use in routine clinical diagnosis of cryptosporidiosis, and comparison with microscopic diagnostic methods

A manuscript based upon this chapter has been published in the Journal of Medical Microbiology.

2.1 Abstract

Despite the rapidly-widening application of molecular methods to clinical diagnostic settings, these technologies have only recently begun to be applied to the diagnosis of enteric parasitic infection. Commercial molecular panels currently available for laboratory-based diagnosis are limited to the most commonly encountered enteric parasites, so there is a continuing dependence on microscopic protocols for less commonly-detected parasites even when molecular methods are used.
2.2 Introduction

Cryptosporidium spp. and Giardia lamblia are, among others such as Blastocystis spp. and Dientamoeba fragilis, two of the most commonly encountered enteric parasites in both developing and developed countries (Osman et al., 2016; Roberts et al., 2013). Nevertheless, incidence rates of Cryptosporidium spp. rarely exceed 2 per 100,000 population in European countries, with similar rates reported in the United States (European Centre for Disease Prevention and Control, 2015a; Painter et al., 2015). Even in Ireland, which has the highest rates of cryptosporidiosis in Europe (12, 11 and 8, per 100,000 in 2012, 2013 and 2014, respectively), incidence is low (European Centre for Disease Prevention and Control, 2015a). The incidence of Giardia lamblia infection, although higher in Europe and the United States than Cryptosporidium spp. (5 per 100,000 population in 2014 and 6 per 100,000 population per 100,000 in 2012, respectively) is still generally low (European Centre for Disease Prevention and Control, 2015b). This low incidence, with obvious public health benefits, does however mean that effective validation of new detection methods remains a challenge because of the relatively small numbers of detections.

Our recent experience when investigating the efficacy of parasitological diagnostic methods is based on two studies conducted in the Medical Microbiology Department of Cork University Hospital (CUH), Ireland. The first was conducted to compare traditional microscopy based methods of detection with the molecular EntericBio Gastro Panel II kit (Serosep, Limerick Ireland). This study was conducted from April to August of 2015, with a view to introduce the EntericBio Gastro Panel II for routine clinical use. An advancement on the GastroPanel I, which had already been in routine use for some years in this laboratory prior to this study (Koziel et al., 2013), this panel was developed to include Cryptosporidium parvum/hominis and G. lamblia targets to the pre-existing bacterial enteric pathogen panel. The adoption of this revised molecular panel by this clinical laboratory replaced routine microscopic screening for these enteric parasites.

A subsequent microscopy-based study, investigating the impact of the replacement of microscopic screening techniques with real-time PCR diagnostic methods on rates of detection of enteric parasites not included in Gastro Panel II, was conducted during the summer months of 2016. Both study periods were chosen in an attempt to capitalise on the increased incidence of the specific enteric parasites encountered in Ireland during
these seasons, as reported in the literature and in all available national and international epidemiological reports (European Centre for Disease Prevention and Control, 2014; Zintl et al., 2006).

2.3 Materials and Methods

2.3.1 Sample Selection

2.3.1.1 Real-time PCR Validation Study
A total of 352 faecal samples submitted to the Medical Microbiology Department of CUH between April and August 2015 and which satisfied the laboratory acceptance criterion, which necessitated samples to correspond to Bristol Stool Chart types 5, 6 and 7, were tested in the first study for EntericBio Gastro Panel II verification.

2.3.1.2 Ova, Cyst and Parasite Microscopy Study
105 samples submitted to the laboratory requesting ova, cyst and parasite (OCP) investigation in patients with a history of foreign travel and presenting with diarrhoeal illness were examined. Samples were amassed during the summer months of 2016. Following faecal concentration, the resulting sample pellet was resuspended in saline and examined. Similarly, microscopy undertaken in this study was based on NEQAS guidelines as described in detail Section 2.3.2.1 below.

2.3.2 Microscopy

2.3.2.1 Real-time PCR Validation Study
Microscopic Methods for Cryptosporidium spp. Detection
Fluorescent auramine phenol staining and confirmatory modified cold Ziehl-Neelsen staining were employed in the diagnosis of Cryptosporidium spp.

The auramine phenol stain protocol was as follows: A medium to thick smear was prepared on a clean microscope slide and air dried. The slide was fixed in methanol for 3 minutes, and stained with auramine-phenol for 10 minutes. The slide was then washed in tap water, decolourised with 1% acid alcohol for 5 minutes and washed once again in tap water. The slide was counterstained using 0.1% potassium permanganate for 30 seconds, washed for a final time in tap water and air dried prior to examination using a fluorescence microscope.
Samples determined to be positive by fluorescent microscopy subsequently underwent confirmatory modified cold Ziehl-Neelsen acid-fast staining, as follows: A medium to thick smear was prepared on a clean microscope slide, air dried and fixed in methanol for 3 minutes. The slide was then stained with cold carbol fuchsin for 10 minutes and washed in tap water. The slide was decolourised with 1% HCl (v/v) in methanol (acid alcohol) for 10 seconds and washed again in tap water. The slide was counterstained using malachite green for 30 seconds, washed in tap water and air dried prior to microscopic examination using the high power objective lens of a standard bright-field microscope.

**Microscopic Methods for *G. lamblia* Detection**

Diagnosis of *Giardia* required samples to undergo modified formal-ether concentration prior to wet mount examination, which was achieved through the use of Midi Parasep faecal parasite concentrators (Apacor Ltd., UK), according to the manufacturer’s recommendations. The sample was resuspended in 0.85% saline prior to microscopic examination, with 1 to 2 drops of resuspended sample cover-slipped and examined under x20 and x40 magnification, respectively, for the presence of *Giardia* cysts.

Traditional wet mount microscopy conducted during this study was based on National External Quality Assessment Service (NEQAS) guidelines, which stipulate that the entire sample deposit should be examined, taking 10 minutes for examination per slide. However, for this study, three slides were examined per sample, amounting to a total of 30 minutes per sample.

**2.3.2.2 Ova Cyst and Parasite Microscopy Study**

OCP microscopic analysis was conducted as described for the detection of *Giardia lamblia* in the preceding study protocol described above. Similarly, microscopy undertaken in this study was based on NEQAS guidelines outlined in Section 2.3.2.1 above. It should be noted that within this study, real-time PCR testing for *Cryptosporidium* spp. and *Giardia lamblia*, via the GastroPanel II assay, was carried out concurrently with wet mount microscopy on all acceptable samples as part of the routine enteric pathogen investigation.
2.3.3 *EntericBio GastroPanel II Molecular Testing*

2.3.3.1 Real-time PCR Validation Study

All samples included in the study were tested using the EntericBio GastroPanel II (Serosep Ltd., Limerick), according to the manufacturer’s instructions. This product is a semi-automated and CE-marked real-time PCR assay that detects *Salmonella enterica* spp., *Shigella* spp., *Campylobacter jejuni/collillari*, verotoxigenic *E.coli* (VTEC), *Cryptosporidium* spp. and *Giardia lamblia* from a single sample within 3 hours without the need for formal nucleic acid extraction. The manufacturers provided two PCR tube strips, denoted A and B, that contained all necessary real-time PCR reagents, aside from the faecally extracted nucleic acid template, lyophilised within the tubes. Strip A contained primers for a multiplex reaction for the detection of *Salmonella enterica* spp., *Shigella* spp., and *Campylobacter jejuni/collillari*. Strip B contained primers for a multiplex reaction for the detection of VTEC, *C. hominis* and *C. parvum* and *G. lamblia* (assemblages A and B).

DNA extraction within this system has been reduced to a single-step process process. Briefly, the whole tip of a FLOQswab (Copan, Italy) was lightly coated with the specimen, which was inoculated into a 4 mL SPS tube and heated at 103°C for 30 min. Having been allowed to cool, the inoculated SPS tubes were transferred to the epMotion 5070 automated pipetting station (Eppendorf, Hamburg), wherein the GastroPanel II specific programme delivered the appropriate aliquots of sample/SPS solution and resuspension buffer to Tube strips A and B.

Upon completion of the pipetting programme, the PCR reaction tubes were vortexed briefly and centrifuged before being transferred to a Roche LightCycler 480 instrument (Roche Molecular Diagnostics, Basel, Switzerland) for real time PCR analysis. A combination of automatic flagging and manual interpretation were employed in identifying positive reactions. Positive reactions qualified as those which were found to produce a sigmoidal amplification curve of fluorescence and which exceeded the noise band.

It should be noted that microscopic slides were examined prior to the results of the real-time PCR. Microscopy was acknowledged as the gold standard in this study, although an expanded gold standard was employed when yielded discordant results arose. When
required, external testing was requested from the Cryptosporidium Reference Unit (CRU), Singleton Hospital, Swansea, Wales, in cases involving *Cryptosporidium* results and the NPRL, Hospital for Tropical Diseases, London, England, for anomalous *Giardia* results. Where samples were found to be positive by routine microscopy methods, but negative by PCR, the sample was re-tested with a fresh SPS tube and the use of an alternative, validated molecular method was employed. Where samples were negative by routine microscopy and positive by PCR, both PCR and microscopic examination were repeated and if deemed negative by microscopy again, the relevant external reference method was employed.

2.3.3.2 Ova, Cyst and Parasite Microscopy Study
Routine testing of samples requesting OCP analysis were analysed as per the EntericBio GastroPanel II protocol as described above.

2.4 Results

2.4.1 Real-time PCR Validation Study
Of the 352 samples, 38 samples were subsequently confirmed to be *Cryptosporidium parvum/hominis* positive and 6 samples were confirmed to be *Giardia lamblia* positive, with 100% concordance between the sensitivity, specificity, positive predictive value and negative predictive value of the expanded gold standard (microscopy and reference laboratory confirmation involving independent microscopy and PCR based methods) and molecular methods for all results.

2.4.2 Ova, Cyst and Parasite Microscopy Study
Overall, only 10 samples were found to contain faecal parasites, and 6 of these detections were of *G. lamblia* cysts. One of these samples also had *Entamoeba dispar* cysts detected. These 6 samples were also found to be positive by routine molecular methods for *G. lamblia*. Of the remaining 4 detections by microscopy, 3 were *Entamoeba coli*, and one other was *Cyclospora cayetanensis*. However, *C. cayetanensis* was only detected via modified cold Ziehl-Neelsen staining, as specifically requested by the resident consultant parasitologist. *Blastocystis* spp. or *Dientamoeba fragilis* were not detected in this study. One sample proved positive for
Cryptosporidium spp. by molecular methods, a diagnosis which was not detected by wet mount microscopy.

2.5 Discussion

This is the first report of the performance of the EntericBio GastroPanel II system for the detection of Cryptosporidium and Giardia. PCR exhibited 100% sensitivity and specificity in the detection of both parasites versus microscopy. This finding is somewhat unprecedented given the consensus regarding the superiority of PCR over microscopy in the detection of Cryptosporidium that has been established in the literature over the past two decades (Stark et al., 2011; Chalmers et al., 2011a). Although the results of this study appear contradictory to the current consensus in the field, several studies have reported either substantial agreement between both methods (Jerez Puebla et al., 2020), or reported that microscopy performs equally as well as, or better than real-time PCR in the detection of Cryptosporidium (Laude et al., 2016; Zheng et al., 2019).

The low numbers of positive cases of Cryptosporidium observed in this validation study may have contributed to the concordance observed between both methods; a higher number of Cryptosporidium positive samples may have allowed for a more robust comparison of methods, potentially producing results more consistent with those generally reported in the literature. Low incidence levels of pathogens of interest among routine samples are a common issue that often complicate many clinical laboratory validation studies. It may also be the case, however, that the dualistic approach of using fluorescent auramine-phenol, and confirmatory mZN staining for Cryptosporidium may have increased the sensitivity and specificity and reduced some of the user-related subjectivity associated with the method. Ultimately, however, the logistical benefits afforded to the Medical Microbiology laboratory through the use of the EntericBio platform, in conjunction with the concordance observed with the gold standard method in routine use, led to the implementation of this molecular method.

In light of the recommendations accompanying microscopic detection in order to arrive at a substantiated diagnosis, in addition to the sheer volume of samples to be tested daily, it would be imprudent to ignore the impact of real-time PCR on the laboratory workload in terms of negative screening. As acceptance criteria vary from laboratory to laboratory (Cacciò and Chalmers, 2016), this lack of standardisation may falsely lower
reported incidence rates. Additionally, large numbers of submitted samples are generally required to be investigated microscopically for either one or several parasites, the diagnostic particulars of which may require the use of multiple staining methods. Staining preparation and slide analysis thus amount to significant burdens in the context of such large sample numbers.

Conversely, although ultimately an inherent limitation of microscopy from an epidemiological standpoint, the inability of this method to distinguish beyond the *Cryptosporidium* genus may offer an advantage over current commercially available molecular platforms. Although molecular methods like real-time PCR are capable of detecting and differentiating between *Cryptosporidium* species (Mary et al., 2013), the vast majority of commercially available molecular panels, including the EntericBio GastroPanel II, identify only the two most prevalent species, *C. parvum* and *C. hominis* (Ryan et al., 2017). Therefore cases of cryptosporidiosis caused by species other than *C. parvum* and *C. hominis* infection will go undetected by these molecular panels. Despite reports of low sensitivity and dependency on user expertise, microscopy remains capable of identifying cases of *Cryptosporidium* infection regardless of species.

It should be noted that Irish epidemiological reports are built upon microscopy-based, genus level data, and no further testing is routinely conducted to identify infecting species. Beyond a handful of independent studies, none of which have been reported species besides *C. parvum* and *C. hominis*, no data pertaining to *Cryptosporidium* spp. implicated in human and livestock infection in Ireland currently exits (Zintl et al., 2009; Zintl et al., 2011). Therefore we remain unaware of what species, if any, other than *C. parvum* and *C. hominis* are currently in transmission in Ireland. The current onus on maximising efficiency and streamlining procedures in combination with a lack of evidence to indicate a more varied array of *Cryptosporidium* spp. in Ireland, Irish clinical laboratories are becoming increasingly reliant on commercial molecular panels for enteric parasite identification. However, an inherent caveat of employing these methods lies in the panels of pathogen targets developed by commercial companies. This limitation is apparent in the current study and will not be remedied until such time as molecule panels incorporate a broader array of *Cryptosporidium* spp. in their target repertoire, or clinical laboratories employ concurrent in-house molecular assays to detect rarer *Cryptosporidium* spp.
Regarding the subsequent OCP study, results such as these are not surprising within the context of enteric parasite epidemiology in Ireland. While Cryptosporidium spp. and G. lamblia predominate, cases of non-notifiable parasites such as Entamoeba spp., Blastocystis sp. and D. fragilis are under-reported and difficult to ascertain. Cryptosporidium spp. incidence markedly increase during early spring (Zintl et al., 2006). This accounts for the low number detected in this summer based study. Incidence of Giardia, the most commonly occurring enteric parasite encountered in this study, remains steady throughout the year (HPSC, 2015a, 2014).

Finding very low numbers of positive samples, as encountered in this study, presents us with a significant challenge when conducting essential validations for any potential addition to diagnostic methods. Parasitology is an integral part of the clinical diagnostic test regimen, which presents a dilemma when attempting continuation and maintenance of clinical expertise in this field.

This problem is further compounded by the lack of an overarching regulatory body to provide guidance and oversee the implementation of international standardisation of enteric parasitology diagnostics. While surveillance systems are active in countries such as the United States, Australia and among certain European countries, these systems are limited to reporting the incidence of enteric parasites, the incidence of which is frequently underestimated (Ryan et al., 2017). Although microscopy remains prevalent broad variation in diagnostic methods and the validation of such methods endures (Chalmers et al., 2010; Manser et al., 2014). However, given the burgeoning nature of molecular diagnostics in this field, the need for robust and reproducible validation methods is inevitable.

One potential solution may lie in the availability of commercially produced cysts and oocysts offer a potential alternative in situations where a paucity of positive samples impacts upon the validation of new diagnostic methods. With these products, it is possible to spike confirmed negative samples with known numbers of cysts/oocysts in order to supplement existing sample groups (Coupe et al., 2005; Robinson et al., 2008).

We also suggest that another potential solution to this dilemma might be the establishment of a pan-European Expert Working Group to consolidate cooperation and standardisation between laboratories, amass and disseminate sufficient, statistically robust laboratory validation materials, as well as educational samples and data, and to
aid in the preparation for planned prospective studies and to further stockpile reference specimens from multiple diagnostic laboratories.

The advantages of real-time PCR are apparent. Real-time PCR may be employed to screen all samples received for commonly encountered parasites during routine enteric pathogen investigation, regardless of acceptance criteria, while simultaneously improving workflow. In the future, it may be advisable to reserve microscopy for cases that warrant investigation when clinical indications are suggestive of enteric parasites not yet detectable by molecular methods.
Chapter III: A longitudinal, epidemiological study of *Cryptosporidium* species in patients presenting with gastroenteritis in Southern Ireland

A manuscript based upon this chapter has been published in the Journal of Clinical Pathology.

3.1 Abstract

_Cryptosporidium_ is a leading cause of gastrointestinal morbidity and mortality worldwide and one of the only enteric parasites to exhibit endemicity in Ireland. Despite the widespread implication of this protozoan parasite in both sporadic and outbreak-related gastrointestinal illness in Ireland, relatively little is known of the associated epidemiological trends and transmission pathways, with clinical epidemiological studies particularly lacking.

Herein, the retrospective, longitudinal study involving the molecular species identification of 163 clinical _Cryptosporidium_ isolates, amassed during routine molecular enteric screening in the Medical Microbiology Department of CUH from August 2015 to August 2018, is described. The isolates underwent species identification _via_ probe based real-time PCR, employing fluorescent probes specific to _C. parvum_ and _C. hominis_, respectively.

Overall, 86.5% (141/163) of clinical cases of cryptosporidiosis were attributed to _C. parvum_, while the remaining 13.5% (22/163) were caused by _C. hominis_. _Cryptosporidium_ infection was found to disproportionately affect patients 14 years of age or younger, with 64% of cases occurring in patients within this demographic. Bimodal seasonal distribution of the disease was observed, with a definitive spring time peak associated with the majority of cases detected annually. A second, autumnal peak of much lesser magnitude was also observed. Species predominance also exhibited seasonality with _C. parvum_ accounting for 100% of all cases detected during the spring. _C. hominis_ occurred predominantly during the late summer and autumn months.

This study, to the best of our knowledge, is the first employing fluorescent probe-based real-time PCR based species identification methodologies to be conducted in Ireland.
3.2 Introduction

Despite the global significance of Cryptosporidium as a major agent of gastrointestinal disease, a thorough understanding of Cryptosporidium epidemiology has only emerged since the advent of molecular techniques (Ryan et al., 2014). Microscopy as a diagnostic method is limited to identification to genus level, as it is not possible to differentiate between the vast majority of Cryptosporidium spp. on the basis of oocyst morphology (Thompson et al., 2016). Therefore, despite being the mainstay in Cryptosporidium diagnosis, microscopy, including bright-field, fluorescent, and immunofluorescent techniques, is largely redundant in detailed epidemiological studies (Thompson et al., 2016; Vejdani et al., 2014).

Prior to the development of molecular detection methods, the ambiguity associated with the speciation and taxonomic classification of Cryptosporidium species led to a large number of nomina nuda erroneously being assigned species status. Revision to the early system, wherein Cryptosporidium species assignment was inferred from host specificity, in favour of the current taxonomic nomenclature based on oocyst morphometric studies, inter-species genetic variation within the 18S rRNA gene, demonstration of host specificity, and compliance with International Code of Zoological Nomenclature (ICZN) guidelines, has produced a more robust taxonomic system (Ryan et al., 2014; Xiao et al., 2004). This system rectifies previously anomalous taxonomic designations and encompasses some 38 valid species, with the number of new species being reported increasing dramatically over the past decade (Feng et al., 2018). Of these 38 species, over 20 have been reported in human disease (Xiao and Feng, 2017). However, C. parvum and C. hominis account for over 90% of cases. Globally, some species including C. meleagrisdis, C. felis, C. canis, C. ubiquitum, C. cuniculus, C. viatorum, C. muris and chipmunk genotype I are commonly implicated in human infection, while species such as C. andersoni, C. suis, C. bovis, C. xiaoi, C. erinacei, C. fayeri, C. scrofarum, C. tyzzeri, horse, skunk and mink genotypes have been reported in fewer than five human cases each (Feng et al., 2018).

Initial molecular species identification methods employed the time consuming and technically laborious method, nested PCR-RFLP, targeting a variety of genes, including the COWP, gp60 and 18S rRNA genes, to yield species specific digestion patterns (Amar et al., 2004; Ruecker et al., 2011; Zavvar et al., 2008; Zintl et al., 2009). Species
identification has also been achieved through the amplification and subsequent sequencing of discrete regions within the abovementioned genes (Zintl et al., 2011). However, novel, streamlined applications of real-time PCR, employing fluorescent probe based technologies and post-PCR applications such as melt curve analysis, used in conjunction with DNA sequencing techniques, have emerged in recent years and are beginning to supersede more manually intensive molecular tools (Hadfield and Chalmers, 2012; Lalonde et al., 2013; Stroup et al., 2012, 2006).

Currently, commercial molecular platforms employed by clinical laboratories are often limited to identification to the genus level, restricted by the broad panels of enteric pathogens which are often targeted through multiplex reactions (Bouzid et al., 2013; Ryan et al., 2017). Such is the case with the clinical isolates identified during routine molecular testing in CUH, the largest university teaching hospital in Ireland and the participating clinical microbiology laboratory in this collaborative, epidemiological study (McAuliffe et al., 2017). Thus, although it is possible to ascertain the general incidence of Cryptosporidium within the community and catchment population of this national centre, the prevalence of individual species remains unknown.

Nationally, epidemiological data is largely confined to retrospective surveillance data accumulated through mandatory notifiable disease reporting (HPSC, 2018a). Similarly, these data are limited by the diagnostic methods employed by clinical laboratories, which do not generally identify the species or gp60-subtype of Cryptosporidium isolates. Consequently, detailed clinical epidemiological studies are lacking. This is particularly pertinent given that Cryptosporidium infection, particularly large-scale, waterborne infection outbreaks, exert a significant clinical and economic impact upon the Irish populace, with an average crude incidence rate (CIR) of 10.8 per 100,000 population reported in 2017 (HPSC, 2018a). 210 outbreaks have been reported since Cryptosporidium was declared a notifiable disease in Ireland in 2004. The 2007 outbreak affecting Galway city and the surrounding area alone is estimated to have cost the state in excess of €19 million (Chyzheuskaya et al., 2017).

The aim of the current study was to further elucidate C. parvum and C. hominis infections among Irish patients presenting with gastroenteritis, using a published fluorescent probe based real-time PCR method. The current epidemiological study was
the first in Ireland to employ real-time PCR based speciation methods, with this method developed by Mary et al. (2013).

3.3 Materials and Methods

3.3.1 Clinical Sample Acquisition

A sample cohort of 163 Cryptosporidium positive faecal samples was amassed, detected upon submission for routine molecular enteric screening to CUH, Ireland, from the centre’s regional service area, Cork City and surrounding county, between August 2015 and August 2018. As there are no mandated acceptance criteria for submitted samples of suspected cases of Cryptosporidium, acceptance criteria are generally established at the discretion of individual laboratories. The Medical Microbiology Department of CUH employed an acceptance criterion necessitating all submitted faecal samples to be designated as being type 5, type 6 or type 7 on the Bristol Stool Chart in order for enteric investigation of any nature to be conducted.

Initial clinical diagnoses were conducted via multiplex real-time PCR employing the CE-marked EntericBio GastroPanel II (Srosep, Limerick, Ireland), a combined platform capable of detecting a total of six bacterial and parasitic enteric pathogen targets, including Salmonella, Shigella, Campylobacter, VTEC, Cryptosporidium spp. and Giardia lamblia. All aspects of this assay were conducted according to manufacturer’s specifications, as previously described in Section 2.3 of Chapter II. Samples determined to be Cryptosporidium positive by real-time PCR underwent confirmatory microscopic testing via acid-fast mZN staining.

The resulting 163 Cryptosporidium positive samples encompassed almost all cases of cryptosporidiosis identified by the laboratory during this period, with 17 samples unavailable for further analysis.

3.3.2 Control Isolates

For control purposes, independently confirmed C. parvum and C. hominis genetic material, as outlined in Table 3.1, was provided by the CRU, (Swansea, Wales) for use in this study. These DNA samples were employed during duplex real-time PCR reactions to act as controls for the fluorophore specific detection and identification of Cryptosporidium spp. in clinical specimens.
In each run, *C. parvum* and *C. hominis* DNA was run separately to represent single infection cases, and mixed to act as control for co-infection detection and identification. *C. parvum* and *C. hominis* DNA, extracted from semi-purified oocysts and provided by the CRU, was mixed in a ratio of 1:1 for the purposes of providing a mixed infection control.

**Table 3.1** Control *Cryptosporidium* spp. used for the real-time PCR based identification in this study

<table>
<thead>
<tr>
<th>Cryptosporidium Species</th>
<th>Cryptosporidium gp60-subtype</th>
<th>Source</th>
<th>UK specimen number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. parvum</em></td>
<td>IlaA15G2R1</td>
<td>Human host</td>
<td>UKP111</td>
</tr>
<tr>
<td><em>C. hominis</em></td>
<td>IeA11G3T3</td>
<td>Human host</td>
<td>UKH50</td>
</tr>
</tbody>
</table>

The study also participated in the CRU’s external quality assessment and feedback scheme, a voluntary programme entailing the routine distribution of *Cryptosporidium* DNA samples to participants for species identification and genotyping. Samples were received at regular intervals and were subject to the same 18S rRNA species typing protocol conducted on clinical isolates. All EQA samples were analysed in duplicate.

### 3.3.3 DNA Extraction

DNA was extracted according to the EntericBio GastroPanel II one-step, heat treatment, extraction protocol as specified by the manufacturers. A 4 mL EntericBio Stool Preparation Solution (SPS) was inoculated with a FloqSwab (Copan, Italy) fully coated with stool sample and heated to 103°C for 30 minutes. Following heat treatment, samples were fully prepared for molecular testing. Samples were then stored at -20°C prior to further testing in this epidemiological study.

### 3.3.4 Primers and Probes

The species designations of all clinical samples were determined via duplex real-time PCR amplification of the 18S rRNA gene. Pan-*Cryptosporidium* specific forward and reverse primers were used, while hybridisation probes targeting a polymorphic region within the target amplicon differentiated between *C. parvum* and *C. hominis* (Table 3.2).
These fluorescent probes employed minor groove binding (MGB) technology and were synthesised by Eurofins Genomics (Ebersberg, Germany). The fluorescent 5’ reported of the *C. parvum* specific probe was altered from VIC as published in Mary *et al.* (2013), to HEX which is detected by the same excitation/emission filter (533-572 nm) of the LightCycler 96 (LC96) (Roche Molecular Diagnostics, Basel, Switzerland), which was used in the species identification of all of isolates in this study.

The limit of detection of the pan-*Cryptosporidium* primer pair was previously determined by Mary *et al.*, to be 300 oocysts per gram of stool (Mary *et al.*, 2013).

<table>
<thead>
<tr>
<th>Table 3.2</th>
<th>Primers and probes used for the real-time PCR based typing of <em>Cryptosporidium</em> samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locus</td>
<td>Sequence (5’ – 3’)</td>
</tr>
<tr>
<td>18S rRNA F</td>
<td>CATGGATAACCGTGTAAT</td>
</tr>
<tr>
<td>18S rRNA R</td>
<td>TACCCTACCGTCTAAAGCTG</td>
</tr>
<tr>
<td><em>C. parvum</em></td>
<td></td>
</tr>
<tr>
<td>18S rRNA</td>
<td>N/A</td>
</tr>
<tr>
<td><em>C. hominis</em></td>
<td></td>
</tr>
</tbody>
</table>

### 3.3.5 Molecular Analysis

In addition to the use of the LC96 System, the LC96 software (Roche Molecular Diagnostics, Basel, Switzerland) was used for amplification curve interpretation and crossing point (Cp) evaluation in all reactions. Real-time PCR amplifications were performed in 20 μl reaction volume and contained 4 μl LightCycler Multiplex DNA Master (Roche Molecular Diagnostics, Basel, Switzerland), constituting all necessary dNTPs and MgCl₂, while also containing qPCR reaction buffer, AptaTaq Polymerase and proprietary additives. Primers and probes were used at a concentration of 4 μM and 0.5 μM, respectively, and 5μl of template DNA was added to each reaction.

Reactions were conducted under the following cycling conditions: initial denaturation at 94°C for 10 min, subsequent 3-step amplification for 45 cycles, including denaturation at 94°C for 10s, annealing at 54 °C for 20s and extension at 72°C for 10s. All reactions were conducted in 96 well plates and overlaid with transparent sealing foils (Roche
Molecular Diagnostics, Basel, Switzerland) compatible with the excitation-emission detection format of the LC96. All reactions were conducted in duplicate.

After real-time PCR amplification, gel electrophoresis was conducted on the resultant amplicons. The 178 bp products were electrophoresed on 2.5% (w/v) agarose gels stained with SYBR Safe DNA gel stain (Biosciences, Ireland) for 45 minutes at 135 V and visualised via UV transilluminator using GelCapture software (DNR Bio-Imaging Systems, Israel). HyperLadder 25bp (Bioline, United States) was electrophoresed concurrently with all clinical sample amplicons for reference as a molecular marker.

Sample amplification curves and Cp values were generated and recorded by the LC96 software following real-time PCR amplification based on fluorophore fluorescence. *Cryptosporidium* species assignment was based on determination of the specific fluorophore detected by the fibre optic module of the LC96 for each sample.

### 3.3.6 Statistical Analysis

Statistical analyses based on temporal, age and sex specific differences within the amassed clinical *Cryptosporidium* isolates were conducted using SPSS software v. 25.0 (IBM, United Kingdom).

In order to ascertain whether or not *Cryptosporidium* infection affected patients equally, the binomial test was conducted on epidemiological data. A test proportion of 50% was employed. In this instance the null hypothesis stated that *Cryptosporidium* was equally likely to affect patients, regardless of gender.

The chi-square test goodness of fit test was employed to evaluate whether or not *Cryptosporidium* infection was equally distributed between different patient age groups. The age groups employed in this statistical analysis were as follows (and also represented in Figure 3.2): <5 years; 5 to 9 years; 10 to 14 years; 15 to 19 years; 20 to 24 years; 25 to 34 years, 35 to 44 years; 45 to 54 years; 45 to 55 years; 55 to 64 years and 65 to 74 years, 75 to 84 years and 85+ years. The chi-square test goodness of fit test was also employed to evaluate whether or not *Cryptosporidium* infection was equally distributed throughout the months of the year.

The chi-square test of independence was conducted in order to examine if an association existed between patient age and *Cryptosporidium* species. This test was also used to
investigate whether a statistically significant temporal association existed between infection incidence and time of year (seasonally and monthly). In these cases the null hypothesis stated that these variables were independent and no association existed between them.

Further statistical analyses were conducted to investigate whether or not a statistically significant association existed between patient age and gender, respectively, and infecting species of *Cryptosporidium*. The chi-square test of independence was utilised in these instances in order to ascertain this, with the null hypothesis asserting that no association existed between these variables. An independent *t*-test was conducted to investigate whether there was a statistically significant difference between Cp values observed for *C. parvum* and *C. hominis*, respectively.

Standard statistical analyses were also conducted in order to determine the mean, median, interquartile range (IQR) and skewness of patient age data. A confidence level of 95% (*α* ≤ 0.05) was employed in all statistical analyses.

### 3.4 Results

#### 3.4.1 18S rRNA Species Identification via Duplex Real-time PCR

Analysis of the 18S rRNA locus of the 163 samples analysed at determined that 141 (86.5%) were infected with *C. parvum*, while *C. hominis* infection accounted for 22 cases (13.5%). No co-infections were detected. The breakdown of species incidence by year is shown in Table 3.3. These incidence rates correlate with those previously reported in Ireland (Zintl et al., 2011, 2009)

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of <em>C. parvum</em> Cases</th>
<th><em>C. parvum</em> Incidence (%)</th>
<th>Number of <em>C. hominis</em> Cases</th>
<th><em>C. hominis</em> Incidence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>9</td>
<td>75%</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>2016</td>
<td>38</td>
<td>77.6%</td>
<td>11</td>
<td>22.4%</td>
</tr>
<tr>
<td>2017</td>
<td>47</td>
<td>94%</td>
<td>3</td>
<td>6%</td>
</tr>
<tr>
<td>2018</td>
<td>47</td>
<td>90.4%</td>
<td>5</td>
<td>9.6%</td>
</tr>
</tbody>
</table>

The amplification curve results seen in Figure 3.1 delineate the species identification of a representative sample set of 14 of the 163 clinical samples that were analysed via
probe based real-time PCR. Within this sample set, 7 *C. parvum* and 7 *C. hominis* isolates were characterised. Cp values were recorded for all real-time PCR amplifications and ranged from to 26-39 cycles. Variability in Cp values is likely to be due to variability in oocyst shedding in patients, which is cyclical in nature, or due to oocyst load of the initial inoculum, prior to DNA extraction.

Appendix I details further data pertaining to the species identity, real-time PCR amplification curves and gel electrophoresis result for each isolates analysed from 2015 to 2018.

**Figure 3.1** Fluorescent probe amplification curves of *Cryptosporidium* isolates. For clarity, 14 isolates, 7 *C. parvum* and 7 *C. hominis*, in addition to positive and negative controls, are shown here, but all 163 clinical isolates were typed via fluorescent probe based real-time PCR. *C. parvum* was detected with a Hex-labelled probe (dashed line curves), while *C. hominis* was detected with a FAM-labelled probe (solid line curves). All reactions were conducted in duplicate, with a single result shown for clarity.

### 3.4.2 Cryptosporidium Reference Unit EQA Participation

Over the course of the study, three batches of EQA sample material were disseminated by the CRU, with each batch consisting of 5 samples. Site participation in these EQA assessments resulted in 100% concordance with results determined by the CRU and compared with 7 other participating clinical laboratories, thus externally validating the specification system used in this study.

### 3.4.3 Epidemiological and Statistical Analyses

Data pertaining to 161 samples were statistically analysed, as patient gender and date of birth were not provided for two (1.2%) of the 163 collected samples. There was no
statistically significant difference between the Cp values observed for \textit{C. parvum} and \textit{C. hominis} \((p = .192)\).

The general age distribution of \textit{Cryptosporidium} infection in this study, as depicted in Figure 3.2, was largely skewed towards patients aged 14 years and under, with an inverse relationship between \textit{Cryptosporidium} incidence and increasing patient age. \textit{Cryptosporidium} infection was not equally distributed among age groups, with a statistically significant difference was found to exist between the actual distribution of cases observed and a theoretical equal distribution of cases \((x^2 =156.578, \text{ df} = 9, p = .000)\). The vast majority of cases, some 64\%, occurred in patients aged 14 years of age or younger, with 53\% of these cases occurring in children under 5 years. The average age of \textit{Cryptosporidium}-infected patients was determined to be 14.5 years and the median patient age as 9 years within an IQR of 3 to 24 years. The maximum patient age observed in the study was 70 years.

There was, however, no statistically significant association found between patient age and specific infecting \textit{Cryptosporidium} species \((x^2 =3.813, \text{ df} = 9, p = 0.923)\). Therefore, although patients under 14 years of age seem to be predisposed towards \textit{Cryptosporidium} infection generally, it seems that patients of any age may be infected by \textit{C. parvum} or \textit{C. hominis}, with no statistically significant age related disparity observed in the studied species.
Of the 163 analysed samples, 76 (46.6%) of specimens were from male patients and 85 (52.1%) specimens were from female patients. No statistically significant difference was identified between in incidence of Cryptosporidium infection in male and female patients (p = 0.529). Infection rates were noted to be slightly higher among females patients within the peak in infection observed in 20 to 34 year olds. Similarly, no statistically significant associations were found to exist between patient gender and specific infecting Cryptosporidium species ($x^2 = 2.421$, df = 1, p = 0.120) (Figure 3.3). Statistical analysis determined that gender does not contribute to an individual’s likelihood of being infected with Cryptosporidium, nor do the studied Cryptosporidium species disproportionately affect one gender over the other. Therefore, it can be deduced that persons are equally likely to be infected by C. parvum or C. hominis regardless of gender.
Figure 3.3  Distribution of Cryptosporidium infection by gender

The seasonal distribution of clinical cryptosporidiosis cases within this study was bimodal, exhibiting two seasonal peaks, with the data presented in Figure 3.4 depicting the average rates seen throughout the study in relation to both national and region (HSE South) incidence rates. The markedly more pronounced peak occurred from April to May and the second, less intense peak was observed during the late summer and early autumn months.

Cryptosporidium infection was not equally distributed throughout the months of the year, with a statistically significant difference found to exist between the actual distribution of observed cases and the theoretical equal distribution of cases ($x^2 = 138.03$, df = 11, $p = .000$), indicating that Cryptosporidium incidence was more prevalent during specific times of year. Similarly a strong statistically significant association was determined to exist between the incidence of specific Cryptosporidium spp and both monthly $x^2 = 65.443$, df = 11, $p = .000$) and seasonal distribution ($x^2 = 50.116$, df = 3, $p = .000$), respectively, and is reflective of the seasonality exhibited by C. parvum and C. hominis infections. C. parvum was the seasonally dominant species during spring months, accounting for 100% of all cases diagnosed from February to April, with 59.8% % of all C. parvum cases in this study occurring during this time. Meanwhile, incidence of C. hominis was confined to the late summer and autumnal
months (from August to October), with 77% of all *C. hominis* cases occurring during this time.

Examination of the impact of travel acquired *Cryptosporidium* infection on Irish incidence rates could not be conducted as travel history was disclosed on only 5 sample request forms (3.1%) of the total 163 samples tested. Although disclosure of patient clinical details and travel history is preferred, it is not mandatory for routine molecular enteric screening in CUH. In all four instances where foreign travel (to Turkey, Germany, and Portugal, and Dubai, respectively) immediately preceded symptom onset, infection was caused by *C. parvum* and all cases were reported from mid-summer to autumn (July to October). However, the lack of data pertaining to foreign travel among patients precludes the opportunity to derive meaningful insights into the relationship between travel acquired *Cryptosporidium* infections among Irish patients.

Despite the seasonal peaks associated with *Cryptosporidium* infection, it was observed that endemic rates outside of these periods remain relatively low.

![Figure 3.4](image)

**Figure 3.4** Average seasonal distribution of identified *C. parvum* and *C. hominis* cases within this study (2015-2018), and average national and regional (HSE South) *Cryptosporidium* cases reported for 2015, 2016, 2017 and 2018 (HPSC, 2018b, 2016b; HPSC, 2015b).
One case of Rotavirus coinfection, 2 cases of *Escherichia coli* coinfection and 6 cases of Verotoxigenic *E. coli* (VTEC) coinfection were observed during the study. Overall, 77.8% (7/9) of all coinfection cases occurred in patients under 14 years, with the 2 remaining coinfection cases (22.2%) occurring in 26 and 44 year old patients, respectively.

### 3.5 Discussion

The incidence rates observed in this study correlate with those previously reported in Ireland, with the national CIR of cryptosporidiosis increasing annually since 2014 (HPSC, 2019; Zintl *et al.*, 2011, 2009). General *Cryptosporidium* incidence in this study was also in keeping with regional incidence rates (HPSC, 2019). These figures are also in concordance, on a wider level, with previously published data pertaining to the epidemiological landscape in European countries such as Sweden, Denmark and France (ANOFEL Cryptosporidium National Network, 2010; Insulander *et al.*, 2013; Stensvold *et al.*, 2015).

*C. parvum* was the most prevalent species among the samples characterised in this study, accounting for of all cases 86.5%, with *C. hominis* accounting for the remaining 13.5%. This is also in concordance with annual epidemiological data published in Ireland and, on a wider level, similar to the epidemiological landscape in European countries such as Sweden, France, Denmark, Portugal and the U.K. and New Zealand (Al Mawly *et al.*, 2015; Alves *et al.*, 2003; Chalmers *et al.*, 2019; Guyot *et al.*, 2001; Insulander *et al.*, 2013; Learmonth *et al.*, 2004; Leoni *et al.*, 2007; McLauchlin *et al.*, 2000; Stensvold *et al.*, 2015). The prevalence of *C. parvum* is indicative of zoonotic transmission as this species is most commonly associated in infected livestock (Feng *et al.*, 2018; Santín, 2013).

No mixed infections were detected in this study. However, it has been reported that PCR is generally insensitive in the detection of minority alleles in the cases of mixed infections where oocysts concentration differs significantly between infecting species (Grinberg and Widmer, 2016). The mixed infection control (a 1:1 mix of *C. parvum* and *C. hominis* DNA) employed in this study may not have accurately represented all real world scenarios where oocyst concentrations can vary considerably between infecting species. Therefore, it is possible that cases of dual infection of *C. parvum* and *C. hominis* may have gone undetected if the oocyst concentration of one species was so
great as to mask the detection of the other. It is recommended that a more rigorous series of standards and ratios of mixed *C. parvum* and *C. hominis* template are employed in future studies of this nature to counteract this and provide a more robust and nuanced mixed infection control.

Additionally, although the primer pair employed in this assay in pan-*Cryptosporidium* spp. specific, the fluorescent probes employed limited detection to *C. parvum* and *C. hominis* specific targets within the primer amplified region. Similarly, the initial screening assay, the EntericBio Gastro Panel II, also detected *C. parvum* and *C. hominis* (but was incapable of differentiating between the two due to the use of a dsDNA binding fluorescent dye instead of species specific probes). As discussed in Chapter II, the lack of additional *Cryptosporidium* spp. targets offered by commercial molecular panels is a factor that prohibits the diagnosis of rarer species.

Therefore, although mixed infections have been previously reported in studies employing Taqman probe based real-time PCR (Peralta *et al.*, 2016), pan-specific probes would have been of little utility in the context of this study as the detection of rarer species by such a probe would have been masked by the confirmed *C. parvum/C. hominis* infections. Ultimately real-time PCR-based differentiation of all species of *Cryptosporidium* known to infect humans (of which approximately 20 have been reported) would have required an individual species-specific fluorophore bound probe for each. This was, unfortunately, not feasible in this study.

The masking of minority alleles associated with PCR has been identified as an issue common among epidemiological studies of *Cryptosporidium* that employ PCR-based methods, which has become a mainstay in the field (Grinberg and Widmer, 2016). The masking of minority alleles associated with PCR may also have also precluded sequence based identification of mixed infection (the pan-*Cryptosporidium* primer pair would have theoretically allowed amplification of all infecting species in the case of a mixed infection). Given the current methodological limitations, within this study and more broadly within the field of *Cryptosporidium* research, the extent of within-host *Cryptosporidium* diversity in nature remains poorly studied.

The age distribution among patients was also reflective of that seen on a European scale (European Centre for Disease Prevention and Control, 2019). Within this study, 64% of cases occurred in patients aged 14 years or younger, with 53% of such cases occurring
in children under 5 years. Similar to trends reported by ECDC data, a slight increase in infection rates was also observed to occur among 20 to 34 year olds, potentially through anthroponotic transmission via caregiver contact with infected children (European Centre for Disease Prevention and Control, 2019).

Cryptosporidiosis is widely regarded as a paediatric disease in Ireland and, indeed, worldwide, where it remains a leading cause of gastrointestinal-related morbidity and mortality in children under 5 years, particularly in low and middle income countries (Kotloff et al., 2019). However, the potential for age-related bias due to the self-selecting nature of sample submission precludes commentary on the incidence of Cryptosporidium species among particular age groups. Although the age profile of this study is concordant with national and European reports (Figure 3.3), it remains unclear whether this is reflective of actual infection incidence or due to under-ascertainment in older patients (Haagsma et al., 2013). Despite mandatory surveillance of this communicable disease, it is widely regarded that Cryptosporidium remains under-reported in Ireland and on a broader European level (Cacciò and Chalmers, 2016).

It remains unclear as to whether the patient age distribution pattern observed in this study, and frequently reported in similar Irish epidemiological reports, is truly reflective of national incidence. While acquired immunity may contribute somewhat to the markedly lower rates associated with increasing age, it is likely that cryptosporidiosis in Ireland is underreported in adult patients due to these individuals choosing to forego medical treatment, given the largely self-limiting nature of this gastrointestinal infection in immuno-comptenat adults (Hunter et al., 2004; Ryan et al., 2016; Scallan et al., 2004).

Studies have reported increased magnitudes of antibody response upon reinfection, with older individuals also possessing higher levels of antibodies (Allison et al., 2011; Priest et al., 2006). This may potentially contribute to some degree to the marked age distribution disparity commonly observed in reported Cryptosporidium cases (Moss et al., 1998; Priest et al., 2006). Overall however, the age distribution pattern observed in this study is concordant with that generally reported in cases of gastrointestinal illness, with the highest burden of illness occurring in patients less than 5 years of age, with a steady decline observed with increasing age (Adlam et al., 2011; Chen et al., 2013; Khalil et al., 2016; Scallan et al., 2005; Scavia et al., 2012; Thomas et al., 2006).
High incidence of Cryptosporidium infection in children is also thought to contribute to a secondary peak among 20 to 34 year olds, corresponding to infection transmission from children to parents. It has also been observed that within this age group, women experience higher levels of infection over their male counterparts (Cacciò and Chalmers, 2016; Chalmers et al., 2009). Both of these trends were observed in this study (Figure 3.3). Higher incidence rates in females have also been associated with acute gastrointestinal illness in general, being largely attributed to pathogen exposure through the child associated care-giving roles predominantly occupied by females (Thomas et al., 2006). An increased risk of contracting cryptosporidiosis has been associated with both symptomatic and asymptomatic or subclinical infection in young children, with asymptomatic carriage potentially acting as a reservoir of infection (Hunter et al., 2004). It is also important to acknowledge the impact of patient reporting behaviours on reported incidence rate. Occurrence of cryptosporidiosis (and gastrointestinal illnesses in general) is widely regarded to be underreported among adults due to the fact that immunocompetent adults generally less likely to seek medical intervention. In contrast, gastrointestinal illness in children, such as that caused by cryptosporidiosis, often results in medical treatment and specimen collection and testing (Garvey and McKeown, 2009). These behaviours may disproportionately skew incidence rates.

Oftentimes C. hominis, due to its association with travel related illness and greater pathogenicity, is more prevalent among adults (Hunter et al., 2004). However, C. hominis incidence was relatively consistent in the majority of patient age groups in this study, as shown in Figure 3.2, with no statistically significant association found to exist between the respective Cryptosporidium species examined in this study and specific patient age groupings.

In terms of seasonal distribution pattern, the bimodality of annual Cryptosporidium incidence, as observed in this study and depicted in Figure 3.5, is well established and reflective of seasonal patterns reported annually in Ireland by the HPSC, which has accrued epidemiological data since Cryptosporidium became a notifiable disease in Ireland in 2004 (HPSC, 2018a, 2017, 2016a, 2015a, 2014, 2013). This seasonal distribution mirrors that observed in the U.K.; which shares similar climatic and agricultural characteristics to those seen in Ireland. This pattern displays a spring time peak coincident with the agricultural practice of calving and lambing season, wherein
80% of all calves produced annually are born during the months of January to April (Berry et al., 2013). *C. parvum* typically predominates during this spring-tide upswing, with 59.8% of all *C. parvum* isolates detected in this study occurring during this period (Zintl et al., 2011, 2009).

It is also noteworthy that no cases of *C. hominis* were detected during the springtime in this study. This is intriguing as a previous Irish study identifying *C. hominis* among human faecal samples observed cases to occur throughout the year, but appearing to cluster during the springtime (although the number of cases were so low in the study that no definite trends could be identified) (Zintl et al., 2009). It may be that the results of this study are an accurate reflection of the incidence and seasonality of *C. parvum* and *C. hominis* during the study period. However, it may also be possible that co-infections of *C. hominis* during this period were masked by more prevalent *C. parvum* infections, remaining undetected by real-time PCR, as discussed previously.

This recurrent surge of infection is widely accepted to be caused by heightened environmental pollution and subsequent water supply contamination due to large numbers of diarrhoeic ruminants infected with *Cryptosporidium* during calving and lambing season. Given the high density of livestock in Ireland, particularly in western, rural areas, the country’s temperate oceanic climate with abundant rainfall and a public water supply obtained from surface water, zoonotic, waterborne transmission is likely to play a significant role during this time (Chalmers et al., 1997; Glaberman et al., 2002; Mahon and Doyle, 2017).

Also supportive of this is the fact that a marked disparity in the geographical distribution of livestock exists in Ireland, with livestock populations heavily concentrated in the west and decreasing eastwards and a concomitant, inverse human population distribution. The role of zoonotic, ruminant driven transmission is evident in national reports, which see the majority of *Cryptosporidium* cases occurring in the west, south-east and midlands and the lowest rates occurring in the east and north-east (HPSC, 2018a, 2017, 2016a)

The markedly smaller, late-summer and autumn occurring peak observed in this study is also typical of Irish seasonal distribution patterns, although significantly less pronounced than the analogous event seen in the U.K (National Health Service (NHS), 2019). In both countries this peak is often reported as travel-related and/or due to oocyst
exposure in recreational water, (Chalmers and Johnston, 2018; Chalmers et al., 2019; HPSC, 2018a, 2017, 2016a; Pedalino et al., 2003). In addition to foreign travel and recreational water based sources of infection, a second calving and lambing season during this period is also thought to be a contributing factor in the U.K. (Zintl et al., 2009). In this study C. hominis incidence was largely confined to this autumnal peak, with only a small number of sporadic C. hominis cases occurring during winter and spring months. Seasonal distribution in Ireland also varies from that reported more widely in Europe, where, although a bimodal distribution pattern is still observed, the autumnal peak far surpasses the springtime peak (European Centre for Disease Prevention and Control, 2018a, 2018b, 2017).

The vast majority of Cryptosporidium outbreaks in Ireland are familial outbreaks, which account for 77.1% of reported outbreaks since Cryptosporidium became a notifiable disease began in 2004 and enhanced cryptosporidiosis surveillance was introduced in 2010 (HPSC, 2018a). In outbreaks such as these, person-to-person transmission arising from contact with an infected family member is a commonly reported route of infection (HPSC, 2018a). Indicative of the role of anthroponotic transmission in Cryptosporidium infection was the patient history of four (2.6%) patients presenting with gastrointestinal symptoms in which contact with a family member known to be infected or previously infected with Cryptosporidium was cited upon sample submission. Globally, anthroponotic transmission of Cryptosporidium is also the dominant route of infection in lower income countries with poor sanitation infrastructures (King et al., 2019).

Direct contact with animal reservoirs is also a well-established transmission route implicated annually in both individual and outbreak infection scenarios. Information pertaining to contact with infected animals was disclosed in the patient history of three (1.8%) samples, highlighting the role of animal reservoirs in direct animal-to-human transmission arising from both agricultural practices and recreational farm visits. Zoonotic transmission via direct physical contact with infected animals was reported as the leading cause of Cryptosporidium outbreaks in Ireland in 2016 and is a prevalent cause of Cryptosporidium outbreaks annually (HPSC, 2018a, 2017, 2016a). It is zoonotic transmission, albeit largely through the contamination of water supplies by animal reservoirs which predominates in higher income countries such as Ireland, England, many mainland European countries, in addition to New Zealand (Efstratiou et al., 2017).
Both of these modes of transmission are also more prevalent among familial outbreaks, with general outbreaks more strongly linked to childcare facilities and contaminated drinking and recreational water and foreign travel (HPSC, 2018a, 2017, 2016a). However, as patient history pertaining to exposure to an infected family member or animal was limited to discrete, singular cases, conclusions pertaining to potential outbreaks arising from these events cannot be inferred from these data. Although disclosure of patient clinical details and travel history is preferred, it is not mandatory for routine molecular enteric screening in CUH. Consequently, considerable variation is observed in the level of detail disclosed in the patient history of faecal samples submitted for routine enteric screening.

In the context of foreign travel, a history of such was disclosed in only four (2.6%) individual cases, to Turkey, Germany, Portugal and Dubai, respectively. In all cases where foreign travel immediately preceded symptom onset, infection was caused by 

\textit{C. parvum} and all cases were reported from July to October, corresponding to reports in the literature that the latter annual peak is at least in part due to infection acquired abroad (Chalmers et al., 2019; Zintl et al., 2009). Again, as with the confirmed reports of direct person-to-person and animal-to-person transmission, the lack of data pertaining to foreign travel among patients precludes the opportunity to derive meaningful insights into the relationship between travel-acquired \textit{Cryptosporidium} infections among patients in Southern Ireland.

While \textit{C. parvum} remains the predominant infective species, particularly in sporadic cases, Irish outbreaks have been attributed to both \textit{C. parvum} and \textit{C. hominis}; with \textit{C. hominis} being the causative agent of one of the most significant outbreaks from a Public Health perspective which occurred in Galway in 2007 and affected approximately 120,000 people (Garvey and Mckeown, 2007). This, however, warrants further environmental and clinical study. It could be posited that Ireland’s geographical isolation from mainland Europe may have thus far conferred protection from the introduction of less frequently encountered \textit{Cryptosporidium} species predominantly acquired through direct contact with commonly domesticated host species, such as \textit{Cryptosporidium meleagridis} (turkeys/birds), \textit{Cryptosporidium ubiquitum} (ruminants/rodents/primates), \textit{Cryptosporidium cuniculus} (rabbits), \textit{Cryptosporidium andersoni} (cattle) and \textit{Cryptosporidium felis} (cats) (Cacciò and Chalmers, 2016; Ryan et al., 2014). A wider variety of \textit{Cryptosporidium} species are increasingly found to be
implicated in human infection in the UK, but also worldwide, transmitted zoonotically by wild and domesticated animal hosts (Cacciò and Chalmers, 2016; Ryan et al., 2014). These reports may indicate future epidemiological shifts in the diversity of *Cryptosporidium* species in Ireland, with the flux of humans and livestock between the UK and Europe having the potential to drive such a change.

Due to the environmental persistence of *Cryptosporidium* oocysts, particularly *C. parvum* oocysts, and subsequent endemicity of this enteric parasite among the livestock population of Ireland, in addition to issues with public water infrastructures, it is likely that *Cryptosporidium* is likely to remain a prominent public health issue in Ireland for the foreseeable future (Carey et al., 2004). Thus, in the context of continued annual outbreaks and the potential for future epidemiological shifts, it is advisable, if not essential that, in addition to rapid, sensitive, high-throughput molecular platforms, capable of detecting a wider range of *Cryptosporidium* spp. for first-line diagnosis, species and subtype (within species) discriminatory platforms are also introduced, in order to support the established national epidemiological frameworks for such a prominent Public Health concern in Ireland.
Chapter IV: Clinical and epidemiological significance of novel and established Cryptosporidium subtypes identified via gp60 sequencing among an Irish patient cohort

A manuscript based upon this chapter has been published in the Experimental Parasitology.

4.1 Abstract
Reported incidence rates of cryptosporidiosis in Ireland are consistently among the highest in Europe. Despite the national prevalence of this enteric parasite and the compulsory nature of incidence surveillance and reporting, in-depth analyses seeking to identify clinical isolates of Cryptosporidium on an intra-species level are rarely undertaken in Ireland. This molecular epidemiology study of 163 clinical Cryptosporidium isolates was conducted in Southern Ireland in order to ascertain population gp60-subtype heterogeneity. Analysis was conducted via real-time PCR amplification and gp60 gene sequencing, which successfully determined the gp60-subtype designation of 149 of the 163 (91.4%) tested isolates.

Overall, 12 C. parvum and five C. hominis gp60-subtypes were identified, with the incidence of the regionally predominant C. parvum species found to primarily occur during springtime months, while C. hominis incidence was largely confined to late summer and autumnal months. Additionally, three C. parvum and four C. hominis regionally novel gp60-subtypes were identified in this study. These gp60-subtypes had not yet been reported in human or livestock infection in Ireland. This study, to the best of our knowledge, is the first to report the implication of the C. parvum gp60-subtypes IIaA17G4R1, IIaA17G3R1, and IIaA15G1R1 and the C. hominis gp60-subtypes IaA14R3, IbA12G3, IeA11G3T3 and IfA12G1R5 in human cryptosporidiosis in Ireland.

These data give insight into the diversification of the Cryptosporidium population and emergent gp60-subtypes, while also allowing comparisons to be made with clinical epidemiological profiles reported previously in Ireland and elsewhere.
4.2 Introduction

Within the Cryptosporidium genus, and more specifically among the predominant human-pathogenic species, *C. parvum* and *C. hominis*, asexual and sexual life cycle stages, genetic recombination and selective pressures, such as parasite-host coevolution, host adaptation and geographic segregation, have led to generation of new genotype families and diverse genetic populations (Abal-Fabeiro *et al.*, 2013; Feng *et al.*, 2002; Garcia and Hayman, 2017). Ultimately, this has also led to the emergence of host-adapted, virulent and hyper-transmissible subtypes of considerable public health significance (Feng *et al.*, 2013).

Analyses of both *C. parvum* and *C. hominis* genomes, which were published in 2003, revealed both species to be completely syntenic, exhibiting approximately 95 to 97% sequence identity (Abrahamsen *et al.*, 2004; Isaza *et al.*, 2015; Widmer and Sullivan, 2012; Xu *et al.*, 2004). Genus-wide genetic diversity was observed to be lower in *Cryptosporidium* than that observed in other apicomplexan parasites, such as *Plasmodium* spp. and *Toxoplasma gondii*; with one genome-wide study indicating *Cryptosporidium* to have diverged from a common apicomplexan progenitor, undergoing large scale gene deletion and extensive cis-rearrangements to produce a considerably modified genome, containing a more limited repertoire of survival genes than those commonly observed among other genera within the apicomplexan phylum (Auburn and Barry, 2017; Mazurie *et al.*, 2013; Su *et al.*, 2012). Additionally, it was found that the protein- and RNA-coding genes of both species were ultimately identical (Widmer and Sullivan, 2012). Consequently, despite what is considered to be a relatively large sequence divergence between *C. parvum* and *C. hominis* (3 – 5%), *Cryptosporidium* exhibits significant structural and compositional conservation, with observed phenotypic differences, such as host range, largely attributed to differential gene regulation and to subtle sequence based variation, such as genetic polymorphisms (micro- and mini- satellite repeat lengths), among proteins that act at the host-parasite interface (Mazurie *et al.*, 2013; Widmer and Sullivan, 2012).

Due to the lack of morphological traits by which to differentiate *Cryptosporidium* spp., length polymorphisms have played a pivotal role in elucidating the taxonomy, epidemiology and population genetics of these and other genera of apicomplexan parasites (De Waele *et al.*, 2013; Sulaiman *et al.*, 2005; Widmer and Sullivan, 2012). Tandem repeats, such as mini- and microsatellite length polymorphisms, are a great
source of genetic variation in evolutionary adaptation, with tandem repeat loci exhibiting a wide range of allelic variables (Gemayel et al., 2010; Widmer, 2018). Unlike single nucleotide polymorphisms (SNPs) and other forms of DNA mutation, which have potentially infinite capacity to generate gross functional alterations and completely novel characteristics, tandem repeat induced variation is modular in nature, insofar as incremental changes in repeat numbers generally result in gradual alteration in gene function (Gemayel et al., 2010). This discrete genetic adaptation has also been documented in other genera of the apicomplexan phylum; while the results of comparative studies of *C. parvum* and *C. hominis* mini- and microsatellite regions are suggestive of the contribution of repeat length evolution at such loci to phenotypic divergences, such as host range, between these species (Ohashi et al., 2002; Tanrıverdi and Widmer, 2006).

Within eukaryotes, many tandem repeats are located within genes encoding surface proteins (Levdansky et al., 2007; Oh et al., 2005; Verstrepen et al., 2004; Zhang et al., 2003; Zhao et al., 2003). One such tandem repeat locus within the *Cryptosporidium* genome, the hyper-variable 60 kDa glycoprotein (gp60) gene, has become the primary marker by which genotypic characterisation of *Cryptosporidium* sample cohorts and populations is denoted and reported within the literature (Abe et al., 2006; Chalmers et al., 2015; de Lucio et al., 2016; Deshpande et al., 2015; Feng et al., 2013; Hijjawi et al., 2010b; Iqbal et al., 2015; Khalil et al., 2017; McKerr et al., 2015; Ng-Hublin et al., 2018; Sow et al., 2016; Xiao, 2010; Zintl et al., 2011, 2009). Found on the surface and apical regions of *Cryptosporidium* zoite stages, the gp60 gene encodes a single precursor protein (gp60) which is subject to proteolytic cleavage to yield two antigenic, mucin like glycoproteins, gp40 and gp15 (Cevallos et al., 2000a; O’Connor et al., 2007; Strong et al., 2000). These virulence factors are implicated in gliding motility, host cell adhesion and invasion, and also contribute to attenuation of the humoral immunity in humans, specifically neutralising antibody response (Bouzid et al., 2013; Cevallos et al., 2000b).

Now the current mainstay in *Cryptosporidium* spp. subtype characterisation, the gp60 nomenclature system was first described by Sulaiman et al. (2005) and further refined by Jex et al. (2007), based on research conducted by Strong et al., (2000), (Jex et al., 2007; Sulaiman et al., 2005). Within this nomenclatural system, species are assigned a species specific prefix (‘I’ denoting *C. hominis*, ‘II’ denoting *C. parvum* and ‘III’
denoting *C. meleagridis* etc.) and a *gp60*-subtype family designation (Ia, Ib, Id, Ie etc. for *C. hominis*; IIA, IIB, IIC, IID etc. for *C. parvum*, IIIA, IIIB, IIIC, IIID etc. for *C. meleagridis*) (Strong *et al.*, 2000).

*gp60*-subtypes are further differentiated on the basis of the number of trinucleotide repeat sequences (TCA, TCG and TCT), which encode the poly-serine tract of the gp40 glycoprotein, within the *gp60* microsatellite region (Sulaiman *et al.*, 2005; Jex *et al.*, 2007). Thus, the name IbA12G3T3 indicates that the parasite is of the *C. hominis* *gp60*-subtype family Ib, containing 12 copies of the TCA repeat, three copies of the TCG repeat and three copies of the TCT repeat in the trinucleotide repeat region of the gene. Oftentimes, additional repeat sequences, designated ‘R’ at the end of the *gp60*-subtype name, are also present in the gp40 region. For example, within the *C. parvum* IIA *gp60*-subtype family, the tandem repeat region of some *gp60*-subtypes is immediately followed by one or several copies of the ACATCA sequence. Considerable variation in downstream repeat sequences associated with a number specific *Cryptosporidium* *gp60*-subtype families have been described (Ryan *et al.*, 2014).

The identification of *Cryptosporidium* spp. at the species and genotype level is essential for the identification of infection sources and gauging the impact on public health in outbreak scenarios. This is particularly pertinent in Ireland, where an average of 15 outbreaks occurs annually (based on surveillance data, 2004-2017) (HPSC, 2019). However, from an epidemiological perspective, *Cryptosporidium* surveillance in Ireland is limited to genus level identification, with typing and subtyping reserved for outbreak analysis. In such situations standard protocol obliges Irish clinical microbiology laboratories to relay samples to the specialist CRU (Swansea, Wales) for typing and sub-genotyping (Mahon and Doyle, 2017). Resultantly, the epidemiological picture of transmission and population genetics of *Cryptosporidium* spp. in Ireland is considerably lacking, particularly in relation to clinical cohorts. Within this study, DNA sequence analysis of the *gp60* gene in the 163 clinical samples of known species designation was employed in order to conform to current conventions in *Cryptosporidium* epidemiological reporting. This permits meaningful comparison with published international epidemiological data.
4.3 Materials and Methods

4.3.1 Clinical Sample Collection

*Cryptosporidium* positive DNA isolates, detected during clinical testing of stool samples from patients presenting with symptoms of acute gastroenteritis, were obtained from the Medical Microbiology Department of CUH over the course of three years, from August 2015 to August 2018, inclusive. During this time 163 clinical samples, satisfying the acceptance criterion of being graded as type 5 – 7 on the Bristol Stool Form Chart (BSFC) (Chumpitazi *et al.*, 2016; Heaton *et al.*, 1992), were amassed.

4.3.2 Molecular Species Determination

The samples were subtyped during an initial epidemiological analysis using real-time PCR, using a pan-*Cryptosporidium* specific 18S rRNA gene targeting primer pair and minor groove binding fluorescent probes specific for *C. parvum* and *C. hominis*, respectively, as described by Mary *et al.*, 2013. For control purposes, verified *C. parvum* and *C. hominis* DNA samples were also provided by the CRU (Swansea, Wales).

4.3.3 Subtyping Primer Design

The tandem repeat region within the 60-kDa glycoprotein encoding gene fragment (gp60) was selected as the subtyping target in this study, as the current convention of *Cryptosporidium* subtyping nomenclature is based up this gene (Sulaiman *et al.*, 2005). As denoted in Table 4.1, separate primers pairs were designed in this study for *C. parvum* and *C. hominis* species, respectively. Sequences belonging to gp60-subtype families prevalent in Europe, specifically the *C. parvum* IIa–IIj families (GenBank accession numbers: AB242224-AB242229, AY382675, AY738185-AY738186, AY738188-AY7381889, AY873780-AY873782, AY738191, AY738195, DQ192502-DQ192508, DQ630514-DQ630515, DQ630517, DQ630519, DQ648531-DQ648537, EU140508), and the *C. hominis* Ia-Ig families (GenBank accession numbers: DQ192510, DQ665689, DQ665692, EF208067, EF576982, EU161648, EU161649, EU161651, EU161652, EU161655, FJ839873-FJ839875, FJ839878, FJ839881-FJ839883, FJ861217, JF681174, JF727781, JF727787, KR296811), were aligned in order to identify homologous regions using BLASTn software.

Table 4.1 Primers designed for gp60-subtyping of Cryptosporidium spp. in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer Sequence (5’ – 3’)</th>
<th>Product size</th>
<th>Annealing Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. parvum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPGP60OUT F</td>
<td>TCTCCGTATAGTCTCCGCTGT</td>
<td>455 bp</td>
<td>65°C</td>
</tr>
<tr>
<td>CPGP60OUT R</td>
<td>TGGGATCTGTTTGGTTTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. hominis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHGP60OUT F</td>
<td>CCACTCAGAGGCACCTTGAAA</td>
<td>964 bp</td>
<td>63°C</td>
</tr>
<tr>
<td>CHGP60OUT R</td>
<td>GACTAGCGGGAGTTTGCT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.3.4 Real-time PCR Amplification and Sequencing of the gp60 Tandem Repeat Region

DNA was extracted according to the EntericBio GastroPanel II (Serosep, Limerick, Ireland) one-step, heat treatment, extraction protocol as previously described.

Real-time PCR amplifications were performed in 20 μl reaction volume and contained 10 μl FastStart Essential DNA Green Master (Roche Molecular Diagnostics, Basel, Switzerland), constituting all necessary dNTPs and MgCl₂, while also containing reaction buffer, FastStart Taq polymerase and SYBR Green I dye. C. parvum and C. hominis primers were used at a concentration of 2 μM and 3 μM, respectively, and 5 μl of genomic template DNA was added to each reaction. Primer selection for each singleplex PCR reaction was based on the previously determined 18S rRNA species result of each sample.

All reactions were conducted on the LightCycler 96 (LC96) instrument (Roche Molecular Diagnostic, Basel, Switzerland). Reactions were conducted under the following cycling conditions: initial denaturation at 95°C for 10 min, subsequent 3-step amplification for 45 cycles, including denaturation at 95°C for 30s, annealing at the temperatures specified in Table 4.1 for 30s, and extension at 72°C for 40s. All reactions were conducted in 96 well plates and overlaid with transparent sealing foils (Roche
Molecular Diagnostics, Basel, Switzerland) compatible with the excitation-emission detection format of the LC96. All reactions were conducted in duplicate.

Sample amplification curves and Cp values were generated and recorded by the LC96 software following real-time PCR amplification based on fluorescence generated by the double stranded DNA binding fluorescent dye, SYBR green I.

The purity and concentration of the resultant DNA amplicons were determined using the NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Delaware, USA) and products were subsequently purified using the High Pure PCR product purification kit (Roche Molecular Diagnostics Basel, Switzerland) as per manufacturer’s instructions. Amplicons were sequenced bidirectionally via Sanger sequencing (Eurofins, Cologne, Germany).

### 4.3.5 Statistical and Phylogenetic Analysis

Statistical analyses based on temporal-, age- and sex-specific differences among the identified Cryptosporidium gp60-subtypes were conducted using SPSS software v. 25.0 (IBM, UK). A confidence level of 95% (α ≤ 0.05) was employed in all statistical analyses. The chi-square test of independence was conducted in order to examine if an association existed between patient age and gender, respectively, and infecting Cryptosporidium gp60-subtype. This test was also used to investigate whether a statistically significant temporal association existed between specific subtypic incidence and time of year (seasonally and monthly). In these cases the null hypothesis asserted that these variables were independent, with no association existing between them.

Phylogenetic analysis was conducted using Mega X: Molecular Evolutionary Genetics Analysis across computing platforms software (Kumar et al., 2018), with maximum likelihood analysis employed to construct evolutionary trees based on sequence data representative of all known gp60-subtypes implicated in clinical Cryptosporidium infection in Ireland and the UK. Evolutionary distance was calculated using the Kimura 2-parameter model, allowing for uniform distribution of variation in mutation rates between codons. Phylogenetic trees were evaluated by bootstrapping consisting of 1000 iterations. Bootstrap confidence levels greater than 70% were reported. The C. hominis gp60 subtype Ib (GenBank accession number JX294571) was used as the outgroup in the phylogenetic analysis of sequences data for C. parvum, while the C. parvum gp60
subtype IIa (GenBank accession number FJ825018) was used as the outgroup in the analysis of the sequence data for *C. hominis*.

### 4.4 Results

#### 4.4.1 Epidemiological Trends

A range of both *C. parvum* and *C. hominis* gp60-subtypes were characterised over the course of the study. Overall, *C. parvum* accounted for 129 (86.6%) of all cases of *Cryptosporidium*, while *C. hominis* constituted the remaining 20 (13.4%) of characterised samples. All cases of *C. parvum* infection were attributed to a single gp60-subtype family; family IIa, with 12 gp60-subtypes identified. Of these 12 gp60-subtypes, IIaA18G3R1 was the predominant gp60-subtype accounting for 59.7% of all cases. Other prevalent *C. parvum* gp60-subtypes included IIaA15G2R1, IIaA20G3R1 and IIaA17G2R1, which accounted for 7.4% (11 cases), 5.4% (8 cases), and 3.4% (5 cases) of all cryptosporidiosis cases, respectively. Incidence of gp60-subtypes IIaA21G3R1, IIaA19G4R1, IIaA19G3R1, IIaA17G1R1, and IIaA10G2R1 was rarer still, accounting for 11 samples (7.4%) in total. Two *C. parvum* cases (1.3%) were attributed to the more nationally infrequent subtypes, IIaA17G4R1 and IIaA17G3R1, respectively. The regionally novel gp60-subtype IIaA15G1R2 was also detected. The sequences of these isolates were submitted to GenBank under accession numbers MT053123 to MT053127, MT093986 to MT093990, and MT113471 to MT113474.

Annually, IIaA18G3R1 was the most prevalent gp60-subtype during each of the four years sampled, with slight inter-annual variation in the second most common IIa gp60-subtypes. The gp60-subtypes encountered also varied annually, with 2017 exhibiting the widest range of 9 confirmed gp60-subtypes and 2015 exhibiting the least, as shown in Figure 4.5.

Several *C. hominis* gp60-subtype families and individual gp60-subtype alleles were encountered over the course of the study, with the majority of samples belonging to the regionally prevalent Ib family. Of the two Ib gp60-subtypes detected, IbA10G2 was the predominant cause of *C. hominis* infection, accounting for 65% (13 cases) of all such infections. The IbA12G3 allele was the second most prevalent gp60-subtype, causing 20% (4 cases) of all *C. hominis* infections. However, cumulatively both of these gp60-subtypes accounted for only 11.3% of all cryptosporidiosis cases. Single cases involving
alleles from the Ia (IaA14R3), Ie (IeA11G3T3) and If (IfA12G1R5) subtype families, respectively, also occurred. However, the combined total of these gp60-subtypes contributed to only 15% of all *C. hominis* infections and 2% of cases of *Cryptosporidium* infection, overall.

Of the *C. hominis* families and gp60-subtypes detected, only one, gp60-subtype IbA10G2, had been previously reported in clinical infection in Ireland (Zintl *et al.*, 2009). The remaining subtypes have only been detected in studies of Irish livestock thus far, with the current study the first to report the implication of these gp60-subtype families in human infections.

The sequences of these clinically novel, Irish isolates were submitted to GenBank under accession numbers MT053128 – MT053134 and MT093991. Data pertaining to the identified *C. parvum* and *C. hominis* gp60-subtypes are summarised in Table 4.2.

The gp60-subtype designation could not be determined for 13 samples, or 8%, of the overall sample cohort due to the insufficient quality of sequence data obtained.
Table 4.2  
*C. parvum* and *C. hominis* gp60-subtypes identified in clinical samples in Southern Ireland from 2015 – 2018 and their previous reports in the literature

<table>
<thead>
<tr>
<th>gp60-subtype</th>
<th>Number of cases</th>
<th>Study prevalence N=149 n=%</th>
<th>Intra-species prevalence N=129 n=%</th>
<th>Literature reports of previous clinical cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIaA18G3R1</td>
<td>89</td>
<td>59.7%</td>
<td>69%</td>
<td>Reported clinically in Ireland (Zintl et al., 2011, 2009), Italy (Del chierico et al., 2011) Australia (Jex et al., 2008; J. S. Ng et al., 2012; O’Brien et al., 2008; Waldron et al., 2011), Canada (Ng-Hublin et al., 2018, 2017) Reported clinically in Ireland (Zintl et al., 2011, 2009), Wales (Chalmers et al., 2019; Chalmers et al., 2011c), Italy (Del chierico et al., 2011) Scotland (Deshpande et al., 2015), The Netherlands (Wielinga et al., 2008), France (Gargala et al., 2017), Spain (de Lucio et al., 2016), Portugal (Alves et al., 2006), Sweden (Insulander et al., 2013), Denmark (Stensvold et al., 2015), Belgium (Geurden et al., 2009), Lebanon (Osman et al., 2015), Jordan (Hijjawi et al., 2016) Australia (O’Brien et al., 2008; Waldron et al., 2011), Mexico (Valenzuela et al., 2014), Canada (Budu-Amoako et al., 2012; Iqbal et al., 2015; Ng-Hublin et al., 2018, 2017), United States (Blackburn et al., 2006; Feltus et al., 2006), Ethiopia (Adamu et al., 2010), Egypt (Helmy et al., 2013), Yemen (Alyousefi et al., 2013) and Malaysia (Iqbal et al., 2012) Reported clinically in Ireland (Mahon and Doyle, 2017; Zintl et al., 2011, 2009), England (Chalmers et al., 2011c), Australia (Jex et al., 2008; Waldron et al., 2011), Slovenia (Soba et al., 2006), Kuwait (Sulaiman et al., 2005) and Jordan (Hijjawi et al., 2010b) Reported clinically in Ireland (Zintl et al., 2011, 2009) England (Chalmers et al., 2011c), Australia (O’Brien et al., 2008; Waldron et al., 2011), Sweden (Insulander et al., 2013), Canada (Ng-Hublin et al., 2017) and United States (Blackburn et al., 2006; Feltus et al., 2006) Reported clinically in Ireland (Zintl et al., 2011, 2009) and Australia (Waldron et al., 2011) Reported clinically in Ireland (Zintl et al., 2011) Australia (Waldron et al., 2011) and Canada (Ng-Hublin et al., 2017) Reported clinically in Ireland (Zintl et al., 2011, 2009) England and Wales (Chalmers et al., 2019; Chalmers et al., 2011c), Scotland (Deshpande et al., 2015), The Netherlands (Wielinga et al., 2008) and Sweden (Insulander et al., 2013) Implication in human cryptosporidiosis has not been previously reported in Ireland. Reported in patients in Australia (Waldron et al., 2011) Implication in human cryptosporidiosis has not been previously reported in Ireland.</td>
</tr>
</tbody>
</table>
### IIaA19G3R1

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>0.67%</td>
</tr>
<tr>
<td>Australia</td>
<td>0.78%</td>
</tr>
</tbody>
</table>

**Reported clinically in**

**Ireland** (Glaberman et al., 2002; Zintl et al., 2011, 2009) England (Chalmers et al., 2011c) and Australia (O’Brien et al., 2008; Waldron et al., 2011)

**Implication in human cryptosporidiosis has not been previously reported in Ireland.**

Reported in patients in England and Wales (Chalmers et al., 2019), Scotland (Deshpande et al., 2015), Spain (Ramo et al., 2015), Australia (Waldron et al., 2011), Canada (Iqbal et al., 2015), Lebanon (Osman et al., 2015), Egypt (Helmy et al., 2013), Jordan (Hijjawi et al., 2010b) and Kuwait (Sulaiman et al., 2005)

### IIaA15G1R2

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>0.67%</td>
</tr>
<tr>
<td>Australia</td>
<td>0.78%</td>
</tr>
</tbody>
</table>

**Reported clinically in**

**Ireland** (Zintl et al., 2011, 2009)

**I.** mplication in human cryptosporidiosis has not been previously reported in Ireland.

Reported in patients in England and Wales (Chalmers et al., 2019), Scotland (Deshpande et al., 2015), Spain (Millán et al., 2019) and Mexico (Urrea-Quezada et al., 2018)

### IbA10G2

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>0.67%</td>
</tr>
<tr>
<td>Wales</td>
<td>0.78%</td>
</tr>
</tbody>
</table>

**Reported clinically in**

**Ireland** (Zintl et al., 2011, 2009)

**C. hominis**

### IbA12G3

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>0.67%</td>
</tr>
<tr>
<td>Spain</td>
<td>5.0%</td>
</tr>
</tbody>
</table>

**Implication in human cryptosporidiosis has not been previously reported in Ireland.**

Reported in patients from Scotland (Deshpande et al., 2015), Spain (Millán et al., 2019), Portugal (Alves et al., 2006), Canada (Ng-Hublin et al., 2018), Australia (Ng-Hublin et al., 2017; Waldron et al., 2011), Mexico (Valenzuela et al., 2014), Jordan (Hijjawi et al., 2010b), Cuba (Pelayo et al., 2008), Brazil and Argentina (Peralta et al., 2016), Peru (Cama et al., 2008), India (Khalil et al., 2017), Qatar (Boughattas et al., 2017), and Nigeria (Ukwah et al., 2017)

### IaA14R3

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>0.67%</td>
</tr>
<tr>
<td>Mexico</td>
<td>5.0%</td>
</tr>
</tbody>
</table>

**Implication in human cryptosporidiosis has not been previously reported in Ireland.**

Reported in patients from Scotland (Deshpande et al., 2015), Spain (de Lucio et al., 2016), Portugal (Alves et al., 2006), Canada (Ng-Hublin et al., 2018, 2017), Mexico (Urrea-Quezada et al., 2018; Valenzuela et al., 2014) and Nigeria (Efunshile et al., 2019)

### IeA11G3T3

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>0.67%</td>
</tr>
<tr>
<td>Mexico</td>
<td>5.0%</td>
</tr>
</tbody>
</table>

**Implication in human cryptosporidiosis has not been previously reported in Ireland.**

Reported in patients from Scotland (Deshpande et al., 2015), Spain (de Lucio et al., 2016), Portugal (Alves et al., 2006), Canada (Ng-Hublin et al., 2018, 2017), Mexico (Urrea-Quezada et al., 2018; Valenzuela et al., 2014), Brazil (Peralta et al., 2016), Peru (Cama et al., 2008), India (Khalil et al., 2017), China (Feng and Xiao, 2017), Thailand (Sannella et al., 2019), Kenya (Mbæ et al., 2015), Ghana (Eibach et al., 2015), Nigeria (Ukwah et al., 2017), and linked to an outbreak in the United States (Fill et al., 2017)

### IfA12G1R5

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>0.67%</td>
</tr>
<tr>
<td>Mexico</td>
<td>5.0%</td>
</tr>
</tbody>
</table>

**Implication in human cryptosporidiosis has not been previously reported in Ireland.**

Reported in patients in Thailand (Sannella et al., 2019) and Kenya (Mbæ et al., 2015)
4.4.2 Allocation of Cryptosporidium to Genotypes and gp60-subtypes based on Phylogenetic Analysis

There is currently no consensus on the exact definition of ‘‘species/genotype’’ within the Cryptosporidium genus, with threshold levels for the differentiation of genotypes from gp60-subtypes also ambiguous (Jex et al., 2008). Herein, a sample was assigned to a particular genotype based on its relationship to a strongly supported clade, following phylogenetic analysis of the gp60 sequence data derived from all study samples, in addition to published sequences representing all currently recognized C. hominis and C. parvum “species/genotypes” to have been reported within Ireland and the U.K.

The 149 clinical samples that yielded sequence data were assigned specific gp60-subtypic identities based on the aforementioned nomenclature system. In order to provide unequivocal support to these preliminary designations, the phylogenetic relationships between study gp60 sequences and published homologous sequences, obtained from GenBank, were established by maximum likelihood.

The sequence data set for each species was subjected to separate phylogenetic analyses, as significant sequence variability within the gp60 locus between C. parvum and C. hominis precluded positional homology during sequence alignment. Phylogenetic analyses of the gp60 locus in C. parvum and C. hominis included 61 and 28 reference sequences, respectively.

Phylogenetic analysis of C. parvum resulted in all genotypes forming monophyletic clades with strong nodal support (maximum likelihood bootstrap proportions [BP] = 98 – 99%), within which several multi-genotypic clades were formed (Figure 4.1). Clade groupings conformed with previously established and published C. parvum phylogenetic analyses.
Figure 4.1  Evolutionary relationships between clinically encountered *C. parvum* gp60-subtypes in this study and *C. parvum* gp60-subtypes reported in Ireland and the U.K. Inferred through Maximum Likelihood analysis of partial gp60 gene sequences. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 iterations) is indicated next to the branches. The evolutionary distances were computed using the Kimura-2 method. The gp60-subtypes identified in this study and submitted to GenBank are denoted by ♦.
Phylogenetic analysis of *C. hominis* also resulted in all genotypes grouping within monophyletic clades with good nodal support (BP = 72-100%) (Figure 4.2). Clade groupings conformed with those previously established and published *C. hominis* phylogenetic analyses. For all *C. parvum* and *C. hominis* isolates encountered in this study, phylogenetic placement of *gp60*-subtypes was within genotypic clades concordant with initial species classification.
Figure 4.2 Evolutionary relationships between clinically encountered *C. hominis* gp60-subtypes in this study and *C. hominis* gp60-subtypes reported in Ireland and the U.K. Inferred through Maximum Likelihood analysis of partial gp60 gene sequences. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 iterations) is indicated next to the branches. The evolutionary distances were computed using the Kimura-2 method. The gp60-subtypes identified in this study and submitted to GenBank are denoted by ♦.
4.4.3  **Seasonal Distribution of gp60-subtypes**

The seasonal distribution of clinical cryptosporidiosis cases within this study was bimodal, exhibiting two seasonal peaks (Figure 4.3). The markedly more pronounced peak occurred from April to May, with the second, less intense peak being observed during the late summer and early autumn months (August to October).

Seasonal shifts in *Cryptosporidium* gp60-subtypes were found to be statistically significant ($\chi^2 = 83.156$, df = 48, $p = 0.001$). Prevalent *C. parvum* gp60-subtypes such as IlaA18G3R1, IlaA15G2R1, IlaA20G3R1, IlaA17G2R1 and IlaA17G3R1 occurred throughout the year, predominantly during the spring and summer, with such cases abating during autumn and winter months. More unusual gp60-subtypes, such as IlaA21G3R1 and IlaA19G3R1, were detected only during spring, while IlaA10G2R1 and IlaA17G4R1 cases were exclusive to summer. Occurrence of *C. hominis* gp60-subtypes were confined predominantly to autumn and winter, with a small number of...
sporadic IbA10G2 cases also occurring during the spring and summer months (Figure 4.4).

Prevalence of the most common Cryptosporidium gp60-subtypes remained relatively consistent, occurring throughout all four years of the study, while rarer gp60-subtypes were more sporadic in occurrence (Figure 4.5).

The C. parvum gp60-subtype, IlaA18G3R1, was the predominant gp60-subtype during every year of the study, constituting approximately 60% of detected cases annually, with the exception of 2015, where it constituted only 40% of samples. However, this is likely to be due to study sampling beginning during August of that year. IlaA15G2R1 and IlaA20G3R1 occurred annually, with the exception of no cases of IlaA15G2R1 during 2015, exhibiting relatively consistent annual rates of approximately 8% and 6%, respectively, aside from a decrease in IlaA20G3R1 cases to 2% in 2017. However, less commonly encountered gp60-subtypes such as IlaA21G3R1, IlaA19G3R1 and IlaA19G4R1, among others, were encountered intermittently, and in some cases,
incidence was confined to small numbers or single cases occurring during a specific year.

Regarding *C. hominis*, IbA10G2 was the only gp60-subtype to arise every year, accounting for approximately 9% of all cases each year. All other *C. hominis* gp60-subtypes occurred during single, disparate years throughout the study.

Overall, 2017 exhibited the most diversity in *Cryptosporidium* gp60-subtype occurrence. However, annual variations in *Cryptosporidium* gp60-subtype incidence were not found to be statistically significant ($\chi^2 = 63.206$, df = 48, $p = 0.70$).

![Figure 4.5](image)

**Figure 4.5** Annual distribution of *C. parvum* and *C. hominis* gp60-subtypes, 2015-2018
4.5 Discussion

Overall, a relatively diverse range of *C. parvum* and *C. hominis* gp60-subtypes were recorded in this study, with 12 and five gp60-subtypes described, respectively. All *C. parvum* gp60-characterised in this study belonged to the gp60-subtype family IIa. Of these, three *C. parvum* gp60-subtypes, IIaA15G1R2, IIaA17G3R1 and IIaA17G4R1 have not been previously described in humans in Ireland, while all but one of the four *C. hominis* gp60-subtypes characterised, IbA10G2, are newly described here in Irish clinical cases of cryptosporidiosis. These gp60-subtypes encompassed a range of *C. hominis* gp60-subtype families, including Ia, Ib, Ie and If. In the present study, all characterised Irish *C. parvum* and *C. hominis* gp60-subtypes formed strongly supported clades consistent with their previous specific identification via other phylogenetic algorithms, as reported in the literature (Jex et al., 2008).

Seasonal distribution of *Cryptosporidium* species was also reflective of previous epidemiological studies conducted in Ireland (Zintl et al., 2011, 2009). *C. parvum* gp60-subtypes were largely confined to the spring time peak, coinciding with the agricultural practices of calving and lambing season, with the exception of IIaA10G2R1 and IIaA17G4R1. There is no previously reported Irish incidence of the seasonal segregation exhibited by these two gp60-subtypes in this study. No cases of *C. hominis* were observed during this period, instead such cases peaked to a lesser degree from late summer through autumn, as previously reported in various Irish and UK studies (Chalmers et al., 2011b; Zintl et al., 2009), maintaining a smaller number of cases throughout the winter months.

The elucidation of IIaA18G3R1 as the predominant *C. parvum* gp60-subtype among those identified in this study is in concordance with the results of previous epidemiological analyses that have also reported IIaA18G3R1 as the predominant gp60-subtype implicated in sporadic cases in human and livestock infection on the island of Ireland (Glaberman et al., 2002; Thompson et al., 2006; Zintl et al., 2011). There is experimental evidence to suggest that the gp60 gene may be subject to host based selective pressures, given the immunogenic nature of the surface glycoproteins encoded by this gene, which play a role in host cell adhesion and invasion (Widmer, 2009). This may have potentially given rise to the extensive polymorphism observed in this gene. Thus, it is reasonable to expect host selective pressures to drive the emergence of new
predominant *gp60*-subtypes over time, as immunity to predominant circulating *gp60*-subtypes develops among the host population.

However, this does not seem to be the case; a 2011 study conducted in Ireland over a 10 year period, from 2000 – 2009, found IIaA18G3R1 to be the most predominant allele (Zintl *et al.*, 2011). The results of the current study suggest that IIaA18G3R1 remains the predominant *gp60*-subtype, which would indicate predominance spanning almost two decades. Consequently, further research is required to delineate the exact relationship between the selective pressures, genetic recombination and evolutionary processes that drive *gp60* gene polymorphism among *Cryptosporidium* species (Abal-Fabeiro *et al.*, 2013; Guo *et al.*, 2015; Garcia-R and Hayman, 2017).

After IIaA18G3R1, IIaA15G2R1 was the second most prevalent *C. parvum* *gp60*-subtype in this study. Although superseded by IIaA18G3R1 in Ireland, Australia and New Zealand (Garcia-R *et al.*, 2020; Waldron *et al.*, 2011), this hyper-transmissible subtype is the most widely distributed and commonly reported *gp60*-subtype in livestock and humans worldwide, including the neighbouring U.K., although studies have yet to elucidate the reasons contributing to its hyper-transmissibility (Brook *et al.*, 2009; Feng *et al.*, 2013; Geurden *et al.*, 2008; Wielinga *et al.*, 2008; Xiao *et al.*, 2007). The species prevalence rates of this study are also reflective of the those seen in other epidemiological studies conducted in humans in Ireland, with similar rates of *C. parvum* also reported among Irish livestock (Thompson *et al.*, 2006; Zintl *et al.*, 2011, 2009). Other *gp60*-subtypes, IIaA20G3R1, IIaA17G2R1 IIaA21G3R1 IIaA19G4R1, IIaA17G1R1 IIaA19G3R1 and IIaA10G2R1, have all been implicated in Irish cases of human cryptosporidiosis in previous studies (Glaberman *et al.*, 2002; Zintl *et al.*, 2011, 2009) and have also been implicated in infection of both humans and livestock internationally (see Table 4.2).

Additionally the *gp60*-subtypes IIaA17G4R1, IIaA17G3R1, and IIaA15G1R2 have not been previously described in human infection in Ireland. Although newly reported in Ireland, these *gp60*-subtypes have also been implicated internationally in both human and livestock infection (Avendaño *et al.*, 2018; Baroudi *et al.*, 2017; Holzhausen *et al.*, 2019; Mammeri *et al.*, 2019; Mercado *et al.*, 2015). However, foreign travel was disclosed in only four cases of *C. parvum* infection; a case involving IIaA17G4R1 wherein the patient had recently returned from Turkey, a case involving IIaA17G2R1 in
a German individual, and a case involving IlaA15G2R1 wherein the patient had recently returned from Portugal. The \textit{gp60}-subtype could not be determined in one case involving travel to Dubai. However, it is difficult to postulate with any certainty the means by which these \textit{gp60}-subtypes have been introduced among the Irish populace, although foreign travel may likely contribute.

Hereofore, IbA10G2 has been the only \textit{C. hominis} \textit{gp60}-subtype to have been described in Ireland. This subtype also accounts for almost all autochthonous \textit{C. hominis} infections in Europe, and is well distributed in both industrialised and developing nations worldwide (Feng \textit{et al}., 2018). It has also been proposed to be more virulent and cause long-term post-symptoms than other \textit{C. hominis} \textit{gp60}-subtypes (Cama \textit{et al}., 2008; Li \textit{et al}., 2013; Cacciò and Chalmers, 2016). Despite the homogeneity of Irish incidences of \textit{C. hominis} in previous reports, in recent years it has been posited that the diversification of the Irish \textit{C. hominis} population in Ireland was highly probable, due to the introduction of a more heterogeneous population through the transmission of travel acquired \textit{gp60}-subtypes (Zintl \textit{et al}., 2009). As is evident in this study, the incidence of IaA14R3, IbA12G3, IfA12G1R5 and IeA11G3T3 may be reflective of the initiation of the process of genotypic diversification \textit{via} anthroponotic transmission of \textit{gp60}-subtypes acquired abroad, or at the very least they represent the potential for such an occurrence, given their disparate temporal distribution.

No details of foreign travel were included in the clinical history for these isolates, so it is possible that infections with these \textit{gp60}- subtypes as the causative agent were acquired locally, having been introduced to the Irish \textit{Cryptosporidium} population from abroad. Reports of diverse, heterogeneous \textit{C. hominis} populations, including the regionally novel \textit{gp60}-subtypes to have occurred in this study, are relatively commonplace internationally, in countries such as Portugal, Canada, Australia, Mexico, China, Nigeria and India, while a more varied array of \textit{gp60}-subtypes has also been reported in human outbreaks in the U.K. in recent years (Chalmers \textit{et al}., 2019; Feng and Xiao, 2017; Khalil \textit{et al}., 2017; Ng-Hublin \textit{et al}., 2017; O’Brien \textit{et al}., 2008; Valenzuela \textit{et al}., 2014).

The detection of new \textit{C. hominis} \textit{gp60}-subtypes may also be indicative of the possibility that anthroponotic routes may be more significant in \textit{Cryptosporidium} transmission within Ireland than previously thought, as zoonotic transmission has long thought to
have been the primary route through which *Cryptosporidium* infection is contracted. These results correlate with previous findings that suggested the gp60 tandem repeat pattern exhibited by Irish *C. parvum* gp60-subtypes was reflective of those observed in regions where anthroponic transmission surpassed zoonotic transmission (Zintl et al., 2011). Possible cases on intra-familial anthroponic transmission were identified in this study, with the clinical history of one patient infected with *C. hominis* IbA12G3 reporting them to have been in recent contact with an infected family member. Clinical histories pertaining to three cases of *C. parvum* IlaA28G3R1 were also indicative of suspected cases of anthroponic transmission, also citing recent contact with a family member previously known to be infected with *Cryptosporidium*.

To conclude, *Cryptosporidium* remains an ever present agent of gastrointestinal illness in Ireland. While IlaA18G3R1 and IbA10G2 remain the most prevalent *C. parvum* and *C. hominis* gp60-subtypes, respectively, the previous genetic homogeneity, particularly within the *C. hominis* population, which was perhaps a result of Ireland’s geographical isolation, has given way to an ever increasing genetic diversity within these species, when compared to previous studies (Mahon and Doyle, 2017; Thompson et al., 2006; Zintl et al., 2011, 2009). This burgeoning diversification and introduction of novel *Cryptosporidium* gp60-subtypes, as identified in the present study, is likely to be indicative of a complex dynamic involving both zoonotic and anthroponic transmission routes and vector-host interactions. Comprehensive re-evaluation of our previous understanding of *Cryptosporidium* transmission routes, host related selective pressures and immunity and their impact on genotypic prevalence in Ireland is required, prompting further study in this field in order to elucidate these interrelated concepts.
Chapter V: Development of a novel high resolution melting analysis-based genotyping method for *Cryptosporidium parvum*

Manuscripts based upon this chapter have been published in MethodsX and The European Journal of Protistology, respectively.


5.1 Abstract

This study employed the post-real-time PCR application, high resolution melting (HRM) analysis, in order to differentiate between characterised clinical and reference *Cryptosporidium parvum* samples obtained from CUH (Cork, Ireland) and the CRU (Swansea, Wales). A sample set composed of 18 distinct *C. parvum* gp60-subtype alleles from the IIa gp60-subtype family (the subtype family accounting for over 80% of all cryptosporidiosis cases in Ireland) was employed. HRM analysis-based interrogation of the *gp60*, MM5 and MS9-Mallon tandem repeat loci was found to completely differentiate between 10 of the 18 studied *gp60*-subtypes. The remaining eight isolates were differentiated into three distinct groupings, with each grouping containing two to three *gp60*-subtypes. All of the *gp60*-subtype isolates within these groupings exhibited indistinguishable amplicon melting characteristics at each of the tandem repeat loci regions studied.

The current study aimed to develop a novel, reproducible, real-time PCR-based multi-locus genotyping method to distinguish between *C. parvum* gp60-subtypes. These preliminary results support the further expansion of the multi-locus panel in order to increase the discriminatory capabilities of this novel method.
5.2 Introduction

Globally, the apicomplexan parasite, *Cryptosporidium*, remains an under-recognised, yet ubiquitous, cause of gastrointestinal illness in both developed and developing countries (Troeger et al., 2017). The *Cryptosporidium* genus encompasses approximately 40 confirmed species, affecting a multitude of mammalian, avian, reptilian, amphibian and piscine hosts (Graczyk and Cranfield, 2000; Nakamura and Meireles, 2015; Ryan, 2010; Ryan et al., 2015, 2014). Among these, *Cryptosporidium parvum* exhibits the broadest host range, infecting a wide array of mammals, but primarily affecting humans and ruminants (Feng et al., 2018). *C. parvum* and *Cryptosporidium hominis* account for >90% of all clinical cases of cryptosporidiosis, with 10% attributable to remaining species (Chalmers et al., 2009). Children, particularly those under the age of five, immunocompromised and immunosuppressed individuals are particularly vulnerable to illness (Florescu and Sandkovsky, 2016; Khalil et al., 2018; Vinayak et al., 2015; Wang et al., 2018; Wolska-Kusnierz et al., 2007). Infection is generally contracted *via* faecal-oral transmission; usually through the consumption of oocyst contaminated food or water (Bouzid et al., 2013).

This protozoan parasite remains a global Public Health concern, having been established as a leading cause of diarrhoea-associated death in young children in the many developing countries to which the protozoan parasite is endemic (Sow et al., 2016). It is also a frequent cause of waterborne, and oftentimes foodborne, outbreaks of infection in many industrialised nations (Efstratiou et al., 2017).

Regional variations in species prevalence are well established, with *C. parvum* predominating in rural areas of New Zealand, much of Europe, such as Ireland and the UK, and Middle-Eastern nations (Areeshi et al., 2007; Iqbal et al., 2011; Learmonth et al., 2001; Meamar et al., 2007); *C. hominis* accounts for the vast majority of human infections in North and South America, Australia, Japan and Africa (Ajjojampur et al., 2007; Cacciò et al., 2005; Gatei et al., 2003; Hira et al., 2011; J. Ng et al., 2010; Pelayo et al., 2008; Sharma et al., 2013; Tiangtip and Jongwutiwes, 2002; Yagita et al., 2001). However, regional variations in *gp60*-subtype prevalence within species are more complex and nuanced; panmixia has been reported to occur in *Cryptosporidium* spp. populations, particularly within *C. parvum* (Feng et al., 2002). The discovery of the capacity for genetic recombination among genetically distinct *Cryptosporidium* genotypes during the gametogenic phase of the life-cycle has resulted in a recent
paradigm shift in current Cryptosporidium spp. genotyping methodologies. Previously, interrogation of a single-locus, most commonly the highly polymorphic 60 kDa glycoprotein gene (gp60), which forms the basis for contemporary Cryptosporidium species identification and subtyping taxonomy (Chalmers et al., 2017; Strong et al., 2000; Sulaiman et al., 2005), was sufficient to characterise Cryptosporidium spp. gp60-subtypes. However, the improved discriminatory capabilities of MLG methods, such as multi-locus sequence typing (MLST) of loci containing micro- and minisatellite repeats (including those that contain variable-number tandem repeats [VNTR]), have resulted in their superseding single-locus genotyping (Chalmers et al., 2018; Tanriverdi et al., 2008). Targeting multiple loci allows for a more comprehensive understanding of regionally autochthonous and allochthonous species and gp60-subtypes, yields insight into population structures and transmission dynamics, as well as aiding in identification of sources of contamination and routes of transmission in outbreak scenarios (Chalmers et al., 2017; Feng et al., 2018).

Particularly pertinent to this is the successful application of high resolution melting (HRM) analysis, a highly sensitive, closed-tube, post-real-time PCR application (Gundry et al., 2003; Wittwer et al., 2003), to the differentiation between a range of Apicomplexan parasites. HRM analysis has been employed to differentiate between Toxoplasma gondii, Sarcocystis spp., Neospora spp. and Cryptosporidium spp., to differentiate between C. parvum and C. hominis, and for intra-species differentiation between two Cryptosporidium meleagridis gp60-subtypes, respectively (Chelbi et al., 2018; Fehlberg et al., 2017; Pangasa et al., 2009). HRM analysis harnesses the dissociation of double stranded DNA with increasing temperature and the resultant decrease in fluorescence, due to the concomitant dissociation of saturating, double-stranded DNA binding fluorophores from denatured DNA strands, to determine amplicon melting temperature (Tm) (Tong and Giffard, 2012).

Herein, we discuss the first attempt at application of HRM analysis of multiple tandem repeat loci to the allelic differentiation of 18 distinct gp60-subtype family IIa isolates of the C. parvum species, which are among the most prevalent in Europe (Cacciò and Chalmers, 2016), in a way that is analogous to the differentiation of C. parvum gp60-subtypes via DNA sequencing of these loci in MLST analyses. The gp60 gene, in addition to the genetic markers, MM5 and MS9-Mallon, both of which also exhibit
VNTR motif regions and have been used in many multi-locus analyses of *C. parvum*, were selected for inclusion in this analysis (Mallon *et al.*, 2003; Mallon *et al.*, 2003; Morrison *et al.*, 2008; Sulaiman *et al.*, 2001; Tanriverdi *et al.*, 2006; Tanriverdi and Widmer, 2006). The purpose of the current study was to develop a method capable of differentiating between *C. parvum* gp60-subtypes via HRM analysis of the tandem repeat loci, with the intention of providing a rapid, cost-effective alternative to multiple locus DNA sequencing with which to conduct detailed, preliminary epidemiological analyses of clinical cryptosporidiosis cases.

5.3 **Materials and Methods**

5.3.1 **Clinical Sample Acquisition**

*Cryptosporidium* positive DNA isolates, detected during clinical testing of stool samples from patients presenting with symptoms of acute gastroenteritis, were obtained from the Medical Microbiology Department of CUH over the course of three years, from August 2015 to August 2018 inclusive, and provided the basis for this study. The 149 amassed isolates were identified to the species and gp60-subtype level via fluorescent probe-based real-time PCR and gp60 gene sequencing, respectively, as described in Chapters III and IV. Isolates representing each of the 12 gp60-subtypes detected within this isolate bank were selected for inclusion in the current study.

For control purposes, six confirmed *C. parvum* DNA samples were also provided by the CRU (Swansea, Wales). Supplementary genomic DNA samples covering an extended array of *C. parvum* gp60-subtypes were also provided by the CRU and incorporated into the study. A single isolate from each gp60-subtype designation was included in this initial study panel.

The gp60-subtype designation and provenance of the isolates included in this study are detailed in Table 5.1.
Table 5.1 gp60-subtype designations and respective provenance details of *C. parvum* study isolates

<table>
<thead>
<tr>
<th>gp60-subtype</th>
<th>Provenance</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. parvum</em> IaA10G2R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA15G1R2</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA15G2R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA16R1</td>
<td>United Kingdom</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA16G1R1</td>
<td>United Kingdom</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA17G1R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA17G2R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA17G3R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA17G4R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA18G1R1</td>
<td>United Kingdom</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA18G3R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA19G3R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA19G4R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA20G1R1</td>
<td>United Kingdom</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA21G2R1</td>
<td>United Kingdom</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA20G3R1</td>
<td>Ireland</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA21G1R1</td>
<td>United Kingdom</td>
</tr>
<tr>
<td><em>C. parvum</em> IaA21G3R1</td>
<td>Ireland</td>
</tr>
</tbody>
</table>

5.3.2 Genetic Marker Selection

Genetic marker selection was based on a number of criteria in order to select those which contained tandem repeat loci optimally suited for HRM analysis application, which requires amplicons ranging from 100 – 150 bp for effective resolution between base pair differences (Tong *et al.*, 2009; Tong and Giffard, 2012). Selection criteria examined prior to locus selection included: tandem repeat sequence; tandem repeat length (repeat regions to be ≤ 150 base pairs in length); tandem repeat %G+C content (an acceptable %G+C content range of a minimum of 33% to a maximum of 83% was established from evaluation of potential candidate loci); combined fragment size range identified between *C. parvum* isolates (amplicons to be ≤ 200 base pairs in length for optimal HRM analysis differentiation); the upstream and downstream sequence conservation (with a preference towards sequence conservation within these regions); and the number of alleles previously described per genetic marker. Markers were selected across several chromosomes to offset the risk of genetic linkage, while those
located on the same chromosome were distant enough to ensure the risk of linkage was excluded (Widmer and Sullivan, 2012)

Based on these criteria, candidate loci were compiled from mini- and microsatellite markers commonly employed in multi-locus genotyping based intra-Cryptosporidium species differentiation (Robinson and Chalmers, 2012). Selection was also based on loci which were reported to most effectively achieve differentiation of 95% of multi locus genotypes among previous MLST studies (Robinson and Chalmers, 2012). Of these loci, *gp60*, MM5 and MS9-Mallon were selected as the most suitable candidates for further analysis.

### 5.3.3 Subtyping Primer Design

Multiple sequence alignment was conducted using MUSCLE software (https://www.ebi.ac.uk/Tools/msa/muscle/) (Madeira *et al.*, 2019) on *gp60*-subtype sequences from *gp60*-subtype families prevalent in Europe (Cacciò and Chalmers, 2016), specifically the *C. parvum* IIa-IIj families (GenBank accession numbers: AB242224-AB242229, AY382675, AY738185, AY738186, AY738188-AY7381889, AY873780-AY873782, AY738191, AY738195, DQ192502-DQ192508, DQ630514-DQ630515, DQ630517, DQ630519, DQ648531-DQ648537, EU140508). This was done to identify homologous regions circumscribing the tandem repeat region of *gp60* between the various subtype families. Multiple sequence alignments were also conducted to identify regions of homology using available sequence data for the MM5 (GenBank accession numbers: KP172504, KP172505, KP265906, KP265907, KP265908, KP265909, KP265910, KP265911) and MS9-Mallon loci (GenBank accession numbers: KT922210, KT922211, KT922212, KT922213, KT922214).

Outer and inner primers were designed within the identified homologous regions using the online primer design tool, Primer-Blast (www.ncbi.nlm.nih.gov/tools/primer-blast/) (Ye *et al.*, 2012). Amplicon sizes for all resultant primer pairs varied, with amplicon length dependent on the number of tandem repeats present.

All primer sets designed for each genetic locus are detailed in Table 5.2.
Table 5.2  Primers used for the real-time PCR based amplification and gp60-subtyping of C. parvum isolates

<table>
<thead>
<tr>
<th>Locus</th>
<th>Chromosome</th>
<th>Tandem Repeat Sequence</th>
<th>Primer Sequence (5’ – 3’)</th>
<th>Amplicon Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gp60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outer primers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPGP60OUT F</td>
<td>6</td>
<td>TC(A/G/T)</td>
<td>TCTCCGTTATAGTCTCCGCTGT</td>
<td>462 – 498</td>
</tr>
<tr>
<td>CPGP60OUT R</td>
<td></td>
<td>TC(G/A/T)</td>
<td>TCGGGGATCTGTGTTGGTCTTT</td>
<td></td>
</tr>
<tr>
<td>Inner primers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPGP60IN F</td>
<td></td>
<td></td>
<td>CAGCCGGTTCCACTCAGA</td>
<td>141 – 177</td>
</tr>
<tr>
<td>CPGP60IN R</td>
<td></td>
<td></td>
<td>GACTGCCTTCTGCGTCTT</td>
<td></td>
</tr>
<tr>
<td>MM5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outer primers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPMM5OUT F</td>
<td></td>
<td></td>
<td>TTTCTGGACTGGATTTGGACTT</td>
<td>334 – 418</td>
</tr>
<tr>
<td>CPMM5OUT R</td>
<td></td>
<td></td>
<td>CTCGCANAATCTCCTGCTCTTTCTT</td>
<td></td>
</tr>
<tr>
<td>Inner primers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPMM5IN F</td>
<td></td>
<td></td>
<td>GGAACGTAACAAAACACAGCG</td>
<td>113 – 170</td>
</tr>
<tr>
<td>CPMM5IN R</td>
<td></td>
<td></td>
<td>TAGCCGAATCTGGCGGTATC</td>
<td></td>
</tr>
<tr>
<td>MS9-Mallon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outer primers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPMS9OUT F</td>
<td></td>
<td></td>
<td>AAAGCAGGAAGGGAGAAGCA</td>
<td>360 – 403</td>
</tr>
<tr>
<td>CPMS9OUT R</td>
<td></td>
<td></td>
<td>AGAAATAGAGCTTTGGCTGGAAC</td>
<td></td>
</tr>
<tr>
<td>Inner primers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPMS9IN F</td>
<td></td>
<td></td>
<td>ACCTGGAGTGATTTGG</td>
<td>165 – 195</td>
</tr>
<tr>
<td>CPMS9IN R</td>
<td></td>
<td></td>
<td>GTTCTTGTGTTAAAAGTAAAAATCC</td>
<td></td>
</tr>
</tbody>
</table>
5.3.4 Nested Real-time PCR Amplification

Nested real-time PCR was employed to standardise the concentration of DNA template prior to HRM analysis conducted in second-round amplification. The variable nature of template concentration in faecal sample extracts precluded the use of a single round of amplification.

Real-time PCR amplifications were performed in 20 μl reaction volumes. First round, outer amplification contained 10 μl FastStart Essential DNA Green Master (Roche Molecular Diagnostics, Basel, Switzerland). *C. parvum gp60*, MM5 and MS9-Mallon outer primers were used at concentrations of 2 μM, 1 μM and 2 μM, respectively; 5μl of genomic template DNA was added to each reaction. All reactions were conducted using the LightCycler 96 (LC96) instrument (Roche Molecular Diagnostics, Basel, Switzerland).

All real-time PCR reactions were conducted under the following cycling conditions: initial denaturation at 95°C for 10 min, subsequent 3-step amplification for 45 cycles, including denaturation at 95°C for 30s, annealing at 65°C for 30s and extension at 72°C for 40s. The resulting DNA amplicons were purified using the High Pure PCR product purification kit (Roche Molecular Diagnostics, Basel, Switzerland), as per manufacturer’s instructions, and were subsequently diluted in a ratio of 1:50 with molecular grade water prior to further analysis.

Second round, inner amplification reactions contained 4 μl of LightCycler 480 High Resolution Melting Master (Roche Molecular Diagnostics, Basel, Switzerland). Magnesium chloride (MgCl₂) was also included at a concentration of 3 mM for all reactions. *gp60*, MM5 and MS9-Mallon primers were used at a concentration of 3 μM, 2 μM and 3 μM, respectively, and 2μl of genomic template DNA was also added to each reaction. All second round reactions were conducted in duplicate.

Real-time PCR reactions were conducted under the following cycling conditions: initial denaturation at 95°C for 10 min, subsequent 3-step amplification for 35 cycles, including denaturation at 95°C for 30s, annealing at 60°C for 30s, and extension at 72°C for 40s.
5.3.5 DNA Sequence Analysis of First Round Amplicons

All purified, first round gp60, MM5 and MS9-Mallon amplicons for the studied C. parvum isolates were sequenced bidirectionally via Sanger sequencing (Eurofins Genomics, Cologne, Germany).

5.3.6 HRM Curve Acquisition and Analysis

HRM analysis was conducted immediately post-PCR amplification on the LC96 by incrementally increasing the temperature by 2.2°C/s from 65°C to 97°C, taking 15 continuous readings/°C. Following high resolution melting-curve acquisition, normalisation regions of 79 to 86 °C were applied for analysis within the LC96 software (Roche Molecular Diagnostics, Basel, Switzerland). The positive/negative threshold was set to the default 0.05 relative fluorescence units (RFU). Delta Tₘ discrimination and curve shape discrimination parameters were set to 100%. Normalised melting curve, differential melting peak, and difference plots were generated and analysed to determine the precise melting temperature of each C. parvum gp60-subtype. To be considered acceptable for HRM analysis, sample Cp values were to be less than 30 cycles, and inter-sample variation in Cp values were to vary by no more than 5 cycle units. Analysis was conducted on all duplicate runs.

5.3.7 Evaluation of Assay Sensitivity and Specificity

Assay sensitivity and specificity testing on genomic Cryptosporidium DNA originating from a faecal environment was conducted for both outer and inner primers using enumerated C. parvum samples. Prior to extraction, oocysts within semi-purified faecal samples were microscopically enumerated with KOVA Glasstic Slides (Medical Supply Company, Dublin, Ireland). The C. parvum sample was determined to contain approximately 1 x 10⁴ oocysts per ml.

DNA was extracted from oocyst samples following the Enteric Bio SPS extraction protocol (Serosep, Limerick, Ireland) employed routinely by CUH in Cryptosporidium DNA extractions. The whole tip of a Copan FLOQswab was lightly coated with the specimen, which was inoculated into a 4mL SPS tube and heated at 103°C for 30 min. A 10-fold serial dilution, ranging from 1 x 10⁴ oocysts per ml to 1 x 10⁻¹ oocysts per ml, was prepared from the C. parvum genomic extract. The limits of detection for gp60,
MM5 and MS9-Mallon outer and inner primer pairs were assessed using melting curve analysis and gel electrophoresis. All reactions were conducted in duplicate.

The specificity of *C. parvum gp60*, MM5 and MS9-Mallon primers sets were also evaluated using DNA extracts from varying enteric pathogens, of various genera, both bacterial and protozoan in nature. *Salmonella, Shigella, Campylobacter, Verotoxigenic Escherichia coli (VTEC)* and *G. lamblia* positive faecal DNA extracts detected during routine molecular enteric screening in CUH were tested. Genomic DNA extracted from *B. hominis* cysts (ATCC, United States of America) employing the DNeasy Blood and Tissue Kit (Qiagen, German) as per the manufacturer’s instructions, was also evaluated. This genomic DNA was exposed to identical real-time PCR conditions and primer concentrations as those outlined above. All reactions were conducted in duplicate.

### 5.3.8 Reproducibility of HRM Analyses

Intra- and inter-experiment reproducibility was assessed by calculating the mean, standard deviation, and coefficients of variation of the duplicate Tₘ peaks generated for each *gp60*-subtype isolate the HRM analysis software. To assess intra-experimental reproducibility, duplicate reactions were conducted for each sample within a single run. Inter-experimental variation was assessed in separate, duplicate runs of identical *gp60*-subtype sample composition.
5.4 Results

5.4.1 DNA Sequencing of Genetic Marker Subtypes

Multiple alleles were identified via DNA sequencing for the gp60, MM5 and MS9-Mallon loci, respectively. The DNA sequences alignments of the heterogeneous regions within the gp60-subtypes, and MM5 and MS9-Mallon alleles encountered in this study are presented in Tables 5.3, 5.4 and 5.5, respectively.

The gp60-subtype designations of the 18 isolates analysed in this study were initially determined during a prior epidemiological study (as detailed in Chapter IV). Consequently, gp60-subtype numbers were designated based on ascending number of TCA, TCG and R repeats present within gp60-subtypes, while HRM analysis groups were designated based on ascending amplicon T_m value. The ‘R’ repeat designation indicates additional repeat sequences that often appear at the end of the gp60 repeat region. Within the C. parvum IIa subtype family, the tandem repeat region of some gp60-subtypes is immediately followed by one or several copies of the ACATCA)

DNA sequence analysis of the MM5 and MS9-Mallon amplicons identified two and five alleles associated with these loci, respectively, among the 18 gp60-subtype isolates analysed. 11 isolates possessed MM5 allele 1, while seven possessed MM5 allele 2. MS9-Mallon alleles 1 and 3 were present in only one isolate each, respectively, while MS9-Mallon alleles 2 and 4 were present in two isolates, respectively. The remaining 12 isolates possessed MS9-Mallon allele 5.

The nature of the inner primer pairs designed for each locus ensured sequence homology both up- and downstream of the tandem repeat region within an amplicon. Consequently, only variations within the tandem repeat region of each locus contributed to differences in amplicon length and T_m and, ultimately, to HRM analysis results.
<table>
<thead>
<tr>
<th>gp60</th>
<th>Nucleotide position (fragment length = 141 – 177 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>45</td>
</tr>
<tr>
<td>Seq 1</td>
<td>G</td>
</tr>
<tr>
<td>Seq 2</td>
<td>G</td>
</tr>
<tr>
<td>Seq 3</td>
<td>G</td>
</tr>
<tr>
<td>Seq 4</td>
<td>G</td>
</tr>
<tr>
<td>Seq 5</td>
<td>G</td>
</tr>
<tr>
<td>Seq 6</td>
<td>G</td>
</tr>
<tr>
<td>Seq 7</td>
<td>G</td>
</tr>
<tr>
<td>Seq 8</td>
<td>G</td>
</tr>
<tr>
<td>Seq 9</td>
<td>G</td>
</tr>
<tr>
<td>Seq 10</td>
<td>G</td>
</tr>
<tr>
<td>Seq 11</td>
<td>G</td>
</tr>
<tr>
<td>Seq 12</td>
<td>G</td>
</tr>
<tr>
<td>Seq 13</td>
<td>G</td>
</tr>
<tr>
<td>Seq 14</td>
<td>G</td>
</tr>
<tr>
<td>Seq 15</td>
<td>G</td>
</tr>
<tr>
<td>Seq 16</td>
<td>G</td>
</tr>
<tr>
<td>Seq 17</td>
<td>G</td>
</tr>
<tr>
<td>Seq 18</td>
<td>G</td>
</tr>
</tbody>
</table>

**Table 5.3** DNA sequence alignments detailing *C. parvum* sequence variations within the tandem repeat region of the gp60 locus

**Abbreviations:** Seq = Sequence

Sequence 1: IlaA1062R1; Sequence 2: IlaA15G1R2; Sequence 3: IlaA15G2R1; Sequence 4: IlaA16R1; Sequence 5: IlaA16G1R1; Sequence 6: IlaA17G1R1; Sequence 7: IlaA17G2R1; Sequence 8: IlaA17G3R1; Sequence 9: IlaA17G4R1; Sequence 10: IlaA18G1R1; Sequence 11: IlaA18G3R1; Sequence 12: IlaA19G3R1; Sequence 13: IlaA19G4R1; Sequence 14: IlaA20G1R1; Sequence 15: IlaA20G2R1; Sequence 16: IlaA20G3R1; Sequence 17: IlaA21G1R1; Sequence 18: IlaA21G3R1
Table 5.4  DNA sequence alignments detailing *C. parvum* sequence variations within the tandem repeat region of the MM5 locus

<table>
<thead>
<tr>
<th>MM5</th>
<th>Nucleotide position (fragment length = 128 – 155 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60  61  62  63  64  65</td>
</tr>
<tr>
<td>Seq 1</td>
<td>T   A   T   C   T   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   T   C   T   T</td>
</tr>
<tr>
<td>Seq 2</td>
<td>T   A   T   C   T   T   C   T   T   C   T   C   C   T   C   T   T   C   T   C   T   C   T   C   T   C   T   C   T   C   T   C   T   T</td>
</tr>
</tbody>
</table>

Abbreviations: Seq = Sequence
Sequence 1: IlaA10G2R1, IlaA15G1R2, IlaA16R1, IlaA16G1R1, IlaA17G2R1, IlaA17G3R1, IlaA18G3R1, IlaA19G3R1, IlaA20G1R1, IlaA20G3R1, IlaA21G3R1; Sequence 2: IlaA15G2R1, IlaA17G1R1, IlaA17G4R1, IlaA18G1R1, IlaA19G4R1, IlaA20G2R1, IlaA21G1R1
## Table 5.5
DNA sequence alignments detailing *C. parvum* sequence variations within the tandem repeat region of the MS9-Mallon locus

<table>
<thead>
<tr>
<th>MS9- Mallon</th>
<th>Nucleotide position (fragment length = 126 – 183 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq 1</td>
<td>A C C T G G A G T G T G A T T T T G G A T T T T G G A T T T T G G</td>
</tr>
<tr>
<td>Seq 2</td>
<td>A C C T G G A G T G T G A T T T T G G A T T T T G G A T T T T G G</td>
</tr>
<tr>
<td>Seq 3</td>
<td>A C C T G G A G T G T G - - - - - - - G A T T T T G G A T T T T G G</td>
</tr>
<tr>
<td>Seq 4</td>
<td>A C C T G G A G T G T G - - - - - - - - - - - - - G A T T T T G G A T T T T G G</td>
</tr>
<tr>
<td>Seq 5</td>
<td>A C C T G G A G T G T G - - - - - - - - - - - - - G A T T T T G G A T T T T G G</td>
</tr>
<tr>
<td>Seq 6</td>
<td>A C C T G G A G T G T G - - - - - - - - - - - - - G A T T T T G G A T T T T G G</td>
</tr>
</tbody>
</table>

Abbreviations: Seq = Sequence
Sequence 1: IlA20G1R1; Sequence 2: IlA16R1, IlA16G1R1; Sequence 3: IlA15G1R2; Sequence 4: IlA21G1R1, IlA21G3R1; Sequence 5: IlA10G2R1, IlA15G2R1, IlA17G1R1, IlA17G2R1, IlA17G3R1, IlA17G4R1, IlA18G1R1, IlA18G3R1, IlA9G3R1, IlA19G4R1, IlA20G2R1, IlA20G3R1
5.4.2 Interpretation of Multi-locus HRM Analysis Curve Profiles and Categorisation of *C. parvum* gp60-subtypes

The melting characteristics of the respective gp60-subtype amplicon groups generated for each of the three tandem repeat loci, gp60, MM5 and MS9-Mallon, were assessed through plots generated by the HRM analysis function of the LC96 software (Roche Molecular Diagnostics, Basel, Switzerland). Normalised melting curves, presenting data as relative fluorescent units versus temperature, are depicted for each of the three loci in Figs 5.1, 5.4 and 5.7, respectively. They include normalised melting curve data for each gp60-subtype. In these plots, pre- and post-melt signals were uniformly set to relative values of 100% and 0%, respectively, to aid interpretation and analysis of the data. Normalised melting peaks, which are based upon the first negative derivative of the change in fluorescence over time (\(-dF/dT\)) versus temperature, are depicted in Figs 5.2, 5.5 and 5.8, respectively, and represent the exact $T_m$ of each gp60-subtype, visually depicting the spread of melting temperature exhibited by each gp60-subtype grouping. Finally, difference plots, which are based upon the difference in fluorescence versus temperature, are observed in Figs 5.3, 5.6 and 5.9, and depict melting curve data following subtraction of baseline curves in order to better highlight distinct gp60-subtype groupings.
Figure 5.1  Normalised melting curves of *C. parvum* gp60-subtypes at the *gp60* locus

Figure 5.2  Normalised melting peaks of the *C. parvum* gp60-subtypes at the *gp60* locus
Figure 5.3  Difference plot of *C. parvum* gp60 -subtypes at the gp60 locus.
Differentiation between each of the 18 individual *C. parvum* gp60-subtypes was not fully achievable through HRM analysis of the gp60 tandem repeat motif alone. However, based on variations in melting temperature exhibited by the gp60 amplicons, it was possible to segregate these 18 gp60-subtypes into 8 distinct, predictive T\textsubscript{m} groupings, as highlighted in Table 5.6. By virtue of their distinct T\textsubscript{m} values, gp60-subtypes IIaA10G1R1, IIaA16R1, IIa17G4R1 were clearly distinguishable from the other gp60-subtypes. While the gp60-subtypes IIaA20G1R1 and IIaA21G1R1; IIaA18G1R1, IIaA20G2R1 and IIaA17G1R1; IIaA15G1R2 and IIaA16G1R1; IIaA17G2R1, IIaA19G3R1, IIaA20G3R1, IIaA21G3R1; and IIaA15G2R1, IIaA18G3R1, IIaA17G3R1 and IIaA19G4R1 were grouped together, respectively, as the T\textsubscript{m} values of the individual gp60-subtype amplicons within these groupings occupied overlapping temperature ranges, precluding definitive gp60-subtype
discrimination. The allelic and HRM analysis grouping designations for the *gp60* locus are outlined in Table 5.6 and depicted in Figs 5.1, 5.2 and 5.3.

The number of HRM analysis groupings within the MM5 locus corresponded to the two alleles detected among the 18 sequenced *gp60*-subtypes. All 10 *gp60*-subtypes found to possess allele 1 were also assigned to Group 1, while the seven *gp60*-subtypes that possessed the alternative allele, which exhibited a higher T_m than Group 1, were assigned to Group 2 (amplification of IIaA10G2R1 was unsuccessful at the MM5 locus). The MM5 allele and HRM analysis grouping breakdown among the studied isolates are outlined in Table 5.7 and depicted in Figs 5.4, 5.5 and 5.6.

**Figure 5.4** Normalised melting curves of *C. parvum* *gp60*-subtypes at the MM5 locus;
Figure 5.5  Normalised melting peaks of the *C. parvum* gp60 -subtypes at the MM5 locus

Figure 5.6  Difference plot of *C. parvum* gp60 -subtypes at the MM5 locus
Table 5.7  HRM analysis based MM5-subtype group designations and subtype amplicon characteristics

<table>
<thead>
<tr>
<th>HRM Analysis Group</th>
<th>Subtypes</th>
<th>%G+C Content</th>
<th>Sequence Length (bp)</th>
<th>Upper $T_m$ Limit (°C)</th>
<th>Lower $T_m$ Limit (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>IlA15G1R2</td>
<td>42.97</td>
<td>128</td>
<td>79.11</td>
<td>79.36</td>
</tr>
<tr>
<td></td>
<td>IlA16R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA16G1R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA17G2R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA17G3R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA18G3R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA19G3R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA20G1R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA20G3R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA21G3R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>IlA15G2R1</td>
<td>43.87</td>
<td>155</td>
<td>79.90</td>
<td>80.15</td>
</tr>
<tr>
<td></td>
<td>IlA17G1R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA17G4R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA18G1R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA19G4R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA20G2R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IlA21G1R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The number of HRM analysis groupings within the MS9-Mallon locus, which were numbered in relation to the ascending $T_m$ values, also corresponded exactly to the five alleles detected among the 18 sequenced isolates within the study. One $gp60$-subtype was assigned to Group 1, two to Group 2, one to Group 3, one to Group 4, and 12 to Group 5. The allelic and HRM analysis grouping designations for the MS9-Mallon locus are outlined in Table 5.8 and depicted in Figs 5.7, 5.8 and 5.9.
Figure 5.7  Normalised melting curves of *C. parvum* subtypes at the MS9-Mallon locus.

Figure 5.8  Normalised melting peaks of the *C. parvum* subtypes at the MS9-Mallon locus;
Figure 5.9  Difference plot of *C. parvum* subtypes at the MS9-Mallon locus.
The combination of *gp60* HRM analysis data with that obtained for the MM5 and MS9-Mallon loci further aided differentiation both within and between *gp60*-subtype groupings. *gp60*-subtypes IIaA20G1R1 and IIaA21G1R1, although presenting the same *gp60* HRM analysis grouping, were assigned to distinct groups for both the MM5 and MS9-Mallon loci. Consequently, these *gp60*-subtypes were completely distinguishable from the remaining 16 *gp60*-subtypes. The isolates assigned to *gp60* Group 5 were differentiated further, but not completely, by HRM analysis of the MS9-Mallon locus. Within this group, IIaA21G3R1 produced a distinct MS9-Mallon group designation to the remaining three *gp60*-subtypes (IIaA17G2R1, IIaA19G3R1 and IIaA20G3R1). Several *gp60*-subtypes assigned to *gp60* Group 6, IIaA15G2R1, IIaA17G3R1, IIaA18G3R1 and IIaA19G4R1, were assigned to different MM5 HRM analysis groups, with IIaA18G3R1 and IIaA17G3R1 assigned to MM5 Group 1 and IIaA15G2R1 and
IIaA19G4R1 assigned to Group 2. This allowed for further differentiation between these four \( gp60 \)-subtypes.

It was also possible to further distinguish between \( gp60 \)-subtypes IIaA15G2R1 and IIaA19G4R1 based on melting peak shape. Both of these displayed the same HRM analysis grouping pattern, due to the distinctively asymmetrical, humped melting peak produced by IIaA15G2R1 at the \( gp60 \) locus. IIaA15G2R1 was the only \( gp60 \)-subtype to produce a morphologically aberrant peak at this locus.

As outlined in Table 5.9, no further differentiation was possible between two separate groups of three \( gp60 \)-subtypes (IIaA17G1R1, IIaA18G1R1 and IIaA20G2R1; and IIaA17G2R1, IIaA19G3R1 and IIaA20G3R1, respectively) that exhibited identical multi-locus HRM analysis patterns.
### Table 5.9 Multi-locus HRM analysis grouping and sequence designations of C. parvum gp60-subtype isolates

<table>
<thead>
<tr>
<th>Isolate</th>
<th>gp60 HRM Analysis Grouping</th>
<th>gp60 Sequence Designation</th>
<th>MM5 HRM Analysis Grouping</th>
<th>MM5 Sequence Designation</th>
<th>MS9-Mallon HRM Analysis Grouping</th>
<th>MS9-Mallon Sequence Designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>IlaA20G1R1</td>
<td>1</td>
<td>14</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IlaA21G1R1</td>
<td>1</td>
<td>17</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>IlaA16G1R1</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>IlaA17G1R1</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA18G1R1</td>
<td>3</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA20G2R1</td>
<td>3</td>
<td>15</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA15G1R2</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>IlaA16G1R1</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>IlaA17G2R1</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA19G3R1</td>
<td>5</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA20G3R1</td>
<td>5</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA21G3R1</td>
<td>5</td>
<td>18</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>IlaA15G2R1</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA17G3R1</td>
<td>6</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA18G3R1</td>
<td>6</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA19G4R1</td>
<td>6</td>
<td>13</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA17G4R1</td>
<td>7</td>
<td>9</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>IlaA10G2R1</td>
<td>8</td>
<td>1</td>
<td>n/a(^a)</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

\(^a\) Differentiated at the *gp60* locus based on melting peak shape

\(^b\) No amplification occurred at this locus

### 5.4.3 Detection Sensitivity and Specificity of C. parvum Tandem Repeat Locus Targets

The evaluation of detection sensitivity of the *C. parvum* inner and outer primer pair sets was conducted on 10-fold serial dilutions of genomic DNA extracted from enumerated *C. parvum* oocysts. Both dilution series spanned a range of $10^4$ to $10^1$ oocysts per ml. Although extracted from semi-purified faecal samples, the complex microbial nature of such samples, and the lack of commercially available pure genomic DNA samples of *C. parvum* precluded an accurate determination of the exact genomic DNA concentration. Thus, DNA concentration is given in terms of serially diluted, enumerated oocysts. The
limit of detection (LOD) for all outer *C. parvum* genetic marker primer pairs was $10^2$ oocysts per ml. The LOD for the *C. parvum* inner *gp60* primer pairs was determined to be $10^3$ oocysts per ml, while the LOD for both inner MM5 and MS9-Mallon primer pairs was determined to be $10^2$ oocysts per ml.

Clinical faecal samples that were determined to not contain evidence of infection by bacterial, fungal or parasitic gastrointestinal pathogens were tested to ascertain the specificity of *gp60*, MM5 and MS9-Mallon outer and inner primer pairs, respectively. Additionally, a range of samples known to contain prokaryotic (*Salmonella*, *Shigella*, *Campylobacter*, VTEC), and eukaryotic (*G. lamblia* and *B. hominis*) organisms were also tested. No unintended amplification occurred in these samples.

Appendix II details the real-time PCR and gel electrophoresis data in support of the results outlined above.

**5.4.4 Assessment of Intra- and Inter-assay Variabilities**

Upon confirmation of the sensitivity and specificity of assay primer sets and PCR conditions, the melting characteristics of the *C. parvum* *gp60*, MM5 and MS9-Mallon amplicons were assessed. Critical evaluations of intra-assay and inter-assay variabilities were carried out by calculating the mean, standard deviation and coefficient of variation of melting temperatures obtained within run and between run duplicates for each of the 18 *C. parvum* *gp60*-subtypes studied.

The coefficient of variation among and within experiments was determined to be $<0.5\%$, indicating a high level of reproducibility. The results of this critical evaluation are outlined in Table 5.10.
<table>
<thead>
<tr>
<th>C. parvum gp60-subtype</th>
<th>gp60</th>
<th>MM5</th>
<th>MS9-Mallon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intra-assay</td>
<td>Inter-assay</td>
<td>Intra-assay</td>
</tr>
<tr>
<td></td>
<td>Reproducibility</td>
<td>Reproducibility</td>
<td>Reproducibility</td>
</tr>
<tr>
<td></td>
<td>Mean T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>CV</td>
<td>Mean T&lt;sub&gt;m&lt;/sub&gt;</td>
</tr>
<tr>
<td>IIAA10G2R1</td>
<td>83.31 ± 0.02</td>
<td>0.03%</td>
<td>83.30 ± 0.08</td>
</tr>
<tr>
<td>IIAA15G2R1</td>
<td>82.81 ± 0.01</td>
<td>0.01%</td>
<td>82.77 ± 0.08</td>
</tr>
<tr>
<td>IIAA15G1R2</td>
<td>82.45 ± 0.04</td>
<td>0.05%</td>
<td>82.43 ± 0.11</td>
</tr>
<tr>
<td>IIAA16G1R1</td>
<td>82.19 ± 0.04</td>
<td>0.05%</td>
<td>82.15 ± 0.07</td>
</tr>
<tr>
<td>IIAA16G1R1</td>
<td>82.42 ± 0.00</td>
<td>0.00%</td>
<td>82.43 ± 0.05</td>
</tr>
<tr>
<td>IIAA17G1R1</td>
<td>82.30 ± 0.04</td>
<td>0.04%</td>
<td>82.34 ± 0.08</td>
</tr>
<tr>
<td>IIAA17G2R1</td>
<td>82.59 ± 0.05</td>
<td>0.06%</td>
<td>82.61 ± 0.07</td>
</tr>
<tr>
<td>IIAA17G3R1</td>
<td>82.75 ± 0.00</td>
<td>0.00%</td>
<td>82.80 ± 0.08</td>
</tr>
<tr>
<td>IIAA17G4R1</td>
<td>83.14 ± 0.00</td>
<td>0.00%</td>
<td>83.13 ± 0.08</td>
</tr>
<tr>
<td>IIAA18G1R1</td>
<td>82.32 ± 0.04</td>
<td>0.05%</td>
<td>82.26 ± 0.04</td>
</tr>
<tr>
<td>IIAA18G3R1</td>
<td>82.81 ± 0.00</td>
<td>0.00%</td>
<td>82.74 ± 0.07</td>
</tr>
<tr>
<td>IIAA19G3R1</td>
<td>82.65 ± 0.04</td>
<td>0.05%</td>
<td>82.56 ± 0.11</td>
</tr>
<tr>
<td>IIAA19G4R1</td>
<td>82.85 ± 0.05</td>
<td>0.06%</td>
<td>82.87 ± 0.12</td>
</tr>
<tr>
<td>IIAA20G1R1</td>
<td>82.02 ± 0.00</td>
<td>0.00%</td>
<td>82.03 ± 0.11</td>
</tr>
<tr>
<td>IIAA20G2R1</td>
<td>82.29 ± 0.00</td>
<td>0.00%</td>
<td>82.26 ± 0.08</td>
</tr>
<tr>
<td>IIAA20G3R1</td>
<td>82.55 ± 0.00</td>
<td>0.00%</td>
<td>82.59 ± 0.15</td>
</tr>
<tr>
<td>IIAA21G1R1</td>
<td>81.99 ± 0.04</td>
<td>0.05%</td>
<td>81.97 ± 0.12</td>
</tr>
<tr>
<td>IIAA21G3R1</td>
<td>82.52 ± 0.04</td>
<td>0.05%</td>
<td>82.43 ± 0.05</td>
</tr>
</tbody>
</table>
5.5 Discussion

This is the first study to investigate the application of HRM analysis to the subtypic differentiation of *C. parvum* isolates. The method developed in this study was capable of differentiating between certain *C. parvum gp60*-subtypes at the *gp60*, MM5 and MS9-Mallon loci. However, overlapping T<sub>m</sub> values precluded absolute differentiation between all 18 studied *gp60*-subtypes. Several *gp60*-subtype groupings were established, reflective of these overlapping T<sub>m</sub> values, thus allowing for differentiation to these subtypic categories. Overall, this method successfully differentiated between 10 of the studied *C. parvum gp60*-subtypes: IlaA10G2R1, IlaA15G1R2, IlaA15G1R2, IlaA15G2R1, IlaA16R1, IlaA16G2R1, IlaA17G4R1, IlaA19G4R1, IlaA20G1R1, IlaA21G1R1, and IlaA21G3R1. Meanwhile, distinct differentiation groupings were established for the remaining eight *gp60*-subtypes, encompassing IlaA17G1R1, IlaA18G1R1 and IlaA20G2R1; IlaA17G2R1, IlaA19G3R1, and IlaA20G3R1, and IlaA17G3R1 and IlaA18G3R1, respectively.

Differentiation of *gp60*-subtype IlaA18G3R1, the most regionally prevalent *gp60*-subtype in Ireland, which accounts for approximately 60% of all clinical cases of cryptosporidiosis, was grouped with *gp60*-subtype IlaA17G3R1. However, regional incidence of the latter is rare, with only two cases of IlaA17G3R1 reported as a cause of human cryptosporidiosis by epidemiological studies over the past two decades (Zintl *et al.*, 2011). HRM analysis also successfully differentiated the second most regionally prevalent *gp60*-subtype, IlaA15G2R1, which accounts for approximately 9% of human cases, from the 15 remaining *gp60*-subtypes. However, differentiation of IlaA20G3R1, which accounts for approximately 5% of cases, was complicated by a HRM analysis grouping pattern identical to that of IlaA17G2R1 and IlaA19G3R1. The remaining nine Irish *gp60*-subtypes are generally encountered sporadically in both human and livestock infection, the combined total of which accounts for less than 15% of all cases (Zintl *et al.*, 2011). The *gp60*-subtypes IlaA16R1, IlaA16G1R1, IlaA18G1R1, IlaA20G1R1, IlaA20G2R1 and IlaA21G1R1, all of which were isolated in the UK and, with the exception of IlaA21G2R1, were fully differentiated from the remaining *gp60*-subtypes by multi-locus HRM analysis, were included to provide a more comprehensive understanding of the capabilities and limitations of the proposed method.

There remains a strong empirical aspect to the kind of base changes detectable via HRM analysis. Amplicon T<sub>m</sub> is generally a composite result of several factors, including
amplicon length, DNA sequence and %G+C content. The correlation of $T_m$ with %G+C content is due to the presence of an additional third hydrogen bond between GC pairs, requiring a higher temperature to achieve strand dissociation than with double hydrogen bonded AT pairs. Thus, it is more difficult to detect base differences that do not appreciably change the %G+C content an amplicon than those that do via HRM analysis, particularly with increasing amplicon length, given the asymptotic relationship exhibited by these two amplicon characteristics (Tong and Giffard, 2012).

This provides the basis for explanation of the limited discrimination between distinct gp60-subtypes observed in this study. It is proposed that the general configuration of the TCA/TCG/TCT repeats, the predominance of TCA repeats, and the limited presence of TCG repeats within the gp60 microsatellite motif are attributable to lack of discriminatory power exhibited by HRM analysis at this genetic marker. Tong et al. (2009) report that length related increases in $T_m$ become more difficult to detect as amplicon size increases and are largely rendered asymptomatic beyond 100 bp (Tong et al., 2009; Tong and Giffard, 2012). Consequently, variations in gp60 tandem repeat motif sequence did not alter %G+C content to sufficiently overcome the negating effect of variations in amplicon length in excess of 150 bp, as highlighted in Table 5.6.

However, despite the overlapping $T_m$ values between different gp60-subtypes observed in this study, the C. parvum groupings defined herein, although not capable of distinguishing between all gp60-subtypes, still allow for the distinct identification of 10 gp60-subtypes. They also serve as a useful, indicative epidemiological tool that narrows the focus to gp60-subtypes assigned to larger groupings. Additionally, melting peak morphology is a function of the precise DNA sequence and resultant base pair interactions (Tong and Giffard, 2012). Regions of sequence that contain different %G+C contents may influence melting peak shape by giving rise to multiple melt domains within an amplicon (Pangasa et al., 2009; Tong and Giffard, 2012). This is likely to be the cause of the alternative melting peak morphology exhibited by IIaA15G2R1 at the gp60 locus.

With regard to the MM5 and MS9-Mallon tandem repeat loci, unlike the gp60 locus, the number of alleles identified among the gp60-subtypes isolates studied via DNA sequencing corresponded to the number of observed HRM analysis groupings, with two and five alleles families determined for these loci, respectively. The inclusion of these
loci allowed for the definitive differentiation between several isolates. Two alleles were observed within the MM5 locus, which when used in conjunction with \textit{gp60} HRM analysis data, differentiated between the IlaA15G2R1/IlaA19G2R1 and IlaA173R1/IlaA18G3R1 groupings. Five alleles were observed within the MS9-Mallon locus, which when used in conjunction with \textit{gp60} HRM analysis data, resulted in definitive differentiation between \textit{gp60}-subtypes IlaA15G1R2 and IlaA16R1, and \textit{gp60}-subtypes IlaA20G1R1, and IlaA21G1R1. The MS9-Mallon locus also contributed to the differentiation of IlaA21G3R1 from the IlaA17G2R1/IlaA19G3R1/IlaA20G3R1 grouping. However, the limited number of alleles present at these genetic markers limited their discriminative utility to fully differentiate between the studied \textit{gp60}-subtypes.

The intra- and inter-assay coefficients of variation were less than 0.5\% for all isolates analysed at the \textit{gp60}, MM5 and MS9-Mallon loci, as outlined in Table 5.10, indicating the method to be highly reproducible. The real-time PCR conditions within this method were designed to allow for concurrent real-time PCR amplification and HRM analysis of all three loci. These results indicate that, with expansion of the panel of genetic markers to a size reflective of those employed in standard MLST and multiple locus VNTR analysis (MLVA) studies (Amer et al., 2013; Tanriverdi et al., 2008), multi-locus HRM analysis has the potential to provide a reliable, rapid, cost-effective alternative to such DNA sequencing based multi-locus analysis methods.

It has long been established that sequence analysis of the \textit{gp60} locus alone does not adequately represent intra-species diversity among \textit{Cryptosporidium} isolates (Widmer and Lee, 2010). While linkage disequilibrium is often observed among clonal populations of \textit{Cryptosporidium}, the recombination of genetically distinct \textit{Cryptosporidium} parasites that occurs in nature often gives rise to linkage disequilibrium decay and alternative panmictic and epidemic populations, which are best appraised via multi-locus sub-typing approaches (Feng et al., 2013; Li et al., 2013; Tanriverdi et al., 2008). The limited number of alleles observed at the MM5 locus and the prevalence of MS9-Mallon allele 5, particularly among isolates of Irish provenance, indicates that further expansion of the multi-locus HRM analysis panel is necessary to accommodate successful differentiation of all regional \textit{gp60}-subtypes. It is also not uncommon for in excess of seven polymorphic loci or more to be analysed in
Cryptosporidium spp. based MLST and MLVA sequencing studies (Chalmers et al., 2017).

This preliminary study was conducted to investigate the potential application of HRM analysis to MLVA based studies of C. parvum gp60-subtype. Although the results indicate the necessity for further expansion of the genetic marker panel to improve resolution of this method, they also support the inherent ability of HRM analysis to differentiate C. parvum gp60-subtypes based on variations within multiple polymorphic loci. Future work would include the expansion and further refinement of the HRM analysis panel to include additional genes known to possess highly variable tandem repeat motifs. Selected markers would span a wider variety of chromosomes in order to exclude genetic linkage (Widmer and Sullivan, 2012). It should be highlighted that this method is intended to provide a sensitive and rapid adjunct to DNA sequencing based gp60-subtype identification, and genus-level diagnostic molecular panels. The qualitative data generated by such panels are not amenable to the more detailed population genetic analyses that are possible via multi-locus or whole-genome based DNA sequencing studies.

This multi-locus HRM analysis method remains of epidemiological utility when employed in conjunction with a panel of established gp60-subtype controls, representative of the regional gp60-subtype population, for comparison and meaningful interpretation of T\textsubscript{m} results. This is especially pertinent given the prevalence of Cryptosporidium in Ireland, particularly in waterborne outbreak scenarios, and the lack of any established, national Cryptosporidium spp. subtyping framework. Thus, with further refinement, this method could potentially provide a rapid, highly sensitive and accessible method for preliminary epidemiological subtyping of both sporadic and outbreak related isolates in Irish clinical microbiology laboratories.

Appendix IV provides an addendum to this chapter wherein statistical analyses of gp60, MM5 and MS9-Mallon HRM analysis groupings and T\textsubscript{m} data are presented and discussed.
Chapter VI: Development of a novel high resolution melting analysis-based genotyping method for Cryptosporidium hominis

A manuscript based upon this chapter has been published in Experimental Parasitology.

6.1 Abstract

Cryptosporidiosis remains the leading protozoan induced cause of diarrhoea-associated mortality worldwide. *Cryptosporidium hominis*, the anthroponotically transmitted species within the *Cryptosporidium* genus, contributes significantly to the global burden of infection, accounting for the majority of clinical cases in many countries. This study applied high resolution melting analysis, a post-real-time PCR application, to the differentiation of six globally prevalent *C. hominis* gp60-subtypes. This novel method targeted three microsatellite, tandem repeat containing genetic markers, *gp60*, the genetic marker upon which current *Cryptosporidium gp60*-subtype nomenclature is based, MSB, and MSE, by which to differentiate between *C. hominis* isolates. This multi-locus approach successfully differentiated between all six *C. hominis* gp60-subtypes studied, some of which, such as IbA10G2, are known to exhibit global ubiquity. Thus, this method has the potential to be universally employed as a sensitive, cost effective and highly reproducible means to rapidly differentiate between *C. hominis gp60*-subtypes. Such a method would be of particular utility in epidemiological studies and outbreak scenarios, providing cost effective, clinically accessible alternative to DNA sequencing. The success of this preliminary study also supports further analysis of an expanded *C. hominis gp60*-subtype range and the potential expansion of the multi-locus panel in order to improve the discriminatory power of this approach.
6.2 Introduction

A recent assessment of the global burden of gastrointestinal disease found *Cryptosporidium* spp. to account for in excess of 1 million deaths and over 71 million DALYs between 1990 and 2015 (Troeger et al., 2017). The impact of this globally ubiquitous, protozoan parasite remains most profound in low income countries, where endemicity, poor sanitation, and limited access to clean water and health care culminate in these high morbidity and mortality rates, particularly in children under 5 years of age (King et al., 2019; Wazny et al., 2013).

*C. hominis* is the predominant cause of human cryptosporidiosis; responsible for 70% of such cases in many regions (Xiao, 2010). Although endemic in many countries in sub-Saharan Africa and South Asia, *C. hominis* is also a well-documented cause of cryptosporidiosis outbreaks in western regions, such as North America, Australia and many European countries (Chalmers et al., 2010; Fournet et al., 2013; Glaberman et al., 2002; Khalil et al., 2018; Leoni et al., 2007; Ng et al., 2010; Segura et al., 2015; Troeger et al., 2017; Zhou et al., 2003).

Previously thought to be human-adapted and exhibit a narrow host tropism (Feng et al., 2018), a recent analysis of literature pertaining to natural and experimental infection of animals with *C. hominis* has found the pathogen to be capable of successfully infecting a wide range of mammalian species, in addition to human and non-human primates, indicating the need for further epidemiological and transmission studies (Widmer et al., 2020). Although capable, *C. hominis* has been found to be more restricted than *C. parvum* in its capacity to undergo genetic recombination, which is limited to the more virulent subtypes (Li et al., 2013), with most studies showing *C. hominis* populations to exhibit a clonal structure in industrialised countries (Gatei et al., 2007; Mallon et al., 2003; Widmer and Sullivan, 2012). These factors are likely to contribute to the lower diversity of gp60-subtypes observed in *C. hominis*, with over 10 gp60-subtype families recognised in this species, in contrast to over 20 gp60-subtype families recognised in *C. parvum*.

Despite the global prevalence and clinical impact of *Cryptosporidium*, there remains no internationally mandated, standardised system by which to identify *Cryptosporidium* species subtypes (Chalmers et al., 2018). The original *Cryptosporidium* spp. subtyping method and related nomenclature were based on DNA sequencing of the tandem repeat
region of the 60-kDa glycoprotein gene (gp60), which encodes for surface glycoproteins involved in host cell adhesion (Strong et al., 2000; Sulaiman et al., 2005). This system is typified by species designations (Ia, Ib etc. for C. hominis; IIa, IIb etc. for C. parvum; IIIa, IIIb etc. for Cryptosporidium meleagridis), applied in conjunction with a gp60-subtype specific sequence denoting the number and nature of TCA/TCG/TCT repeats contained in the tandem repeat region (Strong et al., 2000; Sulaiman et al., 2005). For example, the designation, IbA10G2, denotes a C. hominis species gp60-subtype, containing 10 TCA and two TCG repeats within the gp60 tandem repeat motif. Globally, the gp60-subtype families Ia, Ib, Id, Ie, If and Ig are the most common C. hominis gp60-subtype families found in humans (Feng et al., 2018).

Although the gp60-subtyping nomenclature endures, limitations arising in the genotyping resolution of gp60 locus interrogation alone have led to the employment of MLST as an alternative means to conduct detailed epidemiological studies by which to make inferences regarding population structures of Cryptosporidium. A consensus list of genetic markers, however, has yet to be defined. Additionally, in many regions C. hominis exhibits a clonal population structure, with almost all autochthonous cases of C. hominis infections in Europe attributable to IbA10G2. Thus, the tracking of the clonal expansion of this C. hominis gp60-subtypes in such regions remains a challenge for conventional MLST tools (Xiao and Feng, 2017).

Herein, we use high resolution melting (HRM) analysis, a highly sensitive, fluorescence based, post-real-time PCR application capable of distinguishing sequence differences between amplicons based on melting temperature (Tm) (Tong and Giffard, 2012), to differentiate between C. hominis gp60-subtypes belonging to gp60-subtype families Ia, Ib, Id and Ie. This multi-locus approach interrogated the tandem repeat regions of the microsatellite markers, gp60, MSB and MSE, in order to differentiate between a panel of globally encountered C. hominis gp60-subtypes including; IaA14G3, IbA9G3, IbA10G2, IbA12G3, IdA21 and IeA11G3T3 (Alves et al., 2006; Cama et al., 2008; Chalmers et al., 2010; Chalmers et al., 2019; de Lucio et al., 2016; Deshpande et al., 2015; Efunshile et al., 2019; Eibach et al., 2015; Feng and Xiao, 2017; Flecha et al., 2015; Fournet et al., 2013; Glaberman et al., 2002; Khalil et al., 2017; Leoni et al., 2007; Mbae et al. 2015; Millán et al., 2019; Ng-Hublin et al., 2016, 2017; J. S. Y. Ng et al., 2010; Peralta et al., 2016; Sannella et al., 2019; Segura et al., 2015; Ukwah et al.,
This is the first such study to apply HRM analysis to the epidemiological differentiation of *C. hominis* gp60-subtypes. Although additional analysis of the capacity of this method to resolve differences within clonal *C. hominis* populations is required, with further development, this method may provide an accessible alternative to DNA sequencing by which to rapidly and sensitively identify *C. hominis* gp60-subtypes in clinical microbiology laboratories at the point of diagnosis. This would enable real-time collation of epidemiological data pertaining to this enteric pathogen. This could prove highly advantageous in advancing the current knowledge pertaining to *C. hominis* population heterogeneity in both developing and developed nations.

6.3 Materials and Methods

6.3.1 Clinical Sample Acquisition

*Cryptosporidium* positive DNA isolates, detected during clinical testing of stool samples from patients presenting with symptoms of acute gastroenteritis, were obtained from the Medical Microbiology Department of CUH over the course of three years, from August 2015 to August 2018, inclusive. The 149 amassed isolates were differentiated via fluorescent probe based real time PCR targeting the 18S rRNA gene, designed by Mary *et al.* (2013), and subtyped via gp60 gene sequencing, as discussed in Chapter III. Isolates representing four gp60-subtypes detected within this isolate bank were selected for inclusion in this study.

For control purposes, differentiated *C. hominis* DNA samples were also provided by the CRU (Swansea, Wales). Supplementary genomic DNA samples covering an extended array of *C. hominis* gp60-subtypes were also provided by the CRU and incorporated in the study.

The gp60-subtype designation and provenance of the isolates included in this study are detailed in Table 6.1.
6.3.2 Genetic Marker Selection

Genetic marker selection was based on a number of criteria, in order to select those which contained tandem repeat loci optimally suited for HRM analysis application, which requires amplicons ranging from 100 – 150 bp for effective resolution between base pair differences (Tong and Giffard, 2012). Selection criteria examined prior to locus selection included: tandem repeat sequence, tandem repeat length, tandem repeat %G+C content, combined fragment size range identified between C. hominis isolates, the up- and downstream sequence conservation and the number of alleles previously described per genetic marker. Markers were selected across several chromosomes to offset the risk of genetic linkage, while those located on the same chromosome were distant enough to ensure the risk of linkage was minimal (Widmer and Sullivan, 2012).

Based on these criteria, candidate loci were compiled from mini- and microsatellite markers commonly employed in multi-locus genotyping based intra-Cryptosporidium species differentiation (Robinson and Chalmers, 2012). Selection was also based on loci which were reported to most effectively achieve differentiation of 95% of multi locus genotypes among previous MLST studies (Robinson and Chalmers, 2012). Of these loci, gp60, MSB and MSE were selected as the most suitable candidates for further study. The respective related genes, region and function of these loci are outlined in Table 6.2.

6.3.3 Subtyping Primer Design

Multiple sequence alignment was conducted using MUSCLE software (https://www.ebi.ac.uk/Tools/msa/muscle/) upon subtype sequences from gp60-subtype families prevalent in Europe (Cacciò and Chalmers, 2016), specifically the C. hominis Ia-Ig (GenBank accession numbers: DQ192510, DQ665689, DQ665692, EF208067,
EF576982, EU161648, EU161649, EU161651, EU161652, EU161655, FJ839873-FJ839875, FJ839878, FJ839881-FJ839883, FJ861217, JF681174, JF727781, JF727787, KR296811) in order to identify homologous regions circumscribing the tandem repeat region of gp60 between the various subtypes. Multiple sequence alignments were also conducted to identify regions of homology using available sequence data for the MSB (GenBank accession numbers: KJ186959, KJ186960, KJ186961, KJ186962, KJ186963, KJ186964, KJ186971, KM222509, KM222510, KM222511, MG924442, MG924443, XM662457) and MSE loci (GenBank accession numbers: NW667260, XM661389).

Outer and inner primers were designed within the identified homologous regions using the online primer designing tool, Primer-Blast (www.ncbi.nlm.nih.gov/tools/primer-blast/). Amplicon sizes for all resultant primer pairs varied, with amplicon length dependent on the number of tandem repeats present. Designed primer pairs are detailed in Table 6.2.
Table 6.2 Primers used for the real-time PCR based amplification and subtyping of *C. hominis* samples

<table>
<thead>
<tr>
<th>Locus</th>
<th>Chromosome</th>
<th>Tandem Repeat Sequence</th>
<th>Primer Sequence (5’ – 3’)</th>
<th>Amplicon Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gp60</td>
<td>Outer primers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHGP60OUT F</td>
<td></td>
<td>TCTCCGTATATAGTCATCGCTGT</td>
<td>1003 –</td>
<td></td>
</tr>
<tr>
<td>CHGP60OUT R</td>
<td></td>
<td>TGCAGGATCTTTGTTGTCCTT</td>
<td>1089</td>
<td></td>
</tr>
<tr>
<td>Inner primers</td>
<td>6</td>
<td>TC(A/G/T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHGP60IN F</td>
<td></td>
<td>CCACTCAGGACCTTTGAAA</td>
<td>130 – 158</td>
<td></td>
</tr>
<tr>
<td>CHGP60IN R</td>
<td></td>
<td>CTTCTCGACATCTGTCTTCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSB</td>
<td>Outer primers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHMSBOUT F</td>
<td></td>
<td>TTCCACGCGCTTTTATGTAAG</td>
<td>583 – 601</td>
<td></td>
</tr>
<tr>
<td>CHMSBOUT R</td>
<td></td>
<td>GCAGATATGAGGGGCGCAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inner primers</td>
<td>1</td>
<td>CTATCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHMSBIN F</td>
<td></td>
<td>ATGATCTCGCTTTTGATCGATAACC</td>
<td>238 – 256</td>
<td></td>
</tr>
<tr>
<td>CHMSBIN R</td>
<td></td>
<td>AATTTTCATGGGAGCGTGAAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>Outer primers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHMSEOUT F</td>
<td></td>
<td>GGTTCGGGTAGAGGAGGTT</td>
<td>593 – 659</td>
<td></td>
</tr>
<tr>
<td>CHMSEOUT R</td>
<td></td>
<td>CTTCTCGATCTGATTTTCTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inner primers</td>
<td>4</td>
<td>TCAAGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHMSEIN F</td>
<td></td>
<td>ATACTACTCGGTATCTCTGAGC</td>
<td>206 – 272</td>
<td></td>
</tr>
<tr>
<td>CHMSEIN R</td>
<td></td>
<td>TGAGCTTTATCTGATTGACT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.3.4 Nested Real-time PCR Amplification

Nested real-time PCR was employed to standardise the concentration of DNA template prior to HRM analysis conducted in second-round amplification. The variable nature of template concentration in faecal sample extracts precluded the use of a single round of amplification.

Real-time PCR amplifications were performed in 20 μl reaction volumes for all reactions. First round, outer amplification contained 10 μl FastStart Essential DNA Green Master (Roche Molecular Diagnostics, Basel, Switzerland). *C. hominis* gp60, MS9 and MSE outer primers were used at concentrations of 3 μM, 2 μM and 2 μM, respectively; 5μl of genomic template DNA was added to each reaction. All reactions
were conducted using the LightCycler 96 (LC96) instrument (Roche Molecular Diagnostics, Basel, Switzerland).

All real-time PCR reactions were conducted under the following cycling conditions: initial denaturation at 95°C for 10 min, subsequent 3-step amplification for 45 cycles, including denaturation at 95°C for 30s, annealing at 63°C, 67°C and 60°C for 30s, for gp60, MSB and MSE reactions, respectively, and extension at 72°C for 40s. The resulting DNA amplicons were purified using the High Pure PCR product purification kit (Roche Molecular Diagnostics, Basel, Switzerland), as per manufacturer’s instructions, and subsequently diluted in a ratio of 1:50 with molecular grade water prior to further analysis.

Second round, inner amplification reactions contained 4 µl of LightCycler 480 High Resolution Melting Master (Roche Molecular Diagnostics, Basel, Switzerland). Magnesium chloride (MgCl₂) (25 mM) was included at a concentration of 3 mM for both gp60 and MSE primer pair reactions, and 3.5 mM for MSB reactions. All inner reactions contained inner primer concentrations of 2 μM. 2 µl of genomic template DNA was added to each reaction.

All real-time PCR reactions were conducted under the following cycling conditions: initial denaturation at 95°C for 10 min, subsequent 3-step amplification for 35 cycles, including denaturation at 95°C for 30s, annealing at 64°C for 30s and extension at 72°C for 40s.
6.3.5  **DNA Sequence Analysis of First Round Amplicons**

All purified, first round gp60, MSB and MSE amplicons for the studied *C. hominis* isolates were sequenced bidirectionally via Sanger sequencing (Eurofins Genomics, Cologne, Germany).

6.3.6  **HRM Curve Acquisition and Analysis**

HRM analysis was conducted immediately post-PCR amplification on the LC96 by incrementally increasing the temperature by 2.2°C/s from 65°C to 97°C, taking 15 continuous readings/°C. Following high resolution melting-curve acquisition, a normalisation region of 79 to 88°C was applied to *C. hominis* isolates for analysis within the LC96 software (Roche Molecular Diagnostics, Basel, Switzerland). The positive/negative threshold was set to the default 0.05 relative fluorescence units (RFU). Delta T<sub>m</sub> discrimination and curve shape discrimination parameters were set to 100%. Normalised melting curve, differential melting peak, and difference plots were generated and analysed in order to determine the precise melting temperature of each *C. hominis* subtype. To be considered acceptable for HRM analysis, sample Cp values were to be less than 30 cycles, and inter-sample variation in Cp values were to vary by no more than 5 cycle units. Analysis was conducted on all duplicate runs.

6.3.7  **Evaluation of Assay Sensitivity and Specificity**

Assay sensitivity and specificity testing on genomic *Cryptosporidium* DNA, extracted from a faecal environment, was conducted for both outer and inner primers using enumerated *C. hominis* oocyst samples. Prior to extraction, oocysts within semi-purified faecal samples were microscopically enumerated with KOVA Glasstic Slides (Medical Supply Company, Dublin, Ireland). The *C. hominis* sample was determined to contain approximately 1 x 10<sup>5</sup> oocysts per ml.

DNA was extracted from oocyst samples following the Enteric Bio SPS extraction protocol (Serosep, Limerick, Ireland) employed routinely by CUH in *Cryptosporidium* DNA extractions. The whole tip of a FLOQswab (Copan, Italy) was lightly coated with the specimen, which was inoculated into a 4 mL SPS tube and heated at 103°C for 30 min. A 10-fold serial dilution, ranging from 1 x 10<sup>5</sup> oocysts per ml to 1 x 10<sup>0</sup> oocysts per ml, was prepared from the *C. hominis* genomic DNA extract. The limits of detection
for gp60, MSB and MSE outer and inner primer pairs were assessed using melting curve analysis and gel electrophoresis. All reactions were conducted in duplicate.

The specificity of C. hominis gp60, MSB and MSE primers sets were also evaluated using DNA extracts from varying enteric pathogens, of various genera, both bacterial and protozoan in nature. Salmonella, Shigella, Campylobacter, VTEC and G. lamblia positive faecal DNA extracts detected during routine molecular enteric screening in CUH were tested. Genomic DNA extracted from B. hominis cysts (ATCC, United States of America) employing the DNeasy Blood and Tissue Kit (Qiagen, German) as per the manufacturer’s instructions, was also evaluated. This genomic DNA was exposed to identical real-time PCR conditions and primer concentrations as those outlined above. All reactions were conducted in duplicate.

6.3.8 Reproducibility of HRM Analyses

Intra- and inter-experiment reproducibility was assessed by calculating the mean, standard deviation, and coefficients of variation of the duplicate T_m peaks generated for each gp60-subtype isolate by the HRM analysis software. To assess intra-experimental reproducibility, duplicate reactions were conducted for each sample within a single run. Inter-experimental variation was assessed in separate, duplicate runs of identical gp60-subtype sample composition.
6.4 Results

6.4.1 DNA Sequencing of Genetic Marker gp60-subtypes

Multiple alleles were identified via DNA sequencing at the gp60, MSB and MSE loci, respectively. The gp60-subtype designations of the six isolates analysed in this study were initially determined during a prior epidemiological study (as detailed in Chapter IV). Consequently, gp60-subtype numbers were designated based on gp60-subtype family designation (Ia, Ib, Id and Ie, respectively) and by ascending number of TCA/TCG/TCT and R repeats present within gp60-subtypes. HRM analysis groups were designated based on ascending amplicon T\textsubscript{m} value. The DNA sequences alignments of the heterogeneous regions within the gp60-subtypes, and MSB and MSE alleles encountered in this study are presented in Tables 6.3, 6.4 and 6.5, respectively.

DNA sequence analysis of the MSB and MSE amplicons identified three and four alleles associated with these loci, respectively, among the six isolates analysed. One isolate each possessed MSB alleles 1 and 5, respectively, while four possessed MSE allele 2. Two isolates each possessed MSE alleles 1 and 2, respectively, while MSE alleles 1 and 4 were present in a single isolate each, respectively.

This study included C. hominis gp60-subtypes from a variety of gp60-subtype families, which exhibit non-conserved regions within the target sequence. Therefore, at the gp60 locus, inter-family sequence variations both up- and downstream of the tandem repeat regions also contribute to differences in amplicon G+C content, T\textsubscript{m}, and, ultimately, to HRM analysis results.
DNA sequence alignments detailing *C. hominis* sequence variations within the tandem repeat region of the gp60 locus

<table>
<thead>
<tr>
<th>gp60</th>
<th>Nucleotide position (fragment length = 130 – 154 bp)</th>
</tr>
</thead>
</table>

*Table 6.3*
### Table 6.4 DNA sequence alignments detailing *C. hominis* sequence variations within the tandem repeat region of the MSB locus

<table>
<thead>
<tr>
<th>MSB</th>
<th>Nucleotide position (fragment length = 238 – 268 bp)</th>
<th>Abbreviations: Seq = Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq 1</td>
<td>G C G G G G A G A G</td>
<td>Sequence 1: laA14G3; Sequence 2: IbA9G3; Sequence 3: IdA10G2; Sequence 4: IbA12G3; Sequence 5: IdA21; Sequence 6: laA11G3T3</td>
</tr>
<tr>
<td>Seq 2</td>
<td>A T C T C T C T T</td>
<td></td>
</tr>
<tr>
<td>Seq 3</td>
<td>G T C T C T A</td>
<td></td>
</tr>
<tr>
<td>Seq 4</td>
<td>C T C T A T C C T T T</td>
<td></td>
</tr>
<tr>
<td>Seq 5</td>
<td>C T C T A T C T C T T T</td>
<td></td>
</tr>
<tr>
<td>Seq 6</td>
<td>G G G G G G T G A G</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Seq = Sequence

Sequence 1: laA14G3; Sequence 2: IbA9G3; Sequence 3: IdA21; Sequence 4: IbA12G3; Sequence 5: IdA21; Sequence 6: laA11G3T3

---

<table>
<thead>
<tr>
<th>MSB</th>
<th>Nucleotide position (fragment length = 238 – 268 bp)</th>
<th>Abbreviations: Seq = Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq 1</td>
<td>G C G G G G A G A G</td>
<td>Sequence 1: laA14G3; Sequence 2: IbA9G3; Sequence 3: IdA21; Sequence 4: IbA12G3; Sequence 5: IdA21; Sequence 6: laA11G3T3</td>
</tr>
<tr>
<td>Seq 2</td>
<td>A T C T C T C T T</td>
<td></td>
</tr>
<tr>
<td>Seq 3</td>
<td>G T C T C T A</td>
<td></td>
</tr>
<tr>
<td>Seq 4</td>
<td>C T C T A T C C T T T</td>
<td></td>
</tr>
<tr>
<td>Seq 5</td>
<td>C T C T A T C T C T T T</td>
<td></td>
</tr>
<tr>
<td>Seq 6</td>
<td>G G G G G G T G A G</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Seq = Sequence

Sequence 1: laA14G3; Sequence 2: IbA9G3, IdA21, laA11G3T3; Sequence 3: IbA10G2; Sequence 4: IbA12G
### Table 6.5 DNA sequence alignments detailing *C. hominis* sequence variations within the tandem repeat region of the MSE locus

<table>
<thead>
<tr>
<th>MSE</th>
<th>Nucleotide position (fragment length = 206 – 272 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>108</td>
</tr>
<tr>
<td>Seq 1</td>
<td>G</td>
</tr>
<tr>
<td>Seq 2</td>
<td>G</td>
</tr>
<tr>
<td>Seq 3</td>
<td>G</td>
</tr>
<tr>
<td>Seq 4</td>
<td>G</td>
</tr>
<tr>
<td>Seq 5</td>
<td>G</td>
</tr>
<tr>
<td>Seq 6</td>
<td>G</td>
</tr>
</tbody>
</table>

|     | 136 | 137 | 138 | 139 | 140 | 141 | 142 | 143 | 144 | 145 | 146 | 147 | 148 | 149 | 150 | 151 | 152 | 153 | 154 | 155 | 156 | 157 | 158 | 159 | 160 | 161 | 162 | 163 | 164 | 165 | 166 | 167 | 168 | 169 | 170 | 171 |
| Seq 1 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | C | A | A |
| Seq 2 | G | A | T | C | A | A | G | A | T | C | A | A | G | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | A | T | C | A | A |
| Seq 3 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | A | T | C | A | A |
| Seq 4 | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A |
| Seq 5 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | C | A | T | C | A | A |
| Seq 6 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |

|     | 172 | 173 | 174 | 175 | 176 | 177 | 178 | 179 | 180 | 181 | 182 | 183 | 184 | 185 | 186 | 187 | 188 | 189 | 190 | 191 | 192 | 193 | 194 | 195 | 196 | 197 | 198 | 199 | 200 | 201 | 202 | 203 | 204 | 205 | 206 | 207 |
| Seq 1 | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A |
| Seq 2 | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A |
| Seq 3 | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A |
| Seq 4 | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A |
| Seq 5 | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A | G | A | T | C | A | A |
| Seq 6 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | C | A | G | A | T | C | A | A |

Abbreviations: Seq = Sequence

Sequence 1: IaA14G3; Sequence 2: IbA9G3; Sequence 3: IbA10G2; Sequence 4: IbA12G3; Sequence 5: IdaA21; Sequence 6: IeA11G3T3
6.4.2 Interpretation of HRM Curve Profiles and Categorisation of C. hominis gp60-subtypes

The melting characteristics of the respective $gp60$-subtype amplicon groups generated for each of the three tandem repeat loci, $gp60$, MSB and MSE, were assessed through plots generated by the HRM analysis function of the LC96 software (Roche Molecular Diagnostics, Basel, Switzerland). Normalised melting curves, presenting data as relative fluorescent units versus temperature, are depicted for each of the three loci in Figures 6.1, 6.4 and 6.7, respectively. They include normalised melting curve data for each $gp60$-subtype. In these plots, pre- and post-melt signals were uniformly set to relative values of 100% and 0%, respectively, to aid interpretation and analysis of the data.

Normalised melting peaks, which are based upon the first negative derivative of the change in fluorescence over time ($-dF/dT$) versus temperature, are depicted in Figures 6.2, 6.5 and 6.8, respectively, and represent the exact $T_m$ of each $gp60$-subtype, visually depicting the spread of melting temperature exhibited by each $gp60$-subtype grouping. Finally, difference plots, which are based upon the difference in fluorescence versus temperature, are observed in Figures 6.3, 6.6 and 6.9, and depict melting curve data following subtraction of baseline curves in order to better highlight distinct $gp60$-subtype groupings.
Figure 6.1  Normalised melting curves of *C. hominis* gp60 subtypes at the gp60 locus

Figure 6.2  Normalised melting peaks of *C. hominis* gp60 subtypes at the gp60 locus
Six *C. hominis* gp60-subtypes were analysed in this study and HRM analysis of the gp60 locus was capable of differentiating between four of these six gp60-subtypes. IaA14R3, IbA9G3, IbA10G2 and IbA12G3 exhibited distinct melting curves, melting peaks and difference plot characteristics, as demonstrated in Figures 6.1, 6.2 and 6.3, respectively. However, as outlined in these HRM analysis plots and Table 6.6, gp60-subtypes IdA21 and IeA11G3T3 exhibited overlapping melting temperatures of between 82.40 °C and 82.56 °C. Consequently, these gp60-subtypes could not be differentiated at this locus.

Three groupings were established, corresponding to three of the four alleles that were detected at the MSB locus. These groupings are shown in Figures 6.4, 6.5 and 6.6. IbA10G2 and IaA9G3 exhibited distinct melting temperature ranges of 78.21°C, and 78.34°C to 78.41°C, respectively. DNA sequence analysis of these two isolates found them to possess distinct MSB gp60-subtype alleles. The remaining gp60-subtype
isolates were grouped together, exhibiting a melting temperature range from 78.21 °C to 78.34°C, which lay between those that defined the distinct IbA9G3 and IbA10G2 groupings. Three of the four isolates that comprised this HRM analysis grouping were found, by DNA sequence analysis, to possess the same MSB amplicon sequence, also distinct from those found in IaA9G3 and IbA10G2, respectively. The gp60-subtype IaA12G3 possessed a distinct MSB amplicon sequence that could not be resolved from that of the remaining MSB HRM analysis Group 2 isolates. Both IdA21 and IeA11G3T3 possessed the same MSB allele. Therefore, as with the gp60 locus, differentiation between these two gp60-subtypes was also not possible at the MSB locus. The MSB sequence and melting characteristics of each amplicon are outlined in Table 6.6.

![Figure 6.4](image)

**Figure 6.4** Normalised melting curves of *C. hominis* gp60- subtypes at the MSB locus
Figure 6.5  Normalised melting peaks of the *C. hominis* gp60- subtypes at the MSB locus

Figure 6.6  Difference plot of the *C. hominis* gp60- subtypes at the MSB locus
Four HRM analysis groupings were determined among the six studied isolates at the MSE locus. These groupings are also shown in Figures 6.7, 6.8 and 6.9. However, DNA sequence analysis identified 6 MSE alleles. Two groupings, Group 1 and Group 2 contained two gp60-subtypes each, IaA9G3 and IbA12G3, and IbA10G2 and IdA21, respectively. Within Group 1, IaA9G3 contained two fewer repeat units (TCAAGA) than IbA12G3. Likewise within Group 2, gp60-subtype IbA10G2 contained one less MSE repeat unit than IdA21. Groups 3 and 4 contained the distinct gp60-subtypes, IaA14R3 and IeA11G3T3, respectively. It is evident that the differences in DNA sequence between the isolates within Groups 1 and 2, respectively, were not substantial enough to be differentiated via HRM analysis of the MSE locus alone. The sequence and melting characteristics of each MSE allele amplicon are presented in Table 6.6. The respective HRM analysis based groupings and DNA sequence designations for each gp60-subtype at each locus are detailed in Table 6.7.

![Figure 6.7](image)

**Figure 6.7** Normalised melting curves of *C. hominis* gp60- subtypes at the MSE locus
Figure 6.8  Normalised melting peaks of the *C. hominis* gp60- subtypes at the MSE locus

Figure 6.9  Difference plot of *C. hominis* gp60- subtypes at the MSE locus
Table 6.6  *gp60*-subtype groupings and amplicon characteristics for studied *C. hominis* isolates

<table>
<thead>
<tr>
<th>Group</th>
<th><em>gp60</em>-subtypes</th>
<th>%G+C Content</th>
<th>Sequence Length (bp)</th>
<th>Upper T&lt;sub&gt;m&lt;/sub&gt; Limit (°C)</th>
<th>Lower T&lt;sub&gt;m&lt;/sub&gt; Limit (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>gp60</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>IdA21</td>
<td>46.20%</td>
<td>158</td>
<td>82.40</td>
<td>82.49</td>
</tr>
<tr>
<td></td>
<td>IeA11G3T3</td>
<td>47.20%</td>
<td>142</td>
<td>82.40</td>
<td>82.56</td>
</tr>
<tr>
<td>Group 2</td>
<td>IaA14R3</td>
<td>47.10%</td>
<td>136</td>
<td>83.06</td>
<td>83.15</td>
</tr>
<tr>
<td>Group 3</td>
<td>IbA10G2</td>
<td>49.20%</td>
<td>130</td>
<td>83.26</td>
<td>83.34</td>
</tr>
<tr>
<td>Group 4</td>
<td>IbA12G3</td>
<td>48.90%</td>
<td>139</td>
<td>83.32</td>
<td>83.41</td>
</tr>
<tr>
<td>Group 5</td>
<td>IbA9G3</td>
<td>50.00%</td>
<td>130</td>
<td>83.65</td>
<td>83.74</td>
</tr>
<tr>
<td><strong>MSB</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>IbA10G2</td>
<td>36.13%</td>
<td>238</td>
<td>78.21</td>
<td>78.21</td>
</tr>
<tr>
<td></td>
<td>IaA14R3</td>
<td>35.94%</td>
<td>256</td>
<td>78.27</td>
<td>78.34</td>
</tr>
<tr>
<td></td>
<td>IbA12G3</td>
<td>36.40%</td>
<td>250</td>
<td>78.28</td>
<td>78.34</td>
</tr>
<tr>
<td></td>
<td>IdA21</td>
<td>36.33%</td>
<td>256</td>
<td>78.27</td>
<td>78.28</td>
</tr>
<tr>
<td></td>
<td>IeA11G3T3</td>
<td>36.33%</td>
<td>256</td>
<td>78.21</td>
<td>78.28</td>
</tr>
<tr>
<td>Group 3</td>
<td>IbA9G3</td>
<td>36.57%</td>
<td>268</td>
<td>78.34</td>
<td>78.41</td>
</tr>
<tr>
<td><strong>MSE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>IbA9G3</td>
<td>36.03%</td>
<td>272</td>
<td>79.32</td>
<td>79.41</td>
</tr>
<tr>
<td></td>
<td>IbA12G3</td>
<td>35.92%</td>
<td>284</td>
<td>79.32</td>
<td>79.35</td>
</tr>
<tr>
<td>Group 2</td>
<td>IbA10G2</td>
<td>36.44%</td>
<td>236</td>
<td>79.45</td>
<td>79.55</td>
</tr>
<tr>
<td></td>
<td>IdA21</td>
<td>36.52%</td>
<td>230</td>
<td>79.51</td>
<td>79.55</td>
</tr>
<tr>
<td>Group 3</td>
<td>IaA14R3</td>
<td>36.61%</td>
<td>224</td>
<td>79.58</td>
<td>79.61</td>
</tr>
<tr>
<td>Group 4</td>
<td>IeA11G3T3</td>
<td>36.89%</td>
<td>206</td>
<td>79.65</td>
<td>79.68</td>
</tr>
</tbody>
</table>
Table 6.7 Multi-locus HRM analysis grouping and allele designations of *C. hominis* gp60-subtype isolates

<table>
<thead>
<tr>
<th>Isolate</th>
<th>gp60 HRM Analysis Grouping</th>
<th>Sequence Designation</th>
<th>MSB HRM Analysis Grouping</th>
<th>Sequence Designation</th>
<th>MSE HRM analysis grouping</th>
<th>Sequence Designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>IaA14R3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>IbA9G3</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>IbA10G2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>IbA12G3</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>IdA21</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>IeA11G3R3</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

6.4.3 Detection Sensitivity and Specificity of *C. hominis* Tandem Repeat Locus Targets

The evaluation of detection sensitivity of the *C. hominis* outer and inner primer pair sets was conducted on 10-fold serial dilutions of genomic DNA extracted from enumerated *C. hominis* oocysts. The dilution series spanned a range of $10^5$ to $10^0$ oocysts per ml. Although extracted from semi-purified faecal samples, the complex microbial nature of such samples, and the lack of commercially available pure genomic DNA *C. hominis* samples precluded the ability to accurately determine the exact genomic DNA concentration. Thus, DNA concentration is given in terms of serially diluted, enumerated oocysts. The LOD of the *C. hominis* gp60 outer and inner primer sets was determined to be $10^4$ and $10^2$ oocysts per ml, respectively. The LOD of the *C. hominis* MSB outer and inner primer sets was $10^3$ and $10^3$ oocysts per ml, respectively, while the LOD for the *C. hominis* MSE outer and inner primer sets was $10^4$ oocysts per ml.

Clinical faecal samples that were determined to not contain evidence of infection by bacterial, fungal or parasitic gastrointestinal pathogens were tested to ascertain the specificity of gp60, MSB and MSE outer and inner primer pairs, respectively. Additionally, a range of samples known to contain prokaryotic (bacterial) organisms and eukaryotic (protozoan) organisms, as outlined previously, were also tested. No unintended amplification was detected in these samples.
Appendix III details the real-time PCR and gel electrophoresis data in support of the results outlined above.

6.4.4 Assessment of Intra- and Inter-assay Variabilities

Upon confirmation of the sensitivity and specificity of assay primer sets and PCR conditions, the melting characteristics of the *C. hominis* gp60, MSB and MSE amplicons were assessed. Critical evaluations of intra-assay and inter-assay variabilities were conducted by calculating the mean, standard deviation and coefficient of variation of melting temperatures obtained during, within and between run duplicates for each of the six *C. hominis* isolates studied.

The coefficient of variation among and within experiments was determined to be < 0.5%, indicating an acceptable level of reproducibility. The results of this critical evaluation are outlined in Table 6.8.
Table 6.8  Intra- and inter-assay reproducibility of *C. hominis* isolate HRM analysis of the *gp60*, MSB and MSE tandem repeat loci

<table>
<thead>
<tr>
<th><em>C. hominis</em> gp60-subtype</th>
<th><em>gp60</em></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intra-assay Reproducibility</td>
<td>Inter-assay Reproducibility</td>
<td>Intra-assay Reproducibility</td>
<td>Inter-assay Reproducibility</td>
<td>Intra-assay Reproducibility</td>
<td>Inter-assay Reproducibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>CV</td>
<td>Mean T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>CV</td>
<td>Mean T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>CV</td>
<td>Mean T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>CV</td>
<td>Mean T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>CV</td>
<td>Mean T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>CV</td>
</tr>
<tr>
<td>IdA21</td>
<td>82.51 ± 0.00</td>
<td>0.00%</td>
<td>82.45 ± 0.05</td>
<td>0.06%</td>
<td>78.28 ± 0.00</td>
<td>0.00%</td>
<td>78.28 ± 0.01</td>
<td>0.01%</td>
<td>79.51 ± 0.00</td>
<td>0.00%</td>
<td>79.53 ± 0.03</td>
<td>0.04%</td>
</tr>
<tr>
<td>IeA11G3T3</td>
<td>82.44 ± 0.00</td>
<td>0.00%</td>
<td>83.48 ± 0.08</td>
<td>0.09%</td>
<td>78.25 ± 0.05</td>
<td>0.06%</td>
<td>78.28 ± 0.01</td>
<td>0.01%</td>
<td>79.62 ± 0.05</td>
<td>0.06%</td>
<td>79.67 ± 0.02</td>
<td>0.03%</td>
</tr>
<tr>
<td>IaA14R3</td>
<td>83.19 ± 0.04</td>
<td>0.05%</td>
<td>83.12 ± 0.05</td>
<td>0.06%</td>
<td>78.34 ± 0.00</td>
<td>0.00%</td>
<td>78.31 ± 0.05</td>
<td>0.06%</td>
<td>79.58 ± 0.00</td>
<td>0.00%</td>
<td>79.60 ± 0.02</td>
<td>0.03%</td>
</tr>
<tr>
<td>IbA10G2</td>
<td>83.39 ± 0.05</td>
<td>0.06%</td>
<td>83.31 ± 0.04</td>
<td>0.05%</td>
<td>78.31 ± 0.04</td>
<td>0.05%</td>
<td>78.21 ± 0.00</td>
<td>0.00%</td>
<td>79.48 ± 0.04</td>
<td>0.05%</td>
<td>79.56 ± 0.07</td>
<td>0.09%</td>
</tr>
<tr>
<td>IbA12G3</td>
<td>83.45 ± 0.04</td>
<td>0.05%</td>
<td>83.40 ± 0.01</td>
<td>0.01%</td>
<td>78.31 ± 0.05</td>
<td>0.06%</td>
<td>78.31 ± 0.04</td>
<td>0.05%</td>
<td>79.32 ± 0.00</td>
<td>0.00%</td>
<td>79.34 ± 0.02</td>
<td>0.03%</td>
</tr>
<tr>
<td>IbA9G3</td>
<td>83.68 ± 0.10</td>
<td>0.11%</td>
<td>83.68 ± 0.05</td>
<td>0.06%</td>
<td>78.41 ± 0.00</td>
<td>0.00%</td>
<td>78.38 ± 0.05</td>
<td>0.06%</td>
<td>79.35 ± 0.04</td>
<td>0.05%</td>
<td>79.40 ± 0.02</td>
<td>0.03%</td>
</tr>
</tbody>
</table>
6.5 Discussion

This is the first study to investigate the application of HRM analysis to the subtypic differentiation of *C. hominis* isolates. Heretofore, studies have employed HRM analysis in order to differentiate *Cryptosporidium* spp. from other apicomplexan parasites, such as *Toxoplasma gondii*, *Sarcocystis* spp. and *Neospora* spp. (Fehlberg et al., 2017), to identify and differentiate between *Cryptosporidium* isolates to the species level, and to differentiate between *C. cuniculus* Va and Vb subtype families (Chelbi et al., 2018; Hadfield and Chalmers, 2012; Pangasa et al., 2009). The method developed in this study was capable of differentiating between *C. hominis* gp60-subtypes IaA14R3, IbA9G3, IbA10G2, IbA12G3, IdA21 and IeA11G3T3 using real-time PCR based interrogation of the gp60, MSB and MSE loci. This novel HRM analysis application has successfully resulted in the development of a new multi-locus subtyping tool capable of differentiating between a series of globally prevalent *C. hominis* gp60-subtypes.

Among the six gp60-subtypes tested, HRM analysis identified five Tm groupings at the gp60 locus. Four of the six gp60-subtypes displayed unique melting temperature characteristics, with the exception of IdA21 and IeA11G3T3, which exhibited overlapping Tm ranges (Table 6.6). Tm is a composite characteristic dictated by amplicon %G+C content, length and DNA sequence (Tong and Giffard, 2012). The gp60 tandem repeat unit is generally followed downstream by a polymorphic region, with this region specific to gp60-subtype family designation (Chalmers et al., 2017). The DNA sequences both up- and downstream of the gp60 tandem repeat region exhibited familial conservation within all studied Ib family gp60-subtypes. Therefore, only variations within the tandem repeat region contributed to altering the %G+C content and length of these amplicons, resulting in discreet, incremental differences in amplicon characteristics. Among the Ia, Id and Ie gp60-subtypes, the region proceeding the tandem repeat region was conserved, although differing slightly in sequence to that observed among Ib family gp60-subtypes. The region downstream of the gp60 tandem repeat unit exhibited inter-familial polymorphism within the Ia, Id and Ie gp60-subtypes, respectively. These gp60-subtype family specific, downstream polymorphisms also differed from the conserved flanking sequence observed among all the three Ib gp60-subtypes.

The length of individual amplicons significantly impacted upon HRM analysis based determination of and differentiation between Tm values in this study. %G+C content
exhibits an asymptotic relationship with amplicon length, with the impact of increasing %G+C content made increasingly negligible as amplicon length also increases (Tong and Giffard, 2012). Both *gp60*-subtypes IdA21 and IeA11G3T3 exhibited the longest *gp60* amplicon lengths observed among the studied *gp60*-subtype panel, reducing the impact of amplicon %G+C content. Relatedly, both of these *gp60*-subtypes also exhibited the lowest *gp60* amplicon %G+C values observed within the study. Thus, despite exhibiting *gp60*-subtype family specific polymorphisms within the amplicon sequence, these variations contributed little to observed T_m values. It is evident that the asymptotic relationship existing between %G+C content and T_m resulted in inconclusive, overlapping T_m characteristics between these *gp60*-subtypes that could not be definitively resolved by HRM analysis as amplicon length increased. The *gp60*-subtype or grouping specific T_m ranges observed in this study were the result of the dynamic between amplicon length and %G+C content, as determined by tandem repeat region sequence variations and subtype familial specific polymorphisms.

HRM analysis successfully differentiated between two of the four detected alleles among tested isolates at the MSB locus. However, given that only two *gp60*-subtypes, IbA9G3 and IbA10G2, possessed distinct alleles, the remaining four isolates could not be differentiated at this locus, including IdA21 and IeA11G3T3, which were also assigned to the same *gp60* HRM analysis grouping.

The issue of overlapping T_m values exhibited by genetically distinct *gp60*-subtypes was also observed at the MSB locus. Both the up- and downstream regions flanking the tandem repeat region within all MSB amplicons were conserved. Therefore, given the amplicon length range determined for this locus, the single tandem repeat difference between MSB alleles 1 and 4 did not confer a T_m difference detectable by HRM analysis. Therefore, although of some use, the utility of the MSB locus was limited by the number of alleles observed among the *gp60*-subtypes analysed in this study. Therefore, this locus warrants further study with an expanded *gp60*-subtype panel.

The four T_m groupings identified by HRM analysis at the MSE locus did not correspond directly with the six alleles identified by DNA sequencing of the *gp60*- subtype panel. The *gp60*-subtypes IbA9G3 and IbA12G3, assigned to Group 1, differed by two sequence repeats, while Group 2 isolates, IbA10G2 and IdA21, differed by a single sequence repeat. As seen previously at both the *gp60* and MSB loci, the lengthier MSE
amplicons within Groups 1 and 2, which ranged from 230 to 284 bp in length, negated the impact of %G+C content on amplicon T_m and impeded the ability of HRM analysis to differentiate between gp60-subtypes within these allelic groupings. However, gp60-
subtypes IaA14R3 and IeA11G3T1 were assigned to separate T_m groupings, Groups 3 and 4, respectively, at the MSE locus, thereby enabling successful differentiation of all studied C. hominis gp60-subtypes via multi-locus HRM analysis.

The coefficient of variation of C. hominis subtypes among and between experiments for all studied loci was less than 0.5%, as outlined in Table 6.8, revealing the method to be highly reproducible. Inner primer real-time PCR conditions within this method were designed to allow for concurrent real-time PCR amplification and HRM analysis of all three loci. These results indicate that, with expansion of the panel of genetic markers to a size reflective of those employed in standard MLST and MLVA studies (Amer et al., 2013; Tanriverdi et al., 2008), MLVA-HRM analysis has the potential to provide a reliable, rapid, cost-effective alternative to such DNA sequencing based multi-locus analysis methods.

This method is intended for use as a subtype discriminatory tool following species identification. In this study, a probe based real-time PCR assay targeting the 18S rRNA gene was used to differentiate between C. parvum and C. hominis samples prior to HRM analysis (Mary et al., 2013). A separate, but complementary, discriminatory panel for C. parvum was discussed in Chapter V. In the event of mixed species infection, both discriminatory panels would be employed to produce data based on their respective target species within a single sample. In the event of mixed infections of the same species of differing gp60-subtype designations, the presence of dual peaks within the amplification curves and melting peaks of a locus for a single sample would be indicative of the presence of mixed gp60-subtype infection. The panel is intended to be run simultaneously with a control panel of regionally prevalent C. hominis and/or C. parvum gp60-subtypes to aid in HRM analysis data interpretation and identification.

This method, however, is not without issue and requires further refinement, including analysis of a wider range of C. hominis gp60-subtype families and individual gp60-
subtypes, and expansion of the multi-locus panel to include a wider array of discriminatory loci. This study also highlighted the somewhat counterintuitive impact of increasing amplicon length, which diminishes the discriminative utility of this method
when tandem repeat variations do not significantly impact the %G+C content of larger amplicons.

Future research targeting a wider range of genetic markers, while also limiting amplicon sizes to ensure maximal discriminatory efficiency of the method is also required. Selected markers will be located across a wider variety of chromosomes than those included in this study in order to exclude genetic linkage (Widmer and Sullivan, 2012). The ability of this method to adequately resolve clonal expansion within *C. hominis* IbA10G2 populations also requires further investigation. If warranted, recently reported single nucleotide allele (SNV) containing loci, capable of improved resolution between IbA10G2 isolates, may be incorporated (Beser et al., 2017).

Overall, this study highlights the viability of HRM analysis as a means to *gp60*-subtype *C. hominis* isolates. It is pertinent to note that a paradigm shift is presently underway in the field of clinical enteric parasitology, with the supersession of the current diagnostic gold standard, microscopy, by commercially available molecular diagnostic methods, including a wide variety of real-time PCR based panels (Ryan et al., 2017). The compatibility of HRM analysis with real-time PCR based diagnostic panels and with established clinical laboratory infrastructure and protocols, further increases the viability of this multi-locus based method as an initial epidemiological screening tool. DNA sequencing could then provide further resolution of case linking, if warranted.

Multi-locus HRM analysis subtyping may provide a streamlined, low discriminatory screening tool capable of rapid characterisation of *C. hominis* isolates and greater epidemiologic case linking, which, when employed in conjunction with a panel of control or reference isolates, could significantly advance our understanding of regional *C. hominis* populations implicated in clinical cryptosporidiosis cases, particularly in countries lacking standardised epidemiological protocols or surveillance.

Appendix V provides an addendum to this chapter wherein statistical analyses of *gp60*, MSB and MSE HRM analysis groupings and $T_m$ data are presented and discussed.
Chapter VII: Thesis Summary
7.1 Concluding Remarks
The endemicity of Cryptosporidium spp. in Ireland remains a considerable Public Health challenge. Despite the prevalence of this enteric pathogen, however, clinical Cryptosporidium diagnoses have primarily relied upon microscopic methods for the past several decades. Progression of the field towards molecular detection platforms has been slow. Consequently, despite the national prevalence of cryptosporidiosis, even with the introduction of real-time PCR diagnostic platforms, national epidemiological reports are limited to genus level temporal and spatial analyses. Additionally, detailed molecular epidemiology studies of clinical cases are rarely undertaken in Ireland.

The fundamental objective of this study was to validate and implement a multiplex real-time PCR based diagnostic method into routine clinical practice, analyse the epidemiological output from this diagnostic platform and, ultimately, develop a functionally compatible subtyping method. The initial validation study, which involved screening 352 clinical samples, both microscopically and molecularly, found 100% concordance between the sensitivity and specificity of real-time PCR and fluorescent auramine-O and acid-fast staining microscopy (Chapter II). Real-time PCR subsequently superseded microscopy in routine Cryptosporidium testing in this laboratory (Medical Microbiology Department, CUH). This transition was informed, in part, by the research conducted in this thesis.

This body of research provides contemporary insight into the epidemiological breakdown of Cryptosporidium species and gp60-subtypes in Ireland. Cryptosporidium parvum was determined to account for 86.5% of cases, while 13.5% were attributed to Cryptosporidium hominis (Chapter III). Incidence of C. parvum and C. hominis was found to be consistent with previously published studies and with annual reports released by the Irish notifiable infectious disease surveillance body, the Health Protection Surveillance Centre (HPSC, 2019; Zintl et al., 2011, 2009).

Most significantly, in addition to previously reported C. parvum and C. hominis gp60-subtypes, further gp60-subtyping analysis identified the newly reported emergence of one C. parvum and four C. hominis gp60-subtypes within the national Cryptosporidium population (Chapter IV). These results highlight the diversification of an already highly heterogeneous population. This heterogeneity is predominantly observed among C. parvum, with 12 gp60-subtypes described in this study. C. parvum also accounts for
the majority of cases in Ireland (Zintl et al., 2011, 2009). The potential for genetic recombination within and between Cryptosporidium spp., and the selective pressures known to be exerted on the gp60 locus may contribute to explaining the existing diversity and recent emergence of new C. parvum gp60-subtypes in Ireland (Widmer, 2009; Abal-Fabeiro et al., 2013; Guo et al., 2015; Widmer et al., 2015; Garcia-R and Hayman, 2017).

Until the present study only one C. hominis gp60-subtype, IbA10G2, which is reported to frequently exhibit a clonal population structure, had been reported in Ireland. The identification of four regionally novel gp60-subtypes, in addition to IbA10G2, is a significant development (Chapter IV). It is unlikely that these gp60-subtypes emerged through genetic recombination, as the reported novel gp60-subtypes belonged to a variety of distinct C. hominis allele families. Genetic recombination is also more restricted in C. hominis than among C. parvum gp60-subtypes, usually confined to IbA10G2 (Feng et al., 2018). This diversification is more likely to have been caused by introduction via anthroponotic or zoonotic transmission of gp60-subtypes introduced to Ireland from abroad.

The need for a rapid molecular method with which Irish clinical microbiology laboratories can subtype clinical isolates is apparent, particularly in light of the frequency of Cryptosporidium outbreaks in Ireland. The concept of such a rapid, first-line molecular subtyping tool for use in routine clinical cryptosporidiosis diagnosis provided the basis for the real-time PCR based high resolution melting (HRM) analysis method developed during the course of this research (Chapters V and VI). DNA sequence analysis of the gp60 gene currently remains the current mainstay in Cryptosporidium spp. subtype determination. However, DNA sequencing of routine samples currently remains unfeasible for many clinical microbiology laboratories. Consequently, when clinically warranted, Irish Cryptosporidium isolates are genotyped by the CRU (Swansea, Wales).

The hypothesis upon which this research was based proposed that differences conferred by sequence variation within a tandem repeat locus known to vary between distinct Cryptosporidium spp. gp60-subtypes would be detectable via real-time PCR based HRM analysis. Several loci per Cryptosporidium sp. were analysed to increase the discriminatory power of the method (Chapter V). The gp60, MM5 and MS-9 Mallon
loci interrogated among *C. parvum* subtypes fully differentiated 10 of the 18 subtypes studied. The remaining eight *gp60*-subtypes were defined into distinct allelic grouping based on melting temperature similarities. This provided lower level discrimination between these subtypes, resolving to a level of two to three potential groupings. The wealth of variation in the Irish *C. parvum* population proved challenging. Variations between the eight unresolved subtypes were not substantial enough to be differentiated at these three loci alone.

Future avenues of study for this method would involve expansion of the panel through the identification of further tandem repeat loci. This would improve the discriminatory power of the method allowing for differentiation of a broader scope of *C. parvum* subtypes. Potential loci to be investigated include, MM18, MM19 and TR14, among others, which have all reported in the literature in DNA sequencing MLVA-MLST panels (Chalmers et al., 2017; Robinson and Chalmers, 2012). The %G+C content and number of previously reported locus alleles reported in the literature would play a significant role in the selection of the most discriminatory loci.

HRM analysis based interrogation of the *gp60*, MSB and MSE loci among *C. hominis* subtypes fully differentiated between all of the isolates studied (*Chapter VI*). A more limited spectrum of *C. hominis* *gp60*-subtypes are encountered in Ireland than those observed among *C. parvum*. The existing method developed by this research would therefore be highly applicable to this species identification in Irish clinical laboratories. Further refinement of this method and expansion of the locus panel would increase the scope of *C. hominis* subtype identification, allowing for the use of this application in countries which exhibit broader variation in their *C. hominis* subtype population.

The results of this thesis have demonstrated that HRM analysis is capable of differentiation between *Cryptosporidium* subtypes. Further refinement of this method would result in a comprehensive subtyping tool compatible with many molecular platforms currently employed in routine clinical diagnosis of cryptosporidiosis. The incorporation of such a tool to routine testing would provide a means of accessing heretofore unobtainable epidemiological knowledge pertaining to the incidence of *Cryptosporidium* in Ireland. This method would also assist in outbreak identification and patient management.
Appendices
Appendix I – Chapter III Supplementary Data

A1.1 Real-time PCR Species Identification Results of Clinical Cryptosporidium DNA Samples

The data below (Table A1.1) outlines the species designation of all Cryptosporidium isolates identified in routine diagnostic testing in CUH from August 2015 to August 2018. Species identification was achieved through fluorescent probe based real-time PCR targeting the 18S rRNA gene in Cryptosporidium spp. These data are summarised in Section 3.4.1, and discussed in Section 3.5 of Chapter III.
Table A1.1 Molecular species identification results and average real-time PCR Cp values obtained from duplicate reactions conducted on all studied clinical isolates

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>18S rRNA gene Species Identification Result</th>
<th>Average Cp Value</th>
<th>Sample Number</th>
<th>18S rRNA gene Species Identification Result</th>
<th>Average Cp Value</th>
<th>Sample Number</th>
<th>18S rRNA gene Species Identification Result</th>
<th>Average Cp Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRYPTO_001/15</td>
<td>C. parvum</td>
<td>30.59</td>
<td>CRYPTO_001/16</td>
<td>C. parvum</td>
<td>26.39</td>
<td>CRYPTO_017/16</td>
<td>C. parvum</td>
<td>34.21</td>
</tr>
<tr>
<td>CRYPTO_002/15</td>
<td>C. hominis</td>
<td>29.19</td>
<td>CRYPTO_002/16</td>
<td>C. hominis</td>
<td>28.62</td>
<td>CRYPTO_018/16</td>
<td>C. parvum</td>
<td>31.01</td>
</tr>
<tr>
<td>CRYPTO_003/15</td>
<td>C. parvum</td>
<td>32.23</td>
<td>CRYPTO_003/16</td>
<td>C. parvum</td>
<td>29.86</td>
<td>CRYPTO_019/16</td>
<td>C. parvum</td>
<td>33.93</td>
</tr>
<tr>
<td>CRYPTO_004/15</td>
<td>C. parvum</td>
<td>29.41</td>
<td>CRYPTO_004/16</td>
<td>C. parvum</td>
<td>35.80</td>
<td>CRYPTO_020/16</td>
<td>C. parvum</td>
<td>29.06</td>
</tr>
<tr>
<td>CRYPTO_005/15</td>
<td>C. hominis</td>
<td>32.65</td>
<td>CRYPTO_005/16</td>
<td>C. parvum</td>
<td>30.21</td>
<td>CRYPTO_021/16</td>
<td>C. parvum</td>
<td>32.05</td>
</tr>
<tr>
<td>CRYPTO_006/15</td>
<td>C. hominis</td>
<td>29.24</td>
<td>CRYPTO_006/16</td>
<td>C. parvum</td>
<td>32.34</td>
<td>CRYPTO_022/16</td>
<td>C. parvum</td>
<td>33.57</td>
</tr>
<tr>
<td>CRYPTO_007/15</td>
<td>C. parvum</td>
<td>35.87</td>
<td>CRYPTO_007/16</td>
<td>C. parvum</td>
<td>32.03</td>
<td>CRYPTO_023/16</td>
<td>C. parvum</td>
<td>26.28</td>
</tr>
<tr>
<td>CRYPTO_008/15</td>
<td>C. parvum</td>
<td>29.47</td>
<td>CRYPTO_008/16</td>
<td>C. parvum</td>
<td>30.08</td>
<td>CRYPTO_024/16</td>
<td>C. parvum</td>
<td>34.18</td>
</tr>
<tr>
<td>CRYPTO_009/15</td>
<td>C. parvum</td>
<td>27.13</td>
<td>CRYPTO_009/16</td>
<td>C. parvum</td>
<td>27.87</td>
<td>CRYPTO_025/16</td>
<td>C. parvum</td>
<td>32.58</td>
</tr>
<tr>
<td>CRYPTO_010/15</td>
<td>C. parvum</td>
<td>31.42</td>
<td>CRYPTO_010/16</td>
<td>C. parvum</td>
<td>29.63</td>
<td>CRYPTO_026/16</td>
<td>C. parvum</td>
<td>32.44</td>
</tr>
<tr>
<td>CRYPTO_011/15</td>
<td>C. parvum</td>
<td>33.66</td>
<td>CRYPTO_011/16</td>
<td>C. parvum</td>
<td>31.87</td>
<td>CRYPTO_027/16</td>
<td>C. parvum</td>
<td>29.40</td>
</tr>
<tr>
<td>CRYPTO_012/15</td>
<td>C. parvum</td>
<td>29.35</td>
<td>CRYPTO_012/16</td>
<td>C. parvum</td>
<td>27.90</td>
<td>CRYPTO_028/16</td>
<td>C. parvum</td>
<td>33.64</td>
</tr>
<tr>
<td>CRYPTO_013/16</td>
<td>C. parvum</td>
<td>32.43</td>
<td>CRYPTO_013/16</td>
<td>C. parvum</td>
<td>32.43</td>
<td>CRYPTO_029/16</td>
<td>C. parvum</td>
<td>29.76</td>
</tr>
<tr>
<td>CRYPTO_014/16</td>
<td>C. parvum</td>
<td>30.31</td>
<td>CRYPTO_014/16</td>
<td>C. parvum</td>
<td>30.31</td>
<td>CRYPTO_030/16</td>
<td>C. parvum</td>
<td>35.07</td>
</tr>
<tr>
<td>CRYPTO_015/16</td>
<td>C. parvum</td>
<td>32.18</td>
<td>CRYPTO_015/16</td>
<td>C. parvum</td>
<td>32.18</td>
<td>CRYPTO_031/16</td>
<td>C. parvum</td>
<td>30.50</td>
</tr>
<tr>
<td>CRYPTO_016/16</td>
<td>C. parvum</td>
<td>29.74</td>
<td>CRYPTO_016/16</td>
<td>C. parvum</td>
<td>29.74</td>
<td>CRYPTO_032/16</td>
<td>C. parvum</td>
<td>31.64</td>
</tr>
<tr>
<td>Sample Number</td>
<td>18S rRNA gene Species Identification Result</td>
<td>Average Cp Value</td>
<td>Sample Number</td>
<td>18S rRNA gene Species Identification Result</td>
<td>Average Cp Value</td>
<td>Sample Number</td>
<td>18S rRNA gene Species Identification Result</td>
<td>Average Cp Value</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------</td>
<td>------------------</td>
<td>---------------</td>
<td>---------------------------------------------</td>
<td>------------------</td>
<td>---------------</td>
<td>---------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>CRYPTO_033/16</td>
<td>C. parvum</td>
<td>30.12</td>
<td>CRYPTO_001/17</td>
<td>C. parvum</td>
<td>29.37</td>
<td>CRYPTO_021/17</td>
<td>C. parvum</td>
<td>30.45</td>
</tr>
<tr>
<td>CRYPTO_034/16</td>
<td>C. parvum</td>
<td>33.04</td>
<td>CRYPTO_002/17</td>
<td>C. parvum</td>
<td>31.41</td>
<td>CRYPTO_022/17</td>
<td>C. parvum</td>
<td>27.53</td>
</tr>
<tr>
<td>CRYPTO_035/16</td>
<td>C. hominis</td>
<td>33.86</td>
<td>CRYPTO_003/17</td>
<td>C. parvum</td>
<td>32.21</td>
<td>CRYPTO_023/17</td>
<td>C. parvum</td>
<td>28.99</td>
</tr>
<tr>
<td>CRYPTO_036/16</td>
<td>C. parvum</td>
<td>34.78</td>
<td>CRYPTO_004/17</td>
<td>C. parvum</td>
<td>29.65</td>
<td>CRYPTO_024/17</td>
<td>C. parvum</td>
<td>33.7</td>
</tr>
<tr>
<td>CRYPTO_037/16</td>
<td>C. hominis</td>
<td>34.94</td>
<td>CRYPTO_005/17</td>
<td>C. parvum</td>
<td>29.80</td>
<td>CRYPTO_025/17</td>
<td>C. parvum</td>
<td>28.11</td>
</tr>
<tr>
<td>CRYPTO_038/16</td>
<td>C. parvum</td>
<td>34.88</td>
<td>CRYPTO_006/17</td>
<td>C. parvum</td>
<td>31.82</td>
<td>CRYPTO_026/17</td>
<td>C. parvum</td>
<td>28.83</td>
</tr>
<tr>
<td>CRYPTO_039/16</td>
<td>C. hominis</td>
<td>29.87</td>
<td>CRYPTO_007/17</td>
<td>C. parvum</td>
<td>34.3</td>
<td>CRYPTO_027/17</td>
<td>C. parvum</td>
<td>31.88</td>
</tr>
<tr>
<td>CRYPTO_040/16</td>
<td>C. hominis</td>
<td>35.07</td>
<td>CRYPTO_008/17</td>
<td>C. parvum</td>
<td>30.91</td>
<td>CRYPTO_028/17</td>
<td>C. parvum</td>
<td>28.64</td>
</tr>
<tr>
<td>CRYPTO_041/16</td>
<td>C. parvum</td>
<td>31.06</td>
<td>CRYPTO_009/17</td>
<td>C. parvum</td>
<td>33.89</td>
<td>CRYPTO_029/17</td>
<td>C. parvum</td>
<td>31.83</td>
</tr>
<tr>
<td>CRYPTO_042/16</td>
<td>C. hominis</td>
<td>31.80</td>
<td>CRYPTO_010/17</td>
<td>C. parvum</td>
<td>32.44</td>
<td>CRYPTO_030/17</td>
<td>C. parvum</td>
<td>36.87</td>
</tr>
<tr>
<td>CRYPTO_043/16</td>
<td>C. hominis</td>
<td>31.90</td>
<td>CRYPTO_011/17</td>
<td>C. parvum</td>
<td>30.84</td>
<td>CRYPTO_031/17</td>
<td>C. parvum</td>
<td>35.07</td>
</tr>
<tr>
<td>CRYPTO_044/16</td>
<td>C. hominis</td>
<td>34.38</td>
<td>CRYPTO_012/17</td>
<td>C. parvum</td>
<td>26.46</td>
<td>CRYPTO_032/17</td>
<td>C. parvum</td>
<td>33.59</td>
</tr>
<tr>
<td>CRYPTO_045/16</td>
<td>C. parvum</td>
<td>29.76</td>
<td>CRYPTO_013/17</td>
<td>C. parvum</td>
<td>29.00</td>
<td>CRYPTO_033/17</td>
<td>C. parvum</td>
<td>31.76</td>
</tr>
<tr>
<td>CRYPTO_046/16</td>
<td>C. hominis</td>
<td>35.08</td>
<td>CRYPTO_014/17</td>
<td>C. parvum</td>
<td>32.28</td>
<td>CRYPTO_034/17</td>
<td>C. parvum</td>
<td>31.94</td>
</tr>
<tr>
<td>CRYPTO_047/16</td>
<td>C. hominis</td>
<td>34.02</td>
<td>CRYPTO_015/17</td>
<td>C. parvum</td>
<td>32.07</td>
<td>CRYPTO_035/17</td>
<td>C. parvum</td>
<td>30.78</td>
</tr>
<tr>
<td>CRYPTO_048/16</td>
<td>C. parvum</td>
<td>32.99</td>
<td>CRYPTO_016/17</td>
<td>C. parvum</td>
<td>31.89</td>
<td>CRYPTO_036/17</td>
<td>C. parvum</td>
<td>28.43</td>
</tr>
<tr>
<td>CRYPTO_049/16</td>
<td>C. hominis</td>
<td>29.87</td>
<td>CRYPTO_017/17</td>
<td>C. parvum</td>
<td>30.8</td>
<td>CRYPTO_037/17</td>
<td>C. parvum</td>
<td>33.25</td>
</tr>
<tr>
<td>CRYPTO_050/16</td>
<td>C. parvum</td>
<td>28.16</td>
<td>CRYPTO_018/17</td>
<td>C. parvum</td>
<td>28.16</td>
<td>CRYPTO_038/17</td>
<td>C. parvum</td>
<td>31.95</td>
</tr>
<tr>
<td>CRYPTO_051/16</td>
<td>C. parvum</td>
<td>31.32</td>
<td>CRYPTO_019/17</td>
<td>C. parvum</td>
<td>31.32</td>
<td>CRYPTO_039/17</td>
<td>C. parvum</td>
<td>32.32</td>
</tr>
<tr>
<td>CRYPTO_052/16</td>
<td>C. parvum</td>
<td>29.97</td>
<td>CRYPTO_020/17</td>
<td>C. parvum</td>
<td>29.97</td>
<td>CRYPTO_040/17</td>
<td>C. parvum</td>
<td>34.84</td>
</tr>
<tr>
<td>Sample Number</td>
<td>18S rRNA gene Species Identification Result</td>
<td>Average Cp Value</td>
<td>Sample Number</td>
<td>18S rRNA gene Species Identification Result</td>
<td>Average Cp Value</td>
<td>Sample Number</td>
<td>18S rRNA gene Species Identification Result</td>
<td>Average Cp Value</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>-------------------------------------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>-------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>CRYPTO_041/17</td>
<td>C. parvum</td>
<td>38.26</td>
<td>CRYPTO_001/18</td>
<td>C. hominis</td>
<td>37.04</td>
<td>CRYPTO_019/18</td>
<td>C. parvum</td>
<td>32.76</td>
</tr>
<tr>
<td>CRYPTO_042/17</td>
<td>C. parvum</td>
<td>35.83</td>
<td>CRYPTO_002/18</td>
<td>C. parvum</td>
<td>30.55</td>
<td>CRYPTO_020/18</td>
<td>C. parvum</td>
<td>34.28</td>
</tr>
<tr>
<td>CRYPTO_043/17</td>
<td>C. parvum</td>
<td>35.19</td>
<td>CRYPTO_003/18</td>
<td>C. parvum</td>
<td>38.01</td>
<td>CRYPTO_021/18</td>
<td>C. parvum</td>
<td>28.48</td>
</tr>
<tr>
<td>CRYPTO_045/17</td>
<td>C. parvum</td>
<td>27.71</td>
<td>CRYPTO_004/18</td>
<td>C. parvum</td>
<td>39.07</td>
<td>CRYPTO_022/18</td>
<td>C. parvum</td>
<td>29.33</td>
</tr>
<tr>
<td>CRYPTO_046/17</td>
<td>C. parvum</td>
<td>32.85</td>
<td>CRYPTO_005/18</td>
<td>C. parvum</td>
<td>35.36</td>
<td>CRYPTO_023/18</td>
<td>C. parvum</td>
<td>36.92</td>
</tr>
<tr>
<td>CRYPTO_047/17</td>
<td>C. hominis</td>
<td>32.37</td>
<td>CRYPTO_006/18</td>
<td>C. parvum</td>
<td>34.28</td>
<td>CRYPTO_024/18</td>
<td>C. parvum</td>
<td>38.10</td>
</tr>
<tr>
<td>CRYPTO_048/17</td>
<td>C. hominis</td>
<td>36.19</td>
<td>CRYPTO_007/18</td>
<td>C. parvum</td>
<td>37.97</td>
<td>CRYPTO_025/18</td>
<td>C. parvum</td>
<td>36.04</td>
</tr>
<tr>
<td>CRYPTO_049/17</td>
<td>C. hominis</td>
<td>36.98</td>
<td>CRYPTO_008/18</td>
<td>C. parvum</td>
<td>38.50</td>
<td>CRYPTO_026/18</td>
<td>C. parvum</td>
<td>34.38</td>
</tr>
<tr>
<td>CRYPTO_050/17</td>
<td>C. parvum</td>
<td>34.98</td>
<td>CRYPTO_009/18</td>
<td>C. parvum</td>
<td>33.28</td>
<td>CRYPTO_027/18</td>
<td>C. parvum</td>
<td>37.60</td>
</tr>
<tr>
<td>CRYPTO_010/18</td>
<td>C. parvum</td>
<td>36.60</td>
<td>CRYPTO_010/18</td>
<td>C. parvum</td>
<td>36.60</td>
<td>C. parvum</td>
<td></td>
<td>36.12</td>
</tr>
<tr>
<td>CRYPTO_011/18</td>
<td>C. parvum</td>
<td>34.14</td>
<td>CRYPTO_011/18</td>
<td>C. parvum</td>
<td>34.14</td>
<td>C. parvum</td>
<td></td>
<td>36.12</td>
</tr>
<tr>
<td>CRYPTO_012/18</td>
<td>C. parvum</td>
<td>28.19</td>
<td>CRYPTO_012/18</td>
<td>C. parvum</td>
<td>28.19</td>
<td>C. parvum</td>
<td></td>
<td>29.60</td>
</tr>
<tr>
<td>CRYPTO_013/18</td>
<td>C. parvum</td>
<td>34.53</td>
<td>CRYPTO_013/18</td>
<td>C. parvum</td>
<td>34.53</td>
<td>C. parvum</td>
<td></td>
<td>33.78</td>
</tr>
<tr>
<td>CRYPTO_014/18</td>
<td>C. parvum</td>
<td>32.31</td>
<td>CRYPTO_014/18</td>
<td>C. parvum</td>
<td>32.31</td>
<td>C. parvum</td>
<td></td>
<td>37.93</td>
</tr>
<tr>
<td>CRYPTO_016/18</td>
<td>C. parvum</td>
<td>36.62</td>
<td>CRYPTO_016/18</td>
<td>C. parvum</td>
<td>36.62</td>
<td>C. parvum</td>
<td></td>
<td>37.03</td>
</tr>
<tr>
<td>CRYPTO_017/18</td>
<td>C. parvum</td>
<td>39.91</td>
<td>CRYPTO_017/18</td>
<td>C. parvum</td>
<td>39.91</td>
<td>C. parvum</td>
<td></td>
<td>35.04</td>
</tr>
<tr>
<td>CRYPTO_018/18</td>
<td>C. parvum</td>
<td>37.81</td>
<td>CRYPTO_018/18</td>
<td>C. parvum</td>
<td>37.81</td>
<td>C. parvum</td>
<td></td>
<td>32.91</td>
</tr>
<tr>
<td>Sample Number</td>
<td>18S rRNA gene Species Identification Result</td>
<td>Average Cp Value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018 (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_037/18</td>
<td>C. parvum</td>
<td>33.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_038/18</td>
<td>C. parvum</td>
<td>39.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_039/18</td>
<td>C. parvum</td>
<td>34.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_040/18</td>
<td>C. parvum</td>
<td>36.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_041/18</td>
<td>C. parvum</td>
<td>35.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_042/18</td>
<td>C. parvum</td>
<td>33.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_043/18</td>
<td>C. parvum</td>
<td>35.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_044/18</td>
<td>C. parvum</td>
<td>32.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_045/18</td>
<td>C. parvum</td>
<td>38.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_046/18</td>
<td>C. parvum</td>
<td>34.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_047/18</td>
<td>C. hominis</td>
<td>39.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_048/18</td>
<td>C. hominis</td>
<td>39.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_049/18</td>
<td>C. hominis</td>
<td>32.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_050/18</td>
<td>C. parvum</td>
<td>32.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_051/18</td>
<td>C. parvum</td>
<td>31.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYPTO_052/18</td>
<td>C. hominis</td>
<td>34.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A1.2 Amplification Curves and Electrophoresis Gels

The data below outlines the real-time PCR amplification curves and gel electrophoresis results of the fluorescent probe based real-time PCR species identification assay conducted on the CUH Cryptosporidium isolates. These data are summarised in Section 3.4.1, and discussed in Section 3.5 of Chapter III.

2015 Cryptosporidium Species Identification Data

Figure A1.1

Figure A1.2

Figure A1.1 - A1.2 Original fluorescent probe amplification curves and gel electrophoresis image of 2015 Cryptosporidium spp. isolates
2016 Cryptosporidium Species Identification Data

Figure A1.3

Figure A1.4

Figure A1.5
Figure A1.3 – A1.7.  Original fluorescent probe amplification curves and gel electrophoresis images of 2016 Cryptosporidium spp. isolates
2017 Cryptosporidium Species Identification Data

Figure A1.8

Figure A1.9
Figure A1.8–A1.12  Original fluorescent probe amplification curves and gel electrophoresis images of 2017 *Cryptosporidium* spp. isolates.
Figure A1.14
Figure A1.15
Figure A1.13 – A1.17  Original fluorescent probe amplification curves and gel electrophoresis images of 2018 Cryptosporidium spp. isolates
Appendix II – Chapter V Supplementary Data

The data below outlines the real-time PCR amplification curves and gel electrophoresis results of the sensitivity and specificity analyses conducted on *C. parvum* gp60-subtypes in the development of the novel genotyping assay detailed in Chapter V. These data are summarised in Section 5.4.3, and discussed in Section 5.5 of Chapter V.

### A2.1 Sensitivity Analysis

*Real-time PCR Amplification Curves of C. parvum Tandem Repeat Loci*

**gp60**

*gp60 Outer Primer Pair*

![Figure A2.1](image)

*Figure A2.1*  Real-time PCR amplification curves of sensitivity analysis of *gp60* outer primer pair. Limit of detection – 10^2 oocysts/ml
gp60 inner primer pair

Figure A2.2 Real-time PCR amplification curves of sensitivity analysis of gp60 inner primer pair. Limit of detection – $10^2$ oocysts/ml

MM5

MM5 Outer Primer Pair

Figure A2.3 Real-time PCR amplification curves of sensitivity analysis of MM5 outer primer pair. Limit of detection – $10^2$ oocysts/ml
MM5 Inner Primer Pair

Figure A2.4 Real-time PCR amplification curves of sensitivity analysis of MM5 outer primer pair. Limit of detection – $10^2$ oocysts/ml

MS9

MS9-Mallon Outer Primer Pair

Figure A2.5 Real-time PCR amplification curves of sensitivity analysis of MS9-Mallon outer primer pair. Limit of detection – $10^2$ oocysts/ml
MS9-Mallon Inner Primer Pair

Figure A2.6 Real-time PCR amplification curves of sensitivity analysis of MS9-Mallon outer primer pair. Limit of detection – $10^2$ oocysts/ml

*Gel electrophoresis of C. parvum Tandem Repeat Loci*

*gp60*

*gp60 Outer Primer Pair*

Figure A2.7 Gel electrophoresis image of sensitivity analysis of *gp60* outer primer pair. Limit of detection – $10^2$ oocysts/ml. HL25bp = HyperLadder 25 bp; PC = positive control; NTC = negative template control.
gp60 Inner Primer Pair

Figure A2.8  Gel electrophoresis image of sensitivity analysis of gp60 inner primer pair. Limit of detection – $10^2$ oocysts/ml. HL25bp = HyperLadder 25 bp; PC = positive control; NTC = negative template control

MM5

MM5 Outer Primer Pair

Figure A2.9  Gel electrophoresis image of sensitivity analysis of MM5 outer primer pair. Limit of detection – $10^2$ oocysts/ml. HL25 = HyperLadder 25 bp; PC = positive control; NTC = negative template control
MM5 Inner Primer Pair

![Image of MM5 Inner Primer Pair](image)

**Figure A2.10** Gel electrophoresis image of sensitivity analysis of MM5 inner primer pair. Limit of detection – $10^2$ oocysts/ml. HL25 = HyperLadder 25 bp; PC = positive control; NTC = negative template control

MS9-Mallon

MS9-Mallon Outer Primer Pair

![Image of MS9-Mallon Outer Primer Pair](image)

**Figure A2.11** Gel electrophoresis image of sensitivity analysis of MS9-Mallon outer primer pair. Limit of detection – $10^2$ oocysts/ml. HL25 = HyperLadder 25 bp; PC = positive control; NTC = negative template control
Figure A2.12  Gel electrophoresis image of sensitivity analysis of MS9-Mallon inner primer pair. Limit of detection – $10^2$ oocysts/ml. HL25 = HyperLadder 25 bp; PC = positive control; NTC = negative template control.
A2.2 Specificity Analysis

Real-time PCR Amplification Curves of C. parvum Tandem Repeat Loci

gp60

gp60 Outer Primer Pair

Figure A2.13  Real-time PCR amplification curves of specificity analysis of gp60 outer primer pair.

Figure A2.14  Real-time PCR amplification curves of specificity analysis of gp60 inner primer pair.
MM5

MM5 Outer Primer Pair

Figure A2.15  Real-time PCR amplification curves of specificity analysis of MM5 outer primer pair.

MM5 Inner Primer Pair

Figure A2.16  Real-time PCR amplification curves of specificity analysis of MM5 inner primer pair.
MS9-Mallon

MS9-Mallon Outer Primer Pair

Figure A2.17  Real-time PCR amplification curves of specificity analysis of MS9-Mallon outer primer pair.

MS9-Mallon Inner Primer Pair

Figure A2.18  Real-time PCR amplification curves of specificity analysis of MS9-Mallon inner primer pair.
Gel electrophoresis of *C. parvum* Tandem Repeat Loci

gp60

gp60 Outer Primer Pair

**Figure A2.19**  Gel electrophoresis image of specificity analysis of gp60 outer primer pair. Lane 1 *Salmonella*.; Lane 2. *Shigella*; Lane 3. *Campylobacter*; Lane 4. VTEC; Lane 5. *G. lamblia*; Lane 6. *B. hominis*; Lane 7. *C. parvum* (positive control); Lane 8. NTC. HL100bp = HyperLadder 100bp.
**Figure A2.20** Gel electrophoresis image of specificity analysis of *gp60* inner primer pair. Lane 1 *Salmonella*; Lane 2 *Shigella*; Lane 3 *Campylobacter*; Lane 4 VTEC; Lane 5 *G. lamblia*; Lane 6 *B. hominis*; Lane 7 *C. parvum* (positive control); Lane 8 NTC.HL25bp = HyperLadder 25 bp.
MM5 & MS9-Mallon

*MM5 and MS-Mallon Outer Primer Pairs*

![Gel electrophoresis image of specificity analysis of MM5 (lanes 1-8) and MS9-Mallon (lanes 9-16) outer primer pairs. Lane 1/9 *Salmonella*; Lane 2/10. *Shigella*; Lane 3/11. *Campylobacter*; Lane 4/12. VTEC; Lane 5/13. *G. lamblia*; Lane 6/14. *B. hominis*; Lane 7/15. *C. parvum* (positive control); Lane 8/16. NTC.HL50bp = HyperLadder 50 bp.]

*MM5 and MS9-Mallon Inner Primer Pairs*

![Gel electrophoresis image of specificity analysis of MM5 (lanes 1-8) and MS9-Mallon (lanes 9-16) inner primer pairs. Lane 1/9 *Salmonella*; Lane 2/10. *Shigella*; Lane 3/11. *Campylobacter*; Lane 4/12. VTEC; Lane 5/13. *G. lamblia*; Lane 6/14. *B. hominis*; Lane 7/15. *C. parvum* (positive control); Lane 8/16. NTC.HL25bp = HyperLadder 25 bp.]

193
Appendix III– Chapter VI Supplementary Data

The data below outlines the real-time PCR amplification curves and gel electrophoresis results of the sensitivity and specificity analyses conducted on *C. hominis* gp60-subtypes in the development of the novel genotyping assay detailed in Chapter VI. These data are summarised in Section 6.4.3, and discussed in Section 6.5 of Chapter VI.

**A3.1 Sensitivity Analysis**

*Real-time PCR Amplification Curves of C. hominis Tandem Repeat Loci*

*gp60*

*gp60 Outer Primer Pair*

![Real-time PCR amplification curves of sensitivity analysis of gp60 outer primer pair. Limit of detection – 10^3 oocysts/ml](#)

**Figure A3.1** Real-time PCR amplification curves of sensitivity analysis of *gp60* outer primer pair. Limit of detection – 10^3 oocysts/ml
**gp60 Inner Primer Pair**

![Graph showing real-time PCR amplification curves for gp60 inner primer pair.](image)

*Figure A3.2*  Real-time PCR amplification curves of sensitivity analysis of *gp60* inner primer pair. Limit of detection – $10^4$ oocysts/ml

**MSB**

**MSB Outer Primer Pair**

![Graph showing real-time PCR amplification curves for MSB outer primer pair.](image)

*Figure A3.3*  Real-time PCR amplification curves of sensitivity analysis of MSB outer primer pair. Limit of detection – $10^3$ oocysts/ml
MSB Inner Primer Pair

Figure A3.4  Real-time PCR amplification curves of sensitivity analysis of MSB outer primer pair. Limit of detection – $10^3$ oocysts/ml

MSE

MSE Outer Primer Pair

Figure A3.5  Real-time PCR amplification curves of sensitivity analysis of MSE outer primer pair. Limit of detection – $10^3$ oocysts/ml
MSE Inner Primer Pair

**Figure A3.6** Real-time PCR amplification curves of sensitivity analysis of MSE inner primer pair. Limit of detection – $10^3$ oocysts/ml

*Gel electrophoresis of C. hominis Tandem Repeat Loci*

*gp60*

*gp60* Outer Primer Pair

**Figure A3.7** Gel electrophoresis image of sensitivity analysis of gp60 outer primer pair. Limit of detection – $10^3$ oocysts/ml. HL1kb = HyperLadder 1kb PC = positive control; NTC = negative template control.
gp60 Oer Primer Pair

**Figure A3.8**  Gel electrophoresis image of sensitivity analysis of gp60 outer primer pair. Limit of detection – $10^4$ oocysts/ml. HL25bp = HyperLadder 25bp PC = positive control; NTC = negative template control.

**MSB & MSE**

MSB and MSE Outer Primer Pairs

**Figure A3.9**  Gel electrophoresis image of sensitivity analysis of MSB (lanes 1-9) and MSE (lanes 10-18) outer primer pairs. Limit of detection MSB = $10^5$ oocysts/ml; Limit of Detection MSE = $10^3$ oocysts/ml. HL100bp = HyperLadder 100bp PC = positive control; NTC = negative template control.
MSB and MSE Inner Primer Pairs

Figure A3.10  Gel electrophoresis image of sensitivity analysis of MSB (lanes 1-9) and MSE (lanes 10-18) inner primer pairs. Limit of detection MSB = $10^2$ oocysts/ml; Limit of Detection MSE = $10^3$ oocysts/ml. HL25bp = HyperLadder 25bp PC = positive control; NTC = negative template control.
A3.2 Specificity Analysis

*Real-time PCR Amplification Curves of C. hominis Tandem Repeat Loci*

*gp60*

*gp60 Outer Primer Pair*

![Real-time PCR amplification curves of specificity analysis of gp60 outer primer pair.](image)

**Figure A3.11** Real-time PCR amplification curves of specificity analysis of *gp60* outer primer pair.
gp60 Inner Primer Pair

![Graph showing real-time PCR amplification curves of specificity analysis of gp60 inner primer pair.]

**Figure A3.12** Real-time PCR amplification curves of specificity analysis of gp60 inner primer pair.

MSB

MSB Outer Primer Pair

![Graph showing real-time PCR amplification curves of specificity analysis of MSB outer primer pair.]

**Figure A3.13** Real-time PCR amplification curves of specificity analysis of MSB outer primer pair.
MSB Inner Primer Pair

**Figure A3.14** Real-time PCR amplification curves of specificity analysis of *MSB* inner primer pair.

MSE

*MSE* Outer Primer Pair

**Figure A3.15** Real-time PCR amplification curves of specificity analysis of MSE outer primer pair.
**MSE Inner Primer Pair**

![Graph](image1.png)  
**Figure A3.16** Real-time PCR amplification curves of specificity analysis of MSE outer primer pair.

**Gel electrophoresis of C. hominis Tandem Repeat Loci**

**GP60**

*gp60 Outer Primer Pair*

![Gel Image](image2.png)  
**Figure A3.17** Gel electrophoresis image of specificity analysis of *gp60* outer primer pair. Lane 1 *Salmonella*; Lane 2 *Shigella*; Lane 3 *Campylobacter*; Lane 4 VTEC; Lane 5 *G. lamblia*; Lane 6 *B. hominis*; Lane 7 *C. hominis* (positive control); Lane 8 NTC.HL100bp = HyperLadder 100 bp.
Figure A3.18  Gel electrophoresis image of specificity analysis of gp60 inner primer pair. Lane 1 Salmonella.; Lane 2. Shigella; Lane 3. Campylobacter; Lane 4. VTEC; Lane 5. G. lamblia; Lane 6. B. hominis; Lane 7. C. hominis (positive control); Lane 8. NTC. HL100bp = HyperLadder 100 bp.

MSB & MSE

MSB and MSE Outer Primer Pairs

Figure A3.19  Gel electrophoresis image of specificity analysis of MSB (lanes 1-8) and MSE (lanes 9-16) outer primer pairs. Lane 1/9 Salmonella.; Lane 2/10. Shigella; Lane 3/11. Campylobacter; Lane 4/12. VTEC; Lane 5/13. G. lamblia; Lane 6/14. B. hominis; Lane 7/15. C. hominis (positive control); Lane 8/16. NTC. HL100bp = HyperLadder 100 bp.
MSB and MSE Inner Primer Pairs

Figure A3.20  Gel electrophoresis image of specificity analysis of MSB (lanes 1-8) and MSE (lanes 9-16) inner primer pairs. Lane 1/9 Salmonella.; Lane 2/10. Shigella; Lane 3/11. Campylobacter; Lane 4/12. VTEC; Lane 5/13. G. lamblia; Lane 6/14. B. hominis; Lane 7/15. C. hominis (positive control); Lane 8/16. NTC.HL25bp = HyperLadder 25 bp.
Appendix IV – Addendum to Chapter V

A4.1 Introduction

Further to the publication of the original method and *gp60*-subtype groupings detailed in Chapter V, additional statistical analyses and interpretation of HRM analysis *gp60*-subtype grouping data for *C. parvum* loci were undertaken. These analyses were conducted to ascertain whether the established groups were statistically supported.

Herein, the assessment of these data is detailed as an addendum in support of the original published research.

A4.2 Materials and Methods

A4.2.1 Statistical Analysis – One-way ANOVA

One-way ANOVA testing was performed to determine whether the *gp60*-subtypes within the putative HRM analysis groups defined in Chapter V for *gp60*, MM5 and MS9-Mallon were statistically supported. Analysis aimed to determine if the intra-group T_m values segregated together, exhibiting a statistically significant group effect. Statistical analysis was conducted on the putative groups for *gp60*, MM5 and MS9-Mallon.

ANOVA testing determined differences between groups based on replicate T_m values for each *gp60*-subtype isolates within the respective groups for each locus. A confidence level of 95% (α ≤ 0.05) was employed in all statistical analyses.

Where a statistically significant difference was identified, post-hoc analysis in the form of Tukey’s honest significance post-hoc test was conducted to determine which groups differed significantly.

A4.2.2 Re-interpretation of HRM analysis groups

In cases where existing HRM analysis groups were found to not be statistically supported, groupings were re-interpreted to yield statistically robust groupings. Revised groupings for MS9-Mallon were inferred through visual interpretation of distances between distinct and overlapping subtype isolate melting curves and peaks (using LC96 software), in conjunction with post-hoc analysis data identifying groups that were not statistically different. Statistical analyses (ANOVA and post-hoc analysis) were
conducted on re-interpreted groupings to determine whether these new groups were statistically supported.

A4.3 Results

A4.3.1 Statistical analyses – One-way ANOVA

gp60

One-way ANOVA analysis of gp60 data revealed that a statistically significant difference existed between groups (F[7, 81] = 573.849, p = 0.000).

Tukey’s HSD test for multiple comparisons identified statistically significant differences between all gp60 groups. The resulting data indicating these results are summarised as follows:

- Group 1 and Group 2 (p = 0.000; 95% C.I. = [-.2957, -.0815]), Group 1 and Group 3 (p = 0.000; 95% C.I. = [-.3431, -.1965]), Group 1 and Group 4 (p = 0.000; 95% C.I. = [-.5015, -.3576]), Group 1 and Group 5 (p = 0.000; 95% C.I. = [-.6017, -.4641]), Group 1 and Group 6 (p = 0.000; 95% C.I. = [-.8454, -.7079]), Group 1 and Group 7 (p = 0.000; 95% C.I. = [-1.1989, -.10357]), and Group 1 and Group 8 (p = 0.000; 95% C.I. = [-2.0074, -.6126])

- Group 2 and Group 3 (p = 0.016; 95% C.I. = [-.1870, -.0114]), Group 2 and Group 4 (p = 0.015; 95% C.I. = [-.3457, -.1723]), Group 2 and Group 5 (p = 0.000; 95% C.I. = [-.4464, -.2782]), Group 2 and Group 6 (p = 0.000; 95% C.I. = [-.6901, -.5220]), Group 2 and Group 7 (p = 0.000; 95% C.I. = [-1.0415, -.8518]), and Group 2 and Group 8 (p = 0.000; 95% C.I. = [-1.2401, -.10447])

- Group 3 and Group 4 (p = 0.000; 95% C.I. = [-.2301, -.0895]), Group 3 and Group 5 (p = 0.000; 95% C.I. = [-.3302, -.1961]), Group 3 and Group 6 (p = 0.000; 95% C.I. = [-.5739, -.4398]), Group 3 and Group 7 (p = 0.000; 95% C.I. = [-.9277, -.7673]), and Group 3 and Group 8 (p = 0.000; 95% C.I. = [-1.1267, -.9597])

- Group 4 and Group 5 (p = 0.000; 95% C.I. = [-.1689, -.0377]), Group 4 and Group 6 (p = 0.000; 95% C.I. = [-.4126, -.2815]), Group 4 and Group 7 (p = 0.000; 95% C.I. = [-.7666, -.6088]), and Group 4 and Group 8 (p = 0.000; 95% C.I. = [-.9657, -.8011])
- Group 5 and Group 6 (p = 0.000; 95% C.I. = [-.3058, -.1817]), Group 5 and Group 7 (p = 0.000; 95% C.I. = [-.6604, -.5803]), and Group 5 and Group 8 (p = 0.000; 95% C.I. = [-.8597, -.7705]).
- Group 6 and Group 7 (p = 0.000; 95% C.I. = [-.4167, -.2646]), and Group 6 and Group 8 (p = 0.000; 95% C.I. = [-.6159, -.4597])
- Group 7 and Group 8 (p = 0.000; 95% C.I. = [-.2866, -.1048])

**MM5**

One-way ANOVA analysis of MM5 data revealed that a statistically significant difference existed between the two established groups (F [1, 66] = 1505.702, p= 0.000).

No post-hoc analyses were required as there only were two groups identified at this locus. Therefore all existing subtype groups were supported by statistical analysis.

**MS9-Mallon**

One-way ANOVA analysis of MS9-Mallon data revealed that a statistically significant difference existed between groups (F [4, 67] = 441.792, p= 0.000).

Tukey’s HSD Test for multiple comparisons identified statistically significant differences between all MS9-Mallon groups with the exception of Group 3 and Group 4.

Statistically significant differences were identified between the following groups:

- Group 1 and Group 2 (p = 0.000; 95% C.I. = [-.5149, -.2576]), Group 1 and Group 3 (p = 0.000; 95% C.I. = [-1.1136, -.8164]), Group 1 and Group 4 (p = 0.000; 95% C.I. = [-1.2586, -.9614]), and Group 1 and Group 5 (p = 0.000; 95% C.I. = [-1.3527, 1.1346])
- Group 2 and Group 3 (p = 0.000; 95% C.I. = [-.7074, -.4501]), Group 2 and Group 4 (p = 0.000; 95% C.I. = [-.8524, -.5951]), and Group 2 and Group 5 (p = 0.00109% C.I. = [-.9372, -.7776])
- Group 3 and Group 5 (p = .000; 95% C.I. = [-.3877, -.1696])
- Group 4 and Group 5 (p = 0.000; 95% C.I. = [-.2427, -.0246]).

No statistically significant differences existed between Group 3 and 4 (p = .059; 95% C.I. = (-.2936, .0036)).
A4.3.2 Re-interpretation of HRM analysis groups

MS9-Mallon
Based on the results of one-way ANOVA testing, it was determined that putative Group 3 and Group 4, previously established via HRM analysis at the MS9-Mallon locus, were not statistically supported. Revised groupings for MS9-Mallon were inferred through visual interpretation of distances between distinct and overlapping gp60-subtype melting curves and peaks (using LC96 software), in conjunction with post-hoc analysis data identifying groups that were not statistically different.

Re-interpretation of MS9-Mallon T<sub>m</sub> data resulted in the merging of the original Group 3 and Group 4, now collectively termed Group 3. Subsequently, Group 5 was renamed appropriately as Group 4, but the gp60-subtype assignment to this group remained the same.

Figures A4.1-A4.3 depict the newly established, statistically supported groups following re-interpretation of the data and additional statistical testing in support of these newly established groups. The gp60-subtype assignments in the revised MS9-Mallon grouping system are detailed in Table A4.1.
Figure A4.1  Normalised melting curves of statistically supported MS9-Mallon HRM analysis groups

Figure A4.2  Normalised melting peaks of statistically supported MS9-Mallon HRM analysis groups
Figure A4.3  Difference plot of statistically supported MS9-Mallon HRM analysis groups
A4.3.3 Statistical analysis of re-interpreted HRM analysis groups

MS9-Mallon

One-way ANOVA analysis of re-interpreted MS9-Mallon $T_m$ data revealed that a statistically significant difference existed between all four established groups ($F [3, 68] = 535.476; p= 0.000$).

Tukey’s HSD Test for multiple comparisons identified statistically significant differences between all MS9-Mallon groups. The resulting data indicating these results are summarised as follows:

- Group 1 and Group 2 ($p = 0.000; 95\% \text{ C.I.} = [-.5128, -.2597]$), Group 1 and Group 3 ($p = 0.000; 95\% \text{ C.I.} = [-1.1640, -.9110]$), and Group 1 and Group 4 ($p = 0.000; 95\% \text{ C.I.} = [-1.3508, -1.1365]$)
- Group 2 and Group 3 (p = 0.000; 95% C.I. = [-.7545, -.5480]), and Group 2 and Group 4 (p = 0.000; 95% C.I. = [-.9359, -7790]), and
- Group 3 and Group 4 (p = .000; 95% C.I. = [-.2846, .1277]).

Therefore all four re-interpreted \textit{gp60}-subtype groups were supported by statistical analysis.

\textbf{A4.3.4 Impact on the differentiation of \textit{C. parvum} gp60-subtype isolates}

The merging of MS9-Mallon Groups 3 and 4 involved only three \textit{gp60}-subtype isolates, IIaA15G1R2, IIaA16G1R1, and IIaA21G3R1. IIaA15G1R2 and IIaA16G1R1 exhibited different \(T_m\) characteristics at both \textit{gp60} and MM5 loci, while IIaA21G3R1 exhibited the same profile at these loci as IIaA17G2R1, IIaA19G3R1, and IIaA20G3R1 at the \textit{gp60} and MM5 loci. However, IIaA21G3R1 was still distinguishable from the IIaA17G2R1, IIaA19G3R1, and IIaA20G3R1 isolates at the MS9-Mallon locus after grouping re-interpretation.

These results are summarised in Table A4.2.
Table A4.2  Revised multi-locus HRM analysis grouping and sequence designations of *C. parvum* gp60-subtype isolates

<table>
<thead>
<tr>
<th>Isolate</th>
<th>gp60 HRM Analysis Grouping</th>
<th>MM5 HRM Analysis Grouping</th>
<th>MS9-Mallon HRM Analysis Grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIaA20G1R1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IIaA21G1R1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>IIaA16R1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>IIaA17G1R1</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>IIaA18G1R1</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>IIaA20G2R1</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>IIaA15G1R2</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>IIaA16G1R1</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>IIaA17G2R1</td>
<td>5</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>IIaA19G3R1</td>
<td>5</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>IIaA20G3R1</td>
<td>5</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>IIaA21G3R1</td>
<td>5</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>IIaA15G2R1</td>
<td>6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>IIaA17G3R1</td>
<td>6</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>IIaA18G3R1</td>
<td>6</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>IIaA19G4R1</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>IIaA17G4R1</td>
<td>7</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>IIaA10G2R1</td>
<td>8</td>
<td>n/a&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4</td>
</tr>
</tbody>
</table>

<sup>a</sup> Differentiated at the *gp60* locus based on melting peak shape

<sup>b</sup> No amplification occurred at this locus
A4.4 Discussion

Initial analysis and grouping of the studied gp60-subtype isolate panel was primarily based on visual interpretation of melting curves, melting peaks and difference plots generated by LC96 software. This subsequent re-interpretation of HRM analysis data has produced more statistically robust results. Re-interpreted groupings were also based on visual interpretation of melting curves, melting peaks and difference plots generated by LC96 software and informed by post-hoc analysis data of original grouping categories. The resulting revised groupings for MS9-Mallon locus allow for more reliable and reproducible Tm-based group assignment of isolates.

The gp60 and MM5 gp60-analysis subtype groupings determined in Chapter V of this body of work were statistically supported and required no further analysis.

MS9-Mallon groups were revised and reduced from five groups to four. Within the new grouping system, the putative Groups 3 and 4 were amalgamated and the original Group 3 numbering retained. Re-interpretation of these HRM analysis groups exerted a minor impact on the results reported in Chapter V. Merging of MS9-Mallon Groups 3 and 4 involved three gp60-subtypes, IIA15G1R2, IIA16G1R and IIA21G3R1. IIA15G1R2 and IIA16G1R already exhibited different Tm characteristics at both the gp60 and MM5 loci and were not impacted by the revised grouping. IIA21G3R1 exhibited the same HRM analysis grouping pattern as IIA17G2R1, IIA19G3R1, and IIA20G3R1 at the gp60 and MM5 loci. However, IIA21G3R1 remained distinguishable from these isolates at the MS9-Mallon locus after grouping re-interpretation. Therefore the number of gp60 isolates differentiated by this method was not altered.

Ultimately, revision of the MS9-Mallon grouping system did not exert a major effect on the results reported in Chapter V. Statistical analysis and subsequent re-examination of the HRM data of MS9-Mallon resulted in more robust groupings although with slightly poorer resolution that those of the putative groupings described previously.

It is important to note that DNA sequence induced shifts in Tm detected by HRM analysis can be miniscule. For example, an A/T base change results in a Tm difference of >0.2°C (Liew et al., 2004). Given the subtle nature in Tm difference observed in HRM analysis, it is advised that visual interpretation of Tm data is conducted in
conjunction with T$_m$ data of regionally prevalent control isolates and statistically supported grouping systems. Further analysis of replicate T$_m$ data a broader array of gp60-subtypes within Cryptosporidium spp., and indeed investigating a larger panel of tandem repeat loci would be useful in gaining a more comprehensive understanding of grouping effect and subtype identification at these loci. However, this has been discussed in greater detail in Chapter V.
Appendix V – Addendum to Chapter VI

A5.1 Introduction

Further to the publication of the original method and gp60-subtype groupings detailed in Chapter VI, additional statistical analyses and interpretation of HRM analysis gp60-subtype grouping data for C. hominis loci were undertaken. These analyses were conducted to ascertain whether the established groups were statistically supported.

Herein, the re-assessment of these data is detailed as an addendum in support of the original published research.

A5.2 Materials and Methods

A5.2.1 Statistical Analysis – One-way ANOVA

One-way ANOVA testing was performed to determine whether the gp60-subtypes within the putative HRM analysis groups defined in Chapter VI for gp60, MSB and MSE were statistically supported. Analysis aimed to determine if the intra-group Tm values segregated together, exhibiting a statistically significant group effect. Statistical analysis was conducted on the putative groups for gp60, MSB and MSE.

ANOVA testing determined differences between groups based on replicate Tm values for each gp60-subtype isolates within the respective groups for each locus. A confidence level of 95% (α ≤ 0.05) was employed in all statistical analyses.

Where a statistically significant difference was identified, post-hoc analysis in the form of Tukey’s honest significance post-hoc test was conducted to determine which groups differed significantly.

A5.2.2 Re-interpretation of HRM analysis groups

In cases where existing HRM analysis groups were found to not be statistically supported, groupings were re-interpreted to yield statistically robust groupings. Revised groupings for gp60 were inferred through visual interpretation of distances between distinct and overlapping gp60-subtype melting curves and peaks (using LC96 software), in conjunction with post-hoc analysis data identifying groups that were not statistically different. Statistical analyses (ANOVA and post-hoc analysis) were conducted on re-
interpreted groupings to determine whether these new groups were statistically supported.

A5.3 Results

A5.3.1 Statistical analyses – One-way ANOVA

gp60

One-way ANOVA analysis of gp60 data revealed that a statistically significant difference existed between groups (F [4, 19] = 681.540, p = 0.000).

Tukey’s HSD Test for multiple comparisons identified statistically significant differences between the following groups:

- Group 1 and Group 2 (p = 0.000; 95% C.I. = [-.7398, -.5802]), Group 1 and Group 3 (p = 0.000; 95% C.I. = [-.9323, -.7727]), Group 1 and Group 4 (p = 0.000; 95% C.I. = [-1.0173, -.8577]), and Group 1 and Group 5 (p = 0.000; 95% C.I. = [-1.2948, 1.1352])
- Group 2 and Group 3 (p = 0.000; 95% C.I. = [-.2846, -.1004]), Group 2 and Group 4 (p = 0.000; 95% C.I. = [-.3696, -.1854]), and Group 2 and Group 5 (p = 0.000; 95% C.I. = [-.6471, -.4629])
- Group 3 and Group 5 (p = 0.000; 95% C.I. = [-.4546, -.2704]), and
- Group 4 and Group 5 (p = 0.000; 95% C.I. = [-.2427, -.0246]).

No statistically significant differences existed between Group 3 and 4 (p = .079; 95% C.I. = (-1.771, .0071)).

MSB

One-way ANOVA analysis of gp60 data revealed that a statistically significant difference existed between groups (F [2, 2] = 40.140, p = 0.000).

Tukey’s HSD Test for multiple comparisons identified statistically significant differences between all MSB groups. The resulting data indicating these results are summarised as follows:

- Group 1 and Group 2 (p = 0.000; 95% C.I. = [-.1299, -.0476]), and Group 1 and Group 3 (p = 0.000; 95% C.I. = [-.2371, -.1329]), and
- Group 2 and Group 3 (p = 0.000; 95% C.I. = [-.1374, -.0551]).
Therefore all three subtype groups were supported by statistical analysis.

**MSE**

One-way ANOVA analysis of \( gp60 \) data revealed that a statistically significant difference existed between groups (\( F[3, 20] = 84.283, p= 0.000 \)).

Tukey’s HSD Test for multiple comparisons identified statistically significant differences between all MS9-Mallon groups. The resulting data indicating these results are summarised as follows:

- Group 1 and Group 2 (\( p = 0.000; 95\% \text{ C.I.} = [-.2198, -.1208] \)), Group 1 and Group 3 (\( p = 0.000; 95\% \text{ C.I.} = [-.2978, -.1772] \)), and Group 1 and Group 4 (\( p = 0.000; 95\% \text{ C.I.} = [-.3678, -.2475] \))
- Group 2 and Group 3 (\( p = 0.000; 95\% \text{ C.I.} = [-.1278, -.0072] \)), and Group 2 and Group 4 (\( p = 0.000; 95\% \text{ C.I.} = [-.9359, -.7790] \)), and
- Group 3 and Group 4 (\( p = 0.000; 95\% \text{ C.I.} = [-.1396, .0004] \))

Therefore all four subtype groups were supported by statistical analysis.

**A5.3.2 Re-interpretation of HRM analysis groups**

**gp60**

Based on the results of one-way ANOVA testing, it was determined that two of the five putative groups that were previously established in Chapter VI via HRM analysis of the \( gp60 \) locus were not statistically supported.

Re-interpretation of the \( gp60 \) \( T_m \) data saw the original 5 groups reduced to 4 in the revised grouping scheme. Groups 3 and 4 merged into a single group.

Figures A5.1-A5.3 depict the newly established, statistically supported groups following re-interpretation of the data and additional statistical testing in support of these newly established groups. The \( gp60 \)-subtype assignments in the revised system are detailed in Table A5.1.
Figure A5.1  Melting curves of statistically supported *C. hominis* gp60 HRM analysis groups

Figure A5.2  Melting peaks of statistically supported *C. hominis* gp60 HRM analysis groups
Table A5.1  gp60-subtype groupings and amplicon characteristics for studied C. hominis isolates

<table>
<thead>
<tr>
<th>Group</th>
<th>gp60-subtypes</th>
<th>%G+C Content</th>
<th>Sequence Length (bp)</th>
<th>Upper T&lt;sub&gt;m&lt;/sub&gt; Limit (°C)</th>
<th>Lower T&lt;sub&gt;m&lt;/sub&gt; Limit (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>IdA21</td>
<td>46.20%</td>
<td>158</td>
<td>82.40</td>
<td>82.49</td>
</tr>
<tr>
<td></td>
<td>IeA11G3T3</td>
<td>47.20%</td>
<td>142</td>
<td>83.40</td>
<td>82.56</td>
</tr>
<tr>
<td>Group 2</td>
<td>IaA14R3</td>
<td>47.10%</td>
<td>136</td>
<td>83.06</td>
<td>83.15</td>
</tr>
<tr>
<td>Group 3</td>
<td>IbA10G2</td>
<td>49.20%</td>
<td>130</td>
<td>83.26</td>
<td>83.34</td>
</tr>
<tr>
<td></td>
<td>IbA12G3</td>
<td>48.90%</td>
<td>139</td>
<td>83.32</td>
<td>83.41</td>
</tr>
<tr>
<td>Group 4</td>
<td>IbA9G3</td>
<td>50.00%</td>
<td>130</td>
<td>83.65</td>
<td>83.74</td>
</tr>
</tbody>
</table>

A5.3.3 Statistical analysis of re-interpreted HRM analysis groups

gp60

One-way ANOVA analysis of re-interpreted gp60 T<sub>m</sub> data revealed that a statistically significant difference existed between all four established groups (F [3, 20] = 678.734; p= 0.000).
Tukey’s HSD Test for multiple comparisons identified statistically significant differences between all gp60 groups. The resulting data indicating these results are summarised as follows:

- Group 1 and Group 2 (p = 0.000; 95% C.I. = [-.7458, -.5742]), Group 1 and Group 3 (p = 0.000; 95% C.I. = [-.9650, -.8250]), and Group 1 and Group 4 (p = 0.000; 95% C.I. = [-1.3008, -1.1292])
- Group 2 and Group 3 (p = 0.000; 95% C.I. = [-.3208, -.1492]), and Group 2 and Group 4 (p = 0.000; 95% C.I. = [-.6541, -.4559]), and
- Group 3 and Group 4 (p = .000; 95% C.I. = [-.4058, -.2342]).

Therefore all four re-interpreted subtype groups were supported by statistical analysis.

A5.3.4 Impact on the differentiation of C. hominis gp60-subtype isolates

The revision of the C. hominis gp60 HRM analysis groupings did not impact the ability of the method to differentiate between the gp60-subtype isolates examined. All six of the distinct gp60-subtypes were successfully differentiated by the revised grouping scheme, in concordance with the results detailed in Chapter VI. These results are summarised in Table A5.2.

**Table A5.2** Revised multi-locus HRM analysis grouping and sequence designations of C. hominis gp60-subtype isolates

<table>
<thead>
<tr>
<th>Isolate</th>
<th>HRM Analysis Grouping</th>
<th>HRMA Analysis Grouping</th>
<th>HRM analysis grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td>IaA14R3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>IbA9G3</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>IbA10G2</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>IbA12G3</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>IdA21</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>IeA11G3R3</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>
A5.4 Discussion

Initial analysis and grouping of the studied gp60-subtype panel was primarily based on visual interpretation of melting curves, melting peaks and difference plots generated by LC96 software. This subsequent re-interpretation of HRM analysis data has produced more statistically robust results. Re-interpreted groupings were also based on visual interpretation of melting curves, melting peaks and difference plots generated by LC96 software and informed by post-hoc analysis data of original grouping categories. The resulting revised groupings for both the gp60 and MS9-Mallon loci allow for more reliable and reproducible T\textsubscript{m}-based group assignment of isolates.

Subsequent to ANOVA and post-hoc analysis, C. hominis gp60 groups were reduced from 5 groups to 4, with the original Groups 3 and 4 consolidated in a newly designated Group 3, and Group 5 accordingly renamed as Group 4. The merging of these two groups did not impact the differentiation of the C. hominis gp60-subtypes analysed via HRM analysis in this study. This single group consolidation impacted upon only two gp60-subtypes, IbA10G2 and IbA12G3, which exhibited differing T\textsubscript{m} characteristics at the MSB and MSE loci, respectively.

Additionally, the respective MSB and MSE gp60-subtype groups determined in Chapter VI of this body of work were statistically supported and required no further analysis.
References


Aghamolaie, S., Rostami, A., Fallahi, S., Tahvildar Biderouni, F., Haghighi, A., Salehi,


Aldeyarbi, H.M., Karanis, P., 2016a. The fine structure of sexual stage development and sporogony of *Cryptosporidium parvum* in cell-free culture. Parasitology 143, 749–761. doi: 10.1017/S0031182016000275


Auburn, S., Barry, A.E., 2017. Dissecting malaria biology and epidemiology using population genetics and genomics. Int. J. Parasitol. 47, 77–85. doi:


Barker, I.K., Carbonell, P.L., 1974. *Cryptosporidium agni* sp.n. from lambs, and *Cryptosporidium bovis* sp.n. from a calf, with observations on the oocyst. Z Parasitenkd. 44, 289-98. doi: 10.1007/BF00366112


Baptista, R.P.2, Cooper, G.W., Kissinger, J.C., 2021. Challenges for *Cryptosporidium*


Brisson, D., 2018. Negative Frequency-Dependent Selection Is Frequently


Cacciò, S., Spano, F., Pozio, E., 2001. Large sequence variation at two microsatellite
loci among zoonotic (genotype C) isolates of Cryptosporidium parvum. Int. J. Parasitol. 31, 1082–1086. doi: 10.1016/s0020-7519(01)00233-8


234


Detect Cryptosporidium Species. Pathog. 10. doi: 10.3390/PATHOGENS10060647


Eibach, D., Krumkamp, R., Al-Emran, H.M., Sarpong, N., Hagen, R.M., Adu-Sarkodie,


European Centre for Disease Prevention and Control, 2018b. Annual epidemiological


Flecha, M.J., Benavides, C.M., Tissiano, G., Tesfamariam, A., Cuadros, J., de Lucio,
classification of *Giardia duodenalis*, *Cryptosporidium* spp. and *Entamoeba* spp.
among patients with gastrointestinal symptoms in Gambo Hospital, Oromia

Florescu, D.F., Sandkovsky, U., 2016. *Cryptosporidium* infection in solid organ
transplantation. World J. Transplant. 6, 460. doi: 10.5500/wjt.v6.i3.460

Fournet, N., Deege, M.P., Urbanus, A.T., Nichols, G., Rosner, B.M., Chalmers, R.M.,
Gorton, R., Pollock, K.G., van der Giessen, J.W.B., Wever, P.C., Dorigo-Zetsma,
J.W., Mulder, B., Mank, T.G., Overdevest, I., Kusters, J.G., van Pelt, W.,
Kortbeek, L.M., 2013. Simultaneous increase of *Cryptosporidium* infections in the
Netherlands, the United Kingdom and Germany in late summer season, 2012. Euro
Surveill. 18, 20348. doi: 10.2807/ese.18.02.20348-en

G.R., Dickinson, J., Drews, S.J., Eltorki, M., Graham, T., Jiang, X., Johnson,
D.W., Kellner, J., Lavoie, M., MacDonald, J., MacDonald, S., Svenson, L.W.,
Pediatric EnTeric Infection TEam (APPETITE): epidemiology, emerging

Fretz, R., Svoboda, P., Ryan, U.M., Thompson, R.C.A., Tanners, M., Baumgartner, A.,
2003. Genotyping of *Cryptosporidium* spp. isolated from human stool samples in
Switzerland. Epidemiol. Infect. 131, 663–667. doi: 10.1017/s0950268803008781

identified by high-throughput screening against parasite fatty acyl-CoA binding

2003. Cryptosporidiosis outbreak in British tourists who stayed at a hotel in

*Cryptosporidium* species causing human infections. Parasitology 147, 1532–1537.


Karanis, P., 2018. The truth about in vitro culture of *Cryptosporidium* species. Parasitology. doi: 10.1017/S0031182017001937


255


Mary, C., Chapey, E., Dutoit, E., Guyot, K., Hasseine, L., Jeddi, F., Menotti, J., Paraud,


265


characterisation of *Cryptosporidium* outbreaks in Western and South Australia. Exp. Parasitol. 125, 325–328. doi: 10.1016/j.exppara.2010.02.012


Patel, S., Pedraza-dõô, S., McLaughlin, J., 1999. The identification of Cryptosporidium species and Cryptosporidium parvum directly from whole faeces by analysis of a multiplex PCR of the 18S rRNA gene and by PCR / RFLP of the Cryptosporidium


274


Ryan, U., 2010. Cryptosporidium in birds, fish and amphibians. Exp. Parasitol. 124,


281


Publications