MOLECULAR EPIDEMIOLOGY AND

METABOLOMIC

CHARACTERISATION OF

CRYPTOSPORIDIUM

Josephine Su Yin Ng
BSc. Biomedical Science (Hon)

Thesis presented for the degree of

Doctor of Philosophy
School of Veterinary Life Sciences
Murdoch University 2017
DECLARATION

I declare that this thesis is my own account of my research and contains as its main content, work that has not previously been submitted for a degree at any tertiary institution.

[Signature]

Josephine Su Yin Ng (also known as: Josephine S.Y. Ng-Hublin; Josephine S. Y. Hublin)
ABSTRACT

This thesis examined the molecular epidemiology of the important enteric parasite, Cryptosporidium in Australia with particular reference to cryptosporidiosis in Aboriginal communities, outbreaks, zoonotic transmission and also conducted the first metabolomics analysis of Cryptosporidium.

Chapter 3 revealed striking differences in the epidemiology of Cryptosporidium between Aboriginal and non-Aboriginal people, with notification rates among Aboriginal people up to 50 times higher. Aboriginal people were predominantly infected with C. hominis subtype IdA15G1 and non-Aboriginal people were predominantly infected with C. hominis subtype IbA10G2. Chapters 4 and 5 explored the epidemiology of outbreaks with the C. hominis IbA10G2 subtype, the major subtype identified in all outbreaks.

Chapter 6 examined zoonotic transmission of Cryptosporidium species in rural NSW. Three species of Cryptosporidium were detected in calves; C. parvum, C. bovis and C. ryanae and two in humans; C. parvum and C. bovis. Subtyping identified the concurrence of C. parvum subtypes between calves and humans and this coupled with the identification of the cattle-specific C. bovis in humans and calves provides supportive evidence of zoonotic transmission.

Chapter 7 developed a reproducible faecal extraction method for untargeted gas chromatography-mass spectrometry (GC-MS) analysis and identified distinct differences in faecal metabolite profiles between infected and un-infected individuals. However, as the metabolome is sensitive to external perturbations, a more controlled metabolomics analysis of faecal metabolite profiles was conducted in Chapter 8 using experimentally infected mice. Despite the differences in faecal metabolite profiles between Cryptosporidium infected humans and mice, metabolomic analysis in both studies was still able to clearly differentiate
between infected and uninfected hosts, as well as provide information on the metabolic activity of the parasite during the infection based on faecal metabolite profiles.

The finding of this thesis will greatly assist in our understanding of molecular epidemiology of Cryptosporidium in Australia and the biochemistry of this important parasite.
ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincere gratitude and appreciation to my PhD supervisors, Professor Una M. Ryan, Associate Professor Robert Trengove and Dr. Garth Maker whose enthusiasm, invaluable insights and tireless support has seen me through this journey. Una, you infected me with your love and passion for Cryptosporidium and all the work that you do 15 years ago, and I am always ever so thankful for the boundless opportunities you have given me since then. Thank you for being my mentor; for believing in me; and for being always so patient, kind and understanding. You have been an absolute inspiration to me; I can only dream to be more like you and hope to one day, have the opportunity to “pay it forward”. To Rob and Garth, I am grateful to have had the opportunity to work with such experts in the field of metabolomics! Thank you for your optimism, constant guidance and encouragement through all its intricacies and more.

I would also like to acknowledge and thank Dr. Barry Combs from the Communicable Disease Control Directorate at the Western Australia Department of Health (Chapters 3, 4 and 5) for his close collaboration, guidance, inputs and advice.

A special acknowledgement for the support and friendship extended to me by the wonderful staff, colleagues and friends at the Murdoch University Department of Veterinary and Life Science; particularly to those in the Molecular Epidemiology group, Linda McInnes, Dr. Rongchang Yang, Therese Weldt, Dr. Annika Estcourt and Dr. Andrea Paparini; Parasitology group, Tanya Armstrong, especially for her help and advice on experimental mice work, Louise Pallant and Aileen Elliot; and to the great support from the technical and administrative staff in the department, Ken Chong, Frances Briggs, Dr. Kirsty Townsend, Aileen Elliot, Peter Fallon and Pat Marshall.
To the wonderful team at the Murdoch University Separation Science and Metabolomics Laboratory, Catherine Rawlinson, Dr. Joel Gummer, Hayley Abbis, Hui Wong and Katherine Roussety, thank you for all your help, guidance and their friendship. A special thanks also to Dr. Bong Sze How, for his invaluable insights and for always being a friend.

Last but definitely not least, are the contributions of my family and friends as a whole, for keeping my sanity and for without them, this accomplishment would not have been possible. To my dad (Harry Ng), mom (Lucia Pin), and my brothers (Christopher and Benedict Ng) thank you for your enduring love, support and encouragement. To my wonderful husband, Trevor Hublin - my deepest thank you for your boundless love, patience, support and encouragement; for sacrifices made to ensure that I cross the finish line; and for our beautiful daughters Marie and Madeleine, who had become part of our lives and joined me on this journey to PhD completion. I hope that I have made all of you proud.
SYMBOLS AND ABBREVIATIONS

Symbols

~  approximately
=  equals
>  greater than
<  less than
-  minus
%  percent
x  times
±  plus-minus sign

Abbreviations

µg  microgram
µL  microlitre
µM  micro Molar
AIDS  Autoimmune deficiency syndrome
AMP  Ampicillin
ANOVA  Analysis of variance
bp  Base pair
°C  Celcius centigrade
CDCD  Communicable Disease Control Directorate
CI  Confidence interval
CO₂  Carbon dioxide
DNA  Deoxyribonucleic acid
dNTP  Deoxynucleoside triphosphate
EDTA  Ethylenediamine tetra acetic acid, tri-potassium salt
Eg or e.g.  Exempli gratia – for example
ELISA  Enzyme-linked immunosorbent assay
et al.  And others  

g  unit of gravitation field  
g  gram  

GC-MS  Gas chromatography-mass spectrometry  
gp60  60 kDa glycoprotein  
H<sub>2</sub>O  Water  
HCl  hydrogen chloride  
i.e.  id est – in other words  
IMS  immunomagnetic separation  
kb  Kilobase  
kDa  kilodalton  
Km  Kilometer  
KPHU  Kimberly Public Health Unit  
L  Liter  
m  meters  
m<sup>3</sup>  cubic meters  
M  Molar concentration  
mg  Milligram  
MgCl<sub>2</sub>  Magnesium chloride  
min  Minute  
ML  Maximum likelihood  
mL  Millilitre  
mM  Milli molar  
mm  Millimetre  
MP  Maximum parsimony  
m/z  mass to charge ratio  
n  Number  
NA  Nucleus to anterior measurement  
Na+  Sodium  
NCBI  National Center for Biotechnology Information
ng  Nanograms
NJ  Neighbour Joining; algorithm for constructing phylogenetic trees
NSW New South Wales
ODOO  optimal date of onset
p  Probability of an event due to chance alone
PCA  Principal Component analysis
PCR  Polymerase chain reaction
PCR-RFLP  Polymerase chain reaction-restriction fragment length polymorphism
pH  Negative log of hydrogen ion concentration?
PHU  public health uni
pmol  Picomoles
ppm  parts per million
Qld  Queensland
qPCR  Quantitative polymerase chain reaction
rDNA  Ribosomal deoxyribonucleic acid
RNA  Ribonucleic acid
RPM  Revolutions per minute
RSD  Relative standard deviation
SD  Standard deviation
SE  Standard error
sec  seconds
sp.  Unknown species
sp. n  Novel species
spp.  Several species
SPSS  Statistics Package for Social Studies
18S rRNA  18S ribosomal RNA
TAE  Tris-acetate-EDTA buffer
Taq  *Thermus aquaticus* deoxyribonucleic acid polymerase
TEM  Transmission electron microscopy
USA  United States of America
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>UV</td>
<td>Ultra violet light</td>
</tr>
<tr>
<td>vs.</td>
<td>versus</td>
</tr>
<tr>
<td>WA</td>
<td>Western Australia</td>
</tr>
<tr>
<td>WANIDD</td>
<td>Western Australia Notifiable Infectious Disease Database</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>
PUBLICATIONS AND CONFERENCES

Publications arising from this thesis:


Abstracts in conference proceedings


Ng-Hublin, Hargraves, D., Coombs, B., Ryan, U. 2013. A cryptosporidiosis swimming pool associated outbreak in Kimberley, Western Australia. XXIV International Conference of the World Association for the Advancement of Veterinary parasitology (WAAVP); 25-29 Aug 2013; Perth, WA, Australia. p. 44.

Ng-Hublin, Coombs, B., Mackenzie, B., Ryan, U. 2013. Human cryptosporidiosis in Western Australia – A case of mixed *Cryptosporidium* infection. XXIV International Conference of the World Association for the Advancement of Veterinary parasitology (WAAVP); 25-29 Aug 2013; Perth, WA, Australia.
# TABLE OF CONTENTS

## CHAPTER 1

- Literature review ................................................................. 1
  - 1.1 General introduction .................................................... 1
  - 1.2 Historical timeline ....................................................... 2
  - 1.3 *Cryptosporidium* taxonomy ........................................... 3
  - 1.4 *Cryptosporidium* species ............................................. 5
  - 1.5 *Cryptosporidium* life cycle .......................................... 9
  - 1.6 Clinical symptoms and treatment .................................... 13
  - 1.7 Transmission of *Cryptosporidium* ................................ 15
  - 1.8 Detection and Characterisation ...................................... 17
    - 1.8.1 Microscopic methods ............................................... 17
    - 1.8.2 Immunological Methods .......................................... 17
    - 1.8.3 Molecular detection and characterisation methods .... 18
  - 1.9 *Cryptosporidium* metabolism and metabolomics ............. 22
    - 1.9.1 Metabolomics ..................................................... 23
  - 1.10 Aims and objectives .................................................. 25

## CHAPTER 2.0

- General Materials and Methods ........................................... 26
  - 2.1 Genomic DNA extraction from faecal samples .................. 26
  - 2.2 Polymerase Chain Reaction .......................................... 27
  - 2.3 Agarose gel electrophoresis ........................................ 27
  - 2.4 Gel purification of PCR products .................................. 29
  - 2.5 Sequencing .............................................................. 29
    - 2.5.1 Ethanol precipitation of sequencing reaction ............ 30
  - 2.6 Analysis of sequence chromatogram .............................. 30
  - 2.7 Phylogenetic analysis ................................................ 30

## CHAPTER 3

- Differences in cryptosporidiosis among Aboriginal and non-Aboriginal humans in Western Australia ........................................ 32
  - 3.1 Introduction ............................................................. 32
  - 3.2 Materials and Methods ................................................. 33
    - 3.2.1 Case definition .................................................... 33
    - 3.2.2 Data source and extraction .................................... 33
    - 3.2.3 Molecular typing .................................................. 34
    - 3.2.4 Data analysis ...................................................... 35
CHAPTER 4  
Comparison of three cryptosporidiosis outbreaks in Western Australia – 2003, 2007 and 2011  

4.1 Introduction  
4.2 Methods  
4.2.1 Outbreak periods, data source and analysis  
4.2.2 Molecular identification and analysis  
4.2.3 Data analysis  
4.2.4 Ethical clearance  
4.3 Results  
4.3.1 2003 Outbreak: December 2002 – June 2003  
4.3.2 2007 Outbreak: November 2006 – August 2007  
4.3.3 2011 Outbreak: January 2011 – May 2011  
4.4 Discussion  

CHAPTER 5  
Investigation of a community cryptosporidiosis outbreak in the Kimberley - November/December 2012  

5.1 Introduction  
5.2 Materials and Methods  
5.2.1 Case definition  
5.2.2 Epidemiological Investigation  
5.2.3 Water samples  
5.2.4 Clinical samples  
5.2.5 Molecular characterisation and analysis DNA Extraction, PCR amplification and Sequence analysis  
5.3 Results  
5.3.1 Epidemiological Investigations  
5.3.2 Molecular analysis of clinical samples  
5.3.3 Molecular analysis of water samples  
5.4 Discussion
CHAPTER 6 ................................................................. 91
Evidence of *Cryptosporidium* transmission between cattle and humans in northern New South Wales ................................................................. 91
6.1 Introduction ...................................................................... 91
6.2 Material and Methods ................................................... 91
   6.2.1 Sample selection .................................................. 91
   6.2.2 DNA extraction .................................................. 92
   6.2.3 PCR amplification and sequence analysis .................. 93
   6.2.4 Statistical analysis ............................................. 93
6.3 Results ........................................................................... 94
   6.3.1 *Cryptosporidium* species and subtypes in calves .......... 94
   6.3.2 *Cryptosporidium* species and subtypes in humans .......... 94
   6.3.3 Epidemiological analysis .................................... 98
6.4 Discussion ...................................................................... 99

CHAPTER 7 ........................................................................ 103
Development of an untargeted metabolomics method for the analysis of human faecal samples using *Cryptosporidium*-infected samples .................................................. 103
7.1 Introduction .................................................................... 103
7.2 Materials and methods ................................................. 104
   7.2.1 Chemicals ......................................................... 104
   7.2.2 Samples .......................................................... 104
   7.2.3 Sample preparation .......................................... 104
   7.2.4 Sample preparation for optimised method ............... 105
   7.2.5 Derivatisation .................................................. 106
   7.2.6 Analysis ......................................................... 106
   7.2.7 Data analysis ................................................... 106
7.3 Results ........................................................................... 107
   7.3.1 Optimisation of faecal weight ............................ 107
   7.3.2 Extraction solvents ......................................... 107
   7.3.3 Reproducibility of metabolite recovery .................. 108
   7.3.4 Metabolites identified in *Cryptosporidium* positive and *Cryptosporidium* negative faecal samples .......................................................... 108
7.4 Discussion ...................................................................... 112

CHAPTER 8 ........................................................................ 116
Metabolomic profiling of faecal extracts from *Cryptosporidium parvum* infection in experimental mouse models ............................................ 116
8.1 Introduction .................................................................... 116
**LIST OF FIGURES**

**Figure 1.1:** Diagrammatic representation of the Cryptosporidium life cycle (Barta and Thompson, 2006).

**Figure 1.2:** Subtyping at the gp60 locus. Example of a IIaA17G2R1 repeat.

**Figure 3.1:** Map of PHU’s in Western Australia – urban Perth is divided into North and South Metropolitan regions.

**Figure 3.2:** Comparison of cryptosporidiosis notification rates (cases/100,000 population) between Aboriginal and non-Aboriginal people, stratified by age groups.

**Figure 3.3:** Phylogenetic relationships of C011 with other Cryptosporidium spp. at the actin gene locus inferred by neighbour-joining analysis based on genetic distances calculated using Kimura-2 parameters. Bootstrap values from 1000 pseudoreplicates of >60% are shown.

**Figure 3.4:** Phylogenetic relationships of C011 and other Cryptosporidium spp. gp60 subtype families inferred by neighbour-joining (NJ) analysis based on genetic distances calculated using Kimura-2 parameters. Cryptosporidium fayeri gp60 sequence was used as an outgroup to generate the NJ tree. Bootstrap values from 1000 pseudoreplicates of >50% are shown.

**Figure 4.1:** Epidemiological curve showing the number of cryptosporidiosis cases notified from January 2002 to December 2011 in metropolitan, rural and remote areas in WA.

**Figure 4.2:** Cryptosporidiosis notifications by race in each age group for the 2003, 2007 and 2011 outbreaks.

**Figure 7.1:** Mass spectra from gas chromatography-mass spectrometry (GC-MS) analysis of pooled faecal specimens; (A) extraction of 250 mg of samples with a final concentration of 200 mg/ml, (B) 500 mg samples with a final concentration of 100 mg/ml, (C) 500 mg samples with a final concentration of 200 mg/ml and (D) 1000 mg samples with a final concentration of 200 mg/ml.

**Figure 7.2:** Principal component analysis generated scores plot of Cryptosporidium positive and negative faecal metabolite profiles showing a distinct pattern differentiating the two groups. Cryptosporidium positive samples are represented by ■ and Cryptosporidium negative samples are represented by ●.

**Figure 8.1:** Principal component analysis of mice faecal metabolite profiles. Score plots showed distinct clustering patterns of metabolite profiles of C. parvum infected mice and the uninfected control mice, differentiating the two groups. Cryptosporidium infected mouse samples are represented by ● and uninfected control mouse samples are represented by ■.

**Figure 8.2:** Faecal metabolites that contributed to the variance between infected and uninfected mice. Metabolites contributing to the variance between faecal metabolite profiles of mice infected with C. parvum and the uninfected control mice were compared to metabolites which contributed to the variance of faecal metabolite profiles of humans infected and not infected with Cryptosporidium from Chapter 7). Heat map was generated using the mean of normalised peak area of the metabolite. All mice faecal metabolites showed statistically significant (p < 0.05) contributions to the variance except for those denoted by ^, where p > 0.05.
LIST OF TABLES

Table 1.1: Valid Cryptosporidium species confirmed by molecular analysis.

Table 2.1: Primers used in the characterisation of Cryptosporidium sp. at the 18S rRNA, actin and gp60 loci.

Table 3.1: The annual number, notification rate, NR (notifications per 100,000/year) and notification rate ratio (NRR) of cryptosporidiosis notification in Aboriginal and non-Aboriginal people in Western Australia (2002-2012).

Table 3.2: The age distribution, notification rate, NR (notifications per 100,00/year) and notification rate ratio (NRR) of cryptosporidiosis notification in Aboriginal and non-Aboriginal people in Western Australia (2002-2012).

Table 3.3: Average number of notifications and notification rates (NR) of cryptosporidiosis cases in each PHU from 2002-2012 by Aboriginal status.

Table 3.4: Cryptosporidium species and subtypes identified in Aboriginals and non-Aboriginals in WA.

Table 3.5: Distribution of C. hominis and C. parvum gp60 genotypes in remote, rural and metropolitan regions of WA by Aboriginality.

Table 4.1: Summary of case notification numbers and rates (per 100,000 population) by population groups in the 2003 (December 2002 – June 2003), 2007 (November 2006 – August 2007) and 2011 (January 2011 – May 2011) outbreaks, and a 5-year average (of 7-month periods) in previous, non-outbreak years.

Table 4.2: Cryptosporidium species and gp60 subtypes identified in cases of cryptosporidiosis during two outbreaks (2007 and 2011) in Western Australia, stratified by Public Health Unit.

Table 4.3: Cryptosporidium gp60 subtypes identified in cases of cryptosporidiosis from the 2007 and 2011 outbreak in Western Australia stratified by Indigenous status; Aboriginal people, non-Aboriginal people and unknown.

Table 5.1: Details of possible Cryptosporidium exposure of cases and Cryptosporidium species and subtypes identified from cases where molecular genotyping was carried out.

Table 6.1: Prevalence of Cryptosporidium species and subtypes in cattle and farm workers on the 20 farms analysed in the upper Hunter Valley and Tamworth NSW.

Table 6.2: Associations between risk factors examined and the presence of Cryptosporidium in humans or report of humans with diarrhoea.

Table 7.1: Metabolites identified in human faecal samples that contributed to the variance observed between Cryptosporidium positive and negative faecal samples. Metabolites that showed statistical significance in abundance, where p < 0.05, are denoted by *.

Table 8.1: Mean of normalised peak area of metabolites (compounds) identified in faecal
samples of mice infected with *C. parvum* and the uninfected control mice compared to faecal metabolite profile of humans infected and not infected with *Cryptosporidium* from Chapter 7).
CHAPTER 1

Literature review

1.1 General introduction

*Cryptosporidium* is a parasitic protozoan belonging to the phylum Apicomplexa that infects a wide range of hosts including humans, birds, reptiles and fish (Ryan et al. 2014; Zahedi et al., 2016). In immunocompetent humans, the infection is usually self-limiting (Ryan et al., 2016a). However, in compromised immune systems, it can be the cause of chronic diarrhoea, cachexia, lack of appetite, malnutrition, and can lead to death (Ryan et al., 2016a).

*Cryptosporidium* is the second most common cause of diarrhoea worldwide and death in children, after rotavirus (Kotloff et al. 2013; Striepen 2013). It is also associated with retarded cognitive and functional development in children for up to 7 years post-infection, during which time they have no immunity against additional *Cryptosporidium* parasitisation (Kirkpatrick et al. 2002; Kotloff et al. 2013; Valenzuela et al. 2014).

Although human cryptosporidiosis is more prevalent in developing countries, sporadic episodes do occur in Westernised societies, costing millions in public health burdens (Xiao, 2010). The largest outbreak, which occurred in Milwaukee, Wisconsin (MacKenzie et al., 1995), was estimated to cost $96.2 million, with HIV sufferers accounting for 74% of infections and 69 deaths in the HIV population relating directly to the *Cryptosporidium* outbreak (Corso et al, 2003). The ongoing cost of waterborne disease was estimated at $21.9
billion (annually), and Cryptosporidium forms a substantial percentage of this (Corso et al, 2003). Once an outbreak occurs in the water supplies, it can be extremely difficult to eradicate due to the parasite’s resistance to disinfection procedures including, chlorination (Ryan et al., 2014).

Cryptosporidium can also result in major economic losses in livestock, particularly newborn calves, causing retardation of growth and death. It is mostly a problem in ruminants and the poultry industry, affecting turkeys and chickens (Ryan et al., 2014).

1.2 Historical timeline

Cryptosporidium developmental stages were first discovered in the gastric glands of laboratory mice by Ernest Edward Tyzzer (1907, 1910). The species, which was given the name Cryptosporidium muris, was later found to infect the gastric glands of several other mammals, but not humans. Tyzzer identified a second species in 1912, which was found to infect the small intestine of laboratory mice and was named C. parvum (Tyzzer, 1912). Differentiation of the two species, C. muris and C. parvum, was based on differences in oocyst morphology and infection site. Cryptosporidium muris has large ovoid oocysts (6 - 8 µm) and infects the epithelial cells of the gastric glands in the stomach, whereas, C. parvum has smaller oocysts measuring 4.5 - 5.0 µm and infects the epithelial cells of the small intestine. In 1955, a new species was reported infecting the terminal one-third of the small intestine in turkeys and was named C. meleagridis based on cytology and measurements of oocysts and other life cycle stages (Slavin, 1955).

The medical, veterinary and economic importance of Cryptosporidium was not recognised until the 1970’s, more than half a century after Ernest Edward Tyzzer first discovered it. In 1971 it was found to be associated with bovine diarrhea (Panciera et al.,
1971) and in 1976 saw the first two reports of cryptosporidiosis in immuno-compromised human patients (Miesel et al., 1976). Six years later, the U.S. Centers for Disease Control reported 21 men in six cities with concurrent cryptosporidiosis and AIDS (Goldfarb et al., 1982). *Cryptosporidium* was first reported in reptiles by Brownstein et al., in 1977 and subsequently numerous reports world-wide have revealed *Cryptosporidium* infections to be associated with acute and clinical disease characterised by diarrhoea in humans and various domestic and wild animal species including birds (O’Donoghue, 1995; Fayer et al., 2000; Xiao, 2010; Ryan and Xiao, 2014; Zahedi et al., 2016).

### 1.3 *Cryptosporidium* taxonomy

Until recently, *Cryptosporidium* was classified as a coccidian parasite. However, it has long been speculated that *Cryptosporidium* represents a ‘missing link’ between the more primitive gregarine parasites and coccidians (Ryan et al., 2016b). The similarities between *Cryptosporidium* and gregarines have been supported by extensive microscopic, molecular, genomic and biochemical data (Pohlenz et al., 1978; Bull et al., 1998; Carreno et al., 1999; Beyer et al., 2000; Hijjawi et al., 2002; Leander et al., 2003a; Hijjawi et al., 2004; Rosales et al., 2005; Barta and Thompson, 2006; Butaeva et al., 2006; Valigurová et al., 2007; Boxell et al., 2008; Karanis et al., 2008; Zhang et al., 2009; Borowski et al., 2008; 2010; Hijjawi, 2010; Hijjawi et al., 2010; Templeton et al., 2010; Karanis and Aldeyarbi, 2011; Boxell, 2012; Koh et al., 2013, 2014; Huang et al., 2014; Clode et al., 2015; Valigurová et al., 2015; Aldeyarbi and Karanis, 2016a; 2016b; 2016c; Edwinson et al., 2016; Paziewska-Harris et al., 2016), which have served as the basis for the formal transfer of *Cryptosporidium* from subclass Coccidia, class Coccidiomorphea to a new subclass, Cryptogregararia, within class Gregarinomorphea (Cavalier-Smith, 2014). The genus *Cryptosporidium* is currently the sole
member of Cryptogregaria and is described as comprising epicellular parasites of vertebrates possessing a gregarine-like feeder organelle but lacking an apicoplast (Cavalier-Smith, 2014).

Gregarines (phylum Apicomplexa; class Gregarinomorphea) are a very diverse group of large, single-celled “primitive” apicomplexan parasites that primarily infect the intestines and other extracellular spaces of invertebrates and lower vertebrates (mainly arthropods, molluscs and annelids), which are abundant in natural water sources (Leander et al., 2003a; 2003b; Barta and Thompson, 2006; Leander, 2007; Valigurová et al., 2007). The transmission of gregarines to new hosts usually takes place by oral ingestion of oocysts in both aquatic and terrestrial environments. Four or more sporozoites (depending on the species) escape from the oocysts, find their way to the appropriate body cavity and attach to, or penetrate, the host cells. The sporozoites emerge from a host cell, begin to feed and develop into large trophozoites (Rueckert and Leander, 2008).

Many gregarines do not exhibit intracellular stages and are mostly epicellular parasites. The gregarine life cycle typically only consists of gametogony and sporogony and only a few species exhibit merogony. The sporozoites will generally develop into large trophozoites and attach to the host cell with a specialised attachment apparatus (epimerite, mucron, modified protomerite) (MacMillan, 1973). These specialised structures are derived from the conoid at the apical end. This attachment to the host cell also functions in feeding in that the cytoplasm of the host is taken up by the attached parasite (i.e., myzocytosis) (Valigurová et al., 2007). Two mature trophozoites eventually pair up in a process called syzygy and develop into gamonts. The orientation of gamonts during syzygy differs depending on the species (e.g. side-to-side and head-to-tail). A gametocyst wall forms around each pair of gamonts, which then begins to divide into hundreds of gametes (gametogeny). Pairs of gametes fuse and form zygotes, each of which becomes surrounded by an oocyst wall. Within the oocyst, meiosis occurs to yield four or more spindle-shaped sporozoites.
(sporogony). Hundreds of oocysts accumulate within each gametocyst, and are usually released via host faeces or via host death and decay (Vivier and Desportes, 1990; Kuriyama et al., 2005; Rueckert and Leander, 2008).

The gregarines are thought to be the earliest lineage of apicomplexans (Rueckert and Leander, 2008) and were previously subdivided into three orders; Archigregarinida, Eugregarinida and Neogregarinida (Grassé, 1953; Adl et al., 2012). However, the taxonomy has recently been revised (Cavalier-Smith, 2014), on the basis that it was phylogenetically unsound (Rueckert et al., 2011). In this new classification, the class name Gregarinomorphea has been adopted to broadly refer to all its members (i.e. gregarines, *Cryptosporidium* and *Histogregaria*) (Cavalier-Smith, 2014). Within the various subclasses of Gregarinomorphea are Cryptogregaria, discussed above, and Orthogregarinia (comprising the orders Vermigregarida and Arthrogregarida), for gregarines most closely related to *Cryptosporidium* (Cavalier-Smith, 2014).

### 1.4 *Cryptosporidium* species

The first species of *Cryptosporidium* to be described were *C. muris* (Tyzzer, 1910) and *C. parvum* (Tyzzer, 1912). For many years, following the initial discovery of these species, there was much confusion surrounding *Cryptosporidium* taxonomy and numerous species were described, many erroneously. Initially, the parasite was commonly confused with other apicomplexan genera, particularly the coccidian genus *Sarcocystis*, resulting in a variety of named and un-named species being incorrectly assigned to the genus (Triffitt, 1925; Bearup, 1954; Anderson et al., 1968; Dubey and Duszynski, 1969; Pande et al., 1972). Following this, the concept of strict host specificity was applied and multiple new species were named based on host occurrence including: *C. agni* in sheep, *C. bovis* in calves (Barker
and Carbonell, 1974), C. anserinum in geese (Proctor and Kemp, 1974), C. cuniculus in rabbits (Inman and Takeuchi, 1979), C. garnhami in humans (Bird, 1981), and C. rhesi in monkeys (Levine, 1980). However, cross transmission studies demonstrated that isolates from different animals could be transmitted from one host species to another, which ended this practice. Many of these new Cryptosporidium species became a synonymous group referred to as C. parvum resulting in the widespread use of this name for Cryptosporidium spp. infecting a wide range of mammals including humans. However, the following species; C. meleagridis in turkeys (Slavin, 1955), C. wrairi in guinea pigs (Vetterling et al., 1971), and C. felis in cats (Iseki, 1979) retained their name due to demonstrated biological differences from the established species C. parvum and C. muris. Following this, other Cryptosporidium species were named based on biological differences, such as C. baileyi in chickens (Current et al., 1986), C. serpentis in reptiles (Levine, 1980) and C. varanii in lizards (Pavlásek et al., 1995). More recently, it was recognised that C. parvum in fact consists of two distinct species; C. parvum and C. hominis (Morgan et al., 1997; Morgan-Ryan et al., 2002).

The application of molecular characterisation tools has helped to clarify confusion surrounding Cryptosporidium taxonomy and the existence of multiple species has been validated (Xiao, 2010). Xiao et al., (2004) devised a taxonomic framework used as a guide in the description of new species of Cryptosporidium in order to minimise confusion and the creation of invalid names for species.

Currently, 33 Cryptosporidium species are considered valid (Ryan and Xiao., 2014; Ryan et al., 2014; Holubová et al., 2016; Kváč et al., 2016; Jezkova et al., 2016; Zahedi et al., 2016) (Table 1.1). In humans, more than 20 species have been identified but C. hominis and C. parvum account for more than 90% of human cases of cryptosporidiosis (Xiao, 2010; Ryan and Power, 2012; Ryan et al., 2014), with C. hominis the main cause of human
cryptosporidiosis in most countries (Xiao, 2010; Ryan et al., 2014). Worldwide, \textit{C. hominis} and \textit{C. parvum} infects 8–10 million people per year (Seeber and Steinfeld, 2016).

Many of the \textit{Cryptosporidium} species are morphologically identical but information on host range and other phenotypic characteristics together with molecular phylogenetic data is gradually being collated from different parts of the world, supporting the taxonomic validity of the 33 species currently recognised (Table 1.1).

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|l|l|}
\hline
Species name & Author(s) & Type host(s) & Major host(s) & Reports in humans \\
\hline
\textit{C. ducismarci} & Traversa 2010; Jezkova et al. 2016 & Marginated tortoise (\textit{Testudo marginata}) & Tortoises & None reported \\
\hline
\textit{C. testudinis} & Jezkova et al. 2016 & Russian tortoise (\textit{Testudo horsfieldii} Gray) & Tortoises & None reported \\
\hline
\textit{C. avium} & Holubová et al., 2016 & Red-crowned parakeets (\textit{Cyanoramphus novaезealandiae}) & Birds & None reported \\
\hline
\textit{C. proliferans} & Kváč et al., 2016 & East African mole rat (\textit{Tachyoryctes splendens}) & Rodents & None reported \\
\hline
\textit{C. rubeyi} & Li et al., 2015 & \textit{Spermophilus beecheyi} (California ground squirrel) & Squirrels & None reported \\
\hline
\textit{C. scophthalmi} & Alvarez-Pellitero et al., 2004; Costa et al., 2016 & \textit{Scophthalmus maximus} (turbot) & Turbot & None reported \\
\hline
\textit{C. huwi} & Ryan et al., 2015 & \textit{Poecilia reticulata} (guppy), \textit{Paracheirodon innesi} (neon tetra) and \textit{Puntius tetrazona} (tiger barb) & Fish & None reported \\
\hline
\textit{C. erinacei} & Kváč et al., 2014b & \textit{Erinaceus europaeus} (European hedgehog) & Hedgehogs, horses & Kváč et al., 2014a \\
\hline
\textit{C. scrofarum} & Kváč et al., 2013a & \textit{Sus scrofa} (pig) & Pigs & Kváč et al. et al., 2009a; Kváč et al. \\
\hline
\end{tabular}
\caption{Valid \textit{Cryptosporidium} species confirmed by molecular analysis.}
\end{table}
<table>
<thead>
<tr>
<th>Species</th>
<th>Authors/Year</th>
<th>Host Species</th>
<th>Host(s)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. viatorum</em></td>
<td>Elwin et al., 2012a</td>
<td><em>Homo sapiens</em> (human)</td>
<td>Humans</td>
<td>Elwin et al., 2012a, 2012b; Insulander et al., 2013</td>
</tr>
<tr>
<td><em>C. tyzzeri</em></td>
<td>Tyzzer, 1912; Ren et al., 2012</td>
<td><em>Mus musculus</em> (mouse)</td>
<td>Rodents</td>
<td>Rasková et al., 2013</td>
</tr>
<tr>
<td><em>C. cuniculus</em></td>
<td>Robinson et al., 2010</td>
<td><em>Oryctolagus cuniculus</em> (European rabbit)</td>
<td>Rabbits</td>
<td>Chalmers et al., 2009; Anon, 2010; Molloy et al., 2010; Chalmers et al., 2011a; Koehler et al., 2014a; Koehler et al., 2014b</td>
</tr>
<tr>
<td><em>C. ubiquitum</em></td>
<td>Fayer et al., 2010</td>
<td><em>Bos taurus</em> (Cattle)</td>
<td>Ruminants, rodents, primates</td>
<td>Commonly reported (cf. Fayer et al., 2010; Elwin et al., 2012b)</td>
</tr>
<tr>
<td><em>C. xiaoi</em></td>
<td>Fayer et al., 2009</td>
<td><em>Ovis aries</em> (Sheep)</td>
<td>Sheep and goats</td>
<td>Adamu et al., 2014</td>
</tr>
<tr>
<td><em>C. ryanae</em></td>
<td>Fayer et al., 2008</td>
<td><em>Bos taurus</em> (Cattle)</td>
<td>Cattle</td>
<td>None reported</td>
</tr>
<tr>
<td><em>C. macropodum</em></td>
<td>Power and Ryan, 2008</td>
<td><em>Macropus giganteus</em> (Kangaroo)</td>
<td>Marsupials</td>
<td>None reported</td>
</tr>
<tr>
<td><em>C. fragile</em></td>
<td>Jirků et al., 2008</td>
<td><em>Duttaphrynus melanostictus</em> (Toad)</td>
<td>Toads</td>
<td>None reported</td>
</tr>
<tr>
<td><em>C. fayeri</em></td>
<td>Ryan et al., 2008</td>
<td><em>Macropus rufus</em> (Kangaroo)</td>
<td>Marsupials</td>
<td>Waldron et al., 2010</td>
</tr>
<tr>
<td><em>C. bovis</em></td>
<td>Fayer et al., 2005</td>
<td><em>Bos taurus</em> (Cattle)</td>
<td>Cattle</td>
<td>Khan et al., 2010; Helmy et al., 2013</td>
</tr>
<tr>
<td><em>C. suis</em></td>
<td>Ryan et al., 2004</td>
<td><em>Sus scrofa</em> (Pig)</td>
<td>Pigs</td>
<td>Xiao et al., 2002; Leoni et al., 2006; Cama et al., 2007; Wang et al., 2013</td>
</tr>
<tr>
<td><em>C. galli</em></td>
<td>Pavalasek, 1999; Ryan et al., 2003b</td>
<td>Spermestidae, Frangillidae, <em>Gallus gallus</em>, <em>Tetrao urogallus</em>, <em>Pinicola enucleator</em> (birds)</td>
<td>Birds</td>
<td>None reported</td>
</tr>
<tr>
<td><em>C. hominis</em></td>
<td>Morgan-Ryan et al., 2002</td>
<td><em>Homo sapiens</em> (human)</td>
<td>Humans</td>
<td>Most common species in humans</td>
</tr>
<tr>
<td><em>C. molnari</em></td>
<td>Alvez-Pellitero and Sitja-Bobadilla, 2002</td>
<td><em>Sparus aurata</em> (gilt-head sea bream) and <em>Dicentrarchus labrax</em> (European seabass)</td>
<td>Fish</td>
<td>None reported</td>
</tr>
<tr>
<td><em>C. canis</em></td>
<td>Fayer et al., 2001</td>
<td><em>Canis familiaris</em> (Dog)</td>
<td>Dogs</td>
<td>Many reports (cf. Lucio-Forster et al., 2010)</td>
</tr>
<tr>
<td><em>C. andersoni</em></td>
<td>Lindsay et al., 2000</td>
<td><em>Bos taurus</em> (Cattle)</td>
<td>Cattle</td>
<td>Leoni et al., 2006; Morse et al., 2007; Waldron et al., 2011a; Agholi et al., 2013; Jiang et al., 2014; Liu et al., 2014</td>
</tr>
<tr>
<td>Cryptosporidium</td>
<td>Host</td>
<td>Reference</td>
<td>Host Species</td>
<td>Host Category</td>
</tr>
<tr>
<td>----------------</td>
<td>------</td>
<td>-----------</td>
<td>--------------</td>
<td>---------------</td>
</tr>
<tr>
<td>C. varanii</td>
<td>Pavlasek et al., 1995</td>
<td>Varanus prasinus (Emerald Monitor)</td>
<td>Lizards</td>
<td>None reported</td>
</tr>
<tr>
<td>C. baileyi</td>
<td>Current et al., 1986</td>
<td>Gallus gallus (Chicken)</td>
<td>Birds</td>
<td>None reported</td>
</tr>
<tr>
<td>C. parvum</td>
<td>Tyzzer, 1912</td>
<td>Bos taurus (Cattle)</td>
<td>Ruminants</td>
<td>Commonly reported in humans</td>
</tr>
<tr>
<td>C. meleagridis</td>
<td>Slavin, 1955</td>
<td>Meleagris gallopavo (Turkey)</td>
<td>Birds and humans</td>
<td>Commonly reported in humans</td>
</tr>
<tr>
<td>C. serpents</td>
<td>Levine, 1980</td>
<td>Elaphe guttata, E. subocularis, Sanzinia madagascarensus (Snakes)</td>
<td>Snakes and lizards</td>
<td>None reported</td>
</tr>
<tr>
<td>C. felis</td>
<td>Iseki, 1979</td>
<td>Felis catis (Cat)</td>
<td>Cats</td>
<td>Many reports (cf. Lucio-Forster et al., 2010)</td>
</tr>
<tr>
<td>C. wraithi</td>
<td>Vetterling et al., 1971</td>
<td>Cavia porcellus (Guinea pig)</td>
<td>Guinea pigs</td>
<td>None reported</td>
</tr>
<tr>
<td>C. muris</td>
<td>Tyzzer, 1907; and 1910</td>
<td>Mus musculus (House mouse)</td>
<td>Rodents</td>
<td>Many reports - Guyot et al., 2001; Gatei et al., 2002; Tiangtip and Jongwutiwes, 2002; Gatei et al., 2003; Palmer et al., 2003; Gatei et al., 2006a; Leoni et al., 2006; Muthusamy et al., 2006; Azami et al., 2007; Al-Brik et al., 2008; Neira et al., 2012; Hasajová et al., 2014; Petrinčová et al., 2015; Spanakos et al., 2015</td>
</tr>
</tbody>
</table>

1.5 Cryptosporidium life cycle

The complex, monoxenous life cycle of Cryptosporidium consists of several developmental stages involving both sexual and asexual cycles which are demonstrated in Figure 1.1. Until recently, it was thought that Cryptosporidium were obligate intracellular parasites that completed their life-cycle in an intra-cellular but extra-cytoplasmic (epicellular) location by pulling the host cell membrane around it as an extracytoplasmic “parasitophorous
“sac/membrane” that sequestrated the parasite from the intestinal lumen and the host cell’s cytoplasm (Tzipori and Ward, 2002; Dumenil, 2011). However, the initial description of the complete development of Cryptosporidium in axenic culture (without attachment to host cells) by Hijjawi et al. in 2004, revealed that Cryptosporidium is not an obligate epicellular parasite and this has been confirmed by subsequent studies (Boxell et al., 2008; Hijjawi et al., 2010; Boxell, 2012; Yang et al., 2015; Aldeyarbi and Karanis, 2016a; 2016b; 2016c; Ryan and Hijjawi, 2015). Even when Cryptosporidium is cultivated with host cells, it has been reported that as C. parvum progresses through its life cycle, it becomes more extracellular with no evidence of attachment to cell lines found (Borowski et al., 2010).

The primary site of infection with C. hominis and C. parvum is the small intestine, with the ileum above the ceecal junction being favored in some animals such mice and calves (Xiao and Fayer, 2008). Other species favour the gastric mucosa such as C. muris (Tyzzer, 1910), C. serpentis (Levine, 1980), C. andersonii (Lindsay et al., 2000), C. fragile (Jirků et al., 2008) and C. molnari (Alverez-Pellittero et al., 2002). Cryptosporidium baileyi in chickens favours the respiratory tree and cloaca (Current et al., 1986).

Environmentally-resistant oocysts representing the infective life cycle stage of the parasite are excreted in the faeces (Xiao and Fayer, 2008). These oocysts are extremely resilient, due to their thick trilaminar walls, are able to survive many months in a watery environment and are resistant to disinfectants including chlorine in drinking water (Fayer, 2004). Once the oocysts are excreted into the environment, they can be ingested by a host through the faecal-oral route. Excystation occurs in the gastrointestinal tract; this results in the release of four infective sporozoites through a suture opening, which then attaches to the apical membrane of the epithelial host cell and undergo successive rounds of asexual and sexual reproduction (Xiao and Fayer, 2008) (Figure 1.1).
The motile sporozoites approach a potential host cell anterior end first and actively invade the cell (Fayer et al., 1997; Tzipori and Widmer, 2000). Following the adherence of the anterior end of the sporozoite to the luminal surface of an epithelial cell in the microvilli, each sporozoite matures into a trophozoite. Trophozoites undergo asexual proliferation by merogony to form meronts, marking the beginning of the asexual part of the life cycle (Fayer
et al., 1997; Hijjawi et al., 2010). Two types of meronts have been described in the *Cryptosporidium* life cycle (Current and Reese, 1986). Type I meronts develop six or eight nuclei, each incorporated into a merozoite, which are released from the parasitophorous vacuole once mature. The mature type I merozoites infect other host cells and either recycle as type I meronts and merozoites, or develop into a type II meront, which produces four merozoites (Current, 1990; Current and Garcia, 1991; Tzipori and Widmer, 2000; Hijjawi et al., 2010). Following the release of mature type II merozoites, a new host cell is invaded and the sexual phase in the life cycle (gametogeny) is initiated. Type II merozoites either enlarge and develop into a uni-nucleate macrogamont or undergo cellular fission forming a multi-nucleated microgamont containing 14-16 non-flagellated microgametes.

Microgametes are released from ruptured microgamonts; they penetrate host cells containing macrogamonts and subsequently fertilise the macrogamont forming a zygote (Current, 1990; Current and Garcia, 1991; Hijjawi et al., 2010). The zygote undergoes sporogony during which, both thin-walled and thick-walled oocysts are formed, each containing four potentially infective sporozoites. Thin walled oocysts remain within the host leading to autoinfection and persistent infections, thick walled oocysts are shed in the faeces into the environment, for ingestion by a new host.

The presence of gamont-like extracellular stages in the life cycle of *Cryptosporidium* was first observed in a study by Hijjawi et al. (2002) and has since been reported by several investigators (Hijjawi et al., 2004; Rosales et al., 2005; Karanis et al., 2008; Borowski et al., 2010; Koh et al., 2013, 2014; Huang et al., 2014; Aldeyarbi and Karanis, 2016a). Where these gamont stages occur in the life cycle and what stages they develop into is not clearly understood (Clode et al., 2015).
1.6 Clinical symptoms and treatment

The severity of a *Cryptosporidium* infection can vary from an asymptomatic shedding of oocysts to a severe and life-threatening disease. In immunocompetent individuals, the symptoms of cryptosporidiosis are usually self-limiting but, can be chronic in compromised immune systems, such as in children under 5 years, HIV patients and cancer patients (Chalmers and Davies, 2010). The general symptoms associated with cryptosporidiosis include diarrhoea, vomiting, nausea, lack of appetite and cramps (Chalmers and Davies, 2010). In immunocompromised individuals such as in HIV patients, *Cryptosporidium* is often found in other organs including the pancreas, spleen biliary tract and lungs (Sponseller et al, 2014).

The recent Global Enteric Multicenter Study (GEMS) and other studies to identify the aetiology and population-based burden of paediatric diarrheal disease revealed that *Cryptosporidium* is second only to rotavirus as a contributor of moderate-to-severe diarrheal disease during the first 5 years of life (Kotloff et al., 2013). Malnutrition, which impairs cellular immunity, is an important risk factor for cryptosporidiosis (Gendrel et al., 2003) and *Cryptosporidium* infection in children is associated with malnutrition, persistent growth retardation, impaired immune response and cognitive deficits (Mølbak et al., 1997; Guerrant et al., 1999; Mondal et al., 2009). The mechanism by which *Cryptosporidium* affects child growth seems to be associated with inflammatory damage to the small intestine (Kirkpatrick et al., 2002). Undernutrition (particularly in children) is both a sequela of- and a risk factor for- cryptosporidiosis (MacFarlane and Horner-Bryce, 1987; Sallon et al., 1988; Checkley et al., 1997; Bushen et al., 2007; Mondal et al., 2009; 2012; Quihui-Cota et al., 2015).

The immune status of the host, both innate and adaptive immunity, has a major impact on the severity of cryptosporidiosis and its prognosis (Ryan et al., 2016a). Immunocompetent individuals typically experience self-limiting diarrhea and transient gastroenteritis lasting up
to 2 weeks and recover without treatment, suggesting an efficient host anti-Cryptosporidium immune response. Immunocompromised individuals including HIV/AIDS patients (not treated with antiretroviral therapy) often suffer from intractable diarrhea, which can be fatal (Current and Garcia, 1991). HIV status is therefore an important host risk factor for cryptosporidiosis and although Cryptosporidium is an important pathogen regardless of HIV-prevalence (Kotloff et al., 2013), HIV-positive children are between 3 and 18 times more likely to have cryptosporidiosis than those who are HIV-negative (Tumwine et al., 2005; Mbae et al., 2013; Tellevik et al., 2015).

Although hundreds of drugs have been tested for prophylaxis and treatment of cryptosporidiosis in animals and humans, only one, Nitazoxanide (Alinia®), has been approved for use in humans by the US Food and Drug Administration (FDA). This drug, however, is currently not recommended for use in infants <12 months of age, exhibits only moderate clinical efficacy in children and immunocompetent people, and none in people with HIV (Abubakar et al, 2007; Amadi et al, 2009).

The use of ART (antiretroviral therapy) to maintain a high CD4+ immune cell titre, resulting in a strengthened immune system, has helped combat opportunistic Cryptosporidium in HIV+ and AIDS patients (Miao et al, 2000; Carr et al, 1998; Mengist et al., 2015). In the livestock industry, paramomycin has shown to reduce oocyst shedding and diarrhoea in lambs as well as improve their growth rate (Quilez et al, 2000; Johnson et al, 2000). Nasir et al. (2013) reported that azithromycin was also effective in reducing C. parvum oocyst shedding in calves.

There are currently no vaccines available for cryptosporidiosis (Mead, 2014), and due to the infective and widespread nature of Cryptosporidium, the most effective control of transmission is appropriate sanitation procedures (Ryan et al, 2014); guidelines are available for control strategies and drinking water safety, following the recommendations of the 2002
1.7 Transmission of Cryptosporidium

Infection may be acquired through direct contact with infected persons (person-to-person transmission) or animals (zoonotic transmission) or, through ingestion of contaminated food (foodborne transmission) or water (waterborne transmission) (Xiao, 2010; Zahedi et al., 2016). Numerous studies have demonstrated that respiratory cryptosporidiosis may occur commonly in both immunocompromised and immunocompetent individuals and that Cryptosporidium may also be transmitted via respiratory secretions (Sponseller et al., 2014). The symptoms associated with this route are respiratory (laryngotracheitis) and can be accompanied by mild diarrhea (Bouzid et al., 2013). Several studies also suggest that flies may play an important role in the mechanical transmission of Cryptosporidium including human infectious species (Graczyk et al., 2000; 2003; 2004; Szostakowska et al., 2004; Graczyk et al., 2005; Conn et al., 2007; Getachew et al., 2007; Fetene and Worku, 2009; Fetene et al., 2011; El-Sherbini and Gneidy, 2012; Adenusi and Adewoga, 2013; Zhao et al., 2014).

Cryptosporidium is particularly suited to waterborne transmission as the environmental stage, the oocyst, is highly resistant to oxidant-based disinfectant (Korich et al., 1990; Chauret et al, 2001; Painter et al, 2015) and is excreted in large quantities ($10^8$-$10^9$ oocysts in a single bowel movement) (Cordell, 2001; Yoder and Beach, 2007; Yoder et al., 2012) for up to 60 days after cessation of gastrointestinal symptoms (Jokipii and Jokipii, 1986; Stehr-Green et al., 1987). In addition, the low infectious dose (10-100 oocysts) (DuPont et al, 1995; Chappell et al, 2006), means that ingestion of a relatively small amount of contaminated water is sufficient to initiate infection in a susceptible individual. The long
incubation period of cryptosporidiosis (averaging 7 days) (Chalmers and Davies, 2010),
delays the identification of the source and implementation of interventions to de-contaminate
the source, thus perpetuating transmission.

A quantitative risk assessment has estimated that ingestion of a single oocyst of the C.
parvum IOWA isolate will result in clinical disease in 2.79% of immunologically normal
persons (Pouillot et al., 2004). Given that the 50% infective dose (ID50) for C. hominis is less
than one tenth that of the C. parvum IOWA isolate (DuPont et al., 1995), ingestion of only
one oocyst of a more infectious isolate may lead to a higher incidence of infection in the
general immunocompetent population.

Cryptosporidium is a major cause of waterborne outbreaks and was the etiological
agent in 60.3% (120) of outbreaks reported worldwide between January 2004 and December
2010 (Baldrusson and Karanis, 2011). The largest waterborne disease outbreak ever
documented in humans for any pathogen was the cryptosporidiosis outbreak in Milwaukee in
1993, where approximately one quarter of the population (~403,000 people) were reported
suffering from watery diarrhea. There were 90 deaths associated with this outbreak and it cost
over 90 million US dollars (MacKenzie et al., 1995). Australia appears to be particularly
vulnerable to outbreaks of waterborne disease; of 199 documented waterborne disease
outbreaks between 2004 and 2010, 46.7 % occurred in Australia; most of which, were
associated with swimming pools (Baldrusson and Karanis 2011).

Swimming is a very popular recreational activity worldwide and the majority of
recreational water outbreaks have been linked to swimming pools, with others documented at
water parks, water slides and fountains (Ryan et al. 2017). The most common cause of
recreational water outbreaks is a faecal accident (Ryan et al., 2017). In public pools, the
combination of frequent faecal contamination, oocyst resistance to chlorine, low infectious
dose and high bather densities facilitates transmission with the likelihood of transmission
increasing when recreational waters are used by diapered children, toddlers and incontinent persons (Fayer et al., 2000; Fayer, 2004; Ryan et al., 2016b).

1.8 Detection and Characterisation

1.8.1 Microscopic methods

Prior to the advent of molecular techniques, the only way to identify Cryptosporidium species was microscopic examination of purified or un-purified oocysts (Fayer et al., 2000). This technique, however, lacks sensitivity, is labour intensive and prone to human error (Fayer et al., 2000; McHardy et al., 2014). The most common staining techniques included differential staining methods (e.g. Methylene Blue and Ziehl-Neelson), fluorochrome staining and negative staining techniques (e.g. Malachite green and merbromide) (Kawamoto et al., 1987; Cambell et al., 1992; Chichino et al., 1991; Elliot et al., 1999). Due to lack of distinguishing oocyst features however, morphology cannot be used to definitively identify species and therefore molecular characterisation is required to identify to species level (Morgan et al., 1998; Fayer et al., 2000; Fall et al., 2003; Checkley et al., 2015).

1.8.2 Immunological Methods

Immunologicalbased methods include polyclonal fluorescent antibody tests, latex agglutination reactions, immunofluorescence (IF) with monoclonal antibodies (mAbs), enzyme-linked immunosorbent assays (ELISA), reverse passive haemagglutination (RPH), immunoserology using immunofluorescence detection and ELISA, and solid-phase qualitative immunochromatographic assays (Fayer et al., 2000; Chalmers and Katzer, 2013; Checkley et al., 2015). Cross-reactivity with other microorganisms can however occur due to
the non-specific nature of antibody-based methods and this can limit their use (Fayer et al., 2000; Checkley et al., 2015).

1.8.3 Molecular detection and characterisation methods

The application of advanced molecular techniques has led to an improved taxonomy and systematics, and better understanding of *Cryptosporidium* phylogeny (Ryan et al., 2014). Given the morphological similarity of oocysts by microscopy, these advances are crucial for confident identification, description of host/parasite interactions and a better understanding of the molecular epidemiology (Xiao, 2010; Zahedi et al., 2016).

1.8.3.1 18S ribosomal RNA (18s rRNA)

The 18S rRNA locus is considered the most reliable locus for detection and identification of all *Cryptosporidium* species and genotypes (Xiao, 2010; Ryan et al., 2014). It is multi-copy and has both hypervariable and semi-conserved regions which facilitate specific and sensitive detection (Xiao, 2010). In particular, a PCR-RFLP tool that targets an ~830-bp fragment of the gene and uses SspI and VspI restrictions for genotyping (Xiao et al., 1999; Xiao et al., 2001), is commonly used in the differentiation of *Cryptosporidium* species in humans, animals, and environmental samples. Another method commonly used for genotyping is DNA sequencing of PCR products (Koinari et al. 2013). In many *Cryptosporidium* species and genotypes, minor intra-isolate sequence variations are present among different copies of the 18S rRNA gene. Thus, new genotypes should not be named based on one or two nucleotide substitutions or insertions/deletions in the gene (Xiao, 2010).
1.8.3.2 Other loci

The actin gene is a widely distributed and highly conserved single copy microfilament protein that is thought to play a role in sporozoite gliding motility and penetration into host cells (Kim et al, 1992). Sequencing of the actin gene is another commonly employed characterisation technique. Sequence polymorphisms occur along this whole gene, caused by its ability to evolve rapidly (Xiao et al., 2004). Recently, a genus-specific Cryptosporidium qPCR based on the actin gene has been developed. Although it currently can only be used in screening of Cryptosporidium spp., genotyping can be done subsequently using species-specific qPCR (Yang et al., 2014).

PCR tools based on other genes in general only amplify DNA of C. parvum, C. hominis, C. meleagris, and species/genotypes closely related to them. Thus, studies that have used these tools have usually showed lower Cryptosporidium species diversity (Abd El Kader et al., 2012; Berrilli et al., 2012). These tools have limited usefulness in genotyping Cryptosporidium spp. of animals because of their narrow specificity. They, nevertheless, can be used in the identification of mixed infections with C. hominis or C. parvum in humans in developing countries that have been infected with rare Cryptosporidium species based on PCR analysis of the 18S rRNA gene (Cama et al., 2006). A few other markers such as the 90 kDa heat shock protein and A135 genes have been used in the development of genus-specific PCR-RFLP tools for genotyping Cryptosporidium (Feng et al. 2009; Tosini et al. 2010).

1.8.3.3 Subtyping tools for Cryptosporidium

Subtyping tools are essential to better understand the transmission dynamics of Cryptosporidium species and C. hominis and C. parvum in humans, animals and wildlife in particular (Xiao, 2010). One of the most common subtyping tools is the DNA sequence
analysis of the 60 kDa glycoprotein (gp60, also known as gp40/15) gene. This gene encodes a precursor protein, that is cleaved to produce mature cell surface glycoproteins (gp45/gp40 and gp15) implicated in zoite attachment to, and invasion of, enterocytes (Xiao, 2010; Ryan et al., 2014). Most of the genetic heterogeneity in this gene is the variation in the number of a tri-nucleotide repeat (TCA, TCG or TCT) in the 5’ end (gp40) of the coding region, although extensive sequence polymorphism is also present in the rest of the gene. The latter is used in defining subtype families within a species, whereas the former is used in identifying subtypes within a subtype family.

Figure 1.2: Subtyping at the gp60 locus. Example of a IIaA17G2R1 repeat.

An established subtype nomenclature is used in identifying gp60 subtype family. A subtype name starts with the species and subtype family designation (Ia, Ib, Id, Ie, If, etc. for C. hominis; IIa, IIb, IIc, IId, etc. for C. parvum; IIIa, IIIb, IIIc, IIId, etc. for C. meleagridis; followed by the number of TCA (represented by the letter A), TCG (represented by the letter G), or TCT (represented by the letter T) repeats (Sulaiman et al., 2005; Xiao, 2010; Feng et al. 2011a;). Thus, the name IeA11G3T3 indicates that parasite belongs to C. hominis subtype family Ie and has 11 copies of the TCA repeat, 3 copies of the TCG repeat, and 3 copies of the TCT repeat in the trinucleotide repeat region of the gene. Similarly, the subtype name IIaA17G1R1, indicates that the subtype belongs to C. parvum IIa subtype family, A17
indicates that the subtype has 17 copies of the TCA repeat, and G1 indicates that the subtype has one copy of the TCG repeat (Figure 1.2).

A few subtype families have variations in copy numbers of other repeat sequences (designed as R at the end of the subtype name). Thus, in the *C. parvum* IIa subtype family, some subtypes have two or three copies of the ACATCA sequence right after the trinucleotide repeats, which are represented by “R2” or R3 (R1 for most subtypes). Therefore, the only difference between subtypes IIaA15G1R1 and IIaA15G2R1 is the number of TCG repeats; one has one copy of the repeat, while the other has two copies. Currently, this locus is the most polymorphic marker known in the genome of *Cryptosporidium* (Ryan et al, 2014) and because of its high sequence heterogeneity, *gp60* has become the most widely used gene in *Cryptosporidium* spp. subtyping. It is currently not possible to amplify more distantly related species like *C. felis* or *C. canis* at the *gp60* locus but recently *gp60* primers have been developed for other species such as *C. ubiquitum*, *C. viatorum*, *C. muris* and *C. andersoni* (Feng et al., 2011b; Li et al., 2014; Stensvold et al., 2015).

It has been suggested that genetic variation at *gp60* locus might be essential for the parasite’s long-term success (Abal-Fabeiro et al., 2013), which renders this locus very suitable for analysing genetic and epidemiological associations of cryptosporidiosis. Thus, an advantage of using *gp60* for subtyping is the potential association between subtype families and phenotypes of *C. parvum* and *C. hominis*. This could be due to the biologic importance of the *gp60* protein, which is located on the surface of the apical region of invasive stages of the parasite, and is one of the dominant targets for neutralising antibody responses in humans (O’Connor et al., 2007). Some of the *C. parvum* subtype families, such as IIa and IId, are found in both humans and ruminants, responsible for zoonotic cryptosporidiosis (Xiao, 2010). There are also significant differences in clinical presentations and virulence among some common *C. hominis* or *C. parvum* subtype families in cryptosporidiosis-endemic areas (Cama
et al., 2007, 2008; Del Chierico et al. 2011; Feng et al., 2012). Some gp60 subtypes of C. hominis and C. parvum, such as IbA10G2 of C. hominis and IIaA15G2R1 of C. parvum, are widely distributed in the world probably due to their biologic fitness (Budu-Amoako et al. 2012; Feng et al. 2013; Li et al. 2013).

The whole genome sequencing (WGS) of C. parvum, C. hominis, and C. muris has allowed the identification of microsatellite and minisatellite sequences in Cryptosporidium genomes and other targets that are highly polymorphic between C. parvum and C. hominis (Li et al., 2013; Robinson and Chalmers, 2012). They are frequently used in either multilocus fragment typing (MLFT) or multilocus sequence typing (MLST) to increase the subtyping resolution offered by gp60 sequence analysis (Muthusamy et al. 2006; Ryan et al., 2014; Pérez-Cordón et al., 2016; Tang et al., 2016).

1.9 Cryptosporidium metabolism and metabolomics

Cryptosporidium genomes from multiple species have now been sequenced (e.g. Abrahamsen et al., 2004; Xu et al., 2004; Hadfield et al., 2015; Isaza et al., 2015; Ifeonu et al., 2016). Analysis of these genomes has facilitated a better understanding of the metabolic activities of this parasite and revealed many unexpected and unique biochemical activities. The genome of this parasite is highly compact (~9.2 million bases Mb) and encodes a streamlined metabolism with Cryptosporidium highly reliant on its host/environment for nutrients as it is missing key metabolic pathways and lacks the ability for de novo synthesis of nucleosides, fatty acids and amino acids (Abrahamsen et al., 2004; Xu et al., 2004). Thus, Cryptosporidium has evolved to maximise the number of biosynthetic molecules that are salvaged from the host and has retained the capacity for biosynthesis of only those metabolites that are unavailable, or require too much energy to transport (Abrahamsen et al.,
In contrast to other apicomplexans, Cryptosporidium lacks both mitochondrion and apicoplast genomes. Nuclear genes for DNA replication and translation are also absent (Abrahamsen et al., 2004; Isaza et al., 2015; Ifeonu et al., 2016). A number of putative mitochondrial proteins were identified, however, the parasite lacks a tricarboxylic acid (TCA) cycle and much of the oxidative phosphorylation pathway, indicating it does not rely on complete oxidation and respiratory chains for synthesising adenosine triphosphate (ATP). A complete glycolytic pathway is present, however, atypical enzymes are used to maximise ATP conversion (Abrahamsen et al., 2004).

Cryptosporidium is unable to synthesise most basic metabolites, including amino acids, fatty acids and nucleotides, and instead relies on scavenging from the host to meet its basic metabolic requirements via an extensive collection of amino acid and sugar transporters and salvage pathways. Genome analysis (Abrahamsen et al., 2004) confirmed previous hypotheses that this parasite relies solely on cytosolic glycolysis to fulfill its energy needs (Denton et al., 1996; Entrala and Mascaro, 1997). Cryptosporidium can utilise amylopectin during glycolysis (Thompson et al., 2005) and also possesses an alternative oxidase (AOX) pathway (Roberts et al., 2004; Suzuki et al., 2004). An in silico, genome-scale metabolic model of C. hominis identified 540 reactions performed by 213 enzymes (Vanee et al., 2010). Of these reactions, 514 were metabolic biochemical reactions involving intracellular metabolites and 26 were transport reactions representing the movement of metabolites across the cell membrane (Vanee et al., 2010).

### 1.9.1 Metabolomics

Metabolomics, the study of the metabolome, is the analysis of all endogenous and
exogenous metabolites. As these metabolites play important roles in anabolic and catabolic pathways, the metabolome will change in response to any physiological and pathological alteration, and hence may provide vital information about the health status of an organism. The field of metabolomics thus provides a novel approach to examine parasites and their interactions with their host through the analyses of the entire metabolite or small molecule (<1 kDa) composition of a biological sample. Perturbations in the profiles of these metabolites reflect changes to cellular regulation and physiological processes, as a result of these parasitic infections (Fiehn, 2001; Wang et al., 2004; Li et al., 2008a; Scheltema et al., 2010; Watson, 2010). Currently little is known about the *Cryptosporidium* metabolome, but metabolomics analysis provides an avenue for biomarker discovery, drug targets and improved diagnostics.

Metabolites can be profiled from tissues and a range of biofluids, with urine and blood being the most common samples (Bezabeh et al., 2009). Faecal metabolite profiles have also proven informative, enabling investigation of gastrointestinal diseases and the complex metabolic interactions between the host and its commensal microbial communities in the gastrointestinal tract (Saric et al., 2007; Li et al., 2008b; Jansson et al., 2009; Monleón et al., 2009; Han et al., 2010; Ponnusamy et al., 2011).

Faecal metabolite profiling is still in its infancy and, to date, the majority of studies have utilised nuclear magnetic resonance (NMR) spectroscopy (Saric et al., 2007; Li et al., 2008a; Bezabeh et al., 2009; Jansson et al., 2009; Monleón et al., 2009; Le Gall et al., 2011; Ponnusamy et al., 2011; Deda et al., 2016; Loftfield et al., 2016). Gas chromatography-mass spectrometry (GC-MS), a technique commonly used in metabolome analysis, is a sensitive and reproducible method that enables complex mixtures of metabolites to be resolved (Villas-Bôas et al., 2005). This makes GC-MS an ideal tool for analysis of faecal samples, which contain metabolites from the host, commensal microorganisms and pathogenic organisms.
infected the gastrointestinal tract.

An optimised extraction method is essential to ensure maximum yield of metabolites from a sample, while minimising manipulation and biochemical alteration of the metabolites extracted. Metabolite extractions from faecal samples for GC-MS analysis have been described in several studies (Gao et al., 2009; 2010; Ponnusamy et al., 2011; Deda et al., 2016; Loftfield et al., 2016). The extraction method from human faecal water described in Gao et al. (2009; 2010) allowed for qualitative and quantitative GC-MS analysis; however, a large amount of starting faecal material was required. Using a much smaller amount of faecal sample, Ponnusamy et al. (2011) provided a brief description of extracting metabolites from faecal samples of IBS patients without prior manipulation. Although results from GC-MS analysis showed differences between IBS and non-IBS patients, no data on method optimisation was presented and it did not address differences in faecal consistencies of patients, especially those with gastrointestinal infections (Ponnusamy et al., 2011).

1.10 Aims and objectives

The molecular epidemiology of Cryptosporidium in Australia is relatively poorly understood and very little is known about the metabolomics of this parasite. In order to advance the knowledge of this parasite, the specific aims of this thesis therefore were to:

1. Analyse the molecular epidemiology of Cryptosporidium in Western Australia
2. Review and analyse cryptosporidiosis outbreaks in Western Australia.
3. Analyse the transmission dynamics of Cryptosporidium in rural settings
4. Conduct metabolomics analysis of Cryptosporidium infected versus uninfected faecal samples to better characterise metabolic changes associated with Cryptosporidium infection.
CHAPTER 2.0

General Materials and Methods

2.1 Genomic DNA extraction from faecal samples

Total genomic DNA from faecal samples was extracted using either the QiAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany) or the PowerSoil® DNA Isolation Kit (MoBio, Carlsbad, California). The QiAamp DNA Stool Mini Kit was used in DNA extraction of faecal samples from humans whereas the MoBio Powersoil® DNA Isolation Kit was used primarily to extract total genomic DNA from livestock and wildlife faecal samples. The MoBio Powersoil® DNA Isolation Kit enables removal of environmental artefacts which are often present in faeces of livestock and wildlife and may inhibit downstream analysis by Polymerase Chain Reaction (PCR).

All faecal DNA was extraction was performed with minor modifications to the manufacturer’s protocol. Briefly, after faecal samples were weighed into the allocated microcentrifuge tubes, five cycles of freeze-thaw were carried out where following freezing in liquid nitrogen, the samples were thawed at 100°C. The samples were then incubated at 90°C to ensure the lysis of the thick-walled Cryptosporidium oocyst and the release of DNA material. The final elution volume was adjusted to 50 µL from the manufacturer’s recommended volume of 200 µL for the QiAamp DNA Stool Mini Kit and the 100 µL for the MoBio Powersoil® DNA Isolation Kit to increase DNA concentration.
2.2 Polymerase Chain Reaction

PCR was generally carried out in volumes of 25 µL per reaction where each reaction mixture consists of 1-2 µL of DNA template, 250 µM of each dNTP (Promega Biosciences, LLC., California, U.S.A.), 1x KAPA Taq buffer (KAPA Biosystems, Cape Town, South Africa), 0.5 units KAPA Taq (KAPA Biosystems, Cape Town, South Africa), 12.5 µM of each primer and Ultra Pure PCR grade water (Fisher Biotec, Western Australia, Australia). Concentration of MgCl$_2$ of the reaction is dependent on the primer set used (see Table 2.1).

Reactions were performed on the Applied Biosystems Gene Amp PCR System 2700 thermal cycler (LifeTechnologies, California, U.S.A.) where, each sample was subjected to an initial denaturation at 95°C for 4 mins. Thermal cycling conditions such as denaturation and annealing times as well as annealing temperature varied based on the primer sets used and specific conditions are listed in Table 2.1. All reactions were then subjected to a post PCR extension of 72°C for 7 mins.

2.3 Agarose gel electrophoresis

PCR products were visualised using agarose gel electrophoresis. Separation of the PCR products were carried out on a 1-2 % agarose gel in 1x TAE buffer (40 mM Tris-HCl, 20 mM acetate, 2 mM EDTA, pH adjusted to 7.9). Gels were stained prior to casting using 1x SYBR® Safe DNA Gel Stain (Invitrogen, Victoria, Australia) and visualised using the Dark Reader blue light transilluminator (Clare Chemical Research, Inc., Colorado, U.S.A.). PCR fragment sizes were estimated using 100bp ladder DNA marker (Axygen Biosciences, California, U.S.A.).
Table 2.1: Primers used in the characterisation of *Cryptosporidium* sp. at the 18S rRNA, actin and *gp60* gene loci.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequence (5’→3’)</th>
<th>MgCl₂ concentration (mM)</th>
<th>Annealing temperature (°C)</th>
<th>Approx. amplicon size (bp)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>18S rRNA</td>
<td>F1: TTC TAG AGC TAA TAC ATG CG</td>
<td>3.0 mM</td>
<td>55°C</td>
<td>1,325 bp</td>
<td>(Xiao et al., 1999, 2001)</td>
</tr>
<tr>
<td></td>
<td>R1: CCC TAA TCC TCC GAA ACA GGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Or - R1: CCC ATT TCC TCC GAA ACA GGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2: GGA AGG GTT GTA TTT ATT AGA TAA AG</td>
<td>3.0 mM</td>
<td>55°C</td>
<td>826-864 bp</td>
<td>(Ryan et al., 2003a)</td>
</tr>
<tr>
<td></td>
<td>R2: AAG GAG TAA GGA ACA ACC TCC A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F1: GAC ATA TCA TTC AAG TTT CTG ACC</td>
<td>1.5 mM</td>
<td>58°C</td>
<td>763 bp</td>
<td>(Yang et al., 2001)</td>
</tr>
<tr>
<td></td>
<td>R1: CTG AAG GAG TAA GGA ACA ACC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2: CCT ATC AGC TTT AGA CGG TAG</td>
<td>1.5 mM</td>
<td>58°C</td>
<td>587 bp</td>
<td>(Ng et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>R2: TCT AAG AAT TTC ACC TCT GAC TG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actin</td>
<td>F1: ATG AGG ATG AAG AAG ATA AGC TAT CAA GC</td>
<td>2.5 mM</td>
<td>55°C</td>
<td>1,095 bp</td>
<td>(Sulaiman et al., 2002)</td>
</tr>
<tr>
<td></td>
<td>R1: AGA AGA CAC TTT TCT GTG TGA CAA T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2: CAA GCA TTT GAG TTG TTG ATC AA</td>
<td>2.5 mM</td>
<td>55°C</td>
<td>1,066 bp</td>
<td>(Ng et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>R2: TTT CTG TGT GAC AAT ATG CAT TGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F1: ATG CCA GGA RTW ATG GTD GGT ATG</td>
<td>2.5 mM</td>
<td>55°C</td>
<td>830 bp</td>
<td>(Ng et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>R1: GGD GCA ACR ACY TTR ATC TTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2: GAY GAR GCH CAR TCV AAR AGR GGT AT</td>
<td>2.5 mM</td>
<td>55°C</td>
<td>818 bp</td>
<td>(Ng et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>R2: TTD ATY TTC ATD GTH GAH GGC GC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60 kDa-glycoprotein</td>
<td>F1: GCT TAG ATT CTG ATA TGG ATC TAG</td>
<td>4.0 mM</td>
<td>54°C</td>
<td></td>
<td>(Strong et al., 2000, Sulaiman et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>R1: AGC TTA CTG GTG CTG TAT CAG TT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2: ACC CCA GAA GGC GGA CCA AGG TT</td>
<td>3.0 mM</td>
<td>55°C</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R2: GTA TCG TGG CGT TCT GAA TTA TCA A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.4 Gel purification of PCR products

PCR amplified product fragments were separated by gel electrophoresis on a 1.5% agarose gel and viewed under a Dark Reader blue light transilluminator platform. The appropriate sized band was excised using a scalpel blade, with excision of each band carried out using a new scalpel blade to avoid contamination of amplicons. The bands were each placed in sterile 1.5 ml centrifuge tubes and purified from the gel by either the freeze-squeeze method as described in (Ng et al., 2006) or by using the UltraClean®15 DNA Purification Kit (MoBio, Carlsbad, California) in accordance to the manufacturer’s protocol.

2.5 Sequencing

Purified PCR products were sequenced using the Big Dye version 3.1 Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, California) with each sequencing reaction containing 1 μl dye terminator mix, 1.5 μl of 5x sequencing buffer (Applied Biosystems), 3.25 pmol/μl of the primer used to generate the PCR product, and 30-50 ng of purified PCR product. The reaction was made up to a final volume of 10 μl with PCR grade water (Fisher Biotech). Sequencing reactions were carried out on an Applied Biosystems Gene Amp PCR System 2700 thermal cycler (LifeTechnologies, California, U.S.A.) using the following thermal cycling conditions: 96°C for 2 min, followed by 30 cycles of denaturation at 96°C for 10 sec, annealing at the temperature optimal for the primer used for 5 sec and extend at 60°C for 4 mins. Sequencing reactions were then ethanol precipitated (see 2.5.1 Ethanol precipitation of sequencing reaction) and analysed on an Applied Biosystems 3730 DNA Analyzer (Applied Biosystems Inc.).
2.5.1 Ethanol precipitation of sequencing reaction

Ethanol precipitation of sequencing reaction products were carried out to remove salts and remaining dye terminators. For each sequencing reaction, the full 10 μl sequencing reaction was transferred to the 0.6 ml Eppendorf tube where, as in the sequence listed, 1 μl of 125 mM EDTA, 1 μl of 3M sodium acetate and 25 μl of 100% ethanol was added to the reaction and mixed by pipetting and left at room temperature for 20 min. The solution was then centrifuged at 20,000 x g for 30 min and the supernatant discarded. The resulting pellet was then rinsed with 125 μl of 70% ethanol and centrifuged at 20,000 x g for 5 min. The supernatant was then discarded and the pellet vacuum dried using a speed vac concentrator.

2.6 Analysis of sequence chromatogram

Nucleotide sequences were analysed using Chromas v2.3 (Technelysium) and compared to sequence data available from GenBank™ using the BLAST (Basic Local alignment Search Tool) program (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.7 Phylogenetic analysis

Additional Cryptosporidium sequences as well as sequences for outgroups were downloaded from GenBank nucleotide database (http://www.ncbi.nlm.nih.gov/genbank/). Sequences were aligned with Clustal W (Larkin et al., 2007). Phylogenetic analyses were performed using TREECON (Van de Peer and De Wachter, 1997) with tree topologies inferred by neighbour-joining method based on distances calculated using Kimura-2
parameters. Statistical confidence of the inferred tree was assessed by conducting a bootstrap
test using 1,000 replicates. Distance matrices based on distance calculated with the Kimura-2
model were generated using MEGA v.5 (Tamura et al., 2011).
CHAPTER 3

Differences in cryptosporidiosis among Aboriginal and non-Aboriginal humans in Western Australia

3.1 Introduction

Human cryptosporidiosis is a notifiable disease in several, mainly European, countries (ECDC, 2014). In Australia, the National Notifiable Diseases Surveillance System (NNDSS) was established in 1990, under the auspices of the Communicable Diseases Network Australia and in 2001, cryptosporidiosis was listed as a national notifiable disease (Blumer et al., 2003). At the time of study, only confirmed cases (based on microscopy or PCR) are reported to the NNDSS. As mentioned in section 1.8, Chapter 1, differentiation between species and genotypes of Cryptosporidium is not possible by microscopy due to lack of distinguishing oocyst features and requires the use of molecular tools. However, molecular identification of Cryptosporidium species is not routinely carried out, and therefore the Cryptosporidium species responsible for the infections reported to the NNDSS are generally unknown.

To date, in humans in Australia, C. hominis, C. parvum, C. meleagridis, C. fayeri, C. andersoni, C. bovis, C. cuniculus, a novel Cryptosporidium species most closely related to C. wrairi and the Cryptosporidium mink genotype have been reported (see Case study Chapter 3 Section 3.3.3) (Morgan et al. 1997; Morgan et al., 1998; Gasser et al., 2004; Chalmers et al., 2005; Jex et al., 2007; Ng et al., 2008; Jex et al., 2008; Alagappan et al., 2008; Pangasa et al., 2009; Waldron et al., 2009; Ng et al., 2010a; 2010b; Waldron et al., 2010; Power et al., 2011; Waldron et al., 2011a; Ng et al., 2012; Jex et al., 2012; Ng-Hublin et al., 2013; Sari et al., 2013 unpublished - KF279538; Yang et al., 2013; Koehler et al., 2013; 2014a; 2014b; Ebner et al., 2015).
Very little is known about cryptosporidiosis in Aboriginal populations although previous studies, reported a high rate of cryptosporidiosis in Aboriginal children (Assadamongkol et al., 1992; Lal et al., 2015b). The present study, therefore, aimed to provide insight into the prevalence, disease patterns of cryptosporidiosis and species and subtypes of Cryptosporidium in Aboriginal versus non-Aboriginal populations in Western Australia (WA), from the first full year of reporting in 2002 through to 2012.

3.2 Materials and Methods

3.2.1 Case definition

The Australian case definition for cryptosporidiosis notification, which is also adopted by WA, requires laboratory definitive evidence through the detection of Cryptosporidium oocysts (NNDSS, 2016).

3.2.2 Data source and extraction

Cryptosporidiosis notification data used in the present study were obtained from the Western Australian Notifiable Infectious Disease Database (WANIDD) by the data custodian, the Communicable Disease Control Directorate (CDCD) at the WA Department of Health. The WA surveillance system relies on mandatory reporting of notifiable diseases by laboratories and medical practitioners. WANIDD, is an intranet-based real-time application with notifications from laboratories uploaded electronically and notifications from medical practitioners added manually.

WANIDD data from January 2002 to December 2012 were extracted and included in the present study. The data was saved in Microsoft Excel and de-identified through removal
of case name, address and date of birth, under Murdoch Human Ethics approval number 2012/208 and the WA Department of Health Ethics approval number 2009/48. The ODOO is defined as the ‘true’ date of onset provided by the notifying doctor or obtained during case follow-up, the date of specimen collection for laboratory notified cases, and when neither of these dates is available, the date of notification by the doctor or laboratory, or the date of receipt of notification, whichever is earliest.

Variables exported from WANIDD include public health unit (PHU), sex, country of birth, race (Aboriginality), age, travel history, location where infection was acquired and hospitalisation. Additional information such as clinical symptoms, duration of diarrhea and the infecting *Cryptosporidium* species based on molecular typing data (when available), were included in variables exported.

### 3.2.3 Molecular typing

Molecular typing was conducted on a total of 324 *Cryptosporidium* positive human faecal samples provided by WA clinical pathology laboratories. These samples were matched to the basic demographic data of cases from WANIDD. Molecular genotyping was carried out on isolates with ODOO from 2005 to 2008 and have been detailed in two previous studies Ng et al. (2010a) (n = 198) and Ng et al. (2010b) (n = 42). Molecular genotyping for the remaining 84 samples were carried out from June 2010 to December 2012 as part of the present study.

Extraction of DNA from stools and molecular analysis was carried out as described in Chapter 2.0 (*see Material and Methods*). Initial genotyping of the samples was carried out by PCR-RFLP of an 830 bp fragment of the *Cryptosporidium* 18S gene locus as described in Chapter 2, using restriction analysis of the PCR product by the VspI (Promega, USA) to
discriminate between *C. hominis* and *C. parvum* as described by Xiao et al. (2001). For samples that failed to amplify or produced ambiguous banding patterns, a two-step nested PCR and sequencing of a 540 bp product the 18S gene locus was carried out (as described in Chapter 2). *Cryptosporidium hominis* and *C. parvum* positive samples were sub-typed at the *gp60* gene locus using a two-step nested PCR that amplifies an 830 bp fragment (as described in Chapter 2). Secondary PCR products were purified and sequenced as described in Chapter 2 (see Section 2.4).

**Figure 3.1**: Map of public health units (PHU’s) in Western Australia – urban Perth is divided into North and South Metropolitan regions.

![Map of public health units in Western Australia](image)

### 3.2.4 Data analysis

Statistical analyses were conducted using Microsoft Excel 2010 and OpenEpi version 3.01 ([http://openepi.com/v37/Menu/OE_Menu.htm](http://openepi.com/v37/Menu/OE_Menu.htm)) and STATA 13. Notification rates were
calculated using annual census population data available through the Rates Calculator software (Department of Health, WA) and age standardised to the Australian population. Notification Rate Ratios were calculated by dividing the notification rate in Aboriginal people by the notification rate in non-Aboriginal people. WA population estimates for public health unit (PHU) areas, year, sex, age and Aboriginality status were used to calculate notification rates and described as number of notified cases per 100,000 populations/year. Public Health Units in Western Australia are divided into different health administrative regions – Metropolitan (North and South), Kimberley, Pilbara and Goldfields region (classified as remote) and the Midwest, Wheatbelt, South West and Great Southern (classified as rural) (Figure 3.2).

3.2 Results

A total of 2,988 cryptosporidiosis cases, with an average notification rate of 12.7 cases/100,000 population, were reported in WA during the 11-year period from January 2002 to December 2012. Individuals that were not residents of WA accounted for 17 of the cases notified during this period, of which three cases were Aboriginal people, nine were non-Aboriginal and the Aboriginality status for five of the remaining cases was not recorded.

3.3.1 Cryptosporidium notification rates in Aboriginal and non-Aboriginal people

Of the 2,988 cryptosporidiosis notifications, 24.6% (734/2,988) were Aboriginal people and 53.1% (1,585/2,988) were non-Aboriginal people and Aboriginal status was not recorded for 22.4% of cases. Notification rates in Aboriginal and non-Aboriginal people over the 11-year period ranged from 42.9-135.8 and 2.4-20.4 per 100,000 population per year, respectively (Table 3.1). The highest difference in notification rate between Aboriginal
people and non-Aboriginal people occurred in 2004, with a notification rate ratio of 53.1, and
the lowest in 2007, with a notification rate ratio of 5.6 (Table 3.1).

**Table 3.1:** The annual number, notification rate, NR (notifications per 100,000/year) and
notification rate ratio (NRR) of cryptosporidiosis notification in Aboriginal and non-
Aboriginal people in Western Australia (2002-2012).

<table>
<thead>
<tr>
<th>Year</th>
<th>Aboriginal people</th>
<th>Non-Aboriginal people</th>
<th>NRR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>NR</td>
<td>Numbers</td>
</tr>
<tr>
<td>2002</td>
<td>62</td>
<td>92.8</td>
<td>45</td>
</tr>
<tr>
<td>2003</td>
<td>92</td>
<td>135.8</td>
<td>130</td>
</tr>
<tr>
<td>2004</td>
<td>51</td>
<td>74.3</td>
<td>27</td>
</tr>
<tr>
<td>2005</td>
<td>93</td>
<td>133.6</td>
<td>67</td>
</tr>
<tr>
<td>2006</td>
<td>61</td>
<td>86.1</td>
<td>148</td>
</tr>
<tr>
<td>2007</td>
<td>82</td>
<td>114.2</td>
<td>414</td>
</tr>
<tr>
<td>2008</td>
<td>60</td>
<td>82.3</td>
<td>83</td>
</tr>
<tr>
<td>2009</td>
<td>66</td>
<td>89.4</td>
<td>113</td>
</tr>
<tr>
<td>2010</td>
<td>37</td>
<td>49.3</td>
<td>89</td>
</tr>
<tr>
<td>2011</td>
<td>97</td>
<td>127.8</td>
<td>343</td>
</tr>
<tr>
<td>2012</td>
<td>33</td>
<td>42.9</td>
<td>126</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>734</strong></td>
<td><strong>1,585</strong></td>
<td></td>
</tr>
</tbody>
</table>

A multimodal distribution of notifications was observed for both non-Aboriginal and
Aboriginal people age groups. For non-Aboriginal people, peaks were observed in the 0-4 years age group (33.3% of cases, notification rate = 36.5 notifications per 100,000/year), followed by a second peak in the 30-34 year age group (9% of cases, notification rate = 8.7 notifications per 100,000/year) (cluster age range = 20-39 years) and a third peak in the 70-74 years age group (1.6% of cases, notification rate = 3.7 notifications per 100,000/year) (cluster age range = 65-79 years). In Aboriginal people, peaks in the distribution were observed in the 0-4 years age group (93.2% of cases, notification rate = 749.6 cases per 100,000/year) with a second peak in the 20-24 years age group (0.7% of cases, notification rate = 7.0 notification
per 100,000/year) (cluster age range = 20-34 years) and a third peak in the 55-59 years age group (0.5% of cases, notification rate = 17.9 cases per 100,000/year) (cluster age range = 45-64 years) (Figure 3.2; Table 3.2). The notification rate ratio in the 0-4 years age group in Aboriginal and non-Aboriginal people was 20.5 (Table 3.2). Within the 0-4 years age group, Aboriginal children ≤ 1 years of age had the highest average rate at 1,549 cases/100,000 population, followed by 2 year olds with a rate of 564 cases/100,000 population. Similarly, non-Aboriginal children in the 0-4 years age group with the highest number of notification rates were children ≤ 1 years of age and at 2 years of age with average rates of 41 cases/100,000 population.

**Figure 3.2:** Comparison of cryptosporidiosis notification rates (cases/100,000 population) between Aboriginal and non-Aboriginal people, stratified by age groups.
Aboriginal people had higher notification rate ratios compared to non-Aboriginal people in the 45-49 (notification rate ratio = 4.8), 50-54 (notification rate ratio = 3.7), 55-59 (notification rate ratio = 5.8) and 60-64 (notification rate ratio = 4.9) age groups (Table 3.2). Male Aboriginal people had a higher notifications rate (95.3 per 100,000 population, notification rate ratio = 14.7) compared to male non-Aboriginal people (7.5 per 100,000 population). Similarly, notification rates among female Aboriginal people were higher (90.5 per 100,000 population, notification rate ratio = 12.1) compared to female non-Aboriginal people (7.5 per 100,000 population).

Table 3.2: The age distribution, notification rate, NR (notifications per 100,000/year) and notification rate ratio (NRR) of cryptosporidiosis notification in Aboriginal and non-Aboriginal people in Western Australia (2002-2012).

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Aboriginal (Aborg)</th>
<th>Non-Aboriginal (Non-Aborg)</th>
<th>NRR</th>
<th>Aborg:Non-Aborg</th>
<th>Non-Aborg: Aborg</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>749.6</td>
<td>36.5</td>
<td>20.5</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>5-9</td>
<td>15.5</td>
<td>12.2</td>
<td>1.3</td>
<td>0.8</td>
<td>-</td>
</tr>
<tr>
<td>10-14</td>
<td>3.2</td>
<td>5</td>
<td>0.6</td>
<td>1.6</td>
<td>-</td>
</tr>
<tr>
<td>15-19</td>
<td>3.6</td>
<td>3.1</td>
<td>1.2</td>
<td>0.9</td>
<td>-</td>
</tr>
<tr>
<td>20-24</td>
<td>7</td>
<td>6.3</td>
<td>1.1</td>
<td>0.9</td>
<td>-</td>
</tr>
<tr>
<td>25-29</td>
<td>6.4</td>
<td>6.3</td>
<td>1.0</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>30-34</td>
<td>5.2</td>
<td>8.7</td>
<td>0.6</td>
<td>1.7</td>
<td>-</td>
</tr>
<tr>
<td>35-39</td>
<td>3.7</td>
<td>6.8</td>
<td>0.5</td>
<td>1.8</td>
<td>-</td>
</tr>
<tr>
<td>40-44</td>
<td>4.3</td>
<td>3.8</td>
<td>1.1</td>
<td>0.9</td>
<td>-</td>
</tr>
<tr>
<td>45-49</td>
<td>10.1</td>
<td>2.1</td>
<td>4.8</td>
<td>0.2</td>
<td>-</td>
</tr>
<tr>
<td>50-54</td>
<td>9.6</td>
<td>2.6</td>
<td>3.7</td>
<td>0.3</td>
<td>-</td>
</tr>
<tr>
<td>55-59</td>
<td>17.9</td>
<td>3.1</td>
<td>5.8</td>
<td>0.2</td>
<td>-</td>
</tr>
<tr>
<td>60-64</td>
<td>13.3</td>
<td>2.7</td>
<td>4.9</td>
<td>0.2</td>
<td>-</td>
</tr>
<tr>
<td>65-69</td>
<td>0</td>
<td>3.2</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>70-74</td>
<td>0</td>
<td>3.7</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>75-79</td>
<td>0</td>
<td>2.6</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>80-85+</td>
<td>0</td>
<td>2.3</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Rates of cryptosporidiosis notifications were higher among Aboriginal people compared to non-Aboriginal people across all PHUs, ranging from 11.9-246.6 per 100,000 population (Table 3.3). The remote regions (Goldfields, Kimberley and Pilbara), had the highest average number of cases notified among the Aboriginal people, particularly in the Kimberley, with an average of 38.5 cases notified each year between 2002-2012, at a rate of 246.6 per 100,000 per year (Table 3.3). Among non-Aboriginal people, the highest number of notifications was reported from the Metropolitan region (91.5 per year; notification rate = 5.6 per 100,000 per year), while the highest notification rate was in the Kimberley (notification rate = 51.2 per 100,000 per year (Table 3.3).

Table 3.3: Average number of notifications and notification rates (NR) of cryptosporidiosis cases in each PHU from 2002-2012 by Aboriginal status.

<table>
<thead>
<tr>
<th>PHUs</th>
<th>Aboriginal people</th>
<th>Non-Aboriginal people</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average no. of</td>
<td>NR (per 100,000)</td>
</tr>
<tr>
<td></td>
<td>notifications</td>
<td></td>
</tr>
<tr>
<td>Wheatbelt</td>
<td>1.6 (0-6)</td>
<td>44.9</td>
</tr>
<tr>
<td>Goldfields</td>
<td>5.6 (2-17)</td>
<td>96.8</td>
</tr>
<tr>
<td>Great Southern</td>
<td>0.36 (0-1)</td>
<td>17.1</td>
</tr>
<tr>
<td>Kimberley</td>
<td>38.5 (12-59)</td>
<td>246.6</td>
</tr>
<tr>
<td>Metropolitan</td>
<td>3.2 (0-9)</td>
<td>11.9</td>
</tr>
<tr>
<td>Midwest</td>
<td>4.9 (0-14)</td>
<td>67.9</td>
</tr>
<tr>
<td>Pilbara</td>
<td>11.3 (4-16)</td>
<td>151.8</td>
</tr>
<tr>
<td>South West</td>
<td>1.0 (0-3)</td>
<td>31.7</td>
</tr>
</tbody>
</table>
Hospitalisation data was available for 1,268/2,320 (55%) of cases with recorded Aboriginal status, comprising 251 Aboriginal people and 1,017 non-Aboriginal people. A total of 244 cases of cryptosporidiosis were hospitalised between 2002 and 2012. A significantly (P<0.01) higher proportion of Aboriginal people were hospitalised (53% 133/251) compared to non-Aboriginal people (10.8% 110/1,017). Children aged 0-4 years comprised the majority of both Aboriginal and non-Aboriginal cases hospitalised, at 89.5% (119/133) and 34.2% (38/111) respectively.

3.3.2 Molecular Epidemiology

Data for molecular analyses carried out was available for the specimens tested from April 2005 to January 2008 and from June 2010 to December 2012 (n=324 cases), with Aboriginal status recorded in only 290 of these cases. In total, *C. hominis* was identified in 259/324 (79.9.0%) cases, *C. parvum* in 54 (16.6%) and *C. meleagridis* in 10 (3.1%) cases, with one case (0.3%) identified with a mixed infection of *C. meleagridis*, *Cryptosporidium* Mink genotype and an unknown *Cryptosporidium* genotype (Table 3.4). The latter two genotypes have been analysed in more depth in a case study in section 3.3.3 of this chapter.

Of the cases identified with *C. hominis* (n= 259), 58.7% (152/259) were non-Aboriginal people, 32% (83/259) were Aboriginal people and the status of 9.3% were unknown (Table 3.4). However, Aboriginal people were 1.3 times more at risk (RR 1.3; 95% CI, 1.21-1.44, p < 0.05) than non-Aboriginal people of being infected with *C. hominis*. Only a small proportion of Aboriginal people (2.4%; 2/85) were diagnosed with *C. parvum*, with 97.6% (83/85) of cases diagnosed with *C. hominis* (Table 3.4). Non-Aboriginal people were 8.7 times more likely to be infected with *C. parvum* compared to Aboriginal people (RR 8.7; 95% CI, 2.16-35.16, p < 0.05). All cases infected with *C. meleagridis* were non-Aboriginal
people. Thus, zoonotic Cryptosporidium (C. parvum, C. meleagris and the mink genotype) were identified in 25.8% (53/205) of non-Aboriginal people.

Table 3.4: Cryptosporidium species and subtypes identified in Aboriginals and non-Aboriginals in WA.

<table>
<thead>
<tr>
<th></th>
<th>Aboriginal</th>
<th>Non-Aboriginal</th>
<th>Unknown</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C. hominis</strong></td>
<td>83</td>
<td>152</td>
<td>24</td>
<td>259</td>
</tr>
<tr>
<td>IbA10G2</td>
<td>7</td>
<td>94</td>
<td>14</td>
<td>112</td>
</tr>
<tr>
<td>IdA15G1</td>
<td>49</td>
<td>23</td>
<td>4</td>
<td>75</td>
</tr>
<tr>
<td>IdA16</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>IdA17</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>IeA11G3T3</td>
<td>1</td>
<td>3</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>IfA12G1</td>
<td>1</td>
<td>3</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>IgA17</td>
<td>25</td>
<td>25</td>
<td>3</td>
<td>53</td>
</tr>
<tr>
<td><strong>C. parvum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIaA15G2R1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>IIaA17G2R1</td>
<td>-</td>
<td>9</td>
<td>-</td>
<td>10</td>
</tr>
<tr>
<td>IIaA18G3R1</td>
<td>1</td>
<td>28</td>
<td>9</td>
<td>37</td>
</tr>
<tr>
<td>IIaA19G4R1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>IIaA21G2R1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>IIdA15G1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td><strong>ND</strong></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C. meleagris</strong></td>
<td></td>
<td>11^</td>
<td>-</td>
<td>10</td>
</tr>
<tr>
<td>Mink genotype</td>
<td>-</td>
<td>1^</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>85</td>
<td>205</td>
<td>34</td>
<td>324</td>
</tr>
</tbody>
</table>

*ND – subtype not determined
^Mixed infection – C. meleagris and Cryptosporidium Mink genotype was identified from one case.

The most common C. hominis subtype identified in Aboriginal people was the IdA15G1 subtype (59%; 49/83), followed by IgA17 (30.1%; 25/83) and IbA10G2 (8.4%; 7/83). In non-Aboriginal people, C. hominis IbA10G2 was the most common (61.8%; 94/152), followed by IgA17 (16.4%; 25/152) and IdA15G1 (15.1%; 23/152) (Table 3.4).
Aboriginal people were 3.9 times more likely to be infected with the *C. hominis* IdA15G1 subtype (RR 3.90; 95% CI, 2.57-5.92, p < 0.05) and 1.8 times more likely to be infected with IgA17 (RR 1.8; 95%CI, 1.13-2.98, p < 0.05) compared to non-Aboriginal people. Compared to Aboriginal people, non-Aboriginal people were 7.2 times more likely to be infected with the IbA10G2 subtype (RR 7.24; 95%CI, 3.52-14.87, p < 0.05). The two most common *C. parvum* subtypes identified in non-Aboriginal people were the IIA17G2R1 subtype (21%; 9/42) and IIA18G3R1 (66.6%; 28/42).

The majority of typed cases from Aboriginal people came from remote areas, particularly the Kimberley. Of the 113 typed cases from remote areas, 70.8% (80/113) were from the Kimberley, and Aboriginal people consisted of 67.5% (54/80) of these. Very few typed Aboriginal cases came from metropolitan Perth (2% - 2/101) compared to Non-Aboriginal cases (98% - 99/101). Similarly, 87.1% (954/62) of non-Aboriginal cases came from rural areas compared to Aboriginal people (12.9% - 8/62).

Distribution of gp60 subtypes indicated that in the remote regions of WA, the majority (60.8% - 45/74) of infections in Aboriginal people were attributed to the IdA15G1 subtype, followed by the IgA17 subtype (29.7% - 22/74), whereas in non-Aboriginal people, infection was attributed to the IbA10G2 subtype (46.2% - 18/39) and the IdA15G1 subtype (30.8% - 12/39) (Table 3.5). In the rural regions of WA, 37.5% (3/8) of Aboriginal people were infected with the IdA15G1 subtype, followed by the IbA10G2 and IgA17 subtypes in 25% (2/8) of cases each. Non-Aboriginal people were mainly infected with the IbA10G2 subtype, (22.2% - 12/54), followed by the IgA17 and IdA15G1 subtypes in 16.7% (9/54) and 14.8% (8/54) of cases respectively (Table 3.5). In the Metropolitan area, the IbA10G2 subtype was identified in 63.3% (62/99) of non-Aboriginal people, followed by the *C. parvum* IIA18G3R1 subtype (13.3% - 13/99) and IgA17 (11.1% - 11/99) (Table 3.5). Only two Aboriginal people were identified from the Metropolitan area and the IbA10G2 and
IgA17 subtypes were identified in a case each (Table 3.5).

Table 3.5: Distribution of *C. hominis* and *C. parvum gp60* genotypes in remote, rural and metropolitan regions of WA by Aboriginality

<table>
<thead>
<tr>
<th></th>
<th>REMOTE</th>
<th>RURAL</th>
<th>METROPOLITAN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aboriginal</td>
<td>Non-Aboriginal</td>
<td>Aboriginal</td>
</tr>
<tr>
<td><em>C. hominis</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IbA10G2</td>
<td>4</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>IdA15G1</td>
<td>45</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>IdA16</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IdA17</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IeA11G3T3</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>IfA12G1</td>
<td>1</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>IgA17</td>
<td>22</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td><em>C. parvum</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IlaA15G2R1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IlaA17G2R1</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>IlaA18G3R1</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>IlaA19G4R1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IlaA21G2R1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IldA15G1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Unknown</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>74</strong></td>
<td><strong>39</strong></td>
<td><strong>8</strong></td>
</tr>
</tbody>
</table>

Molecular genotyping data was available for 168 hospitalised cases, for which Aboriginal status was known. Aboriginal people comprised 27/168 of cases, of which 63.0% (17/27) were hospitalised, whilst non-Aboriginal people comprised 141/168 of cases, of which 19.9% (28/141) were hospitalised. All 17 hospitalised Aboriginal people were infected with *C. hominis* and 16 of these were children (aged 0-4 years). The IdA15G1 subtype was identified in 10/17 cases, the IbA10G2 and IgA17 subtypes identified in three cases each and the IeA11G3T3 subtype identified in one case. *Cryptosporidium hominis* was detected in 15/28 hospitalised non-Aboriginal people, followed by *C. parvum* in 11/28 cases and *C.
meleagridis in two cases, one of which was a mixed infection with Mink genotype. The C. hominis IbA10G2 subtype was the most common (11/15), followed by the IdA15G1 subtype in two cases and the IfA12G1 subtype in one case. For cases infected with C. parvum, the IIaA18G3R1 subtype was the most common subtype, identified in eight cases, followed by the IIaA17G2R1 subtype in two cases and the IIaA15G2R1 subtype in one case.

3.3.3 Case study: Human cryptosporidiosis diagnosed in Western Australia – a mixed infection with Cryptosporidium meleagridis, Cryptosporidium mink genotype and an unknown Cryptosporidium species.

As part of the Cryptosporidium case control study conducted by the Western Australia Department of Health, a 24 year-old male resident tested positive for Cryptosporidium and was enrolled in the study. Screening for the presence of other gastroenteric pathogens carried out showed that the patient was negative for the presence of Salmonella spp., Campylobacter spp., Shigella spp., Vibrio spp., verotoxigenic E. coli, Yersinia spp., Plesiomonas spp. and rotavirus. The patient was interviewed using a 20 minute-long questionnaire on his illness and known risk factors for Cryptosporidium. The case definition of diarrhoea was three or more loose stools in any 24-hour period, and at the time of the interview, the patient had experienced intermittent diarrhoea for 12 days that was ongoing and had contained blood. The most number of stools the patient had in a 24 hour period was 6-10. Other symptoms included fever, chills, vomiting, stomach cramps, nausea, headaches and muscle body aches. The patient reported no history of heart disease, high blood pressure, asthma, cancer, arthritis or any other chronic illness that may weaken the immune system and was not on any medications (i.e. antibiotics or antacids), suggesting that the patient was immunocompetent. The patient also reported being HIV negative, but it was unknown if a HIV test was actually
The faecal sample (C011) was submitted to Murdoch University for molecular analysis. Total faecal DNA was extracted using a QIAamp stool DNA extraction kit (Qiagen, Germany). PCR amplification and bi-directional sequencing of the *Cryptosporidium* 18S rRNA, actin and 60 kDa-glycoprotein (*gp60*) gene loci were carried out as previously described (see Chapter 2). Nucleotide sequence analysis was carried out using ChromasPro version v2.3 (http://www.technelysium.com.au) and aligned with reference sequences from GenBank using Clustal W (http://www.genome.jp/tools/clustalw). Phylogenetic analysis, based on evolutionary distances calculated using Kimura-2 parameters and grouped using neighbour joining, was conducted using MEGA version 5.05 (Tamura et al., 2011). Sequences generated from the present study have been submitted to GenBank under accession nos. JX471002-JX471005.

Sequence and phylogenetic analysis of the 18S rRNA gene locus, showed 100% similarity to the *Cryptosporidium* mink genotype (GenBank accession no. EF428191) (data not shown). However, sequence and phylogenetic analysis at the actin locus identified a mixed infection; *C. meleagridis* and an unknown *Cryptosporidium* species, which formed a group with *C. parvum*, *C. erinacei*, *C. hominis*, *C. cuniculus* and *C. tyzzeri* with genetic similarities of 97.7%, 98.1%, 97.7%, 97.7% and 98.5%, respectively (Figure 3.3). Sequence and phylogenetic analysis at the *gp60* gene locus identified a single novel sequence that exhibited 92.7% similarity to the *C. parvum* IId subtype (Figure 3.4).

In the twelve days prior to onset of diarrhoea, the patient travelled to Cairns and Brisbane in Queensland and then overseas to Papua New Guinea (PNG). It is understood the patient travelled alone and no secondary cases were reported. In PNG, he trekked in a remote highland region where he drank the river water. During the PNG trip he also stayed on a rural property where he drank unboiled water, which was sourced from the local town’s main
water supply. At the time he was not aware of anyone else with diarrhoea living in the same residence but did share a toilet with 25 other people. He did not report any contact with domestic, farm or wild animals. In his exposure period he ate raw fruits, fruit juices and raw vegetables and at least some of these were home grown. He also swam in public and commercially operated swimming pools.

In the present case study, it is likely that this patient became infected with Cryptosporidium during his travels to PNG through ingestion of contaminated water; however, this could not be confirmed. As three species of Cryptosporidium were identified in this patient, it is difficult to determine if the clinical symptoms were due solely to infection with C. meleagridis, or if the Cryptosporidium mink genotype, the unknown Cryptosporidium species, or other unidentified pathogens were also contributing to the spectrum of symptoms reported. This case is the first report of the Cryptosporidium mink genotype in a human as previously, it had only been identified in mink (Mustela vison) (Wang et al., 2008). More recently, another Australian human have been identified with the Cryptosporidium Mink genotype in Northern Australia (Ebner et al., 2015). The unknown Cryptosporidium species identified at the actin gene locus in the present case study, was genetically closest to C. tyzzeri (previously Cryptosporidium mouse genotype I), which mainly infects domestic mice and small rodents (Ren et al., 2012). Further studies are required to elucidate the identity of this unknown Cryptosporidium species and to determine if it is capable of causing disease in humans. As the gp60 sequence obtained did not match with either the mink genotype or C. meleagridis, it is likely that it corresponds to the unknown genotype identified at the actin locus, however this remains to be confirmed.
**Figure 3.3:** Phylogenetic relationships of C011 with other Cryptosporidium spp. at the actin gene locus inferred by neighbour-joining analysis based on genetic distances calculated using Kimura-2 parameters. Bootstrap values from 1000 pseudoreplicates of >60% are shown.
**Figure 3.4:** Phylogenetic relationships of C011 and other *Cryptosporidium* spp. *gp60* subtype families inferred by neighbour-joining (NJ) analysis based on genetic distances calculated using Kimura-2 parameters. *C. fayeri* *gp60* sequence was used as an outgroup to generate the NJ tree. Bootstrap values from 1000 pseudoreplicates of >50% are shown.
3.4 Discussion

Rates of reported cryptosporidiosis in Australia are higher than many comparable developed countries (Lal et al., 2015a), with an average rate of 12.5 cases/100,000 population/year across Australia between 2001-2016 (NNDSS, 2016). However, the actual number of cryptosporidiosis cases in Australia and worldwide is likely to be much higher as routine surveillance systems detect only a small fraction of the pathogen infections that occur in the community because: (1) it is estimated that less than 10% of individuals with gastroenteritis visit their local doctor, and of these, less than 10% have a faecal specimen collected (2) not all individuals presenting with gastroenteritis will have faecal samples tested for microorganisms and (3) identification of Cryptosporidium and other pathogens via microscopy lacks specificity and sensitivity (cf. Ryan et al., 2017). For example, a national survey of gastroenteritis in Australia in 2002, suggested a ratio of about 500 community cases to one notified (Hall et al., 2006). A more recent study reported that in 2010 alone, cryptosporidiosis was estimated to be responsible for 195,495 cases of acute gastroenteritis (AGE) and 333 disability-adjusted life years (DALYs) in Australia (Gibney et al., 2014).

Under-reporting of cryptosporidiosis is even more pronounced in Aboriginal communities due to the difficulty they have accessing health services. For example, in 2008, the National Aboriginal and Torres Strait Islander Social Survey (NATSISS) found that about 26% of Indigenous people aged 15 and over living in non-remote areas, had difficulty accessing health services (AIHW, 2011); in contrast, only 2.6% of the general population had difficulty (ABS, 2012). The main reasons cited were long waiting times, services not being available when needed, difficulties with transport and health-care costs. Less commonly reported reasons included lack of engagement, fear of discrimination and poor treatment
arising from previous experiences, and the lack of culturally appropriate services (AIHW, 2011).

The disparity of health status between Aboriginal people and non-Aboriginal people in Australia has been well documented, with diarrhoea caused by gastrointestinal parasites recognised as issues of major importance in Aboriginal populations (Gracey and Cullinane, 2003; Bramley et al., 2004; Holt et al., 2010; Hotez, 2014; Shield et al., 2015). As the majority of Aboriginal people reside in remote areas, the significantly higher cryptosporidiosis notification rates in Aboriginal people compared to non-Aboriginal people (up to 50 times higher), especially in children 0-4 years of age, is likely due to factors associated with remote living such as poor socioeconomic status and household conditions, as well as limited access to nutritious food, potable drinking water, adequate sewage and solid waste disposals; all of which facilitate the faecal-oral transmission of Cryptosporidium, thus increasing the risk of infection (Bailie et al., 2004, Bailie et al., 2010, McDonald and Bailie, 2010, Becker et al., 2015).

Among both Aboriginal and non-Aboriginal people, the largest number of notifications occurred in children aged 0-4 years, in particular those aged 1 year or less, consistent with studies worldwide (Ryan et al., 2016; Kotloff et al., 2016). The multimodal age related pattern observed in non-Aboriginal cases is similar to the general trend observed in previous studies in WA, other states in Australia, as well as in the US, Canada and UK, where there is a cluster of cases among the 20-39 years age groups (Laupland and Church, 2005; Ng et al., 2010b; Yoder and Beach, 2010; Chalmers et al., 2011; Kent et al., 2011; Waldron et al., 2011a). However a third, smaller cluster of cases observed in this study in the 65-79 years age group, suggests that the elderly, likely grandparents, also play a role in care giving for young children. Although a similar distribution was observed in Aboriginal people in this study, the higher notification rates among the 45-64 years age group, compared to the
20-34 years age group, reflects what is known about childcare in Aboriginal communities, where grandparents, instead of parents, are the main care givers to children. This age distribution pattern and the higher cryptosporidiosis rates in females compared to males, highlights the likely occurrence of transmission between children and their caregivers (family members, childcare staff, household contact), possibly through child minding and nappy changing activities (Robertson et al., 2002; Hunter et al., 2004).

Aboriginal people comprise 30-40% of the population in rural and remote areas, in comparison to just 1.5% of the population in metropolitan areas, which may partly explain the increased rates of cryptosporidiosis cases among Aboriginal people in the remote regions of WA, particularly in the Kimberley. Notifications were highest during the Kimberley wet season (data not shown), which occurs from November to April (late spring to autumn), when there is flooding and swimming in natural water holes or catchment is a common practice, resulting in increased potential for transmission.

The predominantly anthroponotically transmitted *C. hominis* was the most common species identified among Aboriginal and non-Aboriginal people, suggesting person to person transmission. A higher proportion of zoonotic *Cryptosporidium* species were detected in non-Aboriginal cases (25.8% v 2.5%), with *C. parvum* and *C. meleagridis* identified mainly from the metropolitan area, but also from rural regions (e.g. South West and Great Southern), where there is greater density of livestock. Four different *C. parvum* subtypes were identified from cases in this region; the IlaA17G2R1 and IlaA18G3R1, which have previously been identified from cattle in the South West region (Ng et al., 2011) and the IlaA19G4R1 and IlaA21G2R1 subtype. There have been no previous reports of the *C. parvum* IlaA19G4R1 and IlaA21G2R1 in livestock in WA, although these subtypes have been identified from cattle in New South Wales (NSW) (Ng et al., 2008). These and the other *C. parvum* Ila subtypes identified in the present study, particularly the IlaA18G3R1 subtype, are common
C. parvum subtypes identified in humans and cattle in Australia (Jex et al., 2007; Jex et al., 2008; Ng et al., 2008; O’Brien et al., 2008; Waldron et al., 2009; Ng et al., 2011; Ng et al., 2012). In the present study, the C. parvum IIdA15G1 subtype was identified in one case, which has been reported in buffalo from the Northern Territory (Zahedi et al., 2016a). As not all C. parvum and C. meleagridis infections are attributed to zoonotic transmission, it is difficult to deduce whether the source of these infections were zoonotic or anthroponotic, without carrying out extensive molecular genotyping of both humans and animals in the regions. However, the identification of these C. parvum genotypes in both humans and livestock (Xiao, 2010), indicates that animals may play a role in the dissemination of Cryptosporidium oocysts. Therefore, further investigation should be carried to determine the extent of zoonoses and the risk to public health as well as to catchment water quality in the area.

Aboriginal people were mainly infected with the C. hominis IIdA15G1 subtype (59%) whereas non-Aboriginal people, were predominantly infected with the IbA10G2 subtype (61.8%). The IIdA15G1 subtype was previously identified as the most common subtype in WA (Ng et al., 2010) and has also been identified in NSW and Victoria (VIC) (Jex et al., 2007; Ng et al., 2008). It has been responsible for waterborne outbreaks in the U.S. (Feng et al., 2014). The IbA10G2 subtype is recognised as a major cause of sporadic cryptosporidiosis in WA, VIC, South Australia (SA), and NSW (Jex et al., 2007; Jex et al., 2008; Waldron et al., 2009; Ng et al., 2010; Waldron et al., 2011a), is highly virulent and has been associated with numerous waterborne cryptosporidiosis outbreaks in Australia (Ng et al., 2010; Mayne et al., 2011; Waldron et al., 2011b; Ng-Hublin et al., 2013) and worldwide (cf. Li et al. 2013).

The identification of distinct C. hominis gp60 subtypes circulating among Aboriginal and non-Aboriginal populations highlights the unique endemicity of cryptosporidiosis in Aboriginal people. Due to low numbers of cases from Aboriginal people genotyped from
other regions in WA, it is difficult to ascertain from this study, whether this endemicity of the
*C. hominis* IdA15G1 subtype is unique to the Kimberley area, as direct comparisons could
not be made in the metropolitan or other rural or remote regions. This finding however,
provides an avenue for more focussed public health attention, better tracking of sources of
infection and implementation of intervention measures to prevent transmission of the disease
and warrants further investigations for a better understanding of the *gp60* subtypes circulating
among the populations and communities of Aboriginal and non-Aboriginal people in the
Kimberley and other areas in WA.

Overall, the findings from the present study shows that the risk of cryptosporidiosis in
Aboriginal people, especially those < 5 years of age is disproportionately greater than in non-
Aboriginal people, with the risk of the disease increasing with remoteness in WA. The
predominance of *C. hominis* infections in both Aboriginal and non-Aboriginal people
indicates that the main transmission route of the disease is person to person; however,
infection with distinct subtypes within the population suggests that the source of infection
and transmission may differ. Factors associated with remote living such as poverty, access to
nutritious food and sanitation resources that may impede the transmission of
*Cryptosporidium*, are likely the underlying factor for the disparities observed in the present
study. More focussed attention on interventions for disease control that includes handwashing
with soap and water, sanitation and hygiene promotion, are most likely to reduce diarrhoeal
illness, especially in children (McDonald et al., 2008) and are currently being carried out by
environmental health practitioners working in Aboriginal communities in remote and rural
zoonotic transmission of cryptosporidiosis in WA however should not be discounted,
particularly in light of the identification of *C. hominis* in kangaroos in Australia (Zhaedi et
al., 2016b) and in cattle in New Zealand (Abeywardena et al., 2012). Further studies and
molecular genotyping involving larger human, livestock and wildlife populations, particularly in livestock dense areas, as well as animals in Aboriginal communities, should be carried out to examine the extent of zoonoses in these communities.

Limitations of the present study include the use of passive surveillance data, which is subjected to under-reporting especially of asymptomatic cases (Lal et al., 2015b). Molecular data was also not available for every case analysed and specimens genotyped were sent from various pathology laboratories from different regions in WA, depending on whether adequate amount for analysis was available, which did not allow for standardisation of sample sizes for cross regional comparisons. However, the use of molecular data to complement existing surveillance data was one of the highlights of the present study, as it provided supporting evidence, and a better explanation and understanding of the surveillance data. Molecular surveillance should be routinely conducted for a better understanding of cryptosporidiosis epidemiology in WA and across Australia, to allow for better detection and more targeted public health interventions.
CHAPTER 4

Comparison of three cryptosporidiosis outbreaks in Western Australia – 2003, 2007 and 2011

4.1 Introduction

Analysis of national notification rates revealed an increased in total notification rates in 2003, 2005, 2006, 2009, 2012, 2013 and 2015, with the highest number of cases reported in 2009 (4,623) (NNDSS, 2016), suggesting that outbreaks may have occurred. The pattern of cryptosporidiosis in Western Australia (WA) however, is somewhat different, with increased notification rates in 2003, 2007 and 2011 (NNDSS, 2016).

Cryptosporidiosis outbreaks in Australia have predominantly been attributed to contaminated recreational waters (Hellard et al., 2000; Puech et al., 2001; Markey, 2002; Black and McAnulty, 2006; Dale et al., 2010; Ng et al., 2010b; Waldron et al., 2011b; Ng-Hublin et al., 2015; Ryan et al., 2016a), with one report of an outbreak due to contaminated milk (Harper et al. 2002), another outbreak linked to contact with animals at a nursery (Ashbolt et al., 2003) and several others, for which the cause was unknown (Dentith and Alexander, 1999; Peacock, 2001).

Of the three cryptosporidiosis epidemic peaks in WA, analysis was only carried out on the outbreak in 2007 (Ng et al., 2010b). For this 2007 community-wide outbreak, molecular characterization identified C. hominis in the samples related to the outbreak. No specific point source outbreaks were identified, although a high proportion (64%) of cases did swim in a pool and 16% reported contact with a person with diarrhoea (Ng et al., 2010b). However, little is known about the other cryptosporidiosis outbreaks that occurred in 2003
and 2011 in WA.

To better understand the transmission, trends and nature of cryptosporidiosis outbreaks in WA, this study aimed to retrospectively compare and review these three cryptosporidiosis outbreaks in 2003, 2007 and 2011 in WA. Analyses was carried out based on case notification data available on the Western Australia Notifiable Infectious Disease Database (WANIDD) and matching molecular genotyping analyses (where available) to better understand the epidemiology of these outbreaks.

4.2 Methods

4.2.1 Outbreak periods, data source and analysis

Data for all reported cases of cryptosporidiosis was extracted from WANIDD for the period of January 2001 to December 2011 by optimal date of onset (ODOO). The ODOO is defined as the ‘true’ date of onset, provided by the notifying doctor or, obtained during case follow-up or the date of specimen collection for laboratory notified cases, and when none of these dates are available, the date of notification by the doctor or laboratory, or the date of receipt of notification, whichever is earliest. Variables exported from WANIDD included public health unit (PHU), sex, race, age, travel history, location where infection was acquired, hospitalization and the infecting Cryptosporidium species based on molecular typing data (where available).

The start and end of each outbreak period was defined as at least a two fold increase in notifications above expected (compared to non-outbreak years) for consecutive months and, capturing the start of the increase and return to expected levels as indicated by the epidemic curve (Figure 4.1). A case was defined as a person diagnosed with cryptosporidiosis
with an onset date of illness within the respective outbreak period. Cases that were not residents in Western Australia were excluded from the analyses.

For information on case travel history, analysis was based on cases which reported ‘yes’ or ‘no’ and any data where history was listed as unknown was excluded from the analysis.

4.2.2 Molecular identification and analysis

Molecular typing was conducted on a total of 167 samples. These included typing conducted with specimens from the 2007 outbreak as part of a previous study by Ng et al. (2010a) \(n = 83\) and Ng et al. (Ng et al., 2010b) \(n = 43\). Molecular typing of the 2011 outbreak \(n = 41\), was carried out as part of the present study. Molecular typing data was unavailable for the 2003 outbreak.

Methods used for molecular identification was as described in Chapter 2. Total DNA extraction was as described in section 2.1. PCR analyses and initial genotyping of the samples was carried out as described in Chapter 3 section 3.2.4. Sequencing analyses was carried out as described in Section 2.5 and analysis of molecular data was conducted as described in Section 2.6 and 2.7.

4.2.3 Data analysis

Data analysis was conducted using the Microsoft Excel 2010 add in, OpenEpi (version 3.01. [http://openepi.com/v37/Menu/OE_Menu.htm](http://openepi.com/v37/Menu/OE_Menu.htm)) and Stata® v13 ([https://www.stata.com](https://www.stata.com)). Incidence rates were calculated using annual census population data available through the Rates Calculator software (Department of Health, WA) and age-
standardized to the Australian population. All rates were annualised so rates could be compared. Public health units (PHUs) in Western Australia were divided into three different regions; Metropolitan (North and South), remote areas (Kimberley, Pilbara and Goldfields PHUs) and rural areas (Midwest, Wheatbelt, South West and Great Southern PHUs) (see Chapter 3; Figure 3.2). WA population estimates for PHU areas, year, sex, age and aboriginality was used to calculate incidence rates and described as number of notified cases per 100,000 population.

The 7-month average of non-outbreak periods was calculated based on notifications from the month of December from the previous year to June, over a 5-year period (See Table 4.1). Notification rate ratios were calculated by dividing the notification rate of metropolitan, rural or remote areas with the notification rate of the 7-month average of non-outbreak years (Table 4.1). As hospitalisation data was only available from 2005 onwards, comparison with non-outbreak years was based on the average number of cases that reported hospitalization per month in 1-year multiplied by the number of months in the corresponding outbreak period. The average was calculated based on the 7-month average of non-outbreak periods over a 4-year period from December 2005-June 2006, December 2007-June 2008, December 2008-June 2009, and December 2009-June 2010.

4.2.4 Ethical clearance

This study was performed with approval from the Murdoch University Human Ethics Committee (Permit number 2012/208) and WA Department of Health Human Ethics Committee (Permit number 2009/48).
**Figure 4.1**: Epidemiological curve showing the number of cryptosporidiosis cases notified from January 2002 to December 2011 in metropolitan, rural and remote areas in WA.

### 4.3 Results

#### 4.3.1 2003 Outbreak: December 2002 – June 2003

During the outbreak period from December 2002 to June 2003, a total of 405 cryptosporidiosis cases were notified. One case (1/405), was not a resident of WA and was therefore excluded from further analyses. The number of cases notified in this outbreak period was 3.1 times higher than the average number of notifications from a similar period in previous non-outbreak year (Table 4.1, Figure 4.1). The annualized rate for this outbreak period was 33 cases/100,000 population. A total of 56.4% (228/404) of notified cases were from the metropolitan area (rate 24 cases/100,000 population), followed by 29.0% (117/404) from remote areas (rate 149 cases/100,000 population) and 16.6% (59/404) from rural areas.
Compared to a similar period in previous non-outbreak years (Table 4.1), the metropolitan area had a 5.2-fold increase in total case notifications (notification rate ratio 4.8). In the remote areas, case notifications were 2.1 times higher than the average number of cases notified for a similar period in non-outbreak years (notification rate ratio 2.1) (Table 4.1) with Pilbara and Goldfields reporting a 2.3-fold and 5.9-fold increase in notifications respectively, and a 1.6-fold increase in case notifications in the Kimberley. In the rural areas, case notifications were 2.0 times higher than the average number of cases notified for a similar period in non-outbreak years (notification rate ratio 2.1) (Table 4.1) with a 4.2-fold increase in cases in the Wheatbelt, 1.4-fold increase in the Midwest, 1.3-fold increase in the South West and 3.0-fold increase in the Great Southern PHU, when compared to a similar period in non-outbreak years.

### 4.3.1.1 Age, Sex, Race

Children in the 00-04 years age group, comprised 58.2% (235/404) (320 cases/100,000 population) of cases, followed by individuals aged 30-34 years (8.9%) (42 cases/100,000 population), 05-09 (6.9%) (36 cases/100,000 population), 35-39 (6.2%) (29 cases/100,000 population) and 10-14 (4.4%) (22 cases/100,000 population). Individuals aged 25-29 and 40-44 years comprised 4.2% of cases each. Those in the other remaining age groups comprised < 2% of cases (Figure 4.2). Overall, there were similar proportions of infection in males (48%) and females (52%). When examining differences between age groups and gender, females aged 30-34 years were 2 times more at risk of infection compared to males in the same age group (RR 2.3; 95% CI, 1.1-5.2, \( p = 0.009 \)). No correlation was found between genders within other age groups.
Table 4.1: Summary of case notification numbers and rates (per 100,000 population) by population groups in the 2003 (December 2002 – June 2003), 2007 (November 2006 – August 2007) and 2011 (January 2011 – May 2011) outbreaks, and a 5-year average (of 7-month periods) in previous, non-outbreak years.

<table>
<thead>
<tr>
<th>Population groups</th>
<th>Outbreaks: Number and (rate per 100,000 population)</th>
<th>*7 month average of non-outbreak periods: December previous year to Jun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total cases</td>
<td>404 (33)</td>
<td>607 (35)</td>
</tr>
<tr>
<td>Metro</td>
<td>228 (24)</td>
<td>405 (30)</td>
</tr>
<tr>
<td>Rural</td>
<td>59 (31)</td>
<td>70 (25)</td>
</tr>
<tr>
<td>Remote</td>
<td>117 (149)</td>
<td>132 (118)</td>
</tr>
<tr>
<td>00-04 years age group Aboriginal people</td>
<td>235 (320)</td>
<td>264 (334)</td>
</tr>
<tr>
<td>Aboriginal people</td>
<td>85 (215)</td>
<td>86 (144)</td>
</tr>
<tr>
<td>Non-Aboriginal people</td>
<td>123 (11)</td>
<td>418 (25)</td>
</tr>
<tr>
<td>Indigenous status unknown</td>
<td>196</td>
<td>103</td>
</tr>
</tbody>
</table>


Of the 404 cases notified in this period, 21% (85/404) (215 cases/100,000 population) of cases were identified as Aboriginal people and 30.4% (123/404) (11 cases/100,000 population) were non-Aboriginal people, with the Indigenous status of 48.5% (196/404) of cases unknown (Table 4.1). Comparing between outbreak and non-outbreak years, the notification rate ratio in Aboriginal people was 2.0 and in non-Aboriginal people was 2.75. Aboriginal people cases were mainly residents of the Kimberley (48.2% - 41/85), whereas more than half (52.8% - 65/123) of non-Aboriginal people cases were residents of the metropolitan area. Amongst Aboriginal people, a high proportion of notifications were from
children in the 00-04 years age group (96.4% - 82/85), compared to 46.3% (57/123) of non-Aboriginal cases in the 00-04 years age group (Figure 4.2). Analysis showed that Aboriginal people in the 00-04 years age group were 10.4 times more at risk of being infected (RR 10.4; 95% CI, 7.3-14.8, p < 0.001) compared to non-Aboriginal people.

Analysis of travel history revealed that 69.3% (283/404) of cases acquired the infection in WA, whilst only 1.5% (6/404) acquired the infection overseas. The place of acquisition of infection was unknown for 28.4% (115/404) of cases. For the 6 cases that acquired the infection overseas, travel to Indonesia (n = 3), Malaysia (n = 1) and New Zealand (n = 1) was recorded, while for one case, no country was listed. Due to incomplete records, no hospitalization data was available for the notifications in this outbreak period. Genotyping data was not available for any of these isolates.

**4.3.2 2007 Outbreak: November 2006 – August 2007**

During the outbreak period from November 2006 to August 2007, a total of 611 cryptosporidiosis notifications were received; four of these cases were not residents of WA and were excluded from further analyses. The number of cases notified in this outbreak was 4.7 times the mean average number of notifications from a similar period in previous non-outbreak years (Table 4.1; Figure 4.1). The annualized rate for this outbreak period was 35 cases/100,000 population. The metropolitan area reported the highest number of notifications, with 66.7% (405/607) of cases (30 cases/100,000 population); this was 9.3 times higher than the average number of cases notified for a similar period in non-outbreak years (notification rate ratio 6.0) (Table 4.1). Cases residing in remote areas constituted 21.7% (132/607) of cases notified (rate 118 cases/100,000 population), which was 2.4 times higher than the average number of cases notified for a similar period in non-outbreak years (notification rate
ratio 1.7) (Table 4.1); increases in total case notifications were reported in the Pilbara (2.2-fold), Kimberley (2.2-fold) and Goldfield (4.7-fold). Rural areas constituted 11.5% (70/607) of cases notified (rate 25 cases/100,000 population); this was 2.3 times higher than the average number of cases notified for a similar period in non-outbreak years (notification rate ratio 1.7) (Table 4.1), with increases in case notifications observed in all areas of the region (Wheatbelt - 2.5-fold; Great Southern - 3.7-fold; South West - 1.9-fold; Midwest - 2.7-fold).

4.3.2.1 Age, Sex, Race

The number of notifications in each age group, stratified by Indigenous status, is presented in Figure 4.2. The highest proportion of cryptosporidiosis notifications were in the 00-04 years age group (235 cases/100,000 population), comprising 43.5% (264/607) of cases followed by those in the 05-09 years (9.1%) (48 cases/100,000 population), the 30-34 years (8.1%) (40 cases/100,000 population), 35-39 years (6.8%) (31 cases/100,000 population) and 20-24 years (4.6%) (22 cases/100,000 population) age groups. Individuals in the 10-14 and 25-29 age groups comprised 4.4% and a rate of 22 cases/100,000 population of cases each. Notifications in the other remaining age groups comprised < 4% of cases notified in this outbreak period. Overall, there was a similar proportion of infection in males (48%) and females (52%). In the 20-24 and 30-34 year age groups, females were 2.7 times (RR 2.69; 95% CI, 1.1-7.1, p < 0.05) and 2.8 times more likely to have cryptosporidiosis (RR 2.8; 95% CI, 1.5-5.5, p < 0.05) compared to males in the same age groups.

A total of 14.2% of the 607 notifications in this time period were Aboriginal people (144 cases/100,000 population) and 68.9% (418/607) were non-Aboriginal people (25 cases/100,000 population) (Table 4.1). Indigenous status was not recorded for 17.0% (103/607) of cases. Comparing the outbreak period with a similar period in non-outbreak
years, Aboriginal people had a notification rate ratio of 1.35, whilst non-Aboriginal people had a notification rate ratio of 6.25.

**Figure 4.2:** Cryptosporidiosis notifications by race in each age group for the 2003, 2007 and 2011 outbreaks.
The majority of notifications in Aboriginal people were from the Kimberly area (70.9% - 61/86), and in non-Aboriginal people the majority of cases were from the metropolitan area (76.1% - 318/418). Among Aboriginal people, 91.2% of cases (79/86) occurred in the 00-04 years age group (Figure 4.2). Aboriginal people in the 00-04 years age group had 8.1 times higher risk of being a notification compared to non-Aboriginal people in the same age group (RR 8.05; 95% CI, 6.0-10.6, p < 0.05).

Place of acquisition data was available for 485/607 cases and showed that 87.4% (424/485) of cases acquired cryptosporidiosis in WA during the 2007 outbreak, with 1.2% (6/485) and 11.3% (55/485) of cases acquiring the disease interstate and overseas respectively. For the remaining 122/607 cases, data on place of disease acquisition was either not reported or unknown. Cases that acquired the disease overseas reported travelling to 25 different countries in Africa, Asia, Antarctica, Europe, North and South America and Oceania. The most frequently reported country of travel was Indonesia (n = 14), followed by Thailand (n = 7). For 4/55 cases, the country of travel was not listed.

Within this outbreak period, a total 46/607 (7.58%) cases were hospitalized and 414/607 (68.2%) cases reported no hospitalization. Hospitalization status for 147/607 was not known. The number of cases that were hospitalized was 1.67 times higher when compared to a similar period in non-outbreak years, where on average, 27.5 cases were hospitalized. Children in the 00-04 age group comprised 63% of hospitalized cases.

Molecular typing of the 2007 outbreak

Species identity data was available for 126 cases notified in the 2007 outbreak period (Ng et al., 2010a; Ng et al., 2010b), of which, two cases were not residents of WA and excluded from this analysis. Molecular characterization at the 18S rRNA and gp60 gene loci
identified *C. hominis* in 123 cases and *C. parvum* in one case. Subtyping of the *C. hominis* and *C. parvum* isolates at the *gp60* gene locus identified five *C. hominis* subtypes families: Ib (n = 57), Id (n = 47), Ie (n = 1), If (n = 2) and Ig (n = 16). Within the *C. hominis* Id subtype family, three subtypes were identified (IdA15G1, IdA16 and IdA17). The one *C. parvum* isolate was identified as the IaA18G3R1 subtype. The most common subtype identified was the IbA10G2 subtype in 46.3% of the *C. hominis* isolates, followed by the IdA15G1 (33.3%) and the IgA17 subtype (13.0%).

The highest subtype diversity was observed in the metropolitan area, where all seven *C. hominis* subtypes and one *C. parvum* subtype was identified (Table 4.2). The IbA10G2 was the most common subtype and identified predominantly in cases from the metropolitan area, in 66.7% (34/51) of cases (Table 4.2). In remote areas (Goldfields, Kimberley, Pilbara), cases were mainly attributed to infection with the IdA15G1 subtype, particularly in the Kimberley (79.4% - 27/34), with no other Id subtypes identified in this area (Table 4.2). Data analysis showed that prior to the outbreak, no cases of IbA10G2 were identified in the remote areas, after which, 11 cases were identified with this subtype late in the outbreak (February - May 2007).

Non-Aboriginal cases were mostly infected with the IbA10G2 subtype (60.6% - 43/71), followed by the IdA15G1 subtype (21.1% - 15/71) (Table 4.3). The IdA15G1 subtype was also the dominant subtype in 66.7% (24/36) of cases in Aboriginal people (Table 4.3). Non-Aboriginal people were 5.5 times more likely to be infected with the IbA10G2 subtype (RR 5.5; 95% CI, 2.1-14.0, p < 0.05) compared to Aboriginal people, but the latter were 3.2 times more likely to be infected with the IdA15G1 subtype, compared to non-Aboriginal people (RR 3.2; 95% CI, 1.9-5.2, p < 0.05).
Table 4.2: Cryptosporidium species and gp60 subtypes identified in cases of cryptosporidiosis during two outbreaks (2007 and 2011) in Western Australia, stratified by Public Health Unit.

<table>
<thead>
<tr>
<th>Public Health Unit</th>
<th>Number of cases in the 2007 Outbreak</th>
<th>Number of cases in the 2011 Outbreak</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C. hominis</td>
<td>C. parvum</td>
</tr>
<tr>
<td></td>
<td>IbA10G2</td>
<td>IdA15G1</td>
</tr>
<tr>
<td>Wheatbelt</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Goldfields</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Great Southern</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Kimberley</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>Metropolitan</td>
<td>34</td>
<td>3</td>
</tr>
<tr>
<td>Midwest</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Pilbara</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>South West</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 4.3: *Cryptosporidium* gp60 subtypes identified in cases of cryptosporidiosis from the 2007 and 2011 outbreak in Western Australia stratified by Indigenous status; Aboriginal people, non-Aboriginal people and unknown.

<table>
<thead>
<tr>
<th>Species</th>
<th>gp60 subtype</th>
<th>No. of cases in 2007 Outbreak</th>
<th>No. of cases in 2011 Outbreak</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Aboriginal</td>
<td>Non-Aboriginal</td>
</tr>
<tr>
<td><em>C. hominis</em></td>
<td>IbA10G2</td>
<td>4</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>IdA15G1</td>
<td>24</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>IdA16</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>IdA17</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>IeA11G3T3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>HaA12G1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>IgA17</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td><em>C. parvum</em></td>
<td>HaA15G2R1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>HaA18G3R1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Place of acquisition data was available for 101/124 genotyped cases. There was significant subtype diversity among cases that acquired the infection in WA, with the identification of six different *C. hominis* subtypes (IbA10G2, IdA15G1, IdA16, IeA11G3T3, IfA12G1 and IgA17) and one *C. parvum* IIaA18G3R1 subtype. For cases that acquired the infection overseas, genotyping data was only available for six cases; four were infected with the IbA10G2 subtype and reported travel to Cambodia (n = 1) and Indonesia (n = 3) respectively; one was infected with the IdA17 subtype and reported travel to Vietnam; and one was infected with the IdA15G1 subtype, with no country of travel reported.

Genotyping data available for 17 hospitalized cases showed that eight were infected with the IbA10G2 subtype, six with the IdA15G1 subtype and three with the IgA17 subtype. Due to the small sample size, no correlation between subtype and hospitalization could be determined.

### 4.3.3 2011 Outbreak: January 2011 – May 2011

A total of 356 cryptosporidiosis notifications were received during the outbreak period from January 2011 to May 2011, of which, one case was not a resident of WA and excluded from further analyses. The number of case notifications in this outbreak was 2.8 times higher than the mean average number of notifications for a similar period in the previous five non-outbreak years (Table 4.1; Figure 4.1). The annualized rate for this outbreak period was 36 cases/100,000 population. The metropolitan area reported the highest number of notifications (53.5% - 190/355) (28 cases/100,000 population), which was 4.3-fold higher than the average number of cases notified for a similar period in non-outbreak years (notification rate ratio 5.6) (Table 4.1). This was followed by the rural areas (23.1% - 82/355) (56 cases/100,000 population), which showed a 2.8-fold increase in case notifications.
(notification rate ratio 3.7), and remote areas (23.4% - 83/355) (125 cases/100,000 population), which showed a 1.5-fold increase in case notifications (notification rate ratio 1.8), when compared to the average number of cases notified for a similar period in non-outbreak years. An increase in total case notifications was observed for all rural (Midwest - 5.3-fold; Wheatbelt - 2.5-fold; Great Southern - 3.3-fold; South West - 1.6-fold) and remote (Pilbara - 1.7-fold; Goldfields - 2.2-fold; Kimberley - 1.3-fold) areas.

**Age, Sex, Race**

The highest number of cryptosporidiosis notifications was observed in the 00-04 years age group (51.8% - 184/355) (283 cases/100,000 population), followed by the 05-09 years (10.3%) (60 cases/100,000 population), 35-39 years (7.3%) (37 cases/100,000 population), 30-34 years (7.0%) (36 cases/100,000 population) and 25-29 years (4.2%) (19 cases/100,000 population) age groups (Figure 4.2). Notifications in the other remaining age groups comprised < 4% of cases notified in this outbreak period (Figure 4.2). There were no significant difference in the rate of notifications for males and females in each of the age groups.

There was a higher number of notifications in non-Aboriginal people (73.6% - 262/355) (28 cases/100,000 population) compared to Aboriginal people (22.8% - 81/355) (256 cases/100,000 population). Indigenous status was not recorded for 3.4% (12/355) of notified cases. Comparing the outbreak period and a similar period in non-outbreak years, Aboriginal people had a notification rate ratio of 2.39, whilst non-Aboriginal people had a notification rate ratio of 7.00. Notifications in non-Aboriginal people were primarily from the metropolitan area (67.9% - 178/262), whereas notifications in Aboriginal people were primarily from the Kimberley area (50.6% - 41/81), followed by the Midwest (16.0% - 13/81)
and Pilbara areas (14.8% -12/81). The proportion of Aboriginal cases from children in the 00-04 years age group was significantly higher when compared to non-Aboriginal people (93.8% vs. 39.3%) (Figure 4.2). Analysis showed that Aboriginal people in the 00-04 years age group were 14 times more at risk of being infected (RR 13.5; 95% CI, 9.9-18.3, p < 0.05) compared to non-Aboriginal people in the same age group.

Of the 355 cases notified within the 2011 outbreak period, place of acquisition was recorded for 312 of the 355 cases notified in the 2011 outbreak period. Of these, 93.3% (291/312) of cases were acquired in WA, with only 6.7% of cases acquired overseas. For cases that acquired the infection overseas, travel to seven different countries were reported, with Indonesia being the most frequently reported country (n = 11), followed by Malaysia (n = 3) Bangladesh (n = 1), Ethiopia (n = 1), Morocco (n = 1), Thailand (n = 1) and Vietnam (n = 1). Country of travel for two cases was not listed.

Hospitalization data was available for 344/355 cases, with the hospitalization status for 91/355 of cases unknown. Hospitalization was reported for 50/344 (14.5%) cases while, 214/344 (62.2%) cases reported no hospitalization. The number of cases that were hospitalized was 3.64 times higher when compared to a similar period in non-outbreak years, where on average, 13.75 cases were hospitalized. Children from the 00-04 years age group comprised the highest proportion of hospitalized cases at 72%.

Molecular typing of the 2011 outbreak

Molecular characterization of 42 cases in the 2011 outbreak period identified *C. hominis* and *C. parvum* in 38 and four cases respectively. Subtyping identified two *C. hominis* subtypes families, Ib and Id and two *C. parvum* subtypes from the IIa family. The IbA10G2 subtype was the most common *C. hominis* subtype, identified in 89.5% (34/38) of
cases, with the IdA15G1 subtype identified in 10.5% (4/38) of *C. hominis* isolates. Subtyping was successful for three *C. parvum* isolates, which were identified as IIA18G3R1 (n = 2) and IIA15G2R1 (n = 1). The subtype for one *C. parvum* isolate could not be determined.

Cases identified with the *C. hominis* IbA10G2 subtype (n = 34) were predominantly from the metropolitan area (30/34), with only two cases identified from the Great Southern and one case each from the Wheatbelt and South West area. The *C. hominis* IdA15G1 subtype was only identified in the Wheatbelt (n = 2) and South West (n = 2) (Table 4.2). *Cryptosporidium parvum* was identified in two cases in the metropolitan area, one of which was the IIA18G3R1 subtype, while the subtype from the other case could not be determined. The *C. parvum* IIA18G3R1 and IIA15G2R1 subtypes were identified in the South West (n = 1) and Great Southern (n = 1) regions respectively (Table 4.2).

Of the 42 cases genotyped in this outbreak period, non-Aboriginal people consisted of 90.4% (38/42) of cases, and Aboriginal people 4.8% (2/42), with no recorded Aboriginal status for 4.8% (2/42) of cases. The *C. hominis* IbA10G2 subtype was identified in 94.1% (32/38) of non-Aboriginal cases, with *C. hominis* IdA15G1 (n = 3), *C. parvum* IIA15G2R1 (n = 1), IIA18G3R1 (n = 1) and an undetermined *C. parvum* subtype (n = 1) identified in the remaining non-Aboriginal cases (Table 4.3). In the two Aboriginal cases typed, the *C. hominis* IbA10G2 and the IdA15G1 was identified respectively. Due to the small number of cases in Aboriginal people genotyped (n = 2), no statistically significant association could be made.

Place of acquisition data was available for 37 of the 42 cases and 97.3% (36/37) of these acquired the infection in WA, with only one case (2.7%) acquired overseas. For those that acquired the infection in WA, *C. hominis* was identified in 89.2% (33/37) of cases, with IbA10G2 the most common subtype (n = 29), followed by the IdA15G1 subtype (n = 3). *Cryptosporidium parvum* was identified in 8.1% (3/37) of cases; IIA15G2R1 (n = 1),
IIaA18G3R1 (n = 1) and undetermined (n = 1). Genotyping analysis carried out on the one case that acquired the infection overseas (Malaysia), identified the IbA10G2 subtype.

Of the 26 cases for which hospitalization status was available, the majority, 80.7% (21/26) reported no hospitalization. Of the remaining 19.2% (5/26) of cases, subtypes IbA10G2 (n = 3), IdA15G1 (n = 1) and *C. parvum* IIaA15G2R1 (n = 1) were identified.

### 4.4 Discussion

Of the three outbreaks compared in the present study, the 2007 outbreak had the highest number of case notifications, followed by the outbreaks in 2003 and 2011. These outbreaks appeared to have occurred predominantly in the urban metropolitan area, where the highest number of case notifications and notification rate ratios were reported. An increase in case notifications were also observed in rural and remote areas, but these areas have non-outbreak rates that are 3-12 fold higher than the metropolitan area. One possible explanation could be that, the metropolitan area, which normally has lower rates than rural or remote areas, has a larger population susceptible to *Cryptosporidium* infection, hence, when an outbreak occurs, it results in a larger proportional increase compared to other areas. *Cryptosporidium hominis* was the most commonly identified species in both the outbreaks where molecular genotyping data was available. This is consistent with previous studies, which have found that infections with *C. hominis* are common in urban areas with high population density, as it provides a stable platform for human to human transmission (Xiao, 2010). However, due to difficulties in distinguishing between outbreak and sporadic cases in a retrospective study, cases notified during the three outbreaks periods however, may also include cases that occurred sporadically.

A previous investigation into the 2007 outbreak revealed that it was likely associated
with swimming pools with subsequent secondary person to person transmission (Dale et al., 2010; Ng et al., 2010b). The 2007 outbreak was also the most prolonged outbreak of the three, beginning in late spring (November 2006) and ending in late winter (August 2007). Previous studies have shown that waterborne cryptosporidiosis outbreaks were more likely to occur during this period, as a result of an increase in recreational water activities (Dale et al., 2010; Yoder and Beach, 2010). The three outbreaks presented in the present study occurred during late spring/early summer periods, when there would have been an increase in recreational water activities such as swimming in swimming pools, water catchments and natural water holes. This period also coincides with summer vacation travel from metropolitan to rural and remote areas, which may have contributed to the spread of the outbreak.

The *C. hominis* IbA10G2 subtype is a major cause of sporadic cryptosporidiosis in WA (Ng et al., 2010a), Victoria (Jex et al., 2007; Jex and Gasser, 2008; Koehler et al., 2013), South Australia (SA) (Ng et al., 2010b), New South Wales (NSW) (Jex et al., 2007; Jex and Gasser, 2008; Waldron et al., 2011a) and Tasmania (Koehler et al., 2014). It has also been associated with various waterborne cryptosporidiosis outbreaks in Australia (Ng et al., 2010b; Waldron et al., 2011b; Mayne et al., 2011) and a number of large waterborne cryptosporidiosis outbreaks in the United States (Zhou et al., 2003), Canada (Ong et al., 2004), France (Cohen et al., 2006), United Kingdom (Glaberman et al., 2002; Chalmers et al., 2010) and Sweden (Widerström et al., 2014). In both the 2007 and 2011 outbreaks, the *C. hominis* IbA10G2 subtype was the most common subtype identified. This subtype is endemic in the WA metropolitan area but is less common in remote and rural areas of WA (Ng et al., 2010a; Ng-Hublin et al., 2017). Interestingly, the IbA10G2 subtype was only identified in cases from remote areas late in the 2007 outbreak. This seems to suggest spread of the outbreak subtype from urban to rural and remote areas. However, genotyping a larger number
of cases from rural and remote areas is needed to provide more conclusive evidence on the
distribution of different subtypes and to confirm the urban-rural movement during the
outbreak.

The observation that notification rates were highest in children aged 00-04 years and
adult females aged 30-34 years is consistent with other studies in WA (Ng et al., 2010a;
2010b; see Chapter 3). This highlights that children ≤ 4 years are most susceptible to
infection, either due to no previous exposure to Cryptosporidium, and/or, have behavior that
encourages infection. This also suggests secondary transmission from infected children to
adult females, who are likely to have roles as caregivers within the family, in schools or in
nurseries, and day care settings (Robertson et al., 2002; Hunter et al., 2004; Yoder and Beach,
2010). Cryptosporidium is resistant to alcohol and hand sanitisers, hence, proper sanitary
practices including washing hands with soap are recommended, especially during outbreaks,
as an important measure to prevent subsequent person to person transmission (Vandenberg et
al., 2012).

Non-Aboriginal people were most affected in all three outbreaks analyzed in the
present study, comprised the majority of cryptosporidiosis cases notified, and compared to
Aboriginal people, had higher notification rate ratios than in non-outbreak years. The higher
number of notifications in non-Aboriginal people may be a reflection of the distribution of
Aboriginal populations in WA. As observed in Chapter 3, the vast majority of non-Aboriginal
people live in urban areas, while most Aboriginal people live in rural and remote areas of
WA. The notification rate in Aboriginal people however, is 6-20 times higher than in non-
Aboriginal people, which highlights the unequal distribution of disease burden in these
communities. This reflects a broader pattern of health inequity that results in significantly
higher burden of disease and lower life expectancies in Aboriginal and Torres Strait Island
peoples (AIHW, 2016).
Infection in Aboriginal people, which constituted the majority of cases in remote areas, was mainly attributed to the IdA15G1 subtype (data not shown). The IdA15G1 subtype is thought to be endemic to the Aboriginal population, particularly in the remote areas of WA (see Chapter 3). This indicates that person to person transmission occurs mainly within, rather than between, these two population groups. However, as many apparently sporadic cases may be part of small outbreaks that go undiagnosed (Briggs et al., 2014), the high proportion of cases from the same area identified with this subtype suggests that, in addition to the outbreak attributed to the IbA10G2 subtype, a second, smaller scaled outbreak attributed to the IdA15G1 subtype, may have occurred concurrently amongst the Aboriginal communities.

In both the 2007 and 2011 outbreaks, a higher number of cases were hospitalized compared to non-outbreak years, with >50% of cases hospitalized being young children < 4 years in age. The 2011 outbreak had a higher proportion of cases hospitalized compared to the 2007 outbreak; however, as information on clinical symptoms and/or immunological status of these cases was not available, it is difficult to ascertain if symptoms in one outbreak might be more severe than the other. Molecular genotyping showed that the majority of the 2007 and 2011 cases hospitalized, were infected with the *C. hominis* IbA10G2 subtype. The IbA10G2 has been reported as the most virulent *C. hominis* subtype, with those infected more likely to present multiple symptoms such as nausea, malaise and vomiting in addition to diarrhea (Cama et al., 2008; Li et al., 2013). However, due to the small number of hospitalized cases that were typed, a significant correlation between subtypes in each outbreak with symptoms and hospitalization could not be examined.

For the 2003 outbreak, travel history was only available for a small proportion of cases (1.7%) and hence, meaningful comparison with the latter two outbreaks could not be carried out. In the 2007 and 2011 outbreaks, most cases (82-87%) acquired their illness in
WA with only 7-11% of cases having acquired their infection overseas. This indicates that local transmission was driving these outbreaks. There have been reports from previous studies of travelers returning from overseas with cryptosporidiosis and subsequently spreading the infection in the community (mainly through swimming pools) (Lazensky et al., 2008; Chalmers et al., 2011). However, the primary cases in these outbreaks were not identified and although the IbA10G2 subtype was identified in cases with overseas travel history in the 2007 outbreak, no conclusions can be made based on the small number of cases genotyped. Further investigation and genotyping of additional cases with overseas travel history is required, to better understand the importance of overseas travel in local transmission within Western Australia.

Early detection and rapid response to an increase in Cryptosporidium incidence is essential for outbreak control. Waterborne outbreaks often remain undetected, as there can be a considerable delay before an outbreak is recognized, and are usually more widespread, hindering timely detection and the ability to ascertain a common source or exposure (Hellard et al., 2000; Dale et al., 2010). This would have been the case for the three cryptosporidiosis outbreaks in WA (2003, 2007 and 2011); inundation of cases from multiple possible exposures (i.e. childcare, schools, swimming pools, family members) and competing priorities, which results in insufficient resources to interview every case, are likely to impede timely identification of risk factors and point source outbreaks. As this is a descriptive retrospective study, the cause of the outbreak or mode of transmission in these three outbreaks cannot be established. The seasonal and age related trends as well as the characteristics of the dominant IbA10G2 subtype detected, does however, suggest that these outbreaks were due to person to person transmission, and that undetected waterborne outbreaks (Hellard et al., 2000; Dale et al., 2010), may have played a part in the transmission as well as the scale and length of these outbreaks. It would also be interesting to compare
meteorological data between the three outbreaks, to examine if there were any correlation with the scale of the outbreak, as well as outbreak versus non-outbreak years to examine if any potential correlations between weather patterns and outbreak triggers or persistence of an outbreak.

An improved understanding of the epidemiology, sources and transmission of cryptosporidiosis is needed to develop better control programs. Additional information on case exposures, subtypes identified from routine molecular genotyping of a larger pool of cases as well as clinical symptoms would certainly provide a more in depth and robust understanding of the pathogenicity and epidemiology of the Cryptosporidium in each of these WA outbreaks. This would assist in identifying key factors that may be used towards prediction or prevention of further cryptosporidiosis outbreaks in WA. Across Australia, species and subtype data should routinely be incorporated into national surveillance programs, which would facilitate a web-based database, enabling epidemiological interpretation of genotype occurrence and distribution trends in both sporadic and outbreak cases.
CHAPTER 5

Investigation of a community cryptosporidiosis outbreak in the Kimberley - November/December 2012

5.1 Introduction

In November and December 2012, the Kimberley Population Health Unit (KPHU) (located in the Kimberley region of Western Australia) reported 18 notified cases of cryptosporidiosis, compared to the five-year average for this period of <1 case. The outbreak was suspected to be waterborne as a number of school aged children that were diagnosed with Cryptosporidium had been swimming at a Broome public swimming pool (BPP). In response to the increase in notifications, the Broome local government environmental health officers alerted the local primary school, child-care centres and operators of other swimming pools to advise on any potential risks and infection control.

Although not routinely carried out in Australia, molecular tools were applied to clinical and environmental samples in this investigation. Environmental and epidemiological investigations were also conducted by the Shire of Broome local government environmental health officers and OzFoodNet Communicable Disease Control Directorate respectively, to better understand the source of the outbreak.
5.2 Materials and Methods

5.2.1 Case definition

A case was defined as a person diagnosed with cryptosporidiosis with an onset date of illness from November to December 2012 and who was a resident of Broome, Kimberley.

5.2.2 Epidemiological Investigation

*Cryptosporidium* notifications in the Kimberley region were referred by the Kimberley PHU to the Kimberley Local Government Shires where the cases lived. Cases were then interviewed using a structured questionnaire on known risk factors for *Cryptosporidium*, by the Local Government Environmental Health Officers (LGEHO). Demographic information of the patients that met with the case definition were exported from WANIDD. Descriptive analysis was conducted based on demographic information of cases and information obtained from the questionnaire survey of the cases to identify common exposures.

5.2.3 Water samples

Water samples were collected post-treatment and sent to Murdoch University for molecular analysis on the 8th of January 2013 in 10L aliquots from three different swimming pool locations; (1) BPP, following reports of additional cryptosporidiosis cases received on the 2nd of January 2013, that indicated that these cases had swum at BPP after superchlorination treatment on the 12th of December, 2012. No further exposure information on these cases was available; (2) A water playground with a small pool located by the beach (WP), which was frequented by children and cases that had been exposed to BPP. Treatment of the pool was by
UV radiation; and (3) a swimming pool from a residence in Roebuck Estate, where one case had reported swimming, prior to becoming ill.

The water samples received were concentrated and purified immediately in accordance with the USEPA 1623 method. Briefly, concentration of water samples were carried out using Envirochek filters (Gelman) and resultant eluates were purified by immunomagnetic separation (IMS) using a Dynabeads®GC Combo kit (Dynal Invitrogen, Norway) according to the manufacturer’s protocol with the resulting pellet subjected to DNA extraction.

5.2.4 Clinical samples

Where sufficient material was available, faecal samples of patients that met with the case definition and were diagnosed as microscopy positive for Cryptosporidium by PathWest Laboratories, Perth, were sent to Murdoch University for molecular analysis. As part of routine diagnostic examination by the pathology laboratory, the faecal samples were also tested for Salmonella spp., Campylobacter spp., Shigella spp., Vibrio spp., Yersinia spp., Plesiomonas spp. and rotavirus.

5.2.5 Molecular characterisation and analysis DNA Extraction, PCR amplification and Sequence analysis

Methods used for molecular characterisation of clinical and water samples in this study was carried out as described in Chapter 2. Details for DNA extraction were described in Section 2.1, PCR analyses were described in Section 2.2, sequencing in Section 2.5 and analysis of molecular data was conducted as described in Section 2.6 and 2.7.
5.3 Results

5.3.1 Epidemiological Investigations

A total of 18 patients were diagnosed positive for the presence of Cryptosporidium by microscopy and met with the case definition with an onset date of illness from the 9th of November, 2012 to the 26th of December, 2012 (Table 5.1). These cases were negative for the presence of Salmonella spp., Campylobacter spp., Shigella spp., Vibrio spp., Yersinia spp., Plesiomonas spp. and rotavirus.

The average age was 11 years with an age range of <1 years to 47 years old. The highest number of cases were from individuals aged 00-04 years (9/18), followed by 05-09 years (5/18). The average duration of diarrhoea was nine days (range 3 to 21 days), with no reports of hospitalisation. A total of 25 family members of the known cases of cryptosporidiosis were also ill with similar symptoms and with onsets of illness at the same time or after the onset of the notified cases.

Eleven of the confirmed cases of cryptosporidiosis (65%) reported swimming at the BPP prior to illness onset from the 9th of November to the 26th of December, 2012 (Table 5.1). Other potential exposures reported for these 11 cases included swimming at WP (n = 2), other siblings in the household that had similar illness and that had also swam at the BPP (prior to the cases becoming ill) (n = 3) and having attended childcare centre A (n = 2) (Table 5.1). For the seven cases (35%) that did not report swimming at the BPP, potential exposures include having attended Primary school B (n = 2), attended childcare centre A (n = 1) and other siblings in the household that had similar illness prior to the cases becoming ill (n = 4) with one of the siblings having attended Primary school B (Table 5.1). None of the confirmed cases reported contact with farm or wild animals. Only one individual reported travel outside of the Broome area, (this individual had returned from India and was unwell on the 17th of
November 2012. This individual (Case 7) swam at the BPP nearly every day, but contrary to the doctor’s advice, had continued to swim at BPP. This person had a housemate who was also unwell with undiagnosed gastroenteritis and swam nearly every day at BPP, and also continued to swim while ill. The onset of illness of the housemate was unknown.

5.3.2 Molecular analysis of clinical samples

Although 18 cases met with the case definition, sufficient faecal material was available for only 8 cases and sent for molecular analysis. PCR amplification and nucleotide analysis at the gp60 gene locus identified the *C. hominis* subtype IbA10G2 subtype in all eight faecal DNA samples analysed (Table 5.1).

5.3.3 Molecular analysis of water samples

Of the water samples collected from three different locations and analysed at the *gp60* locus, *Cryptosporidium* was detected in one water sample collected from the WP and was typed as the *C. hominis* IbA10G2 subtype. No *Cryptosporidium* was detected in the water samples from the BPP swimming pool and the Roebuck Estate household pool. The water samples were collected 3 weeks after treatment of pools at each respective site with the BPP swimming pool being treated by superchlorination (chlorine levels raised to 20 ppm for at least 12.75 hours) and the WP treated with ultraviolet light (UV) irradiation. It is not known if the water sample from the household pool from the Roebuck estate was treated despite recommendation for superchlorination, or if the water sample was collected pre- or post-treatment.
Table 5.1: Details of possible Cryptosporidium exposure of cases and Cryptosporidium species and subtypes identified from cases where molecular genotyping was carried out.

<table>
<thead>
<tr>
<th>Case</th>
<th>Age</th>
<th>Exposure to Cryptosporidium</th>
<th>Genotyping results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Swam at BPP</td>
<td>Swam at WP</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>Y</td>
<td>UK</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>Y</td>
<td>UK</td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>9</td>
<td>47</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>10</td>
<td>36</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>11</td>
<td>35</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>12</td>
<td>&lt; 1</td>
<td>N</td>
<td>UK</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>14</td>
<td>12</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>15</td>
<td>&lt; 1</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>17</td>
<td>11</td>
<td>N</td>
<td>UK</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>N</td>
<td>UK</td>
</tr>
</tbody>
</table>

NB: Y – Yes; N – No; UK – Not known; UD – Not determined; ^Household member who ill and attends Childcare centre A *Household member who is ill and attends Primary school B
5.4 Discussion

In the present study, the identification of the same *C. hominis* subtype IbA10G2 in all eight faecal samples genotyped from an outbreak of cryptosporidiosis in the Shire of Broome and from a water sample from WP suggests that the IbA10G2 subtype was responsible for the outbreak. Previous analysis of *Cryptosporidium* subtypes from a longitudinal study in WA from 2005 to 2008 (with corresponding demographic data from WANIDD), revealed that the IbA10G2 subtype was uncommon in the Kimberley region (Ng et al., 2010a). In that study, the IbA10G2 subtype was responsible for only 5% of cryptosporidiosis cases that were typed, with the majority of cases (>80%) attributed to infection with the *C. hominis* IdA15G1 subtype. Across Australia, the *C. hominis* subtype IbA10G2, is the most common cause of sporadic human cryptosporidiosis in Victoria and New South Wales and the third most common cause of sporadic human cryptosporidiosis in Western Australia. It has also been responsible for previous community-wide outbreaks of cryptosporidiosis in Western Australia and South Australia in 2007 and an outbreak in New South Wales in 2009 (Ng et al., 2010b; Waldron et al., 2011b).

From the investigations conducted, the BPP swimming pool was a likely source of infection (Vandenberg et al., 2012), with 65% of the cases reported swimming at the pool. As part of the public health action taken as a result of the outbreak, superchlorination of the BPP swimming pool was carried out on two separate occasions with all ablution facilities cleaned and sanitised, and ablution cleaning equipment replaced with new equipment. Despite a potential source of infection for this outbreak, molecular typing of water samples collected from the BPP pool water was negative for
Cryptosporidium by PCR. The water sample, however, was collected approximately 1 month after the first superchlorination treatment. This followed reports of additional cryptosporidiosis cases received that indicated exposure to the BPP swimming pools after the initial superchlorination treatment. Subsequently, a second superchlorination treatment was carried out at the BPP. Unfortunately, no exposure information was available for these additional cases although the negative Cryptosporidium pool water result suggests that the BPP may not be the source of infection on this occasion.

Although 35% (7/18) of cases did not report swimming at the BPP, these cases had contact with people with similar illness at the childcare centre and primary school, which suggests that person to person transmission was occurring. There were also family members of cases who had similar illness. Cryptosporidium infection can easily be transmitted from person to person, particularly in the absence of good hygiene measures, such as washing of hands with soap and water after using the toilet or changing diapers and before and after handling of food and tending to someone with diarrhea (Vandenberg et al., 2012). One individual (Case 13), who was thought to have become infected after swimming at the pool in the Roebuck Estate residence also had other members of the household with similar illness. Molecular analysis of the water sample collected from this pool showed that the pool was negative for the presence of Cryptosporidium, providing reason to suggest that the case become ill as a result of person to person transmission amongst household members and not due to swimming in a contaminated pool.

The WP facility did test positive for the C. hominis IbA10G2 subtype. This facility was pre-fitted with UV filter for water disinfection and as a result of this, the Shire of Broome felt that the facility did not pose a risk for Cryptosporidium
contamination. The detection of *Cryptosporidium* at this site does not necessarily indicate that the UV treatment was unsuccessful as the DNA analysis conducted was not a measure of viability (i.e. the positive PCR result may have come from oocysts that had been inactivated by the UV treatment and were no longer viable). Viability testing would require a cell culture viability assay (King et al., 2011), which was beyond the scope of the present study.

A limitation in this investigation was that no water samples were collected prior to treatment of the swimming pools for molecular analysis. Comparison of water sample analysis before and after treatment would provide stronger evidence linking potential source of outbreak to cases identified as well as enable observation of the effectiveness of the water treatment used in these circumstances. However, based on the association of the *C. hominis* IbA10G2 subtype and waterborne cryptosporidiosis outbreaks, its identification in all outbreak case samples in this study, where 65% of cases had exposure to the same swimming pool, as well as in a water sample suggests that this outbreak was likely to be waterborne and person to person transmission. This is in agreement with the outcome of epidemiological investigations conducted by OzFoodNet and the Broome Local Government and Environmental Health staff.

It is well known that accidental faecal releases (AFR’s) in pools by infected swimmers can release large numbers of *Cryptosporidium* oocysts. For example, an infected individual can excrete up to one billion oocysts during an infectious period and 1 ml of faeces can contain as many as $5 \times 10^7$ *Cryptosporidium* oocysts (Polgreen et al., 2012). If a child has a loose bowel movement of 150 ml into a typical 25m x 12m municipal pool of about 450m$^3$, this would result in an average concentration of about
20,000 oocysts/litre (20/ml). When an area has a large number of swimmers, these swimmers will contribute to the mixing process. The average swimmer ingests ~28-51ml of pool water (children and adults respectively) (Schets et al., 2011), therefore, a swimmer swallowing just 10 ml of water would ingest an average of 200 oocysts, which is well above a dose capable of causing infection (<100 oocysts) (Chappell et al., 2011). During this outbreak period, a total of six AFR’s were reported.

The earliest date of onset in this outbreak for any case attending the BPP was the 9th of November, 2012. This suggests that the individual (Case 7), who swam regularly at BPP and had an onset of illness on 17th November, probably did not start the outbreak but may have contributed to cryptosporidiosis transmission after this date. It is important that people with cryptosporidiosis exclude themselves from swimming in pools as even when they become asymptomatic they can still shed oocysts (Ajjampur et al., 2010). The WA Department of Health recommends that cryptosporidiosis cases abstain from swimming pools for at least 2 weeks after diarrhoea has stopped. Unfortunately, many cases of cryptosporidiosis are undiagnosed and so are not given advice to exclude themselves from swimming pools. Public swimming pools could review their general signage recommending people avoid swimming if they are symptomatic or recently recovered from any form of gastroenteritis.

The present study highlights the importance of rapid implementation of public health measures to help reduce the spread of infection. It also highlights the need for further research to determine if swimming pools are a common source of community-acquired cryptosporidiosis in Australia and also the need for more effective Accidental Faecal Release Management (AFRM) guidelines and better education of the public to
avoid swimming in pools when they have had infectious gastroenteritis. Currently, nitazoxanide (CRYPTAZ) is approved for treatment of cryptosporidiosis in children and immunocompetent adults in the U.S.A. (Rossignol, 2010), but is not easily available in Australia. However, the role of nitazoxanide to limit duration of shedding of oocysts (Vandenberg et al., 2012), deserves more attention for its use in outbreaks in Australia.
CHAPTER 6

Evidence of Cryptosporidium transmission between cattle and humans in northern New South Wales.

6.1 Introduction

A previous preliminary study in New South Wales (NSW), which examined the species/genotypes and sub-genotypes of Cryptosporidium in 7 human and 15 cattle cases of sporadic cryptosporidiosis in rural western NSW during the period from November 2005 to January 2006, reported that four of the six C. parvum subtypes found in humans were also found in the cattle, indicating that zoonotic transmission may be an important contributor to sporadic human cases of cryptosporidiosis in rural NSW (Ng et al., 2008). Here, we report a more extensive study of human and cattle faecal samples from farms in rural NSW to more accurately elucidate the transmission dynamics of Cryptosporidium in rural populations.

6.2 Material and Methods

6.2.1 Sample selection

In June 2010, invitation letters were sent to dairy farmers in northern NSW to
recruit for the study. The study sample included twelve farms from the upper Hunter Valley, and eight from around Tamworth. Farms were selected that had herd sizes with more than 100 milking cows. A total of 20 herds meeting these criteria were selected using a rectangular transect approach expanding in area until the desired sample size was reached. Between dates 12/7/2010 and 29/8/2010 veterinarians collected approximately 10 rectal swabs from individual calves on each farm and conducted a survey requesting information on farming practices, scouring history, management of calves and treatment regimes.

A public health epidemiologist and environmental health officer visited the fourteen farms in the Hunter Valley within seven days of the veterinarians collecting specimens. They requested information on farm worker demographics including those who live on the farm property with direct or indirect contact with the dairy calves, diarrhoea in the past month, association with travel, child care centres, public swimming pools, potable water source, consumption of raw milk, eating among the animals, contact with the animals and washing hands after animal contact.

**6.2.2 DNA extraction**

A total of 196 faecal samples were collected from calves on 20 farms and 63 faecal samples were collected from humans from 14 of these farms. All cattle and human specimens were sent to Murdoch University, Western Australia (WA) and refrigerated on receipt (4°C). Total DNA was extracted using a QIAmp DNA Stool Kit (Qiagen, Hilden, Germany) as described in Chapter 2 section 2.1. DNA was stored at -20°C until testing was completed.
6.2.3 PCR amplification and sequence analysis

Samples were initially genotyped to species level at the 18S rRNA locus using a two-step nested PCR to amplify a product size of ~540 bp and *C. parvum* isolates were subgenotyped using a two-step nested PCR to amplify a ~832 bp fragment of the *gp60* gene as described in Chapter 2 section 2.2. Amplified DNA fragments from secondary PCR products purified as described in section 2.3 and 2.4, sequencing reaction as described in section 2.5 and nucleotide analyses carried out as described in section 2.6 and 2.7 in Chapter 2.

Representative sequences have been deposited into GenBank under accession numbers JQ362488–JQ362497.

6.2.4 Statistical analysis

Prevalence and 95% confidence intervals were calculated based on the exact binomial method (Ross, 2003). Analysis of risk factors associated with presence of *Cryptosporidium* was limited to humans who submitted faecal specimens and completed the survey (62/63). Diarrhoea, as the primary clinical symptom of cryptosporidiosis, was used in the case definition for cryptosporidiosis and its association with risk factors was analysed. Statistical analyses were performed using SPSS version 17.0 (SPSS inc. Chicago, USA) to investigate associations between the presence of *Cryptosporidium* sp. and the factors surveyed in the questionnaire. Univariable analyses conducted included chi-square test for independence, Fisher’s exact test for statistical significance with p-value cutoff points when p=0.2-0.25 and odds ratio (OR) with 95% confidence intervals.
6.3 Results

6.3.1 Cryptosporidium species and subtypes in calves

The overall prevalence of Cryptosporidium in cattle was 73.5% (144/196) (95% CI: 66.7, 79.5). The prevalence on the different farms ranged from 30% to 100% (Table 6.1). 18S sequences were obtained for 142 of the 144 positives; 85 were C. parvum, 29 were C. bovis and 14 were C. ryanae. Eleven were mixed C. parvum/C. bovis infections and 3 were mixed C. parvum/C. ryanae infections. Sub-typing analysis at the gp60 locus was successful for 84 of the 143 positives. The most common subtype identified was IlaA18G3R1 (n = 57), followed by IlaA19G3R1 (n = 11), IlaA17G2R1 (n = 7), IlaA19G2R1 (n = 6), IlaA16G3R1 (n = 2) and IlaA20G3R1 (n = 1).

6.3.2 Cryptosporidium species and subtypes in humans

A total of 63 faecal samples were collected from humans from 14 of these farms. The overall prevalence of Cryptosporidium in humans was 23.8% (15/63) (95% CI: 14.0, 36.2) (Table 6.1). 18S sequences were obtained for 14 of the 15 positives; 12 were C. parvum and 2 were C. bovis. Of those that were positive for Cryptosporidium, 4/15 humans reported having diarrhoea, all 4 of which were infected with C. parvum. No clinical symptoms were reported from those infected with C. bovis. Sub-typing analysis at the gp60 locus was successful for 7 of the 12 C. parvum positives. Subtype IlaA18G3R1 was identified in 5 humans, IIdA18G2 in one human and a mixed subtype of IlaA18G3R1 and IIdA19G2 in one human.
Table 6.1. Prevalence of *Cryptosporidium* species and subtypes in cattle and farm workers on the 20 farms analysed in the upper Hunter Valley and Tamworth NSW.

<table>
<thead>
<tr>
<th>Farm no</th>
<th>Calf age</th>
<th>No. positive</th>
<th>Cryptosporidium prevalence (95% CI)</th>
<th>Species</th>
<th>gp60 subtype</th>
<th>No. positive</th>
<th>Species</th>
<th>gp60 subtype</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;2 months</td>
<td>9/10</td>
<td>90% (55.5, 99.7)</td>
<td>C. parvum (n = 7)</td>
<td>Iia A17G2R1 (n = 4)</td>
<td>3/5</td>
<td>C. parvum (n = 3)</td>
<td>Iia A18G3R1 (n = 2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C. bovis (n = 2)</td>
<td>Iia A18G3R1 (n = 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C. bovis (n = 2)</td>
<td>Iia A19G2R1 (n = 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C. bovis (n = 2)</td>
<td>Iia A19G3R1 (n = 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>&lt;2 months</td>
<td>10/10</td>
<td>100% (69.2, 100)</td>
<td>C. parvum (n = 10)</td>
<td>Iia A18G3R1 (n = 7)</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C. parvum (n = 7)</td>
<td>Iia A19G2R1 (n = 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>&lt;2 months</td>
<td>9/10</td>
<td>90% (55.5, 99.7)</td>
<td>C. bovis (n = 7)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C. ryanae (n = 2)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>&lt;2 months</td>
<td>10/10</td>
<td>100% (69.2, 100)</td>
<td>C. parvum (n = 2)</td>
<td>Iia A16G3R1 (n = 2)</td>
<td>0/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C. ryanae (n = 6)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mixed C. parvum/C. ryanae (n = 1)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ND (n = 1)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1-2 months</td>
<td>7/10</td>
<td>70% (34.8, 93.3)</td>
<td>C. parvum (n = 3)</td>
<td>Iia A17G2R1 (n = 3)</td>
<td>0/2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C. bovis (n = 3)</td>
<td>Iia A18G3R1 (n = 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mixed C. parvum/C. bovis (n = 1)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>Eggs</td>
<td>% (95% CI)</td>
<td>Species</td>
<td>Genotype(s)</td>
<td>Frequency</td>
<td>PCR Detection</td>
<td>Notes</td>
</tr>
<tr>
<td>---</td>
<td>-----------</td>
<td>------</td>
<td>------------</td>
<td>--------------------------</td>
<td>-------------</td>
<td>------------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>6</td>
<td>&lt;2 months</td>
<td>7/10</td>
<td>70% (34.8, 93.3)</td>
<td><em>C. bovis</em> (n = 2)</td>
<td>Ila A19G2R1 (n = 4)</td>
<td>1/10</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mixed <em>C. parvum/C. bovis</em> (n = 5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>&lt;2 months</td>
<td>7/9</td>
<td>77.8% (40.0, 97.2)</td>
<td><em>C. parvum</em> (n = 7)</td>
<td>Ila A18G3R1 (n = 7)</td>
<td>0/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>&lt;2 months</td>
<td>7/10</td>
<td>70% (34.8, 93.3)</td>
<td><em>C. parvum</em> (n = 7)</td>
<td>Ila A18G3R1 (n = 6)</td>
<td>0/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1-2 months</td>
<td>3/10</td>
<td>30% (6.7, 65.2)</td>
<td><em>C. parvum</em> (n = 2)</td>
<td>-</td>
<td>5/10</td>
<td></td>
<td><em>C. parvum</em> (n = 4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mixed <em>C. parvum/C. ryanae</em> (n = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>&lt;2 months</td>
<td>10/10</td>
<td>100% (69.2, 100)</td>
<td><em>C. parvum</em> (n = 10)</td>
<td>Ila A18G3R1 (n = 10)</td>
<td>0/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>&lt;2 months</td>
<td>5/10</td>
<td>50% (18.7, 81.3)</td>
<td><em>C. parvum</em> (n = 2)</td>
<td>Ila A18G3R1 (n = 3)</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>&lt;2 months</td>
<td>9/10</td>
<td>90% (55.5, 99.7)</td>
<td><em>C. bovis</em> (n = 4)</td>
<td>Ila A18G3R1 (n = 1)</td>
<td>1/6</td>
<td></td>
<td><em>C. parvum</em> (n = 1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. ryanae</em> (n = 4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mixed <em>C. parvum/C. ryanae</em> (n = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>&lt;2 months</td>
<td>9/10</td>
<td>90% (55.5, 99.7)</td>
<td><em>C. parvum</em> (n = 9)</td>
<td>Ila A18G3R1 (n = 6)</td>
<td>1/2</td>
<td><em>C. parvum</em> (n = 1)</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1-2 months</td>
<td>9/10</td>
<td>90% (55.5, 99.7)</td>
<td><em>C. parvum</em> (n = 9)</td>
<td>Ila A19G3R1 (n = 8)</td>
<td>0/2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>&lt;2 months</td>
<td>10/10</td>
<td>100% (69.2, 100)</td>
<td><em>C. parvum</em> (n = 4)</td>
<td>Ila A18G3R1 (n = 9)</td>
<td>2/5</td>
<td><em>C. parvum</em> (n = 2)</td>
<td>Ila A18G3R1 (n = 1)</td>
</tr>
<tr>
<td>No.</td>
<td>Age</td>
<td>C. parvum/ C. bovis</td>
<td>C. parvum</td>
<td>C. bovis</td>
<td>C. ryanae</td>
<td>Other</td>
<td>Total</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>16</td>
<td>&lt;2 months</td>
<td>3/10 30% (6.7, 65.2)</td>
<td>C. parvum (n=2)</td>
<td>IIa A18G3R1 (n=1)</td>
<td>0/3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>&lt;2 months</td>
<td>8/10 80% (44.4, 97.5)</td>
<td>C. parvum (n=2)</td>
<td>IIa A18G3R1 (n=1)</td>
<td>2/8 C. bovis (n=1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>&lt;2 months</td>
<td>5/10 40% (12.2, 73.8)</td>
<td>C. parvum (n=3)</td>
<td>IIa A18G3R1 (n=2)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>&lt;2 months</td>
<td>4/10 40% (12.2, 73.8)</td>
<td>C. parvum (n=3)</td>
<td>IIa A18G3R1 (n=1)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>&lt;2 months</td>
<td>3/7 42.9% (6.7, 65.2)</td>
<td>C. parvum (n=3)</td>
<td>IIa A18G3R1 (n=3)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>144/196 73.5% (66.7, 79.5)</td>
<td>C. parvum (n=3)</td>
<td>IIa A18G3R1 (n=3)</td>
<td>-</td>
<td></td>
<td>15/63</td>
<td></td>
</tr>
</tbody>
</table>
6.3.3 Epidemiological analysis

The survey conducted was completed by 62/63 humans who submitted faecal specimens. The dataset generated from the survey was used for analysing *Cryptosporidium* infections in humans and limited to those who contributed stool samples and were positive for *Cryptosporidium* through screening by PCR (15/62). Humans who were *Cryptosporidium* positive were 8.2 times more likely to have diarrhoea (95% CI: 1.3, 50.5). However, none of those with diarrhoea consulted with a doctor about their symptoms. Univariate analyses of risk factors such as eating close to animals, washing hands after contact with animals, drinking raw milk and contact with calves did not reveal any significant association with the presence of *Cryptosporidium* or having diarrhoea compared with individuals without *Cryptosporidium* and/or without diarrhoea (Table 6.2). All participants reported rainwater as the main potable water source and none of the participants reported travelling overseas. Visiting child care centres and public swimming pools could not be determined as risk factors due to low number of respondents in the questionnaire survey.

Table 6.2: Associations between risk factors examined and the presence of *Cryptosporidium* in humans or report of humans with diarrhoea.

<table>
<thead>
<tr>
<th>Risk Factors</th>
<th><em>Cryptosporidium</em> Infection</th>
<th>Diarrhoea</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>Eating among the animals</td>
<td>0.87 (0.25, 3.02)</td>
<td>0.54</td>
</tr>
<tr>
<td>Washing hands after contact with animals</td>
<td>1.52 (0.42, 5.55)</td>
<td>0.38</td>
</tr>
<tr>
<td>Drinking raw milk</td>
<td>0.45 (0.07, 3.16)</td>
<td>0.38</td>
</tr>
<tr>
<td>Contact with calves</td>
<td>0.31 (0.18, 5.30)</td>
<td>0.43</td>
</tr>
</tbody>
</table>
6.4 Discussion

In the present study, the overall prevalence of Cryptosporidium in cattle was 73.5% (95% CI: 66.7, 79.5). This is much higher than reports from previous studies, one of which reported a prevalence of 17.6% (95% CI: 13.0, 22.9) in pre-weaned calves from 5 farms in WA and a prevalence of 39.1% (95% CI: 23.7, 41.1) in a NSW farm (Ng et al., 2011). Another study reported that the total prevalence of Cryptosporidium in calves from 84 dairy and dairy beef properties across Australia was 58.5% (95% CI: 54.4, 62.4) (Izzo et al., 2011). In the present study, all the calves were < 2 months old whereas in the previous study in NSW and WA, the age of the calves ranged from 1 week to 4 months old (Ng et al., 2011). A study in the US reported that the highest prevalence of infection occurs in calves <8 weeks of age (45.8%), followed by post-weaned calves (3-12 months of age) (18.5%) and heifers (12-24 months of age) (2.2%) (Santin et al., 2008). In the present study, the prevalence on the different farms ranged from 30% to 100%. Most of the calves on these properties were diarrheic, however, insufficient testing was conducted to conclude that Cryptosporidium was the sole cause.

Three species of Cryptosporidium were detected in the calves; C. parvum accounted for the majority (59.4%), followed by C. bovis (20.3%) and C. ryanae (9.8%). Mixed C. parvum/C. bovis and C. parvum/C. ryanae infections accounted for 9.8% of infections. Pig genotype II and C. andersoni were not identified in the cattle in the present study. The prevalence of C. parvum was higher than a recent previous study in NSW, which reported a prevalence of 42% for C. parvum amongst typed species in pre-weaned cattle (Ng et al., 2011). There also appears to be geographical differences in the prevalence of C. parvum in calves in Australia with lower C. parvum rates in WA compared to Eastern states. For example, a previous study reported that the prevalence of C. parvum in pre-weaned cattle in 5
WA farms was 17.6%, whereas another study in Victoria, identified *C. parvum* in 46.3% (124/268) of calves on pasture-based dairy farms in three regions sampled (Nolan et al., 2009).

The overall prevalence of *Cryptosporidium* in humans in this study was 23.8% (15/63) (95% CI: 14.0, 36.2) which is higher than the overall prevalence reported in dairy farm workers in India at 11.8% (Khan et al., 2010). Different methods of *Cryptosporidium* detection however, were used where in this study, PCR-based molecular tools were used compared with microscopy and ELISA, which were used to detect for the presence of *Cryptosporidium* in dairy farm workers in India (Khan et al., 2010). PCR based molecular tools however, have been shown to be more sensitive to microscopy or ELISA when detecting *Cryptosporidium* in faecal samples, which may have contributed to the difference in prevalences (Morgan et al., 1998; Nair et al., 2008). Of the 14 that were sequenced at the 18S gene locus, 12 were *C. parvum* and 2 were *C. bovis*, with the latter identified in a 3 year-old child and a 23 year-old adult, from separate farms. *Cryptosporidium bovis* was first described in 2005 and is morphologically indistinguishable from *C. parvum* (Fayer et al., 2005). It has a narrow host range and has previously been described in cattle and sheep. To date there has only been one report of *C. bovis* in a dairy farm worker in India, where the infection was asymptomatic (Khan et al., 2010). In the present study, infection of *C. bovis* in both humans was asymptomatic. Both individuals had reported drinking raw milk and had regular contact with calves. One of the individuals reported not regularly washing their hands after contact with calves. Whether the two humans involved were actually infected or passing oocysts as a result of mechanical transmission is unknown but highlights the potential for transmission between cattle and humans in rural areas.

Reports of cryptosporidiosis outbreaks have identified direct or indirect contact with animals, washing hands after handling animals, and drinking raw (unpasteurised) milk as risk
factors for Cryptosporidium transmission to humans (Harper et al., 2002; Ashbolt et al., 2003; Kiang et al., 2006). In the present study however, we could not infer any statistically significant associations between the risk factors examined with the presence of Cryptosporidium or reports of diarrhoea in humans on these farms, due to the small sample size. Further studies with a larger sample are required to provide stronger evidence associating these risk factors with Cryptosporidium transmission in rural populations.

Subtype analysis identified that C. parvum IlaA18G3R1 was identified in 69% of the 80 cattle isolates subtyped. This was also the most common subtype identified in six of the seven human isolates successfully subtyped. IlaA18G3R1 is a common subtype in both humans and cattle worldwide and has been reported in both calves and humans in NSW (Ng et al., 2008; Waldron et al., 2009b) and in WA (O’Brien et al., 2008; Ng et al., 2010a, 2010b; Ng et al., 2011), humans and cattle in Victoria (Jex et al., 2007; Nolan et al., 2009) and humans in South Australia (Jex et al., 2008). A recent study, which included sporadic cryptosporidiosis in the Hunter Valley human population between January 2008 to December 2010, identified seven C. hominis and five C. parvum isolates, with IlaA18G3R1 the most common C. parvum subtype (Waldron et al., 2011a). Although the epidemiological analysis was inconclusive, this finding suggests that zoonotic transmission may be occurring. However, a much larger study is required to confirm this.

Interestingly, subtypes IIdA18G2 and IIdA19G2 were identified from two humans who lived on the same farm but these subtypes were not found in any of the cattle isolates genotyped. The IId subtype family is much less commonly identified in cattle, sheep, goats and humans in European studies and dairy calves in Egypt (Amer et al., 2010; Xiao, 2010), and has never been found in humans or cattle in the United States and Canada (Xiao, 2010). IId subtypes (IIdA15G1, IIdA24G1) have been reported previously in humans but not in cattle in NSW (Waldron et al., 2009). Both the IIdA18G2 and IIdA19G2 subtypes have not
been reported previously in humans or cattle and represent novel subtypes. The finding of these subtypes in humans in the present study and the finding of *C. hominis* subtypes in sporadic human cases from a recent study in the Hunter Valley (which included the present study area) (Waldron et al., 2011a), suggests that human to human transmission is also common in NSW.

In conclusion, the present study identified *C. parvum* subtypes IIdA18G2 and IIdA19G2 for the first time in humans, showing the genetic diversity of *C. parvum* subtypes infecting humans in NSW. The concurrence of *C. parvum* subtypes and the cattle-specific *C. bovis* in humans and calves provides evidence of zoonotic transmission and the possible association of infected calves and human infection with *Cryptosporidium*. Studies involving more extensive sampling of both calves and farm workers and collection of more extensive epidemiological data are needed for a better understanding of the sources of *Cryptosporidium* infections in humans in rural areas.
CHAPTER 7

Development of an untargeted metabolomics method for the analysis of human faecal samples using *Cryptosporidium*-infected samples

7.1 Introduction

As nothing is known about the metabolomics of *Cryptosporidium*, in the present study, a method for total metabolite extraction and gas chromatography-mass spectrometry (GC-MS) analysis from faecal samples of humans infected with *Cryptosporidium* was developed. Cryptosporidiosis ranges in seriousness from mild to severe with signs and symptoms depending on the site of infection and nutritional and immune status of the host (Chalmers and Davies, 2010). The loss of absorptive epithelium due to the parasite infection results in a reduction of villous surface area, crypt cell hyperplasia and an inflammatory response of the underlying mucosa. Impaired intestinal epithelial permeability is often associated with *Cryptosporidium* infection, and with loss of absorptive epithelial cells, disaccharidase activities are disrupted (Argenzio et al., 1990; Griffiths et al., 1994; Farthing, 2000; Rosenthal and Glew, 2009). All of these factors would be expected to contribute to biochemical changes occurring in *Cryptosporidium*-positive individuals.

Clinical diagnosis is generally undertaken by testing stools for the presence of oocysts by microscopy (Buret, 2009). Microbiological examination of a single sample from a patient cannot exclude *Cryptosporidium*, since oocysts may fall below detectable numbers even during symptomatic episodes (Jokipii and Jokipii, 1986). Faecal samples submitted also vary in amount and consistency. It is therefore important to optimise a faecal extraction method so
that it is sufficiently robust to account for these variations prior to GC-MS analysis.

7.2 Materials and methods

7.2.1 Chemicals

All chemicals were purchased from Sigma Aldrich (Australia) at a purity of > 99% unless otherwise noted. Methanol and n-hexane (> 95%) were purchased from LabScan (Australia).

7.2.2 Samples

A total of 16 human faecal samples sourced from a local diagnostic laboratory were used. Of these, 8 were confirmed microscopy-positive Cryptosporidium samples and 8, which were microscopy negative for Cryptosporidium, were used as negative controls. Samples were confirmed as Cryptosporidium positive and negative using PCR and sequence analysis of the 18S rRNA gene as described in sections 2.2, 2.4, 2.5 and 2.6 in Chapter 2. Samples were stored at 4°C prior to analysis.

7.2.3 Sample preparation

The following parameters of sample preparation were optimised: amount of faecal material, faecal consistency (water content) and extraction solvent. All tests were performed in triplicate on pooled Cryptosporidium positive (n=8) and negative samples (n=8) unless noted. Trimethylsilylated ribitol (10 µg/ml) was used as an internal standard in all analyses.

To determine the optimum minimal amount of faecal material required for extraction,
a range of amounts and ratios were tested. 250, 500 and 1000 mg of faeces were diluted in methanol to a final concentration of 200 mg/ml. Samples containing 500 mg faeces were also diluted to a concentration of 100 mg/ml. Solutions were vortexed and centrifuged at 2,530 x g at 10°C for 15 min. A 50 µl aliquot of each extract, which equated to 10 mg of faeces for samples at the 200 mg/ml concentration and 5 mg for the 100 mg/ml samples, was dried in an Eppendorf rotary vacuum concentrator and stored at -80°C until derivatisation.

Three extraction solvents were studied: methanol (room temperature), methanol (-40°C) and chloroform: methanol: water (CMW) (1:3:3). 250 mg of faeces was added to a concentration of 200 mg faeces/ml. For the CMW extraction, 250 mg of faeces was considered equivalent to 250 µl of water. Extraction solutions were subsequently handled as described above.

It was apparent from initial experiments that the variable water content of the faecal samples had a substantial impact on the analytical outcomes. To account for this variability in consistency, samples were freeze-dried prior to extraction, and extraction was based on dry weight. Analysis of freeze-drying on 4 samples with varied consistency identified that 200 mg wet weight corresponded to approximately 40 mg dry weight.

### 7.2.4 Sample preparation for optimised method

The final optimised method was as follows: 250 mg (wet weight) of faecal material was freeze-dried overnight and the dry weight measured. Methanol was added to each sample to give a final concentration of 40 mg dry faeces/ml. The resulting solution was vortexed and centrifuged at 2,530 x g at 10°C for 15 min. A 50 µl aliquot of this extract, which corresponded to 2 mg of dry faecal weight or 10 mg of weight faecal weight, was dried in a rotary vacuum concentrator and stored at -80°C until derivatisation.
7.2.5 Derivatisation

Methoximation was achieved by the addition of 40 µl of methoxyamine-HCl (20 mg/ml in pyridine), followed by incubation for 90 minutes at 30°C with shaking at 1,200 rpm. Trimethylsilyl derivatisation was carried out with the addition of 20 µl N-trimethylsilyl-N-methyl trifluoroacetamide (MSTFA) and incubation at 37°C for 90 minutes. An n-alkane mix (C₁₀-C₃₂) was added to each sample following derivatisation to allow the calculation of Kovat’s Indices.

7.2.6 Analysis

Derivatised samples were analysed by splitless injection (1 µl) onto an Agilent 6890 gas chromatograph coupled to an Agilent 5973N mass spectrometer. A 30 m Varian Factor Four VF-5ms capillary column (ID = 0.25 mm, Df = 0.25 µm) with 10 m Ezi-Guard was used for separation. The injector was set at 230°C, transfer line at 300°C and the ion source at 230°C. Helium carrier gas flow rate was retention-time locked to elute mannitol-TMS at 30.6 minutes. The initial oven temperature was set to 70°C, increasing at 1°C/min for 5 minutes, and subsequently increasing at a rate of 5.63°C/min to 300°C. The mass spectrometer was set to scan the range m/z 45 to 600 (Abbiss et al., 2012).

7.2.7 Data analysis

Data were deconvoluted, normalised and matched to an in-house library using AnalyzerPro v2.2.0.7 (SpectralWorks, UK). Data were subsequently exported to The Unscrambler v 9.8 (CAMO, Norway) for multivariate analysis.
7.3 Results

7.3.1 Optimisation of faecal weight

GC-MS analysis of samples extracted from 250, 500 and 1000 mg identified a mean of 328 (± 1), 271 (± 7) and 224 (± 3) compounds respectively, with an overall significant difference in the mean number of compounds identified in the three weight groups (p < 0.05). Analysis of 250 mg extracts (at 200 mg/ml) showed that more compounds were detected and a better resolution of lower level analytes was achieved compared to extraction using 500 mg and 1000 mg of faeces (at 200 mg/ml) (Figure 7.1A, C and D). This was most likely due to overload of the GC-MS with the higher extracts. This is supported by comparison of the extracts of 500 mg faeces at two different final concentrations. Extracts at a final concentration of 100 mg/ml yielded an average of 346 (± 7) compounds, which was significantly higher than 500 mg extracts at 200 mg/ml, which yielded an average of 271 (± 7) compounds (p < 0.05). Mass spectra of the 100 mg/ml extract from 500 mg faeces were similar to that of the sample extracted from 250 mg faeces to a final concentration of 200 mg/ml (Figure 7.1), which was expected based on the amount of material analysed. As there was no significant difference between the average number of compounds identified from 250 mg faeces (at 200 mg/ml) and 500 mg faeces (at 100 mg/ml) (p = 0.05), 250 mg faeces was extracted to a final concentration of 200 mg/ml in all subsequent experiments.

7.3.2 Extraction solvents

GC-MS analysis of 50 µL faecal extracts was compared for three different solvents: methanol (room temperature), cold methanol (-40°C) and a mixture of chloroform: methanol: water (CMW) (1:3:3). Methanol at room temperature yielded a mean of 356 (± 30)
compounds, which was the highest average yield compared to samples extracted in -40°C methanol or CMW, which yielded averages of 352 (± 5) and 327 (± 6) compounds respectively. There was no significant difference in metabolite yield between the extraction solvents based on the number of compounds detected (p > 0.05).

7.3.3 Reproducibility of metabolite recovery

GC-MS analysis of 16 human faecal samples using the optimised extraction conditions detected a total of 135 metabolites belonging to classes of amino acids, carbohydrates, organic acids, alcohols, amines, amino ketones, nucleosides, bile acids, fatty acids and indoles in both Cryptosporidium positive and negative samples. In order to measure method reproducibility, the relative standard deviation (RSD) of the internal standard, trimethylsilylated ribitol, was calculated, giving a mean RSD of 9.14%.

7.3.4 Metabolites identified in Cryptosporidium positive and Cryptosporidium negative faecal samples

Multivariate analysis was applied in order to study the variation between the Cryptosporidium positive and negative faecal samples. Principal component analysis (PCA) of the GC-MS data, normalised to the internal standard, showed a pattern differentiating between the two sample groups. As a result, the score plots generated revealed distinct clustering patterns where the negative samples grouped closely together and the positive samples showed variability within the group (Figure 7.2). Principal component (PC) 1 explained 30% of the variance while PC-2 explained 23%. X-loadings were used to determine the 30 compounds that contributed most to the variance between the two groups. These metabolites included amino acids (n = 10), carbohydrates (n = 5), organic acids (n =
alcohol (n = 4) and metabolites from various other classes (n = 6) (Table 7.1). Higher metabolite abundance was generally observed in the Cryptosporidium positive samples when compared to the negative samples. Statistical analysis revealed that 9 of these metabolites showed significant differences in level between the two sample groups (denoted by * in Table 7.1).

**Table 7.1:** Metabolites identified in human faecal samples that contributed to the variance observed between Cryptosporidium positive and negative faecal samples. Metabolites that showed statistical significance in abundance, where p < 0.05, are denoted by *.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Carbohydrates</th>
<th>Organic Acids</th>
<th>Alcohols</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-Alanine *</td>
<td>Trehalose</td>
<td>Gluconic acid</td>
<td>Mannitol</td>
<td>Hexdecanoic acid</td>
</tr>
<tr>
<td>L-Phenylalanine *</td>
<td>N-Acetyl mannosamine</td>
<td>Glutaric acid</td>
<td>Xylitol</td>
<td>Cholesterol</td>
</tr>
<tr>
<td>L-Aspartic acid</td>
<td>Xylose</td>
<td>Benzoic acid</td>
<td>Thrietol*</td>
<td>Phosphoric acid</td>
</tr>
<tr>
<td>L-Isoleucine*</td>
<td>Sucrose</td>
<td>Succinic acid*</td>
<td>Glycerol</td>
<td>Putrescine</td>
</tr>
<tr>
<td>L-Glutamic acid</td>
<td>Glucuronic acid</td>
<td>Lactic acid</td>
<td></td>
<td>Indole-3-acetic acid</td>
</tr>
<tr>
<td>L-Valine*</td>
<td></td>
<td></td>
<td></td>
<td>Urea*</td>
</tr>
<tr>
<td>L-Serine*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-Acetylglutamic acid*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyroglutamic acid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GABA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 7.1: Mass spectra from gas chromatography-mass spectrometry (GC-MS) analysis of pooled faecal specimens; (A) extraction of 250 mg of samples with a final concentration of 200 mg/ml, (B) 500 mg samples with a final concentration of 100 mg/ml, (C) 500 mg samples with a final concentration of 200 mg/ml and (D) 1000 mg samples with a final concentration of 200 mg/ml.
Differences in carbohydrates levels between the two groups were not statistically significant ($p > 0.05$), although the level of the disaccharide, trehalose, was higher in negative samples, while sucrose was higher in the positive samples. For the monosaccharides, N-acetyl mannosamine and xylose were higher in the negative samples, and glucuronic acid was higher in the positive samples.

For the organic acids, succinic acid, which is a tricarboxylic acid (TCA) cycle intermediate, was significantly higher in the positive samples compared to the negative samples ($p < 0.05$). No significant difference ($p > 0.05$) in abundance of sugar alcohols was detected between positive and negative samples except for threitol, which was significantly higher in the positive samples ($p < 0.05$). In addition, significantly higher levels of urea were
present in positive samples (p < 0.05).

7.4 Discussion

In the present study, a faecal metabolite extraction method for GC-MS analysis from small amounts of faeces was developed. When designing extraction protocols, a key compromise is between the need to avoid overloading the capacity of the instrument and the desire to profile metabolites at low levels, which would require extraction of greater amounts of sample. It was expected that extraction of greater amounts of material would allow better profiling of metabolites at low abundances. However, due to overloading of the instrumentation, the opposite was observed. Analysis of 250 mg extracts (at 200 mg/ml) showed that more compounds were detected and produced a better resolution of lower level analytes compared to extraction using 500 mg and 1000 mg of faeces (at 200 mg/ml). Contact between faecal material and extraction solvent may have been reduced when larger amounts of material was used, thus resulting in decreased extraction efficiency, especially that of the lower level analytes. Therefore, in all subsequent experiments, 250 mg faeces were extracted to a final concentration of 200 mg/ml.

Comparison of three different solvents: methanol (room temperature), cold methanol (-40°C) and a mixture of chloroform: methanol: water (CMW) (1:3:3) revealed no significant difference in metabolite yield between the extraction solvents or number of compounds detected (p > 0.05). To date, no single solvent or combination of solvents has been demonstrated as optimal for all quenching and extraction methods. The preferred solvent should be able to rapidly inactivate metabolism and extract the maximum number of metabolites without any chemical or physical alteration of the metabolites (Villas-Bôas et al., 2005). Methanol is one of the most frequently used solvents for metabolite quenching and
extraction, and has been used for extraction of microbial cells and animal tissues at low temperatures (cold, -40°C), room temperature or coupled with chloroform to allow for simultaneous extraction of polar and non-polar metabolites (Prasad Maharjan and Ferenci, 2003; Villas-Bôas et al., 2005). It has also been reported to produce better metabolite extraction results compared to other solvents (Prasad Maharjan and Ferenci, 2003; Villas-Bôas et al., 2005). As there were no significant differences between the three different extraction solvents tested, and as it is widely used in metabolite extractions, methanol at room temperature was chosen as the preferred solvent to reduce handling time and avoid the risk of chloroform toxicity (Villas-Bôas et al., 2005). The optimised extraction method (250 mg faeces extracted in methanol to a final concentration of 200 mg/ml) showed good reproducibility with <10% variance of the internal standard recorded across all samples (median RSD of the internal standard = 9.14%).

Using this optimised method, higher metabolite abundance was generally observed in Cryptosporidium positive versus negative samples. These differences in abundance may be a result of Cryptosporidium’s mechanism of invasion, where through penetration and proliferation in the luminal enterocytes, it causes alterations to the physiological structure of the small intestine and permeability of nutrients from the lumen (Chalmers and Davies, 2010). This may have also resulted, or contributed to the changes in amino acid, nitrogen and energy metabolism as observed in this study.

Six amino acids showed significant differences (p < 0.05) in metabolite levels between the two sample groups. The higher abundance of phenylalanine, valine, isoleucine and serine in the positive samples may be due to increased intestinal epithelium permeability in Cryptosporidium infected individuals, as well as intracellular protein catabolism (Topouchian et al., 2001; Topouchian et al., 2003; Topouchian et al., 2005). This not only leads to a change in amino acid metabolism, but also to alterations in nitrogen metabolism.
and the urea cycle, as observed with significantly higher N-acetylglutamic acid and urea present in positive samples ($p < 0.05$). Synthesis of N-acetylglutamic acid is stimulated by the presence of proteins and amino acids, and hence the higher abundance observed may correspond to the higher amino acid abundance that was detected in the positive samples. This may have resulted in increased activation of key enzymes, thus up-regulating the urea cycle.

The differences observed in carbohydrate levels between the two groups, although not statistically significant, are consistent with expected changes in the brush border structures of the small intestine as a result of the infection. Trehalase, the disaccharidase responsible for hydrolysis of trehalose, has its greatest activity in the small intestine (Richards et al., 2002). Unlike any other disaccharidase, trehalase is tightly bound to the external surface of microvilli enterocytes (Richards et al., 2002). Continuous disruption of the brush borders caused by the mechanism of Cryptosporidium infection may have resulted in increased trehalase activity resulting in increased trehalose hydrolysis and hence the lower abundance observed in the positive samples.

The higher levels of succinic acid in Cryptosporidium positive samples may indicate a shift in energy metabolism. High levels of succinic acid without oxygen inhibit the tricarboxylic acid cycle and cause a shift towards anaerobic glycolysis in the intestine (Vejchapipat et al., 2000). However, further studies are required to better understand these metabolic changes in the small intestine during Cryptosporidium infection.

The sugar alcohol threitol was significantly higher in the positive samples ($p < 0.05$). Threitol is the main end product of xylose metabolism which occurs in the liver (Pitkänen, 1977; Craig and Ehrenpreis, 1999), which suggests that, in the positive samples, higher amounts of xylose may have permeated through the intestinal epithelium, which corresponds to the lower abundance of xylose detected in the positive samples compared to the negative
samples (p > 0.05). This was also observed for mannitol which was lower in positive samples, compared to the negative samples, although the difference was not significant (p > 0.05). Studies of the effects of *Cryptosporidium* infection on the permeability of the intestinal epithelium have reported both increased and decreased carbohydrate permeability (Griffiths et al., 1994; Goodgame et al., 1995; Capet et al., 1999; Klein et al., 2008). Although results from the present preliminary study suggest that there may be an increase in permeability, more studies need to be conducted and a more accurate method, as suggested by Gookin et al. (2002), would be to measure permeability with respect to mucosal surface area.

In conclusion, we have developed a reproducible faecal extraction method for untargeted GC-MS analysis that allows for the analysis of varying faecal amounts and consistencies as a result of infection by the gastrointestinal parasite *Cryptosporidium*. The assay was validated using clinically diagnosed samples and the faecal metabolite profiles generated were able to distinguish between infection positive and negative individuals. The differences in metabolites detected between these two groups indicated changes in amino acid, nitrogen and carbohydrate metabolism and demonstrated the effects of *Cryptosporidium* infection and its association to changes in intestinal epithelial permeability, however further research is required to confirm this. The faecal extraction method developed in the present study can be used in the extraction of faecal samples for further studies of *Cryptosporidium* infection, as well as other gastrointestinal diseases.
CHAPTER 8

Metabolomic profiling of faecal extracts from Cryptosporidium parvum infection in experimental mouse models

8.1 Introduction

The metabolomics study on Cryptosporidium in Chapter 7, developed a faecal metabolite extraction method for untargeted gas chromatography-mass spectrometry (GC-MS) analysis using Cryptosporidium positive and negative human faecal samples. In that study, faecal metabolite profiles of cryptosporidiosis positive patients could be differentiated from cryptosporidiosis negative patients, suggesting that metabolic homeostasis and intestinal permeability were affected as a result of the infection. However, as the metabolome is sensitive to external perturbations, it was difficult to examine the extent to which the changes to the human faecal metabolites observed were attributed to infection with Cryptosporidium or to other factors such as age, diet and immune susceptibility etc. which may differ between patients. Hence the aim of the present study was to carry out a more controlled metabolomics analysis of faecal metabolite profiles using experimentally infected mice, to better characterise metabolic changes associated with Cryptosporidium infection.

8.2 Materials and Methods

8.2.1 Experimental infection in mice

The C. parvum (S26) isolate used in the present study to infect mice was originally
obtained from a naturally infected calf from the Institute of Parasitology, University of Zurich, Switzerland and was passaged through mice at Murdoch University using methods described by Meloni and Thompson (1996). Briefly, 1-day old neonatal Swiss mice (n = 10) obtained from the Animal Resource Centre (Perth, Western Australia) were inoculated individually with approximately 100,000 oocysts. As all the faecal pellets from this litter were required to produce consistent metabolomics data, a second litter (n=10) was inoculated as described above and the mice were confirmed as positive for Cryptosporidium by microscopy. A third litter (n=10) was not inoculated and used as an uninfected control for the experiment. The mice were then returned to their dams and euthanised 8 days post-inoculation by CO₂ exposure. Faecal samples from each mouse were collected directly from the rectum into 1.5 ml microcentrifuge tubes, placed in liquid nitrogen then on dry ice to arrest metabolism. Samples were then stored at -80°C until analysis. Mice infection and metabolomics experiments were repeated twice for reproducibility. All research on animals was conducted under Murdoch University animal ethics permit no: R2351/10.

8.2.2 Chemicals

All chemicals were purchased from Sigma Aldrich (Australia) at a purity of > 99% unless otherwise noted. Methanol and n-hexane (> 95%) were purchased from LabScan (Australia).

8.2.3 Metabolite extraction and sample preparation

Metabolite extraction was carried out as described in Chapter 7 with minor modifications, to allow for the small amount of faeces produced by each mouse. Briefly, faecal samples from each mouse were freeze-dried overnight to remove all moisture and their
subsequent dry weight measured. Methanol containing the internal standard ribitol (10 µg/ml) was added to each sample to a final concentration of 50 mg dry faeces/ml, vortexed until completely homogeneous and then centrifuged at 2,530 x g at 10°C for 15 minutes. From each sample, a 50 µl aliquot of the extract was dried in a rotary vacuum concentrator and stored at -80°C until derivatisation. Prior to analysis, methoximation and trimethylsilyl derivatisation was carried out as described in Chapter 7.

8.2.4 GC-MS instrumentation and analysis

Gas chromatography (GC)-mass spectrometry (MS) analysis was performed on 1 µL of each derivatised sample in splitless mode on an Agilent 6890 gas chromatograph coupled to an Agilent 5973N mass spectrometer as described in Chapter 7. An Agilent FactorFour VF-5ms capillary column (ID = 0.25 mm, Df = 0.25 µm) measuring 30 m with 10 m EZ-Guard was used for separation. Helium was used as a carrier gas with retention time locked to elute mannitol-TMS at 30.6 minutes. Initial oven temperature was set to 70°C, with a temperature ramp set to increase at a rate of 1°C/min for 5 minutes, and subsequently 5.63°C/min to 300°C, holding for 10 minutes. The injector was set at 230°C, the transfer line to the MS at 300°C and the MS ion source at 230°C. The mass spectrometer was set to scan the range m/z 45 to 600 at 1 scan per second.

8.2.5 Data processing and statistical analysis

The GC-MS data files were deconvoluted and normalised to the internal standard, ribitol. Compounds were matched against an in-house library or reference compounds from the National Institute of Standards and Technology (NIST) (http://www.nist.gov) using AnalyzerPro v2.7.0 (SpectralWorks, UK). Data was then exported to The Unscrambler® X
CAMO, Norway) where principal component analysis (PCA) of the data was performed. Univariate statistical analysis of the metabolites identified was carried out and an unpaired T-test performed to compare metabolite intensities between the infected and non-infected groups with p ≤ 0.05 considered statistically significant.

8.3 Results

8.3.1 Analysis of the metabolites identified in faecal samples from C. parvum infected mice and uninfected mice

GC-MS analysis of faecal samples from 8-day old C. parvum infected and uninfected mice (n =20) detected approximately 220 compounds. A total of 101 of the compounds detected were matched against the in-house library and reference database (Table 8.1), of which, 87 were identified to belong to various classes, including amino acids, carbohydrates, organic acids, amines, nucleosides and fatty acids, while the remaining 14 were matched to known ‘unknowns’, compounds that are routinely detected but have not been positively identified. These ‘unknowns’ were named according to their retention index followed by its base peak. Normalised to the internal standard, trimethylsilylated ribitol, the relative standard deviation (RSD) was 18% of the mean.

Principal component analysis (PCA) of the MS data normalised to the internal standard showed distinct clustering of the C. parvum infected mice and the uninfected control mice. The score plot generated showed that the infected mice samples clustered closely together, whereas samples from uninfected mice showed variation within the group (Figure 8.1). Principal component 1 (PC1) explained 55% of the variance, with PC2 explaining 16% of the variance observed. Based on the X-loadings, a total of 40 compounds, or metabolites that contributed most to the variance between the two groups were identified. These
compounds, or metabolites, included amino acids \( (n = 17) \), carbohydrates \( (n = 8) \), lipids \( (n = 7) \), organic acids \( (n = 3) \), and metabolites from various other compound classes \( (n = 5) \) (Figure 8.2). Statistical analysis revealed that 33 of the 40 metabolites showed significant differences in levels \( (p > 0.05) \) between the group of infected and uninfected mice (Figure 8.2). Levels of metabolites observed in the infected mice were generally lower, when compared to the uninfected mice, with a mean ratio of 1:3.9.

**Figure 8.1:** Principal component analysis of mice faecal metabolite profiles. Score plots showed distinct clustering patterns of metabolite profiles of *C. parvum* infected mice and the uninfected control mice, differentiating the two groups. *Cryptosporidium* infected mouse samples are represented by ● and uninfected control mouse samples are represented by ■.
Table 8.1: Mean of normalised peak area of metabolites (compounds) identified in faecal samples of mice infected with *C. parvum* and the uninfected control mice compared to faecal metabolite profile of humans infected and not infected with *Cryptosporidium* from Chapter 7).

<table>
<thead>
<tr>
<th>Metabolites</th>
<th>Mice Uninfected</th>
<th>Mice Infected</th>
<th>Human Uninfected</th>
<th>Human Infected</th>
</tr>
</thead>
<tbody>
<tr>
<td>9,12 Octadecadienoic acid</td>
<td>1.02857</td>
<td>0.20211</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9-Octadecenamide (Z)</td>
<td>0.00000</td>
<td>0.00266</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9-Octadecenenitrile (Z)</td>
<td>0.00000</td>
<td>0.00361</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>0.09813</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabinose</td>
<td>0.06795</td>
<td>0.01505</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>0.03258</td>
<td>0.04954</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Azelaic acid</td>
<td>0.00345</td>
<td>0.01257</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzoic acid</td>
<td>0.01703</td>
<td>0.02933</td>
<td>0.07116</td>
<td>0.03384</td>
</tr>
<tr>
<td>beta-Alanine</td>
<td>0.00028</td>
<td>0.00116</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cellobose</td>
<td>0.00000</td>
<td>0.00028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0.83152</td>
<td>0.75479</td>
<td>1.29944</td>
<td>0.85916</td>
</tr>
<tr>
<td>D(-)-Ribose</td>
<td>0.23501</td>
<td>0.11145</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-Pinitol</td>
<td>0.00194</td>
<td>0.00281</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fructose</td>
<td>0.22061</td>
<td>0.11492</td>
<td>0.00279</td>
<td>2.76036</td>
</tr>
<tr>
<td>Fumaric acid</td>
<td>0.11235</td>
<td>0.06616</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GABA</td>
<td>0.01766</td>
<td>0.00036</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galactose</td>
<td>0.47900</td>
<td>0.08688</td>
<td>0.75856</td>
<td>1.29266</td>
</tr>
<tr>
<td>Galactose</td>
<td>0.62402</td>
<td>0.10529</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g-Hydroxyglutaric acid</td>
<td>0.01600</td>
<td>0.00056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucopyranose</td>
<td>0.00000</td>
<td>0.00071</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose</td>
<td>0.01213</td>
<td>0.00189</td>
<td>0.45676</td>
<td>0.07358</td>
</tr>
<tr>
<td>Glutaric acid</td>
<td>0.06168</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucuronic acid</td>
<td>0.00020</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyceric acid</td>
<td>0.03987</td>
<td>0.00074</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycerol</td>
<td>0.45593</td>
<td>0.04332</td>
<td>0.03898</td>
<td>4.36919</td>
</tr>
<tr>
<td>Glycerol 3-Phosphate</td>
<td>0.03289</td>
<td>0.03980</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexdecanoic acid</td>
<td>1.88391</td>
<td>3.21622</td>
<td>7.12470</td>
<td>4.80394</td>
</tr>
<tr>
<td>Hypoxanthine</td>
<td>0.01141</td>
<td>0.01946</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactose</td>
<td>0.00000</td>
<td>0.00028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Alanine</td>
<td>1.19349</td>
<td>0.11553</td>
<td>3.07196</td>
<td>4.41755</td>
</tr>
<tr>
<td>L-Asparagine</td>
<td>0.01213</td>
<td>0.00189</td>
<td>0.45676</td>
<td>0.07358</td>
</tr>
<tr>
<td>L-Aspartic acid</td>
<td>0.00037</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Glutamic acid</td>
<td>0.05682</td>
<td>0.00032</td>
<td>0.20789</td>
<td>0.81873</td>
</tr>
<tr>
<td>Substance</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
<td>Value 4</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>L-Glutamic acid</td>
<td>0.22357</td>
<td>0.12511</td>
<td>0.23186</td>
<td>1.18186</td>
</tr>
<tr>
<td>L-Glycine</td>
<td>0.19238</td>
<td>0.07508</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Isoleucine</td>
<td>0.16561</td>
<td>0.00056</td>
<td>0.87763</td>
<td>1.75428</td>
</tr>
<tr>
<td>L-Lysine</td>
<td>0.34528</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Methionine</td>
<td>0.03741</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Phenylalanine</td>
<td>0.16680</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.69200</td>
</tr>
<tr>
<td>L-Proline</td>
<td>0.15992</td>
<td>0.00147</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Proline</td>
<td>0.01467</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Serine</td>
<td>0.15947</td>
<td>0.00979</td>
<td>0.18630</td>
<td>2.00585</td>
</tr>
<tr>
<td>L-Serine</td>
<td>0.19654</td>
<td>0.00252</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Threonine</td>
<td>1.11461</td>
<td>0.00230</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Threonine</td>
<td>0.34061</td>
<td>0.00591</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Tyrosine</td>
<td>0.11709</td>
<td>0.00564</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Valine</td>
<td>0.79384</td>
<td>0.01195</td>
<td>1.65323</td>
<td>3.62800</td>
</tr>
<tr>
<td>L-Valine</td>
<td>0.26370</td>
<td>0.00141</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lysine</td>
<td>0.00000</td>
<td>0.00023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malic acid</td>
<td>0.14626</td>
<td>0.00486</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maltose</td>
<td>0.00000</td>
<td>0.00156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mannose</td>
<td>0.16015</td>
<td>0.00283</td>
<td></td>
<td></td>
</tr>
<tr>
<td>myo-Inositol</td>
<td>1.24033</td>
<td>1.03768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>myo-inositol</td>
<td>0.04207</td>
<td>0.01482</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-Acetyl mannosamine</td>
<td>0.80440</td>
<td>0.01258</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-Acetyl mannosamine</td>
<td>0.38392</td>
<td>0.00470</td>
<td>0.04638</td>
<td>0.00000</td>
</tr>
<tr>
<td>N-Acetylglutamic acid</td>
<td>0.63399</td>
<td>0.08983</td>
<td>0.15029</td>
<td>0.83735</td>
</tr>
<tr>
<td>N-Acetyleneuraminic acid</td>
<td>0.11003</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonadecanoic acid</td>
<td>0.00863</td>
<td>0.02710</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Octadecanoic acid</td>
<td>1.60698</td>
<td>1.73383</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oleic acid</td>
<td>1.19663</td>
<td>0.57099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ornithine</td>
<td>0.08115</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ornithine/Arginine</td>
<td>0.04499</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pantothenic acid</td>
<td>0.01983</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pentanoic acid</td>
<td>0.11422</td>
<td>0.27086</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphoric acid</td>
<td>0.19807</td>
<td>0.41548</td>
<td>0.09436</td>
<td>0.00000</td>
</tr>
<tr>
<td>Phosphoric acid</td>
<td>0.00973</td>
<td>0.01107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-Hydroxybenzoic acid</td>
<td>0.00264</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Putrescine</td>
<td>0.06820</td>
<td>0.06379</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyroglutamic acid</td>
<td>0.13045</td>
<td>0.12352</td>
<td>0.40310</td>
<td>0.85106</td>
</tr>
<tr>
<td>Siloxane</td>
<td>0.03971</td>
<td>0.06942</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorbitol</td>
<td>0.00219</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.68246</td>
</tr>
<tr>
<td>Succinic acid</td>
<td>0.04901</td>
<td>0.12651</td>
<td>0.20056</td>
<td>2.40465</td>
</tr>
<tr>
<td>Sucrose</td>
<td>0.00753</td>
<td>0.00387</td>
<td>0.00000</td>
<td>0.87304</td>
</tr>
<tr>
<td>Tetradecanoic acid</td>
<td>0.00132</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tetradecanoic acid</td>
<td>1.53994</td>
<td>0.33194</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymine</td>
<td>0.02137</td>
<td>0.00012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metabolite</td>
<td>Infected Mean</td>
<td>Uninfected Mean</td>
<td>50% CI Low</td>
<td>50% CI High</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Tryptamine</td>
<td>0.00480</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyramine</td>
<td>0.04344</td>
<td>0.00177</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uracil</td>
<td>0.33669</td>
<td>0.04564</td>
<td>0.50018</td>
<td>0.77988</td>
</tr>
<tr>
<td>Urea</td>
<td>0.01857</td>
<td>0.01261</td>
<td>0.00000</td>
<td>5.49976</td>
</tr>
<tr>
<td>Uric acid</td>
<td>0.26681</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xylitol</td>
<td>0.00457</td>
<td>0.00000</td>
<td>0.00181</td>
<td>0.00606</td>
</tr>
<tr>
<td>Xylose</td>
<td>0.03934</td>
<td>0.00432</td>
<td>0.02144</td>
<td>0.00000</td>
</tr>
<tr>
<td>Unknown amide_1210_228</td>
<td>0.02602</td>
<td>0.03632</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_1270_174</td>
<td>0.43404</td>
<td>0.30429</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_1339_170</td>
<td>0.03800</td>
<td>0.07740</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_1410_234</td>
<td>0.00000</td>
<td>0.00256</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown amide_1548_115</td>
<td>0.03724</td>
<td>0.03600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_1616_142</td>
<td>0.03440</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_1673_116</td>
<td>0.00057</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_1708_173</td>
<td>0.00136</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_1871_319</td>
<td>0.00000</td>
<td>0.00062</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_1943_318</td>
<td>0.00038</td>
<td>0.01056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_1951_217</td>
<td>0.00065</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_2021_204</td>
<td>0.00338</td>
<td>0.00079</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_2306_204</td>
<td>0.00662</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown_2333_217</td>
<td>0.00047</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 8.3.2 Metabolites identified in faecal samples of *C. parvum* infected mice and uninfected mice

Levels of the 17 amino acids that contributed most to the variance between the mice groups were lower in the infected mice when compared to the uninfected mice (Figure 8.2). In addition, L-lysine, L-phenylalanine, glutamine, ornithine and N-acetylneuraminic acid were not detected in the infected mice samples. Mean levels of amino acids between infected and uninfected mice showed significant differences (p < 0.05), with the exception of L-tyrosine, L-phenylalanine, glutamine and pyroglutamic acid, which although detected in lower amounts in the infected mice compared to the uninfected mice, were not statistically different (p > 0.05).
The 8 carbohydrates which contributed to the variance included the monosaccharides, galactose, mannose, fructose, and glucose; the sugar alcohols, glycerol and myo-inositol and an amino sugar, N-acetylmannosamine (Figure 8.2). In the infected mice, lower levels of these carbohydrates were detected compared to the uninfected mice with significant differences (p < 0.05) in the levels detected, except for the levels of fructose, myo-inositol and glucose, which were not significantly different (p > 0.05).

For the lipids, significant differences (p < 0.05) were observed between the infected and uninfected mice groups with higher levels of hexadecanoic acid (ratio 1.69:1) and nonadecanoic acid (ratio 3.12:1), and lower levels of 9, 12-octadecadienoic acid (1:5.1), tetradecanoic acid (ratio 1:6.2) and cholesterol (1:1.43) detected in the infected mice group. Other lipid metabolites which contributed to the variance between the mice groups included pentanoic acid and oleic acid. Although higher levels of pentanoic acid (2:1) and lower levels of oleic acid (1:2.5) were detected in the infected mice compared to the uninfected mice, the differences were not significant (p > 0.05).

The 3 organic acids that contributed most to the variance between infected and uninfected mice groups showed significant differences in the levels detected (p < 0.05). Lower levels of malic acid were detected in the infected mice (ratio of 1:30), whereas for succinic acid and benzoic acid, higher levels were detected in the infected mice with a ratio of 2.56:1 and 1.72:1, respectively.

Levels of the pyrimidine uracil were significantly lower (p < 0.05) in the infected mice with a ratio of 1:7.38. There was no statistical difference (p > 0.05) in the levels of hypoxanthine, detected, which was 1.5 times higher in the infected mice compared to the uninfected mice. Comparison of the two unknown compounds that contributed to variance between the infected and uninfected mice groups against the NIST library revealed chemical structure matches to the amide class of compounds. The first unknown amide (with a
retention index of 1210 and base peak of 228), was detected in higher levels in the infected mice (p < 0.05), whereas there was no statistical difference between the levels of the second unknown amide (with a retention index of 1548 and base peak of 115) in the two mice groups.

**Figure 8.2:** Faecal metabolites that contributed to the variance between infected and uninfected mice. Metabolites contributing to the variance between faecal metabolite profiles of mice infected with *C. parvum* and the uninfected control mice were compared to metabolites which contributed to the variance of faecal metabolite profiles of humans infected and not infected with *Cryptosporidium* from Chapter 7). Heat map was generated using the mean of normalised peak area of the metabolite. All mice faecal metabolites showed statistically significant (p < 0.05) contributions to the variance except for those denoted by ^, where p > 0.05.

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Control Mice</th>
<th>Infected Mice</th>
<th>Infection -ve</th>
<th>Infection +ve</th>
<th>Key</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amino acid</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GABA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Glutamine^</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Alanine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Aspartic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Isoleucine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Lysine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Phenylalanine^</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Proline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Serine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Threonine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Tyrosine^</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>L-Valine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>N-Acetylglutamic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>N-Acetylenorvaline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ornithine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ornithine/Arginine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Pyroglutamic acid^</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><strong>Carbohydrate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D(-)-Ribose</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Fructose^</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Galactose</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Glucose</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Glycolar</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Mannose</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>myo-Inositol^</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>N-Acetylmannosamine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>9,12-Octodecadienoic acid [2,2]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Cholesterol</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Hexadecanoic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Nonadecanoic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Oleic acid^</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Pentanoic acid^</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Tetracosanoic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><strong>Lipid</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzoic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Malic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Succinic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Hypoxanthine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Phosphoric acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Uracil</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Unknown amide 1210_228</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Unknown amide 1548_115</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Key: 0.00 - 0.05, 0.05 - 0.10, 0.10 - 0.15, 0.15 - 0.20, 0.20 - 0.25, 0.25 - 0.30, 0.30 - 0.35, 0.35 - 0.40, 0.40 - 0.45, 0.45 - 0.50, 0.50 - 0.60, 0.60 - 0.70, 0.70 - 0.80, 0.80 - 0.90, 1.00 - 1.25, 1.25 - 1.50, 1.50 - 1.75, 1.75 - 2.00, 2.00 - 2.25, 2.25 - 2.50, 2.50 - 2.75, 2.75 - 3.00, 3.00 - 3.25, 3.25 - 3.50, 3.50 - 3.75, 3.75 - 4.00, 4.00 - 4.25, 4.25 - 4.50, 4.50 - 4.75, 4.75 - 5.00, 5.00 - 7.00, > 7.00.
8.4 Discussion

The mechanism of invasion of Cryptosporidium, from its initial attachment to enterocytes to subsequent development of an intracellular yet extracytoplasmic parasitophorous vacuole, results in changes to the physiological structure of the small intestinal luminal brush border, impairing mucosal absorption and permeability (Gookin et al., 2002; Leav et al., 2003). With impaired intestinal absorption and permeability due to the infection, studies conducted have shown a reduction in the uptake of nutrients such as glucose, amino acids, lipids and vitamin A (Capet et al., 1999; Topouchian et al., 2003; Topouchian et al., 2005). It can be hypothesised that, with a reduction in uptake, these nutrients would be excreted, and that a higher abundance of metabolites would be detected in the faeces of Cryptosporidium infected mice. Results from the present study however, showed that this was not the case as analyses of the faecal metabolite profiles from both Cryptosporidium infected and uninfected mice showed that lower faecal metabolite abundances were generally recorded from the infected mice compared to that of the uninfected mice.

Genome sequencing and biochemical data has revealed that Cryptosporidium is highly reliant on its host/environment for nutrients, as it is missing key metabolic pathways and lacks the ability for de novo synthesis of nucleosides, fatty acids and amino acids (Abrahamsen et al., 2004; Xu et al., 2004; Zhu, 2008; Vanee et al., 2010; Xu et al., 2010). Hence, to supplement its streamlined metabolism, Cryptosporidium scavenges nutrients from the host intestinal lumen which, in effect, may have resulted in the lower faecal metabolite abundances detected in the infected mice in the present study. Although metabolites identified were synonymous between the two groups, metabolites that were found to contribute to the variance between the Cryptosporidium infected and uninfected mice included precursors to amino acids, carbohydrates, fatty acids, as well as derivatives of
purine and pyrimidine metabolism, hypoxanthine and uracil.

A total of 17 different amino acids detected in the mice faecal extracts contributed to the variance between the infected and uninfected mice. Abundance of these amino acids was significantly lower in the infected mice compared to the uninfected mice (p < 0.05), except for L-tyrosine, L-phenylalanine, glutamine and pyroglutamic acid, which showed no significant difference in mean abundance (p > 0.05) (Figure 8.2). The lower abundance of all 17 amino acids identified from the infected mice compared to uninfected may be a result of amino acid scavenging by *Cryptosporidium* from the intestinal lumen, as it has at least 11 amino acid transporters to enable scavenging from the host to supplement for its lack of ability to synthesise amino acids (Abrahamsen et al., 2004). Although *C. parvum* shows capability in interconverting a number of amino acids, for example serine and glycine interconversion, conversion of aspartate to asparagine and reverse conversion of glutamate to glutamine (Rider Jr and Zhu, 2010), these amino acids were detected in lower abundance in the infected mice indicating that such interconversion may not have taken place.

The highly-streamlined metabolism of *Cryptosporidium*, which lacks genes encoding for mitochondria and tricarboxylic acid cycle enzymes, suggests that the parasite’s main energy pathway is glycolysis, where these mono- and poly-saccharides are scavenged and used as fuel for energy production (Abrahamsen et al., 2004; Rider Jr and Zhu, 2010). This may have resulted in lower abundances of metabolites and intermediaries involved in energy metabolism being detected in infected mice when compared to uninfected mice. These metabolites include the complex sugar mannose, the monosaccharide glucose, glycerol and malic acid, which suggest an increase in both glycolysis and gluconeogenesis. The higher abundance of succinic acid, an intermediate of the tricarboxylic acid cycle (TCA), detected in both *Cryptosporidium* infected mice and humans, suggests a decrease in TCA activity during an infection (discussed in Chapter 7). This could be caused by an increase in demand for the
TCA intermediates, oxaloacetate and α-ketoglutarate, to act as precursors for amino acid synthesis. This higher abundance of succinic acid may also result in a shift towards anaerobic glycolysis in the intestine as in the absence of oxygen, high levels of succinic acid may also inhibit the tricarboxylic acid cycle (Vejchapipat et al., 2000).

Despite only a narrow difference in ratio of benzoic acid abundance detected in Cryptosporidium infected mice and uninfected mice, the level of benzoic acid detected from the infected mice was significantly higher compared to the uninfected mice (p < 0.05) and contributed to the variance between these two groups. This was also observed in Cryptosporidium infected humans where higher abundance of benzoic acid was detected compared to the uninfected humans (Chapter 7). In mammals, benzoic acid is absorbed in the small intestine where it binds a conjugate with glycine, and is excreted as hippuric acid in the urine (Bridges et al., 1970; Krähenbühl et al., 1997). With Cryptosporidium’s dependency on the host for amino acids, glycine, which was detected in lower abundance in infected mice, may not have been available to conjugate with benzoic acid to supplement the metabolic demand from the infection, resulting in a decrease in acylation of benzoic acid. This could be confirmed by examining urine samples from infected mice for hippuric acid abundance, which should be lower in infected mice compared to uninfected mice.

The faecal metabolite profile of lipids detected from both mouse groups in the present study consisted mainly of saturated fatty acids of 14 to 18 carbons in length, which plays important roles in energy metabolism and cell membrane biosynthesis. Studies into the lipid metabolism of C. parvum revealed that the parasite lacks the ability to synthesise fatty acids de novo, but is capable of fatty acid chain elongation via the enzyme elongase CpLCE1, which is involved in elongating carbon C\textsubscript{12:0} to C\textsubscript{16:0} saturated fatty acid substrates to C\textsubscript{18:0} products (Fritzler et al., 2007; Rider Jr and Zhu, 2010). In the study conducted by Fritzler et al. (2007), it was observed that when the C\textsubscript{14:0} fatty acid was used as a substrate, the longest
The elongation product synthesised was C\textsubscript{16:0} fatty acid, which then did not serve as a substrate for further elongation. This may explain for the lower abundance of tetradecanoic acid (C\textsubscript{14:0}), which could have been used as a substrate of elongase, resulting in the higher abundance of hexadecanoic acid (C\textsubscript{16:0}) but lower abundance of octadecanoic acid (C\textsubscript{18:0}), as there was no further elongation of C\textsubscript{16:0} fatty acid substrates synthesised from C\textsubscript{14:0} in the infected mice.

The lower levels of cholesterol detected in the infected mice compared to uninfected mice was similar to that observed in the human faecal metabolite profile where levels of cholesterol were lower in those infected with \textit{Cryptosporidium} compared to those uninfected (Table 8.1) (Chapter 7). Important to the development and life cycle of the parasite, a recent study showed that \textit{Cryptosporidium} scavenges cholesterol from the intestinal lumen and enterocytes of its host (Ehrenman et al., 2013), which may explain for the lower levels of cholesterol detected in both infected mice and humans.

Of the metabolites that contributed most to the variance between \textit{Cryptosporidium} infected and uninfected humans in the previous study in Chapter 7 and mice in this present study, the metabolites L-alanine, L-aspartic acid, L-valine, L-isoleucine, L-serine, L-phenylalanine, N-acetylglutamic acid, glutamic acid, pyroglutamic acid, δ-aminobutyric acid (GABA), N-acetyl mannosamine, glycerol, cholesterol, hexadecanoic acid, benzoic acid, succinic acid and phosphoric acid were identified in both faecal metabolite profiles of humans and mice. In the previous study however, higher levels of metabolites were generally detected in \textit{Cryptosporidium} positive patients, differing from the results of the present study, where lower metabolite levels were generally detected in faecal samples from \textit{Cryptosporidium} infected mice (Table 8.1). Differences in metabolite profiles between different host types have been previously reported by Saric et al. (2007). In that study, a comparison of faecal metabolite profiles from mice, rats and humans showed that the levels of metabolites differed between the species, presumably as a result of different endogenous
and exogenous perturbations and differences in the gut microbiome between species (Saric et al., 2007). Despite the differences in faecal metabolite profiles between Cryptosporidium infected humans and mice, metabolomic analysis in both studies was still able to clearly differentiate between infected and uninfected hosts, as well as provide information on the metabolic activity of the parasite during the infection based on faecal metabolite profiles.

Metabolites that contributed to the differences between infected and uninfected groups in both the present study and in Chapter 7 are key nutrients scavenged by Cryptosporidium to supplement its streamlined metabolic pathway, which is supported by results from recent Cryptosporidium transcriptomic studies. These studies have shown that genes that annotated for transporters and metabolic enzymes for key nutrients (sugars, nucleotides, amino acids and lipids) were highly expressed at the time of infection (Xu et al., 2010; Mauzy et al., 2012). Although no metabolite unique to Cryptosporidium infection was identified in the present study, the abundance of metabolites detected in the faeces of mice infected with C. parvum compared to uninfected mice reflects not only on the effects of the infection on the host environment where absorption capabilities in the small intestine have been impaired, but also the fate of metabolites that were not absorbed.

Metabolomics may be useful for the diagnosis of Cryptosporidium infections, as it allows for detection based on differences as a result of physiological changes caused by the infection rather than detecting oocysts in faeces by microscopy or PCR, where sensitivity is limited by both the numbers of oocysts present and the intermittent shedding of oocysts. However, whether the profiles observed are unique to Cryptosporidium infection or are seen in infections with other gastrointestinal pathogens needs to be ruled out along with extensive analysis of the metabolites profiled before this method could be developed as a diagnostic tool. Metabolite profiling of Cryptosporidium infection should also be expanded to include profiles from other biofluids, such as blood and urine, as well as tissue samples, which will
provide a more comprehensive understanding of Cryptosporidium host-parasite interaction and its impact on host metabolism. This could potentially result in the identification of unique metabolites as biological markers for better diagnosis and open new avenues for the development of anti-Cryptosporidium therapies.
CHAPTER 9

General Discussion

9.1 Introduction

The focus of this thesis was the characterisation of sporadic and outbreaks cases of cryptosporidiosis in humans, zoonotic transmission from animals to humans and the first metabolomics analysis of Cryptosporidium-infected human and mice faecal samples. Analysis included examining the differences in cryptosporidiosis epidemiology among Aboriginal and non-Aboriginal people in Western Australia, analysis of three cryptosporidiosis outbreaks in Western Australia (2003, 2007 and 2011) and an outbreak in Broome, WA in 2012 and molecular genotyping of faecal samples from cattle and their farmers in New South Wales. In addition, a reproducible faecal extraction method for untargeted GC-MS analysis that allows for the analysis of varying faecal amounts and consistencies as a result of infection by Cryptosporidium, was developed, and used to analyse Cryptosporidium positive and negative human faecal samples and experimentally infected mice samples.

The findings presented in this thesis has provided more evidence of the complex epidemiology of cryptosporidiosis in Australia; it confirms that specific Cryptosporidium subtypes are associated with outbreaks and showed that although anthropogenic transmission cycles dominate in Australia, that zoonotic transmission can occur, particularly when humans are in close contact with livestock. In addition, metabolomics analysis identified distinct profiles between Cryptosporidium infected and uninfected faecal samples and provided new insights into the metabolome and biochemistry of this parasite. Interestingly, higher
metabolite profiles were observed in *Cryptosporidium* infected human faecal samples compared to un-infected samples, but in mice, lower faecal metabolite abundances were detected in infected versus uninfected faecal samples. The difference in metabolite abundance between humans and mice is an important finding and has been reported in previous studies, thus raising questions about the appropriateness of animal models when studying the biochemistry of *Cryptosporidium*.

9.2 The epidemiology of sporadic cryptosporidiosis cases

Cryptosporidiosis, is one of the most common gastrointestinal diseases in Australia, responsible for 44,836 notified cases between 2001-2006 (NNDSS, 2016). However, the actual number of cryptosporidiosis cases is much higher as the majority of gastroenteritis cases due to cryptosporidiosis and other enteric diseases are never identified because: (1) it is estimated that less than 10% of individuals with gastroenteritis visit their local doctor, and of these, less than 10% have a faecal specimen collected (Garthright et al. 1988; Hellard and Fairley, 1997; Majowicz et al. 2005; Tam et al. 2012), (2) not all individuals presenting with gastroenteritis will have faecal samples tested for microorganisms (Majowicz et al. 2005; Tam et al. 2012) and (3) identification of *Cryptosporidium* and other pathogens via microscopy lacks specificity and sensitivity (Checkley et al. 2015). For example, a national survey of gastroenteritis in Australia in 2002, suggested a ratio of about 500 community cases to one notified (Hall et al., 2006). A more recent study reported that in 2010 alone, cryptosporidiosis was estimated to be responsible for 195,495 cases of acute gastroenteritis (AGE) and 333 disability-adjusted life years (DALYs) in Australia (Gibney et al., 2014).

Lack of notification of cryptosporidiosis is even more pronounced in Aboriginal communities due to difficulty accessing health services. For example, in 2008, the National
Aboriginal and Torres Strait Islander Social Survey (NATSISS) found that about 26% of Indigenous people aged 15 and over living in non-remote areas had difficulty accessing health services (AIHW, 2011); in contrast, only 2.6% of the general population has difficulty (ABS, 2012). The main reasons cited were long waiting times, services not being available when needed, difficulties with transport and health-care costs. Less commonly reported reasons included lack of engagement, fear of discrimination and poor treatment arising from previous experiences, and the lack of culturally appropriate services (AIHW, 2011). This under-reporting of cryptosporidiosis is an important factor that should be taken into account when calculating baselines rates of cryptosporidiosis in Australia.

A striking finding of Chapter 3 was the distinct differences in epidemiology between Aboriginal and non-Aboriginal patients in Western Australia, with notification rates among Aboriginals up to 50 times higher compared to non-Aboriginals and the highest prevalence of cryptosporidiosis in Aboriginal children in remote areas such as the Pilbara and Kimberley. Aboriginals were also found to be 3.9 times more at risk at being infected with the *C. hominis* IdA15G1 subtype, while non-Aboriginals were 7.2 times more at risk of being infected with the IbA10G2 subtype, reflecting the unique endemicity of cryptosporidiosis in Aboriginal communities.

It is well known that children in remote Australian Aboriginal communities experience exceptionally high rates of gastroenteritis and other common childhood infections including otitis media and skin and respiratory infections (Li et al., 2007; Clucas et al., 2008). These infections have serious health consequences, leading to impaired growth and development and consequent poor educational outcomes (Bailie et al., 2010). It is therefore particularly important for health organisations serving remote Aboriginal Australians to determine and monitor the prevalence of gastroenteritis and other important diseases in the population in order to ensure that health resources are being directed for greatest benefit.
Given the lack of effective treatment options for cryptosporidiosis, understanding transmission routes is central to control. Sources and transmission of Cryptosporidium can be difficult to establish because of the ubiquitous nature and the multiple routes of transmission of this parasite. The high prevalence of C. hominis subtypes in both Aboriginal and non-Aboriginal population indicates that anthroponotic transmission dominates. However, it is also important to note that although C. hominis was originally thought to be exclusively anthroponotic, it has been reported in cattle in New Zealand (Abeywardena et al., 2012), Scotland (Smith et al., 2005), India (Feng et al., 2007) and Korea (Park et al., 2006) and therefore cattle may be reservoirs for human infection. Well-designed case control studies, with detailed collection of data on water and food sources and animal contact, are essential in identifying risk factors involved in the acquisition of Cryptosporidium infection in both Aboriginal and non-Aboriginal communities.

Key risk factors reported in previous case-control studies, the majority of which were reviewed by Hunter and Thompson (2005) and Yoder and Beach (2010), are the ingestion of contaminated drinking or recreational water, contact with infected persons or animals, travel to areas where the disease is endemic, and contact with children < 6 years old (especially but not exclusively with diarrhea). It has been suggested that repeated exposures to small numbers of Cryptosporidium oocysts may provide protective immunity, as higher levels of anti-Cryptosporidium antibodies were associated with reduced rates of diarrhea (Frost et al., 2005). Frost et al. (2005) suggested that the lower levels of illness in people with increased levels of antibodies may be due to enhanced immunity from repeated exposures to oocysts in drinking water. Similarly, consumption of raw vegetables and ice cream have also shown negative association with Cryptosporidium infection, possibly through the conference of protective immunity through repeated exposures to low levels of oocysts (Robertson et al., 2002; Hunter et al., 2010).
In Australia however, data indicates that recreational exposure and not drinking water is a major source of community cryptosporidiosis cases (Dale et al. 2010). For example, a case control study in Melbourne and Adelaide, which assessed risk factors for sporadic cryptosporidiosis in Australia, concluded that consumption of plain tap water was not found to be associated with disease and that swimming in public pools and contact with a person with diarrhoea were the most important risk factors (Robertson et al., 2002). Another double-blinded, randomised, controlled trial carried out in Melbourne also reported no evidence of waterborne disease in Melbourne (Hellard et al., 2001). In Western Australia, the situation is more complicated because despite the fact that across Eastern Australia, recreational access is not allowed on direct supply sources, in Western Australia, limited recreational access to drinking water catchments has been allowed, although only in the outer catchment. Significantly, a study in WA has shown that recreational water catchments, which allowed swimming and camping, had a higher prevalence of Cryptosporidium, with the human-associated Cryptosporidium identified in the majority of samples (Loganthan et al., 2012). Non-recreational catchments had a lower prevalence and all the samples genotyped were C. parvum. This suggests that recreational access to drinking water catchments may be a source of cryptosporidiosis in WA (Loganthan et al., 2012).

9.3 The epidemiology of cryptosporidiosis in outbreaks

Diagnosis of cryptosporidiosis outbreaks is particularly problematic as in addition to the reasons discussed in section 9.2 for the lack of detection of cryptosporidiosis, current outbreak detection methods lack sensitivity, specificity and timeliness i.e. there is a considerable lapse of time before an outbreak is recognised (Hellard et al., 2000; Dale et al., 2010). For example, in the largest Cryptosporidium outbreak which occurred in Milwaukee
in 1993, where an estimated 403,000 individuals contracted cryptosporidiosis via contaminated drinking water (MacKenzie et al., 1995), only a small number of cases were investigated prior to the identification of the outbreak and confirmation of its cause (MacKenzie et al., 1995). Another outbreak of cryptosporidiosis in California involving more than 250 persons, whose common exposure was a waterpark, went unreported to the local public health department for weeks, despite ~75% of the waterpark’s 26 employees developing a gastrointestinal illness (Wheeler et al., 2007). In fact, most incident cases of cryptosporidiosis are not confirmed to be due to *Cryptosporidium* (Chalmers, 2008; Smith et al. 2010b; Tam et al. 2012; Briggs et al. 2014), and many apparently sporadic cases may be part of small outbreaks, where the remainder of the cases go undiagnosed.

In all of the outbreaks typed in this thesis (Chapters 4 and 5), the *C. hominis* IbA10G2 subtype was the most common subtype identified. Overall, the Ib subtype family is the major subtype responsible for waterborne and foodborne outbreaks of cryptosporidiosis in many countries. Subtype IbA10G2 has been found in ~50% of *C. hominis*–associated outbreaks in the United States, including the massive outbreak in Milwaukee, Wisconsin, USA, in 1993 (Zhou et al., 2003; Xiao, 2010). Close to a decade later, this subtype was still the most common *Cryptosporidium* parasite detected in raw Milwaukee wastewater (Zhou et al., 2003). It is the only subtype identified in cryptosporidiosis outbreaks by *C. hominis* in countries in Europe (Li et al., 2013) and was responsible for an outbreak in NSW (Waldron et al., 2011b). In a longitudinal birth-cohort study of cryptosporidiosis in a peri-urban shantytown in Lima, Peru, the IbA10G2 subtype was more virulent than other *C. hominis* subtypes (Cama et al., 2008).

The *gp60* gene encodes glycoproteins gp15 and gp45, which are implicated in attachment to and invasion of host cells (Strong et al., 2000). Because attachment of sporozoites to epithelial cells and invasion of the host cell membrane are crucial steps in the
pathogenesis of cryptosporidiosis, these 2 glycoproteins are presumed to be surface-associated virulence determinants that may be under host immune selection, which might explain the extensive polymorphism in the *gp60* gene (Leav et al., 2002).

*Cryptosporidium* has a sexual phase in its lifecycle, during which sexual recombination can occur between genetically distinct strains (Tanriverdi et al., 2007; Li et al., 2013; Takumi et al., 2015). While recombination has been demonstrated for *C. parvum* (Mallon et al., 2003), *C. hominis* was previously thought to have a clonal population structure with genetic recombination rare or nonexistent (Li et al., 2013). However, more recent analysis using a much larger number of genetic markers (*n* = 32) has demonstrated that while *C. hominis* has essentially a clonal population structure, limited recombination does occur, mostly at the *gp60* locus or loci around *gp60* (Li et al., 2013). It has been suggested that recombination within subtype IbA10G2 around the *gp60* locus may be involved in pathogenicity (Li et al., 2013), as the common occurrence of its parental subtypes and biologic fitness of the recombinant subtype with the IbA10G2 trait have probably facilitated genetic exchange and spread of this virulent subtype (Li et al., 2013). Other factors that facilitate the rapid spread of *C. hominis* IbA10G2 subtype during outbreaks include its higher oocyst shedding intensity compared to *C. parvum*. For example, in the swimming pool associated outbreak in NSW, patients infected with *C. hominis*, IbA10G2 subtype, were shedding higher numbers of oocysts (10^5 to 10^7 oocysts/g) than those infected with *C. parvum* (10^3 to 10^4 oocysts/g) (Waldron et al. 2011b). Based on the dominance, persistence, distribution, and virulence of the IbA10G2 subtype in human populations and its association with waterborne outbreaks, this subtype should be considered a significant threat to global human health (Waldron et al. 2011b).
9.4 Zoonotic transmission of cryptosporidiosis

Epidemiological and microbiological evidence for the zoonotic transmission of cryptosporidiosis comes from outbreaks involving veterinary students and researchers who had contact with infected young calves and involving children visiting farms where the same strain type was identified in animals and people (Bouzid et al., 2013). Evidence supporting zoonotic transmission, based on gp60 sequence data, is much more abundant for *C. parvum* than for *C. hominis* (Xiao, 2010; Jex and Gasser, 2010).

Of the zoonotic sources of infection, cattle are recognised as a major contributor, because *C. parvum* subtypes infecting humans have also been isolated from cattle (Xiao and Fayer, 2008; Xiao, 2010). In Chapter 6, the potential for zoonotic transmission between cattle and their farmers was examined in NSW. The *C. parvum* IIaA18G3R1 subtype was identified in 69% of the 80 cattle isolates subtyped. This was also the most common subtype identified in six of the seven human isolates successfully subtyped, providing supportive evidence for zoonotic transmission. The IIa subtype family is the most common subtype found in cattle worldwide (Xiao, 2010; Abeywardena et al., 2015). Intriguingly, although the *C. parvum* IIaA18G3R1 subtype has been reported with high prevalence in Australia, it is rarely reported elsewhere in the world (Jex and Gasser, 2010) suggest unique transmission patterns in Australia. Why this genotype should be so common in Australia and not elsewhere is not clear, but may indicate a predilection for Australian cattle.

In Chapter 6, further support for zoonotic transmission came from the identification of *C. bovis* (a largely cattle and sheep specific species) in two humans; in a 3 year-old child and a 23 year-old adult, from separate farms. Both individuals had reported drinking raw milk and had regular contact with calves and one of the individuals reported not regularly washing their hands after contact with calves. Further molecular studies should focus on understanding the epidemiology of zoonotic infections in rural regions in Australia and in
developing countries, where animals live in close proximity to humans.

9.5 *Cryptosporidium* metabolomics

Metabolomics, which is the study of the metabolome comprising intracellular and extracellular metabolites that are consumed and produced as a result of biological activity, is still in its infancy for *Cryptosporidium*. A major advance in our understanding of the biochemistry of *Cryptosporidium* has arisen from the sequencing of *Cryptosporidium* genomes (Abrahamsen et al., 2004; Xu et al., 2004; Hadfield et al., 2015; Isaza et al., 2015; Ifeonu et al., 2016) and the development of a genome-scale metabolic model of *C. hominis* comprising of 213 gene-associated enzymes, involved in 540 reactions among the major metabolic pathways, which provides a link between the genotype and the phenotype of the organism (Vanee et al. 2010). However, these approaches are limited by the fact that a majority of genes remain functionally uncharacterized and metabolomic approaches are increasingly being used to highlight differences, as well as common features, in the metabolism of different parasite species and developmental stages that were not anticipated from genomic or transcriptomic studies (Kloehn et al., 2016). Validation of metabolite profiles is complicated by measurement accuracy, selectivity, linearity, reproducibility, robustness, and limits of detection. The statistical challenges include analysis, interpretation, and description of the vast amount of data generated. Despite these drawbacks, metabolomics provides great opportunities and the potential to understand and manage cryptosporidiosis and other protozoan diseases.

In Chapters 7 and 8, the first metabolomics analysis of *Cryptosporidium* was conducted on *Cryptosporidium*–infected humans and mice faecal samples versus un-infected faecal samples using an untargeted approach. Untargeted metabolomics, also called discovery
metabolomics, involves the comparison of the metabolome between the control and test groups, to identify differences between their metabolite profiles which may be relevant to specific biological conditions. In the workflow of untargeted metabolomics, analytical reproducibility is critical and the optimised faecal extraction method developed in Chapter 7 showed good reproducibility with <10% variance of the internal standard recorded across all samples. Metabolomics analysis in Chapters 7 and 8, characterised metabolic changes associated with Cryptosporidium infection and differences in metabolic consumption between different hosts, which further informs our understanding of the metabolic requirements of this parasite.

The differences in metabolite profiles observed can be as a result of variation in metabolite consumption and production in Cryptosporidium, which may be influenced by external factors that the parasite is exposed to, such as pH, temperature, other organisms and external stresses. As the gut microbiota have been shown to be essential in maintaining the intestinal epithelial barrier function (van den Elsen et al. 2017), infection with Cryptosporidium, regardless of the Cryptosporidium isolate used in the infection, causes physiological damage to the intestinal lining, and likely alters the gut microbiota (Ras et al., 2015). This in turn is likely to result in changes in metabolite requirements in the host and possibly, that of the parasite. Hence, apart from the differences in human and mice gut microbiota, changes in the host gut microbiota as a result of Cryptosporidium infection, may also be one of the factors influencing the distinct differences observed between faecal samples of Cryptosporidium infected individuals and uninfected individuals.

Another study has demonstrated a very practical use for metabolomics analysis and was able to differentiate between viable and non-viable samples in aquatic environments (Beale et al., 2013). A number of key metabolites including aromatic and non-aromatic amino acids, carbohydrates, fatty acids, and alcohol type compounds explained the
differences between the viable and non-viable oocysts. This is important, as critical information needed to make public health decisions regarding Cryptosporidium includes oocyst viability (NHMRC, 2011). Thus this latter study and the data generated from Chapters 7 and 8, show that metabolomics is a potential pathway for the development of rapid detection methods for Cryptosporidium that utilise specific metabolic biomarkers.

9.6 Future studies

Despite the considerable progress in the last 20 years on molecular epidemiology of Cryptosporidium, significant research gaps remain. Current molecular characterisation methods for Cryptosporidium generally rely on PCR followed by Sanger sequencing, with the 18S rRNA, actin and gp60 being the most commonly characterised loci (Xiao, 2010). Most of the commonly adopted molecular protocols involve nested-PCR amplifications to generate sufficient material for subsequent Sanger sequencing (Xiao, 2010). A limitation of this approach, however, is that rare species and/or mixed infections are generally not detected due to low relative abundance (Grinberg et al., 2013). This leads to a possible underestimation of the prevalence of mixed infections and/or zoonotic Cryptosporidium species. Nested PCR approaches also have an inherent risk of contamination and can exhibit strong PCR bias and/or stochastic variation (Park and Crowley, 2010). High throughput amplicon-based Next Generation Sequencing (NGS) has been successfully applied to Cryptosporidium (Paparini et al., 2015) and the superior depth of coverage offered by NGS and the high throughput capacity will facilitate the large scale molecular epidemiological analyses that are required to better understand mixed infections and the transmission dynamics of Cryptosporidium.
Studies on the transmission of *Cryptosporidium* in humans and domesticated animals are also hampered by the lack of suitable subtyping tools for *Cryptosporidium* species that are genetically distant from *C. parvum* and *C. hominis*. The discriminatory power of MLST tools, which have greatly contributed to the understanding of *Cryptosporidium* population genetics, geographic distribution and tracking of the parasite, were developed based on microsatellites from the genome of *C. parvum* and *C. hominis*, and can only be applied to genetically related species (Tanriverdi and Widmer, 2003; Gatei et al., 2006b; 2007; 2009; Xiao, 2010; Herges et al., 2012; Feng et al., 2013; 2014; Wang et al., 2014). The ability to subtype all major *Cryptosporidium* species at the *gp60* is also important for understanding the transmission dynamics, particularly as recent research suggests that *gp60* plays an active and essential role in the life cycle of the parasite and that genetic variation at this locus might be essential for the parasite's long-term success (Abal-Fabeiro et al. 2013; Feng et al. 2013, 2014; Li et al. 2014). However, currently only *C. parvum*, *C. hominis*, *C. meleagrindis*, *C. tyzzeri*, *C. cuniculus*, *C. fayeri*, *C. ubiquitum* and *C. viatorum* can be subtyped at this locus (Chalmers et al. 2009; Lv et al. 2009; Power et al. 2009; Xiao, 2010; Feng et al. 2011a; Kváč et al. 2013b; Li et al. 2014; Stensvold et al., 2015). Using NGS, a recent study sequenced the whole genome of *C. ubiquitum* and developed a *gp60*-based tool, which showed that the *gp60* gene of *C. ubiquitum* has extensive sequence differences from the *gp60* gene of other *Cryptosporidium* spp. (Li et al. 2014). The availability of the whole genome sequence for *C. ubiquitum*, enabled the recent development of MLST tools for this species, which revealed greater genetic diversity within *C. ubiquitum* subtypes and broad host adapted groups in this species (Tang et al., 2016).

These findings highlight the need for Whole Genome Sequencing (WGS) analysis of diverse *Cryptosporidium* species to develop *gp60* and multi-locus sequence typing tools (MLST) for high-resolution subtyping of other common *Cryptosporidium* species in humans.
and domestic animals. Extensive WGS of Cryptosporidium spp. will likely lead to improved understanding of virulence factors in C. parvum and C. hominis and the genetic basis for host specificity and human infectivity of various Cryptosporidium species. This in turn will promote the development of vaccines and new therapies to help control the spread of Cryptosporidium (Striepen, 2013).

WGS of Cryptosporidium spp. has previously relied on propagation of the parasite in animals to generate enough oocysts from which to extract DNA of sufficient quantity and purity for analysis. More recent methods have developed methods for preparation of genomic Cryptosporidium DNA suitable for WGS directly from human stool samples and used it to generate 10 high-quality whole Cryptosporidium genome assemblies (Hadfield et al., 2015). Coupled with the higher sequencing power of NGS, this dramatically reduces the cost of WGS and extends the potential for this technology to be applied to the rapid and accurate genomic characterisation of clinical isolates, for the purposes of control of transmission, and treatment during outbreaks.

Metabolite profiling approaches have been used to identify novel or unanticipated metabolites and metabolic pathways in other protozoan parasites (MacRae et al., 2012; Millerioux et al., 2013; Arroyo-Olarte et al., 2013), to functionally characterise protein function (Oppenheim et al., 2014; de Macêdo et al., 2015), to generate defined minimal medium for cultivation of specific parasite stages (Creek et al., 2013) and to investigate the mode of action of anti-parasite drugs and resistance mechanisms (Lewis et al., 2014; Trochine et al., 2014). Future studies should also apply these approaches to Cryptosporidium. Currently, a major hurdle for research laboratories to facilitate biological, pathological, immunological and molecular and drug evaluation studies on Cryptosporidium has been the inability to continuously propagate Cryptosporidium in vitro (Ryan and Hijjawi, 2015). Although a continuous culture system has been described for Cryptosporidium (Morada et
al., 2016), in practice, the system is expensive and does not support the growth of *C. hominis*, the main species infecting humans. The modification of established cell culture methods to optimise growth is traditionally a laborious task. Metabolomics is a useful tool for the rational optimisation of cell culture medium and also offers a rapid and unbiased method for the analysis of cellular metabolism in different culture media. For example, a previous study successfully used untargeted semi-quantitative and targeted quantitative metabolomic analyses of fresh and spent media to identify the major nutritional requirements for the growth of bloodstream form of *Trypanosoma brucei* (Creek et al., 2013). A similar metabolomics approach could be used to improve the growth of *C. hominis* in culture, which would be a major advancement.

### 9.7 Conclusions

Overall this thesis has revealed novel insights into the epidemiology and metabolome of *Cryptosporidium* and cryptosporidiosis. Experiments conducted throughout this PhD have resulted in many more questions arising, which need to be answered, highlighting the complexity of this parasite and lack of knowledge of its biochemistry in particular.
REFERENCES


http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1301.0~2012~Main%20Features~Health%20service%20usage%20and%20experiences%20of%20care~234


Bouzid, M., Hunter, P. R., Chalmers, R. M., and Tyler, K. M. (2013). *Cryptosporidium*


Chappell, C. L., Okhuysen, P. C., Langer-Curry, R. C., Akiyoshi, D. E., Widmer, G., and


Protozoology, 33, 289-296.


Protozoology, **16**, 581-585.


El-Sherbini, G.T., and Gneidy, M.R. (2012). Cockroaches and flies in mechanical transmission of medical important parasites in Khaladyia Village, El-Fayoum,


sequence typing and genetic structure of Cryptosporidium hominis from children in Kolkata, India. *Infection, Genetics and Evolution*, 7, 197–205.


regional, seasonal and demographic variation. *Epidemiology and Infection*, 134, 111-118.


meleagridis isolate UKMEL1, C. baileyi isolate TAMU-09Q1, and C. hominis isolates TU502_2012 and UKH1. Pathogens and Diseases, In press.


Giardia sp. and Cryptosporidium sp. infections in a large Canadian health region. 

*BMCM Infectious Diseases*, **5**, 72.


Li, J.V., Wang, Y., Saric, J., Nicholson, J. K., Dirnhofer, S., Singer, B. H., Tanner, M.,


A community outbreak of cryptosporidiosis in Sydney associated with a public swimming facility: a case-control study. *Interdisciplinary Perspectives on Infectious Disease*, **2011**, 341065


Ryan, U., and Power, M. (2012). *Cryptosporidium* species in Australian wildlife and
domestic animals. *Parasitology*, **139**, 1673-1688.


Scheltema, R. A., Decuypere, S., T’kindt, R., Dujardin, J.C., Coombs, G. H., and Breitling, R.


Slavin, D. (1955). *Cryptosporidium meleagridis* (sp. nov.). *Journal of Comparative Pathology*, **65(3)**, 262-266.


Tyzzer, E.E. (1912). *Cryptosporidium parvum* (sp. nov.) a coccidian found in the small intestine of the common mouse. *Archive fur Prtistenkunde, 26*, 394-418.


Xu, P., Widmer, G., Wang, Y., Ozaki, L.S., Alves, J.M., Serrano, M.G., Puiu, D., Manque, P.,


Zhao, Z., Dong, H., Wang, R., Zhao, W., Chen, G., Li, S., Qi, M., Zhang, S., Jian, F., Zhao,
