MOLECULAR EPIDEMIOLOGY OF GIARDIA DUODENALIS INFECTIONS AMONG INDIGENOUS COMMUNITIES IN RURAL MALAYSIA

CHOY SEOW HUEY

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF MEDICINE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2016
UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Choy Seow Huey (I.C/Passport No: 870408-14-5318 )
Registration/Matric No: MHA 110015
Name of Degree: Doctor of Philosophy
MOLECULAR EPIDEMIOLOGY OF GIARDIA DUODENALIS INFECTIONS AMONG INDEGENOUS COMMUNITIES IN RURAL MALAYSIA

Field of Study: Parasitology

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature Date:

Subscribed and solemnly declared before,

Witness’s Signature Date:

Name:
Designation:
ABSTRACT

*Giardia duodenalis* is a protozoan parasite that can cause significant diarrheal diseases and is the most common intestinal protozoan parasite worldwide. It affects especially children from the rural areas, who are also the most vulnerable group that suffers from nutritional disorders that have been linked to this parasite. Being part of the complex group of parasitic, bacterial and viral diseases that debilitate the susceptible communities in developing regions from achieving full development potential, *Giardia* was included in the ‘Neglected Diseases Initiative’ in 2004. In Malaysia, information on the epidemiology of *Giardia* infection among different indigenous communities is limited. *Orang Asli*, the indigenous people that live in West Malaysia are the communities most at risk of acquiring parasitic infections. Meanwhile, the prevalence of the infection among indigenous people in East Malaysia has not been well explored. In addition, there is a scarcity of information on the genetic diversity and the dynamics of transmission of *G. duodenalis*. This cross-sectional study was carried out to investigate the prevalence and risk factors of *Giardia* infection among indigenous people in rural Malaysia. It also aims to identify *G. duodenalis* assemblages and sub-assemblages present in these communities based on multilocus genotyping approach. Moreover, the genetic data obtained by the present study were combined with a larger global sequence data for genetic diversity analyses. Faecal samples were collected between April 2011 and February 2013 from 1,330 participants from seven states of Malaysia. The samples were examined by wet mount and formalin-ether sedimentation methods while demographic, socioeconomic and environmental information was collected using a pre-tested questionnaire. Samples positive for *Giardia* were genotyped by using markers targeting the glutamate dehydrogenase (*gdh*), beta-giardin (*bg*) and triose phosphate isomerase (*tpi*) genes. The *tpi* sequences obtained by the present study as well as sequences from the global data obtained from the NCBI GenBank were used to analyse the population structure of *G. duodenalis*. The overall prevalence of *Giardia* infection was 11.6%. The prevalence was found to be significantly higher among the aboriginal population in West Malaysia (13.6%) when compared to the indigenous people in East Malaysia (5.8%). Multivariate logistic regression identified age of ≤ 12 years, lacking of toilet at household, not washing hands before eating, not washing hands after playing with animals, not boiling water before consumption, bathing in the river, and not wearing shoes when outside as the significant risk factors of *Giardia* infection among these communities. A significant association between *Giardia* infection and diarrhoea
among the studied population was reported. The frequency of diarrhoeal cases was significantly higher among *Giardia*-infected participants from West Malaysia when compared to their counterparts from East Malaysia. Of the 154 positive samples, 138 successfully yielded amplification by at least one of the markers (*gdh*, *bg* and *tpi*). Genotyping result showed that 69 of the isolates were classified as assemblage A and 69 were classified as assemblage B. Mixed infections were detected in 49 samples using a *tpi*-based assemblage specific protocol. At the sub-assemblages level, isolates belonged to assemblage A were AII. High nucleotide variation found in isolates of assemblage B made subtyping difficult to achieve. Infection with *Giardia* assemblage A was significantly associated with the age of ≤ 12 years, not boiling water before drinking and had close contact with domestic animals. With regard to assemblage B, large family size, bathing in river, practicing indiscriminate/open defecation, not washing hands before eating, and playing with soil were the associated factors. No association between the assemblages and the presence of symptoms was found. Analysis of the Malaysian and global data showed that assemblages A, B, and E (the most prevalent assemblages in humans and animals), have different level of genetic diversity. Assemblage B had the highest level of both haplotype diversity and nucleotide diversity, followed by assemblage E. The analysis also revealed population expansion and high gene flow in all assemblages. In conclusion, the present study shows that the prevalence of *Giardia* infection is still high and of public health concern among indigenous populations in rural Malaysia. The findings of assemblage B and the anthropomorphic genotype AII implicate human-to-human transmission as the most possible mode of transmission among Malaysian indigenous people. Meanwhile, the population genetic study provides new insight into the genetic diversity of *Giardia* assemblages in different geographical regions and should have brought enlightenment to the dynamics and distribution of *Giardia* infection. In view of the significant difference in the prevalence of *Giardia* infection among the different indigenous communities, implemented policies that may help in controlling the infection should be identified. Providing proper sanitation, as well as provision of clean drinking water and proper health education regarding good personal hygiene practices will help significantly in reducing the prevalence and burden of *Giardia* infection in these communities.
**ABSTRAK**

ACKNOWLEDGEMENTS

First and foremost, I would like to express my heartfelt gratitude to my beloved family, especially to my mom and dad, who didn’t understand much the topic of my study, but still showed their unconditional love through words of encouragement. There would always be home-cooked meals full of nutrition awaiting me when I got home and made me felt so blessed and loved.

Deepest gratitude to my supervisors, Prof. Johari Surin, Assoc. Prof. Dr. Hesham Al-Mekhlafi, and Assoc. Prof. Dr. Mohammed Mahdy. Successful completion of this project depends largely on their guidance and encouragement. Admittedly there were times when I could not meet their expectations, I am grateful that they would always be there to give me their constant advice and cheerful support.

This work required a lot of field visiting to rural areas in West and East Malaysia. I would like to extend my gratitude to two graceful ladies in helping me to gain access as well as permission to these areas. One of them is my consultant, Dr. Maria Suleiman, whose team of friendly and helpful staffs from the Department of Health had assisted me greatly during my field trips in Sabah. Another lady is my lecturer in the department, Prof. Dr. Yvonne Lim who not only allowed me to join her collaboration work with other departments and JAKOA (Jabatan Kemajuan Orang Asli) in conducting projects related to Orang Asli in West Malaysia but also gave me a lot of spiritual motivation.

My special thanks to all the villagers from the communities I have visited for their kind cooperation and patience. They could have just ignored or left after knowing my intention on sample collection, but most of them were willing to participate and contribute to this study. As a result of the supporting teams and understanding
participants, I had not only got to collect samples and data for my project, but also got the chance to gain eye-opening and invaluable experiences through visiting to areas that I was previously not familiar with.

I am grateful to have an abundance of caring friends. Some of them were there to share my grouses when I met obstacle in my work; some of them were there to remind or criticize me, which I have no doubt about their good intentions to motivate me; some of them were there to assist me in my project like data and sample collection. My sincere gratitude to my friends, Romano, Li Li, Hany, Wahib, Nabil, Joon, Soo Ching, Xiang Ting, Nan Jiu, Lorainne, Reena, Yuee Teng, Mian Zi, Vinnie, Stanley, Lucas, Azlan, Aidil, Lian Chee, and Yue Shin.

I would like to express my greatest appreciation to the head of Department of Parasitology, Prof. Rohela Mahmud for providing me the needed support and the opportunity to work in a good and harmony department. I am thankful to all staff members especially Kak Mazni and Encik Affifudin who rendered their help throughout the duration of my programme. I am also grateful to have other friends in the department, although with background different races, religions, and countries, have left me with a lot heartwarming and joyful memories, sometimes through festival celebration, sometimes through motivating words and, sometimes just through the flashing of a big smile when passing by in the corridor.

Last but not least, I wish to thank Ministry of Education and University of Malaya for providing scholarship under program MyBrain 15 and funding the research under High Impact Research grant, UMRG grants and student grant. The scholarship has lightened my financial burden which allows me to focus more on my study. The research grants have not only allowed me to buy necessary research materials but also covered my expenses in travelling to the different places during sample collection.
# TABLE OF CONTENTS

Abstract ......................................................................................................................................................... iii  
Abstrak ......................................................................................................................................................... v  
Acknowledgements ............................................................................................................................................ vii  
Table of Contents ........................................................................................................................................... ix  
List of Figures ................................................................................................................................................ xiv  
List of Tables ................................................................................................................................................ xvi  
List of Symbols and Abbreviations ................................................................................................................ xvii  
List of Appendices .......................................................................................................................................... xix

## CHAPTER I: INTRODUCTION

1.1 General Introduction ................................................................................................................................... 1  
1.2 Problem statement ....................................................................................................................................... 4  
1.3 General objective ....................................................................................................................................... 6  
1.4 Specific objectives ....................................................................................................................................... 6  
1.5 Significance of study ................................................................................................................................... 7

## CHAPTER II: LITERATURE REVIEW

2.1 Intestinal parasites ....................................................................................................................................... 8  
2.1.1 Intestinal parasitic infections and their global burden ........................................................................ 8  
2.1.2 Intestinal parasitic infections and the indigenous people in Malaysia ............................................ 10  
2.2 Historical background of *Giardia* ........................................................................................................ 13  
2.2.1 Discovery and description .................................................................................................................. 13  
2.2.2 Nomenclature ................................................................................................................................... 14  
2.2.3 Evolution of Eukaryotes ................................................................................................................... 15
2.3 *Giardia* and parasitology.............................................................................................17

2.3.1 Classification...........................................................................................................17

2.3.2 Assemblages and sub-assemblage of *G. duodenalis*..............................................18

2.3.3 Morphology and life cycle.......................................................................................19

2.3.4 Epidemiology..........................................................................................................25

2.3.5 Clinical manifestation and pathogenesis.................................................................31

2.3.6 Diagnosis................................................................................................................40

2.3.7 Treatment................................................................................................................41

2.4 Molecular characterization.........................................................................................43

2.4.1 Molecular tools for discrimination of genetic variation within

*Giardia duodenalis*........................................................................................................44

2.4.2 Distribution of genotypes and its epidemiological implications.........................50

---

**CHAPTER III: MATERIALS AND METHODS**

3.1 Malaysia profile...........................................................................................................67

3.2 Study areas................................................................................................................68

3.3 Study population.......................................................................................................71

3.4 Sample size and sampling strategy...........................................................................72

3.5 Questionnaire survey...............................................................................................73

3.6 Faecal samples collection and examination.............................................................74

3.7 DNA extraction..........................................................................................................74

3.8 Genotyping of *Giardia duodenalis* using *gdh* gene..............................................76

3.9 Genotyping of *Giardia duodenalis* using *tpi* gene...............................................76

3.10 Detection of mixed infection using *tpi* gene.........................................................77

3.11 Genotyping of *Giardia duodenalis* using *bg* gene.............................................78

3.12 Determination of PCR results..................................................................................79

3.13 Phylogenetic analysis...............................................................................................80
CHAPTER IV: RESULTS

4.1 Prevalence and Associated Risk Factors of *Giardia* Infection among Indigenous Communities in Rural Malaysia

4.1.2 Study cohort and socioeconomic profile

4.1.3 Prevalence of *Giardia* infection

4.1.4 Associated factors with *Giardia* infection

4.1.5 Risk factors of *Giardia* infection

4.1.6 Association of *Giardia* infection with diarrhoea

4.2 Molecular genotyping of *Giardia duodenalis*

4.2.1 Preliminary assessment of nucleotide sequences and identification of assemblages

4.2.2 Analysis of molecular data of respective markers and their distribution

4.2.2.1 Assessment of *tpi* sequences

4.2.2.2 Assessment of *gdh* sequences

4.2.2.3 Assessment of *bg* sequences

4.2.3 Molecular epidemiology

4.2.3.1 Mixed infections

4.2.3.2 Association of genotypes with risk factors

4.2.4 Phylogenetic analyses and subtyping

4.2.4.1 Subtyping of *G. duodenalis* assemblages A

4.2.4.2 Subtyping of *G. duodenalis* assemblages B

4.2.4.3 Multilocus genotyping
4.2.5 Population genetic study.................................................................142

4.2.5.1 Genetic diversity and haplotype networks of *G. duodenalis* of the Malaysian population.................................................................142

4.2.5.2 Genetic diversity and haplotype networks of *G. duodenalis* of the worldwide population.................................................................146

4.2.5.3 Test for neutrality........................................................................147

4.2.5.4 Genetic differentiation and gene flow.......................................153

CHAPTER V: DISCUSSION

5.1 Prevalence and Epidemiology of *Giardia* Infection...............................155

5.2 Molecular characterization of *Giardia duodenalis*......................................166

5.2.1 Molecular epidemiology........................................................................166

5.2.2 Distribution of genotypes in different localities......................................170

5.2.3 Subtyping and its implication for zoonotic transmission.....................171

5.3 General opinions and interpretation of the molecular data obtained using multilocus genotyping.................................................................172

5.3.1 Variation in efficiency of amplification by different loci......................172

5.3.2 Discordant genotyping results............................................................174

5.3.3 High polymorphism and heterogeneous nucleotides in assemblage B..176

5.3.4 Allelic sequence heterozygosity and mixed infection..........................177

5.3.5 Debate related to genetic recombination..............................................179

5.4 Population genetic study........................................................................187

CHAPTER VI: CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions.............................................................................................191

6.2 Recommendations......................................................................................193

6.3 Perspective future studies........................................................................194
References.........................................................................................................................................196

List of Publications and Papers Presented ......................................................................................229

Appendices.......................................................................................................................................231
LIST OF FIGURES

Figure 2.1: Trophozoites and cyst of *Giardia* .......................................................... 21

Figure 2.2: Life cycle of *Giardia* ............................................................................. 24

Figure 3.1: A map shows the study areas................................................................... 84

Figure 3.2: Workflow chart .......................................................................................... 85

Figure 4.1: Phylogram constructed by BI analysis, based on the representative sequences of *tpi*, *gdh*, and *bg* and reference sequences from Genbank... 138

Figure 4.2: Phylogram constructed by NJ analysis, based on the representative sequences of *tpi*, *gdh*, and *bg* and reference sequences from Genbank... 140

Figure 4.3: Phylogram constructed by BI analysis, based on the sequences amplified at all three loci.......................................................................................... 141

Figure 4.4: Median joining network of assemblage A sequences from Malaysian population ............................................................................................................. 144

Figure 4.5: Median joining network of assemblage B sequences from Malaysian population ............................................................................................................. 145

Figure 4.6: Median joining network of assemblage A sequences of the worldwide population ............................................................................................................. 149

Figure 4.7: Median joining network of assemblage B sequences of the worldwide population ............................................................................................................. 150

Figure 4.8: Median joining network of assemblage E sequences of the worldwide population ............................................................................................................. 151

Figure 4.9: Observed and expected mismatch distribution for *Giardia duodenalis* based on *tpi* gene.................................................................................. 152

Figure 5.1: Event of recombination ............................................................................ 185

Figure 5.2: Relationship revolving the question of ASH and mixed infection........ 186
LIST OF TABLES

Table 2.1: Members and their characteristics of the genus *Giardia* ........................................22

Table 2.2: Host distribution of *Giardia duodenalis* assemblages and their formerly used names .............................................................................................................23

Table 2.3: Prevalence of human *Giardia* infection from 1970s-2010s in different populations of Malaysia ..............................................................................................................29

Table 2.4: Correlation between assemblages A and B with symptoms ....................................34

Table 2.5: Nucleic acid-based methods .........................................................................................47

Table 2.6: Genotyping markers .....................................................................................................49

Table 2.7: Prevalence and genotype of human Giardia infection in different regions of the world .................................................................................................................................59

Table 2.8: Prevalence of subtypes of assemblage A in humans and animals and their distribution in different regions of the world .........................................................66

Table 3.1: Villages and their main tribal group involved in the study ........................................70

Table 4.1: General characteristics of the indigenous communities that participated in this study .................................................................................................................................................87

Table 4.2: Prevalence and distribution of intestinal parasitic infections among the indigenous communities that participated in this study ........................................................................89

Table 4.3: Univariate analysis of factors associated with *Giardia* infection among the indigenous communities that participated in this study .................................................92

Table 4.4: Multivariate analysis of risk factors associated with *Giardia* infection between Peninsular Malaysia and Sabah ..........................................................................................96

Table 4.5: Multivariate analysis of risk factors associated with *Giardia* infection among the indigenous communities that participated in this study .................................99

Table 4.6: The distribution of assemblages A and B based on the different loci and mixed infection ........................................................................................................................................101

Table 4.7: The distribution of assemblages A and B and mixed infection for all the isolates ........................................................................................................................................102

Table 4.8: Accession number and isolates of tpi ........................................................................107
Table 4.9: Multiple alignments of *tpi* sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblages A and B ................................................................. 109

Table 4.10: Distribution of *tpi* subtypes in different states................................. 111

Table 4.11: Accession number and isolates of *gdh* ........................................ 115

Table 4.12: Multiple alignments of *gdh* sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblages A and B .................................................................................. 117

Table 4.13: Distribution of *gdh* subtypes in different states ................................ 119

Table 4.14: Accession number and isolates of *bg* ............................................. 122

Table 4.15: Multiple alignments of *bg* sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblages A and B .................................................................................. 123

Table 4.16: Distribution of *bg* subtypes in different states .................................. 125

Table 4.17: Univariate analysis of factors associated with *Giardia* infection according to assemblages ................................................................................................................. 130

Table 4.18: Genetic diversity of *G. duodenalis* isolates from Malaysia among three loci ....................................................................................................................................... 143

Table 4.19: Genetic diversity and test of neutrality of *G. duodenalis* for the worldwide population ......................................................................................................................... 148

Table 4.20: Genetic differentiation (FST) and gene flow (Nm) ............................. 154
## LIST OF SYMBOLS AND ABBREVIATIONS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>≈</td>
<td>Almost equal to</td>
</tr>
<tr>
<td>&amp;</td>
<td>And</td>
</tr>
<tr>
<td>asb</td>
<td>Assemblage</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>BI</td>
<td>Bayesian Inference</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>cm</td>
<td>Centimeter</td>
</tr>
<tr>
<td>CDC</td>
<td>Centre for Disease Control and Prevention</td>
</tr>
<tr>
<td>$X^2$</td>
<td>Chi-square test</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>°C</td>
<td>Degree centigrade</td>
</tr>
<tr>
<td>DALYs</td>
<td>Disability-Adjusted Life Years</td>
</tr>
<tr>
<td>dNTPs</td>
<td>Deoxynucleotide triphosphates</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxynucleotide acid</td>
</tr>
<tr>
<td>=</td>
<td>equal</td>
</tr>
<tr>
<td>et al,</td>
<td><em>et alia</em> (others)</td>
</tr>
<tr>
<td>e.g.,</td>
<td><em>exempli gratia</em> (example)</td>
</tr>
<tr>
<td>FERG</td>
<td>Foodborne Disease Burden Epidemiology Reference Group</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>g/L</td>
<td>Gram per liter</td>
</tr>
<tr>
<td>i.e.,</td>
<td><em>id est</em> (that is)</td>
</tr>
<tr>
<td>≥</td>
<td>Greater than or equal to</td>
</tr>
<tr>
<td>JAKOA</td>
<td><em>Jabatan Kemajuan Orang Asli</em></td>
</tr>
<tr>
<td>km</td>
<td>Kilometer</td>
</tr>
</tbody>
</table>
\( p \) : Level of significance

\( \text{MgCl}_2 \) : Magnesium chloride

NCBI : National Centre for Biotechnology Information

\( \mu g \) : Microgram

\( \mu L \) : Microliter

\( \mu M \) : Micromolar

ml : Milliliter

mm : Micrometer

MOH : Ministry of Health

OD : Odd ratio

\% : Percent

PCR : Polymerase Chain Reaction

RAPD : Random Amplification of Polymorphic DNA

RPM : Revolution Per Minutes

RM : Ringgit Malaysia

RFLP : Restriction Fragment Length Polymorphism

SSCP : Single-Stranded Conformation Polymorphism

\~ : Similar to (tilde)

STH : Soil-Transmitted Helminth

spp : Species

SD : Standard Deviation

SPSS : Statistical Package for the Social Sciences

\( \leq \) : Smaller or equal to

U : Unit

UV : Ultraviolet

WHO : World Health Organization
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Consent form</td>
<td>231</td>
</tr>
<tr>
<td>II</td>
<td>Questionnaire</td>
<td>232</td>
</tr>
<tr>
<td>III</td>
<td>Distribution of the samples from different states and tribes</td>
<td>234</td>
</tr>
<tr>
<td>IV</td>
<td>Photos of the villages</td>
<td>235</td>
</tr>
<tr>
<td>V</td>
<td>Gel photos</td>
<td>238</td>
</tr>
</tbody>
</table>