MOLECULAR CHARACTERIZATION OF ROTAVIRUS ISOLATES IDENTIFIED FROM CHILDREN UNDER FIVE YEARS WITH ACUTE GASTROENTERITIS IN KOTA KINABALU AND KUNAK DISTRICTS, SABAH

LIA NATASHA AMIT

FACULTY OF MEDICINE AND HEALTH SCIENCES
UNIVERSITI MALAYSIA SABAH
2022

MOLECULAR CHARACTERIZATION OF ROTAVIRUS ISOLATES IDENTIFIED FROM CHILDREN UNDER FIVE YEARS WITH ACUTE GASTROENTERITIS IN KOTA KINABALU AND KUNAK DISTRICTS, SABAH

LIA NATASHA AMIT

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

FACULTY OF MEDICINE AND HEALTH SCIENCES
UNIVERSITI MALAYSIA SABAH
2022

UNIVERSITI MALAYSIA SABAH
BORANG PENGESAHAN STATUS TESIS

JUDUL : **MOLECULAR CHARACTERIZATION OF ROTAVIRUS ISOLATES IDENTIFIED FROM CHILDREN UNDER FIVE YEARS WITH ACUTE GASTROENTERITIS IN KOTA KINABALU AND KUNAK DISTRICTS, SABAH**

IJAZAH : **DOKTOR FALSAFAH SAINS PERUBATAN**

BIDANG : **SAINS PERUBATAN**

Saya **LIA NATASHA AMIT**, Sesi **2018-2022**, mengaku membenarkan tesis Doktoral ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis ini adalah hak milik Universiti Malaysia Sabah
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/):

[ ]	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA 1972)
[ ]	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
[/]	TIDAK TERHAD	

Disahkan Oleh,

_____ _____

LIA NATASHA AMIT (Tandatangan Pustakawan)
DM1821011T

_____

Tarikh : 30 September 2022 (Prof. Dr. Kamruddin Ahmed)
Penyelia Utama

DECLARATION

I hereby declare that the material in this thesis is my own except for quotations, equations, summaries, and references, which have been duly acknowledged.

21 April 2022

Lia Natasha Amit
DM1821011T

CERTIFICATION

NAME : **LIA NATASHA AMIT**

MATRIC NUM. : **DM1821011T**

TITLE : **MOLECULAR CHARACTERIZATION OF ROTAVIRUS ISOLATES IDENTIFIED FROM CHILDREN UNDER FIVE YEARS WITH ACUTE GASTROENTERITIS IN KOTA KINABALU AND KUNAK DISTRICTS, SABAH**

DEGREE : **DOCTOR OF PHILOSOPHY IN MEDICAL SCIENCE**

FIELD : **MEDICAL SCIENCE**

VIVA DATE : **21 APRIL 2022**

CERTIFIED BY;

Signature

1. **MAIN SUPERVISOR**
   Prof. Dr. Kamruddin Ahmed
   _____

2. **CO-SUPERVISOR**
   Prof. Dr. Mohammad Saffree Jeffree
   _____

3. **CO-SUPERVISOR**
   Dr. Fong Siat Yee @ Alison
   _____

4. **CO-SUPERVISOR**
   Dr. Daisuke Mori
   _____

ACKNOWLEDGEMENT

Foremost, I would like to express my sincere gratitude to my supervisor Prof Kamruddin Ahmed for his continuous support, guidance, insightful comments, and the opportunity to engage in the writing and lab activities that foster the research skills. The experiment works could not have been possible without the help of the two laboratory experts, Dr. Daisuke Mori and Mrs Jecelyn John. My sincere thanks also go to my co-supervisors, Dr. Alison Fong Siat Yee and Prof Dr. Mohammad Saffree Jeffree. I will always remember my fellow colleagues for all the exciting time and educational discussion that we had. Last but not least, debt of gratitude is owed to my families for their motivation and understanding.

Lia Natasha Amit
21 April 2022

ABSTRACT

Globally, rotavirus is the leading cause of acute gastroenteritis (AGE) in children under five years age. The highest prevalence of rotavirus-positivity was reported in South East Asia. In Malaysia, AGE cases increased exponentially between 1990 and 2017. However, the major etiologic agent was unknown. The nationwide surveys in 2006 and 2016 showed that Bumiputera children in Borneo were prone to AGE. To date, rotavirus surveillance was conducted between 1977 and 2017 in West Malaysia states and Sarawak. Meanwhile, in Sabah there were two preliminary studies conducted between 2005 and 2018, respectively. This study aimed to determine the burden, distribution of rotavirus G/P genotypes, genomic diversity, and commercial vaccines efficacy towards circulating rotavirus strains. Watery stools were collected from children with AGE under five years old who went to four Sabah government healthcare facilities. Rotavirus was detected by commercially available Rotaclone kit. Viral genomic RNA was extracted from Rotaclone-positive sample. Genotyping was performed by reverse-transcriptase PCR and the result was confirmed by sequencing of outer capsid genes. Genomic diversities were determined by PAGE. Multiple sequence alignment was done by ClustalW. Phylogenetic analyses and amino acid sequence similarity calculations were conducted in MEGA. Between January 2018 and February 2020, 422 watery stool samples were collected at Hospital Likas (n=294), Hospital Kunak (n=104), Klinik Kesihatan Menggatal (n=15), and Klinik Kesihatan Telipok (n=9). The rotaviral-AGE positivity was (96/422) (22.7%). Children aged 12-23 month, male, and Bajau ethnicity were more likely to be infected with rotavirus. Most common G/P genotypes were (n=47)," with "G/P genotypes identified by ELISA in the 69 rotavirus were G3P[8] (n=47), G9P[8] (n=10), G1P[8] (n=7), G12P[6] (n=3), G8P[8] (n=1), and GXP[8] (n=1). Novel rotavirus strains that had never reported in Malaysia before, such as equine-like G3P[8], G12P[6], and bovine-like G8P[8], were identified in this study. Sabahan strains exhibited a large genomic diversity as indicated by the presence of various electrophoretype patterns and G/P genotypes. Phylogenetic analysis showed that Sabahan strains originated from various sources or evolved independently for a long time since then they were introduced to Sabah. Sabahan strains had considerable amino acid variations on VP7 antigenic epitopes, indicating the vaccines would not be effective. Continuous surveillance in all Sabah districts is crucial to capture the overall picture of rotaviral-AGE burden and rotavirus G/P genotype distribution. Whole genome sequencing analysis is required to elucidate the genomic diversity, origin, and evolution of Sabahan strains. Further exploration of sociodemographic and risk factors data is essential to have a better understanding of the rotavirus transmission.

ABSTRAK

PENGKELASAN MOLEKULAR ISOLAT-ISOLAT ROTAVIRUS YANG DIKENALPASTI DARIPADA KANAK-KANAK DI BAWAH UMUR LIMA TAHUN YANG MENGHIDAP GASTROENTERITIS AKUT DI DAERAH KOTA KINABALU DAN KUNAK, SABAH

Rotavirus ialah penyebab utama gastroenteritis akut kanak-kanak di bawah umur lima tahun di seluruh dunia. Kepositifan gastroenteritis akut-rotavirus tertinggi dilaporkan di Asia Tenggara. Di Malaysia, berlaku peningkatan kes gastroenteritis di antara tahun 1990 dan 2017. Walau bagaimanapun, agen penyebabnya tidak diketahui. Soal selidik peringkat kebangsaan pada tahun 2006 dan 2016 mendapati bahawa kanak-kanak Bumiputera cenderung dijangkiti gastroenteritis akut. Pengawasan berterusan rotavirus telah dilakukan di antara tahun 1977 dan 2010 di Malaysia Barat dan Sarawak. Sementara itu, di Sabah terdapat dua kajian awal pada tahun 2015 dan 2018. Kajian ini bertujuan menentukan beban, taburan genotip-genotip G/P rotavirus, kepelbagaian genomik, dan keberkesanan vaksin komersial terhadap strain rotavirus yang beredar. Najis cair telah dikumpul daripada kanak-kanak yang menghidap gastroenteritis akut dan berumur lima tahun ke bawah yang telah menerima rawatan di empat fasiliti kesihatan kerajaan. Rotavirus telah dikesan menggunakan kit komersial Rotaclone yang sedia ada. RNA genomik virus telah diekstrak daripada sampel positif-Rotaclone. Pengkelasan genotip dilakukan dengan amplifikasi PCR transcriptase berbalik dan keputusan disahkan dengan penjujukan gen-gen kapsid luar. Kepelbagaian genomik ditentukan oleh PAGE. Penjajaran jujukan berbilang dilakukan dengan ClustalW. Analisis filogenetik dan pengiraan persamaan jujukan asid amino dilakukan dengan MEGA. Dari Januari 2018 hingga Februari 2020, 422 sampel najis cair telah dikumpul dari Hospital Likas (n=294), Hospital Kunak (n=104), Klinik Kesihatan Menggatal (n=15), dan Klinik Kesihatan Telipok (n=9). Kepositifan gastroenteritis akut-rotavirus ialah (96/422) (22.7%). Kanak-kanak yang berumur 12-23 bulan, lelaki, dan etnik Bajau adalah lebih cenderung dijangkiti rotavirus. Genotip-genotip G/P yang biasa dikenalpasti ialah G3P[8] (n=47), diikuti G9P[8] (n=10), G1P[8] (n=7), G12P[6] (n=3), G8P[8] (n=1), GXP[8] (n=1). Kajian ini menemui strain-strain rotavirus baru yang pertama di Malaysia iaitu strain serupa rotavirus kuda-G3P[8], strain G12P[6], dan strain serupa rotavirus lembu-G8P[8]. Strain-strain Sabah mempunyai kepelbagaian genomik seperti yang ditunjukkan melalui variasi corak elektroforetip dan genotip G/P. Analisis filogenetik menunjukkan bahawa strain-strain Sabah berasal daripada pelbagai sumber atau telah lama dibawa masuk ke Sabah pada masa lampau sebelum berevolusi sendiri. Strain-strain Sabah mempunyai beberapa variasi asid amino pada epitop antigen VP7, menunjukkan bahawa vaksin tidak akan berkesan. Pengawasan berterusan melibatkan semua daerah di Sabah adalah penting untuk mendapatkan gambaran keseluruhan beban gastroenteritis akut-rotavirus dan taburan genotip G/P. Analisis penjujukan genom diperlukan untuk menentukan kepelbagaian genomik, asal-usul, dan evolusi strain-strain Sabah. Lanjutan eksplorasi sosiodemografi dan faktor-faktor risiko adalah penting untuk lebih memahami penyebaran rotavirus.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

cm	-	Centimeter
Hz	-	Hertz
L	-	Liter
mA	-	Milliampere
mL	-	Milliliter
mm	-	Millimeter
rpm	-	Revolutions per minute
V	-	Volt
µL	-	Microliter
°C	-	Degree Celcius
µM	-	Micromolar

LIST OF ABBREVIATIONS

AGE	-	Acute gastroenteritis
APS	-	Ammonium persulphate
bp	-	Base pair
cDNA	-	Complementary DNA
DNA	-	Deoxyribonucleic acid
dsRNA	-	Double-stranded RNA
ELISA	-	Enzyme-linked immunosorbent assay
EM	-	Electron Microscopy
GTase	-	Guanylyltransferase
GRSN	-	Global Rotavirus Surveillance Network
IFN	-	Interferon
kb	-	Kilobase
mRNA	-	Messenger RNA
NHMS	-	National Health Morbidity Survey
PAGE	-	Polyacrylamide gel electrophoresis
PBS	-	Phosphate buffer saline
PCR	-	Polymerase chain reaction
pp	-	Pyrophosphate
qRT-PCR	-	quantitative PCR
RNA	-	Ribonucleic acid
RT-PCR	-	Reverse transcriptase PCR
RVA	-	Rotavirus A
TEMED	-	Tetramethylethylenediamine
WHO	-	World Health Organization

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background

Gastroenteritis is an inflammation of gastrointestinal tract indicated by three or more episodes of watery stool in 24 hours. Gastroenteritis is classified into three types depending on symptoms and the length it lasts: acute gastroenteritis (AGE), acute bloody gastroenteritis, and persistent gastroenteritis. AGE lasts for <14 days. Persistent gastroenteritis lasts for ≥14 days (Guerrant *et al.*, 2001; World Health Organization, 2005). Globally, AGE has become the fifth leading cause of mortality in children below five years age, accounting for 500,000 deaths (Troeger *et al.*, 2017). Most of AGE mortalities occur in developing countries of Africa and South East Asia (Sidoti *et al.*, 2015). A variety of bacteria, parasites, and viruses are responsible for AGE but 70-90% of them are caused by viruses. Rotavirus, norovirus, and sapovirus are the most common gastroenteritis viruses. However, 40% of the global AGE is caused by rotavirus (Orenstein *et al.*, 2020; Franco *et al.*, 2012; Operario *et al.*, 2017). According to a study conducted by Global Rotavirus Surveillance Network (GRSN) from 2008-2016, the highest rotaviral-AGE positivity was reported in South East Asia (Aliabadi *et al.*, 2019). In Malaysia, rotavirus is responsible for 31,000 hospitalizations, 145,000 home-treated cases, 41,000 outpatient visits, and 27 death cases, annually (Loganathan *et al.*, 2016). Despite the increasing trend of AGE cases from 1990 until 2017, rotavirus surveillance in Malaysia has been conducted only between 1977 and 2010 exclusively in West Malaysian states and Sarawak. In Sabah, preliminary rotavirus studies were conducted in 2005 and 2018, respectively (Goh *et al.*, 2009; World Health Organization, 2018).

Rotavirus was initially discovered in duodenum epithelia in electron microscopy. The name of rotavirus is originated from the Latin word 'rota', which means wheel, because rotavirus appears like a wheel in the observation using a transmission electron microscope. Structurally, a complete rotavirus particle is made of 11 double-stranded RNA segments enclosed by three layered particles: inner core, intermediate layer, and outer capsid. Inner core comprises of VP1, VP2, and VP3 proteins. Intermediate layer is mainly composed of VP6 protein. Outer capsid is made of VP7 and VP4 protein (Martella *et al.*, 2010; Desselberger, 2014). From 11 RNA segments, six of them encode of structural proteins (VP1, VP2, VP3, VP4, VP6, VP7), while remaining five encode of nonstructural proteins (NSP1, NSP2, NSP3, NSP4, NSP5/NSP6) (Mattion *et al.*, 1991). VP1 and VP3 act as RNA-dependent RNA polymerase (RdRP) and methyltransferase during viral replication, respectively (Matthijnssen *et al.*, 2008a). VP7 and VP4 are major rotavirus neutralizing antigens thus become the ideal targets for vaccine development (Schoondermark *et al.*, 2013). Nonstructural proteins function in combating host immune response (Qin *et al.*, 2011; Piron *et al.*, 1998). NSP4 protein has multiple functions. During replication, it facilitates transportation of newly synthesised double-layered particle rotavirus to endoplasmic reticulum for outer capsid synthesis. It also elicits the activation of enteric nervous system (Trask *et al.*, 2012; Ramig, 2004; Hagbom *et al.*, 2011).

Rotavirus is distinguished into eight Groups/Species based on VP6 antigenicity. Among eight known Groups, Rotavirus Group A significantly infect human and various animal species. Rotavirus Group A is further classified into G/P genotypes (binomial classification). G and P genotypes are defined by VP7 and VP4 proteins, respectively (Hoshino *et al.*, 1985; Estes & Cohen, 1989). Up to date, there are 36 G and 51 P genotypes have been discovered (Sadiq *et al.*, 2018; Pasittungkul *et al.*, 2021; Donato *et al.*, 2021). G1P[8], G3P[8], G4P[8], G9P[8], and G2P[4] are the globally predominant human rotavirus genotypes (Kaplon *et al.*, 2018; Ide *et al.*, 2015). Rare G/P genotypes, which are uncommon in human rotavirus, are mostly found in developing countries (Chakraborty *et al.*, 2016; Yahiro *et al.*, 2018; Wangchuk *et al.*, 2014). They are often naturally maintained in non-human hosts. For example, G12 and P[6] genotypes are common in porcine rotaviruses. G8 and G10 genotypes are common in bovine rotaviruses (Luchs *et al.*, 2016). Lately, these rare genotypes have been found in human rotaviruses, suggesting the zoonotic potential

of rotaviruses (Uchida *et al*., 2006; Tacharoenmuang *et al*., 2020; Luchs *et al*., 2016; Kondo *et al*., 2017).

Rotaviruses can be classified into different groups by testing electrophoretype pattern (RNA profile) (Tacharoenmuang *et al*., 2020; Wangchuk *et al*., 2014). Electrophoretype pattern is the migration pattern of 11 RNA segments which is developed by polyacrylamide gel electrophoresis (PAGE) (Matsui *et al*., 1990; Wani *et al*., 2003). In addition, three genogroups (gene groups); Wa-like, DS-1-like, and Au-like (Nakagomi *et al*., 1989). Wa-like genogroup harbors G1/G3/G4/P[8] genotype, long electrophoretype pattern, and gene backbone type 1. DS-1-like genogroup harbors G2P[4] genotype, short electrophoretype pattern, and gene backbone type 2. Au-like genogroup harbors G3P[9] genotype, long electrophoretype pattern, and gene backbone type 3 (Sadiq *et al*., 2018). RNA-RNA hybridization studies showed that Wa-like, Ds-1-like, Au-like genogroup are closely related to rotaviruses of porcine, bovine, feline origin, respectively (Desselberger *et al*., 2001; Matthijnssens *et al*., 2010a).

Rotavirus strains carrying similar G/P genotype and identical electrophoretype pattern are inferred to be the same strain. However, further whole genome sequencing of 11 RNA segments is required to confirm it(Tacharoenmuang *et al*., 2020; Komoto *et al*., 2016). Evolutionary relationship of Group A rotavirus in human and animal species can be elucidated by whole genome sequencing and phylogenetic analysis of genotypes representing 11 RNA segments (Komoto *et al*., 2021; Sircar *et al*., 2021). Illumina (MiSeq desktop sequencer) is the common Next Generation Sequencing platform used for sequencing rotavirus genome (Tacharoenmuang *et al*., 2016; Fujii *et al*., 2019). Rotavirus genome constellation is represented as Gx-Px-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx according to VP7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5/NSP6, respectively. This whole genome-based nomenclature is designated by Rotavirus Classification Working Group (RCWG). According to RCWG, assignment of a new genotype for each RNA segment is determined according to the nucleotide identity percentage cut-off value (Matthijnssens *et al*., 2011a).

Rotavirus is a versatile microorganism due to its ability to evolve by several different mechanisms such as accumulation of point mutation, gene rearrangement and gene reassortment (Gómara *et al*., 2001). Accumulation of point mutation (antigenic drift) occurs during replication with the aid of error prone RdRP (Jain *et al*., 2014). Gene rearrangement occurs by duplication and deletion of RNA segment. For example, RNA segment 11 undergoes partial duplication at the 3' terminal of untranslated region which causes slower migration of RNA segment 11 during PAGE, resulting a formation of short electrophoretype pattern (Schnepf *et al*., 2008). Rapid evolution of rotavirus occurs by gene reassortment (antigenic shift). Gene reassortment can occur between rotaviruses of different genogroups (intergenogroup) or between human and animal/animal-like rotaviruses, in the host cell they co-infect, due to independent segregation of RNA segments Equine-like G3P[8], G12P[6], and bovine-like G8P[8] are likely to have emerged due to gene reassortment events, individually (Bányai *et al*., 2012; Matthijnssens *et al*., 2012; Komoto *et al*., 2015; Nakagomi *et al*., 2017a; Tacharoenmuang *et al*., 2016; Tacharoenmuang *et al*., 2020).

Intergenogroup gene reassortment occurred between Wa-like G1P[8] strain and DS-1-like G2P[4] strain, which led to the emergence of a reassortant strain, DS-1-like G1P[8] strain (Fuji *et al*., 2019). Then, DS-1-like G1P[8] strain underwent gene reassortment with equine strain carrying G3 genotype VP7 gene segment, causing the emergence of equine-like G3P[8] strain (Komoto *et al*., 2016). Bovine-like G8P[8] strain emerged by gene reassortment between DS-1-like G1P[8] strain and bovine/bovine-like strain carrying G8 genotype VP7 gene segment (Kondo *et al*., 2017). Previously, G12 and P[6] genotypes were commonly found in porcine rotavirus strains (Desselberger *et al*., 2001; Miyazaki *et al*., 2011). The emergence of G12P[6] strain was mediated by gene reassortment between porcine/porcine-like G12 strain and porcine/porcine-like P[6] strain (Rahman *et al*., 2007; Martella *et al*., 2010; Patton *et al*., 2012). The globally predominant G9P[8] strain emerged by gene reassortment between globally predominant G1/G3/G4P[8] strains and porcine strain carrying G9 genotype VP7 gene segment (Iturriza-Gómara *et al*., 2000a). Several studies showed that gene reassortment events are highly frequently occurring among rotavirus strains in developing countries as indicated by the presence of diverse strains harboring different combination of G/P genotypes and