

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



**LIA NATASHA AMIT**

**UMMS**  
UNIVERSITI MALAYSIA SABAH

**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**



**UMS**

**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 Doktorai 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**  
**DM1821011T**

---

(Tandatangan Pustakawan)

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



**UMS**  
UNIVERSITI MALAYSIA SABAH

## 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;

UMS  
Signature  
UNIVERSITI MALAYSIA SABAH

**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



UMMS  
UNIVERSITI MALAYSIA SABAH

## 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**

### **PENKELASAN 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. Pengelasan 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

	Page
<b>TITLE</b>	i
<b>DECLARATION</b>	ii
<b>CERTIFICATION</b>	iii
<b>ACKNOWLEDGEMENT</b>	vi
<b>ABSTRACT</b>	v
<b><i>ABSTRAK</i></b>	vi
<b>LIST OF CONTENTS</b>	vii
<b>LIST OF TABLES</b>	x
<b>LIST OF FIGURES</b>	xiii
<b>LIST OF SYMBOLS</b>	xvi
<b>LIST OF ABBREVIATIONS</b>	xvii
<b>LIST OF APPENDICES</b>	xviii
<b>CHAPTER 1: INTRODUCTION</b>	1
1.1 Background	1
1.2 Problem statement	7
1.3 Hypothesis	9
1.4 Research Objectives	9
1.5 Significance of Research	9
<b>CHAPTER 2: LITERATURE REVIEW</b>	10
2.1 Global AGE burden	10
2.2 Global rotavirus burden	10
2.3 AGE burden in Malaysia	11
2.4 Rotavirus burden in Malaysia	11
2.5 Structure of rotavirus	12
2.5.1 Rotavirus genome	14

2.5.2	Rotavirus RNA segment	14
2.5.3	Rotavirus proteins	16
2.6	Rotavirus Classification	16
2.6.1	Classification by serotype (species)	17
2.6.2	G&P genotypes (binary classification)	17
2.6.3	Electrophoretype pattern	18
2.6.4	Genogroup classification	19
2.6.5	Genotype constellation classification	20
2.7	Rotavirus evolution	20
2.7.1	Point mutation	20
2.7.2	Gene reassortment	21
2.7.3	Gene rearrangement	22
2.7.4	Intragenic recombination	22
2.8	Rotavirus transmission	23
2.9	Rotavirus pathogenesis	24
2.10	Rotavirus clinical symptoms	24
2.11	Rotavirus diagnosis	25
2.11.1	ELISA (enzyme linked immunosorbent assay)	26
2.11.2	PAGE (Polyacrylamide gel electrophoresis)	26
2.11.3	Reverse transcriptase PCR (RT-PCR)	27
2.11.4	Full genome sequencing	28
2.12	Rotavirus prevention and treatment	29
2.13	Rotavirus vaccines	29
2.14	Distribution of rotavirus genotypes in worldwide and Malaysia	31
2.15	Rotavirus genotypes distribution in animals	33
<b>CHAPTER 3: RESEARCH METHODOLOGY</b>		<b>37</b>
3.1	Research framework, research design, duration of research, selection of research area, sampling, summary of research methodology	37

3.2	Ethical approval and consent of participants	41
3.2.1	Type of sample collected and sample size calculation	41
3.2.2	Collection of patient's watery stool sample and patient information	42
3.3	Rotavirus antigen detection by using commercially available ELISA kit	44
3.4	RNA extraction from ELISA-positive sample using QIAamp Viral RNA Mini Kit	45
3.5	RT-PCR for detection of rotavirus VP7 and VP4 genes	46
3.6	Polyacrylamide gel electrophoresis (PAGE)	53
3.6.1	Electrophoretyping	59
3.7	Phylogenetic analysis	60
3.8	Comparison of VP7 antigenic epitopes of circulating Sabahan rotavirus strains and vaccine strain	60
<b>CHAPTER 4: RESULT AND DISCUSSION</b>		61
4.1	Total number of watery stool samples collected during 2018-2020	61
4.2	Detection of rotavirus antigen from watery stool samples by ELISA commercial kit (Rotaclone)	64
4.3	The total number of rotavirus-positive samples identified during 2018-2020	64
4.4	Detection of rotavirus VP7 and VP4 genes	68
4.5	Electrophoretype of rotaviruses identified in this study	76
4.6	Phylogenetic analyses of Sabahan rotavirus strains	80
4.7	Amino acid differences in the neutralizing epitopes of VP7 protein of circulating strains and vaccine strains	93
<b>CHAPTER 5: CONCLUSION</b>		100
<b>REFERENCES</b>		101
<b>APPENDICES</b>		137

## LIST OF TABLES

	Page
Table 2.1 : RNA segments of rotavirus and proteins encoded in respective segments	15
Table 2.2 : Rotavirus proteins with their respective functions	16
Table 2.3 : Rotavirus species and their animal hosts	17
Table 2.4 : Information of WHO prequalified vaccines	31
Table 2.5 : Distribution of rotavirus G/P genotypes in Malaysia from 1977-2010	33
Table 2.6 : Distribution of Group A Rotaviruses G genotypes identified in human and animals	34
Table 2.7 : Distribution of Group A Rotaviruses P genotypes identified in human and animals	35
Table 3.1 : Primers used for PCR to detect rotavirus VP7 gene	47
Table 3.2 : Primers and their volume for a mixture of Master Mix in one PCR tube	47
Table 3.3 : Reagents and their volume for a mixture of AccessQuick Master Mix, AMV Reverse Transcriptase, and H <sub>2</sub> O	47
Table 3.4 : Total volume of reagents in one PCR tube for PCR of VP7 gene	48
Table 3.5 : Primers used for PCR to detect rotavirus VP4 gene	48
Table 3.6 : Primers and their volume for a mixture of Master Mix in one PCR tube	48
Table 3.7 : Reagents and their volume for a mixture of AccessQuick Master Mix, AMV Reverse Transcriptase, and H <sub>2</sub> O	49
Table 3.8 : Total volume of reagents in one PCR tube for PCR of VP4 gene	49
Table 3.9 : Primers used for rotavirus G-genotyping	50
Table 3.10 : Reagents and volume of Master Mix and the first PCR product of VP7 gene in one PCR tube	50
Table 3.11 : Primers used for rotavirus P-genotyping	51
Table 3.12 : Reagents and volume of Master Mix and the first PCR product of VP4 gene in one PCR tube	52

Table 3.13	Reagents and their volume for preparation of 10% separating gel	56
Table 3.14	Reagents and the volume for preparation of 4% upper gel	57
Table 3.15	Preparation of 1X running buffer	58
Table 4.1	Total number of watery stool samples collected from AGE patients involving all sampling sites	61
Table 4.2	Gender of AGE patients	62
Table 4.3	Number of rotavirus-positive samples detected from AGE cases	65
Table 4.4	Gender of rotavirus patients	65
Table 4.5	Rotavirus G genotype identified based on sequencing	71
Table 4.6	Rotavirus P genotype identified based on sequencing	71
Table 4.7	Changes of rotavirus G/P genotype distribution based on study phase	73
Table 4.8	RotaTeq™ (WI78-9) (depending on amino acid residues 87, 213, 238, and 148 in VP7[indicated by red triangle]) may not be sufficiently protective against Sabahan G3 strains	93
Table 4.9	Rotavac (depending on residues 87, 100, 145, and 221 in VP7[indicated by red triangle]) may not be sufficiently protective against Sabahan G9 strains	95
Table 4.10	Rotasiiil (depending on amino acid residues 87 and 100 in VP7[indicated by red triangle]) may not be sufficiently protective against Sabahan G9 strains	96
Table 4.11	Rotarix and Rotateq (depending on residues 94, 96, 213, 238, 145, 147, and 148 in VP7[indicated by red triangle]) may not be sufficiently protective against Sabahan G8 strain	96
Table 4.12	Comparison of VP7 antigenic epitope between Rotarix and Rotateq vaccine strains and Sabahan G1 strains. Amino acid alterations which associate with the failure of neutralizing antibody binding are indicated by red triangle	97

Comparison of VP7 antigenic epitope between Rotarix and Rotateq vaccine strains and Sabahan G12 strains. Amino acid alterations which associate with the failure of neutralizing antibody binding are indicated by red triangle



UMS  
UNIVERSITI MALAYSIA SABAH

## LIST OF FIGURES

	Page
Figure 2.1 : A & B - Rotavirus particle structure	13
Figure 2.2 : Rotavirus VP2 structure	13
Figure 2.3 : 1) Schematic diagram of rotavirus virion, 2) RNA segments of rotavirus and proteins encoded in respective segments	15
Figure 2.4 : Comparison of short electrophoretype and long electrophoretype	19
Figure 2.5 : Full length of VP7 gene (1062 bp) is amplified using Beg9 and End9 primer. For G genotyping, primers aAT8, aBT1, aCT2, aDT4, aET3, aFT9 amplifies variable regions within VP7 gene that are specific to G genotype (Gouvea et al., 1990)	27
Figure 3.1 : Research framework	38
Figure 3.2 : Sampling in this study	39
Figure 3.3 : A summary of the research methodology	40
Figure 3.4 : Stool collecting bottle. Each bottle contains 30 ml of PBS to which stool sample is suspended	43
Figure 3.5 : The cap of the stool collecting bottle is designated to be attached to the scoop	43
Figure 3.6 : The volume of the watery stool sample collected by the scoop should be sufficient according to the illustrated picture which is represented with 'O' symbol. The volume of the sample should not be less or overflow the scoop	44
Figure 3.7 : Polyacrylamide gel electrophoresis	53
Figure 3.8 : Cleaned glass plates were put on the tissue layer	54
Figure 3.9 : The spacer positioning cardboard was placed on a cleaned glass plate. Then, a rubber was placed side to side of the spacer positioning cardboard	54
Figure 3.10 : The glass tightener was put at side to side. The spacer positioning cardboard was now sandwiched by two cleaned glasses and positions are locked by the glass tightener	55

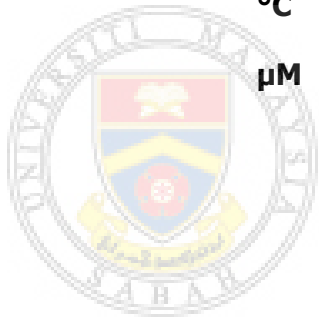
Figure 3.11	: The position of the glass plates, the rubber, and the glass stand	56
Figure 4.1	: AGE patients' age	62
Figure 4.2	: AGE patients' ethnicity	63
Figure 4.3	: Sample A12, B1, B2, B3, B5, B7, and B10 were defined as rotavirus-positive as they generated greater blue colour than the positive control (red-circled well)	64
Figure 4.4	: Rotavirus patients' age	66
Figure 4.5	: Rotavirus patients' ethnicity	67
Figure 4.6	: Gel electrophoresis shows the PCR products of VP7 gene (1062 bp)	68
Figure 4.7	: Gel electrophoresis shows the PCR products of VP4 gene	68
Figure 4.8	: Gel electrophoresis shows the PCR products of VP7 gene for G3 (374 bp) and G8 genotypes (885 bp)	69
Figure 4.9	: Gel electrophoresis shows the PCR products of VP7 gene for G9 genotype (306 bp)	70
Figure 4.10	: Gel electrophoresis shows the PCR products of VP4 gene for P[6] genotype (413 bp) and P[8] genotype (280 bp)	70
Figure 4.11	: Rotavirus G/P genotype identified from 2018-2020 in all study sites	72
Figure 4.12	: Rotavirus G/P genotype distribution during the first phase study in Sabaheast coast (Kunak) and west coast (Kota Kinabalu)	74
Figure 4.13	: Rotavirus G/P genotype distribution during the second phase study in Sabah east coast (Kunak) and west coast (Kota Kinabalu)	75
Figure 4.14	: Electrophoretype of rotavirus strains identified in the first phase study strains	77
Figure 4.15	: Electrophoretype of rotavirus strains identified in the second phase study	78
Figure 4.16	: All the rotavirus electrophoretotypes identified in this	79



Figure 4.17	: Phylogenetic analysis of VP7 gene of Sabahan G3P[8] strains constructed by Neighbour-Joining method in MEGA X software. First phase strains are indicated by bold circle. Second phase strains are indicated by bold triangle	80
Figure 4.18	: Phylogenetic analysis of VP7 gene of Sabahan G9P[8] strains constructed by Neighbour-Joining method in MEGA X software	84
Figure 4.19	: Phylogenetic analysis of VP7 gene of Sabahan G12P[6] strains constructed by Neighbour-Joining method in MEGA X software	86
Figure 4.20	: Phylogenetic analysis of VP7 gene of a Sabahan G8P[8] strain constructed by Neighbour-Joining method in MEGA X software	87
Figure 4.21	: Phylogenetic analysis of VP7 gene of Sabahan G1P[8] strains constructed by Neighbour-Joining method in MEGA X software. First phase study strains and second phase study strains are represented by bold circle and bold triangle, respectively	89
Figure 4.22	: Phylogenetic analysis of VP4 gene of Sabahan G1P[8], G3P[8], G8P[8], G9P[8], GXP[8] constructed by Neighbour-Joining method in MEGA X software. First phase study strains and second phase study strains are represented by circle and triangle, respectively. Rotavirus G types are represented by color; maroon (G1P[8]), orange (equine-like G3P[8]), blue (G8P[8] and GXP[8]), purple (human/wild type G3P[8]), green (G9P[8])	91
Figure 4.23	: Phylogenetic analysis of VP4 gene of Sabahan G12P[6] strains constructed by Neighbour-Joining method in MEGA X software	92

## LIST OF SYMBOLS

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



UMMS  
UNIVERSITI MALAYSIA SABAH

## LIST OF ABBREVIATIONS

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

## LIST OF APPENDICES

	Page
Appendix A : Patient information sheet	137
Appendix B : Informed consent form	139
Appendix C : Age distribution of children with AGE	143
Appendix D : Ethnicity distribution of AGE patient	146
Appendix E : Age distribution of children with rotavirus infection	148
Appendix F : Sample collection date and patient details for G3P [8] rotavirus strains (L68L69,L70,L71) identified from Hospital Likas	150
Appendix G : Percentage of nucleotide similarity on VP7 gene between Sabahan G3P [8]-lineage I and G3P [8]-lineage III	151
Appendix H : Percentage of nucleotide similarity on VP7 gene between Sabahan G3P [8] strains and other countries strains (lineage I)	152
Appendix I : Percentage of nucleotide similarity on VP7 gene between Sabahan G3P [8] strains and other countries strains (lineage III)	154
Appendix J : Percentage of nucleotide similarity on VP7 gene between Sabahan G9P [8] strains and other countries strains	155
Appendix K : Percentage of nucleotide similarity on VP7 gene between Sabahan G12P [6] strains and other countries strains	156
Appendix L : Percentage of nucleotide similarity on VP7 gene between Sabahan G8P [8] strains and other countries strains	157
Appendix M : Percentage of nucleotide similarity on VP7 gene between Sabahan G1P [8] strains and other countries strains (lineage I & II)	158

# 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  $\geq 14$  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 (Matthijnsen *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 (Kaplun *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; Matthijnsens *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 (Matthijnsens *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 electrophoretic 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; Matthijnsens *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