DIAGNOSIS OF DENGUE INFECTION USING METABOLOME ANALYSIS (BIOMARKERS PROFILING)

by

NURUL SHAHFIZA BT. NOOR

Thesis submitted in fulfillment of the requirements for the degree of Master of Science

April 2014

ACKNOWLEDGEMENTS

Alhamdulillah, thanks to ALLAH ALMIGHTY for valuable experience and knowledge and for giving me this opportunity to finish this study.

First and foremost, I would like to express my gratitude towards my supervisor Prof. Abdel-Hamid Zaki Amer. He had given me a lot of ideas, comments, advice and encouragement and been tolerance through out my research. With his guidance and motivations I would be able to go through this research.

I would also like to thank my supervisory committee members who are Assoc. Prof. Dr. Tang Thean Hock and Assoc. Prof. Dr. Hasnah Osman for all their support and guidance. Without them I am sure that I would have gone through a lot of trouble to complete my work. Besides, I would like to thank Prof. Dr. Khozirah Shaari (UPM), Assoc. Prof. Dr. Alfi Khatib (UPM) and Dr. Baharudin Ibrahim (USM) for their valuable guidance related to statistical analysis on this study. My sincere thanks to Dr. Teh Chin Hoe, Dr. Fang Fang and Dr. Joshua Hicks for providing expertise on the use of NMR for metabolomic study. My sincere thanks to Dr. Chow Ting Soo (Penang General Hospital) for allowing me to do urine collection at her Dengue Ward, Penang General Hospital.

I would like to express my sincere thanks to all staff involve in this study, En. Ahmad Zainuddin Yunus (Pharmaceutical Chemistry Unit, School of Pharmaceutical Sciences, USM), En. Zahari Othman (NMR Unit, School of Chemical Sciences, USM), En. Salahudin Mohd. Raof (Laboratory of Natural Products, Institute of Bioscience, UPM).

Special thanks to all my friends and coursemates of Infectomic Cluster and also in Laboratory of Natural Products, Institute of Bioscience, UPM. It has been a great pleasure working with them. Thank you so much for all your help and support and I really appreciate it.

I wish to express my deepest gratitude to my mother, Pn. Zuraini Md. Ali and my parents in law, En. Abu Bakar Idris and Pn. Maznah Basha, my beloved husband, En. Mohd. Hezry Abu Bakar, my son, Muhammad Adib Ziyaad Ilmaan, my brother and sister-in-law, En. Zamrul Hisyam Noor and Pn. Fatimah Azlina Ahmad, for their patience, undivided support and advice throughout my study. Without them, I would not be able to reach this final stage of my study. Thank you so much to my cousins Nazirul Hafiz Nazri, Muhammad Nazirul Syahmi Nazri and Muhammad Rasyad Razin Ruzlan for helping me collecting the samples.

Finally, I am thankful to all the volunteers involved in this study for their contribution without which none of this work would have been possible.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
LIST OF ABBREVIATIONS	xi
LIST OF SYMBOLS	xiii
ABSTRAK	xiv
ABSTRACT	xvi
CHAPTER ONE : GENERAL INTRODUCTION	1
CHAPTER TWO : LITERATURE REVIEW	6
2.1 Dengue: the global burden of disease	6
2.2 Dengue: the global epidemiology	8
2.2.1 Dengue epidemiology in Malaysia	9
2.3 Dengue case classification	11
2.4 Dengue diagnosis	11
2.4.1 Virus isolation	13
2.4.2 Serological diagnosis	13
2.4.3 Molecular diagnosis	15
2.5 Metabolomics as powerful tool for biomarker discovery	16
2.5.1 Biofluids	17

CHAPTER THREE : MATERIALS AND METHODS 20

3.1 Materials 20

3.1.1 Instrument 20

3.1.2 Subjects 20

3.1.3 Biological samples 22

3.1.4 Chemicals 22

3.2 Methods 23

3.2.1 Ethical approval 23

3.2.2 Sample preparation 23

3.2.2.1 Sample collection and storage 23

3.2.2.2 Sample preparation for 5mm NMR tube 23

3.2.3 NMR file processing 24

3.2.4 Chemometric analysis 24

3.2.4.1 Data reduction 24

3.2.4.2 Pattern recognition (Multivariate analysis) 25

3.2.5 Statistical analysis (Univariate analysis) 25

CHAPTER FOUR : RESULTS AND DISCUSSION 26

4.1 Urinalysis of patient infected with dengue virus versus healthy control 26

4.1.1 ^{1}H NMR urinalysis 26

4.1.2 Chemometric multivariate analysis and univariate statistical analysis 29

4.1.3 Urinary metabolic profiling of patient infected with dengue disease 30

4.1.4 Metabolite identification 36

4.1.5 Biochemistry and interpretation 38

4.1.5.1 Acetoacetic acid, 4-hydroxyphenylpyruvic acid, S-sulfocysteine, valerylglycine, betaine and fructose 39

4.1.5.2 N-acetylglutamic acid, creatine, creatine phosphate, myo-inositol, succinic acid, citrate, 3-hydroxy-3-methylglutarate 42

4.1.6 Conclusion 46

4.2 Chracterization of gender differences in patient infected with dengue virus 47

4.2.1 Differences in total metabolic profiles 48

4.2.1.1 Principle component analysis (PCA) 48

4.2.1.2 Orthogonal partial least square-discriminant analysis (OPLS-DA) 50

4.2.2 Metabolite identification 53

4.2.3 Biochemistry and interpretation 55

4.2.4 Conclusion 58

CHAPTER FIVE : GENERAL CONCLUSION & FUTURE WORKS **59**

REFERENCES **61**

APPENDIX A **73**

APPENDIX B **74**

APPENDIX C **85**

APPENDIX D **86**

LIST OF TABLES

		Page
Table 2.1	Advantages and limitation of dengue diagnostic tests.	**12**
Table 4.1	OPLS-DA model bins selected for assignment of urinary metabolites perturbed in the dengue infected patients compared to healthy individuals	**37**
Table 4.2	OPLS-DA model bins selected for assignment of urinary metabolites perturbed in the male dengue infected compared to female dengue infected.	**54**

LIST OF FIGURES

		Page
Figure 2.1	Number of dengue cases in Malaysia, 1995-2007 (Ministry of Health, 2010)	**10**
Figure 2.2	General time-line of a primary infection with dengue virus, from identification and isolation of the virus to detection of IgM and IgG (Philippe *et al.*, 2006).	**14**
Figure 4.1	Aliphatic region of 1D ^1H NMR spectra of the urine samples (a) healthy individual (b) male patient infected with dengue disease (c) female patient infected with dengue disease.	**28**
Figure 4.2	PCA loadings (up) and scores (below) plot showing the variation between urinary metabolic profiles of dengue-infected individuals and healthy individuals. This PCA model was constructed from ^1H NMR of fifty-two patients with dengue infection and fourty-three healthy individuals, regardless their gender differences. R^2X for the model = 0.3731.	**33**
Figure 4.3	OPLS-DA scores plot showing the variation between urinary metabolic profiles of the dengue-infected patients and healthy individuals. This model was constructed from ^1H NMR data of urine from 52 dengue-infected individuals and 43 healthy individuals, regardless their gender differences. Black square = dengue infected individuals and red circle = healthy individuals. The statistical parameters of the model were as follows R^2X= 0.9081 and Q^2 (cum) = 0.7880. The ellipse is a 95% Hotelling's T^2 ellipse.	**34**

Figure 4.4	Representative OPLS-DA S-plot showing relative contribution of signals to clustering of dengue-infected individuals and healthy individuals. This plot corresponds to Figure 4.3. Each point in the figure represents a bin. The p(corr)[1] axis represents the correlation of the bin towards the predictive variation shown in Figure 4.1. The p[1] axis represents the magnitude of the spectral bins. Bins from region circled in red represents the bins that decreased in the dengue infected individuals. Bins from region circled in black represents the bins that increased in the dengue infected individuals. Among them, bins with high p(corr)[1] (< -0.2 for increased and > 0.5 for decreased, high p[1] and VIP values exceeding 1.0 as point of cut off and $p<0.05$ are selected for metabolite identification.	**35**
Figure 4.5	PCA loadings (up) and scores (below) plot showing the variation between urinary metabolic profiles of dengue-infected patients (males vs females). This PCA model was constructed from spectra of ^1H NMR after pre-treating with SNV. It was constructed from 52 patients with dengue infection (female, n=13 and male, n=39). R^2X for the model = 0.4235.	**49**
Figure 4.6	OPLS-DA scores plot showing the variation between urinary metabolic profiles of the male vs female dengue-infected patients. This model was constructed from ^1H NMR data of urine from 52 dengue-infected patients (female, n=13 and male, n=39). Black square = male individual infected with dengue and red circle = female individual infected with dengue. The statistical parameters of the model were as follows $R^2Y= 0.5560$ and Q^2 (cum) = 0.0767. The ellipse is a 95% Hotelling's T^2 ellipse.	**51**

Figure 4.7 Representative OPLS-DA S-plot showing relative contribution of signals to clustering of dengue-infected individuals and healthy individuals. This plot corresponds to Figure 4.6. Each point in the figure represents a bin. The p(corr)[1] axis represents the correlation of the bin towards the predictive variation shown in Figure 4.1. The p[1] axis represents the magnitude of the spectral bins. Bins from region circled in red represents the bins that decreased in the male individuals infected with dengue. Bins from region circled in black represents the bins that increased in the male individuals infected with dengue. Among them, bins with high p(corr)[1] (< -0.2 for increased and > 0.2 for decreased, high p[1] and VIP values exceeding 1.0 as point of cut off and p<0.05 are selected for metabolite identification. 52

LIST OF ABBREVIATIONS

1.	AGA	L-arginine glycine amidinotranferase
2.	AKI	Acute kidney injury
3.	BHMT	Betaine-hemocysteine methyltransferase
4.	CD	Celiac disease
5.	COD	Chronic obstructive pulmonary disease
6.	Cum	Cumulative
7.	DALYs	Disability-adjusted life years
8.	DENV	Dengue virus
9.	DF	Dengue fever
10.	DHF	Dengue hemorrhagic fever
11.	DSS	Dengue shock syndrome
12.	D_2O	Deuterium oxide
13.	HMDB	Human Metabolome Database
14.	KH_2PO_4	Potassium phosphate
15.	KOH	Potassium hydroxide
16.	Min	Minute
17.	mL	Milliliter
18.	Mm^3	Cubic millimetre
19.	MREC	Medical Review and Ethics Committee
20.	MVA	Multivariate analysis
21.	NaN_3	Sodium azide
22.	NMR	Nuclear magnetic resonance
23.	OPLS-DA	Orthogonal Partial Least Square-Discriminant Analysis

24.	PCA	Principle Component Analysis
25.	PCR	Polymerase chain reaction
26.	PD	Parkinsons' disease
27.	QoL	Quality of Life
28.	RT-PCR	Reverse transcriptase polymerase chain reaction
29.	SNV	Standard Normal Variate
30.	TSP	3-(trimethysilyl) propionic-2,2,3,3-d_4 acid
31.	VIP	Variable Importance of Projection
32.	WHO	World Health Organisation

LIST OF SYMBOLS

No.	Symbol	Meaning
1.	°C	Degree celcius
2.	K	Kelvin
3.	<	Less than
4.	μM	Micro molar
5.	>	More than
6.	%	Percent

DIAGNOSIS JANGKITAN DENGGI BERDASARKAN ANALISIS METABOLOM (PROFIL BIOPENANDA)

ABSTRAK

Denggi merupakan sejenis penyakit yang menjadi ancaman besar dan mendesak di Malaysia. Penyakit ini semakin banyak menunjukkan manisfestasi yang luar biasa. Manifestasi yang dimaksudkan termasuklah ensefalopati, kegagalan hati, kegagalan buah pinggang, miokarditis, dan rabdomiolisis. Punca utama bagi manisfestasi tersebut adalah disebabkan oleh perubahan metabolisme. Oleh yang demikian, pendekatan metabolomik digunakan untuk menyiasat metabolit yang dapat digunakan sebagai penanda bagi penyakit denggi, yang membawa kepada jangkitan denggi yang teruk. Dalam kajian ini, percubaan untuk menentukan respons metabolisme yang berhubung kait dengan jangkitan denggi telah dibuat dengan menggunakan spektrometri metabolomik yang berasaskan ^{1}H NMR. Kami juga mengkaji perbezaan dari sudut jantina pada mangsa yang terkena jangkitan denggi. Lima puluh dua orang pesakit yang didiagnosis menghidap demam denggi di Hospital Besar Pulau Pinang dan empat puluh tiga orang individu yang sihat terlibat dalam kajian ini. Air kencing aliran tengah yang dikumpul kemudian dianalisis dengan menggunakan spektroskopi ^{1}H NMR, diikuti oleh analisis multivariat kimometrik. Dua jenis pengecaman pola digunakan, iaitu Analisis Komponen Utama (PCA) dan Analisis Diskriminan-Kuasa Dua Terkecil Separa Ortogon (OPLS-DA). Isyarat NMR yang disorot di dalam plot-S OPLS-DA kemudian dipilih dan dikenal pasti dengan menggunakan Pangkala Data Metabolom Manusia (HMDB), *Chenomx Profiler*, dan daripada kajian-kajian yang sudah diterbitkan. Profil metabolisme air kencing bagi pesakit demam denggi didapati berbeza daripada profil metabolisme air kencing bagi individu yang sihat. Model yang amat baik ramalannya dibina daripada profil air kencing bagi pesakit demam denggi melawan profil air kencing bagi individu yang sihat dengan nilai R^2Y (kum.) berjumlah 0.935, dan nilai Q^2Y (kum.) berjumlah 0.832. Kami juga mendapati bahawa pemprofilan air kencing NMR mampu untuk

menunjukkan perbezaan metabolisme jantina manusia. Perbezaan ini penting bagi membezakan kelas-kelas individu yang mempunyai keadaan fisiologi yang serupa, khususnya bagi pesakit demam denggi. Perbezaan yang ketara antara pesakit demam denggi dan individu yang sihat, dan perbezaan yang tidak ketara antara pesakit lelaki (n=39) dan pesakit perempuan (n=13) didapati berhubung kait dengan metabolisme asid amino dan kitaran perantaraan asid trikarboksilik. Jangkitan denggi turut dikaitkan dengan *β*-pengoksidaan asid lemak. Variasi yang jelas dalam beberapa metabolit telah dicatatkan, termasuklah asid amino, pelbagai jenis asid organik, betaina, valerilglisina, mio-inositol, glisina. Biomolekul yang diperhati boleh dijadikan petunjuk bagi mana-mana laluan metabolisme yang terjejas. Oleh yang demikian, biomolekul tersebut berpotensi untuk digunakan sebagai biopenanda bagi mendiagnosis penyakit denggi. Keputusan yang diperoleh membolehkan analisis yang pantas dan tidak invasif, selain membolehkan diagnosis penyakit denggi dengan menggunakan profil metabolisme air kencing. Kajian ini mendedahkan metabolit yang mungkin terlibat dan menjadi punca jangkitan denggi menjadi lebih teruk; jangkitan denggi yang semakin teruk jarang berlaku.

DIAGNOSIS OF DENGUE INFECTION USING METABOLOME ANALYSIS (BIOMARKERS PROFILING)

ABSTRACT

Dengue is a major health and pressing threat to Malaysia with increasing atypical manifestations. These including encephalopathy, hepatic failure, renal failure, myocarditis and rhabdomyolysis which.have been attributed chiefly to metabolic alterations. Therefore, the metabolomics approach was used to investigate for metabolite candidates that potentially could be utilized as an indicator for dengue disease; leads to dengue severity. In this study, attempt was made to determine the metabolic response associate with dengue infection by means of ^1H NMR-based metabolomic spectrometry. We also explored the differences in gender infected with dengue disease. Fifty-two patients diagnosed with dengue fever at Penang General Hospital and fourty-three healthy individuals were recruited in this study. The mid-stream urine collected was analyzed with ^1H NMR spectroscopy, followed by chemometric multivariate analysis. Two types of pattern recognitions were employed, namely, Principle Component Analysis (PCA) and Orthogonal Partial Least Square-Discriminant Analysis (OPLS-DA). NMR signals which highlighted in the OPLS-DA S-plot were further selected and identified using Human Metabolome Database (HMDB), Chenomx Profiler and from published literature. The urinary metabolic profiles of dengue infected patients were distinct from those of healthy individuals. A highly predictive model was constructed from urine profile of dengue infected patients versus healthy individuals with the total R^2Y (cum) value is 0.935, and the total Q^2Y (cum) value is 0.832. We also discovered that NMR urine profiling was able to capture human gender metabolic differences that are important for the distinction of classes of individuals of similar physiological conditions; infected with dengue disease. Distinct differences between dengue infected patients versus healthy individuals and subtle differences in male (n=39) versus female (n=13) infected with dengue disease were found to be related to the metabolism of amino acid and tricarboxylic

acid intermediates cycle. Furthermore dengue infection was linked to β-oxidation of fatty acids. Distinct variations in certain metabolites were recorded including amino acids, various organic acids, betaine, valerylglycine, myo-inositol and glycine. The biomolecules observed are indicative of the possible metabolic pathways affected which may have a potential of being used as biomarker for dengue diagnosis. The results open up a possibility of rapid, non-invasive analysis and diagnosis of dengue disease using urine metabolic profile. This study revealed metabolites that might be involved in development of dengue severity which is rarity.

CHAPTER 1

GENERAL INTRODUCTION

1.1. Introduction

Dengue is an endemic viral disease affecting tropical and subtropical regions around the world (Tripathi, 2010), including Malaysia. At present, no effective antiviral available and pathogenesis underlying dengue disease are yet to be fully elucidated. The dengue virus (DENV) is an RNA flavivirus which has four serotypes numbered DENV1-DENV4 (Ashley, 2011) and the most important vector is *Aedes aegypti* mosquito. DENV has re-emerged as a global major health problem is either epidemic or endemic health threat in many tropical and sub tropical areas including South East Asia, South America and Africa. DENV has infects more than fifty million individuals every year and the incidence is increasing (Ashley, 2011). It is estimated that there are half a million cases per year of severe dengue diseases resulting in 22,000 deaths (WHO, 2009).

In Malaysia, dengue was first documented in 1902 and made reportable in 1971 (Fang *et al.*, 1984). After the first outbreak of dengue hemorrhagic fever (DHF) in 1962 (Rudnick *et al.*, 1965), major dengue outbreaks occurred every four years until 1992; since then the disease became endemic with yearly and frequent outbreaks (Poovaneswari, 1993). Malaysia's reported incidence of dengue has remained high, with an average of 125 to 150 per 100,000 people annually from 2002 to 2006. In 2010, the number of reported dengue cases rose by 12 percent to 45,901 and the number of recorded dengue fever show an increment of 54 percent from 2009 made the total number of fatalities to 134 cases (Ahmad Nizal *et al.*, 2012). However this is probably an underestimate since notification is not compulsory and due to lack of awareness. The histories of the disease are traced over the

years and changes of clinical presentation have been noticed, means the diagnosis is often not confirmed or delayed. Therefore, the actual magnitude of dengue infection in Malaysia might be larger than reported (Muhammad Azami *et al.*, 2011).

The diagnosis of dengue is usually made clinically which sometimes can produce a broad spectrum of symptoms that range from mild febrile illness to severe disease or with atypical manifestation. Clinical features are often nonspecific and therefore require laboratory confirmation (WHO, 2009), especially for surveillance and outbreak investigations. Virus isolation provides the most convincing evidence of infection, but facilities for culture are not always available. Detection of virus-specific RNA, by polymerase chain reaction (PCR), provides accurate diagnosis but requires expensive reagents and equipment. Stringent quality control is necessary to avoid false positive results due to contamination. Serological diagnosis is widely available and can provide an alternative to virus isolation or PCR to support the diagnosis of dengue fever (WHO, 2009). However, serum specimens may be negative for the antibodies if collected too early as antibodies can usually be detected approximately five days after onset of fever. To add, the differing antibodies response patterns between primary and secondary dengue infection underscore the need to evaluate the sensitivity and specificity of commercially available tests (WHO, 2009). Therefore, there is a great demand for the rapid detection and differentiation of dengue virus infection in the early phase of illness in order to provide timely clinical treatment and etiologic investigation and disease control.

Various biomarkers have been designated for the purpose of medical needs with relevant to the disease or the pathway affected (van Gool *et al.*, 2010). They can be used to detect a disease in an individual in the early stages before that disease advances to become a serious illness. Biomarkers can also be used to monitor a patient's response to drugs or in other therapeutic areas. Biomarkers discovery effort might focus on identifying novel biomarkers or validate reported biomarkers in the model systems used. There are several

approaches for identifying, selecting and validating novel biomarkers, such as molecular profiling (van Gool *et al.*, 2010) capturing the genomics, transcriptomics, proteomics and metabolomics technologies.

Metabolomics, as the fourth generation of omics technology, is a growing field in recent era which detects/or quantifies the constituents of the metabolome, resulted from perturbation of the biological pathway that vary according to the physiological, developmental or pathologic state of the cell, tissue, organ or organism (Kim and Maruvada, 2008). These variations can be used to elucidate changes of the metabotype associated to disease-related biochemical reactions which are scarcely available from the other omics techniques. Metabolomics profiling can be applied directly in clinical testing (van Gool *et al.*, 2010). For instance, glucose and cholesterol are metabolomic biomarkers that are part of standard clinical chemistry testing. Consequently, metabolomics, is a desirable tool for diagnosing disease as metabolites are very similar across species and represents an easily translatable and system-wide biomarker discovery approach (van der Greef *et al.*, 2006). A recent surge in metabolomic applications which are probably more accurate than routine clinical practice, dedicated to characterizing the biological fluids. The most common and preferable biofluids used for metabolome investigation are cerebrospinal fluid, saliva, serum and urine (Bouatra *et al.*, 2013).

Urine is the most largely viewed as a waste product, but indeed urine has long been a favored biofluid among metabolomics researchers as it contains enormous amount of information. It is sterile, easy-to-obtain in large volumes, largely free from interfering proteins or lipids and chemically complex. Analysis of urine for medical purposes has dates back to ancient Egypt (Echeverry *et al.*, 2010) and continue to be an important cornerstone to modern medical practices. Urinalysis is routinely performed with dipstick tests that can readily measure urinary glucose, bilirubin, ketone bodies, nitrates, leukocyte esterase, specific gravity, hemoglobin, urobilinogen and protein (Bouatra *et al.*, 2013). More detailed

urinalysis can be also used to study a variety of renal conditions, such as bladder, ovarian and kidney diseases (Pasikanti *et al.*, 2010; Kim *et al.*, 2009; Issaq *et al.*, 2008; Kind *et al.*, 2007). However, no study has been conducted in profiling urine for dengue disease. The current samples used in investigating the metabolites being perturbed in dengue disease including cell culture, serum and analysis of lipidomic (Cui *et al.*, 2013; Birungi *et al.*, 2010). Thus, the discussion in this study was majorly referred to these two published literature with referral to dengue manifestations.

Taking into consideration the steadily increasing rate of dengue infection in Malaysia, metabolomics technologies are expected to be a powerful tool for identifying any disturbances of metabolic processes caused by dengue infection that can reveal a variety of health and disease traits. We have applied a combination of NMR spectrometry and multivariate chemometric analysis to identify potential candidate metabolite biomarkers of dengue infection in the infected individual. This was achieved by carrying out metabolic profiling of urine to detect multibiomarkers in the handling sample. NMR spectroscopy is the choice of metabolomics analysis in terms of its specificity, non-destructive, and at the same time does not require pre-selection of the analysis conditions (Dunn *et al.*, 2005). Urine is a popular biological sample used in metabolomic investigations due to its non-invasive collection and complex metabolic nature of fluid. The chemometric multivariate analysis techniques are used for analyzing and interpreting the complex spectral data sets hence is applicable to deal with large data sets to detect differences between physiologically or pathologically distinct states and is useful to identify potential biomarkers that can improve the understanding of the health and disease processes. Therefore, we would be able to identify possible candidate biomarkers, taking into consideration sample preparation methods and spectral data acquisition.

1.2. Objectives

1.2.1 General objectives:

Dengue virus causes a broad spectrum of illnesses which demand a corressponding shift in research focus towards metabolomics approaches. Metabolomics may have the potentiality to revolutinize the diagnosis of dengue disease to prevent a waste of time and effort in spite of its sensitivity. To add, progression of classical dengue fever to dengue hemorrhagic fever or dengue shock syndrome as the most feared complication of dengue infection with no early prognostic markers of this disorder have been defined. Therefore, discovering a potential multi-biomarkers due to dengue infection is a great demand for the rapid detection of dengue virus infection in order to provide a timely clinical treatment and etiological investigation and disease control.

1.2.2 Specific objectives:

The main goal of this study is to reduce the suffering associated with dengue disease. This goal could be achieved through the following:

1. To determine of the urinary metabolic profiles using ^{1}H NMR metabolomic approaches to discriminate between: A) Dengue infected and non-dengue infected subjects; B) Male and female dengue infected subjects.
2. To investigate of the identified potential multi-biomarkers of dengue disease to better define the underlying pathophysiology.

CHAPTER 2

LITERATURE REVIEW

2.1 Dengue: the global burden disease

At the beginning of 21st century, DF/DHF is the most important and rapidly spreading arboviral disease, occurring in the tropic and subtropic countries of the world where almost half of the world's population was estimated to be at risk of infection. It was estimated that between 50 and 100 million cases of DF and several hundred thousand cases of DHF occur each year, depending on the epidemic activity. The case fatality rate varies among countries, but can be as high as 10-15% in some and <1% in others. In 1998, >1.2 million cases of DF/DHF occurred throughout Asia and the Americas and reported to the WHO indicating as a major epidemics (Gubler, 2002).

The primary vectors *Aedes aegypti* and probably less important *Aedes albopictus* have spread throughout the tropics. Annually, an estimated 50 million dengue infections occurring particularly in South-East Asia, the Americas and the Western Pacific islands with 500,000 severe dengue cases are reported (Guzman *et al.*, 2010). Estimates starting from the 1980s show dengue as an emerging disease with disastrous consequences for people's health and household economy and for society in general. Globally the estimated number of disability-adjusted life years (DALYs) lost to dengue in 2001 was 528 (Cattand *et al.*, 2006). In Puerto Rico, between1984 and 1994, an estimated yearly mean of 580 DALYs per million populations were lost to dengue, similar to the cumulative total of DALYs lost to malaria, tuberculosis,intestinal helminthes and the childhood disease cluster in all of Latin America and Caribbean (Meltzer *et al.*, 1998). In a prospective study for over a five year period of

school children in Northern Thailand, the mean annual burden of dengue was 465.3 DALYs per million, with non-hospitalized patients with dengue illness contributing 44-73% of the total (Anderson *et al.,* 2007). Meanwhile, a two years study, 2005-2006, associated with economic losses due to dengue infection were conducted in eight countries: five in the Americas (Brazil, El Salvador, Guatemala, Panama, and Venezuela) and three in Asia (Cambodia, Malaysia, Thailand) revealed a heavy economic cost to the health system and society (Suaya *et al.,* 2009). Nevertheless, underreporting and misdiagnoses are major obstacles to understanding the full burden of dengue (Suaya *et al.*, 2007), as the number of dengue cases varies extensively from year to year.

This disease is threatening the public health in the developing world since they lack of proper diagnostics and inability to control mosquito populations. To make it worst, at present, there are no vaccines or anti-viral treatments available for DENV infections. To add, secondary infections with any other serotypes (heterologous serotypes) of DENV have been hypothesized to cause DHF/DSS, a severe form of the disease as infection with one serotype of DENV could only provided life-long immunity only to the infecting serotype. Thus, the capacity of the disease and the large numbers of persons with symptomatic dengue infection may give a significant impact on the health care systems of the countries involved and on household and labour economies, especially during epidemics.

2.2 Dengue: the global epidemiology

During the 19th century, dengue was considered a sporadic disease, causing epidemics at long intervals. However, this pattern has changed and ranked as the most important mosquito borne viral disease in the world. The incidence has increased 30-fold for the last fifty-years with significant outbreaks occurring in five of six World Health Organisation (WHO) regions (South-East Asia Region, Western Pacific Region, Americas, African Region and Eastern Mediterranean Region) (Malavige *et al.*, 2004).

The first epidemics of dengue-like disease date back to 1779-1780, when outbreaks occurred in Batavia (Jakarta), Cairo and Philadelphia (Mairuhu *et al.*, 2004). The near simultaneous occurrence of dengue outbreaks on three continents indicates that dengue viruses and their mosquito vectors have had a world wide distribution in the tropics for more than 200 years. Until the Second World War, epidemics occurred on almost all continents every 10-30 years, mainly because both the viruses and the mosquito vectors relied on sailing vessels for transport from countries of the tropics (Mairuhu *et al.*, 1997). After 1945, dengue cases complicated by haemorrhage and shock (DHF and DSS) were increasingly documented in South-East Asia. The first reported DHF epidemic occurred in the Philippine Islands in 1953-1954. Throughout 1970s, the dengue disease were spreading in the region and remained confined in South-East Asia and has become epidemics in other regions of the world in 1980s and 1990s, caused by all four dengue viral serotypes (Nimmanniya, 2002). This situation could be observed in 2001-2002 with similar degree of epidemic activity. As a matter of fact, over half the world's population lives in areas potentially at risk for dengue transmission, making dengue the most important human viral disease transmitted by arthropod vectors in terms of morbidity and mortality (Mairuhu *et al.*, 2004).

2.2.1 Dengue epidemiology in Malaysia

Dengue illness is among the most relevant public health problem in Malaysia. An average of 5000 cases was reported dengue cases annually in the early 1990s (Ministry Of Health Malaysia, 2004). **Figure 2.1** visualized the incidence rate showed an upward trend from 44.3 cases/100,000 population in 1999 to 181 cases/100,000 population in 2007 (Ministry Of Health Malaysia, 2010). Malaysia was reported to have higher case fatality rates (4.67 percent) compared with the neighbouring countries like Thailand and Indonesia, with the case fatality rates of 0.3 percent and 0.5 percent, respectively. Between 2003 and 2005 there were 23, 071 cases reported for persons over the age of 15 years and 66 percent of reported cases in Malaysia are adults, the economically active section of the population (Shepard *et al.*, 2012). A recent study in Malaysia showed that dengue has a considerable impact on social function, vitality and well-being, hence represents a 60% lost in quality of life (QoL) at its worst symptom (Lum *et al.*, 2008).

Malaysia, in fact, has a good laboratory-based surveillance system. However, it is basically a passive system and has a little predictive capability (Gubler, 1998). The serologically confirmed cases are approximately 40-50% of cases at the time of notification. Problem may rise if one waits for laboratory confirmation of the case before notification. Delay in notification may lead to delay in control measure, which will further lead to occurrence of outbreaks, since dengue needs optimum time of management but the transformation of DF into severe form of dengue only takes a very short period (WHO, 1985). In order to bring the disease under an adequate control, novel technologies and approaches are requires.

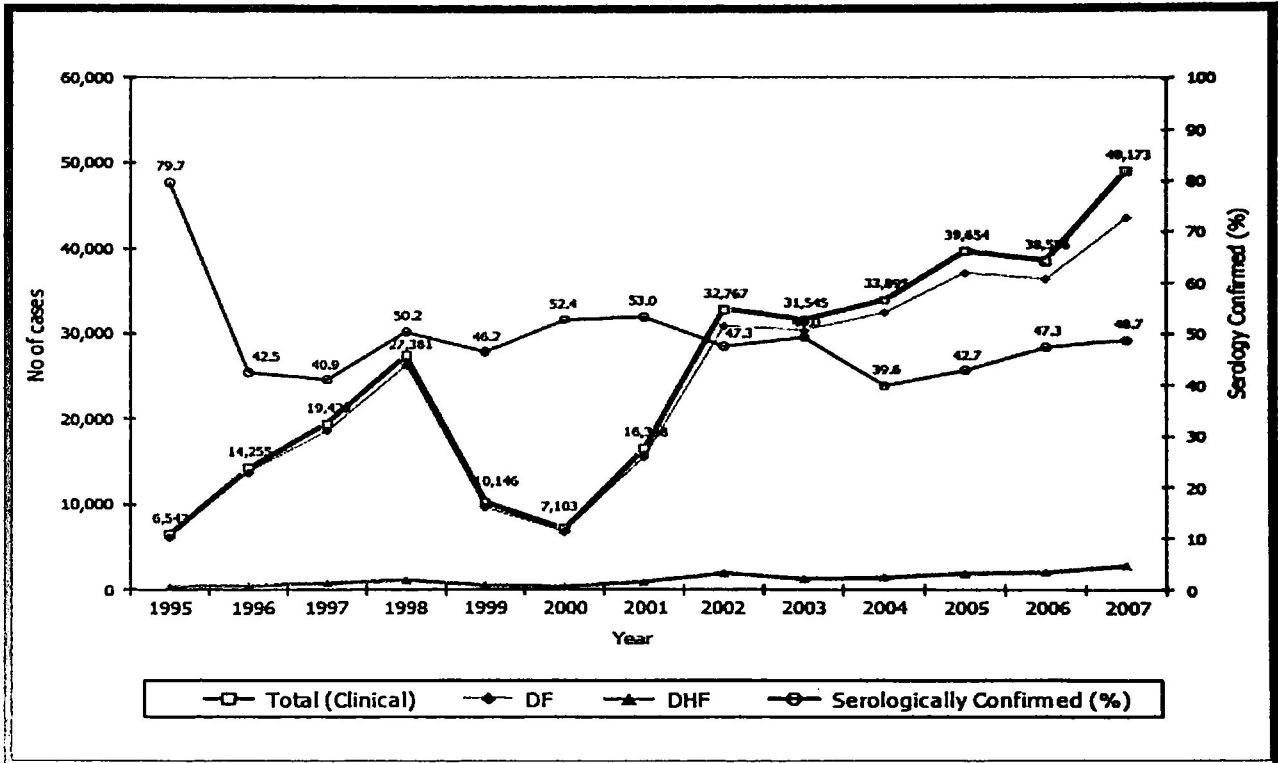

Figure 2.1: Number of dengue cases in Malaysia, 1995-2007 (Ministry of Health, 2010

2.3 Dengue case classification

Dengue was previously categorized using the World Health Organisation classification into dengue, dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS). The case definitions did not always identify patients at risk of severe illness as dengue has a wide spectrum of clinical presentations, often with unpredictable clinical evolution and outcome. Most patients recover following a self-limiting non-severe clinical course, while a small proportion progress to severe dengue disease to DHF or DSS (Balasubramanian *et al.*, 2012). The disease severity progression is difficult to define, but this is an important concern since appropriate treatment may prevent these patients from developing more severe clinical conditions.

2.4 Dengue diagnosis

Since dengue infection may present as a severe disease, characterized by haemorrhage and shock, it is important to make a rapid diagnosis and if possible to identify the serotype involved in the infections. Correct diagnosis and serotype identification assume a greater importance for the severe form of the disease, since there is no marker that predicts the progression from classic dengue fever to the severe forms of the disease (De Paula and da Fonseca, 2004). Currently, dengue diagnosis is based on virus isolation, serology and molecular techniques which have advantages and limitations of the tests **(Table 2.1).**

Table 2.1: Advantages and limitation of dengue diagnostic tests.

Diagnostic tests	Advantages	Limitations
Viral isolation and identification	- Confirmed infection - Specific - Identifies serotypes	- Requires acute sample (0-5 days post onset) - Requires expertise and appropriate facilities - Takes more than one week for the confirmation - Does not differentiate between primary and secondary infection - Expensive
RNA detection	- Confirmed infection - Specific and sensitive - Identifies serotype and genotype - Results in 24-48 hours	- Potential false-positive owing to contamination - Require acute sample (0-5 days post onset) - Requires expertise and expensive laboratory equipment - Does not differentiate between primary and secondary infection
Serological test	- Confirmed infection - Least expensive - Easy to perform	- IgM levels can be low in secondary infection - Confirmation requires two or more serum samples - Can differentiate between primary and secondary infection if samples were taken before day 8-10

2.4.1 Virus isolation

Traditionally, isolation and identification of dengue virus is the only way to diagnose a current dengue infection. Virus can be isolated from blood employing *Aedes* or mammalian cell lines. The intrathoracic inoculation of mosquitoes is the most sensitive system for the isolation of dengue virus, but it needs a particular skill and special containment facilities for direct inoculation of mosquitoes. Isolation of dengue virus provides the most direct and conclusive approach to diagnosis. However, it is useless in early diagnosis because several days are required to isolate and classify the virus (Teles *et al.*, 2005). Therefore, cell culture is preferable for routine diagnosis. Mammalian cell cultures such as Vero cells, LLCMK2 have been employed but with less efficiency (Philippe *et al.*, 2006). This technique is sensitive when there is a presence of high level of infectious particles in the serum, typically starting 2 or 3 days before the onset of fever and lasting until 4 or 5 days of illness. Dengue viraemia is short, thus this technique might be useless if the sample is taken too late or early before the presence of viral infection causing misinterpretation of the infection.

2.4.2 Serological diagnosis

The acquired immune response to infection with dengue virus consists of the production of IgM and IgG antibodies, depending on whether the individual has a primary or a secondary infection. **Figure 2.2** depicts the general time-line of a primary infection from virus isolation/identification to detection of IgM and IgG. A primary infection with dengue is characterized by a slow and low-titre antibody response. IgM antibody is the first immunoglobulin isotype to appear. Anti-dengue IgG at low titre is detectable at the end of the first week of illness, increasing slowly thereafter. High levels of IgG are detectable even in the acute phase and they rise dramatically over the following 2 weeks. The kinetics of the IgM response is more variable. Since IgM levels are significantly lower in secondary dengue infections, some false-negative results in tests for anti-dengue IgM are observed during

secondary infections. IgM antibody-capture enzyme linked immunosorbent assay (MAC-ELISA) has become an important tool in the routine diagnosis of dengue; this technique has a sensitivity and specificity of approximately 90% and 98%, respectively, but only when used 5 or more days after the onset of fever. Serum, blood on filter paper, and saliva, but not urine, can be used for detection of IgM if samples are taken within the appropriate time frame (5 days or more after the onset of fever) (WHO, 2009).

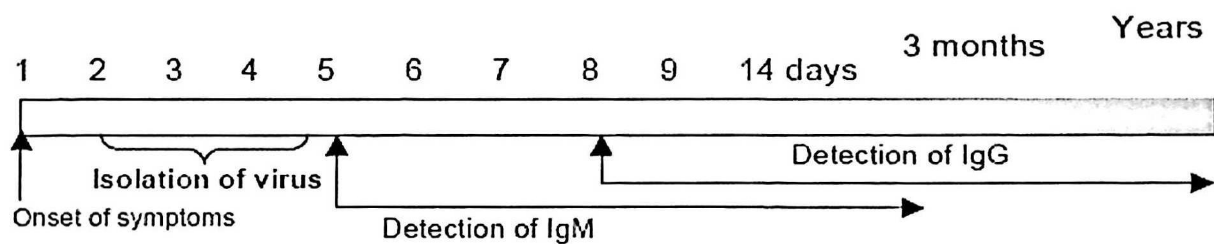

Figure 2.2: General time-line of a primary infection with dengue virus, from identification and isolation of the virus to detection of IgM and IgG (Philippe *et al.*, 2006).

The serological diagnosis of dengue virus infection is rather complicated for the following reasons: (i) patients may have multiple and sequential infections with the four dengue virus serotypes due to a lack of cross-protective neutralization antibodies; (ii) multiple and sequential flavivirus infections make differential diagnosis difficult due to the presence of preexisting antibodies and original antigenic in regions where two or more flaviviruses are cocirculating; (iii) IgG antibodies have high degrees of cross-reactivity to homologous and heterologous flavivirus antigens; and (iv) the serodiagnosis of past, recent, and present dengue virus infections is difficult due to the long persistence of IgG antibodies in many dengue patients with secondary infections (Gubler, 1996; Innis *et al.*, 1989). Thus, among the viral infections that can be diagnosed by serology, dengue virus infection is among the most challenging.

2.4.3 Molecular diagnosis

Molecular diagnosis may assume an important role in dengue diagnosis, since they are able to readily detect the dengue viruses during the acute phase. The various methods are used including nucleic acid hybridization and reverse transcriptase-polymerase chain reaction (RT-PCR) (De Paula and da Fonseca, 2004).

Nucleic acid hybridization specifically a dot blot nucleic acid hybridization test (Henchal *et al.*, 1987) using RNA extracted from dengue virus-infected cell culture supernatants and a pools of infected *A. albopictus* with biotinylated probes or ^{32}P-labelled probes is a sensitive method that has been applied in both diagnostic and epidemiological studies. The test using radiolabelled probes is more sensitive than using biotinylated probes. This RNA technique can only be applied either directly on fresh samples or on retrospective analyses of fixed samples (Khan and Wright, 1987). Therefore, RNA method has been used more often as a research tool rather than a routine diagnostic method (Monath *et al.*, 1989) and it requires an experienced technicians working with it.

Meanwhile, RT-PCR is rapid, sensitive, simple and if correctly standardized, it can be used for genome detection in human clinical samples, biopsies, autopsy tissues or mosquitoes (Deubel *et al.*, 1990). A major challenge for early dengue diagnosis by RT-PCR is the short period available for successful detection of viraemia. Moreover, PCR-based methods need stringent quality control and meticulous sample handling to avoid false-positive results due to unwanted amplification of contaminants. Additionally, serum samples for RT-PCR assays must be stored at extremely low temperature due to intrinsic liability of dengue RNA genome, which is not feasible in many endemic areas (Teles, 2011).

2.5 Metabolomics as a powerful tool for biomarker discovery

Metabolomics is a comprehensive analysis of low molecular weight metabolites in a biological sample, shows a great potential in biomarker discovery, especially in disease diagnosis and pharmaceutical areas. It represents diverse groups of low molecular weight compounds including lipids, amino acids, peptides, nucleic acids and organic acids, vitamins, thiols, carbohydrates and hormones as well. Metabolites alteration may act as an indicator of a disease development in individuals, thus the measurement of metabolites has become an important part of clinical practice (Zhang *et al.*, 2012a). Metabolites can be used to detect a disease in an individual in the early stages before that disease advances to become a serious illness. Meanwhile, traditional markers of conventional clinical chemistry and histopathology method are not region-specific and only increase significantly after serious disease or injury. Particularly, for the early detection of disease, highly sensitive and specific biomarkers as primary indicators in biofluids (blood, urine, saliva etc.) are relatively more useful because they can be used for non-biopsy tests (Zhang *et al.*, 2012a)

Currently, there are no reliable clinical or laboratory indicators that accurately predict the development of classical dengue to DHF. Thus, metabolomics (metabolome analysis) which is representing the metabolite profiles of the cellular processes in a cell, tissue, organ, or organism may be a useful tool for the identification of novel biomarkers in dengue infection capable of acting as an early diagnostic tool.

2.5.1 Biofluids

Monitoring certain metabolite levels in urine sample which is the most commonly used biofluid in metabolomics, has become an important way to detect early stages in disease (Bujak *et al.*, 2011). Ease of collection enables serial sampling to facilitate metabolite identification, quantification and subsequent data analysis. Urine requires easier samples preparation than other biofluids due to its lack of complexity and lower protein content thus making it particularly suitable for metabolomics analysis. Moreover, metabolites in urine as end products of normal and pathologic cellular process are closely linked to phenotypes (Zhang *et al.*, 2012b).

For example, the urinary metabolic signature of celiac disease (CD) was examined by nuclear magnetic resonance (NMR) (Bertini *et al.*, 2009) showing classification accuracy of CD was 69.3% vs healthy control. CD urines showed altered concentration of indoxyl sulfate, meta-hydroxyphenyl propionic acid and phenylacetylglycine. Urine metabolomics has also been used to investigate metabolic signatures in lung cancer (Carrola *et al.*, 2011). Metabolites of interest included β-hydroxyisovalerate, hippurate, α-hydroxyisobutyrate and creatinine.

Recently, there has been interest to identify novel urine biomarkers to diagnose early stages of renal injury. Acute kidney injury (AKI) is an extensively studied clinical problem due to its incidence and high mortality and morbidity. Traditional diagnosis is generally unreliable due to its lack of sensitivity and specificity. Using metabolomics, new AKI biomarkers have been and continue to be discovered (Al-Ismaili *et al.*, 2011). These new biomarkers offer promise for early AKI diagnosis and in assessment of renal injury severity.

The urinary metabolomic method was also applied to the urine profile samples of breast cancer patients and normal persons (Nam *et al.*, 2009). Among nine altered metabolic pathways, four metabolic biomarkers were identified to be different in cancer and normal subjects.

This non-invasive approach has revealed metabolic disorder in biologic systems and has the capability to provide comprehensive information on putative biomarkers for the non-invasive monitoring of disease. The application of urine metabolic analysis to disease is an emerging field and shows great potential in biomarker discovery to facilitate understanding of biochemical mechanisms of pathophysiology for early detection.

Blood maintains a normal homeostasis in the human body by constant regulatory mechanisms, hence metabolic profiling of serum or plasma provides a global view of the instantaneous metabolic status. Moreover, blood perfuses essentially all living cells in the human body and thus is anticipated to carry vital information on virtually every cell. Large scale metabolites of blood plasma are now increasingly gaining attention for their use in the diagnosis of human disease (Gowda *et al.*, 2008).

For instance, Soga and colleagues (2011) had applied serum metabolomics to analyze small metabolites between patients with liver disease from healthy control. They had discovered noninvasive and reliable biomarkers for rapid-screening diagnosis of liver diseases which is more accurate and gain a better understanding of disease mechanisms.

The development of biomarkers for the diagnosis and monitoring disease progression in Parkinson's disease (PD) is of great importance since diagnosis based on clinical parameters has a considerable error rate (Bogdanov *et al.*, 2008). The results showed that serum metabolomics approach appears to be feasible. They were able to separate both medication-free patients, as well as PD patients taking dopaminergic therapies from normal

controls. Thus, metabolomic profiling holds great promise for developing both diagnostic and disease progression markers for PD.

Oral cancer is the eight most common cancers worldwide and represents a significant disease burden. Blood samples of oral cancer patients were analyzed using nuclear magnetic resonance spectroscopy and multivariate chemometric analysis. The results showed a distinct separation between serum samples from cancer patients and from a control group and could also discriminate between different stages of disease. The metabolic profile obtained for oral cancer was significant, even for early stage disease and relatively small tumors. This suggests a systemic metabolic response to cancer, which bears great potential for early diagnosis (Tiziani *et al.*, 2009).

CHAPTER 3

MATERIALS AND METHODS

3.1 Materials

3.1.1 Instrument

All one-dimensional ^1H NMR spectra of the urine samples was acquired on AVANCE III 500 MHz Bruker spectrometer with BBO broadband probe using TSP (δ 0.00ppm) as an internal standard and D_2O as the frequency lock at 300 K. The pulse sequence used included an excitation sculpting routine for the suppression of the water signal (Hicks *et al*, 2009).

3.1.2 Subjects

This study was designed with excluding the effects of race, geography, age, diet intake and sampling time, since they are well-known to contribute to significant variations in the urinary metabolic profiles (Rezzi *et al.*, 2007). Besides, we want to ensure the samples collected were fit with the actual practices.

Sample size is calculated according to the yearly report published by Rancangan Kawalan Penyakit Bawaan Vektor (RKPBV), Penang by using the Power and Sample Size Program (http://www.kck.usm.my/ppsg/stats_resources.htm).

A study of a continuous response variable from independent control and experimental subjects with 1 control(s) per experimental subject. In a previous study the response within each subject group was normally distributed with standard deviation 47. If the true difference in the experimental and control means is 23, 89 experimental subjects and 89 control subjects were need to be able to reject the null hypothesis that the population means of the experimental and control groups are equal with probability (power) 0.9. The Type I error probability associated with this test of this null hypothesis is 0.05.

Ninety-six patients were successfully collected from Penang General Hospital, Penang between February until July 2011 instead of eighty-nine experimental subjects as calculated by Sample Size Program. However only fifty-two patients were serologically confirmed infected with dengue fever, another forty-four patients were IgM negative indicating non-dengue fever. Fifty healthy volunteer individuals were self-identified as healthy without major illnesses and fulfill the listed inclusion and exclusion criteria, recruited from Pangsapuri Pauh Damai (5.371307, 100.420048) and Taman Pauh, Bukit Mertajam, Penang (5.369736,100.422138). However, after initial screening with principle component analysis and consultation of the urine samples revealed the presence of precipitation/cellular debris in seven samples, resulted only forty-three samples of healthy individual were valid to be used for further analysis.

Inclusion criteria include all dengue-infected patients and healthy individual who had voluntarily agreed to participate in this study. While exclusion criteria include 1) menstruated girls; 2) persons who already had or recently had any medical conditions (e.g. surgery, cancer therapy, organ transplantation, serious heart, liver or kidney problems); 3) persons with a recent history of drug or alcohol abuse; 4) pregnant women; 5) persons who have taken any medication (e.g. birth control pills, high or low blood pressure pills); 6) persons who have the genetic or/and auto-immune disease background (e.g. G6PD, thalassemia, leukemia, thyroid, epilepsy).

3.1.3 Biological samples

Urine is a choice of study sample due to its ideal bio-medium for disease study, readily available and less complex than other biofluids, making it suitable for metabolomics versus other biofluids (Zhang *et al.*, 2012). Urine samples were collected from infected dengue patients and healthy volunteers to be used for urinalysis of multibiomarkers in dengue disease.

Fifty-two urine samples were used in this study once the patients have been serologically confirmed with dengue infection and fourty-three urine samples were collected from healthy individuals whom have self-identified as healthy without major illnesses and fulfill the listed inclusion and exclusion criteria. All the samples were collected in the urine container by the subjects. Subjects were asked to fill in the container until half of the container via mid-stream method. All the urine samples both from the patients and healthy volunteers were collected in the morning from 0900 to 1100.

3.1.4 Chemicals

Generally, chemicals used for NMR analysis were of analytical grade and purchased from Sigma-Aldrich (Steinheim, Germany). This including sodium azide (NaN_3), Deuterium oxide (D_2O) and 3-(trimethysilyl) propionic-2, 2, 3, 3-d_4 acid (TSP). Potassium phosphate (KH_2PO_4) was purchased from R & M chemicals (Essex, UK). Potassium hydroxide (KOH) was purchased from Merck (Darmstadt, Germany).

3.2 Methods

3.2.1 Ethical approval

Permission to perform this research has been approved by Medical Review and Ethics Committee (MREC), Ministry of Health Malaysia (Protocol No.: NMRR-10-297-5392). Signed informed consent was obtained from each volunteered patients and healthy individual after a full explanation of the study is provided. All data has been handled confidentially and anonymously.

3.2.2 Sample preparation

3.2.2.1 Sample collection and storage

Urine samples were collected half of the urine container and kept in the cold box prior to be processed. Urine samples were centrifuged at 1500 × g for ten minutes to remove any cellular debris. One milliliter of urine aliquots were transferred to 1.5 milliliter microtubes. The aliquots were stored -80°C prior to NMR spectrometry (Hicks *et al.*, 2009).

3.2.2.2 Sample preparation for 5mm NMR tube

To prepare samples for NMR experiments, 540 microliters aliquot of urine samples were added with 60 microliters of 1.5 M potassium phosphate buffer (Ph 7.4) in D_2O containing 0.1 percent of TSP as frequency reference prior to analysis and 2 mM sodium azide as bacteriostatic reagent. Six hundreds microliters of the mixture was transferred to five millimeter NMR tube (Hicks *et al.*, 2009) prior to be analyzed by proton NMR.

3.2.3 NMR file processing

Proton NMR spectra of urine samples from the subjects were recorded followed by chemometric multivariate analysis. The proton NMR spectra were extracted from δ 0.00 to δ 10.00 and expanded to aliphatic region from δ 0.00-δ 4.70. Presaturation (pre-sat) method was applied for the suppression of the water signal (δ 4.69 ppm-δ 4.97 ppm). The NMR spectra were binned to 0.04 ppm and scaled with Pareto scaling prior to chemometric analysis.

3.2.4 Chemometric analysis

3.2.4.1 Data reduction

The resulting spectra were manually phased and baseline corrected and reduced to ASCII file using Chenomx software (version 5.1, Alberta, Canada). For each spectrum, the spectral region δ 0.52- δ 10.00 was binned into regions of 0.04 ppm width giving a total of 238 integrated regions per NMR spectrum. The signals of δ 4.69- δ 4.97 were excised from the analysis, mainly to eliminate variation in water suppression efficiency peaks. The averaged signals of binned ^1H NMR data from each sub-sample group were subjected to Principle Component Analysis (PCA) and Orthogonal Partial Least Square-Discriminant Analysis (OPLS-DA). PCA and OPLS-DA were performed by SIMCA-P+ version 12.0.1.0 (Umetrics AB, Umeå, Sweden).