# Host Gene Expression and Genetic Susceptibility to Dengue Virus Infection in Vietnamese Children

## Thesis

UNRESTRICTED

# Host Gene Expression and Genetic Susceptibility to Dengue in Vietnam

HOANG TRUONG LONG

A thesis submitted to the Open University (U.K)

for the degree of Doctor of Philosophy in the field of Life and Bio-molecular sciences

May 2010

Oxford University Clinical Research Unit

Hospital for Tropical Diseases

Ho Chi Minh City, Vietnam

DATE OF SUBMISSION : 28 APR 2010

DATE of AWARD : 25 FEB 2011



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

**ABSTRACT**

Dengue has been a public health problem in over 120 tropical and subtropical countries. It is estimated that there are approximately 3.6 billion of the world's population at risk of dengue virus (DENV) infection, and about 70-500 million DENV infections annually. Every year, about 36 million cases of dengue fever, of which about 2.1 million cases are dengue haemorrhagic fever and dengue shock syndrome (DHF/DSS) resulting in 21,000 deaths among children and young adults, are reported [1]. The disease pathogenesis is still poorly understood. Effective antiviral drug, vaccine and animal model for the disease are not available. We conducted two global gene expression microarray studies and a genome-wide association study in order to identify prognostic markers of severe dengue in Vietnamese patients and to identify genetic markers that confer susceptibility to dengue and DSS. The results show that acute DENV infection is associated with over-presentation of transcripts that are related to interferon-inducible transcripts and canonical gene ontology terms that included response to virus, immune response, innate immune response and inflammatory response. Pathway and network analysis identified STAT1, STAT2, STAT3, IRF7, IRF9, IRF1, CEBPB and SP1 as key transcriptional factors mediating the early response. Interestingly, the only differences in the transcriptional signatures of early DSS and uncomplicated dengue cases were the greater abundance of several neutrophil-associated transcripts in patients who progressed to DSS. Host genetic study identified several single nucleotide polymorphisms (SNPs) that show strong evidence for association with dengue or DSS. None of the SNPs was significant at suggestive genome-wide level but ontology analysis of genes implicated by associated SNPs suggests biologically plausible functional relationships between the genes. This thesis provides insights into host gene response in DENV infection and also host genetic susceptibility to dengue which in turn could be useful hint for future studies in the field.

## ACKNOWLEDGEMENT

I would like to express my gratitude to Dr Cameron Simmons and Professor Jeremy Farrar for their constant support and encouragement since I started in research environment. I am especially grateful to my supervisors Dr Martin Hibberd and Dr Sarah Dunstan for their supervision during the research process.

I would like to specially thank Dr Bridget Wills for sharing samples and data that are very necessary for my thesis.

I would like to thank my colleagues at the Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam for their help in the clinical aspect of the work. A special thank you goes to Ms Kien, Ms Quyen, Ms Doan, Huy and Vinh for helping me with important lab work. Thank to all the staffs who worked for the studies at collaborative sites including: Dr Le Thi Phuong, Dr Nguyen Thi Hong Tham, Dr Uyen and Dr Mai Ngoc Lanh at Dong Thap Hospital and doctors and nurses at Hospital for Tropical Diseases and Hung Vuong Hospital. I thank all the participants who agreed to participate to the study.

I am grateful to the collaboration of Dr Martin Hibberd - Genome Institute of Singapore (GIS), Singapore and his staffs including Dr Sonia Davila, Ling Ling, Pauline Aw, Chang Hua, Eileen and Erwin for helping me with the gene expression microarray and the whole genome genotyping and analysis. I also want to thank Dr David Lynn for helping me with the data analysis.

I sincerely appreciate the following institutions for their financial, technical and clinical support: Wellcome Trust of Great Britain; Oxford University Clinical Research Unit and Hospital for Tropical Diseases, Ho Chi Minh City, Vietnam.

Finally, I wish to thank my parents, my sisters and my wife for being by my side all the time.

**ABBREVIATION**

| | |
|---|---|
| 95%CI | 95% Confident Interval |
| ADCC | Antibody-dependent cellular cytotoxicity |
| ADE | Antibody-dependent enhancement |
| ASOs | Allele specific oligos |
| BPI | Bactericidal/permeability-increasing protein |
| BRLMM | Bayesian Robust Linear Model with Mahalanobis distance classifier |
| cDNA | Complementary deoxyribonucleic acid |
| CEU | Utah Residents with Northern and Western European Ancestry |
| CHB | Han Chinese in Beijing |
| Chr | Chromosome |
| DC-SIGN | Dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin |
| DE | Differential Expressed |
| DEF1A | Defensin 1 alpha |
| DENV | Dengue Virus |
| DF | Dengue Fever |
| DHF/DSS | Dengue Haemorrhagic Fever/Dengue Shock Syndrome |
| DM | Dynamic Model |
| DNA | Deoxyribonucleic Acid |
| dNTP | Deoxyribonucleotide Triphosphate |
| EAV | Equine arteritis virus |
| EDTA | Ethylene Diamine Triacetic Acid |
| ELA2 | Elastase 2 |
| ELISA | Enzyme-Linked Immunosorbent Assay |
| FDR | False Discovery Rate |
| FS | Full-sibs |

| | |
|---|---|
| GO | Gene Ontology |
| GTYPE | GeneChip Genotyping Analysis Software |
| GWAS | Genome-Wide Association Study |
| Hct | Hematocrit |
| HLA | Human leukocyte antigen |
| HTD | Hospital For Tropical Diseases |
| HWE | Hardy-Weinberg Equilibrium |
| IBD | Identical By Descent |
| IFN | Interferon |
| IgG | Immunoglobulin G |
| IgM | Immunoglobulin M |
| ISG | Interferon stimulated gene |
| IPA | Ingenuity Pathway Analysis |
| JEV | Japanese Encephalitis Virus |
| JPT | Japanese in Tokyo |
| Kb | Kilobase-pair |
| LD | Linkage Disequilibrium |
| LSO | Locus specific oligo |
| MAC ELISA | The IgM antigen capture ELISA |
| MAF | Minor Allele Frequency |
| MCR | Median call rate |
| MHC | Major histocompatibility complex |
| MPO | Myeloperoxidase |
| mRNA | Messenger Ribonucleic Acid |
| MZ | Monozygous Twins |
| NK cell | Natural Killer Cell |
| NS1 | Non-structural Protein 1 |

| | |
|---|---|
| OD | Optical Density |
| OR | Odds Ratio |
| PBMC | Peripheral Blood Mononuclear Cell |
| PBS | Phosphate Buffer Saline |
| PCA | Principle component analysis |
| PO | Parents-Offspring |
| QC | Quality control |
| RNA | Ribonucleic Acid |
| Rpm | Round per minute |
| RT-PCR | Reverse transcrtiptase Polymerase Chain Reaction |
| SAM | Significance Analysis of Microarray |
| SAPE | Streptavidin Phycoerythin |
| SNP | Single Nucleotide Polymorphism |
| TNF | Tumor Necrosis Factor |
| UTR | Untranslated Region |
| WHO | World Health Organisation |
| YRI | Yoruba in Ibadan, Nigeria |

**Table of contents**

## List of tables

## List of figures

CHAPTER 1

# 1. INTRODUCTION

## 1.1 Dengue

Dengue is an acute disease caused by one of four antigenically distinct serotypes of dengue virus (DENV). It is presently the most common arthropod-borne viral disease of humans. Dengue has been a public health problem in many countries. It is estimated that there are approximately 3.6 billion of the world's population in over 124 nations that are at risk of DENV infection, and about 70-500 million DENV infections annually. Every year, about 36 million cases of dengue fever, of which about 2.1 million cases are dengue haemorrhagic fever and dengue shock syndrome (DHF/DSS) resulting in 21,000 deaths among children and young adults, are reported [1]. Dengue fever (DF) epidemics have occurred not only in Asia, but also in other countries and regions such as the South Pacific, Australia, and Central and South America [3-5]. Figure 1.1 shows the countries and areas that are at risk of dengue. Dengue is characterized by fever, rash, headache, eye pain, arthralgias, myalgias, and haemorrhagic manifestation. These symptoms normally appear within 3-5 days of infection.


Figure 1.1: Countries and areas at risk of dengue transmission [1].

The first well-documented outbreak of DHF took place in Manila in 1953/1954 and was followed by a larger outbreak in Bangkok in 1958 [6]. During the rest of the 20th century, the disease has spread widely to many countries across Southeast Asia and America, increasing both the number of countries affected and case numbers [7, 8]. The number of countries reporting dengue outbreaks has been increasing yearly, from 35 countries from 1990-1999 to nearly 70 countries from 2000 to 2005 [9]. Figure 1.2 shows global burden of dengue in terms of case prevalence and number of countries affected.

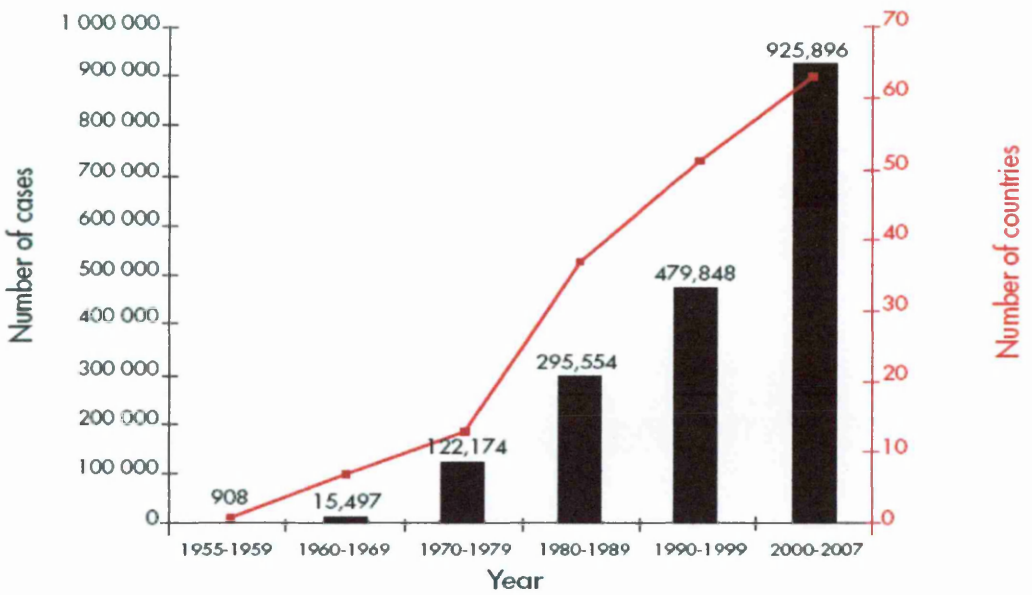



Figure 1.2: The global burden of dengue in terms of case prevalence and number of countries affected. Average yearly number of dengue cases is depicted by bars, and the number of countries reporting dengue is shown in the line.

## 1.2 Dengue epidemiology in Vietnam

### 1.2.1 Historical information

A dengue-like illness was first recorded in Vietnam in 1913 and epidemics occurred in the north and the central provinces [10]. The first DHF outbreaks occurred in 1963. After 1975, many district and commune level health posts were established; doctors and nurses received improved training and acquired experience in diagnosing and treating dengue and

consequently the mortality rate has declined significantly [11]. These efforts, however, could not prevent the geographic expansion of dengue epidemics throughout Vietnam.

### 1.2.2 Dengue epidemiology in the south of Vietnam since 1998

Dengue has been a public health problem in Vietnam for over 30 years with tens of thousands of cases reported every year. The majority of hospitalised cases of dengue occur in the 20 southern provinces. Figure 1.3 summarises the morbidity and mortality of dengue in the last 10 years from 1998 to 2007. The mortality rate has declined from approximately 0.3% to less than 0.1% during this period. The disease is reported every month of the year but the peak of transmission occurs during the rainy season (from May to November). In 2007, patients with DHF or DSS represented 85% of all hospitalised dengue cases reported with the majority of them being children less than 15 years of age [12]. All four serotypes of DENV were found to co-circulate in Vietnam but each serotype has been dominant for different periods of time. DENV3 was dominant in the epidemic in 1998 but DENV2 was the main serotype in 2003-2006 period followed by DENV1 from 2007 until now.

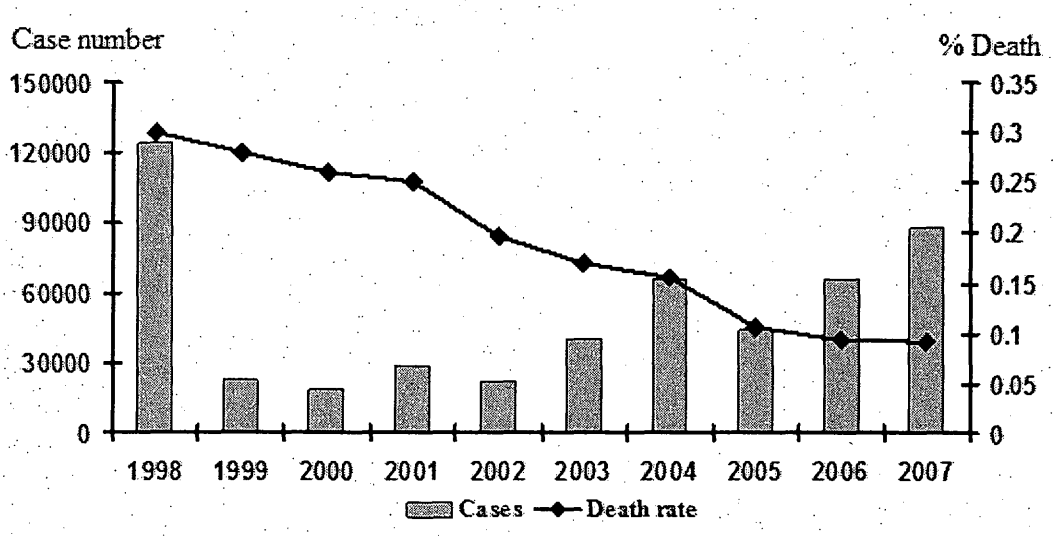



Figure 1.3: Dengue case burden and mortality rates in the southern Vietnam from 1998 to 2007 [12]. The grey bars represent case number by year and the line represents the mortality rate.

### 1.3 Dengue vector

Mosquitoes were demonstrated as vectors of DENV through epidemiologic and experimental observations by Graham in 1903 [13]. Up to now, all the known DENV vectors belong to the genus Aedes, particularly *A. aegypti*- the principal vector, and *A. albopictus*- the secondary vector [13-15]. Mosquitoes belonging to the genus *Aedes* have African, Asian, or Australian origins. Most of them breed in the forest, in tree holes or other natural water containers. The male mosquitoes only suck the juice from plants while the females need a blood meal to lay eggs. Female mosquitoes feed on a variety of animals, monkeys and occasionally humans. In contrast, *A. aegypti* and *albopictus* have adapted to live in peri-urban and urban environments with humans as a major feeding target. Their larvae can be found in artificial and natural water containers indoors or outdoors. Compared to *A. albopictus*, *A. aegypti* is more domesticated [13]. *A. aegypti* prefer to lay their eggs into clean water, where their eggs hatch into larvae within 48 hours. The larvae become pupae – a short resting, non-feeding period before emerging as adult mosquitoes.

This cycle takes approximately 7-14 days. The mosquito lives for 15 days on average though it can be as long as 174 days [13]. *A. aegypti* usually feeds in mid-morning and late afternoon biting 2-3 times for a single blood meal [16]. These feeding habits of *Ae. aegypti* assist in the transmission of DENV, with multiple persons living in the same household at risk of exposure. Although the dispersal of adult mosquitoes from their place of hatching is reported as limited [16], mosquitoes infected with DENV can appear in another distant place or another continent, consistent with movement of people by train, bus, boat or airplane. Persons with dengue may introduce the virus to a new population when they go travelling if suitable vectors are available.

Ambient temperature is considered a critical factor in DENV transmission. It influences the mosquitoes' breeding and their lifespan, and affects the dynamics of virus replication in vectors such that higher temperatures favour more rapid dissemination of DENV in the mosquitoes' body [17]. The density of *A. aegypti* additionally depends on artificial breeding sites because of their domesticated characteristics. After DENV infection, mosquitoes retain the capacity to transmit DENV for life [18]. Evidence for vertical transmission of DENV has been collected in the laboratory and to a limited extent in the field [19, 20], but its wider importance in the epidemiology of dengue is poorly understood.

### 1.4 Dengue case classification [1]

With a wide spectrum of clinical presentations, the outcomes and clinical evolution of dengue are unpredictable. While the majority of patients recover from dengue, a small proportion progress to severe disease which is characterized by plasma leakage with or without haemorrhagic manifestations. Appropriate treatment could help to prevent patients from progressing from non-severe to severe disease. However, to define the group of patients who progress to severe disease is very difficult. WHO classify dengue cases into 2

main groups - severe dengue and mild dengue, which in turn is subdivided into 2 groups – dengue with and without warning signs. Figure 1.4 summarises dengue case classification by WHO.

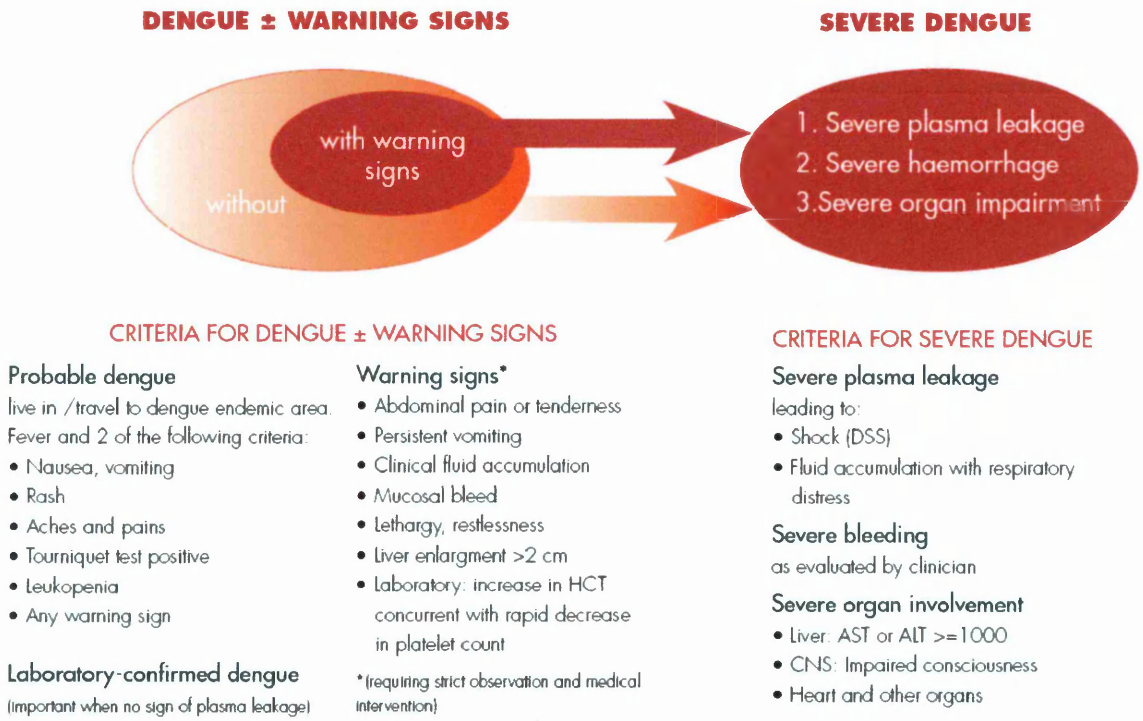



Figure 1.4: WHO Dengue case classification. WHO classify dengue cases into 2 main groups - severe dengue and mild dengue, which in turn is subdivided into 2 groups – dengue with and without warning signs. [1]

After the incubation period of 4 to 10 days after the bite of an infected mosquito, the disease begins abruptly [1]. In the febrile phase, which usually lasts 2-3 days, patients often present themselves with facial flushing, skin erythema, body ache, myalgia, arthralgia and headache, anorexia, nausea and vomiting. Mild haemorrhagic manifestations may also be observed including petechia and mucosal membrane bleeding (e.g. nose and gums). Laboratory findings in this phase will show a progressive decrease in total white cell and platelet count. The critical phase is normally around the time of defervescence (when the body temperature drops to 37.5 -38$^{0}$C) on days 3-7 of illness. The beginning of this phase is associated with a measurable increase in capillary permeability often revealed

by an increased haematocrit level. Plural effusion and acites may be detectable depending on the degree of plasma leakage. When plasma leakage is critical, patients can develop shock. During prolonged shock, progressive organ impairment, metabolic acidosis and disseminated intravascular coagulation may occur as the result of organ hypoperfusion. Organ impairment such as severe hepatitis and encephalitis may also occur without severe plasma leakage or shock. The recovery phase occurs 24 - 48 hours after critical phase and is characterised by the return of laboratory findings to a normal range and clinical improvement. These phases of dengue illness are depicted in figure 1.5.

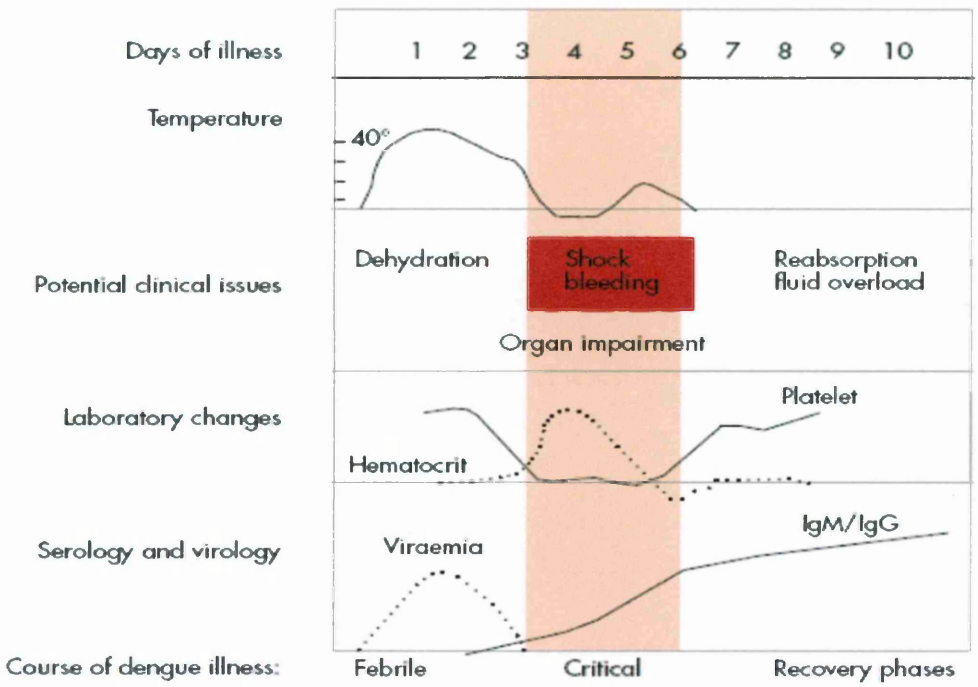



Figure 1.5: The course of dengue illness. There are three main stages of the disease which are febrile phase, critical phase and recovery phase. The febrile phase normally occurs in the first 72 hours which is followed by the critical phase in the next 72 hours.

### 1.5 Clinical management

There are no specific anti-viral therapies available to treat dengue. Supportive care is the mainstay of dengue treatment. Using analgesics with acetaminophen is recommended, but not aspirin to avoid Reye's syndrome associated with dengue. Early plasma volume replacement by oral fluid or fluid transfusion (Ringer's lactate or dextran 40) can modify the severity and prevent shock. Dengue mortality can be reduced with regular monitoring of vital signs, careful fluid management when required, and good nursing care. [21]. Recently, several randomised control trials have been conducted to evaluate the efficiency of the immunomodulators chloroquine [22] and predsinolone (unpublished data) and of the first antiviral drug (a polymerase inhibitor, developed for the treatment of Hepatitis C virus infection) in the treatment of dengue. This indicates that there is a research interest and that are developments in the treatment of dengue, and that treatment for dengue other than supportive care may become available in the near future.

### 1.6 Laboratory diagnosis of DENV infection

Dengue can be diagnosed by different methods which include serology, virology and molecular techniques. A patient can be laboratory confirmed with dengue if he/she has dengue-like clinical symptoms and the presence of a) IgM seroconversion or high and rising levels of DENV-reactive IgM/IgG in paired plasma specimens or b) DENV viremia detected by virus isolation or RT-PCR or c) NS1 antigenemia.

#### 1.6.1 MAC ELISA

The IgM antigen capture (MAC) ELISA is a sensitive and relatively specific method used for detecting DENV-reactive IgM in serum or plasma [21]. The antigens used in this assay are usually inactivated viral particles [23]. Acute dengue is defined by rising levels of IgM antibodies and/or by fourfold or greater increase in IgG in paired sera/plasma samples [21]. Primary and secondary dengue can be defined by the molar ratio of IgM to IgG ratio in

acute illness [23, 24]. In this context, the term "secondary dengue" describes the serological profile of the immune response rather than definitively determining that this was the patient's 2nd DENV infection, since it could also be the 3rd or even 4th exposure [25]. The advantage of the MAC ELISA is that it is relatively inexpensive and simple [23, 24] . The main limitation of serology for dengue is that it is poorly sensitive in the first few days of illness and in places where multiple antigenically-related flaviviruses co-circulate appropriate specificity controls must be included in each assay. The WHO/TDR recently reviewed the performance of several commercial assays for the detection of dengue IgM. The main findings were that the test sensitivities (61.5% to 99%) and specificities (79.9% to 97.8%) vary in different commercial tests [26]

### 1.6.2 DENV nucleic acid detection

Dengue can be diagnosed by detecting DENV RNA in tissue or plasma samples using reverse-transcriptase polymerase chain reaction (RT-PCR) methods [27, 28]. Infecting serotypes of DENV can be detected based on the differences in amplicon sizes [28]. Alternatively, real time RT-PCR using specific primers and fluorescent probes can be used [29, 30]. Primers and probes for DENV detection have targeted conserved sequences in the C-prM junction or NS5 region or the 3' non-coding region [28, 31, 32]. RT-PCR can detect DENV RNA in the early days of illness when the patient is still viraemic [21, 33]. However, it is not suitable for routine diagnosis in all settings because it is expensive in terms of specialized training, equipment and reagents.

### 1.6.3 DENV NS1 protein detection

DENV NS1 can be detected in peripheral blood in the acute phase and has been detected until the 9$^{th}$ day of illness, even when DENV RNA could not be detected by polymerase chain reaction (PCR) [34]. In general, the specificity of NS1 commercial tests is high (~100%) but the sensitivity has varied in different studies. The sensitivity of NS1 tests also

depends on the day of illness the sample was collected [35, 36]. The highest sensitivity was documented within the first three days of illness [36]. Primary dengue is usually associated with higher NS1 detection sensitivity [36, 37] relative to secondary dengue as measurable DENV-reactive IgG/M in plasma may reduce NS1 sensitivity in secondary dengue infection [36, 37]. Combining NS1 and MAC ELISA diagnostic tests can increase the positive rate of diagnosis in the acute phase of illness [38, 39] .

### 1.6.4 Virus isolation in cell culture

DENV from serum, plasma, blood or tissue is isolated by inoculating the clinical specimen into suitable cell lines followed by identification with serotype-specific anti-DENV monoclonal antibodies (mAb) by immuno-fluorescence assay [21]. A mosquito cell line (e.g. C6/36) is most commonly used because of greater sensitivity than mammalian cells (e.g. Vero, BHK) [40, 41]. Although viral isolation gives a definitive diagnosis of dengue, it is not suitable for routine diagnostic laboratories because it is time consuming (taking > 5 days for cultures to become positive) [40, 42].

## 1.7 Dengue virus

There are four antigenically distinct serotypes of dengue virus (DENV1-4) that have been identified so far. Dengue virus belongs to genus Flavivirus within the *Flaviviridae* family. Flavivirus genus consists of about 70 *flaviviruses*, many of which are arthropod-borne human pathogens causing a variety of diseases, such as West Nile virus, Japanese Encephalitis virus, and Yellow fever virus [43]. DENV is a small, enveloped virus with a diameter of about 50 nanometres. The virus is a single-stranded genomic RNA of positive polarity approximately 11kb in length [44]. The genomic RNA of the virus is surrounded by an icosahedral or isometric nucleocapsid about 30nm in diameter (review [45]). The nucleocapsid is surrounded by a lipid bilayer about 10nm deep into which the viral structural protein M and the envelop glycoprotein (E) are embedded [46]. The genome of

DENV contains a 5' untranslated region (UTR; 100 nucleotides), a single open reading frame encoding for the viral polyprotein and a 3' untranslated region (3'-UTR; ~400 nucleotides). The 5' and 3' UTR are very important in both viral replication or translation [47-50]. The open reading frame of DENV encodes for a polyprotein , which is co- and post-translationally processed into three structural proteins (C, M and E) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) [45]. The highly basic 12 -14 kDa capsid (C) protein forms the viral nucleocapsid together with the genomic RNA. This protein is rich in lysine and arginine residues and contains 114 amino acids. During infection, the C protein localises to the nuclei of infected cells [51]; the premature M (PrM), a glycoprotein precursor (18.1 -19kDa, 166 amino acid), prevents the E protein undergoing low-pH induced conformational changes prior to viral maturation by forming prM – E heterodimers [52]. E protein (55 – 60 kDa, 493 – 495 amino acid) which is exposed on the surface of the virion, is associated with virus binding and fusion of viral and cellular membranes during receptor-mediated endocytosis [53, 54]. Most of the non-structural proteins are involved in viral replication. NS5 is a RNA-dependent RNA polymerase and carries an additional N-terminal methyltransferase domain important for viral replication [55]. NS3 acts as the viral serine protease, which requires the cofactor NS2B for full activity. Furthermore, NS3 contains an RNA helicase and nucleotide triphosphatase activity important for replication of the viral RNA [56, 57]. Functions of the other proteins (NS1, NS2A, NS2B, NS4A and NS4B) are poorly understood. NS1 protein (46-50 kDa, 352 amino acid) is secreted by infected cells or associated with the infected cell membrane [58, 59]. NS1 is also assumed to act at early stage in viral RNA replication and may also be important for the pathogenesis of DHF [60-62].

## 1.8 Immunopathogenesis in DENV infections

### 1.8.1 Humoral response in DENV infection

DENV-reactive antibodies may be involved in virus neutralization, complement lysis, and antibody-dependent cellular cytotoxicity (ADCC). Antibodies neutralize the virus through inhibiting viral fusion and viral membrane attachment [63]. Antibodies against E proteins are elicited in the early phase of illness and are amongst the most abundant populations of antibodies following DENV infection [64]. Animal models have revealed that passive transfer of antibodies against structural and non-structural proteins may protect against lethal DENV challenge [65-67]. Antibody can have viral neutralizing activity through inhibiting viral fusion and viral membrane attachment [63] and can involve in antibody-mediated cell lysis through ADCC, a mechanism through which DENV infected cells are recognized and lysed by natural killer cells with promotion by specific antibody and complement. ADCC may be associated with dengue pathogenesis because this mechanism was observed in acute sera from DHF/DSS patients but not from DF patients [68]. ADCC has been postulated to be a host immune mechanism of controlling DENV infection in the very early phase of disease [69].

### 1.8.2 Cellular immune responses in dengue

The innate immune response plays a very important role in antiviral immunity to DENV in both primary and secondary dengue. The role of ADCC in dengue pathogenesis is not well understood but early activation of NK cells was found in DHF/DSS patients [70-72]. Moreover, the percentage of CD8 T cells and NK cells expressing CD69, a transient cell activation marker, was higher in children with DHF than in those with DF [70].

Specific T cell proliferation against DENV and other *flaviviruses* was observed when healthy adult volunteers received a monovalent live-attenuated candidate vaccine. CD8+ cytotoxic T cells can recognise DENV NS1, NS3 and E protein [73]. DENV-reactive

CD4+/ CD8+ T cells can produce IFN-γ, TNF and chemokines that were efficient at lysing DENV-infected cells in-vitro [74]. Memory serotype cross-reactive T cells were hypothesized to contribute to the pathogenesis in secondary infection by producing high levels of inflammatory cytokines during the acute phase [75, 76].

### 1.8.3 Cytokine and Chemokine responses in dengue

Host gene-expression studies of acute dengue patients indicated that cytokine-related genes were up-regulated in patients with severe dengue relative to patients with mild dengue [77, 78]. In-vitro and ex-vivo studies showed that DENV infection can induce the secretion of several cytokines and chemokines such as tumor necrosis factor alpha (TNFα), interferon (IFNγ), interleukin-6/8/10 (IL-6, IL-8, and IL-10), or RANTES [79-81]. The overproduction of pro-inflammatory and anti-inflammatory cytokines was also observed in Vietnamese infants with dengue [82].

Elevated levels of cytokines and chemokines during secondary infection by infected and activated mononuclear phagocytes and T cells, and driven by the viral antigenic burden, can partly explain the vascular leakage phenomenon, the hall mark of severe dengue. In primary dengue in infants, memory DENV-reactive T cells are not relevant to the immediate host response. Nonetheless, elevated levels of inflammatory cytokines have been described in Vietnamese infants with primary dengue [82]. Molecules such as MCP-1 and TNFα, which are elevated in DHF/DSS cases, have well characterized vasodilatory effects in animal models and endothelial cell monolayers [83, 84]. TNFα enhanced endothelial cells to produce reactive nitrogen and oxygen species which may then induce the haemorrhagic manifestation [85]. Interestingly, an in-vitro study has revealed that the chemokine IP-10 competes with DENV in binding heparan sulphate on the cell surface. IP-10 is also abundantly expressed during acute DENV infection [86, 87]. Generally, there is

a widely held view that cytokines and chemokines contribute to the immunopathogenesis of dengue.

## 1.9 Risk factors for severe dengue

### 1.9.1 Viremia

Higher peak viremia titres were observed in children with DHF relative to children with DF, and in secondary infections relative to primary infections [33, 88]. Higher viremia was also observed in adults with DHF relative to those with DF. Furthermore, viremia was demonstrated to persist longer in DHF patients than in DF patients [89]. The kinetics of DENV infection prior to disease onset remains poorly characterised and very difficult to study. Studies in Vietnamese patients with DSS showed that the viremia declined to a very low level at the time of shock (B. Wills personal communication). A majority of studies have described viremia in plasma so that the results may not be representative of infection in other organs of the body [90, 91].

### 1.9.2 Virulence

Dengue virus serotypes are hypothesised to have evolved independently [92]. All four serotypes can cause DHF/DSS but DENV2 is more frequently associated with severe dengue [93-95] followed by DENV1 and DENV3 [96-98] and DENV4 [99]. There are many different genotypes and strains within each genotype, each of which may contribute differently to disease outcomes. For example, the DENV2 American genotype was demonstrated to be associated with DF and very few DHF/DSS cases in Peru during an epidemic in 1995 where the majority of patients had secondary dengue [100]. In comparison, the 1981 epidemic of DENV2 Asian 2 genotype in Cuba resulted in about 300,000 infections and around 30,000 cases of DHF/DSS with 158 deaths [101, 102] . This suggested that certain serotypes or genotypes may be more virulent than the others. This may be because certain serotypes or genotypes have high replication fitness [103]. ADE

may also play a role in this phenomenon as it provides competitive advantages to viruses that undergo enhancement through increasing viral replication [104]. Natural positive selection may also contribute to the fitness of the DENV as it was shown that this selection was different between serotypes [105]. However, the virulence of DENV is very hard to assess because there is no good animal model of DHF/DSS.

### 1.9.3 Antibody-dependent enhancement (ADE)

Antibody-dependent enhancement (ADE) is a phenomenon in which non-neutralizing, cross-reactive antibodies elicited by a primary DENV infection, or acquired passively by vertical transmission, enhance the infectivity of DENV for Fc receptor-bearing cells [106]. This model postulates that sub-neutralizing levels of DENV-reactive IgG enhance the infectivity of virus for Fc receptor bearing cells. This process would then theoretically result in increased viral burdens and more severe clinical symptoms. The ADE phenomenon is well established in vitro and was first described by Halstead et al in 1977 [107]. The target cells in these experiments were usually primate cell lines expressing Fc receptors on their surfaces. Incubating *Flavivirus* antisera with DENV2 could enhance viral infection in human peripheral blood mononuclear phagocytes [108]. ADE activity has also been observed in Fc-receptor positive K562, U937, and Raji-1 cells [109]. ADE activity has been documented in rhesus monkeys which had been passively transferred diluted human anti-DENV immune serum. Subsequent DENV infection of these monkeys showed the ratio of infection enhancement ranged from 2.7 to 51.4 [110]. Similarly, a significantly greater vireamia titre was observed in monkeys which received non-neutralising of monoclonal anti-DENV antibody compared with those receiving PBS [109].

The evidence to support the ADE hypotheses has been mostly inferred from epidemiological data that demonstrates secondary infection is a risk factor for severe

disease. In one prospective study of Thai children, the DENV infection enhancing activity of pre-infection sera for primary human monocytes positively correlated with clinical outcome, i.e. children with pre-illness sera that showed high levels of infection-enhancing activity developed more severe disease [111]. However, this study has not been replicated. In a more recent study of Thai children, there was no correlation between the DENV-infection enhancing activity of pre-infection sera for K562 cells and the clinical phenotype of these children when they were exposed to DENV [112]. In infants, Kliks et al showed a correlation between the age of disease onset in Thai infants (n=13) and the dilution at which cord sera maximally enhanced infection of primary monocytes [113]. It was also shown that infants at 6-9 months old are more vulnerable to DHF/DSS. This may be because of the decline of maternal-derived anti-DENV antibody to a subneutralising level that enhances infectivity of DENV [114]. However, ADE cannot solely explain disease pathogenesis as there are still DHF/DSS cases with primary DENV infections [93] and only a small proportion of all secondary dengue infections develop DHF/DSS.

### 1.9.4 Dengue and genetic predisposition

Genetic polymorphisms include single nucleotide polymorphisms (SNPs), gene copy number variants and genetic deletions. These subtle changes might very well have important consequences for susceptibility to disease [115]. Genome-wide association studies have been used to identify genetic variations that confer susceptibility to viral diseases such as HIV and HCV infection. A polymorphism near the *IL28B* gene was shown to be associated with a two-fold change in response to treatment in both the European and African population [116] and with better clearance of the virus in the same populations [117]. A polymorphism in the IL28B gene was also demonstrated to be associated with the response to treatment in the Japanese population [118]. Polymorphisms in the *HLA-DP* gene were also shown to be associated with susceptibility to HBV infection in Han-Chinese and Thai people [119]. In DENV infection, indirect evidence implicating

host genetic background to DHF/DSS was observed in Cuban dengue epidemics whereby a reduced risk for DHF/DSS was observed in those with an African ancestry compared to those with European ancestry [94, 102, 120-124]. These Cuban observations are of significant epidemiological interest, as the differences in susceptibility to DHF among racial groups in Cuba coincide with the low susceptibility reported in African and Black Caribbean populations [125, 126].

Although dengue virus has been repeatedly isolated in Africa, and DF is known to be presented in 19 countries in this continent, there are only sporadic reports of DHF cases [125, 127, 128]. An additional observation is that in Haiti, despite the presence of viral risk factors, no DHF/DSS cases have been reported [126]. The dengue epidemiological observations related to ethnicity in Africa, Cuba and Haiti could indicate the sharing of a common genetic background that moderates the clinical outcome of DENV infection in individuals who have an African genetic background. Bearing in mind the central role of immunological mechanisms in the pathogenesis of DHF, the genes associated with the immune response must be considered carefully in the context of human genes regulating dengue disease severity, which might be distributed unequally in Black individuals and White individuals.

HLA genes have been among the most studied polymorphic genes in dengue. Early case-control studies of serologically determined HLA class I alleles in DENV-infected patients were performed in Cuba during the 1981 DENV2 outbreak. A positive association of HLA-A1, HLA B and HLA CW1 with DHF was shown [129]. Other serological studies of HLA class I alleles have been performed in ethnically and geographically distinct populations in Thailand, suggesting an increase in the frequency of class I serotype HLA-A2 in DHF patients [130].

The study of the HLA association with dengue illness was reconsidered again in 2001 by Hsin et al. using molecular analysis of HLA class I allele profiles in Vietnamese patients. These authors demonstrated evidence for association of alleles of the HLA-A locus with susceptibility to DHF [131]. More recently, a second and larger case-control study in ethnic Thai cases confirmed the association of classical HLA class I alleles (A2, A*0207, B46, B51) with the clinical outcome, in previously dengue immunologically primed individuals [132].

A molecular approach to HLA case-control studies was reconsidered in Cuba, which represents a unique scenario to perform genetic studies of DHF resistance/susceptibility [121, 122, 133]. A significantly increased frequency of HLA I alleles A*31 and B*15 was found in Cuban individuals with a history of symptomatic DENV infection compared to controls. HLA II alleles DRB1*07 and DRB1*04, on the other hand, showed an elevated frequency in controls compared with dengue cases. Genes in the class III region encode a number of proteins, including complement proteins (C4A, C4B, C2), TNFα and TNFβ and heat shock proteins [134]. Hsin et al. studied promoter polymorphisms in the TNFα gene but did not find an association with DHF, and Fernandez-Mestre et al. reported a significant increase of the TNF-308A allele in patients with DHF [135]. Table 1.1 summarises the association of HLA polymorphisms with the outcomes of dengue.

The number of studies on polymorphisms within non-HLA genes remains low. Hsin and colleagues investigated the association between susceptibility to DHF and polymorphic non-HLA alleles such as vitamin D receptor (VDR), Fcγ receptor II (FcγRII), IL-4), IL-1RA, and mannose-binding lectin [136]. They found that variant at position 352 of the vitamin D receptor (VDR) gene was associated with resistance to severe dengue ($P = 0.03$) and variant at position 131 of Fcγ receptor II may be associated with protection from DHF.

It has been reported that there is a stronger association of the allelic variant of a DC-SIGN1 codifying gene CD209, DCSIGN1-336, and the risk of DF compared to DHF or population controls. DC-SIGN1 is a dendritic cell–specific ICAM-3 grabbing non-integrin which are essential for productive infection of dendritic cells [137].

Recently associations between TAP and HPA gene polymorphism and dengue were described, suggesting that the heterozygous pattern at the TAP1 333 locus and HPA1a/1a and HPA2a/2b genotypes confer susceptibility to DHF. The HPA1a/1b genotype was determined to be a genetic risk factor for DSS [138].

As the aforementioned studies demonstrate, investigations into the genetic basis of susceptibility to dengue have made enormous strides since the completion of the human genome sequencing project. Nevertheless, this field of research is dominated by case-control association studies of small sample sizes, uncertain case ascertainment and with unknown levels of population stratification (genetic admixture in one population that obscures any meaningful study of genetic association). Typically, these studies examined polymorphisms in a limited number of candidate genes for which there was some a priori rationale for their investigation. For the vast majority of case-control association studies, there was never an attempt to replicate the genetic association in either the same or different population.

Improvements in technology that permit high-throughput genotyping of genetic polymorphisms have allowed a genome-wide approach to investigating host genetic susceptibility. Essential to this approach are large sample sizes in order to make statistical correction of the number of comparisons being made. Replication of the genetic association in a second population is now considered an essential validation step.

Table 1.1: Summary of association of HLA genes with outcome of dengue [2].

| HLA of Interest | | Case (*n*) vs Control[a] (*n*) | freq. in case | freq. in control | *p* value | Country | Citation |
|---|---|---|---|---|---|---|---|
| *HLA class I* | | | | | | | |
| HLA-A | | | | | | | |
| HLA-A1 | | $1^0$ DHF (8) vs control (138) | 0.25 | 0.03 | 0.035 | Thailand | Chiewsilp et al., 1981 [57] |
| | | DSS (87) vs control (276) | 0.31 | 0.08 | $<10^{-4}$ | Cuba | Paradoa-Perez et al., 1987 [58] |
| HLA-A2 | | $2^0$ DSS (41) vs control (138) | 0.68 | 0.51 | 0.047 | Thailand | Chiewsilp et al., 1981 |
| | | $2^0$ DF (106) vs control (140) | 0.57 | 0.34 | 0.0005 | Thailand | Stephens et al., 2002 [59] |
| | | $2^0$ DHF (103) vs control (140) | 0.51 | 0.34 | 0.0076 | Thailand | Stephens et al. |
| | HLA-A*0203 | $2^0$ DF (106) vs control (140) | 0.34 | 0.19 | 0.00048 | Thailand | Stephens et al. |
| | HLA-A*0203 | $2^0$ DHF (103) vs $2^0$ DF (106) | 0.17 | 0.34 | 0.01 | Thailand | Stephens et al. |
| | HLA-A*0207 | $2^0$ DHF (103) vs control (140) | 0.28 | 0.14 | 0.012 | Thailand | Stephens et al. |
| HLA-A*24 | | DHF (309) vs control (251) | 0.35 | 0.259 | 0.021 | Vietnam | Loke et al., 2001 [60] |
| HLA-A*24 | | DSS (53) vs control (62) | 0.35 | 0.25 | 0.04 | HCMC Vietnam | Lan et al., 2008 [61] |
| HLA-A*24 | | DSS (37) vs control (50) | 0.375 | 0.25 | 0.029 | Vihn Long Vietnam | Lan et al., 2008 |
| HLA-A*29 | | DSS (87) vs control (276) | 0.62 | 0.01 | <0.001 | Cuba | Paradoa-Perez et al., 1987 |
| HLA-A*31 | | DF and DHF (120) vs control (189) | 0.045 | 0.008 | $<10^{-4}$ | Cuba | Sierra et al., 2007 [62] |
| HLA-A*3 | | DHF (309) vs control (251) | 0.11 | 0.183 | 0.014 | Vietnam | Loke et al., 2001 |
| *HLA-B* | | | | | | | |
| HLA-B blank | | All DSS (41) vs control (138) | 0.07 | 0.01 | <0.02 | Thailand | Chiewsilp et al., 1981 |
| | | DSS (87) vs control (276) | 0.36 | 0.11 | $<10^{-4}$ | Cuba | Paradoa-Perez et al., 1987 |
| | | $2^0$ DSS (41) vs control (138) | 0.07 | 0.23 | <0.05 | Thailand | Chiewsilp et al., 1981 |
| HLA-B13 | | DF and DHF (120) vs healthy control (189) | 0.156 | 0.044 | 0.00005 | Cuba | Sierra et al., 2007 |
| HLA-B15 | | $2^0$ DF (106) vs control (140) | 0.05 | 0.16 | 0.012 | Thailand | Stephens et al., 2002 |
| HLA-B62 | | $2^0$ DF and DHF (209) vs control (140) | 0.005 | 0.05 | 0.016 | Thailand | Stephens et al. |
| HLA-B76 | | $2^0$ DF and DHF (209) vs control (140) | 0.01 | 0.07 | 0.005 | Thailand | Stephens et al. |
| HLA-B77 | | $2^0$ DF (106) vs control (140) | 0.17 | 0.06 | 0.009 | Thailand | Stephens et al. |
| HLA-B5 | | $2^0$ DF (103) vs control (140) | 0.17 | 0.06 | 0.021 | Thailand | Stephens et al. |
| | HLA-B*51 | $2^0$ DHF (103) vs $2^0$ DF (106) | 0.16 | 0.06 | 0.036 | Thailand | Stephens et al. |
| | HLA-B*51 | $2^0$ DHF (103) vs control (140) | 0.16 | 0.04 | 0.0052 | Thailand | Stephens et al. |
| | HLA-B*52 | $2^0$ DF (106) vs control (140) | 0.11 | 0.02 | 0.0067 | Thailand | Stephens et al. |
| | HLA-B*52 | $2^0$ DHF (103) vs $2^0$ DF (106) | 0.01 | 0.11 | 0.0049 | Thailand | Stephens et al. |
| *HLA class II* | | | | | | | |
| HLA-DR | | | | | | | |
| | HLA-DRI | DF (64) vs control (201) | 0.328 | 0.159 | 0.0057 | Brazil | Polizel et al., 2004 [63] |
| | HKA-DRB1*04 | $2^0$ DF and DHF (77) vs control (189) | 0.081 | 0.142 | 0.001 | Cuba | Sierra et al., 2007 |
| | | DF (47) vs DHF (34) | 0064 | 0.35 | 0.011 | Mexico | LaFleur et al., 2002 [64] |
| | HLA-DRB1*07 | DF and DHF (120) vs healthy control (189) | 0.07 | 0.197 | 0.0001 | Cuba | Sierra et al., 2007 |
| | HLA-DRB1*0901 | DSS (14) vs control (61) | 0.15 | 0.32 | 0.0018 | Vietnam | Lan et al., 2008 |
| HLA-DR II | | DF (47) vs DHF (34) | 0.02 | 0.12 | 0.003 | Mexico | LaFleur et al., 2002 |
| *HLA-DQ* | | | | | | | |
| HLA DQI | | DF (64) vs control (201) | 0.796 | 0.577 | 0.0052 | Brazil | Polizel et al., 2004 |

[a] Control groups are comprised of sympatric, healthy, age and sex matched individuals of the same ethnicity as DENV patient cohorts.

### 1.9.5 Global host gene response and prognostic markers of severe dengue

With a wide spectrum of clinical presentations, the outcome and clinical evolution of dengue are unpredictable. While the majority of patients recover from dengue, a small proportion progress to severe disease which is characterized by plasma leakage with or without haemorrhagic manifestations. At the present, there are no specific therapies for dengue and treatment of patients with DSS is limited to supportive care. Early diagnosis which helps to decide appropriate intervention might help to prevent patients from developing DSS. However, to define the group of patients who progress to DSS is very difficult [1]. Complications of dengue (e.g. cardiovascular shock) typically occur between the 4th and 6th days of illness. There is thus a window of time in the first few days of illness where it should be possible to establish a diagnosis [36] and potentially, make a prognosis of disease severity with a host or viral biomarker. The identification and development of tools to detect prognostic markers of DSS could allow for improved patient triage, closer clinical monitoring, or in the future, treatment with anti-viral or immune modulating therapies.

Gene expression microarrays may provide a comprehensive understanding of the host gene response in infectious diseases. Whole blood samples or samples which are a subpopulation of cells (PBMCs, Neutrophiles, B and T cells) were often used for these methods. The use of whole blood involves less manipulation of the specimen. Furthermore, the use of whole blood gives a comprehensive picture of gene expression during the systemic host response to infection and reflects the overall interaction within a complex system. However, two challenges remain when choosing whole blood samples. Firstly, an important gene expression pattern may be undetectable in samples that contain many cell types like whole blood, particularly, a pattern generated by a rare cell population. Secondly, the gene expression data might not be a reflection of the real difference between two groups of samples if there are differences in the number of sub-cell types between the

groups. This means that the resulting gene transcript data may simply be a reflection of the sum of the unique, constitutive transcript profiles from each cell type, rather than an actual increase or decrease in transcript levels within a given cell type in the response to the infection or stimulus [139] [140]. It was also shown that different cell types (Neutrophiles, lymphocytes, etc) in whole blood contribute to variations in gene expression in whole blood samples in healthy people. However, it was also demonstrated that intrinsic individual differences are not the dominant source of variation in gene expression among healthy people and that it is feasible of using gene expression in whole blood as a basis for detection and diagnosis of disease in human patients [141].

Recently, several studies have been performed that describe gene expression profiles in blood samples from dengue patients [77, 78, 142-145]. Simmons et al described whole blood transcriptional signature in Vietnamese adults with DHF enrolled around the time of defervescence [78]. The timing of sample collection is clearly a major factor in studies of the transcriptional signature, with samples collected during the febrile phase having a characteristic anti-viral profile, i.e. with interferon-stimulated genes highly prominent [78]. Studies of PBMC (i.e. minus the neutrophil population) have also been described [142, 143]. Nascimento et al. used PBMC isolated from patients infected with DENV at different times in the disease course. PBMC samples were collected from patients with DF, DHF and non-dengue patients on illness days ranging from day 1 to day 11. All DHF patients were grade I and II according to WHO classification. The author determined 4 genes (PSMB9, MT2A, C3aR1 and HLA-F) that distinguish DF from DHF based on Linear Discrimination Analysis (LDA). Caveats to this result are that the majority of samples used in this study were collected 7 days after onset of disease and patients with life-threatening DSS were not present in the study population.

In a gene expression array study in Thailand, Ubol et al used samples that were collected from 5 DF and 4 DHF patients at various times before defervescence (ranging from day -1 to day -4) ). The standout features were the up-regulation of interferon-related genes, CD59 and IP-10 in PBMC from DF patients compared to PBMC from DHF patients. The author hypothesised that mild dengue (DF) developed because of a more robust IFN response leading to a reduction in the number of virus-infected cells, virus production and cell damage. However, given the very small sample size and inconsistencies in timing of the samples, this study was unable to determine prognostic markers of DSS.

## 1.10 Knowledge Gap

In the context of host genetic susceptibility to dengue, despite many candidate gene studies having been performed, there is still a major lack of understanding from a genome-wide perspective of the genetic determinants of susceptibility. A major limitation of previous studies has been relatively small sample sizes.

In the context of understanding gene expression and identifying prognostic markers, most previous studies (including from Viet Nam) have been small in sample size and rarely included patients with life-threatening DSS. Sample timing is also an issue that must be taken into account. Investigating samples collected on day 4 of illness may only help explain the consequence of what happened well before the patients developed shock and could not be used to predict disease outcomes.

## 1.11 Objectives of the thesis

The pathogenesis of severe dengue remains incompletely understood and effective drugs and vaccines are currently not available. Predicting the outcome of a DENV infection is an important goal that could help to determine disease mechanism and also help to improve treatment. Clinical experience has identified several features (e.g. abdominal pain, persistent vomiting) as being associated with DSS, however many of these manifestations

occur near the time of shock rather than early in the illness. In this thesis, we aim to determine prognostic markers of severe dengue by using global gene expression microarray. The first study (Chapter 3) focuses on exploring differences in acute responses between uncomplicated dengue and DSS on day 4 of illness. This study looks at differences in the whole blood transcriptional profiles between children with DSS and children with uncomplicated dengue on day 4 of illness.

The study described in Chapter 4 was designed to identify prognostic markers of DSS by looking at differences in early transcriptional profiles in the whole blood of dengue patients with uncomplicated dengue and patients with severe dengue (DSS). These patients were classified based on the most up-to-date WHO clinical guideline [1]. Furthermore, the samples were collected at less than 72hrs after disease onset and the DSS samples were very well matched with uncomplicated dengue patients by age, sex and infecting DENV serotype. To our knowledge, this study has the largest sample size to investigate early in vivo transcriptional profiles of dengue patients with different and highly relevant clinical presentations. We also showed in the study that there was no evidence of viral phylogenetic structure that related to clinical outcomes. Key molecules of acute dengue were also identified using network analysis of differential expressed transcripts in acute dengue samples relative to convalescent dengue samples.

In this thesis, we also aimed to identify SNPs that are associated with dengue and also SNPs that distinguish severe dengue from uncomplicated dengue. We performed genome wide association case control study to investigate frequency of 250K SNPs in cases (including DSS patients and uncomplicated dengue patients) and in population controls (cord blood samples). The aim of this study was to determine SNPs that confer the susceptibility to dengue. The most significant SNPs were followed up in a new cohort. In addition, to identify SNPs that are associated with severity of dengue, we conducted a

genome wide association case control study to investigate 500K SNPs in cases (severe dengue patients) and in controls (mild dengue patients). The top significant SNPs were genotyped again in a new set of samples with larger sample size. Despite no genome-wide significant SNP was identified, we identified several SNPs that showed strong evidence for association with susceptibility to dengue and DSS

CHAPTER 2

# 2. MATERIALS AND METHODS

## 2.1 Methods

### 2.1.1 Patient diagnosis

#### 2.1.1.1 Serology for Detection of DENV-reactive IgM/IgG by Panbio ELISA

DENV antigen (Dengue 1-4 recombinant E proteins) was first diluted 250 times with antigen diluent (Phosphate buffered saline) by mixing 10 µl of antigen with 2.5 ml of antigen diluent. The diluted antigen was then mixed with an equal amount of MAb Tracer (Horseradish peroxidase conjugated monoclonal antibody) and kept at room temperature until used. At the same time, plasma samples, positive controls, negative controls and calibrators were diluted 100 times with sample diluent (Tris buffered saline (pH 7.2-7.6)) by adding 10ul of sample to 1000 µl of sample diluent. The diluted samples, controls and calibrators were pipetted into the assay plate (Anti-human IgM/IgG Coated Microwells) and incubated for 1hr at $37^0$C. After incubation, the assay plate was washed 6 times with Wash buffer (20x concentrate of phosphate buffered saline (pH 7.2-7.6) with Tween 20) before 100ul of the antigen-Mab was pipetted into the assay plate which was subsequently incubated for 1hr at $37^0$C and washed 6 times with Wash buffer. After washing, 100ul of TMB (a mixture of 3,3',5,5'-tetramethylbenzidine and hydrogen peroxide in a citric-acid citrate buffer (pH 3.5-3.8)) was added into each well and incubate for 10 minutes at room temperature for blue color to develop. The reaction was stopped by adding 100 µl of Stop solution into each well. Finally, the absorbance of each well was measured at the wavelength of 450 nm.

*Data calculation and interpretation*

$$\text{Panbio units} = \text{Index value} \times 10.$$

Index value was calculated by dividing the sample absorbance by the cut-off value. Cut-off value was calculated by multiplying the average of the triplicates of the Calibrator by the calibration factor, which is specific for each batch.

Table 2.1: Panbio Elisa data interpretation

| IgG Panbio Units | IgM Panbio Units | Results |
|---|---|---|
| < 9 | < 18 | Negative |
| 9 - 11 | 18 – 22 | Equivocal |
| >9 | >22 | Positive |

### 2.1.1.2 Serology for Detection of DENV-reactive IgM/IgG by MAC/GAC ELISA

Supernatant from DENV1-4 or JEV (Japanese Encephalitis virus) infected C6/36 cells and anti-DENV or JEV mouse monoclonal antibodies were supplied by Venture Technologies (UNIMAS Research Park, University Malaysia Sarawak, Malaysia). Briefly, ELISA plates (Maxisorp, Nunc) were coated with 100µl anti-human IgM (A0425-Dako Company) or anti-human IgG (I2136-Sigma Company) at a dilution at 1:2000 overnight at 4°C. Plates were washed and then blocked with 200 µl phosphate buffer saline (PBS) 3% Bovine Serum Albumin (BSA) per well for at least 2 hours at room temperature. Samples, positive and negative controls were diluted at 1:100 in PBS 0.05% Tween 20. After the 2$^{nd}$ wash, 100µl of diluted plasma (1/100) and controls were added and incubated for 2 hours at room temperature. After washing five times, 100µl of antigen (pooled supernatants from C6/36 cultures of DENV1-4 or of JEV) was added to each well and incubated at 4°C overnight. After washing, the assay was continued with a 1-hour incubation with a cocktail of mouse monoclonal antibodies to DENV1-4 E protein and JEV E protein and then bound mAb detected by 100µl of a 1:2000 dilution of anti-mouse Ig Horseradish Peroxidase (HRP) (P260-DAKO Company). After washing, substrate o-phenylenediamine dihydrochloride (OPD) was used to develop colour metric reaction which then being stopped by addition of 50µl 10% $H_2SO_4$. The Optical Density (OD) was read at 490nm.The IgM and IgG positive control sample was a pool of acute plasma from Vietnamese dengue patients. The negative

control sample was a mixture of plasma collected from healthy adult Vietnamese blood donors with no recent history of dengue. Wells with PBS in place of DENV antigen were used to define the background absorbance value. Negative control value (ODNC) was defined by mean OD value of 3 negative controls minus mean OD of 3 background wells (ODB). The assay cut-off (ODCO) was defined as 5x ODNC. The sample ratio (RS) was calculated by taking the sample OD (ODS) minus mean background (ODB) and dividing by the assay cut-off i.e.: RS = (ODS – ODB)/ ODCO]. The interpretation of results was as follows:

- If RS < 0.8, result was interpreted as negative

- If RS > 1.2, result was interpreted as positive

- If RS from 0.8 – 1.2, result was interpreted as equivocal

### 2.1.1.3 NS1 antigen-capture ELISA

#### 2.1.1.3.1 Assay procedure

The Platelia Dengue NS1 Ag kit is a one step sandwich – format microplate enzyme immunoassay for the qualitative or semi-quantitative detection of DENV-NS1 antigen in human plasma or serum. Tests were carried out according to the manufacturer's instructions. Briefly, 50 µl of plasma or control sera [including cut-off (in duplicates), one negative and one positive control sera were incubated directly and simultaneously with 50 µl of diluent and 100 µl of diluted conjugate at $37^0$ C for 90 min in microplate wells. The plate was washed six times with 1x washing solution (TRIS-NaCl – pH 7.4, 1% Tween 20), then immune-complexes were detected by using a color development reaction [160 µl, Mix of TMB Substrate Solution (citric acid and sodium acetate pH 4.0, H2O2 - 0.015%, and DMSO - 4%), and TMB hromogen (0.25% of 3,3',5,5' tetramethylbenzidine)]. After 30 mins of incubation in the dark at room temperature, the enzymatic reaction was stopped by 100 µl of 1N sulfuric acid solution. The OD was read at

dual 450/620 nm. The presence of NS1 antigen was based on comparison of the OD of the sample with the cut-off of control. All samples were tested in a single well.

#### 2.1.1.3.2 Calculation and interpretation of results

The sample ratio (RS) was expressed by the following formula: RS = ODS/ODCO, where

- ODS: Optical density (OD) obtained on the sample

- ODCO: Cut-off value corresponding to the mean OD values of the duplicate control wells, and which must be >0.2.

Result interpretation:

* If sample ratio was <0.5: sample was non-reactive for dengue NS1 antigen, and interpreted as negative

* If sample ratio ranged between 0.5 and 1.0: sample was considered equivocal for dengue NS1 antigen, and interpreted as equivocal.

* If sample ratio was >1.0: sample was reactive for dengue NS1 antigen, and interpreted as positive.

### 2.1.1.4 DENV serotype identification and vireamia quantification

#### 2.1.1.4.1 RNA extraction and cDNA generation

In initial experiments, a diluted suspension of equine arteritis virus (EAV) was added to each plasma sample prior to RNA extraction. The EAV served as an internal control for both cDNA generation and real time RT-PCR. We defined the dilution of EAV suspension and the volume of this diluted suspension following our standard operating procedure. DENV RNA was extracted from patient's plasma using the QIAamp Viral RNA Mini Kit (Catalogue No. 52906, Qiagen, US) [146]. We performed the experiment according to manufacturer's handbook. For each sample, we used 50µl plasma specimen and

approximately 20µl EAV diluted suspension and then harvested 60µl stock of concentrated DENV/EAV RNA. RNA was reverse transcribed to cDNA through two steps with reagents described in table 2.2. Mix 1 (10µl) and RNA (16µl) was incubated at 65°C for 5 minutes and then instantly put on ice while adding mix 2 (14µl) in. The cDNA was made at 50°C for 60 minutes and finally stored in -20°C.

#### 2.1.1.4.2 DENV serotypes identification by real-time RT-PCR

A real-time quantitation Taq Man RT-PCR assay was performed based on a method described by Laue et al [147]. DENV specific primers/probes were also adapted from Laue et al and described in Simmons et al [148]. Briefly, 2µl cDNA from each sample was used in multiplex real-time RT-PCR that contained both DENV and EAV specific primers and probes. Primer sequences for DENV and EAV are described in table 2.3. Reagents (volume/reaction) using in this assay was listed in table 2.4. A linearized plasmid containing the cloned target amplicon was used to make a standard curve for each assay. Results were expressed as cDNA equivalents per milliliter of plasma. The limit of detection for DENV1, DENV2, and DENV3 was 10 copies/ reaction. For DENV4, the detection limit was 100 copies/ reaction. The detectable signal from the internal control amplicon defined the validity of the assay (Ct value range from 30-35). PCR cycling conditions were 95°C for 14 min and then 95°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec for 45 cycles.

Table 2.2: Reagents for reverse transcriptase PCR

| **Mix 1** | **Reagents** | **Volume (concentration)/ reaction** | **Supplier** |
|---|---|---|---|
| | Random hexamers (100μg/ml) | 2μl | Roche, Lewes, UK |
| | 10mM dNTP | 2μl | QIAgene |
| | De-ionized water | 6μl | |
| **Mix 2** | 5X buffer | 8μl | Invitrogen |
| | DTT | 2μl | Invitrogen |
| | Rnase Inhibitor (40U/μl) | 0.8μl | Invitrogen Cat No. 10777019 |
| | Superscript III reverse transcriptase (200U/μl) | 0.4μl | Invitrogen Cat No. 18080-094 |
| | De-ionized water | 2.8μl | |

Table 2.3: Dengue and EAV specific primer sequence and probes

| DENV | Primer sequences (5'-3') | | Probes |
|---|---|---|---|
| | **Forward** | **Reverse** | |
| 1 | ATCCATGCCCACCAYCAATG | GATCARTGGTGTGGATCCCTG | 6-FAM TCAGTGTGGAATAGGGTTTGGATAGAGGAABHQ-1 |
| 2 | ACAAGTCGAACAACCTGGTCCAT | GAGAAGACCAATGGTGCGGC | 6-FAM GTT+T+Tg+T+CT+TC+CA+TCCA-BHQ-1 |
| 3 | TTTCTGCTCCCACCACTTTCAT | AGACTRGCATCCAACGCCA | 6-FAM AAGAAAGTTGGTAGTTCCCTGCAGACCCCA BHQ-1 |
| 4 | GYGTGGTGAAGCCYCTRGAT | ATGAAGGATGGCCGYTCACT | 6-FAM ACTTCCCTCCTCTTYTTGAACGACATGGGA BHQ-1 |
| EAV | CATCTCTTGCTTTGCTCCTTA | AGCCGCACCTTCACATTG | Cy5 GCGCTCGCTGTCAGAACAACATTATTGCCCACAGCGCG BHQ-1 |

Table 2.4: Reagents for real time PCR

| **Reagents (Cat. No., Company)** | **Concentration** | **Volume (µl)** | | | |
|---|---|---|---|---|---|
| | | **DENV1** | **DENV2** | **DENV3** | **DENV4** |
| FW+ RV DENV Primer | 55nM | 3 | 3 | 3 | 3 |
| FW + RV EAV Primer | 55nM | 3 | 3 | 3 | 3 |
| MgCl2 | 25mM | 2.5 | 3.5 | 3.8 | 2.5 |
| Buffer 10X | | 2.5 | 2.5 | 2.5 | 2.5 |
| dNTP (3622614001, Roche) | 10mM | 1 | 1 | 1 | 1 |
| DENV Probe | 1µM | 5 | 2 | 6 | 5 |
| EAV Probe | 5µM | 1.5 | 1.5 | 1.5 | 1.5 |
| Taq (203209, Qiagen) | 5U/µl | 0.2 | 0.2 | 0.2 | 0.2 |
| H20 (W4502, Sigma) | | 4.3 | 6.3 | 2 | 4.3 |

DENV serotype-specific quantitative RT-PCR was performed to determine viral serotype and viremia in the enrolment plasma samples collected from each patient.

#### 2.1.1.5 Algorithm for laboratory confirmation of dengue

The laboratory confirmation of dengue was based on the results from IgM/IgG captured Panbio ELISA. In brief, patients who with negative serology results in paired samples collected at least 4 days apart and with the second sample being collected on day 6 of illness onward were regarded dengue negative. A patient was determined to have recent dengue (but not acute dengue) if there was no IgM response and IgG unit are between 10-20 units with second sample being collected at least on day 6 of illness. A patient is highly suggestive to have primary dengue infection if there is only IgM response (>11 units) and IgG negative (<18 units) with second sample being collected after day 6 of illness. A patient is highly suggestive to secondary dengue infection if he/she is positive with IgG (>22) and negative with IgM (<9).

The laboratory confirmation of dengue could also based on algorithm that were developed based on results from either: 1/ capture IgM and IgG ELISA using DENV/JEV antigens and mAb reagents, which was provided by Venture Technology (Sarawak, Malaysia); 2/ internally-controlled and serotype-specific real time reverse transcriptase PCR (RT-PCR) assays to express the cDNA equivalents per milliliter of plasma; 3/ NS1 Platelia ELISA-based assay provided by BioRad (Hercules, CA) to detect DENV NS1 proteins in plasma.

A summary of the criteria for "laboratory-confirmed acute dengue" acute dengue confirmation is shown in figure 2.1. Primary and secondary infection was distinguished by the magnitude of IgG ELISA Units in early convalescent plasma samples taking into account the day of illness.

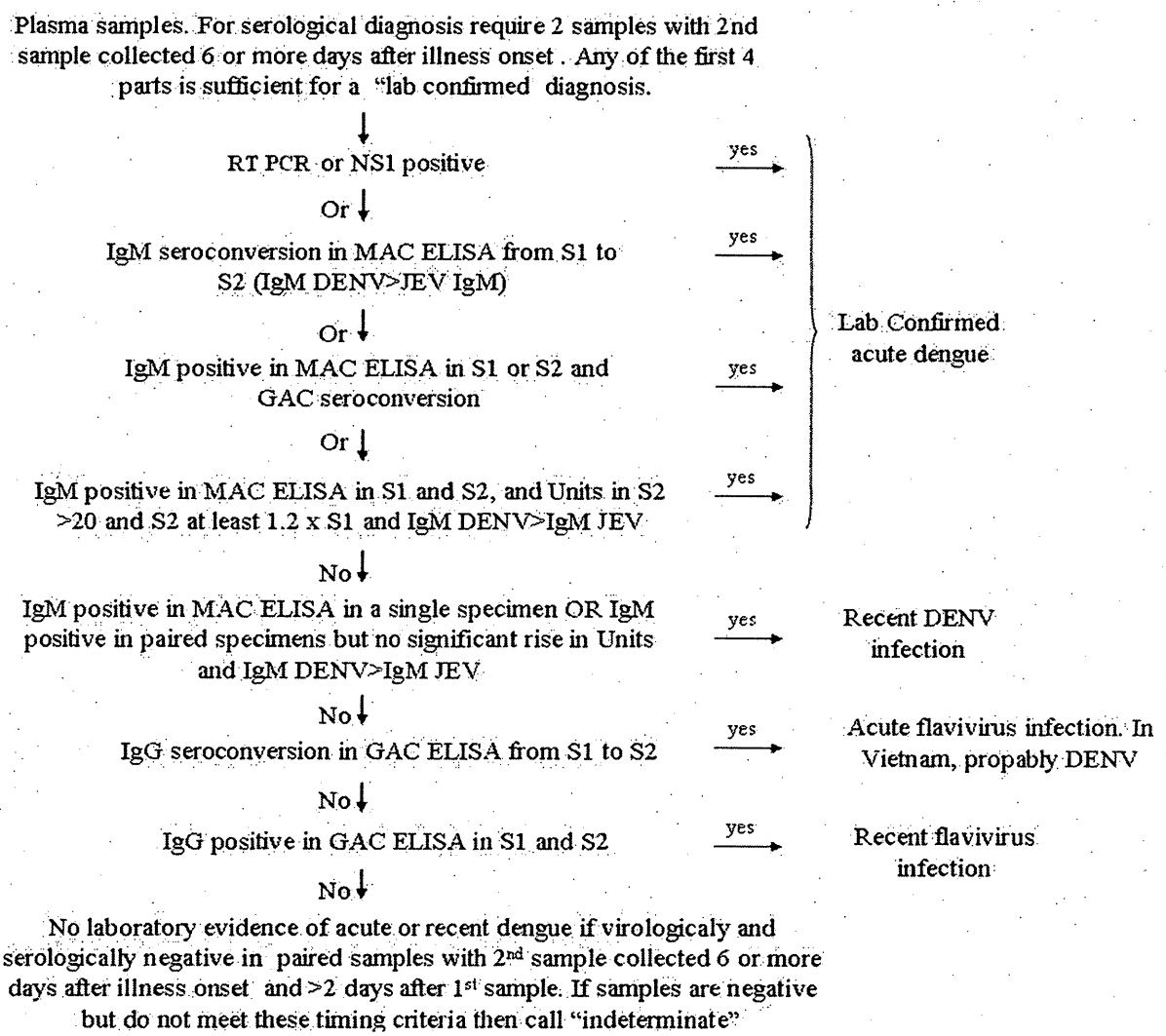



Figure 2.1: Algorithm for laboratory confirmation of dengue

## 2.2 Classification of patients

Disease severity was determined based on WHO classification criteria [1]. Dengue infected patients were divided into 2 groups: uncomplicated dengue and dengue shock syndrome. Uncomplicated dengue was the term used to classify patients with DENV infections that were treated in the general ward and that no special treatment is needed.

CHAPTER 3

# 3. PATTERNS OF GENE TRANSCRIPT ABUNDANCE IN THE BLOOD OF CHILDREN WITH SEVERE OR UNCOMPLICATED DENGUE

## 3.1 Introduction

Dengue haemorrhagic fever (DHF) is a severe presentation of DENV infection. DHF is characterized by fever, myalgia, thrombocytopaenia and a capillary leak syndrome. At its most severe, capillary leak results in hypovolemic shock or dengue shock syndrome (DSS); this requires urgent intravenous fluid resuscitation to restore intravascular volume. A risk factor for DHF/DSS is secondary infection by a DENV serotype distinct from the individual's previous DENV infection history [93, 149, 150]. The basis for secondary infection as a risk factor for severe disease has been ascribed to the infection-enhancing potential of sub-or non-neutralising antibodies from a previous infection [106], cross-reactive memory T cells that secrete vasodilatory cytokines [75, 131, 151], and host genetic pre-disposition [132, 136] . Studies to investigate the pathogenesis of dengue are made more challenging by the absence of a robust animal model of disease.

There remains considerable uncertainty as to how the DENV-host interaction results in vascular leak, the most important clinical characteristic of DHF. Early viral load in vivo appears to be important as DHF patients have significantly higher serum viral loads and NS1 concentrations, possibly driven by antibody dependent enhancement, than DF cases [33, 88]. DHF has also been associated with a robust host inflammatory immune response; significantly greater plasma concentrations of inflammatory cytokines [79, 152, 153] and activated lymphocytes are found in patients with DHF than in cases with DF [75, 152].

Type I interferons (IFN-α/β) are likely to contribute to host defense against DENV infection. Mice deficient in IFN-α/β receptors are more susceptible to DENV infection [154] and in vitro, IFN-α/β can protect cells against DENV infection [155]. Once infected however, DENV non-structural proteins can functionally attenuate the activity of IFN-α/β-mediated antiviral mechanisms, suggesting an evolved viral strategy for escaping early innate immune defenses [156, 157]. Finally, the process of antibody-dependent enhancement of infection may result in attenuation of α/β-interferon-mediated

upregulation of transcription factors important to innate antiviral effector mechanisms [158].

Genome-wide gene expression studies represent an opportunity to identify novel markers associated with dengue immunity and disease pathogenesis. Ubol *et al* [142] demonstrated an early difference in the transcriptome between children with DF and DHF, and we previously found that transcripts from Type 1 interferon stimulated genes (ISGs) were less abundant in adult DSS patients [78].

### 3.2 Purposes of the study

The present cross-sectional study describes global patterns of gene transcript abundance in whole-blood from two groups of pediatric dengue patients who exhibit markedly different clinical phenotypes after identical illness durations. The key goal was to identify transcripts that distinguish severe dengue from uncomplicated dengue.

## 3.3 Materials and Methods

### 3.3.1 Samples collection

Patients for microarray and quantitative PCR experiments were enrolled at the Hospital for Tropical Diseases, Ho Chi Minh City, Viet Nam between July and November 2005. All the patients were children (less than 15 years of age). For each patient, blood samples were collected at four time points; at admission (study day 1), day 2, day 4 and convalescence (1 month after discharge from hospital). Whole blood samples used for microarray were collected into Paxgene blood RNA tube (Qiagen). The sample collection and storage procedures were as instructed by the manufacturer. Samples used for diagnosis and laboratory tests were collected into EDTA tubes (BD). Daily haematocrit and platelet count data were collected from hospital notes and recorded into standard case record forms. WHO classification criteria [21] were applied to each case after review of study notes. Written informed consent was obtained from the parent or guardian of each patient. The study protocol was approved by the Scientific and Ethical Committee of The Hospital for Tropical Disease and the Oxford University Tropical Research Ethical committee.

### 3.3.2 Expression microarray

In this experiment, RNA was converted into cRNA, labelled and hybridized to the gene chip. Figure 3.1 summarises the cRNA synthesis procedure.

#### 3.3.2.1 Extraction of whole blood RNA for gene expression microarray

Whole-blood (2.5ml) was collected directly into PAXgene RNA tubes (Qiagen, Sussex, UK). RNA extraction was performed using Paxgene RNA kits (Qiagen, Sussex, UK). RNA quantity and quality was assessed by spectrophotometry (Nanodrop).

#### 3.3.2.2 Reverse transcription to synthesize first strand cDNA

RNA samples were brought to 11µl (containing approximately 500ng RNA) using RNase-free water and then placed into non-stick, sterile, RNase-free, 0.5ml microcentrifuge tubes.

For a single sample, 9µl of the reverse transcription master mix (1µl T7 Oligo(dT) primer, 2µl 10X first strand buffer, 4µl dNTP mix, 1µl RNase inhibitor, 1µl ArrayScript) was used. After the reaction tube was mixed 2 or 3 times by pipetting up and down and by flicking, followed by a brief centrifugation to collect the reaction in the bottom of the tube, it was incubated for 2 hours at $42^0$C. After incubation, the tube was centrifuged briefly (~5 sec) to collect the reaction mixture at the bottom of the tube.

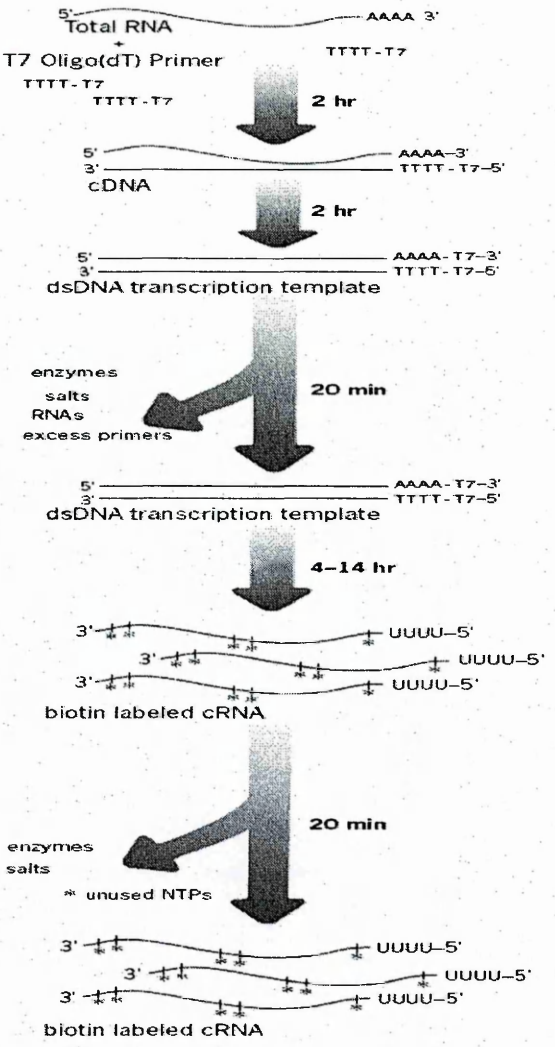



Figure 3.1: cRNA synthesis procedure. mRNA in the total genomic RNA was first reversed transcribed into first strand of cDNA. The first strand cDNA product served as the template for the double- strand cDNA synthesis. After purification, the double-stranded cDNA was in vitro transcribed into antisense RNA which was biotin labelled.

### 3.3.2.3 Second strand cDNA synthesis

For each sample, 80µl of the second strand master mix (63µl nuclease free water, 10µl 10X second strand buffer, 4µl dNTP mix, 2µl DNA polymerase, 1µl Rnase H) was prepared in a nuclease-free tube followed by transfer to each sample and mixed by pipetting up and down 2-3 times and by flicking the tube 3-4 times. The reaction tube was then centrifuged briefly to collect the reaction to the bottom of the tube before incubated at $16^0$C for 2 hours. After incubation, the reaction was put on ice immediately before cDNA purification.

### 3.5.1.1 cDNA purification

The double stranded cDNA product was purified by mixing 250µl with cDNA binding buffer before being passed through a cDNA filter cartridge by centrifugation (1 minute at 10,000 rpm) followed by washing with 500 ml of wash buffer through centrifugation (1 minute at 10,000 rpm). Next, the cDNA filter cartridge was transferred to a cDNA elution tube and eluted by 10µl preheated (at 50 -$55^0$C) nuclease free water by centrifugation for 1.5 minute at 10,000g until the nuclease free water is through the filter. Second elution was performed by using 9µl of nuclease free water and centrifuge for 2 minutes at 10,000g. Finally the purified, double stranded cDNA sample (~17.5 µl) was collected.

### 3.3.2.4 In vitro transcription to synthesize antisense RNA

Antisense RNA was synthesized using in vitro transcription technology. The IVT master mix (2.5µl T7 10X reaction buffer, 2.5µl T7 enzyme mix, 2.5 biotin-NTP mix) (www.ambion.com) was prepared at room temperature into a nuclease-free microcentrifuge tube and mixed by gently vortexing. 7.5µl of IVT master mix was transferred to each sample then the reaction tube was mixed by pipetting up and down 2-3 times or by flicking. The tube was centrifuged briefly to collect the reaction mixture in the bottom of the tube. To enable the in vitro transcription process, the reaction tube was incubated at $37^0$C for 4 to 14 hours using a hybridization oven to prevent condensation.

After incubation, 75µl of nuclease free water was added to the cRNA sample to bring the final volume to 100µl.

### 3.3.2.5 cRNA purification

cRNA product was purified in order to remove enzymes, salts, and unincorporated nucleotides from the cRNA. 350µl of cRNA binding buffer was first added to each sample then 250µl of 100% ethanol was added before the mixture was mixed by pipetting up and down 3 times. The mixture was then transferred to a cRNA filter cartridge (www.Ambion.com) followed by a 1 minute centrifugation at 10,000g. After the flow through was removed, the cartridge was placed into a cRNA collection tube before being washed two times with 650µl wash buffer by centrifugation for 1 minute at 10,000g. After the cRNA filter cartridge was placed to a new cRNA collection tube, cRNA was eluted from the cRNA filter cartridge by applying 100µl preheated Nuclease free water (50 – 60$^0$C) to the centre of the cRNA filter cartridge and centrifugation for 1.5 minutes at 10,000g until the nuclease free water was through the filter. At the end, approximately 100µl of the cRNA was collected in the cRNA collection tube.

### 3.3.2.6 Array hybridization

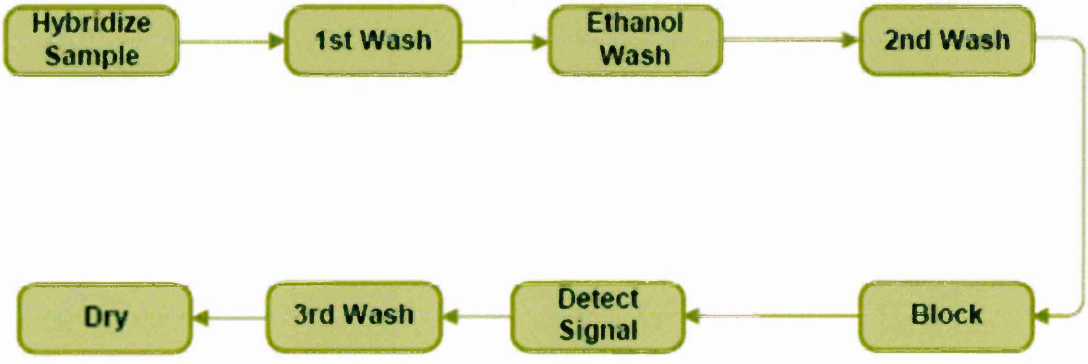



Figure 3.2: Array hybridization, washing, staining and scanning steps

The procedures of hybridization, washing and staining of the bead chip were described in figure 3.2. For each array, 850ng cRNA sample was prepared in 11.3 µl of RNase free water and left at room temperature for 10 minutes to resuspend cRNA. For each bead chip of 8 samples, a hybridization mixture was prepared by mixing 125µl of HYB E1 buffer (preheated in $55^0$C for 10 minutes) with 75µl of formamide. Firstly, 22.7µl of hybridization mix (HYBE1 + formamide) was added to 11.3 µl of cRNA sample. Next, the mixture (34µl) was preheated at $65^0$C for 5 minutes before being loaded onto the centre of the array. After all 8 samples were loaded, the beach chip was loaded into the Hyb cartridge before being placed on the Beadchip Hyb wheel in preheated oven and incubated for 16 to 20 hours at $55^0$C. After the incubation, the chip was washed by using 250 ml of Wash E1BC solution on the orbital shaker for 15 minutes, 250ml of absolute ethanol for 10 minutes, and the second 250 ml of Wash E1BC buffer for 2 minutes. After being washed, the beadchip was blocked with 4 ml of block E1 buffer in 10 minutes, stained with 2 ml of Block E1 buffer and streptavidin-Cy3 (2µl of 1 mg/ml stock per chip) in 10 minutes, washed again by 250 ml of wash E1BC buffer for 5 minutes, and dried for four minutes by centrifugation at 275 rcf at room temperature. Finally the beadchip was scanned on the Illumina BeadArray Reader. The gene expression data was extracted from images collected from Illumina BeadArray Reader.

### 3.3.3 Quantitative RT-PCR

RNA samples used in this experiment were extracted from whole blood using Paxgene blood RNA kit. The RNA concentration was first normalized to the same concentration of 250ng/µl using RNase free water. Next, RNA was reversed transcribed into cDNA by mixing 50µl of 2X of reverse transcription master mix (10µl reverse transcription buffer, 4µl 25X dNTPs, 10µl 10X random primer, 5µl MultiScribe Reverse transcriptase (50U/µl ), and 50µl nuclease free water) with 50µl of 250ng/µl RNA. The reaction tube was then loaded to the thermal cycler. The cDNA cycling conditions were $25^0$C for 10 minutes and

$37^0$C for 120 minutes. The cDNA sample was mixed with Taqman universal PCR master mix and loaded into micro fluidic cards (Applied Biosystems). The micro fluidic card was then run in the ABI Prism 7900HT Sequence Detection System. The data was obtained and analyzed using SDS enterprise software.

### 3.3.4 Microarray data analysis and statistics

#### 3.3.4.1 Data extraction and quality control

The raw expression intensity data was extracted from the arrays using Beadstudio software (Illumina Inc.). This software was used for data quality control (QC) and for obtaining raw intensity data of transcripts included in the array. The raw intensity data exported from Beadstudio includes probe intensity data or gene intensity data. Both of the data formats can be used for analysis. The raw data was normalized to the background by subtracting the background signal.

#### 3.3.4.2 Data normalisation and analysis

After removing the noise by subtracting the background, the data was then normalized using standard normalization procedures (Genespring GX7.3 software, Silicon Genetics, CA) for one colour array data. In brief, data transformation was corrected for low signal, with values recorded at less than 0.01 increased to the minimum (0.01). Per-chip (mean) normalization accounted for chip variability by dividing all of the measurements on each chip by a $50^{th}$ percentile value. Per-gene normalization accounted for variability between probe sets for different genes. Finally, the data was filtered by which only genes that have the detection confidence of greater then 0.999 (or P value < 0.001) in at least one sample will be used for further analysis. ANOVA t-test (unequal variances, non-multiple corrections, false discovery rate less than 5%) built in Genespring software was used for all comparisons. The t-test criteria were: (1) 5% of false discovery rate; (2) Fold change greater or equal to 1.5.

### 3.3.4.3 Ingenuity Pathway Analysis

Ingenuity Pathway analysis (IPA) (www.ingenuity.com) is an approach that helps to determine canonical biological pathways that interested genes are associated. Canonical pathways are well-characterized metabolic and cell signaling pathways. The information contained in canonical pathways comes from specific journal articles, review articles, text books, and KEGG Ligand. These sources were put together and formed the knowledge base. The ratio of canonical pathway was calculated by dividing the number of molecules in a given pathway in the input gene list to the total number of molecules that make up that pathway. To determine whether a canonical pathway formed by the genes in the upload gene list are by chance or biological associated, P value for each pathway was determine based on Right-tailed Fisher's Exact Test. The significance of the test depends on the number of canonical pathway eligible genes in investigating gene list, and the total number of genes that are known to be associated with that canonical pathway in the knowledge base. In our analysis, we use cut-off of P value at 0.05.

### 3.3.4.4 PCR Data analysis

The cycle threshold (Ct) value of each transcript obtained in real-time PCR assay was analyzed manually through checking the amplification curves to remove the spurious signal. Transcripts with too low amplification signals (Ct > 36) were also removed. The expression data was then analyzed using SDS software and RQ manager software (Applied Biosystems). Relative quantification value (RQ value) was calculated as described below: Transcript abundance data (threshold (Ct) value) for each gene was manually checked to remove the false positive values using SDS2.2 software and RQ Manager (Applied Biosystems). Relative transcript abundance was determined using the delta-delta Ct method (Livak *et al* 2001). In brief, this method uses a single sample, termed the calibrator sample, as a comparator for every unknown sample's gene expression level. The calibrator can be any sample chosen to have all of the genes expressed (Ct value of less than 36

cycles). The calibrator is analyzed on every assay plate with the unknown samples of interest. The relative fold difference is calculated using the formula: Fold induction = $2^{-[\Delta\Delta Ct]}$, where $\Delta\Delta Ct$ = [Ct of gene of interest in unknown sample – Ct of 18S gene in unknown sample] – [Ct gene of interest in calibrator – Ct of 18S gene in calibrator]. The fold-difference in transcript abundance in samples derived from patients with different clinical phenotypes was compared using the Mann-Whitney test using Multi-experiment Viewer software (http://www.tm4.org).

## 3.4 Results

### 3.4.1 Gene expression microarray

#### 3.4.1.1 Sample collection

From July to November 2005, we prospectively enrolled 111 children with acute dengue into a study of transcriptional responses in whole-blood at The Hospital for Tropical Diseases (HTD), HCMC, Viet Nam. Using WHO classification criteria, 27 patients had DSS (pulse pressure ≤20 mmHg with poor peripheral perfusion and rapid, weak pulse), and were admitted to the intensive care unit and received appropriate resuscitation. The remaining 84 patients had dengue without evidence of cardiovascular compromise, including patient groups with Dengue Fever, DHF grades I and II. We have used the term uncomplicated dengue to refer to these hospitalized patients as they did not require any significant clinical interventions and were managed throughout on the general dengue ward. From the 27 patients with DSS, we selected RNA samples from 9 patients who presented to hospital with DSS on the 4$^{th}$ day of illness (with the day of onset of symptoms taken as day 1) for microarray analysis. For each DSS patient, an age- and (where possible) sex-matched control patient with uncomplicated dengue, and with an identical duration of illness, was selected. Table 3.1 summarises the baseline characteristics of patients used for microarray study.

Table 3.1: Baseline characteristics of patients included in microarray study

| **Variable** | **Uncomplicated dengue (N = 9)** | **DSS (N = 9)** | ***P-value*[a]** |
|---|---|---|---|
| | N (%) or Median (range) | N (%) or Median (range) | |
| **Male sex** | 5 (55.6%) | 4 (44.4%) | |
| **Age(years)** | 9-14 | 10-14 | |
| **Day of illness** | 4 | 4 | |
| **Febrile** | 9 (100%) | 0 (0%) | |
| **Fever day** [c] | -2.5 (-4 to -1) | 0 | |
| **Infecting serotype**<br>**DENV-1**<br>**DENV-2**<br>**DENV-3**<br>**DENV-4** | <br>4<br>5<br>0<br>0 | <br>1<br>7<br>0<br>1 | |
| **Clinical severity**<br>**DF**<br>**DHF II**<br>**DHF III**<br>**DHF IV** | <br>7<br>2 | <br><br><br>8<br>1 | |
| **Mean viraemia cDNA copies/ml** | 48317364 (19775 – 163000000) | 819735 (24 – 4234125) | ***0.006*** |
| **Secondary infection** | 9 (100%) | 9 (100%) | |
| **Platelet nadir** | 44,000 (34,000-110,000) | 24,000 (11,000-86,000) | |
| **% haemoconcentration** | 15.11 (-9-22) | 28 (13-53) | ***0.001*** |
| **CD3+ T cells** [b] | 649 (262 - 1097) | 845 (599 – 1387) | |
| **CD3+/CD4+ T cells** | 331 (161 - 649) | 371 (155 - 561) | |
| **CD3+/CD8+ T cells** | 207 (67 - 309) | 417 (293 - 752) | ***0.01*** |
| **CD19+ B cells** | 196 (75 - 352) | 229 (109 - 502) | |
| **CD16+/CD56+ NK cells** | 124 (37 - 424) | 148 (51 - 236) | |

Notes:

[a] Mann-Whitney test

[b] The absolute count (cells/μl) of lymphocyte subsets in each sample was determined by flow cytometry

[c] the day of defervescence was regarded as fever day 0

The aim of this study was to describe early whole-blood transcriptional responses in two groups of paediatric patients (n=9 per group) comparing, a) children with uncomplicated dengue, and b) children presenting with DSS. All patients recovered fully. The clinical and haematological characteristics, including absolute counts of the major lymphocyte subsets, of the 18 dengue patients are described in table 3.1. Patients with DSS had lower platelet nadirs and, on the basis of changes in haemoconcentration, significantly greater vascular leakage than uncomplicated dengue patients (table 3.1). On the 4$^{th}$ day of illness, when blood samples for microarray analysis were collected, patients presenting with DSS had significantly higher absolute counts of $CD3^{+}CD8^{+}$ T cells (table 3.1), but significantly lower plasma viraemia levels, than patients with uncomplicated dengue (table 3.1).

### 3.4.1.2 RNA for microarray experiments

The total RNA amount required in gene expression microarray was 500ng in total 11ul of water. RNA were isolated from the whole blood of the patients and measured by Nanodrop. The quality and quantity of RNA samples used for microarray assay is summarized in table 3.2.

Table 3.2: Quantity and quality of RNA in samples used for microarray experiment

| Sample code | Patient group | Time point | RNA (ng/μL) | OD (260/280) |
|---|---|---|---|---|
| DF1093 | DSS | Day 4 | 54.51 | 1.91 |
| DF1097 | DSS | Day 4 | 72.94 | 2.04 |
| DF1105 | DSS | Day 4 | 149.20 | 1.96 |
| DF1106 | DSS | Day 4 | 32.89 | 2.01 |
| DF1107 | DSS | Day 4 | 102.91 | 2.13 |
| DF1108 | DSS | Day 4 | 177.49 | 1.65 |
| DF1110 | DSS | Day 4 | 114.58 | 2.09 |
| DF1114 | DSS | Day 4 | 113.36 | 2.08 |
| DF1115 | DSS | Day 4 | 69.46 | 2.15 |
| DF1093 | DSS | Convalescence | 40.78 | 2.01 |
| DF1097 | DSS | Convalescence | 84.93 | 2.11 |
| DF1105 | DSS | Convalescence | 94.21 | 2.12 |
| DF1106 | DSS | Convalescence | 78.75 | 2.17 |
| DF1107 | DSS | Convalescence | 66.86 | 2.17 |
| DF1114 | DSS | Convalescence | 219.2 | 2.11 |
| MD1589 | UC dengue | Day 4 | 117.74 | 2.11 |
| MD1590 | UC dengue | Day 4 | 66.82 | 1.95 |
| MD1591 | UC dengue | Day 4 | 71.43 | 1.87 |
| MD1594 | UC dengue | Day 4 | 66.56 | 2.10 |
| MD1605 | UC dengue | Day 4 | 49.21 | 2.14 |
| MD1613 | UC dengue | Day 4 | 71.10 | 2.15 |
| MD1639 | UC dengue | Day 4 | 85.07 | 1.92 |
| MD1640 | UC dengue | Day 4 | 114.75 | 1.98 |
| MD1641 | UC dengue | Day 4 | 68.46 | 2.05 |
| MD1589 | UC dengue | Convalescence | 78.26 | 2.16 |
| MD1591 | UC dengue | Convalescence | 59.37 | 2.08 |
| MD1640 | UC dengue | Convalescence | 85.11 | 2.09 |
| MD1639 | UC dengue | Convalescence | 109.84 | 2.17 |
| MD1605 | UC dengue | Convalescence | 127.84 | 2.12 |
| MD1613 | UC dengue | Convalescence | 99.32 | 2.1 |

#### 3.4.1.3 Differentially expressed genes

Gene expression microarray was performed for ~24,000 genes in each patient using Illumina technology. Analysis was focused on identifying transcripts that were differentially abundant between acute and convalescent samples and was restricted to those transcripts that were detected in at least one specimen. For the resulting 13,535 gene transcripts, we applied the following criteria in all comparisons:

a) Gene transcripts were detected in at least 50% of samples. Detection confidence was calculated based on the proportion of probes that were successfully detected relative to the total number of probes for each transcript.

b) Mean fold-change between two conditions was greater than 1.5 fold. The fold difference was calculated based on the difference in mean of average normalised intensity signal in each group for each transcript.

c) False discovery rate was less than 5%.

By these criteria, microarray data analysis revealed 3,092 gene transcripts that were differentially abundant (523 less abundant, 2569 enriched) in acute (day 4) samples from DSS patients compared to six matched convalescent samples, and 2,471 gene transcripts that were differentially abundant (267 less abundant, 2204 enriched) in acute (day 4) samples from uncomplicated dengue patients compared to six matched convalescent samples. There were 1197 genes common to the two gene lists.

#### 3.4.1.4 Uncomplicated dengue: acute versus convalescent samples

Unsupervised pathway analysis of transcripts either significantly enriched (n=2204) or under-represented (n=267) in acute samples (relative to six matched convalescent samples) was used to identify biological themes in the transcriptome of patients with uncomplicated dengue. Amongst transcripts significantly enriched in acute samples, genes in pathways associated with oxidative metabolism and mitochondrial dysfunction were over–represented, suggesting a highly active metabolic state consistent with a host response to infection (figure 3.3). Of immune response pathways, the interferon pathway was most prominent, followed by IL-10, antigen presentation, and IL-6 signaling pathways. In contrast, amongst transcripts that were significantly under-abundant in acute samples, there were no significant molecular themes arising from pathway analysis, most likely because of the small number (n=267) of elements for analysis. Thus, the biological processes

highlighted in the enriched population of transcripts are consistent with a metabolically demanding host response in which interferon-driven immune processes and cytokine networks are prominent.

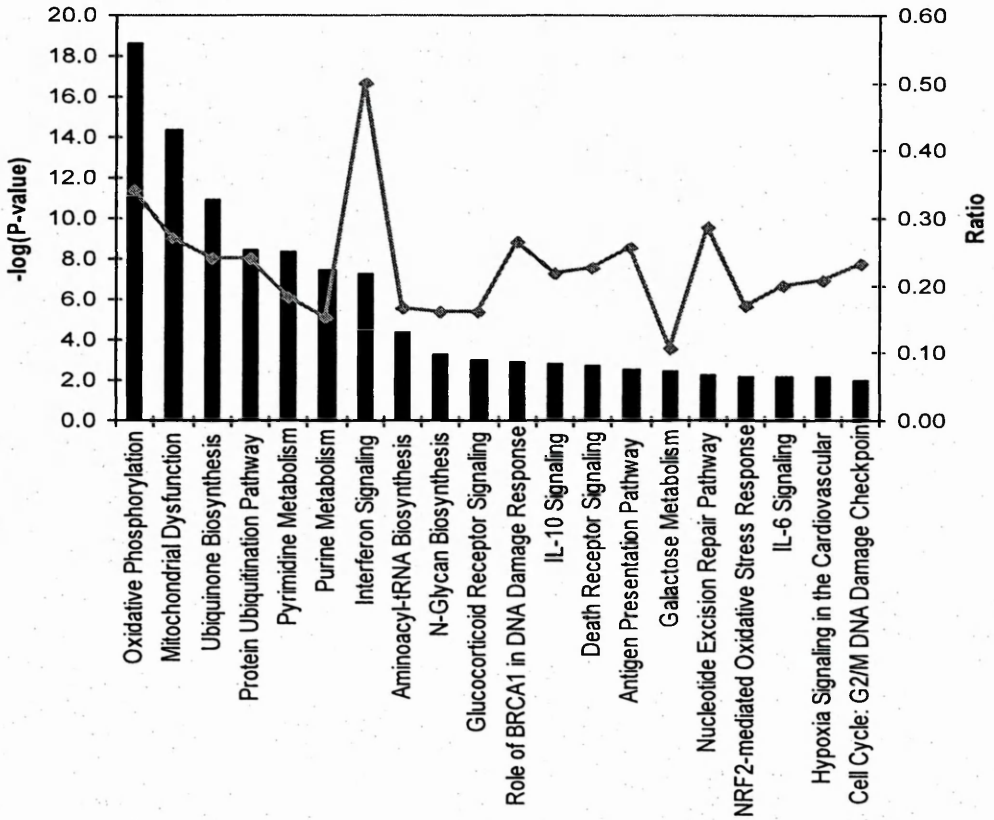



Figure 3.3: Pathway analysis of transcripts enriched in acute samples from uncomplicated dengue patients. Shown are the top twenty canonical pathways identified by unsupervised pathway analysis of filtered microarray data (Ingenuity Systems©) representing transcripts enriched (n=2204) in acute samples (4th day of illness) from uncomplicated dengue patients relative to matched convalescent samples (day 30). The strength of the statistical association is indicated by the length of the bars. The ratio value reflects the proportion of gene elements in the enriched gene list that belong to one of these canonical pathways.

### 3.4.1.5 DSS: acute versus convalescent samples

Surprisingly, unsupervised pathway analysis of transcripts significantly enriched (n=2569) in acute samples from children with DSS did not identify any prominent pathway or biological process. Unsupervised analysis of under-represented transcripts in acute samples (n=523) implicated the death receptor, apoptotic and IL-10 signaling pathways as being features of this "down-regulated" transcriptome (figure 3.4). However, the strength of the association between these pathways and samples was not as robust as was observed between acute and convalescent samples in children with uncomplicated dengue (figure 3.3). Remarkably then, at a time when capillary leakage had precipitated hypovolemic shock in these afebrile DSS patients, the whole-blood transcriptome appeared benign.

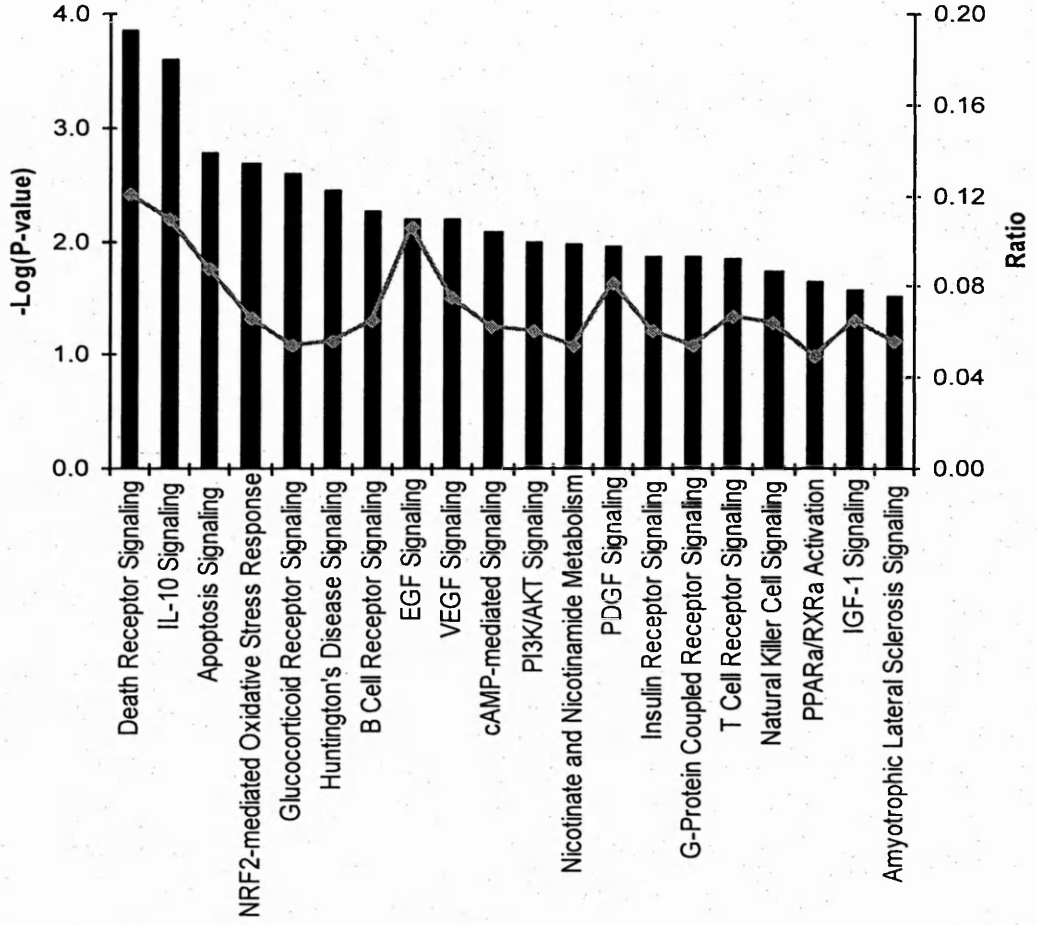



Figure 3.4: Pathway analysis of transcripts under-abundant in acute samples from DSS patients. Shown are the top twenty canonical pathways identified by unsupervised pathway analysis of filtered microarray data (Ingenuity Systems©) representing transcripts under-

abundant (n=523) in acute samples (4th day of illness) from DSS patients relative to matched convalescent samples (day 30). The strength of the statistical association is indicated by the length of the bars. The ratio value reflects the proportion of gene elements in the enriched gene list that belong to one of these canonical pathways.

The absence of a strong transcriptional signature in the blood of children with DSS is in stark contrast to the prominent molecular themes identified in acute samples in febrile children with uncomplicated dengue (figure 3.3). These differences between acute and convalescent samples were not because of intrinsic differences between convalescent samples in the two patient groups; unsupervised comparison of these convalescent samples identified just 138 differentially abundant transcripts amongst 13535 transcripts analyzed. Collectively, these data suggest fundamental differences in the evolution of the transcriptional response between these two groups of children with identical durations of illness, but with different clinical phenotypes.

#### 3.4.1.6 Uncomplicated dengue versus DSS

The transcriptional profiles of acute (day 4) samples from patients with uncomplicated dengue and those with DSS (day 4) were compared directly to identify differentially expressed gene transcripts. The rationale for this comparison was to compare the host response in these two groups of patients with identical illness histories, but distinct clinical phenotypes and disease evolution. In unsupervised analysis, 1,749 transcripts were differentially abundant between acute uncomplicated dengue and DSS patients; transcripts from 1,030 genes were enriched and 719 less-abundant in uncomplicated dengue patients.

Unsupervised pathway analysis of transcripts significantly enriched in uncomplicated dengue patients implicated the death receptor, interferon, apoptosis, IL-6, NF-κB and IL-10 signaling pathways as immune-function related processes distinguishing uncomplicated dengue from DSS (figure 3.5). Unsupervised pathway analysis of transcripts less abundant in uncomplicated dengue patients identified a diverse set of pathways (figure 3.6), with the majority associated with cellular metabolic activities and not immune function.

Collectively, and in the context of the samples available, these data identify apoptotic and immune-function related molecular pathways as defining the transcriptional difference between these two groups of children with identical durations of illness, but with different clinical outcomes.

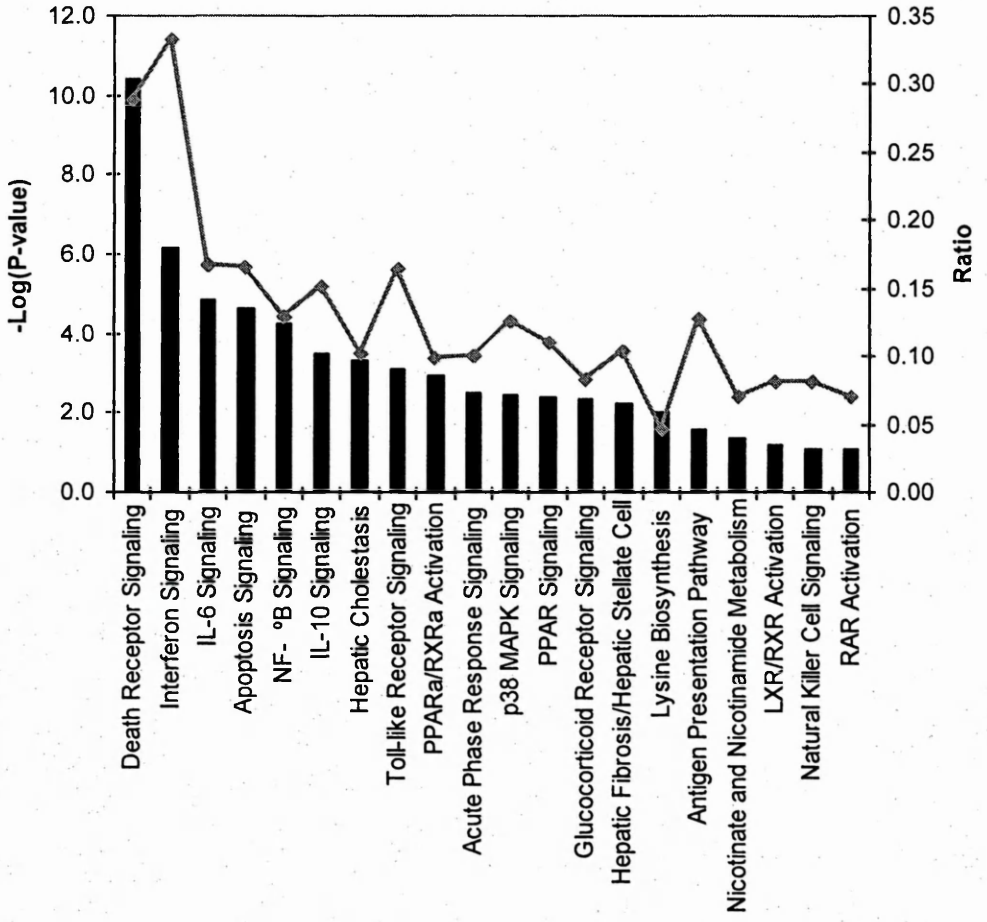



Figure 3.5: Pathway analysis of transcripts enriched in acute (day 4) samples from patients with uncomplicated dengue relative to acute (day 4) samples from DSS patients. Shown are the top twenty canonical pathways identified by unsupervised pathway analysis of filtered microarray data (Ingenuity Systems©) representing transcripts significantly enriched (n = 1030) in acute samples from uncomplicated dengue patients relative to acute (day 4) samples from DSS patients. The strength of the statistical association is indicated

by the length of the bars. The ratio value reflects the proportion of gene elements in the enriched gene list that belong to one of these canonical pathways.

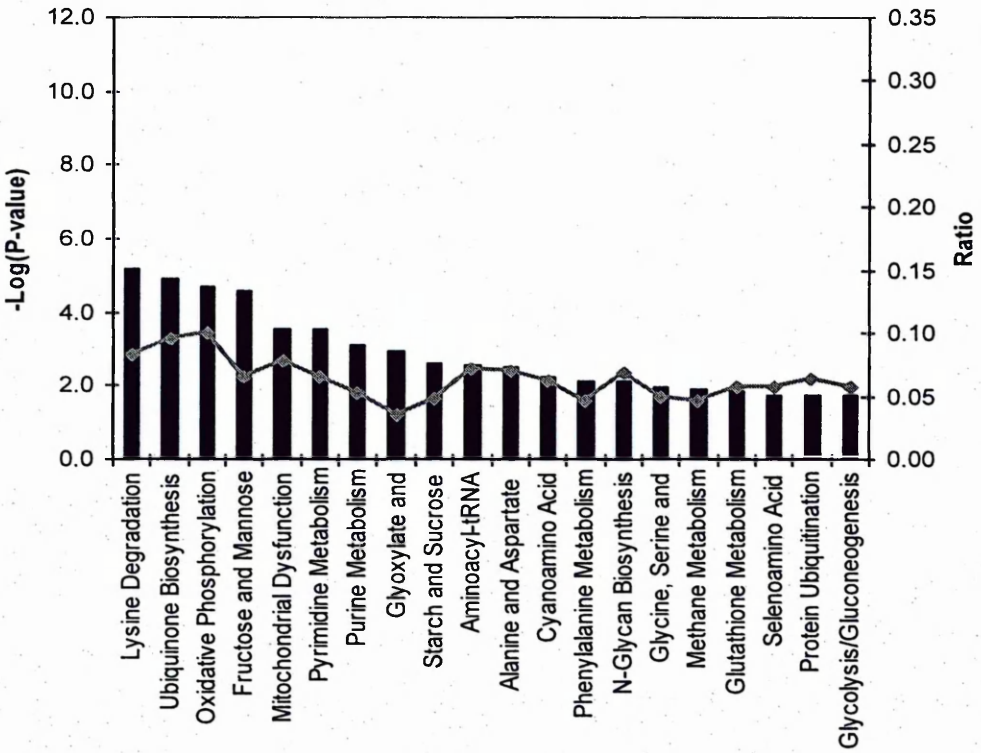



Figure 3.6: Pathway analysis of transcripts under-abundant in acute (day 4) samples from patients with uncomplicated dengue relative to acute (day 4) samples from DSS patients. Shown are the top twenty canonical pathways identified by unsupervised pathway analysis of filtered microarray data (Ingenuity Systems©) representing transcripts significantly under-abundant (n=719) in acute samples from uncomplicated dengue patients relative to acute (day 4) samples from DSS patients. The strength of the statistical association is indicated by the length of the bars. The ratio value reflects the proportion of gene elements in the enriched gene list that belong to one of these canonical pathways.

### 3.4.2 RT-PCR validation of gene transcript abundance

TaqMan RT-PCR assays were used to validate the relative abundance of transcripts identified as differentially expressed between, 1) acute versus convalescent samples from children with DSS and with uncomplicated dengue and 2), acute samples from children with DSS versus uncomplicated dengue. The selection of gene transcripts to validate was influenced by two guiding themes- a) they were biologically replicated in the array (i.e. relative consistent between patient samples) and b) immune-function related genes should be prominent as these pathways distinguished uncomplicated dengue from DSS and in general, a focus on the immune transcriptome would serve the wider purpose of informing models of immunopathogenesis. Validation of gene transcript abundance was performed with acute (day 4) RNA samples from eight DSS patients (6 of who were included in the microarray) and eight matched uncomplicated dengue patients (six were included in microarray).

#### 3.4.2.1 Acute versus convalescent samples in uncomplicated dengue

Sixty-nine transcripts were selected for validation by RT-PCR assays from amongst all transcripts significantly enriched in acute samples from uncomplicated dengue patients relative to convalescent samples. Emphasis was placed on validating canonical ISGs (Interferon stimulated genes) as these were highly enriched in the microarray (figure 3.3). Of the 69 transcripts selected, 64/69(93%) were also differentially abundant when measured by RT-PCR. The heat map in figure 3.7 shows differences in the pattern of gene expressions of samples in DSS and uncomplicated dengue patient. Figure 3.8A and 3.8B show the similarity in fold change between expression microarray and from quantitative PCR. Table 3.3 shows the list of the genes validated by taqman PCR

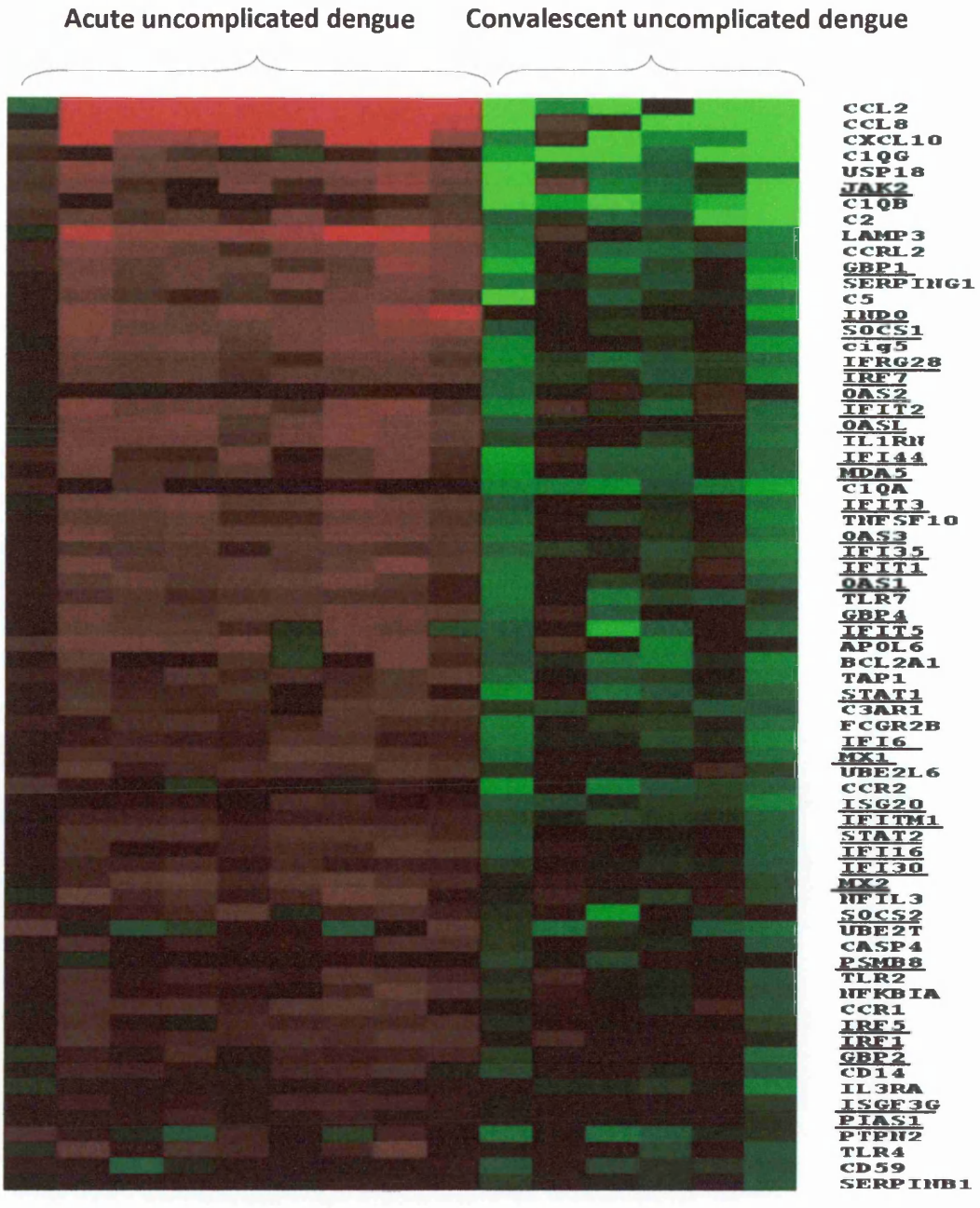
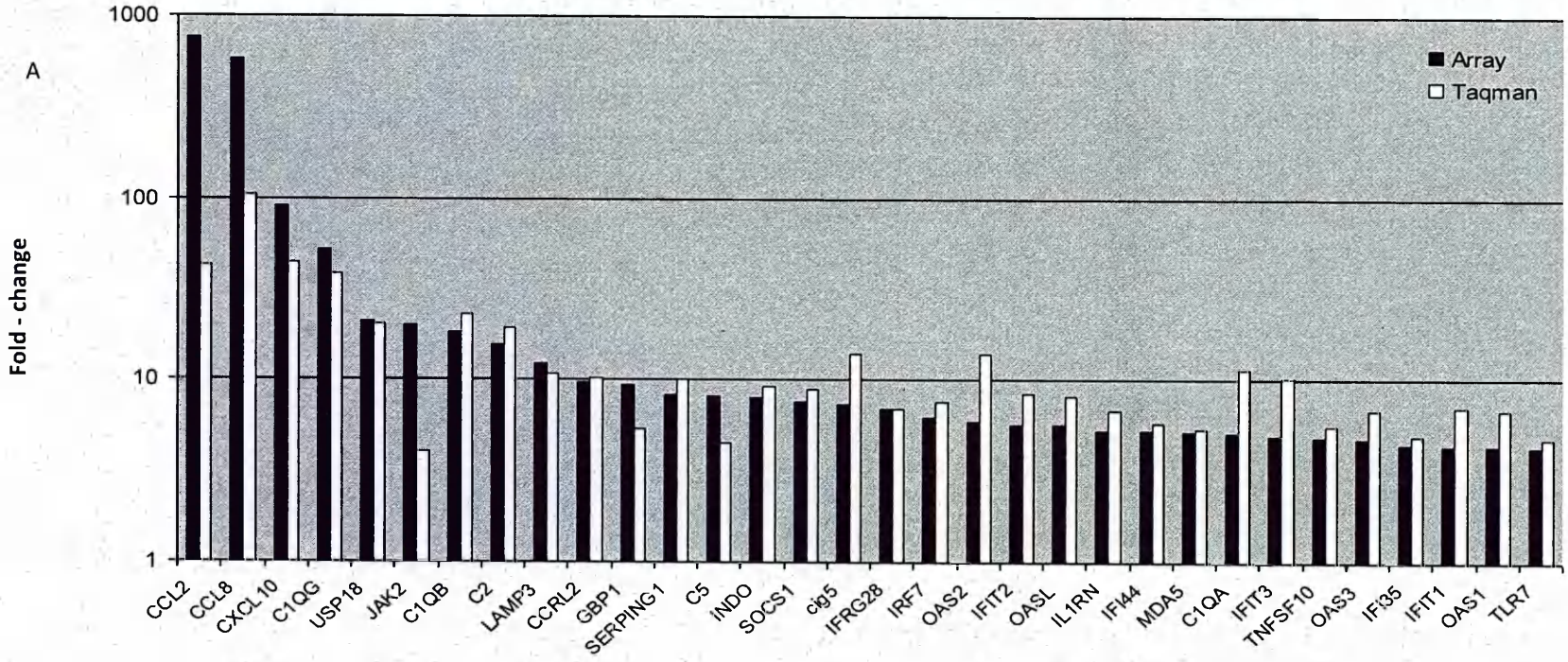



Figure 3.7: Shown is a heat map of individual patients samples filtered on those transcripts enriched in acute samples from patients with uncomplicated dengue relative to matched convalescent samples (n=69) and which were selected for RT-PCR validation. The gene names of canonical ISGs are underlined next to the heat map.

A


Fold - change
1000
100
10
1
Array
Taqman
CCL2
CCL8
CXCL10
C1QG
USP18
JAK2
C1QB
C2
LAMP3
CCRL2
GBP1
SERPING1
C5
INDO
SOCS1
cig5
IFRG28
IRF7
OAS2
IFIT2
OASL
IL1RN
IFI44
MDA5
C1QA
IFIT3
TNFSF10
OAS3
IF35
IFIT1
OAS1
TLR7

B

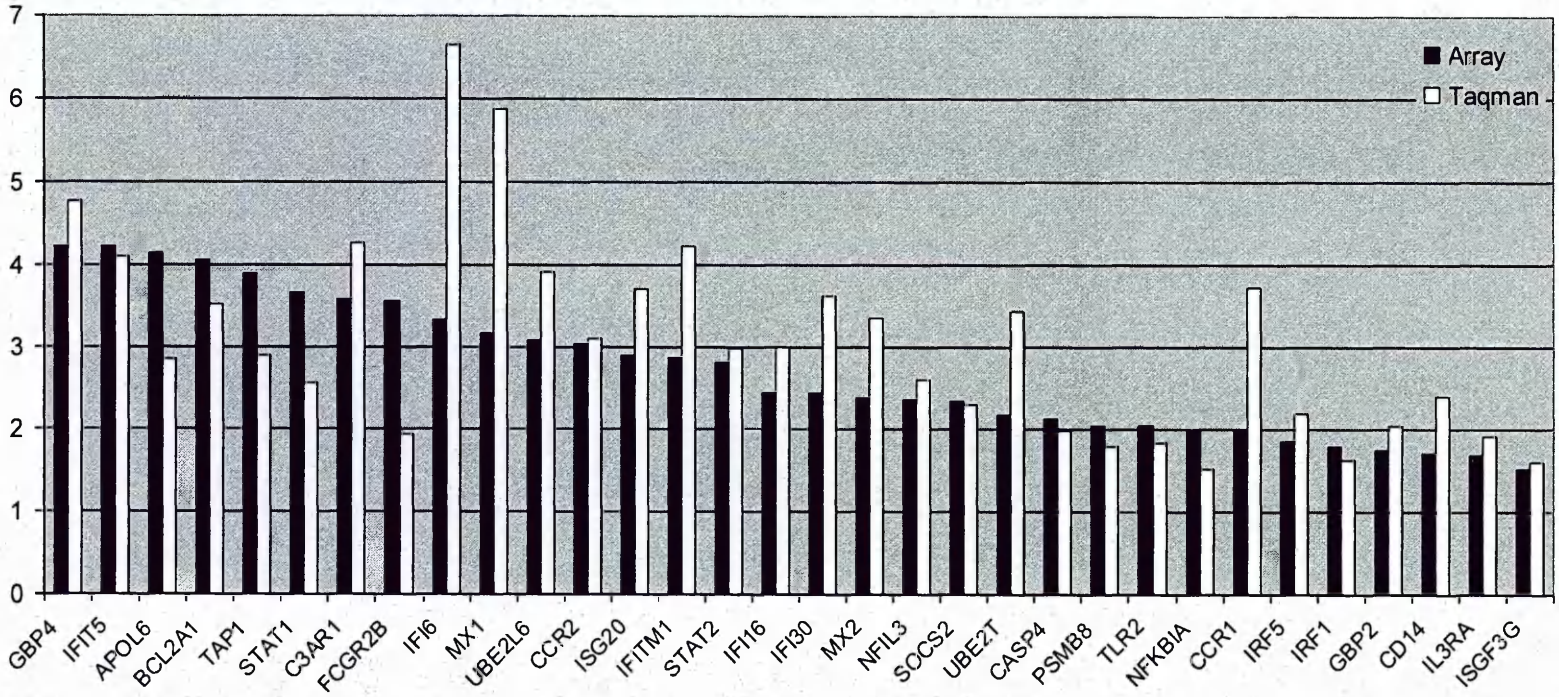



Figure 3.8: RT-PCR validation of transcripts enriched in acute samples from patients with uncomplicated dengue relative to matched convalescent samples. Shown in A) and B) are graphs representing the mean fold-difference in abundance of 64 of the 69 (93%) transcripts that were validated by RT-PCR. The fold difference is shown for both microarray analysis (black bars) and by RT-PCR assay (white bars), with bars above the line indicating greater abundance in acute samples. Five of the 69 transcripts were not validated by RT-PCR.

Table 3.3: RT-PCR validation results of transcripts enriched in acute (day 4) UC dengue patients relative to matched control samples (day 30)

| Genbank | Symbol | Function | Microarray | | Taqman PCR | |
|---|---|---|---|---|---|---|
| | | | FC | P value | FC | P value |
| NM_002982 | CCL2 | chemokine (C-C motif) ligand 2 | 771.0 | 9.23E-05 | 42.5 | 1.19E-03 |
| NM_005623 | CCL8 | chemokine (C-C motif) ligand 8 | 580.4 | 1.22E-03 | 104.5 | 1.19E-03 |
| NM_001565 | CXCL10 | chemokine (C-X-C motif) ligand 10 | 91.3 | 2.30E-03 | 44.1 | 1.19E-03 |
| NM_172369 | C1QG | complement component 1 | 51.9 | 1.52E-03 | 38.2 | 1.19E-03 |
| NM_017414 | USP18 | ubiquitin specific peptidase 18 | 21.1 | 5.01E-03 | 20.3 | 1.19E-03 |
| NM_004972 | JAK2 | Janus kinase 2 (a protein tyrosine kinase) | 19.8 | 2.73E-02 | 4.0 | 2.62E-03 |
| NM_000491 | C1QB | complement component 1 | 18.0 | 9.81E-05 | 22.8 | 1.19E-03 |
| NM_000063 | C2 | complement component 2 | 15.6 | 8.44E-03 | 19.3 | 1.19E-03 |
| NM_014398 | LAMP3 | lysosomal-associated membrane protein 3 | 12.1 | 2.20E-04 | 10.8 | 1.19E-03 |
| NM_003965 | CCRL2 | chemokine (C-C motif) receptor-like 2 | 9.7 | 2.56E-06 | 10.3 | 1.19E-03 |
| NM_002053 | GBP1 | guanylate binding protein 1, interferon-inducible, 67kDa | 9.3 | 1.66E-03 | 5.4 | 1.19E-03 |
| NM_000062 | SERPING1 | serpin peptidase inhibitor, clade G | 8.2 | 2.67E-04 | 10.1 | 1.19E-03 |
| NM_001735 | C5 | complement component 5 | 8.2 | 3.11E-02 | 4.5 | 1.19E-03 |
| NM_002164 | INDO | indoleamine-pyrrole 2,3 dioxygenase | 8.0 | 3.15E-04 | 9.1 | 1.19E-03 |
| NM_003745 | SOCS1 | suppressor of cytokine signaling 1 | 7.5 | 6.01E-06 | 8.8 | 1.19E-03 |
| NM_080657 | cig5 | radical S-adenosyl methionine domain containing 2 | 7.3 | 6.37E-04 | 13.8 | 1.19E-03 |
| NM_022147 | IFRG28 | 28kD interferon responsive protein | 7.0 | 3.01E-06 | 6.9 | 1.19E-03 |
| NM_004030 | IRF7 | interferon regulatory factor 7 | 6.2 | 3.65E-06 | 7.6 | 1.19E-03 |
| NM_016817 | OAS2 | 2'-5'-oligoadenylate synthetase 2, 69/71kDa | 5.9 | 5.28E-04 | 13.9 | 1.19E-03 |
| NM_001547 | IFIT2 | interferon-induced protein with tetratricopeptide repeats 2 | 5.7 | 4.00E-03 | 8.4 | 1.19E-03 |
| NM_198213 | OASL | 2'-5'-oligoadenylate synthetase-like | 5.7 | 4.64E-05 | 8.2 | 1.19E-03 |
| NM_173842 | ILIRN | interleukin 1 receptor antagonist | 5.3 | 4.15E-04 | 6.8 | 1.19E-03 |
| NM_006417 | IFI44 | interferon-induced protein 44 | 5.3 | 5.61E-03 | 5.8 | 1.19E-03 |
| NM_022168 | MDA5 | interferon induced with helicase C domain 1 | 5.2 | 1.57E-03 | 5.4 | 1.19E-03 |
| NM_015991 | C1QA | complement component 1, q subcomponent, alpha polypeptide | 5.1 | 2.63E-05 | 11.3 | 1.19E-03 |
| NM_001549 | IFIT3 | interferon-induced protein with tetratricopeptide repeats 3 | 4.9 | 5.30E-04 | 10.3 | 1.19E-03 |
| NM_003810 | TNFSF10 | tumor necrosis factor (ligand) superfamily, member 10 | 4.8 | 2.36E-04 | 5.5 | 1.19E-03 |
| NM_006187 | OAS3 | 2'-5'-oligoadenylate synthetase 3 | 4.7 | 1.47E-03 | 6.8 | 1.19E-03 |
| NM_005533 | IFI35 | interferon-induced protein 35 | 4.4 | 3.44E-05 | 4.9 | 1.19E-03 |
| NM_001548 | IFIT1 | interferon-induced protein with tetratricopeptide repeats 1 | 4.4 | 3.13E-03 | 7.0 | 1.19E-03 |
| NM_016816 | OAS1 | 2',5'-oligoadenylate synthetase 1, 40/46kDa | 4.3 | 1.04E-03 | 6.8 | 1.18E-03 |
| NM_016562 | TLR7 | toll-like receptor 7 | 4.3 | 4.32E-04 | 4.8 | 1.19E-03 |
| NM_052941 | GBP4 | guanylate binding protein 4 | 4.2 | 1.60E-03 | 4.8 | 1.19E-03 |
| NM_012420 | IFIT5 | interferon-induced protein with tetratricopeptide repeats 5 | 4.2 | 9.48E-03 | 4.1 | 1.19E-03 |
| NM_030641 | APOL6 | apolipoprotein L, 6 | 4.1 | 5.13E-03 | 2.9 | 1.19E-03 |
| NM_004049 | BCL2A1 | BCL2-related protein A1 | 4.1 | 4.95E-04 | 3.5 | 4.91E-02 |
| NM_000593 | TAP1 | transporter 1, ATP-binding cassette, sub-family B (MDR/TAP) | 3.9 | 4.20E-04 | 2.9 | 1.19E-03 |
| NM_007315 | STAT1 | signal transducer and activator of transcription 1, 91kDa | 3.7 | 4.39E-03 | 2.6 | 1.19E-03 |
| NM_004054 | C3AR1 | complement component 3a receptor 1 | 3.6 | 6.33E-06 | 4.3 | 1.19E-03 |
| NM_004001 | FCGR2B | Fc fragment of IgG, low affinity IIb, receptor (CD32) | 3.6 | 1.04E-03 | 1.9 | 2.79E-02 |
| NM_002038 | IFI6 | interferon, alpha-inducible protein (clone IFI-6-16) | 3.3 | 1.13E-03 | 6.6 | 1.19E-03 |
| NM_002462 | MX1 | interferon-inducible protein p78 (mouse) | 3.2 | 3.77E-03 | 5.9 | 1.19E-03 |
| NM_004223 | UBE2L6 | ubiquitin-conjugating enzyme E2L 6 | 3.1 | 2.41E-03 | 3.9 | 1.19E-03 |
| NM_000647 | CCR2 | chemokine (C-C motif) receptor 2 | 3.0 | 7.95E-03 | 3.1 | 1.19E-03 |
| NM_002201 | ISG20 | interferon stimulated exonuclease gene 20kDa | 2.9 | 3.01E-04 | 3.7 | 1.19E-03 |
| NM_003641 | IFITM1 | interferon induced transmembrane protein 1 | 2.9 | 1.97E-06 | 4.2 | 1.19E-03 |
| NM_005419 | STAT2 | signal transducer and activator of transcription 2, 113kDa | 2.8 | 9.92E-04 | 3.0 | 1.19E-03 |
| NM_005531 | IFI16 | interferon, gamma-inducible protein 16 | 2.4 | 9.01E-03 | 3.0 | 1.19E-03 |
| NM_006332 | IFI30 | interferon, gamma-inducible protein 30 | 2.4 | 3.23E-04 | 3.6 | 1.19E-03 |
| NM_002463 | MX2 | myxovirus (influenza virus) resistance 2 (mouse) | 2.4 | 2.18E-03 | 3.4 | 1.19E-03 |
| NM_005384 | NFIL3 | nuclear factor, interleukin 3 regulated | 2.4 | 9.31E-03 | 2.6 | 1.78E-03 |
| NM_003877 | SOCS2 | suppressor of cytokine signaling 2 | 2.3 | 4.72E-02 | 2.3 | 2.79E-02 |

### 3.4.2.2 Acute versus convalescent samples in DSS

Nineteen transcripts were selected for validation by RT-PCR assays from amongst transcripts differentially abundant in acute samples from DSS patients relative to convalescent samples; of these, 11 transcripts were validated by RT-PCR, with transcripts from the complement pathway most prominent. The heat map of the 11 transcripts in DSS patients and UC patients is shown in figure 3.9. Figure 3.10 shows the fold change determined in expression microarray and in taqman PCR. Table 3.4 shows the genes that were validated by taqman PCR.

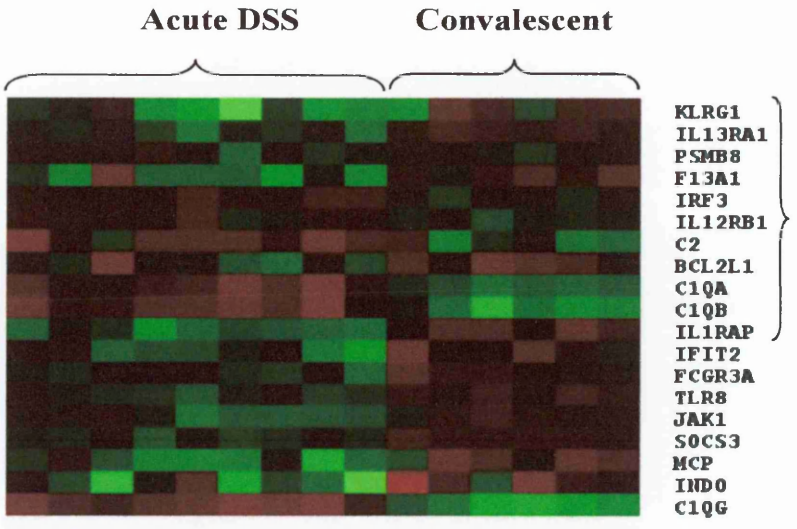



Figure 3.9: Heatmap represents 19 transcripts differentially expressed in acute DSS patients relative to matched convalescent patients

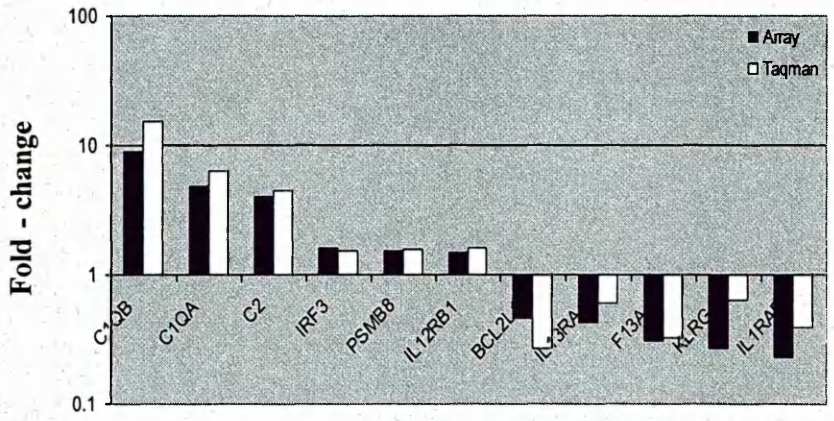



Figure 3.10: Reverse-transcription polymerase chain reaction (RT-PCR) validation of transcripts enriched or less abundant in acute samples from patients with dengue shock syndrome (DSS) relative to that in convalescent samples.

Table 3.4: RT-PCR validation of differential expressed genes in acute DSS patients (day 4) relative to their auologous controls (day 30)

| Symbol | Function | Microarray | | Taqman PCR | |
|---|---|---|---|---|---|
| | | Fold change | P value | Fold change | P value |
| C1QB | complement component 1, q subcomponent, beta polypeptide | 9.0 | 1.48E-04 | 15.4 | 7.78E-04 |
| C1QA | complement component 1, q subcomponent, alpha polypeptide | 4.8 | 1.85E-05 | 6.3 | 7.78E-04 |
| C2 | complement component 2 | 4.0 | 6.48E-03 | 4.4 | 3.28E-03 |
| IRF3 | interferon regulatory factor 3 | 1.6 | 2.40E-03 | 1.5 | 6.32E-03 |
| PSMB8 | proteasome subunit, beta type, 8 | 1.5 | 8.57E-03 | 1.6 | 1.57E-02 |
| IL12RB1 | interleukin 12 receptor, beta 1 | 1.5 | 1.99E-02 | 1.6 | 3.28E-03 |
| BCL2L1 | BCL2-like 1 | 0.5 | 2.75E-02 | 0.3 | 2.32E-03 |
| IL13RA1 | interleukin 13 receptor, alpha 1 | 0.4 | 2.31E-04 | 0.6 | 2.74E-02 |
| F13A1 | coagulation factor XIII, A1 polypeptide | 0.3 | 5.55E-03 | 0.3 | 8.65E-03 |
| KLRG1 | killer cell lectin-like receptor subfamily G, member 1 | 0.3 | 4.24E-02 | 0.6 | 3.57E-02 |
| IL1RAP | interleukin 1 receptor accessory protein | 0.2 | 7.97E-05 | 0.4 | 7.78E-04 |

### 3.4.2.3 Acute uncomplicated dengue versus acute DSS

Fifty-nine transcripts were selected for validation by RT-PCR assays from amongst transcripts significantly enriched in acute samples from uncomplicated dengue patients relative to DSS patient); this included 47 transcripts also investigated in the comparison of acute versus convalescent samples in uncomplicated dengue patients. Of these 59 transcripts, 30/59 (51%) were also differentially abundant when measured by RT-PCR. The expression patents of these 59 genes in each DSS sample and UC dengue sample are shown in figure 3.11. Figure 3.12 and table 3.5 show the fold changes of the 30 genes determined in expression microarray and in taqman PCR. Collectively, these data generally

validate one of the central findings in the microarray analysis; the general lack of a host-defense profile in patients with DSS.

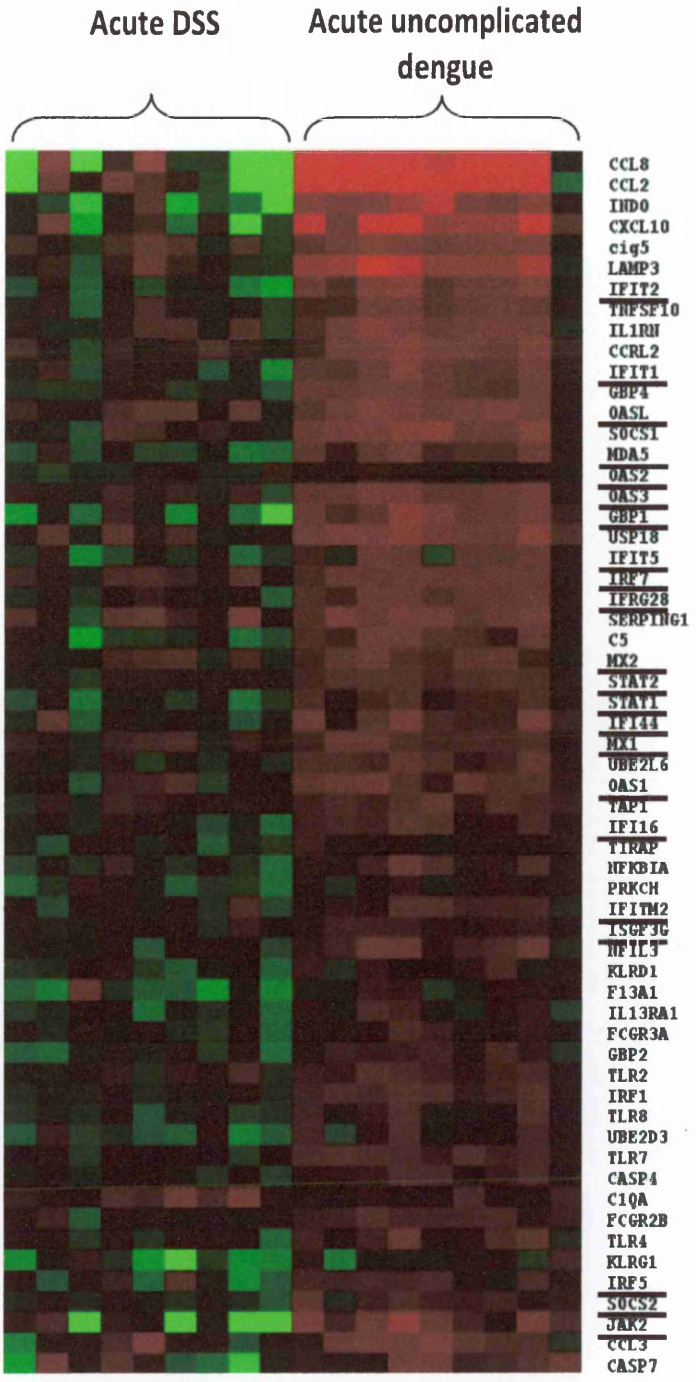



Figure 3.11: Heat map of individual patients samples filtered on those transcripts (n=59) enriched in acute samples from patients with uncomplicated dengue relative to DSS patients and which were selected for RT-PCR validation. The gene names of canonical ISGs are underlined next to the heat map.

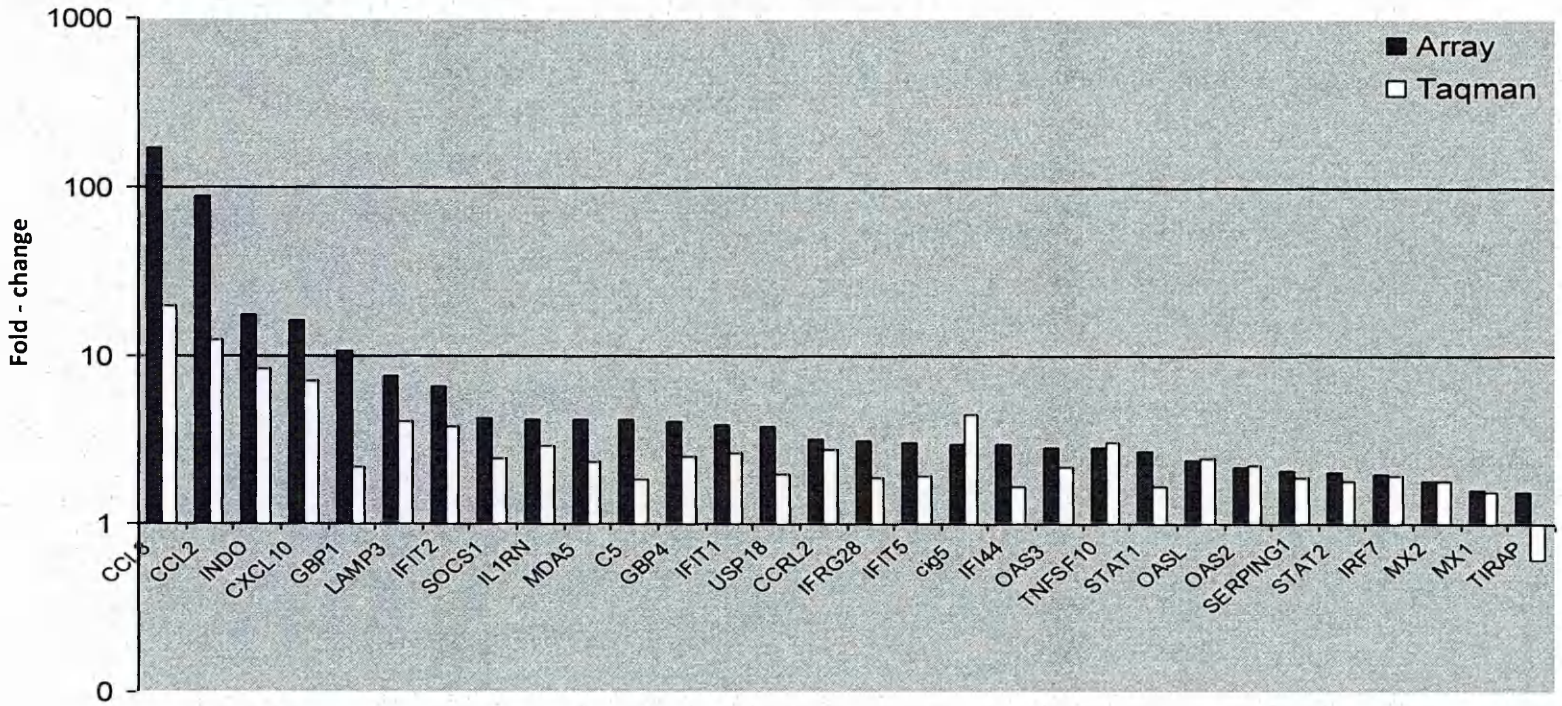



Figure 3.12: Graph representing the mean fold-difference in abundance of 30 of the 59 (51%) transcripts that were validated by RT-PCR. The fold difference is shown for both microarray analysis (black bars) and by RT-PCR assay (white bars), with bars above the line indicating greater abundance in acute samples. Twenty-nine of the 59 transcripts were not validated by RT-PCR.

Table 3.5: RT-PCR validation of transcripts differentially expressed in acute UC dengue patients (day 4) relative to acute DSS patients (day 4).

| Symbol | Function | Microarray | | Taqman PCR | |
|---|---|---|---|---|---|
| | | Fold change | P value | Fold change | P value |
| CCL8 | chemokine (C-C motif) ligand 8 | 170.6 | 6.10E-04 | 19.6 | 2.32E-03 |
| CCL2 | chemokine (C-C motif) ligand 2 | 87.0 | 8.99E-04 | 12.4 | 2.32E-03 |
| INDO | indoleamine-pyrrole 2,3 dioxygenase | 17.4 | 2.67E-04 | 8.4 | 2.32E-03 |
| CXCL10 | chemokine (C-X-C motif) ligand 10 | 16.3 | 1.75E-04 | 7.0 | 1.63E-03 |
| GBP1 | guanylate binding protein 1, interferon-inducible, 67kDa | 10.7 | 1.55E-03 | 2.1 | 1.13E-03 |
| LAMP3 | lysosomal-associated membrane protein 3 | 7.6 | 2.34E-04 | 4.1 | 1.63E-03 |
| IFIT2 | interferon-induced protein with tetratricopeptide repeats 2 | 6.6 | 3.39E-05 | 3.7 | 1.63E-03 |
| SOCS1 | suppressor of cytokine signaling 1 | 4.3 | 1.40E-04 | 2.4 | 7.78E-04 |
| IL1RN | interleukin 1 receptor antagonist | 4.2 | 1.80E-03 | 2.9 | 3.28E-03 |
| MDA5 | interferon induced with helicase C domain 1 | 4.2 | 2.24E-05 | 2.3 | 1.13E-03 |
| C5 | complement component 5 | 4.1 | 2.09E-04 | 1.8 | 1.17E-02 |
| GBP4 | guanylate binding protein 4 | 4.0 | 9.06E-07 | 2.5 | 7.78E-04 |
| IFIT1 | interferon-induced protein with tetratricopeptide repeats 1 | 3.8 | 1.73E-04 | 2.6 | 2.32E-03 |
| USP18 | ubiquitin specific peptidase 18 | 3.7 | 3.39E-05 | 2.0 | 2.32E-03 |
| CCRL2 | chemokine (C-C motif) receptor-like 2 | 3.2 | 1.71E-04 | 2.8 | 1.57E-02 |
| IFRG28 | 28kD interferon responsive protein | 3.1 | 2.38E-05 | 1.9 | 2.32E-03 |
| IFIT5 | interferon-induced protein with tetratricopeptide repeats 5 | 3.0 | 4.07E-03 | 1.9 | 3.28E-03 |
| cig5 | radical S-adenosyl methionine domain containing 2 | 2.9 | 3.09E-03 | 4.5 | 7.78E-04 |
| IFI44 | interferon-induced protein 44 | 2.9 | 1.62E-03 | 1.6 | 1.17E-02 |
| OAS3 | 2'-5'-oligoadenylate synthetase 3 | 2.8 | 3.08E-04 | 2.1 | 3.28E-03 |
| TNFSF10 | tumor necrosis factor (ligand) superfamily, member 10 | 2.8 | 1.93E-04 | 3.0 | 1.13E-03 |
| STAT1 | signal transducer and activator of transcription 1, 91kDa | 2.7 | 1.94E-04 | 1.7 | 1.63E-03 |
| OASL | 2'-5'-oligoadenylate synthetase-like | 2.4 | 6.79E-03 | 2.4 | 7.78E-04 |
| OAS2 | 2'-5'-oligoadenylate synthetase 2, 69/71kDa | 2.2 | 4.87E-04 | 2.2 | 2.32E-03 |
| SERPING1 | serpin peptidase inhibitor, clade G | 2.0 | 4.27E-02 | 1.9 | 2.09E-02 |
| STAT2 | signal transducer and activator of transcription 2, 113kDa | 2.0 | 1.33E-04 | 1.8 | 7.78E-04 |
| IRF7 | interferon regulatory factor 7 | 2.0 | 3.39E-03 | 1.9 | 1.13E-03 |
| MX2 | myxovirus (influenza virus) resistance 2 (mouse) | 1.8 | 2.82E-02 | 1.8 | 8.65E-03 |
| MX1 | interferon-inducible protein p78 (mouse) | 1.6 | 3.13E-02 | 1.5 | 4.57E-03 |
| TIRAP | toll-interleukin 1 receptor (TIR) domain containing adaptor protein | 1.5 | 1.07E-02 | 0.6 | 8.65E-03 |

## 3.5 Discussion

This study attempted to better understand DSS by defining the whole-blood transcriptome at the time of hypovolemic shock and comparing it to matched convalescence samples and to well-matched children with acute uncomplicated dengue. A key and unexpected finding was the relatively benign transcriptional profile of children with DSS relative to their own convalescent samples and more starkly, relative to children with acute uncomplicated dengue but with an identical duration of illness. In particular, transcripts belonging to pathways associated with apoptosis, cytokine signaling (IL-6, IL-10), NFκ-B, and interferon were less abundant in patients with DSS at the time of cardiovascular compromise, despite the presence of a measurable viraemia. These data highlight significant heterogeneity in the type or timing of host immune responses precipitated by DENV infection independent of the duration of illness. Indeed, this study suggests that if a whole-blood transcriptional signature contributes to capillary leakage and DSS, then it occurs before hypovolemic shock manifests.

Capillary leakage leading to DSS is the commonest serious complication of dengue in children living in endemic settings. DSS typically occurs around the time of defervescence and at a time when the dengue viral burden is in steep decline. Host immune responses, driven at least in part by overall viral antigenic mass, have been repeatedly nominated as important in promoting clinically significant capillary leakage (reviewed in [74]). The rationale for the current study was to better understand the host response at the time of vascular decompensation in paediatric dengue patients. The DSS patients in this study were afebrile at the time of hypovolemic shock (day 4) and on average had significantly lower viraemias than the comparison group of children with uncomplicated dengue, who in contrast, were all still febrile. In essence then, although matched by duration of illness, overall serological response, age and sex, these two groups of children had contrasting disease evolutions when samples for microarray analysis were collected.

Consistent with the concept that the patients in this study had different disease evolutions, we observed starkly different whole-blood transcriptional profiles between afebrile children with DSS and febrile children with uncomplicated dengue, with the relative abundance of transcripts from the type I interferon pathway being one of the distinguishing features. For example, transcripts for several canonical ISGs were significantly more abundant in samples from children with uncomplicated dengue than in either samples from patients with acute DSS or matched convalescent samples. Several of the ISGs (OAS1, OAS2, OAS3, IRF7, SOCS1, STAT1) identified in this study were also abundant in the transcriptome of febrile adult Singaporean patients with uncomplicated dengue [144] and in experimentally DENV infected macaque monkeys [159]. Collectively, these data and the observation that acute dengue is associated with elevated IFN-α levels [160], supports the concept that ISG induction in whole-blood is a feature of the febrile phase of dengue. These data are consistent with both in vitro [155-157] and in vivo (murine) [154] studies that suggest the IFN-α/β immune pathway makes an important contribution to host defense against DENV infection.

This study informs models of dengue pathogenesis in several ways. First, the transcriptome of uncomplicated, febrile dengue patients was associated with many of the transcriptional features expected of a systemic viral infection in which the blood is a relevant compartment in which to monitor host responses. Thus, genes and pathways relating to oxidative metabolism, IFN-α/β-stimulation, cytokine expression and apoptosis were all prominent in the transcriptional signature. These prominent genes and pathways may serve as candidate genes in future host genetic studies of dengue susceptibility. In contrast, DSS was associated with an underlying whole-blood transcriptional profile that from an immunological perspective was benign. A possible explanation for these phenomena is that the "host defense" transcriptional profile is observed well before the point of defervescence and cardiovascular decompensation in DSS patients. In this scenario, a very high initial

viral antigen burden drives a rapid and robust "host defense" transcriptional profile that wanes by the time of defervescence and cardiovascular decompensation as viral antigen is cleared by secondary humoral and cellular immune responses. Paradoxically, the rapidity and strength of this antigen-driven secondary response, along with other risk factors, may largely account for the extent of capillary leakage in children with secondary dengue [152]. A conceivable, but perhaps less likely explanation for the missing "host defense" profile at the time of decompensation in DSS patients is that it is attenuated throughout the course of disease. This attenuated response could be related to intrinsic host-genetic influences on gene expression or it might also be a virus-mediated phenomenon, either of which could result in higher initial viral burdens in vivo. In relation to this, the capacity of dengue virus non-structural proteins to inhibit the activation of IFN-stimulated response elements in vitro is well documented [156, 157]. Finally, the absence of an inflammatory transcriptome at the time of cardiovascular decompensation may explain why adjuvant corticosteroid use in patients with established DSS has little efficacy- there is essentially little in the way of an inflammatory transcriptional response to modulate [161].

This current study extends our previous observations in adults with DSS [78]. In contrast to our previous study, the present study was more comprehensive in that it sampled RNA from paediatric patients matched by age, sex and duration of illness and validated a wider panel of transcripts by RT-PCR. A limitation of our current study is the relatively small sample size and lack of samples collected very early in the evolution of disease. In addition, we focused on children with DSS on the 4$^{th}$ day of their illness- these "fast progressors" are not representative of the general population of DSS patients in our hospital setting who on average experience hypovolemic shock on the 5$^{th}$ day of illness. Future studies will focus on the very early host-viral events in febrile children who subsequently progress to DSS. Ultimately, a better understanding of the molecular evolution of DSS could help identify novel biomarkers with prognostic utility and form the rationale for new clinical interventions in this important disease.

CHAPTER 4

# 4. AN EARLY WHOLE BLOOD TRANSCRIPTIONAL SIGNATURE ASSOCIATED WITH PROGRESSION TO DENGUE SHOCK SYNDROME IN VIETNAMESE CHILDREN

## 4.1 Introduction

Dengue is the most significant mosquito born viral infection of humans. The dengue pandemic has spread to the extent that between 70-500 million infections occur each year in over 100 countries resulting in ~40 million clinically apparent cases and ~20,000 deaths [162]. There are no licensed vaccines for the prevention of dengue.

Severe dengue is characterised by a haemorrhagic diathesis, thrombocytopenia and a capillary leakage syndrome. In severe dengue, capillary leakage leads to life-threatening complications that can include cardiovascular shock, called dengue shock syndrome (DSS), severe bleeding or more rarely, other major organ impairment [1]. DSS is the commonest life-threatening complication of dengue in children. Most cases of DSS are associated with secondary dengue virus (DENV) infections but the exact mechanism of disease pathogenesis is not fully elucidated. The prevailing view is that a combination of host and viral characteristics interact and contribute to the expression of the clinical phenotype. Host features implicated in susceptibility to severe disease include genetic variation [2], previous DENV infection history [93, 96, 149, 150], individual propensity for capillary leakage that itself is influenced by age and sex [163]. Viral genetic traits might also be important, with some lineages of DENV being virologically [164] and epidemiologically fitter [165] than others and more frequently associated with severe disease.

Severe dengue is uncommon given the total number of infections [166, 167]. Nevertheless, there are no specific therapies for treatment of severe dengue and management is limited to supportive care. In children, the most common complication is cardiovascular shock and this typically occurs between the $4^{th}$ and $6^{th}$ days of illness, during the so-called critical phase where compensatory mechanisms can no longer maintain an adequate intravascular volume. There is therefore a window of time in the first few days of illness where it should

be possible to establish a diagnosis [36] and potentially, make a prognosis of disease severity with a host or viral biomarker before decompensation occurs. The identification of prognostic markers of disease severity could allow for improved patient triage and management, or in the future, treatment with anti-viral or immune modulating therapies.

## 4.2 Aim of the study

The aim of this study was to identify whole blood transcriptional signatures in Vietnamese children and adults that were associated with progression to DSS. Intriguingly, our findings identified several neutrophil-associated transcripts as more abundant in Vietnamese patients who progressed to DSS, suggesting a hitherto unrecognised association between neutrophil activation, pathogenesis and the development of DSS.

## 4.3 Materials and Methods

### 4.3.1 Sample collection

The study was conducted between 2007 and 2008 at Dong Thap hospital, Dong Thap Province, Vietnam. The study protocol was approved by the Scientific and Ethical committee of Dong Thap Hospital and the Oxford Tropical Research Ethical committee. Subjects for this study were recruited from Dong Thap hospital (Dong Thap province, Viet Nam). Patients underwent screening in the outpatient clinic of Dong Thap Hospital if there was a clinical suspicion of dengue and illness duration of less than 72hrs. Plasma from screened patients was tested with an NS1 ELISA test (Biorad, California, USA). Patients that were NS1 ELISA positive were admitted for further observation and participation in the study protocol if written informed consent was obtained from the patients (patients >14yrs) or from the patient's parents/guardians (patients ≤14yrs). Blood samples for analysis of RNA (Paxgene tube) and plasma were collected at the time of enrolment (study day 1), on study day 3 and at the time of discharge. Haematological data were recorded

daily and all patients received an ultrasound within 24hrs of defervescence. Clinical and haematological assessments were performed daily.

### 4.3.2 Expression microarray (see Chapter 3, section 3.3.2)

### 4.3.3 ELISA assays

Elastase (ELA2) in patient plasma samples was measured using a PMN-Elastase ELISA kit (ImmunDiagnostik, Netherland). Plasma Defensin 1α (DEF1A), Bactericidal/permeability-increasing protein (BPI) and Myeloperoxidase (MPO) were measured using capture ELISA assays (Hbt, Hycult Biotechnology, Netherland). The limit of detection was 120pg/ml for ELA2, 50pg/ml for DEF1A, 250pg/ml for BPI and 0.4ng/ml for MPO. Plasma albumin concentrations were measured by capture ELISA. Briefly, ELISA plates (Maxisorp, Nunc, Denmark) were coated with anti-albumin pAb (polyclonal rabbit anti-human albumin (Dako Co.) overnight and then blocked with 2% skim milk (Merck) in PBS-T 0.05% for 1h at $37^0$C. Next, samples and standards were added into the wells followed by addition of diluted biotinylated anti-albumin antibody solution and incubated for 1hr at $37^0$C. Diluted streptavidine HRP (P0397, Dako Co. Japan) was added and incubated for 1hr at $37^0$C before the OPD substrate was added (Sigma, , USA). After the reactions were stopped by diluted sulphuric acid (10%), the OD value was measured using a plate reader (Biorad, California, USA) and the result was analysed by Microplate software (Biorad, California, USA).

### 4.3.4 DENV genome sequencing

Viral RNA was isolated from patients' plasma samples (QIAamp Viral RNA Kit – Qiagen, cat # 52906) and reverse transcribed to cDNA with superscript III reverse transcriptase (Invitrogen, USA), random hexamers (Roche) and a specific oligonucleotide targeting the 3' end of the target genome sequences. cDNA was then amplified using a high fidelity DNA polymerase (HotstarTaq – Qiagen, Cat # 203205) and a pool of tagged specific

primers to produce 14 overlapping amplicons. The details of primer sequences for whole DENV genome amplification are listed in table 4.1 - 4.3. Table 4.4 shows the Reagents and thermal cycle conditions for whole DENV genome amplification

Each amplicon was ~2Kb in length. This amplification strategy provided 2-3X physical coverage of each nucleotide site. The amplification products were sent to the Broad Institute, USA for sequencing. The procedure for whole DENV genome sequencing is demonstrated in figure 4.1.

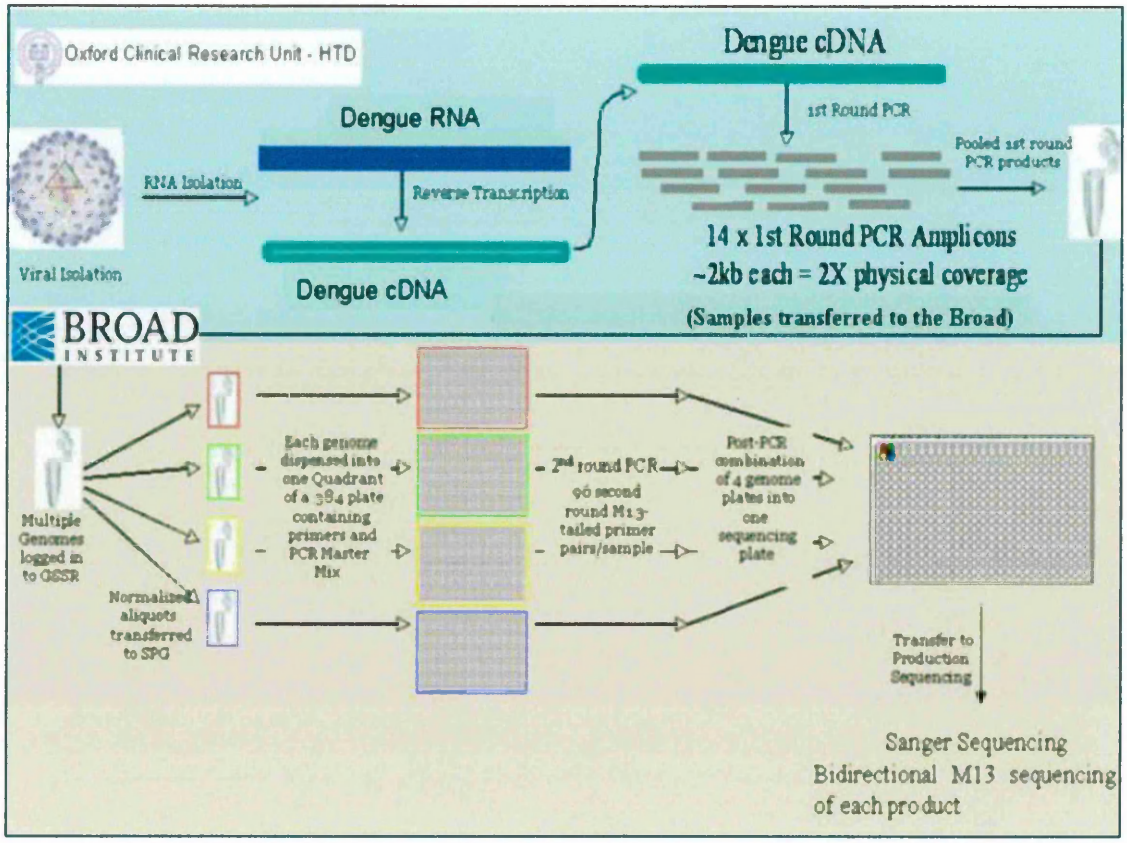



Figure 4.1: Whole DENV genome sequencing strategy. Viral RNA was first isolated from plasma sample, reversed transcribed into cDNA before being amplified into 14 overlapping fragments in the first round of PCR. The products of PCR reaction were sent to Broad Institute for sequencing.

Viral genomes were sequenced using the Broad Institute's ABI directed amplification viral sequencing pipeline (see http://www.broadinstitute.org/annotation/viral/Dengue). Amplicons were then sequenced in the forward and reverse directions using primer panels consisting of 96 specific primer pairs that produce 500-700bp amplicons from the target viral genome. Total coverage delivered post amplification and sequencing was ~ 8X.

Table 4.1: Oligonucleotide primer sequences used for DENV-1 amplification in 14 fragments; Note: Y stands for C/T; R for A/G nucleotide.

| DENV-1 Forward Primer | DENV-1 Reverse Primer | Region |
|---|---|---|
| AGTTGTTAGTCTACGTGGACCG | AYCACGATGTARCTCTCACCAA | 1-2073 |
| TTCTAGCCATACCYCCAACAGC | CTTCCACATTTGAGTTCTCTRCC | 261 - 2469 |
| AGCACATGCCATAGGAACATCC | AYTTGGGTGTAGGAGTCACGCA | 856 - 2953 |
| TGAGACCCAGCATGGAACYGT | YAGCACCGGAAGCCATGTTGTT | 1873 - 4060 |
| CYTGGACCATGAAAATAGGAATAGG | ATRGTTCCATCATCTTGGACCTC | 2289 - 4380 |
| CATATGGYTGAAATTGCGTGACTC | AACACCTCGTCYTCAATCTCTGG | 2917- 5070 |
| ATGGACTTGCAATGGGYATYATG | TGCTTCTGTCCAATGRGCGTGRT | 3849 - 5995 |
| AGTGYTATGGGACACACCYAG | TCCACACTGGCCATCCATARCA | 4525 - 6678 |
| GCYCAAGCTAAAGCATCACARG | AAGGCGAGAAGTGGAACTCCTA | 5015 - 7116 |
| CACGCYCAYTGGACAGAAGCAAA | TCYACCACACTTGGCATGTARG | 5975 - 8145 |
| TCYGAACAAGGAGGAARAGCCTA | TCTGTTGTCCRAAGGGTGTGGT | 6476 - 8627 |
| AAAYTGAGGTGGTTTGTGGAGAG | ACGGCTGAACAGATRGCATTAGC | 7754 - 9918 |
| AGTGGAACCAGAGGTAGCCAAC | GGTTTTTACATCCCCACGATGG | 8368- 10465 |
| ATYCCCATGGTCACACAAATAGC | AGAACCTGTTGATTCAACAGCAC | 8573 - 10735 |

Table 4.2: Oligonucleotide primer sequences used for DENV-2 amplification in 14 fragments; Note: Y stands for C/T; R for A/G; W for A/T and K for G/T nucleotide)

| DENV-2 Forward Primer | DENV-2 Reverse Primer | Region |
|---|---|---|
| AGTAGTTAGTCTACGTGGACCG | GATCCRAAATCCCARGCTGTGTC | 1-2207 |
| CAGATCTCTGATGAATAACCAACG | GATCCRAAATCCCARGCTGTGTC | 87-2207 |
| CCAGAAGACATAGAYTGTTGGTG | TRCCTGCATGATTCCTTTRATGTC | 619-2718 |
| TGCCCAACACAAGGRGAACCYA | ATCTTCCATGTRTCATTGAGTGC | 1156-3056 |
| ATGGTGCARGCYGATAGTGGTT | TAGGCTCCRTCTTCCAGTTCRG | 2410-4589 |
| YATGACAGGAGACATYAAAGGAATC | CTTCCARCCTCCTCCATAYGATA | 2685-4773 |
| ATGCTYAGGACCCGAGTAGGAA | AGRCAAGCTGCTATRTCATTTCC | 3541-5645 |
| GCATGGAARGTGAGYTGCACAA | TATTCRCCRTCAATGGCATCCAC | 3967-6089 |
| GTYACAAGGAGTGGARCATATGT | AGRAGGGGAACTCCRATGTCCA | 4984-7106 |
| GCAGCYGGGATTTTYATGACAG | GTYTCTCCTRTGTTGCCAGTTC | 5446-7592 |
| AGATGGYTGGAYGCTAGGATCTA | TARTGCCATGAYGTTTCATGCTC | 6301-8450 |
| CTAGAWCCAATACCYTATGATCC | GTGATTCTTGTGTCCCAKCCTG | 7288-9197 |
| GTGACATAGGGGARTCRTCACC | GTARTCTGTGTATTCCTCATTGCC | 8003-10215 |
| GCCATATTCACYGATGAGAACAAR | AGAACCTGTTGATTCAACAGCAC | 8800-10723 |

Table 4.3: Oligonucleotide primer sequences used for DENV-3 amplification in 14 fragments; Note: Y stands for C/T; R for A/G nucleotide.

| DENV-3 Forward Primer | DENV-3 Reverse Primer | Region |
|---|---|---|
| TGGACCGACAAGAACAGTTTC | GGCTTTGTCTCCAATTCCAA | 16-2086 |
| GCATGATTGTGGGGAAGAAT | TGTCGACCTGATTCCGCACA | 468-2593 |
| AGCCCTATTTCTTGCCCATT | GTGAAGACTCCGAACCCGTA | 844-2904 |
| TTTCATGCATTGTGATAGGA | GGGTTTGCTTTTGCCAAGTG | 2358-4508 |
| ATGGAGTGTGCGGAATCAGG | TCTCCATCCTCCTCCGTATG | 2568-4765 |
| ACATGATTGCAGGGGTTCTC | ATTGCCTGAATTCCATGAGC | 3561-5560 |
| TGGAGAACAGCCACCCTGATT | TCACCCCTCCTCATGAGTTC | 3962-6138 |
| GTGGGGGTTGGAGTACAAAA | TTGCTTGTCTTGGGGAGTTC | 4619-6769 |
| CATTGAAAGGGCTCCCAATA | TCCAGCAGCTGTCCTTTTTT | 5217-7240 |
| GATGTCTCAAGCCAGTGATT | CACACTTTTCAGGTGGAAGA | 5796-7985 |
| GAGACCTAGGTGGCTTGATG | ATGGTAAGCCCACGTTTTGTA | 6289-8479 |
| CCCTAGCCACAGGACCAATA | GTTCTCGAGGTCTGCCTTCG | 7404-9469 |
| CATGCCAACTGTGATTGAGC | CAAATGGCTCCCTCTGACTC | 8119-10263 |
| ACGAAACCATGGGATGTGGT | CATTTTCTGGCGTTCTGTGC | 8546-10679 |

Table 4.4: Reagents and thermal cycle conditions for whole DENV genome amplification

| Reagents | µl/reaction | Final concentration |
|---|---|---|
| FW Primer (10 µM) | 0.5 | 0.2 µM |
| RV Primer (10 µM | 0.5 | 0.2 µM |
| DMSO | 1.25 | 5% |
| 10x HiFiTaq Buffer | 2.5 | 1x |
| dNTPs 10mM | 1 | 0.4mM |
| HiFi Taq (5U/µl) | 0.2 | 1 unit |
| cDNA | 2 | |
| $H_2O$ | up to 25 µl | |
| **PCR conditions for Whole Genome Amplification in 14 fragments** | | |
| **T° (C)** | **Time** | **Number of cycles** |
| 94 | 2 mins | 1 cycle |
| 94 | 30 sec | 25 cycles |
| 60 | 1 min | |
| 72 | 2 mins | |
| 72 | 10 mins | 1 cycle |
| 4 | hold | |

### 4.3.5 Microarray data normalisation and analysis

Standard normalisation procedures in Genespring were used (ver.10, Agilent, Santa Clara, CA, USA). In brief, array normalisation was performed by dividing the mean of array intensity value by the 75$^{th}$ percentile value of all arrays. Gene normalisation was performed by dividing the mean value of the gene in each array by the 75$^{th}$ percentile value of the gene in all the arrays. Normalised data was then filtered based on expression data in which only transcripts with detection confidence of less than 0.001 in at least one out of 227 samples (total samples arrayed) under analysis were used. The detection P value was calculated by Beadstudio software (Illumina). After normalization, Significance Analysis of Microarray (SAM) was used to detect transcripts that were relatively more or less abundant in one group of samples. Significant genes were those with False Discovery rate of less than 5 percent and fold difference between 2 groups of at least 2 fold.

### 4.3.6 MultiExperiment Viewer (MeV)

MeV version 4.2 was used to analyze microarray data. This software contains Significance Analysis of Microarray (SAM) (Tusher at al. 2001) in its package. SAM was able to provide the False Discovery Rate (FDR) for microarray data, which is the proportion of genes likely to have been identified by chance as being significant. SAM can detect transcripts that were relatively more or less abundant in one group of samples. Significant genes were those with *P* value of less than 0.05 and fold difference between 2 groups of at least 2 fold.

### 4.3.7 InnateDB pathway analysis

#### 4.3.7.1 Pathway & Gene Ontology Analysis of DE Genes using InnateDB

Illumina Probe IDs were mapped to NCBI Entrez Gene IDs and these were used to cross-reference and upload genes represented on the arrays to InnateDB (www.innatedb.com) [168] along with associated gene expression data. A list of pathways mapping to the

uploaded genes was returned and pathway analysis was undertaken to determine which pathways were significantly associated with up- and down-regulated genes using the Hypergeometric distribution test. InnateDB automatically tests for over-representation of differentially expressed (DE) genes in more than 3,000 pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [169], the NCI-Nature Pathway Interaction Database (PID) (http://pid.nci.nih.gov), Integrating Network Objects with Hierarchies (INOH) pathway database (http://www.inoh.org/), NetPath (http://www.netpath.org) and Reactome databases [170]. The Benjamini and Hochberg (BH) FDR correction [171] was applied to correct for multiple testing. Similarly, the InnateDB Gene Ontology analysis tool was used to identify Gene Ontology terms [172] that were significantly associated with DE genes.

#### 4.3.7.2 InnateDB Molecular Interaction Network Analysis

InnateDB pathway analysis can be very powerful in determining which annotated pathways are most significantly associated with DE genes. Pathway analysis, however, relies on using the association of DE genes to known biological pathways, which are often annotated as relatively simple linear cascades. Network analysis has the ability to move the investigation from this simple view of the signalling response to a more comprehensive analysis of the molecular interactions between DE genes and their encoded proteins and RNAs, potentially allowing one to uncover as yet unknown signalling cascades or pathways, functionally relevant sub-networks and the central molecules, or hubs, of these networks. By investigating networks that include interactions between DE genes and their non-differentially expressed interacting partners, one has the potential to identify key regulators of gene expression, even though these regulators themselves may not be differentially expressed but regulated at the posttranscriptional level.

InnateDB contains databases of all human and mouse experimentally-supported molecular interactions and annotation on more than 10,000 manually curated human and mouse innate immunity relevant interactions. InnateDB allows one to upload a gene list of interest along with associated gene expression data, and returns this data integrated in a molecular interaction network context for visualization and further interrogation and analysis.

## 4.4 Results

### 4.4.1 Patient populations

Between June 2006 and Dec-2007, 450 patients with dengue and less than 72 hrs of illness were enrolled into this prospective, hospital-based study. Thirty five patients subsequently progressed to DSS as defined by WHO criteria (pulse pressure $\leq$ 20 mmHg with poor peripheral perfusion and rapid, weak pulse) [1]. The remaining 415 patients were defined as having "uncomplicated dengue" by the recently revised WHO criteria [1] as they did not require any clinical interventions and were managed throughout in/at the general infectious disease ward. Twenty-four DSS patients were chosen for gene expression microarray experiment. For each DSS patient, 1-3 patients with uncomplicated dengue and matched by sex, infecting virus serotype, age (within 2 years) and day of illness were selected as controls for virological investigations and host gene expression profiling. Subsequently, 56 patients with uncomplicated dengue infection were selected as controls. The demographic and virological characteristics of the DSS and matched uncomplicated dengue patients are summarised in table 4.5. The median day of illness was 3 days (range 2-5) and the median time to defervescence was 3 days (range 2-3) (table 4.5). In patients who developed DSS, the median time from enrolment to shock was 2 days (range 1-4).

Table 4.5: Baseline characteristics of patients included in microarray study

| **Clinical phenotypes** | **DSS (n = 24)** | **Uncomplicated dengue (n = 56)** | **P value*** |
|---|---|---|---|
| **Median age (range)** | 11 (2 – 30) | 10.5 (2 - 29) | |
| **Gender (Male (%))** | 16/24 (67%) | 34/56 (61%) | |
| **Median day of illness (range)** | 3 (2 - 5) | 3 (2 - 5) | |
| **Fever day (range)** | -3 (-3 to -2) | -3 (-3 to -2) | |
| **Received i.v. fluids** | 24 (100%) | 0 (0%) | |
| **Median array sample date relative to date of shock (range)** | -2 (-3 – 0) | N/A | |
| **Primary infection**<br>**Unknown** | 1 (4%)<br>1 (4%) | 1 (1.8%)<br>0 | |
| **DENV Serotype**<br>**DENV1**<br>**DENV2**<br>**DENV3**<br>**DENV4**<br>**Unknown** | <br>16 (67%)<br>7 (29%)<br>0 (0%)<br>0 (0%)<br>1 (4%) | <br>41 (73%)<br>15 (27%)<br>0<br>0<br>0 | |
| **Viremia**<br>**Median (range)** | (Missing = 6)<br>3.47E+7<br>(3.11E+4 - 2.15E+9) | (Mising = 14)<br>4.18E+07<br>(1.27E+5 – 1.46E+9) | 0.7 |
| **NS1 ng/ml**<br>**Median (range)** | (Missing= 3)<br>1002 (0.01 – 3842) | (Missing= 17)<br>117 (0.01 – 3415) | 0.048 |
| **Neutrophils**<br>**(x10$^3$ cells/ul)** | (Missing = 1)<br>2.75<br>(0.8 – 5.5) | (Missing = 6)<br>2.48<br>(0.7 – 8.4) | 0.7 |

Note: [a] Fever day is the number of days relative to the time point when the patient became afebrile, defined as fever day 0. * Mann-Whitney test P value.

### 4.4.2 Virological comparisons between patients with DSS versus uncomplicated dengue

Patients with DSS (n=21) had significantly higher plasma NS1 concentrations at the time of enrolment than matched uncomplicated dengue controls (n=39) (table 4.1). Plasma viraemia levels were not significantly different between the two groups (table 4.1). The consensus genome (nt 30-10649) sequences of 15 DENV-1 and 3 DENV-2 from patients with DSS were determined directly from plasma and compared phylogenetically by neighbour joining methods to consensus genome sequences sampled from 31 DENV-1 and 9 DENV-2 from matched patients with uncomplicated dengue (figure 4.2 and figure 4.3).

All DENV-1 sequences belonged to the genotype I lineage and there was no evidence of a phylogenetic structure in the neighbour joining tree that was related to clinical outcome (figure 4.2). All DENV-2 sequences belonged to the Asian-1 lineage and similarly, consensus genome sequences from DSS patients were not phylogenetically distinct from uncomplicated dengue cases (figure 4.3).

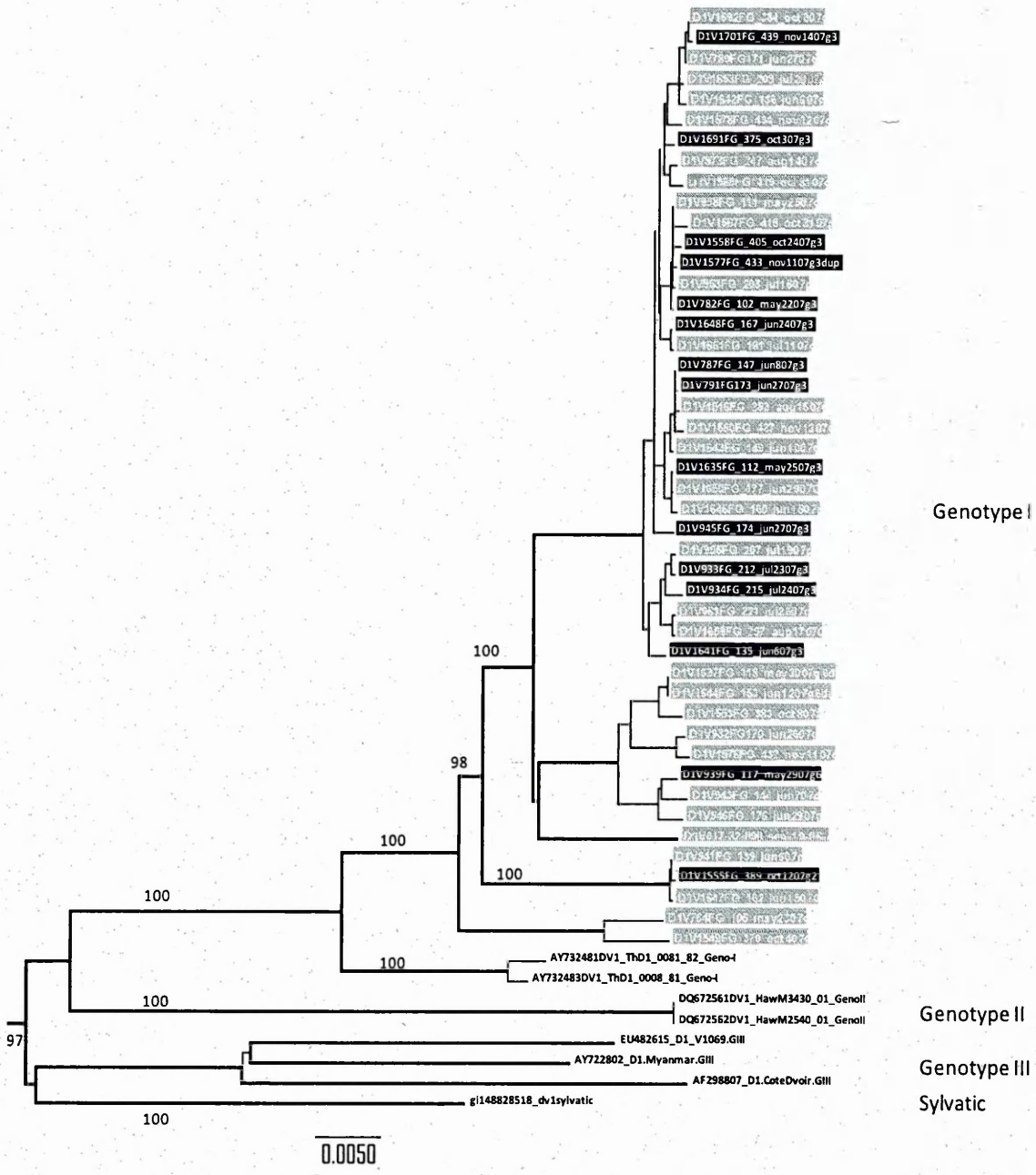



Figure 4.2: Phylogenetic tree (Neighbor Joining method) of DENV-1 genome sequences. The NJ tree contains genome sequences deduced from the plasma of 32 DENV-1 infected patients with uncomplicated dengue (grey highlighted tip labels) and 15 genomes sampled from DENV-1 infected patients with DSS (black highlighted tip labels). Trees are mid-point rooted and contain sequences from other DENV-1 viruses for reference only (black tip labels).

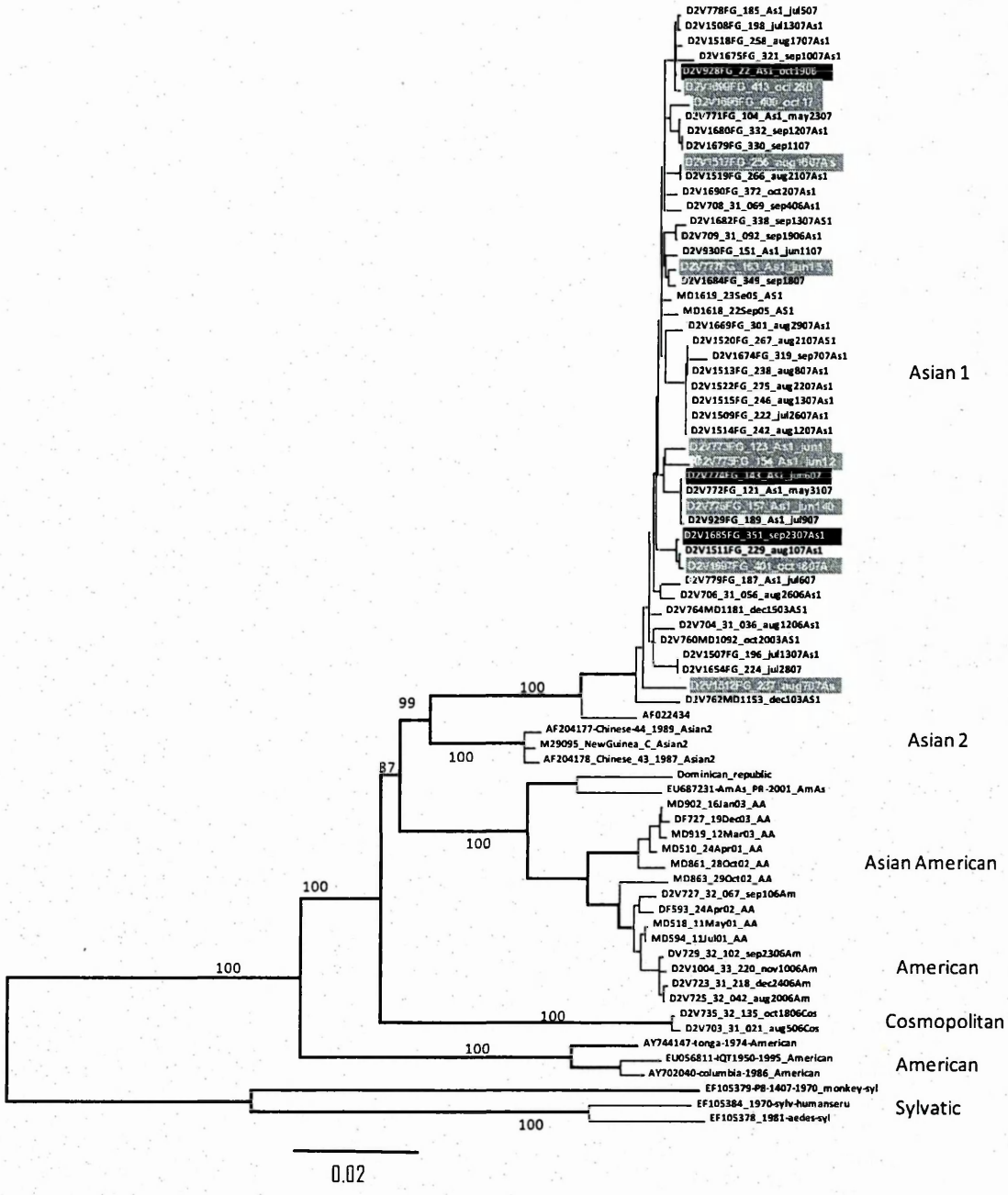



Figure 4.3: Phylogenetic tree (Neighbor Joining method) of DENV-2.The tree contains genome sequences deduced from the plasma of 9 DENV-2 infected patients with uncomplicated dengue (grey highlighted tip labels) and 3 genomes sampled from DENV-2 infected patients with DSS (black highlighted tip labels). Trees are mid-point rooted and contain sequences from other DENV-2 viruses for reference only (black tip labels).

### 4.4.3 Differences in host gene transcript abundance between the early febrile phase of dengue and convalescence

After normalization (see section 4.3.5), we obtained 9870 genes for downstream analysis. We undertook an expansive interrogation of the host transcriptional signature in dengue by comparing transcriptional profiles of 9870 genes in samples collected early in the acute phase (fever day -2 or -3) from 80 acute dengue patients (24 DSS and 56 uncomplicated dengue) with 34 convalescent control samples (18 severe dengue and 16 uncomplicated dengue). By SAM analysis with a FDR <5%, we identified 860 differentially expressed (DE) transcripts (fold change equal to or greater than 2, q value ≤5%). Of the 860 differentially abundant gene transcripts, 309 were less abundant and 551 were more abundant in acute samples relative to convalescent samples. InnateDB pathway, ontology and network analysis of differentially expressed (DE) transcripts was performed to identify the molecular pathways, networks and biological processes that dominate the whole blood transcriptional profile, and in particular, the major signalling hubs involved in the early immune response.

### 4.4.4 Pathways and Gene Ontology terms associated with up-regulated genes

Many of the over-abundant DE transcripts are interferon-inducible. Indeed of the 551 up-regulated genes, 173 (31%) were annotated as Type I interferon-inducible by the Interferome database [173]. Pathways that were significantly over-represented in up-regulated genes after correction for multiple testing (FDR < 5%) included *Systemic lupus erythematosus* (KEGG database); *Classical complement pathway* (PID BIOCARTA) and *Complement and coagulation cascades* (KEGG). The pathway *Systemic lupus erythematosus* is likely significant due to the overlap between this pathway and the complement pathways and because of the large number of histone genes that are up-regulated. A number of complement and coagulation related genes were found to be up-

regulated (C1QA; C1QB; C1QC; C2; C3AR1; C5; F9; PLAUR; PROS1; SERPINA1; SERPING1). Prior to correction for multiple testing (which is conservative given the large number of pathways tested) a range of pathways was identified as being significantly over-represented. These pathways included Toll-like receptor signaling pathway, IL27-mediated signaling events, IL12-mediated signaling events, Chemokine signaling pathway, Cytokine-cytokine receptor interaction, Lysosome, RIG-I-like receptor signaling pathway, IFN-gamma pathway and the JAK-STAT pathway and regulation pathway. Many of these are likely of biological significance (despite being below the statistical threshold). The top 20 most significant pathways are highlighted in table 4.6.

The most significant GO terms associated with up-regulated genes (FDR <5%) included response to virus, immune response, innate immune response and inflammatory response. A full list of GO terms associated with up-regulated genes is described in table 4.7.

### 4.4.5 Pathways and Gene Ontology terms associated with down-regulated genes

Almost all of the significantly down-regulated pathways were related to translation (table 4.8). All these pathways are significant due to the down-regulation of genes encoding ribosomal subunit proteins. The pathway hemoglobins chaperone was also significantly over-represented due to the down-regulation of several genes involved in the heme biosynthesis pathway (ALAD; ALAS2; ERAF; FECH; HBB; HMBS). GO analysis paints a similar picture to the pathway analysis with a number of terms related to translation all being significant (FDR < 5%) (Table 4.9).

Table 4.6: The top 20 pathways that were significantly over-represented in up-regulated genes in acute dengue patients relative to convalescent dengue patients

| Pathway Name | Pathway Id | Genes Ratio (%) | Pathway p-value | Pathway p-value (corrected) |
|---|---|---|---|---|
| Systemic lupus erythematosus | 2805 | 25 | 4.32E-11 | 2.79E-08 |
| Ribosome | 474 | 22 | 4.32E-08 | 1.39E-05 |
| Viral mRNA Translation | 1921 | 19 | 6.34E-07 | 1.37E-04 |
| Eukaryotic Translation Termination | 1706 | 19 | 7.44E-07 | 9.61E-05 |
| Peptide chain elongation | 1248 | 19 | 7.44E-07 | 9.61E-05 |
| Formation of a pool of free 40S subunits | 1923 | 17 | 2.83E-06 | 3.05E-04 |
| L13a-mediated translational silencing of Ceruloplasmin expression | 1958 | 16 | 3.10E-06 | 2.86E-04 |
| GTP hydrolysis and joining of the 60S ribosomal subunit | 1919 | 16 | 3.53E-06 | 2.85E-04 |
| Hemoglobins chaperone | 4185 | 50 | 1.23E-05 | 8.80E-04 |
| Classical complement pathway | 3972 | 36 | 4.66E-04 | 3.01E-02 |
| Complement and coagulation cascades | 456 | 16 | 5.25E-04 | 3.08E-02 |
| IL4 | 3925 | 18 | 9.81E-04 | 5.28E-02 |
| Signaling events mediated by Stem cell factor receptor (c-Kit) | 5696 | 17 | 1.13E-03 | 5.63E-02 |
| Porphyrin and chlorophyll metabolism | 603 | 21 | 1.17E-03 | 5.38E-02 |
| IL4-mediated signaling events | 4210 | 16 | 1.22E-03 | 5.27E-02 |
| JAK degradation signalling | 342 | 60 | 1.23E-03 | 4.68E-02 |
| Negative regulation of (Phosphorylation of JAK) in JAK STAT pathway | 56 | 60 | 1.23E-03 | 4.68E-02 |
| Glypican 2 network | 5681 | 100 | 2.61E-03 | 9.38E-02 |
| Cell surface interactions at the vascular wall | 1687 | 18 | 2.77E-03 | 9.40E-02 |
| BH3-only proteins associate with and inactivate anti-apoptotic BCL-2 members | 1930 | 43 | 3.99E-03 | 1.23E-01 |

Notes: Gene ratio is was calculated by dividing the number of molecules in a given pathway in the input gene list to the total number of molecules that make up that pathway; Pathway P value was calculated based on Hypergeometric distribution test; The Benjamini and Hochberg (BH) FDR was applied to correct for multiple testing

Table 4.7: The top 20 most significant GO terms associated with up-regulated genes in acute dengue patients relative to convalescent dengue patients

| GO term Name | GO term Id | Genes Ratio | GO term p-value | GO term p-value (corrected) |
|---|---|---|---|---|
| Response to virus [biological_process] | GO:0009615 | 24 | 1.01E-15 | 1.19E-12 |
| Immune response [biological_process] | GO:0006955 | 8 | 1.38E-13 | 8.20E-11 |
| Innate immune response [biological_process] | GO:0045087 | 9 | 1.53E-13 | 6.07E-11 |
| Inflammatory response [biological_process] | GO:0006954 | 9 | 1.55E-07 | 4.59E-05 |
| Chromosome [cellular_component] | GO:0005694 | 10 | 2.75E-07 | 6.54E-05 |
| Nucleosome [cellular_component] | GO:0000786 | 9 | 9.31E-07 | 1.84E-04 |
| Nucleosome assembly [biological_process] | GO:0006334 | 8 | 3.78E-06 | 6.42E-04 |
| Chemotaxis [biological_process] | GO:0006935 | 9 | 1.20E-05 | 1.79E-03 |
| Sugar binding [molecular_function] | GO:0005529 | 9 | 1.49E-05 | 1.96E-03 |
| Hematopoietin/interferon-class…* | GO:0005062 | 8 | 5.15E-05 | 6.11E-03 |
| Extracellular space [cellular_component] | GO:0005615 | 5 | 7.26E-05 | 7.83E-03 |
| Cell-cell signaling [biological_process] | GO:0007267 | 7 | 9.06E-05 | 8.96E-03 |
| Integral to plasma membrane [cellular_component] | GO:0005887 | 4 | 1.00E-04 | 9.15E-03 |
| Complement activation, classical pathway [biological_process] | GO:0006958 | 18 | 1.29E-04 | 1.09E-02 |
| Signal transducer activity [molecular_function] | GO:0004871 | 6 | 2.40E-04 | 1.90E-02 |
| I-kappaB kinase/NF-kappaB cascade [biological_process] | GO:0007249 | 19 | 3.78E-04 | 2.80E-02 |
| Complement component C1 complex [cellular_component] | GO:0005602 | 100 | 5.57E-04 | 3.89E-02 |
| Lysosome organization [biological_process] | GO:0007040 | 24 | 5.73E-04 | 3.78E-02 |
| Membrane fusion [biological_process] | GO:0006944 | 22 | 7.23E-04 | 4.51E-02 |
| Protein binding [molecular_function] | GO:0005515 | 3 | 8.30E-04 | 4.92E-02 |

Note: *Hematopoietin/interferon-class (D200-domain) cytokine receptor signal transducer activity [molecular_function]; Gene ratio is was calculated by dividing the number of molecules in a given GO term in the input gene list to the total number of molecules that make up that GO term; GO term P value was calculated based on Hypergeometric distribution test; The Benjamini and Hochberg (BH) FDR was applied to correct for multiple testing

Table 4.8: The top 20 significantly down-regulated pathways in acute dengue patients relative to convalescent dengue patients

| Pathway Name | Pathway Id | Genes Ratio (%) | Pathway p-vaiue | Pathway p-value (corrected) |
|---|---|---|---|---|
| Ribosome | 474 | 21 | 7.34E-15 | 2.25E-12 |
| Viral mRNA Translation | 1921 | 18 | 1.43E-13 | 2.19E-11 |
| Eukaryotic Translation Termination | 1706 | 18 | 1.71E-13 | 1.31E-11 |
| Peptide chain elongation | 1248 | 18 | 1.71E-13 | 1.31E-11 |
| L13a-mediated translational silencing of Ceruloplasmin expression | 1958 | 16 | 3.67E-13 | 2.25E-11 |
| GTP hydrolysis and joining of the 60S ribosomal subunit | 1919 | 15 | 4.28E-13 | 2.18E-11 |
| Formation of a pool of free 40S subunits | 1923 | 16 | 7.86E-13 | 3.44E-11 |
| Hemoglobins chaperone | 4185 | 50 | 2.69E-08 | 1.03E-06 |
| Nitrogen metabolism | 549 | 17 | 8.31E-04 | 2.83E-02 |
| Ribosomal scanning and start codon recognition | 1527 | 9 | 1.01E-03 | 2.81E-02 |
| Translation initiation complex formation | 1685 | 9 | 1.01E-03 | 2.81E-02 |
| Porphyrin and chlorophyll metabolism | 603 | 12 | 2.82E-03 | 7.20E-02 |
| Negative regulation of (Phosphorylation of JAK) in JAK STAT pathway | 56 | 40 | 3.16E-03 | 7.43E-02 |
| Formation of the ternary complex, and subsequently, the 43S complex | 1396 | 9 | 3.32E-03 | 7.26E-02 |
| Heme biosynthesis | 1947 | 33 | 4.68E-03 | 9.54E-02 |
| Aspartate, asparagine, glutamate, and glutamine metabolism | 1638 | 29 | 6.47E-03 | 1.16E-01 |
| IL-4 signaling pathway(JAK1 JAK3 STAT6) | 378 | 29 | 6.47E-03 | 1.16E-01 |
| Bicarbonate transporters | 6581 | 22 | 1.08E-02 | 1.74E-01 |
| Prostanoid ligand receptors | 5579 | 22 | 1.08E-02 | 1.74E-01 |
| Il-2 receptor beta chain in t cell activation | 4011 | 8 | 1.17E-02 | 1.79E-01 |

Table 4.9: Top 20 significant down-regulated GO terms in acute dengue patients relative to convalescent dengue patients

| GO term Name | GO term Id | Genes Ratio (%) | GO term p-value | GO term p-value (corrected) |
|---|---|---|---|---|
| Translational elongation [biological_process] | GO:0006414 | 16 | 2.39E-14 | 1.87E-11 |
| Hemoglobin complex [cellular_component] | GO:0005833 | 50 | 6.40E-09 | 2.50E-06 |
| Structural constituent of ribosome [molecular_function] | GO:0003735 | 6 | 2.27E-08 | 5.93E-06 |
| Ribosome [cellular_component] | GO:0005840 | 6 | 8.86E-08 | 1.73E-05 |
| Translation [biological_process] | GO:0006412 | 5 | 8.39E-07 | 1.31E-04 |
| Oxygen transport [biological_process] | GO:0015671 | 31 | 2.07E-06 | 2.69E-04 |
| Heme biosynthetic process [biological_process] | GO:0006783 | 31 | 2.50E-05 | 2.44E-03 |
| Oxygen transporter activity [molecular_function] | GO:0005344 | 31 | 2.50E-05 | 2.44E-03 |
| Cytosolic large ribosomal subunit [cellular_component] | GO:0022625 | 14 | 2.55E-05 | 2.21E-03 |
| Hemoglobin metabolic process [biological_process] | GO:0020027 | 60 | 2.71E-05 | 1.93E-03 |
| Tetrapyrrole biosynthetic process [biological_process] | GO:0033014 | 60 | 2.71E-05 | 1.93E-03 |
| Oxygen binding [molecular_function] | GO:0019825 | 15 | 9.23E-05 | 6.01E-03 |
| Prostaglandin J receptor activity [molecular_function] | GO:0001785 | 100 | 1.98E-04 | 1.19E-02 |
| Iron ion binding [molecular_function] | GO:0005506 | 4 | 2.56E-04 | 1.43E-02 |
| Cortical cytoskeleton [cellular_component] | GO:0030863 | 30 | 3.09E-04 | 1.61E-02 |
| Heme binding [molecular_function] | GO:0020037 | 6 | 3.97E-04 | 1.94E-02 |
| Prostaglandin D receptor activity [molecular_function] | GO:0004956 | 67 | 5.89E-04 | 2.71E-02 |
| Response to cold [biological_process] | GO:0009409 | 23 | 7.14E-04 | 3.10E-02 |
| Hemoglobin binding [molecular_function] | GO:0030492 | 40 | 1.93E-03 | 7.93E-02 |
| Small ribosomal subunit [cellular_component] | GO:0015935 | 15 | 2.65E-03 | 1.03E-01 |

### 4.4.6 Network Analysis of Differential Gene Expression Profiles in Acute Vs. Convalescent dengue

InnateDB (www.innatedb.com) [168] was used to generate molecular interaction networks. The first was a network consisting of only the interactions between differentially expressed genes and their encoded products (AcuteDengue_DE network, figure 4.4), while the second expanded upon this network by including all non-differentially expressed interacting partners of the DE genes (AcuteDengue_All network, figure 4.5). The AcuteDengue_DE network has 289 nodes and 429 edges, in comparison to the more extensive AcuteDengue_All network, which had 4,364 nodes and 9,094 edges.

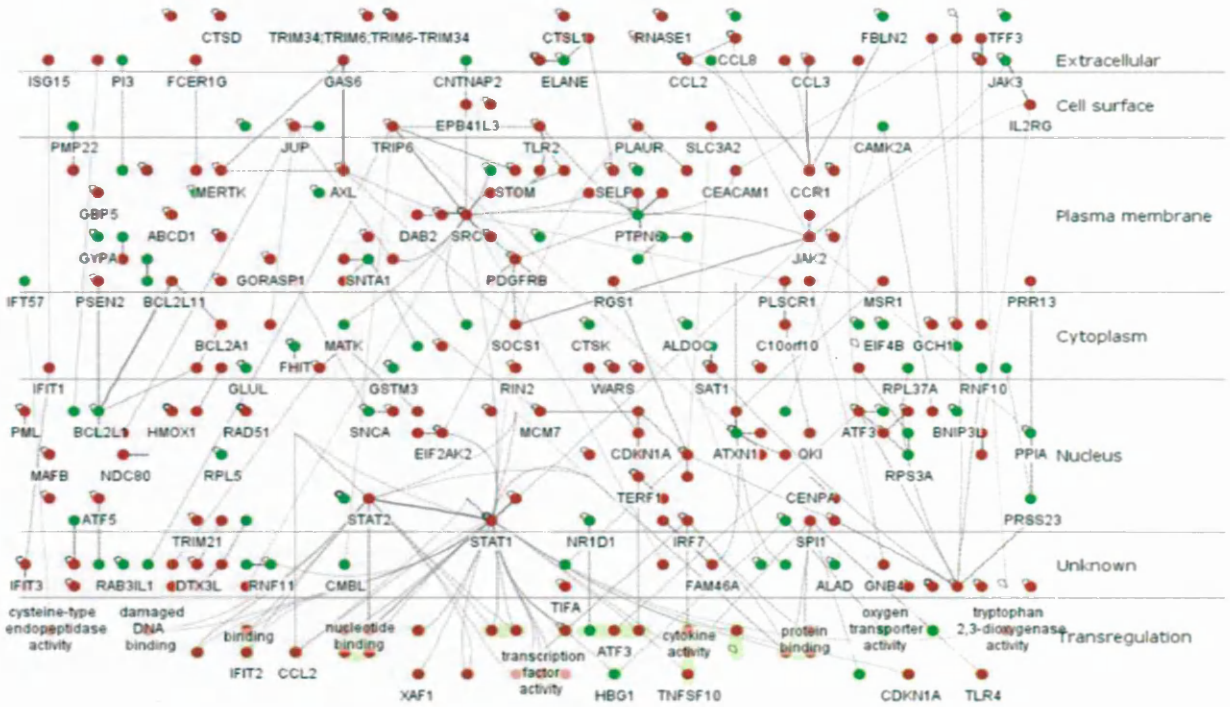



Figure 4.4: AcuteDengue_DE_network - A Network of known protein-protein and protein-DNA interactions encoded by genes differentially expressed in acute Dengue patients in comparison to convalescent samples. Nodes encoded by up-regulated genes are shown in red; down-regulated in green. This network was generated using InnateDB (www.innateDB.com) and was visualised using the Cerebral 2.0 plug-in for Cytoscape 2.6.2 which was developed as part of the InnateDB project. This network has 289 nodes and 429 edges. The top 5 hubs (i.e. genes/proteins that are highly connected to other DE genes) in this network were identified as the transcription factors STAT1 and STAT2 (2X up-regulated), the tyrosine kinase SRC (2X up-regulated), PTPN6 (SHP1) (2.5X down-regulated) and C1orf103 (2X up-regulated).

The AcuteDengue_DE network was analysed to identify network hubs and bottlenecks which may represent the key regulatory nodes in the network. Using the "Degree" algorithm from the cytoHubba plug-in [174] the top 5 hubs (i.e. genes/proteins that are highly connected to other DE genes) in this network were identified as the transcription factors STAT1 and STAT2 (2X up-regulated), the tyrosine kinase SRC (2X up-regulated), PTPN6 (SHP1) (2.5X down-regulated) and C1orf103 (2X up-regulated). The Hubba software also allows one to predict proteins that act as bottlenecks in the network. Bottlenecks are network nodes that are the key connector proteins in a network and have many "shortest paths" going through them, similar to bridges or tunnels on a highway map [175]. Several of the hubs including STAT1, SRC and PTPN6 (SHP1), were identified as bottlenecks in the network, further supporting their central role in signalling. TRIP6 and JAK2 were also identified in the top 5 bottlenecks. Analysis of the AcuteDengue_All network, which consists of all interactions involving molecules encoded by DE genes (regardless of whether interacting molecule is DE), also identified SRC, STAT1, TRIP6, SHP1, C1orf103 and JAK2 in the top 20 hub/bottleneck nodes (figure 4.5). SRC is the highest-ranked hub in this network and SRC and STAT1 are the top 2 bottlenecks.

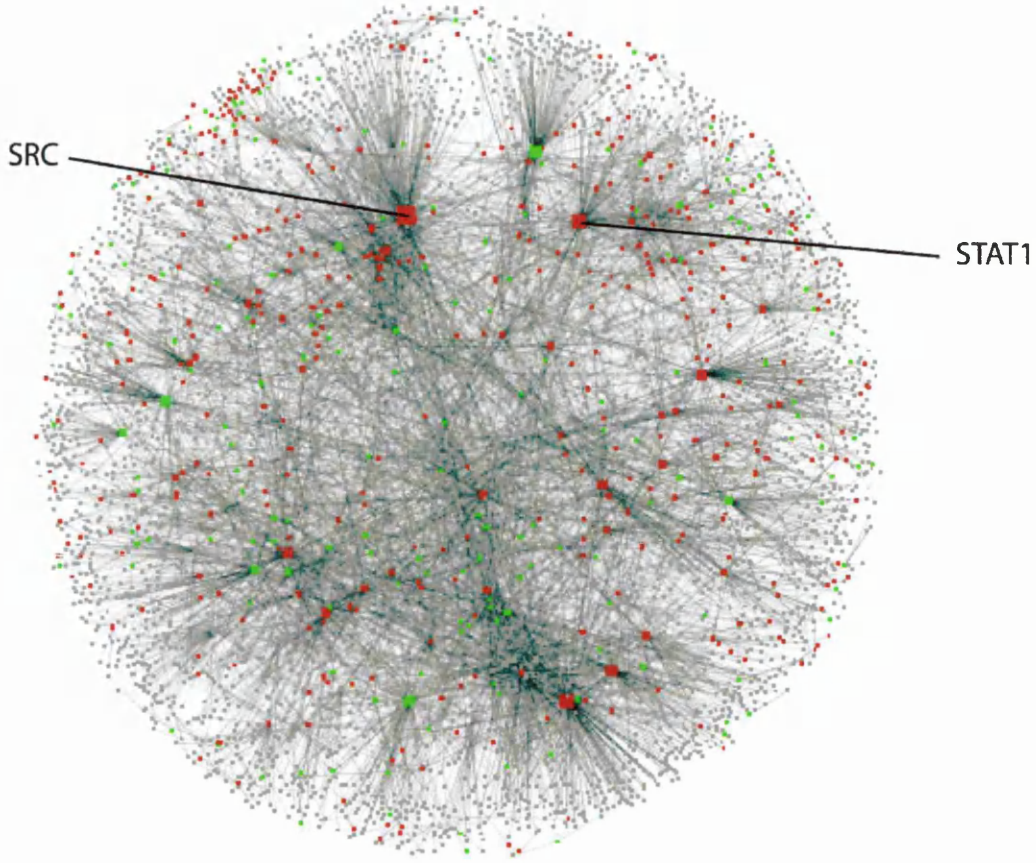



Figure 4.5: AcuteDengue_ALL_network - InnateDB Network of known protein-protein and protein-DNA interactions encoded by DE genes and all known interacting partners in acute Dengue patients in comparison to convalescent control samples. Nodes encoded by up-regulated genes are shown in red; down-regulated in green. Analysis of this network enables the identification of central regulators (hubs/bottlenecks) that are not necessarily regulated at the transcriptional level.

Two major differentially expressed sub-networks were identified in the AcuteDengue_DE network. The top-ranked network consisted of 23 nodes (including JAK2, JAK3, SRC, TLR2, IL2RG, SOCS1, SHP1, TRIP6 and many other JAK/STAT and SRC regulators). The second-ranked sub-network, also consisting of 23 nodes, was enriched for nuclear-localised proteins (13 nodes) and in terms such as transcription co-repressor activity (ATF3; DDIT3 (CEBPZ); ID3; NFIL3) and may represent an important transcriptional regulatory network.

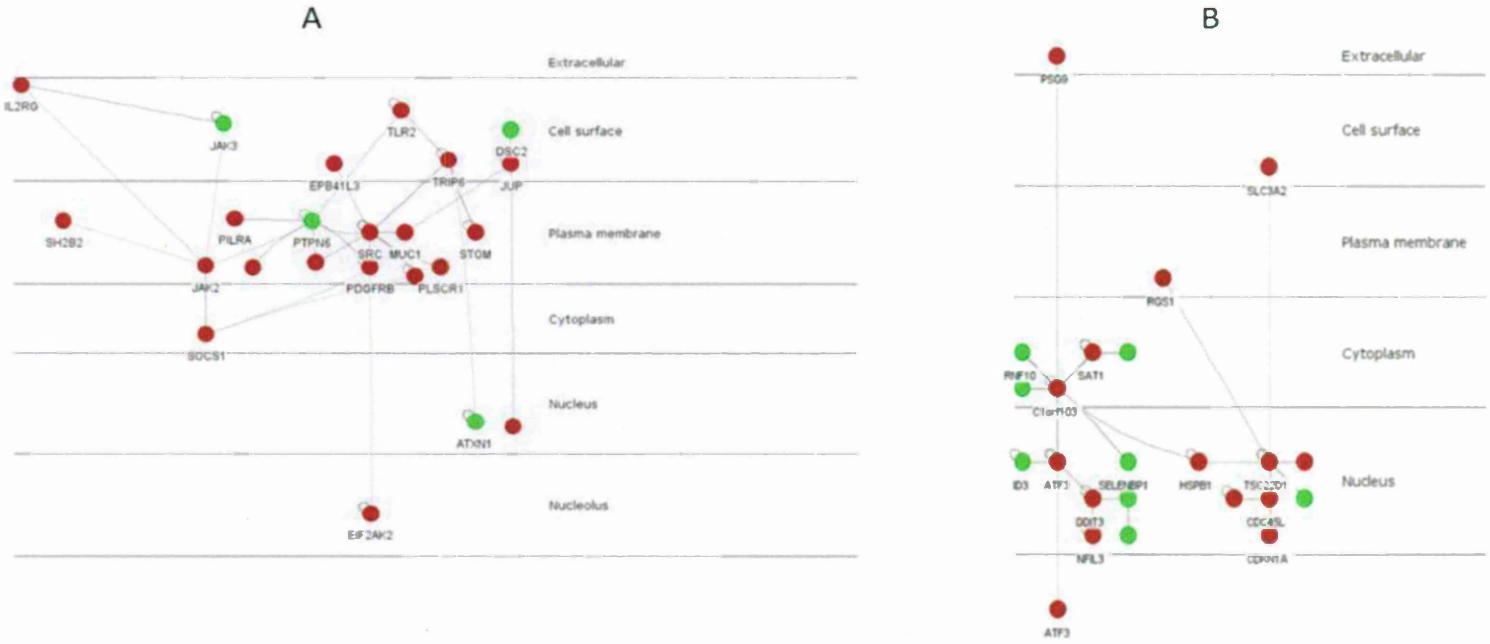



Figure 4.6: Two major differentially expressed sub-networks were identified in the AcuteDengue_DE network. Shown in A) is the top-ranked network consisted of 23 nodes (including JAK2, JAK3, SRC, TLR2, IL2RG, SOCS1, SHP1, TRIP6 and many other JAK/STAT and SRC regulators). In B) is the second ranked sub-network, also consisting of 23 nodes.

### 4.4.7 Combining the Transcriptional Regulatory Network and the Physical Interaction Network

InnateDB was also used to construct a network of all predicted transcription factor (TF) interactions with DE genes (TF_network). This predicted transcriptional regulatory network was then merged with the AcuteDengue_ALL network. This combined network provides a much more comprehensive picture of the connection between signalling and transcriptional regulation (figure 4.7).

Nodes with significantly more interactions with DE genes than expected by chance were highly enriched in transcription factors involved in IFN/NFkB signalling responses including STAT1, STAT2, STAT3, IRF7, IRF9, IRF1, CEBPB and SP1. IRF1, for example, although not DE itself, is predicted to interact with 100 DE genes. Similarly, CEBPB is predicted to interact with 88 genes and is not itself DE. Gene expression analysis alone would not reveal the multitude of transcription factors that are likely the key regulators driving the host response to dengue.

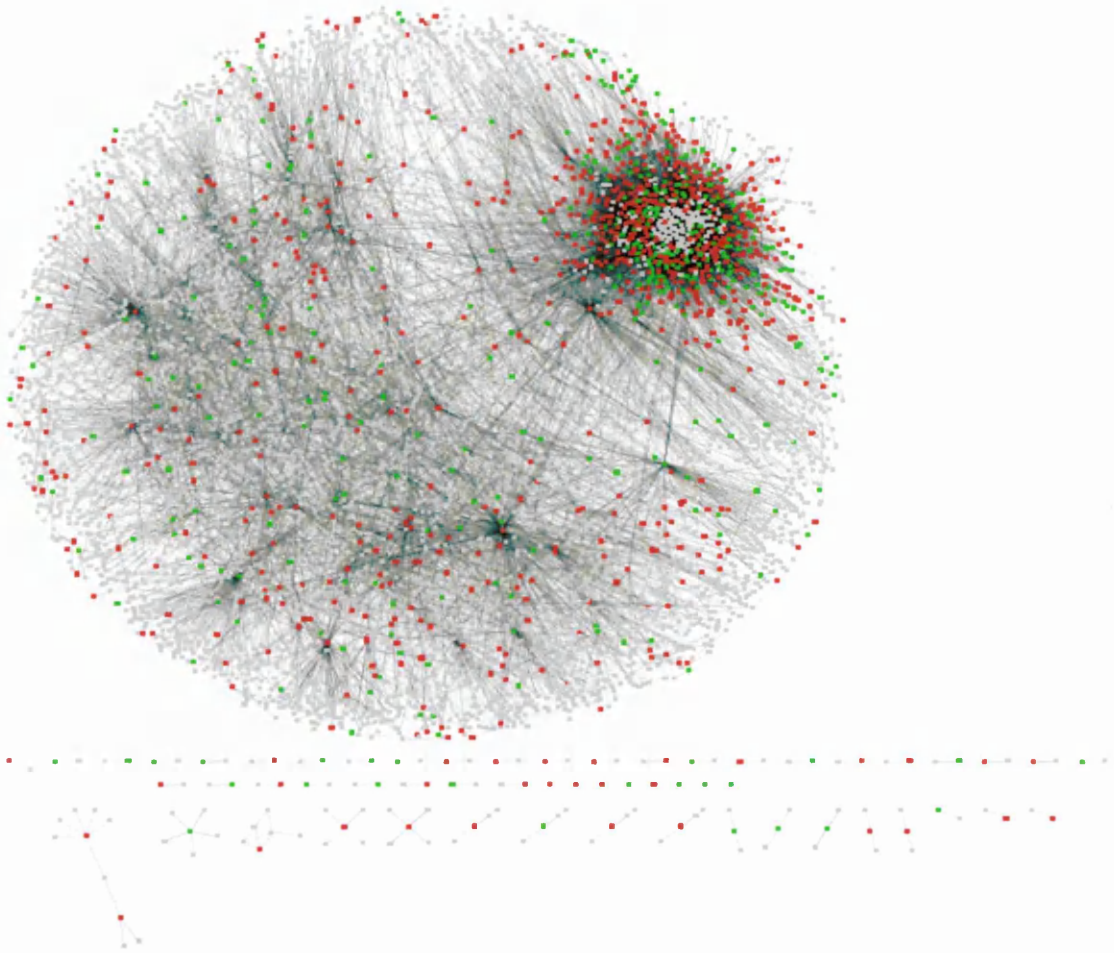


Figure 4.7: InnateDB Network of known protein-protein and protein-DNA interactions encoded by DE genes and all known interacting partners merged with the network of CisRed predicted transcription factor-gene interactions involving DE genes in dengue acute samples in comparison to convalescent control samples. Nodes encoded by up-regulated genes are shown in red; down-regulated in green. Note: Nodes in this network are not necessarily DE.

### 4.4.8 Differences in early host gene transcript abundance between DSS and uncomplicated dengue patients

The identification of prognostic markers of severe dengue is an important goal. We therefore compared gene transcript abundance on fever day -2 or -3 in the 24 DSS patients and their 56 matched controls. By SAM analysis, twenty one transcripts were significantly enriched in DSS patients relative to uncomplicated dengue patients (FDR <5%, >2-fold difference in abundance). There were no significantly down-regulated transcripts in acute DSS patients compared to acute uncomplicated dengue patients. Table 4.10 summarises the list of 21 differentially abundant transcripts. Prominent amongst these were transcripts that could be linked to activated neutrophils, with 12 of the 21 differentially expressed genes associated with activated neutrophils in the shape of membrane-bound integrin receptors (CEACAM6, CEACAM8), cytokine decoy receptor (IL1R2), secreted proteases (CTSG, ELA2), inflammatory molecules (S100A12), secreted anti-microbial proteins/peptides (DEF1, DEF4, BPI, CAMP, PGLYRP1) or oxidative enzymes (MPO). We explored possible functional relationships between the 21 differentially expressed genes using the Ingenuity Pathway analysis (IPA) knowledge base. Unsupervised IPA network analysis identified two clusters of 35 genes each that included 18 of the 21 differentially expressed genes (figure 4.8). These two network interactions were highly unlikely to have occurred by chance ($P$=$10^{-31}$ and $P$ =$10^{-21}$). The first cluster included 12 differentially expressed genes and the second cluster included 9 differentially expressed genes. The neutrophil-associated CAMP and MPO, together with the decoy receptor IL1R2, were found in both networks. Cathepsin G (CTSG) and Elastase (ELA), which were found in the list of 21 DE genes, were the key elements of the first network while TNF-α and β-estrodiol were the key elements of the second network. Strikingly, these data suggest gene expression profiles from members of two overlapping networks discriminate between patients who progress to DSS from those with an uncomplicated disease course.

The differential abundance of neutrophil-associated transcripts in patients who progressed to DSS was not simply a reflection of neutrophil levels in the sample as the mean absolute count of neutrophils in each group was not significantly different from one another (median (95%CI) DSS patients $2.75*10^3/mm^3$ (2.22 – 3.33) versus $2.48*10^3/mm^3$ (2.27–3.10) in uncomplicated dengue cases). These results indicated neutrophils are phenotypically activated in children who subsequently develop DSS and suggests certain neutrophil-associated transcripts could have prognostic value in identifying patients at risk of severe disease.

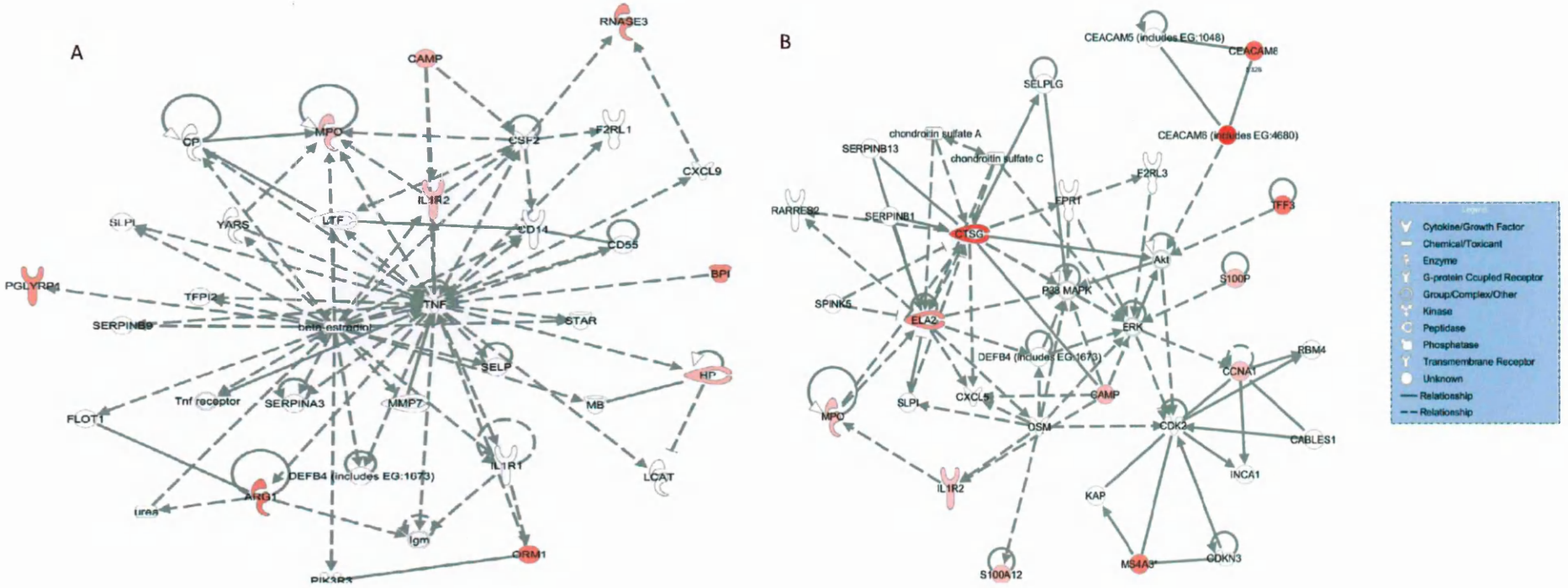



Figure 4.8: Ingenuity Pathway Analysis (IPA) of differentially expressed genes between DSS and uncomplicated dengue patients. Twenty one differentially expressed transcripts were analyzed using IPA. Two significant networks were identified: A) Antigen presentation, cell-mediated immune response, humoral immune response (Network 2, score 21) and B) Cancer, cell cycle, cell-mediated immune response (network 1, score 31). The lines between genes represent known interactions, with solid lines representing direct interactions and dashed lines representing indirect interactions. Differentially expressed genes are highlighted in red, non-highlighted genes were identified by IPA. The high scores associated with these networks indicates they were highly unlikely to be formed by chance.

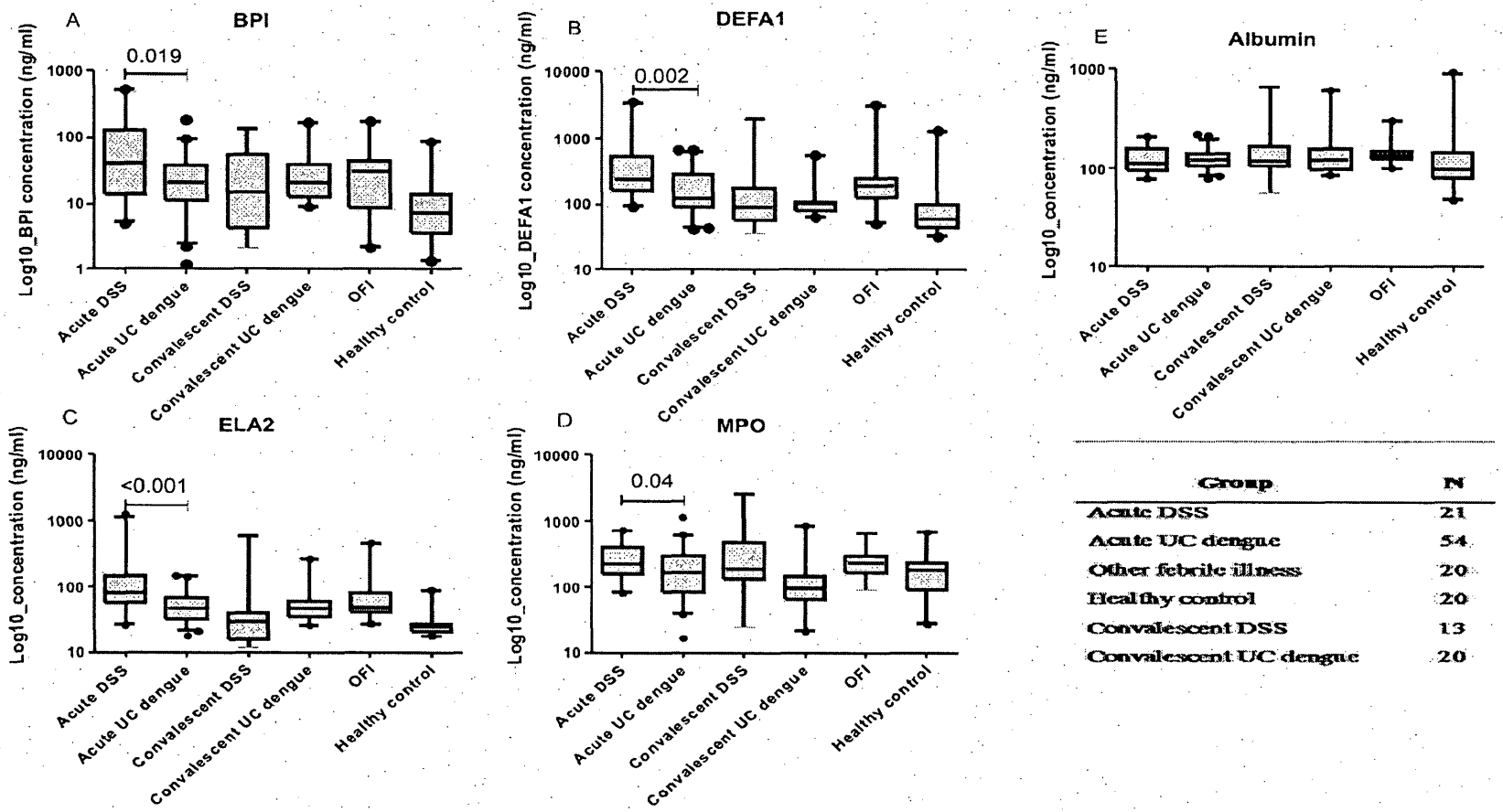
Table 4.10: Differential expressed transcripts in acute DSS patients relative to uncomplicated dengue patients at fever day -2 or -3

| Symbol | Gene description | SAM analysis | | Acute DSS | | | Acute UC dengue | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Fold Change | q-value (%) | # Samples detected (%) | Mean | Std | # Samples detected (%) | Mean | Std |
| DEFA4 | Defensin, alpha 4, corticostatin | 6.7 | 0 | 22 (92) | 5.5 | 17.4 | 39 (70) | 0.8 | 0.8 |
| CEACAM6 | Carcinoembryonic antigen-related cell adhesion molecule 6 | 6.5 | 0 | 12 (50) | 8.3 | 24.4 | 13 (23) | 1.3 | 2 |
| CTSG | Cathepsin G | 6 | 2.6 | 17 (71) | 5.8 | 15.5 | 23 (41) | 1 | 1.1 |
| CEACAM8 | Carcinoembryonic antigen-related cell adhesion molecule | 5.3 | 0 | 22 (92) | 5.5 | 14 | 39 (70) | 1 | 1.3 |
| ARG1 | Arginase, liver | 4.5 | 0 | 20 (83) | 4.9 | 10.8 | 26 (46) | 1.1 | 0.9 |
| MS4A3 | membrane-spanning 4-domains, subfamily A, member 3 | 4.5 | 0 | 20 (83) | 5 | 13.8 | 37 (66) | 1.1 | 1.2 |
| BPI | Bactericidal/permeability-increasing protein | 3.8 | 0 | 21 (88) | 4.8 | 12.1 | 37 (66) | 1.2 | 1 |
| ELA2 | Elastase 2, neutrophil | 3.8 | 0 | 23 (96) | 3.6 | 7.3 | 48 (86) | 0.9 | 0.8 |
| PGLYRP1 | Peptidoglycan recognition protein 1 | 3.7 | 0 | 24 (100) | 3.3 | 7.2 | 56 (100) | 0.9 | 0.8 |
| RNASE3 | Ribonuclease, RNase A family, 3 (eosinophil cationic protein) | 3.6 | 0 | 12 (50) | 4 | 8.8 | 30 (54) | 1.1 | 0.7 |
| MS4A3 | Membrane-spanning 4-domains, subfamily A, member 3 | 3.2 | 4.4 | 20 (83) | 3.5 | 9.6 | 38 (68) | 1.1 | 0.8 |
| LOC728358 | Defensin alpha 1 | 2.7 | 0 | 24 (100) | 2.7 | 4.1 | 56 (100) | 1 | 1.2 |
| HP | Haptoglobin | 2.5 | 0 | 24 (100) | 4.6 | 5.2 | 53 (95) | 1.8 | 1.7 |
| S100P | S100 calcium binding protein P | 2.3 | 2.4 | 24 (100) | 2.9 | 3.4 | 43 (77) | 1.3 | 1.7 |
| CCNA1 | Cyclin A1 | 2.1 | 4.7 | 22(92) | 25.8 | 38.3 | 46 (82) | 12.2 | 13.4 |
| S100A12 | S100 calcium binding protein A12 | 2.1 | 0 | 24 (100) | 2.5 | 2.2 | 54 (96) | 1.2 | 1.1 |
| CAMP | Cathelicidin antimicrobial peptide | 2.4 | 0 | 24 (100) | 2.7 | 3.7 | 56 (100) | 1.1 | 0.9 |
| TFF3 | Trefoil factor 3 (intestinal) | 4.9 | 0 | 11 (46) | 3.3 | 7.8 | 8 (14) | 0.7 | 0.6 |
| ORM1 | Orosomucoid 1 | 4.5 | 0 | 9 (38) | 4.3 | 11.1 | 12 (21) | 1 | 2.4 |
| MPO | Myeloperoxidase | 2.2 | 4.3 | 24 (100) | 3.1 | 6.3 | 46 (82) | 1.4 | 1 |
| IL1R2 | Interleukin 1 receptor, type II | 2 | 2.1 | 21 (88) | 4.1 | 4.1 | 43 (77) | 2 | 1.8 |

### 4.4.9 Plasma concentrations of Bactericidal/permeability-increasing protein (BPI), Defensin-1α, Elastase 2 and Myeloperoxidase (MPO)

To determine if the transcriptional profile of neutrophil activation extended to the plasma proteome, the plasma concentrations of BPI, DEF1A, ELA2 and MPO were measured in plasma samples collected at the same study enrolment time point as the RNA used for expression array analysis. Plasma concentrations of all four proteins were significantly higher at the time of study enrolment in those children who developed DSS relative to children with uncomplicated dengue, although the absolute difference was small (figure 4.9A-D). Plasma concentrations of BPI, DEF1A and ELA2 were also significantly higher in children who developed DSS compared to convalescent samples and healthy donor samples

To understand if elevated concentrations of these neutrophil-associated proteins were independent of early haemoconcentration in children with DSS (i.e. reduced vascular volume leading to higher plasma protein concentrations), plasma albumin levels in all samples were measured as a surrogate marker of the plasma protein concentration (figure 4.9E). Levels of BPI, ELA2 and MPO were not significantly correlated with the plasma albumin concentration in the same sample, suggesting their elevated levels in plasma were not merely a reflection of haemoconcentration at this time point (Pearson's correlation; ([Plasma albumin] vs. [BPI], ($r = -0.04$, $P = 0.6$), vs. ELA2 ($r=-0.41$, $P = 0.6$), or vs. MPO ($r=-0.008$, $P = 0.9$). In contrast, concentrations of DEF1A were weakly correlated with the plasma albumin concentration ([Plasma albumin] vs. DEF1A ($r=-0.176$, $P = 0.014$), suggesting at least some component of this measurement might have been influenced by the reduced vascular volume.




| Group | N |
|---|---|
| Acute DSS | 21 |
| Acute UC dengue | 54 |
| Other febrile illness | 20 |
| Healthy control | 20 |
| Convalescent DSS | 13 |
| Convalescent UC dengue | 20 |

Figure 4.9: Concentrations of secreted neutrophil-associated proteins in plasma. Concentrations of **A)** BPI, **B)** DEF1A, **C)** ELA2, **D)** MPO and **E)** albumin in acute and convalescent dengue cases, and for reference, from patients with other febrile illness and healthy donor plasma samples. The box and whisker plots represent median and inter-quartile ranges.

## 4.5 Discussion

This case-control study investigated the early whole blood transcriptional signature in children who subsequently developed DSS, the most common life-threatening complication of dengue in children. This study was rooted in clinical practice by focusing on the dengue syndrome that always requires a clinical intervention, often in the setting of the intensive care unit. Strikingly, we identified in the first few days of illness, two overlapping gene networks that distinguished patients who developed DSS from those with uncomplicated dengue. A feature of these networks was genes associated with neutrophil activation and degranulation, suggesting a hitherto unrecognised association of neutrophils with pathogenesis and expression of the overall disease phenotype.

Previous studies from our groups have described the whole blood transcriptional signature in dengue patients by microarray analysis [78, 144, 176]. The timing of sample collection is clearly a major factor in the transcriptional signature, with samples collected during the febrile phase having a characteristic anti-viral profile, e.g. with interferon-stimulated genes highly prominent [78, 144, 176], whilst those collected during the afebrile stage had a predominantly metabolic profile [78, 176]. Studies of PBMC (i.e. minus the neutrophil population) have also been described [142, 143]. In short however, all previous microarray studies of the dengue host response have been relatively small in size, collected samples at heterogenous time-points and rarely included patients with DSS, the commonest life-threatening complication in children. Against this backdrop, the current study has a number of strengths, notably the matched case-control design, large sample size, early sampling prior to defervescence and cardiovascular decompensation and inclusion of detailed genomic scale information on the infecting pathogen.

Previous epidemiological, in vitro and in vivo studies have suggested phenotypic differences can exist between virus lineages of the same serotype [100, 164, 177]. We determined virus genome sequences and demonstrated that viruses from DSS patients in this study were not phylogenetically different from those in patients with uncomplicated dengue, implying that host factors were more important determinants of the clinical course. In addition, the viremia (as measured by qRT-PCR) was not significantly higher in children who subsequently developed DSS, although the plasma NS1 concentration was, suggesting that antigen burden may be a better correlate of severe clinical outcomes, as has been alluded to previously [60].

Using a large sample size, we defined the major transcriptional features of the acute response to DENV infection during the febrile period. Complement, TLR and RIG-I signalling pathways, interferon-stimulated genes and cytokine/chemokines and their receptors were the major features of the transcriptional signature, consistent with previous studies in smaller numbers of febrile dengue patients [142, 144, 176]. Utilising a systems biology approach that investigated transcriptional profiles in the context of their molecular interaction networks, we identified the transcription factors STAT1 and STAT2, the tyrosine kinase SRC, SHP1, TRIP6 and JAK2 as key central molecules in these networks. Furthermore, the top-ranked differentially expressed sub-network was enriched with molecules involved in cytokine signalling and JAK/STAT pathways (including JAK2, JAK3, SRC, TLR2, IL2RG, SOCS1, SHP1, TRIP6 and many other JAK/STAT and SRC regulators).

The top hub/bottleneck molecules in these networks form a densely connected network module, with a variety of known interactions between the nodes, SRC phosphorylates STAT1, for example, upon IFNαstimulation [178]. SHP1 also interacts with SRC and regulates its activity [179]. SHP1 has been shown to negatively regulate TLR-mediated production of pro-inflammatory cytokines by inhibiting activation of the transcription factor NF-kappaB and

MAP kinase while promoting TLR- and RIG-I-activated production of type I interferon by inhibiting the kinase IRAK1 [180]. JAK2 also interacts directly with STAT1 and SHP1. The interaction between JAK2 and phosphorylated STAT1 is enhanced upon IFNα stimulation [178] and results in STAT1 homodimer formation [181]. SHP1 (which is down-regulated) may inhibit JAK-STAT signalling by targeting SOCS1 (which is up-regulated) to JAK2 [182, 183]. TRIP6 interacts with TLR2 and is a signalling component of multiple NF-kappaB activation pathways [184] and is phosphorylated by SRC [185]. The interactions between the top hub/bottleneck molecules would support this being a key core signalling module in the network and a central feature in the host response. Analysis of the transcriptional regulatory network also identified the STAT and IRF transcription factors as the key regulators of the transcriptional response.

This provides ex vivo evidence in human infection of the prominence of the STAT mediated signalling pathway where it is likely critical for regulating the expression of most ISGs. The importance of the STAT1 pathway in control of DENV replication in mice and mosquitoes has been demonstrated previously [186, 187] and STAT1/2 may also be targets of DENV-mediated interference in the interferon signalling pathway [156, 157].

We identified 21 genes as differentially expressed (more abundant) in patients who developed DSS compared to matched control patients. Remarkably, almost all of these genes belonged to one of two overlapping networks, in which some of the interconnecting elements have immune response functions. That TNF-α should be a central hub in the 2$^{nd}$ network is striking, given that TNF-α has been repeatedly implicated in the pathogenesis of severe capillary leakage [152, 188]. Clearly, further studies are needed to understand why these two networks, and some of their differentially expressed members, are associated with progression to DSS.

Neutrophil activation and degranulation was a most prominent theme in the DSS-associated differentially expressed gene list. We verified that plasma proteins concentrations of CTSG, BPI, ELA2 and MPO were also higher in early DSS than in control patients, albeit the absolute difference was small and unlikely to be useful for prognosis. Of the differentially expressed genes associated with neutrophil degranulation, ELA2, CTSG and the defensins (DEFA1, DEFA4) are of particular interest. ELA2 and CTSG are serine proteases that can cleave vascular endothelial cadherin and thereby compromise the integrity of the vascular endothelium [189]. These proteases might conceivably play a similar role in perturbing the endothelium in capillary beds during DENV infection. Accordingly, ELA2 has also been detected previously at higher concentrations in serum of patients with DSS compared with patients without shock [190]. DEF1A and DEF4A are neutrophil associated defensins with anti-viral activity [191-193]. These innate anti-microbial peptides may also functionally participate in the innate anti-viral response to DENV infection.

The dengue capillary leakage syndrome begins in children with secondary infections within the first few 1-2 days of illness and can be measured by ultrasound as early as the 3$^{rd}$ day of illness [25]. Commencement of capillary leakage in infants with primary dengue likely occurs with similar kinetics and can lead to DSS between illness days 4-6 [78]. The triggering of capillary permeability early in the disease evolution in both these clinical settings might be mediated by robust innate immune responses rather than acquired responses, particularly since pre-existing immunity does not exist in infants with primary infection. We propose that activated, degranulating peripheral blood neutrophils could contribute to an early triggering of capillary permeability. In this model, neutrophils (which express Fc receptors) are activated by immune complexes and/or by high early virus antigen burdens in blood and tissues, where they secrete cytokines and chemotactic molecules. High viral antigen burdens in vivo could be

arrived at by antibody dependent enhancement in secondary infection or in primary infection of infants born to dengue immune mothers. Neutrophil activation and degranulation alone is highly unlikely to be sufficient to drive capillary leakage to the extent seen in patients with DSS. Instead, neutrophil adherence to endothelial cells and the secretion of soluble mediators of vascular permeability may represent a small step in the inflammatory cascade that synergises with other adaptive host responses to mediate capillary permeability in severe secondary infection. In infants with primary infection and severe dengue, an innate response that includes neutrophil activation, together with an inherently permeable vascular endothelium, might be sufficient to trigger clinically significant vascular leak. Further studies of neutrophils and their response to DENV infection are warranted.

# CHAPTER 5

# 5. HOST GENETIC SUSCEPTIBILITY TO DENGUE

## 5.1 Introduction

Epidemiological evidence hinted that genetic background could contribute to susceptibility to dengue. Studies in Cuba indicated that only patients of European origin presented with DHF/DSS [94, 102, 120-124]. Furthermore, despite the presence of risk factors of DHF/DSS, only DF but not DHF/DSS was known across the continent of Africa [125-128].

Polymorphisms on the HLA region have been examined in host genetic susceptibility studies of dengue [121, 122, 129-133, 194] but unfortunately, no reproducible conclusions has been made, primarily because of small sample sizes. Besides the HLA region, polymorphisms on other genes have also been studied such as DC-SIGN, TAP, HPA, Vitamin D receptor, FcγR [136-138].

Collectively, the majority of studies into host genetic susceptibility to dengue were case-control association studies of small sample size with uncertain case ascertainment and with unknown levels of population stratification (genetic admixture in one population that obscures any meaningful study of genetic association). Typically, these studies examined polymorphisms in a limited number of candidate genes for which there was some a priori rationale for their investigation. For the vast majority of case-control association studies, there was never an attempt to replicate the genetic association in either the same or different populations.

We conducted a genome-wide association case – control study to investigate host genetic susceptibility to dengue. The study was design to have two phases - an exploration and a replication phase with stringent genotyping quality control throughout the study.

## 5.2 Purposes of the study

1. To identify Single Nucleotide Polymorphisms (SNPs) which confer susceptibility or protection to symptomatic dengue. Frequencies of 250,000 SNPs in 366 dengue cases and 166 population controls were investigated using Affymetrix 250K genotyping technique. The most significant SNPs were selected for replication in a second sample set of 1260 cases and 1216 controls using Sequenom genotyping technique.
2. To conduct a GWAS to identify SNPs associated with dengue severity in 376 samples (187 severe dengue cases, 189 mild dengue cases) using Affymetrix 500K (NSP and STY) SNP genotyping technology. SNPs that were potentially associated with severe disease were genotyped in a $2^{nd}$, larger sample set containing 488 severe dengue cases and 542 mild dengue cases using the Illumina GoldenGate genotyping platform.

## 5.3 Materials and Methods

### 5.3.1 Study design

The study's design is depicted in figure 5.1

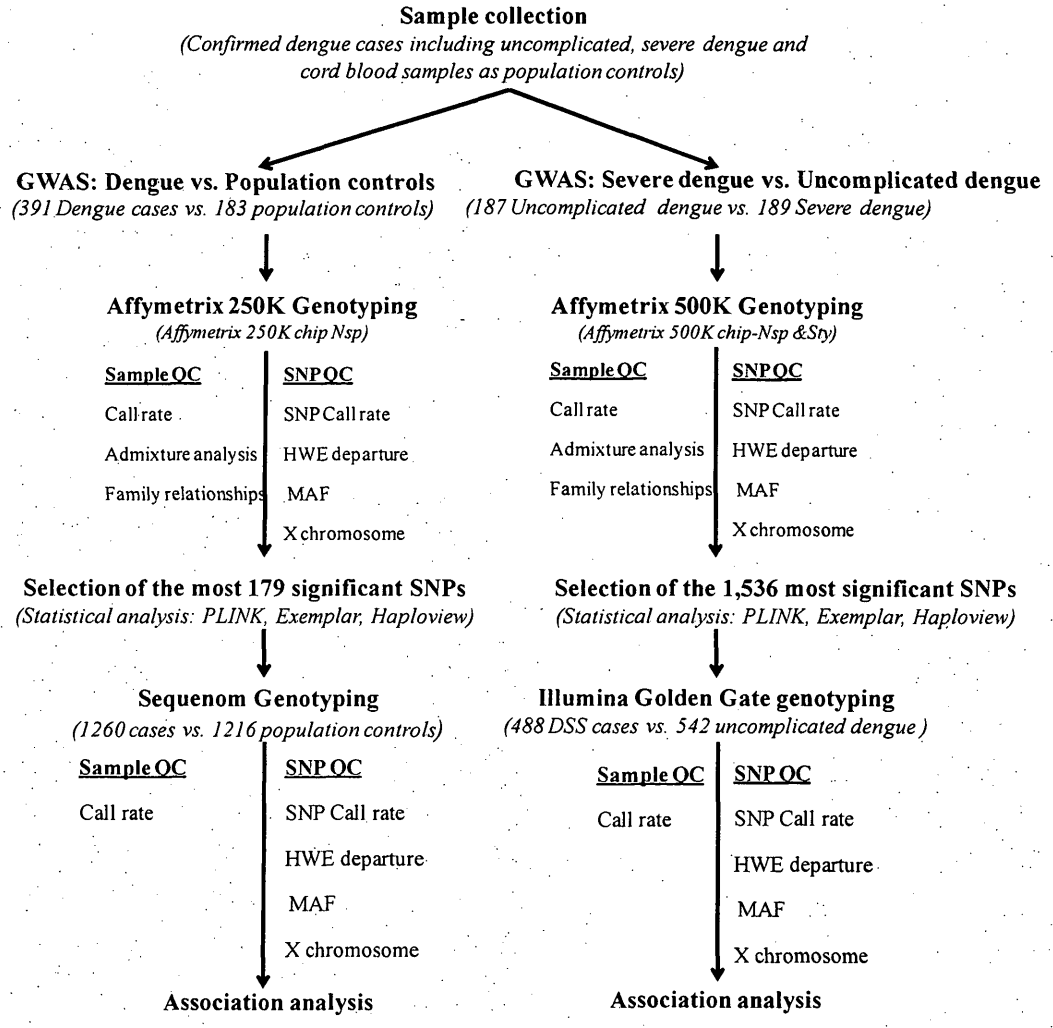



Figure 5.1: Diagram summarising study design. Two GWAS studies were performed to identify SNPs that are associated with susceptibility to dengue or SNPs that are associated with severe dengue. The GWAS were followed by replication studies with independent sample sets and techniques.

### 5.3.2 Samples collection

Blood samples from dengue patients who were enrolled in a prospective descriptive study of dengue at HTD between 1999 and 2007 were used for genetic association studies. Peripheral blood (2.0ml) was collected at the time of study enrolment and again at the time of discharge. Plasma samples were collected after centrifugation and used for diagnosis and research laboratory tests. After plasma was collected, the remaining pellet was kept in a 1.5ml tube and stored at -20$^0$C until used for DNA extraction. Blood pellet samples of controls were collected from cord blood of babies born in Hung Vuong Hospital, Ho Chi Minh city, Vietnam in 2003 and 2007 and stored in -20$^0$C. All protocols were approved by Hospital for Tropical Disease, Hung Vuong hospital, Ho Chi Minh City. Approval for the studies was obtained from the ethical committees of the Hospital for Tropical Diseases, Hung Vuong hospital and Oxford Tropical Research Ethnics committee, UK. All participants provided written informed consent to participate.

### 5.3.3 Affymetrix 500K genotyping

The genotyping process is summarised in figure 5.3. In brief, genomic DNA was digested by restriction enzyme (Nsp or Sty) before being ligated with adaptor, amplified by PCR, fragmented, labelled and hybridised into the beadchip.

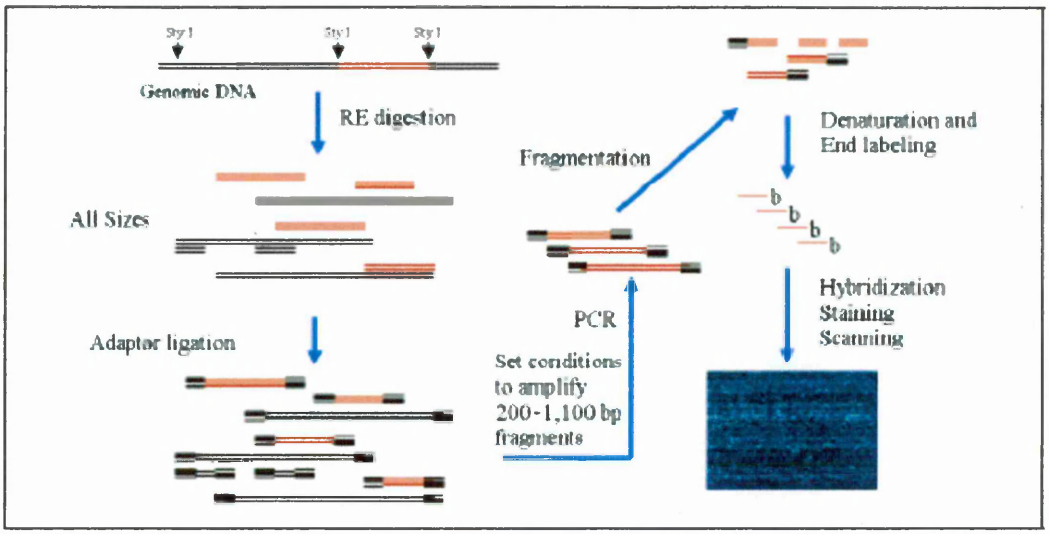



Figure 5.2: GeneChip Mapping Assay – StyI; Genomic DNA was digested by StyI, amplified by PCR, labeled and hybridized to the chip, then scanned to determine the SNP genotypes.

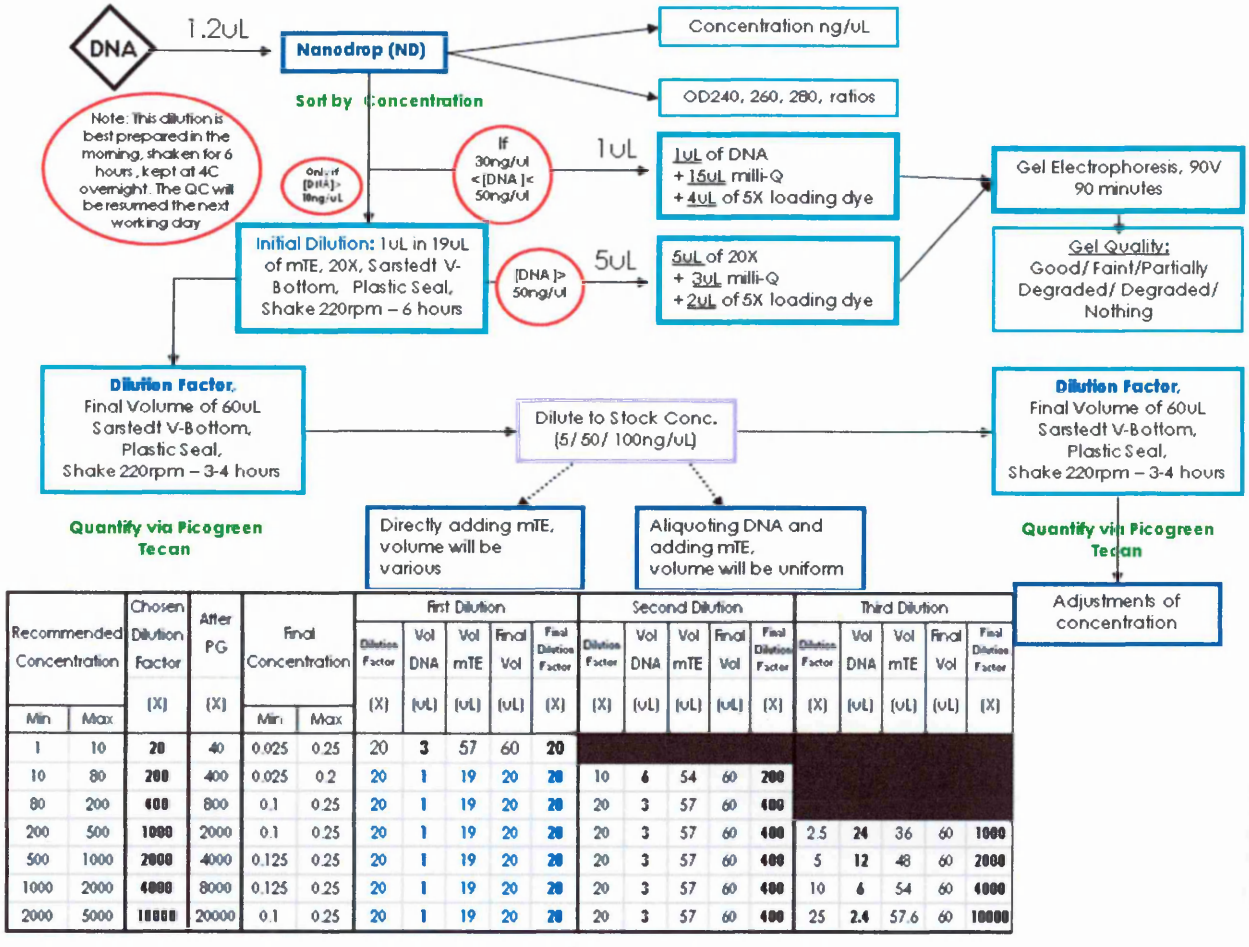
#### 5.3.3.1 DNA extraction for the GWAS

All DNA extractions were performed using a Nucleon kit (GE Healthcare, UK) according to the manufacturers' instructions. The blood pellet was re-suspended in red cell lysis buffer, reagent A (10mM Tris-HCl, 320mM surose, 5mM MgCl2, 1% Triton-X100, pH 8.0), with a ratio 1:4 blood to buffer. The nucleated cells were pelleted after a 10 minute by centrifugation at 1300g. Next, the cell pellet was re-suspended and lysed in 350µl of reagent B (40mM Tris-HCL, 60mM EDTA, 150mM NaCl, 1% SDS, pH 8.0), and incubated for 10 minutes at $70^0$C. The cell lysate was then mixed with 100µl sodium perchlorate, followed by 600µl of chloroform. The tube was mixed by hand and inverted at least seven times. Without remixing the phase, 150µl of nucleon resin was added to the tube before being centrifuged at 350g for one minute. Without disturbing the nucleon resin layer, the upper phase was transferred to a new tube. Next, two volumes of cold absolute ethanol was added and mixed by inversion until the precipitate appeared. The tube was centrifuged at top speed for five minutes to pellet the DNA, and the supernatant was discarded. The pellet was washed with 1ml of cold 70%

ethanol, re-centrifuged, and the supernatant discarded. The pellet was dried for 10 minutes before dissolving in 600 µl TE buffer.

#### 5.3.3.2 DNA Quality Control

1.2µl of genomic DNA was measured by spectrophotometry using a Nanodrop instrument (Thermal scientific, USA). Next, 1 µl of genomic DNA was diluted 20 times with TE buffer (1µl of DNA in 19µl of mTE buffer) in 96-wells plate (Sarstedt V-bottom). The plate was then shaken at 220 RPM overnight at room temperature to resuspend the DNA thoroughly. After shaking overnight, 5µl of diluted DNA was then used for gel electrophoresis to check the quality of the DNA. The DNA was first mixed with 5µl of 2X loading dye then loaded into a 1.2 percent agarose gel. The gel was run at 90 voltages for 90 minutes before a gel picture was taken using a gel imager. 20X diluted genomic DNA was measured using picogreen (Tecan). In brief, each DNA sample was further diluted based on the starting concentration of the DNA samples obtained through Nanodrop measurement. After a second dilution, the plate was then shaken for 4 hours at 220 RPM to dilute the DNA thoroughly. After that, 50µl of diluted DNA was mixed with 50µl picogreen (Invitrogen Quant-iT PicoGreen dsDNA Reagent (solution in DMSO), P7581) before the DNA concentration was measured using a fluorimeter (Tecan,place). Details of DNA QC process is illustrated in figure 5.3




| Recommended Concentration | | Chosen Dilution Factor | After PG | Final Concentration | | First Dilution | | | | | Second Dilution | | | | | Third Dilution | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Dilution Factor | Vol DNA | Vol mTE | Final Vol | Final Dilution Factor | Dilution Factor | Vol DNA | Vol mTE | Final Vol | Final Dilution Factor | Dilution Factor | Vol DNA | Vol mTE | Final Vol | Final Dilution Factor |
| Min | Max | (X) | (X) | Min | Max | (X) | (uL) | (uL) | (uL) | (X) | (X) | (uL) | (uL) | (uL) | (X) | (X) | (uL) | (uL) | (uL) | (X) |
| 1 | 10 | 20 | 40 | 0.025 | 0.25 | 20 | 3 | 57 | 60 | 20 | | | | | | | | | | |
| 10 | 80 | 200 | 400 | 0.025 | 0.2 | 20 | 1 | 19 | 20 | 20 | 10 | 6 | 54 | 60 | 200 | | | | | |
| 80 | 200 | 400 | 800 | 0.1 | 0.25 | 20 | 1 | 19 | 20 | 20 | 20 | 3 | 57 | 60 | 400 | | | | | |
| 200 | 500 | 1000 | 2000 | 0.1 | 0.25 | 20 | 1 | 19 | 20 | 20 | 20 | 3 | 57 | 60 | 400 | 2.5 | 24 | 36 | 60 | 1000 |
| 500 | 1000 | 2000 | 4000 | 0.125 | 0.25 | 20 | 1 | 19 | 20 | 20 | 20 | 3 | 57 | 60 | 400 | 5 | 12 | 48 | 60 | 2000 |
| 1000 | 2000 | 4000 | 8000 | 0.125 | 0.25 | 20 | 1 | 19 | 20 | 20 | 20 | 3 | 57 | 60 | 400 | 10 | 6 | 54 | 60 | 4000 |
| 2000 | 5000 | 10000 | 20000 | 0.1 | 0.25 | 20 | 1 | 19 | 20 | 20 | 20 | 3 | 57 | 60 | 400 | 25 | 2.4 | 57.6 | 60 | 10000 |

Figure 5.3: DNA quality control procedure. DNA quality was determined by Nanodrop, electrophoresis, and picogreen measurements.

### 5.3.3.3 Genomic DNA plate preparation

DNA samples were diluted with mTE buffer before being added into working plates. Two concentrations of genomic DNA were used for this experiment. For DNA samples that had concentrations ranging from 35ug/µl to 80ng/µl, 5µl DNA of each patient or control was added into one well of a 96-well plate. 3µl of samples with DNA concentrations of 80 to 120ng/µl was prepared in another 96 – well plate.

### 5.3.3.4 Restriction enzyme digestion

Restriction enzymes (Nsp I and separately, Sty I) was used to digest the genomic DNA. For one 96-well plate, the digestion master mix was prepared on ice as following: 1275.1µl AccuGen®Water, 220.8µl NE buffer 2 (10X), 22.1µl BSA (100X; 10mg/ml) and 110.4µl of NspI (or StyI) restriction enzyme. The genomic DNA plate was vortexed for 3 seconds, spun for one minute at 2000rpm then kept on a cooling chamber while the digestion master mix was prepared. Next, 14.75µl of the master mix was transferred into each well of the working plate. The plate was then sealed with adhesive film, vortexed at high speed for 3 seconds and spun down at 2000rpm for 30 seconds. Finally, the plate was loaded into a preheated thermal cycler and run with 500K digest program ($37^0$C for 120 minutes, $65^0$C for 20 minutes and hold at $4^0$C). When the program was finished, the plate was spun down at 2000rpm for 2 minutes.

#### 5.3.3.5 Ligation

During this step, the digested DNA samples were ligated using either the Nsp or Sty Adaptor. A ligation master mix was prepared then added into the samples. The plate was then loaded into the thermal cycler and the 500K ligation program was run. The details of these steps are described as below.

The digested DNA plate was put on a cooling chamber while the ligation master mix was prepared in a 2ml tube as follows: 82.8µl Adaptor Nsp I (or Sty Adaptor), 276µl T4 DNA Ligase Buffer (10X) and 220.8µl T4 DNA ligase (400U/µl ). The ligation master mix was then mixed at high speed for one second each for 3 times and spun down for 3 seconds. Next, 5.25µl of ligation master mix was transferred into each well of the digestion plate. The plate was then sealed with adhesive film, vortexed at high speed for 3 seconds and spun down at 2000rpm for 30 seconds before being loaded into the preheated thermal cycler and run with 500K ligation program as follows: 16°C for 180 minutes, $70^0$C for 20 minutes and hold at $4^0$C. After the ligation program was finished, the plate was taken out, spun down at 2000rpm for 30

seconds and put on ice before being diluted with 75µl of cold AccuGenWater. The plate was then sealed with adhesive film, vortexed a high speed for three seconds and spun down at 2000rpm for 30 second in order to keep the samples at the bottom of the wells.

#### 5.3.3.6 PCR for amplification of ligated DNA

During this step, an equal amount of ligated sample was transferred into three new 96-well plates before the PCR master mix being added into each plate and 500K PCR program was run. After the program was finished, the result would be checked by running 3µl of each PCR product on 2% TBE gel.

Firstly, 10µl of the samples in each row of the 96-well plate were transferred into the same rows of 3 new 96-well plates (labeled P1, P2 and P3) using 12-channel P20 pipette (row A of ligated plate in to row A of P1, P2 and P3, respectively). For the 3 new PCR plates, the PCR master mix was prepared into 50ml Falcon tube on ice as follow: 13.082mL AccuGENEwater, 3.312mL TITANIUM Taq PCR buffer (10X), 6.624mL GC-Melt (5M), 4.637mL dNTP (2.5mM each), 1.490mL PCR primer 002 (100mM) and 0.663mL TITANIUM Taq PCR Polymerase (50X). Next, the PCR master mix was vortexed thoughtfully in order to distribute the reagent evenly then poured in to the solution basin. 90µl of the PCR master mix was then transferred into PCR plate (P1, P2 and P3) using 12-channel P200 pipette before the plates were carefully sealed with adhesive films, vortexed at high speed for 3 seconds and spun down at 2000rpm for 30 seconds. Finally, the plates were loaded into the preheated thermal cyclers to run with 500K PCR program as follow: $94^0$C for 3 minutes; 30 cycles of: $94^0$C for 30 seconds, $60^0$C for 45 seconds and $68^0$C for 15 seconds; $68^0$C for 7 minutes and hold at $4^0$C.

### 5.3.3.7 Checking the PCR product

The result of previous steps was checked by running 3µl of PCR products in 2% TBE gels. 3µl of loading dye was added into each well of 3 new plates labeled P1Gel, P2Gel and P3GelBefore 3µl of each PCR product from P1, P2 and P3 was transferred to the corresponding plates P1Gel, P2Gel and P3Gel. The plates were then sealed with adhesive films, vortexed and spun down at 2000rpm for 30 seconds before 6µl of loading dye and sample mixture being loaded into 2% TBE gels. The gels were run at 120V for one hour. Finally, gel pictures were taken and the product was verified to make sure that it is distributed around 250bp to 1100bp.

### 5.3.3.8 PCR product purification and elution

Firstly, 8µl of diluted EDTA (0.1M) was added into the product before the plates was tightly sealed, vortexed and spun at 2000rpm for 30 seconds. The manifold was prepared by connecting the manifold and regulator to a suitable vacuum source that able to maintain at 600mbar. Next, the samples from the same row and well of each PCR product plate was transferred and pooled to the corresponding row and well of the Clean-Up Plate before being covered to prevent contaminations from outside. The vacuum was then turned on and slowly brought up to 600mbar and maintained until the clean-up plate became dry. After that, the PCR product was washed 3 times with AccuGENE water by adding 50µl of water into each well of the clean-up plate for 30 minutes for the first 2 times and 45 minutes for the third time. After the clean-up plate was completely dried, it was removed from the manifold and its bottom was immediately blotted on a thick stack of clean absorbent paper to remove any remaining liquid and dried with absorbent wipe.

#### 5.3.3.9 PCR product elution

45µl RB buffer was added into each well of the clean-up plate. The plate was then tightly sealed and shaken for 10 minutes at 1000rpm at room temperature. PCR product was collected by transferring of 45µl of the product from the clean-up plate to corresponding well of a fresh 96-well.

#### 5.3.3.10 Quantitation and normalization

In this step, three independent dilutions of each PCR product were measured. The average concentration of these 3 dilutions was calculated then normalized with RB buffer to acquire concentration of 2ug/µl. Firstly, 198µl AccuGENE water was added into each well of each optical plates labeled OP1, OP2 and OP3 followed by transferring 2µl DNA product from clean-up DNA plate into corresponding wells of the optical plates. Next, the optical plates were sealed with adhesive films, vortexed thoughtfully for 3 times at high speed for 5 seconds and spun at 2000rpm for 1 minute. This step was repeated 5 times in order to make sure the DNA was mixed evenly in each well. Next, 100µl of diluted DNA from each of the optical plates OP1, OP2 and OP3 were transferred to the new plates for OD measurement at the wavelength of 260, 280 and 320nm. The three OD readings for every sample were subtracted to water blank OD before the average OD of 3 replicates was calculated. Neat DNA concentration was calculated as following:

Sample concentration in µg/µl = average sample OD value x (0.05ug/µl) x 100.

*0.05 µg/µl is the amount of double stranded DNA equivalent to 1 unit of OD at 260nm.*

*The ratio OD260/OD320, which indicates the quality of the DNA sample, should be between 1.8 and 2.0. OD320 should be very close to zero.*

DNA concentration was normalized by mixing 90ug of DNA sample with appropriated amount of RB buffer in order to bring the concentration to 2ug/µl in total 45µl.

#### 5.3.3.11 Fragmentation

In this step, the purified, normalized DNA product was fragmented using fragmentation reagent. Firstly, 5µl of fragmentation buffer was added to each sample followed by 5 µl of the fragmentation reagent (0.05U/µl of DNase I). The plate with mixture of DNA and fragmentation reagent was then vortexed at high speed for 3 sec and spun down at $4^0$C at 2000rpm for 30sec. Finally, the place was then loaded onto preheat block ($37^0$C) thermal cycler and run with 500K fragmentation program ( $37^0$C for 35 minutes, $95^0$C for 15 minutes and hold at $40^0$C).

#### 5.3.3.12 Checking the fragmentation reaction

After the 500K fragmentation program finished, the plate was spun at 2000rpm for 30 sec to keep the fragmentation product at the bottom of the wells. 4µl of the fragmentation product was mixed with 4µl of gel loading dye before being loaded onto 4% TBE gel and run for 30 minutes at 120V. The gel was finally inspected under gel capture machine.

#### 5.3.3.13 Labelling

In this step, fragmented samples were labeled using GeneChip DNA labeling reagent. Firstly, the labelling master mix was prepared into a 15ml centrifuge tube with 1545.6µl of TdT buffer (5X), 220.8µl of GeneChip DNA labeling Reagent and 386.4µl of TdT enzyme (30U/ml). The mixture was then vortexed at high speed 3 times (1 second each time) and spun down for 3 seconds. Next, 19.5µl of the master mix was then transferred to each well of the fragmented plate placed on the other cooling chamber. The reagent and fragmented products was then pipetteted up and down for few times to ensure all the labeling mix is added to the fragmented

product. After tightly sealed, vortexed at high speed for 3 seconds, spun down and 2000rpm for 30 seconds, the plate was then loaded onto a preheated ($37^0$C) thermal cycler block and run with the 500K labeling program ($37^0$C for 4 hours, $95^0$C for 15 minutes and hold at $4^0$C). After the labeling program finished, the place was spun down at 2000rpm for 30 seconds before proceeded to the next step.

#### 5.3.3.14 Hybridization

Firstly, the hybridization master mix was prepared by adding the reagents into a 50ml centrifuge tube then thoroughly vortexed and pulse spun for 3 seconds. Next, 190µl of the hybridization master mix was then added into each of labeled samples (the total volume now is 260µl). The plate was the tightly sealed then the samples were denatured on a thermal cycler ($95^0$C for 10 minutes and held at $49^0$C). After the denaturation, each sample was loaded onto the GeneChip Human Mapping 250K array (Nsp and Sty). The arrays were then placed into a hybridization oven that preheated to $49^0$C for 16 to 18hrs.

#### 5.3.3.15 Washing, staining and scanning

The following hybridization, washing and staining procedures were performed by the Fluidics Station 450 (Affymetrix 2006). In brief, the staining protocol for mapping arrays was a three-stage process consisting of a Streptavidin Phycoerythin (SAPE) stain followed by an antibody amplification step and a final stain with SAPE. After the staining process, the array was then filled with Array Holding Buffer and was ready for scanning. After the wash protocols were completed, the probe array was scanned by the GeneChip Scanner 3000 7G and the scanned probe array image (.dat file) and cell intensity data (.CEL file) were automatically generated by GeneChip Operating Software and made ready for analysis by GeneChip Genotyping Analysis Software (GTYPE).

### 5.3.4 Illumina Golden Gate Assay

After DNA QC, 250ng DNA of 25ng/μl was used for this experiment. The experiment was performed according to the manufacture's instruction (Illumina, Inc. San Diego, CA 92121 USA). In brief, genomic DNA was first activated for binding to paramagnetic particles. The activated DNA was then subjected to a hybridization step in which assay oligonucleotides, hybridization buffer, and paramagnetic particles were combined with activated DNA. Assay oligonucleotides comprise three types of oligonucleotides: two allele specific oligos (ASOs) which hybridize to the alleles and a locus specific oligo (LSO) which hybridizes to several bases downstream of the allele site and contains a unique address sequence that target to a particular bead type. These three types of oligonucleotide sequence contain regions of genomic complementarily and universal primer sites. During hybridization, the ASOs and LSO hybridize to genomic DNA sample bound to paramagnetic particles. The hybridization occurs before any amplification step so as to prevent introduction of bias amplification into the assay. Following the hybridization, several washing steps are performed to reduce noise by removing excess and mis-hybridized oligonucleotides. The extension and ligation of ASOs to the LSO joins the information about the genotype present at the SNP site to the address sequence on the LSO. The joined, full-length products were then used as templates for PCR amplification with cy3-labeled P1 primer and cy5-labeled P2 primer from 3' and P3 primer from 5' site. After downstream processing, the single-stranded, dye-labeled DNAs are hybridized to their complement bead type through their unique address sequence. This hybridization of the products from the solution onto solid surface beadchips allows for individual SNP genotype readout. Finally, the BeadArray Reader is used to analyze fluorescent signal on the BeadChip and analyzed by software for automated genotyping clustering and calling. The diagram in figure 5.4 below summarises the whole procedure.

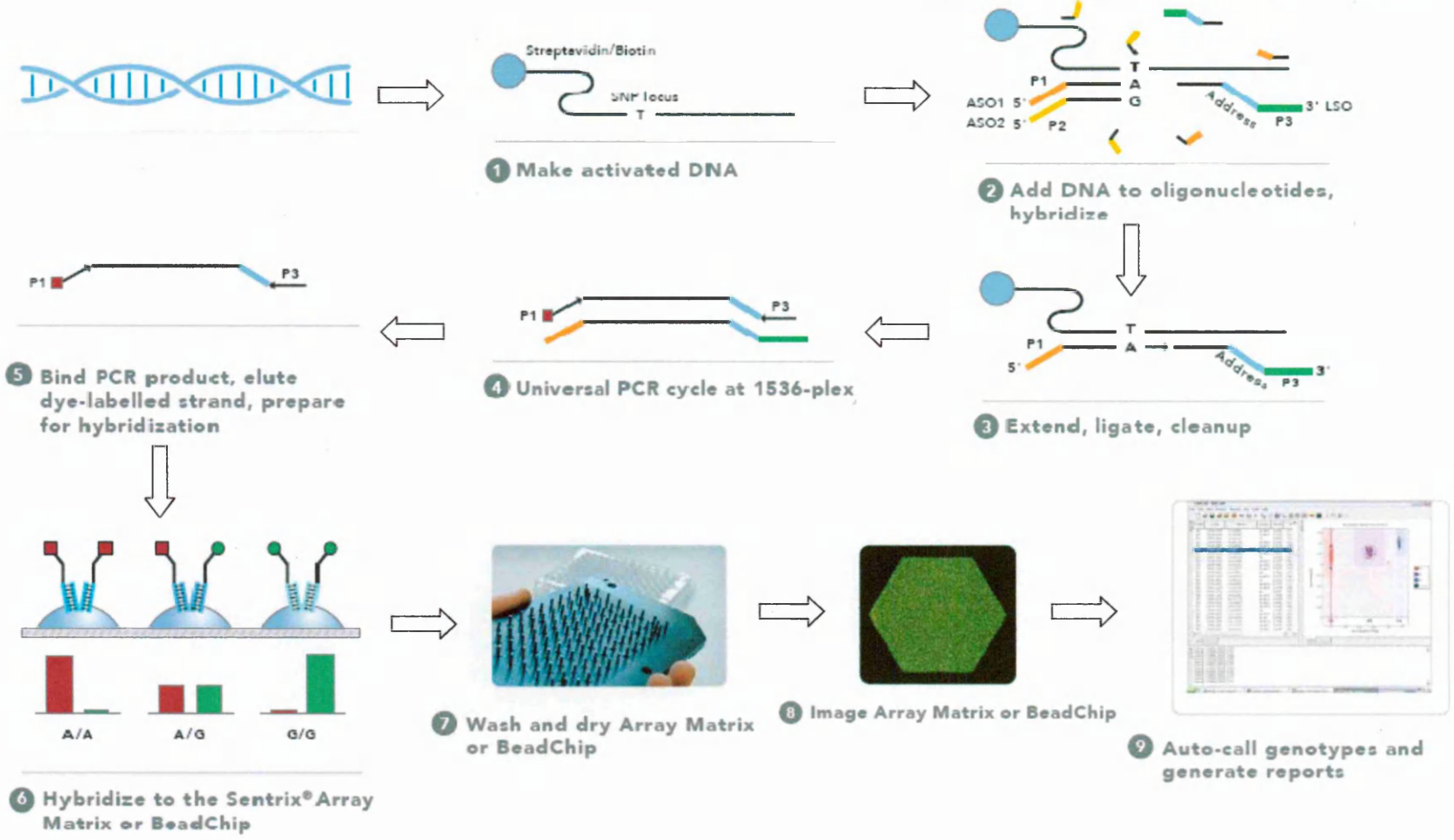



Figure 5.4: Illumina GoldenGate Genotyping flow chart (Illumina,). Genomic DNA was activated, hybridized with probes, amplified by PCR and hybridized to the BeadChip. After washing and drying steps, the BeadChip was scanned by confocal microscope to analyze fluorescent signal on the BeadChip which analyzed by software for automated genotyping clustering and calling.

### 5.3.5 SEQUENOME

This method is a universal method for detecting insertions, deletions, substitutions, and other polymorphisms in amplified DNA. Firstly, the genomic DNA sequence flanking the SNP sequence was amplified followed by the neutralisation of unincorporated dNTPs in amplification products using shrimp alkaline phosphatase (SAP). This enzyme cleaves a phosphate from the unincorporated dNTPs, converting them to dNDPs and rendering them unavailable to future reactions. Next, the iPLEX Gold reaction cocktail (primer, enzyme, buffer, mass-modified nucleotides) is added to the amplification products. The mixture of amplified products and iPLEX Gold reaction cocktail was then themocycled to process the iPLEX Gold reaction which involves the enzymatic addition of mass-modified nucleotides into the diagnostic site. In the reaction mixture, all four mass-modified nucleotides – A, T, C and G – are present. During the iPLEX Gold reaction, the primer is extended by one of the nucleotides, which terminates the extension of the primer. The DNA polymerase which incorporates nucleotide was used in the iPLEX Gold reaction. The reaction produces allelic-specific extension products of different masses depending on the sequence analyzed (figure 5.5).

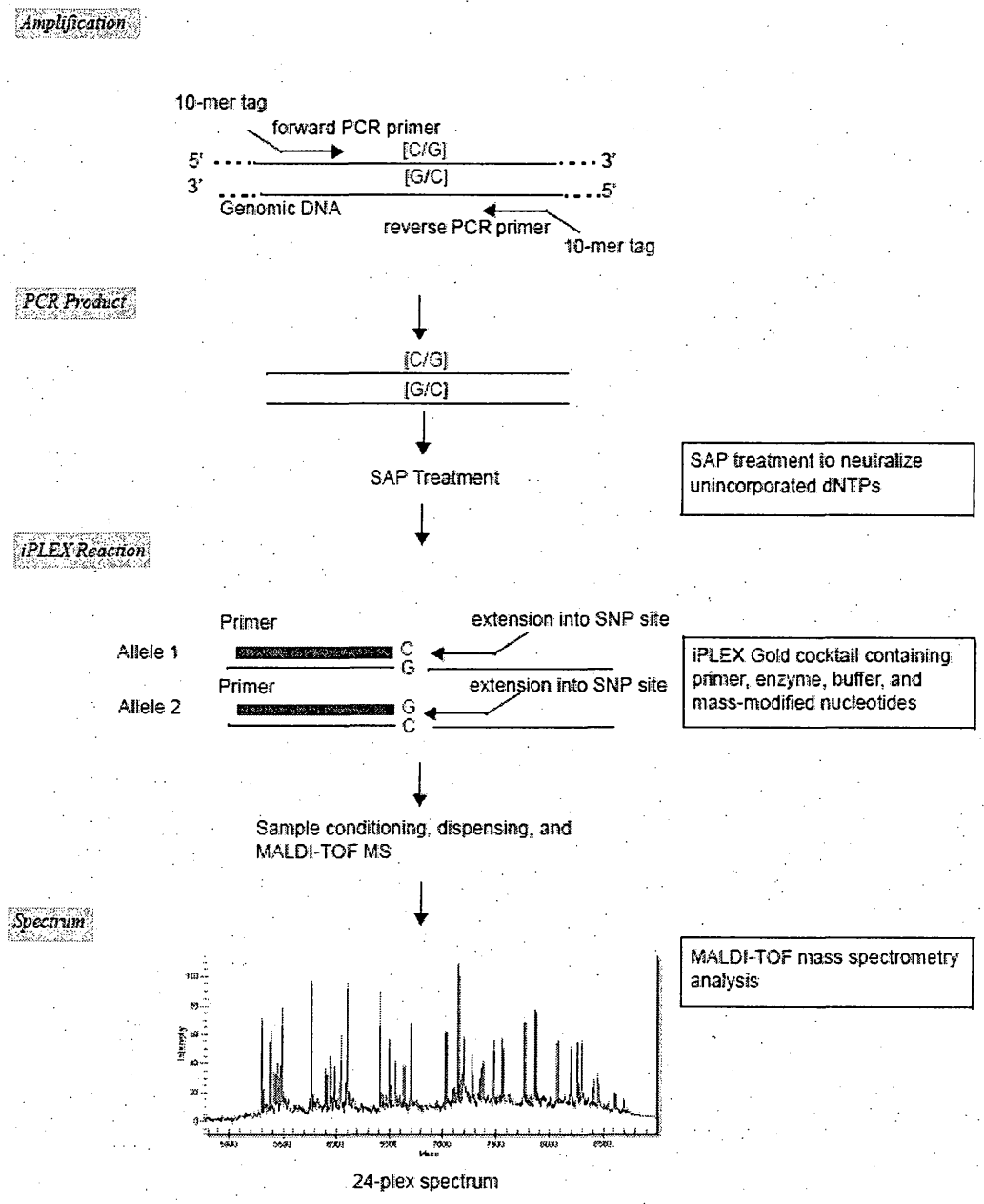



Figure 5.5: Flow chart of Sequenom genotyping platform.

### 5.3.6 Genome-wide association data analysis

#### 5.3.6.1 GeneChip Genotyping Analysis Software (GTYPE)

GCOS version 4.0 was used in a genome-wide scan to automate the control of the GeneChip Fluidics Stations and Scanners. In addition, GCOS also supports GTYPE generating genotype calls for the GeneChip Mapping Arrays. GTYPE enables importation of samples attributes, creates custom file sets, exports batch SNP or linkage data, creates

virtual data sets, assesses sample mismatch and updates SNP annotations from the NetAffx Analysis Center. This software generates .CHP files which display a Dynamic Model (DM) scatter plots containing genotype calls for each SNP with a confident score. A mapping algorithm report is automatically generated in .rpt files when intensity data is analyzed from mapping arrays. Report files summarise data from the samples with SNP call rate (Number of SNP calls/total number of SNPs on the array) and genotype call rate (Number of genotype calls for the sample/total number of SNPs on the array). Call rate is displayed in the DM algorithm and it is an indicator of the overall performance of the assay. Report files also provide quality control (QC) performance with Median PM (Median intensity value for perfect match cells for the QC probe sets), MCR (Median call rate) and common SNP patterns.

#### 5.3.6.2 Data pre-processing

The raw data output from the Affymetrix genotyping consists of measures of probe hybridization intensities. For each individual at each SNP, there are either 6 or 10 probe quartets. Each probe quartets consists of four probe cells which assay for a perfect match or a mismatch to a specific 25-base oligonucleotide sequence for each of the two possible alleles (genetically denoted A and B). The hybridization intensities need to undergo a pre-processing phase to combine the information across the probe quartets to yield a pair of coordinates corresponding to the signal strength for each of the two possible alleles. The .cel files are pre-processed to a format the enables viewing the quality of the genotyping data.

#### 5.3.6.3 QC for genotyping data

Highly accurate and reliable genotype calling is an essential step in a high-throughput SNP genotyping technique. In this study, we used both Dynamic Model (DM) and Bayesian Robust Linear Model with Mahalanobis distance classifier (BRLMM) calling algorithms to

determine genotypes of each SNP. TheDM has been used for the GeneChip Human Mapping 500K. However, the BRLMM was generated to produce a new model which provides a significant improvement over DM. Moreover, we also generated cluster plots (SNPSignal tool, affymetrix) which allowed us to visualise the quality of each SNP. Genotyping data was first assessed by DM algorithm. Samples with DM call rate of less than 90% were subjected to repeat genotyping. This allows us to assess the genotyping quickly and decide samples that need to be re-genotyped. Samples with DM call rate of over 90% were assessed again using the BRLMM algorithm because this algorithm provides improvement in call rate over DM algorithm.

#### 5.3.6.4 BRLMM Analysis tool (BAT) version 2.0

BRLMM (Bayesian Robust Linear Model with Mahalanobis distance classifier) is evolved from the Robust Linear Model with Mahalanobis distance classifier (RLMM, http://www.stat.berkely.edu/users/nrabbee/RLMM/) which provided an improvement over DM algorithm on the Mapping. An extension of the RLMM model, developed for mapping the 500K gene chip, provides an improvement over DM in two important areas – it improves overall performance (call rate an accuracy), and it equalizes the performance on homozygous and heterozygous genotypes. The difference between RLMM and BRLMM is the addition of a Bayesian step which provides improved estimates of cluster centers and variances.

#### 5.3.6.5 Principle component analysis (PCA)

Population stratification – allele frequency differences between cases and controls due to systematic ancestry differences – can cause spurious associations in GWAS studies which normally consist of a lot of samples and SNPs. Population stratification is the problem for large scale association study which includes thousands of samples. PCA analysis allow us to assess how homogenous or heterogeneous are the population under studying (cases and

controls belong to the same population) and how different this population compare to other populations such as CHB, JPT, CEU or YRI.

#### 5.3.6.6 RELPAIR

RELPAIR version 2.0.1 is a FORTRAN 77 program that infers the relationships of pairs of individuals based on genetic marker data, either within families or across entire samples . This is a likelihood-based method for inferring the most likely relationship for putative sib pairs. It calculates the multipoint likelihood of the marker data for each pair conditional on each of four possible relationships: full-sibs, monozygotic twin, half sibs, and unrelated pairs. RELPAIR was written to evaluate the multipoint likelihoods and to assess the most likely relationship between different putative sib pairs. The multipoint likelihood depends on population marker-allele frequencies, inter-marker distances and the presumed relationship of the pair. We have used this program to test all possible pairs of individuals in our sample. It can indentify apparently independent samples that are in fact related and it also indentifies related individuals erroneously classified as unrelated because of sample switches or duplications.

#### 5.3.6.7 Cluster plots

We used SNP Signal tool (Affymetrix) to check the quality of each SNP separately. The SNP Signal Tool generates cluster drawings of SNP Signals, one drawing per SNP over a set of experiments. Each drawing consists of a set of points in "Contrast" vs. "Size" space. The points are colored according to the BRLMM-p genotype calls. Experiments with different genotypes for the SNP should cluster from left to right in contrast space: BB (red), AB (green) and AA (blue). Visual confidence can be assigned to a SNP whose experiments are resolved in Contrast space (the X axis), and have few No Calls (grey points).

#### 5.3.6.8 Quartile – quartile plots (Q-Q plots)

Q-Q plots are constructed by ranking a set of values of a statistic from smallest to largest and plotting them against the expected values (. Under the null hypothesis there should be no significant difference in allele frequencies at genetic markers between cases and controls. If the observed data are consistent with the null hypothesis, the points should fall approximately along the reference line. The greater the departure from the $45^0$ reference line, the greater the evidence that the observed $X^2$ statistic do not follow the expected $X^2$ distribution, suggesting that the measured differences between cases and controls are not random.

#### 5.3.6.9 Hardy-Weinberg Equilibrium (HWE)

Deviation from the HWE can arise in the presence of a common depletion polymorphism, because of a mutant PCR-primer site or because of a tendency to miscall heterozygotes as homozygotes. For a large-scale genotyping study, we assessed the HWE in the control group as a data quality check. We used Pearson goodness-of-fit test to test for the deviations from the HWE. This test was known simply as "the $X^2$ test" because the test statistic has approximately a $X^2$ null distribution. We used PLINK to check for the depature form HWE in the controls

#### 5.3.6.10 Association tests

To identify SNPs that are significantly associated with dengue or with severe dengue in GWAS, we applied five models of chi-square tests which are used to examine the form of relationship between two discrete variables and to determine if the relationship observed was significantly different from what would be expected if there were no relationship between variables.

1. Armitage's trend test was used as a genotype-based test to test the hypothesis of zero slope for the line that fits the three genotyping risk estimate best

2. Allelic test is a 1-df chi-square test which was used to compare allele frequencies in cases and controls
3. Genotypic test is a 2-df chi-square which was used to compare three genotypes (the two homozygote and the heterozygote)
4. Dominant test is a 1-df chi-square test which was used to compare wild type genotypes (the wild type homozygote and the heterozygote) and the mutant genotype (the mutant homozygote)
5. Recessive test

#### 5.3.6.11 Helixtree

We used Helixtree to analyze genotyping data from sequenom. This software allow us to perform quality control of the SNP and the samples such as call rate, HWE, MAF. We can also perform association test using this software.

#### 5.3.6.12 WGAviewer

This software was used to annotate SNPs that found significant form association analysis. It allows us to look at detail annotation of the SNPs such as LD, haplotype, genes that the SNPs located.

#### 5.3.6.13 Haploview

We used haploview to investigate the LD patterns between SNPs and also to plot association data that exported from PLINK.

#### 5.3.6.14 PLINK

PLINK (http://pngu.mgh.harvard.edu) is a free, open-source tool set for whole genome association analysis. It is designed to perform analysis ranging from basic to large-scale analyses in a computationally efficient manner. PLINK can be used to analyze both genotype and phenotype data so that the raw data needed to be processed. We used PLINK to perform both data quality control and association analyses.

## 5.4 Results

### 5.4.1 Quality control (QC)

#### 5.4.1.1 DNA QC

Prior to genotyping, DNA samples were stringently checked for quality and quantity by different methods. The success of high-throughput SNP genotyping using the 500K Chip (250K Nsp and 250K sty) relies on genomic DNA of good integrity and high quality which is free of contaminants that would affect the downstream enzymatic reactions. Figure 5.6 shows an example of DNA samples used in the experiment visualised by gel electrophoresis.

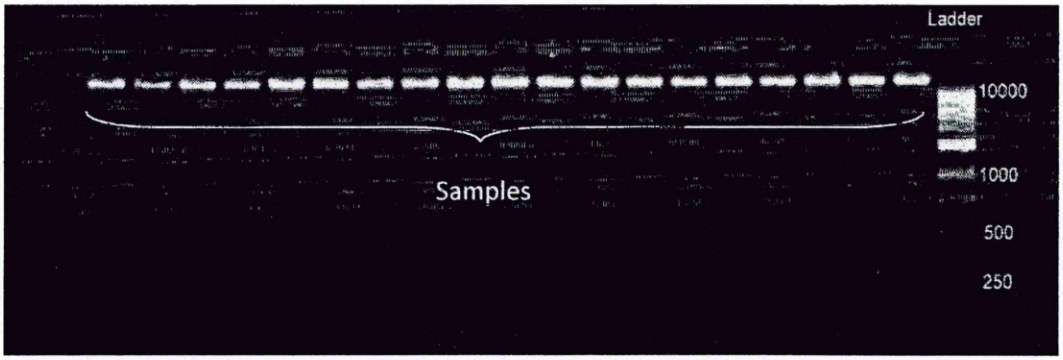



Figure 5.6: Gel electrophoresis on a 1% agarose gel. Lane ladder refers to a Gene Ruler 1Kb DNA ladder (Fermentas, USA) which is composed of 14 DNA fragments measuring up to 10kb. Genomic DNA samples that are >10Kb in size.

#### 5.4.1.2 Quality control for Affymetrix genotyping

##### 5.4.1.2.1 Assessment of PCR

Genomic DNA was digested with Nspl or Sty 1 restriction enzymes, ligated to adaptors and PCR amplified using PCR primers. PCR products were electrophoresed on agarose gels to check their size. A successful PCR reaction was indicated by DNA fragments with

the size of 250 to 1100bp when checked by electrophoresis. Figure 5.7 showed an example of a successful digestion and PCR amplification to yield amplimers of size 250 to 1100bp.

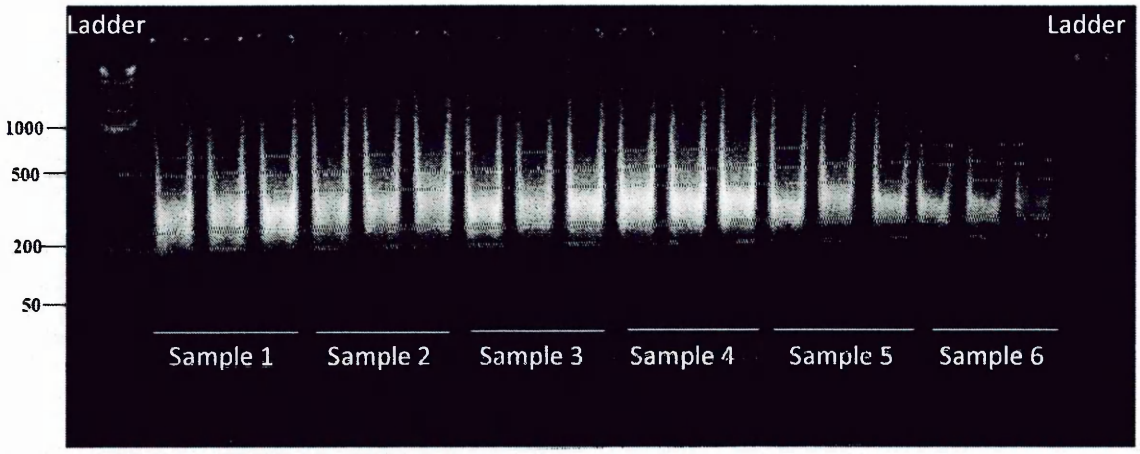



Figure 5.7: PCR products electrophoresed on a 2% TBE agarose gel at 120V for 1hrs. Average product distribution is between ~250 to 1100bp. The first and the last lane are 50bp DirectLoad Wide Range Marker (Sigma, USA) and the remaining lanes are PCR products after amplification

#### 5.4.1.2.2 Assessment of fragmentation

After being purified and normalised to 2µg/µl, 90µg of PCR product was fragmented by using DNase I. The result of the fragmentation was checked by agarose gel electrophoresis. Successful fragmentation was indicated by an excess of DNA less than 180bp. Figure 5.8 shows an example of a successful DNase I fragmentation.

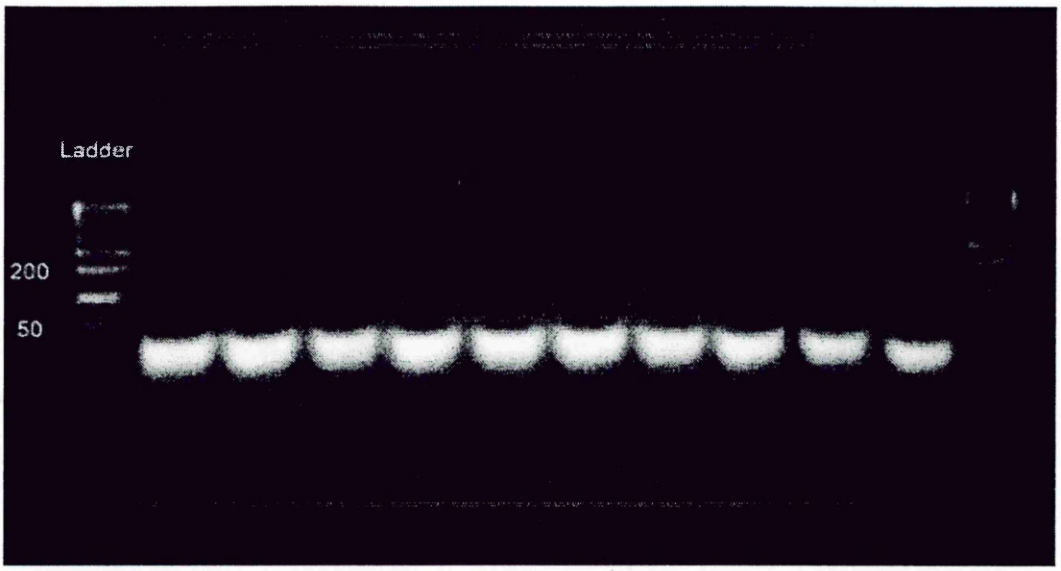



Figure 5.8: PCR products digested with DNase I were electrophoresed on a 4% TBE agarose gel at 120V for 30 minutes. Average size is <180bp. The first and the last lanes are 50bp DirectLoad Wide Range DNA Marker (Sigma, USA) and the remaining lanes are DNA samples after fragmentation.

### 5.4.2 GWAS: Dengue versus population controls

#### 5.4.2.1 Sample collection

For the genome-wide association case control study, we collected a set of samples for use as population controls. These were the cord blood samples from babies born to Vietnamese Kinh women in Hung Vuong Hospital, Ho Chi Minh City and were collected between 2003 and 2007. For Affymetrix 250K genotyping, we used DNA from 183 cord blood samples and 391 DNA samples from dengue cases. For validation study, we used 1,260 DNA samples from dengue cases and 1,216 cord blood control DNA samples.

#### 5.4.2.2 SNP QC

The results of SNP QC is summarised in table 5.1. Briefly, we removed from the data set 11178 SNPs that had call rates of less than 93%, 49472 SNPs that had minor allele

frequencies of less than one percent (monomorphic SNPs), 5710 SNPs on the X chromosome, 710 SNPs for departure from HWE ($P < 10^{-7}$ in the controls), and finally, 6924 SNPs were removed because there was more than 5% differences in SNP call rates between cases and controls. After these QC procedures, 194327 SNPs were left for further analyses.

Table 5.1: SNP QC for GWAS comparing dengue and population controls

| **SNP Filter** | **Number of SNPs** |
|---|---|
| Total number of SNPs | **262264** |
| SNP callrate < 93% | 11178 |
| MAF < 0.01 | 49472 |
| Chromosome X SNPs | 5710 |
| Out of HWE (*P* > 1e-7) in control | 701 |
| SNP call rate difference ≥5% between cases and controls | 6924 |
| **SNPs left for analysis** | **194326** |

### 5.4.2.3 Principle component analysis

The principle component analysis was applied for 526 samples with 194,326 SNPs that passed the QC. Genotyping data of SNPs in all the cases and controls were compared with those of SNPs in Caucasian (CEU), African (YRI) and Asian populations (HCB and JPT). All of the Vietnamese case and control subjects clustered with the Han Chinese (HCB) and Japanese (JPT) (figure 5.9A). The Caucasian and the African SPNs from the HapMap database were then excluded leaving only SNPs from the Asian populations (HCB and JPT) were used. The PCA plot of this analysis using only SNPs from Asian population showed that the subjects in the study are homogeneous and do not belong to other Asian

population (figure 5.9B). The PCA result with PC1 and PC2 indicated that the case and the control samples are homogeneous except four subjects with PC2 >0.19 (Figure 5.9C). The result after excluding the 4 subjects was shown in figure 5.9D.

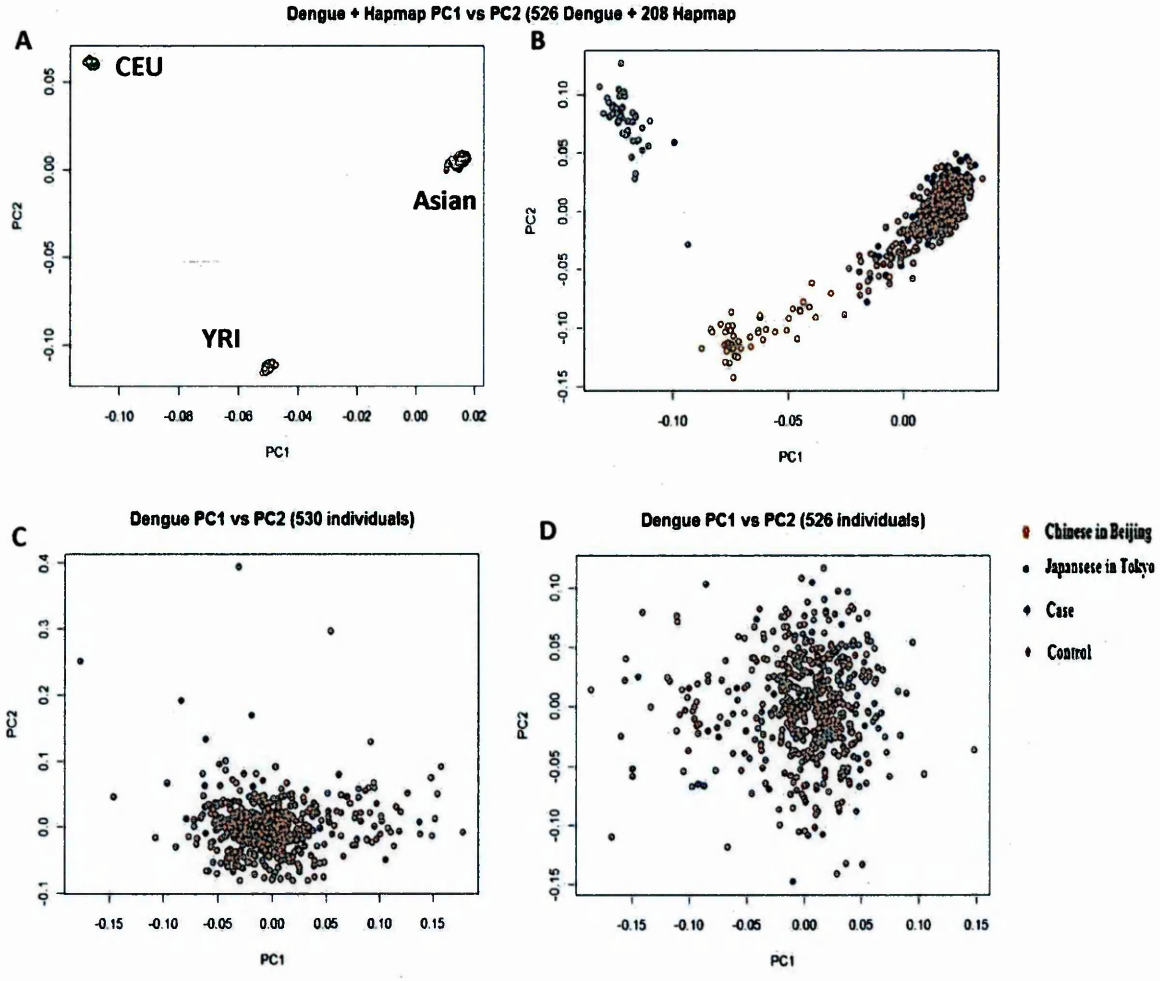



Figure 5.9: Principle components analysis indicates that Vietnamese population belongs to Asian (A) but different to Chinese (HCB) and Japanese (JPT) (B). Analyses of cases and controls showed that they are homogenous. Shown in (C) is the analysis of cases and controls before removing the outliers; and in (D) is the result after removing the outliers. The X axis shows the principle component 1; the Y axis shows the principle component 2

#### 5.4.2.4 QC of DNA samples

DNA samples from 361/391 cases and 170/183 controls were sufficient for genotyping. The details of sample quality control are summarised in table 5.2. We removed 3 samples which had BRLMM call rate less than 93%, 2 samples for being MZ or FS and 4 samples that were outliers as Vietnamese by PCA analysis leaving 522 samples (357 cases and 165 controls) for analysis. The baseline characteristics of the 357 cases are summarised in table 5.3

Table 5.2: SNP QC results

| **Filtering criteria** | **Number of samples** |
|---|---|
| Total number of samples | **531** |
| Cases | 361 |
| Controls | 170 |
| Sample callrate < 93% | 3 |
| MZ<br>FS | 1<br>1 |
| Outlier by PCA | 4 |
| Total samples left | **522** |

Table 5.3: Baseline characteristics of 357 cases used in GWAS comparing dengue cases and population controls

| Variants | DSS<br>N = 179 | Uncomplicated dengue<br>N = 178 |
|---|---|---|
| Sex (% male) | 50.8 % | 60.6 % |
| Age Median (range) years | 10.8 (7 – 14 ) | 11.2 (7 - 15) |
| Max hematocrit (%)<br>Mean (range) | 50<br>39 - 59 | 45.6<br>34.7 - 60 |
| Platelet nadir ($10^3$ cell/ml)<br>Mean (range) | 26242<br>5400 - 110000 | 63707<br>8510 - 290000 |

Note: max hematocrit is the maximum percentage of hematocrit between day 3 and day 8 of illness; platelet nadir is the minimum platelet count between day 3 and day 8.

### 5.4.2.5 QQ plot

QQ plot is a graphical method that can be utilized to explore the population uniformity. The expected P values of the test statistic from the controls were plotted against the observed P values in the cases. Under the null hypothesis, we would expect that all the data will fix to the 45-degree line. The QQ plot in figure 5.10 shows that majority of the data fix with the expected line which means that the population under study is homogenous. The red dots that are deviated from the line indicate either that the theoretical distribution was incorrect, or that the sample was contaminated with values generated by a true association

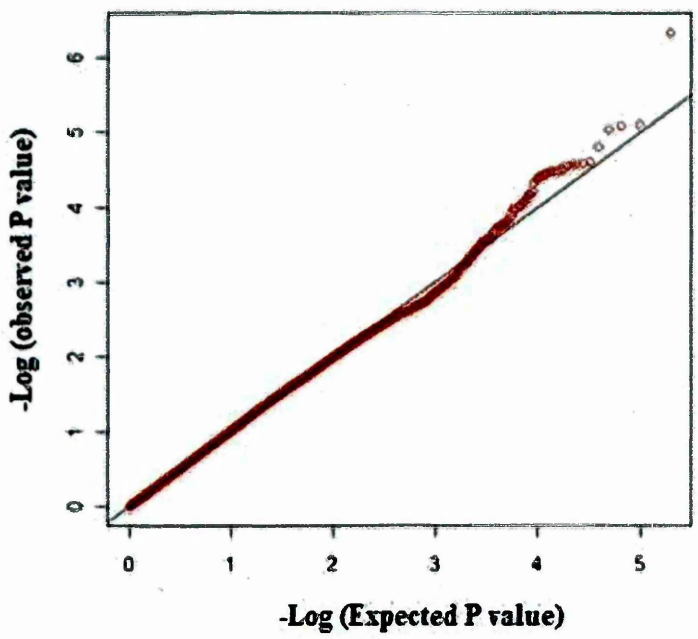



Figure 5.10: Quartile-quartile plot of trend test of the observed P values from cases (Y axis) and expected P values in controls (X axis) of 194326 SNPs. Majority of the dots fit with the 45 degree line indicating that subjects in cases and controls are homogenous. Only the SNPs in the extreme tail are those potentially associated with disease.

#### 5.4.2.6 Association test

We applied Armitage's trend test included in PLINK software to test for association. The tests were carried out to analyse 194326 SNPs in 357 cases and 165 controls. The results were then visualised by using Haploview software. The P values of the trend test from comparison of frequency of all autosomal markers in cases and controls were plotted against the chromosomal SNP position (figure 5.11).

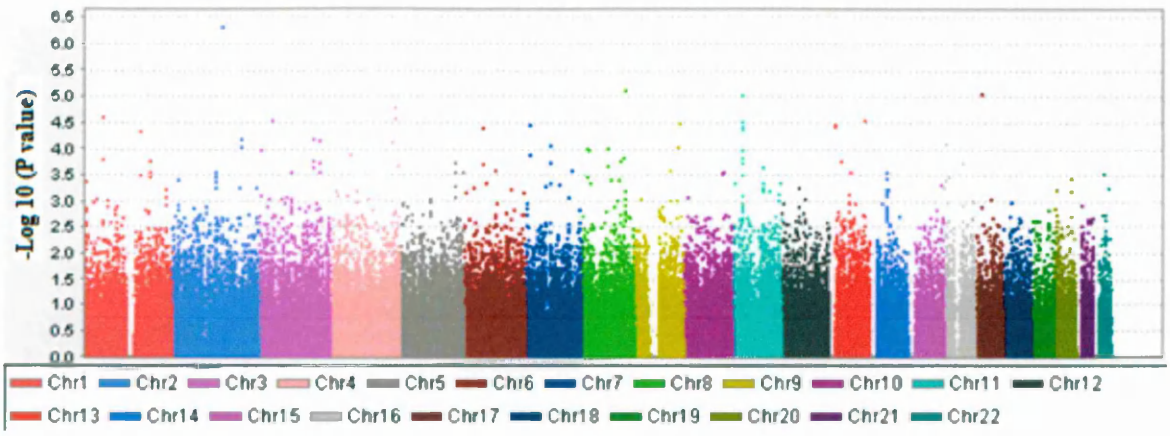



Figure 5.11: Trend test P value of 194326 SNPs plotted against chromosomes. Frequencies of SNPs in 357 cases were compared with 165 controls by Armitage's Trend test using PLINK. The test P values of all SNPs were plotted against the physical position of 22 autosomal chromosomes by Haploview software. The Y axis shows the -log10 scale of the P values. The X axis shows the chromosome clustered into different colours. The most significant SNP in this analysis was lower than $10^{-6}$ and mapped in chromosome 2.

### 5.4.2.7 Cluster plot of SNPs

To avoid the spurious results that may be derived from genotyping errors, we visually checked clusters of potentially associated SNPs. The BRLMM algorithm calls were used for drawing the SNP clusters using The SNP Signal Tool (Affymetrix). This software generates cluster drawings of SNP signals, one drawing per SNP over a set of experiments. It is impossible to check all the SNPs so we only checked the clusters of SNPs in the top hits SNPs (P value $\leq 10^{-3}$). From the top hits, we chose the top 179 SNPs with good genotyping quality for replication in an independent cohort. Figure 5.12 shows the clusters of SNP_A-2205258 which has good quality.

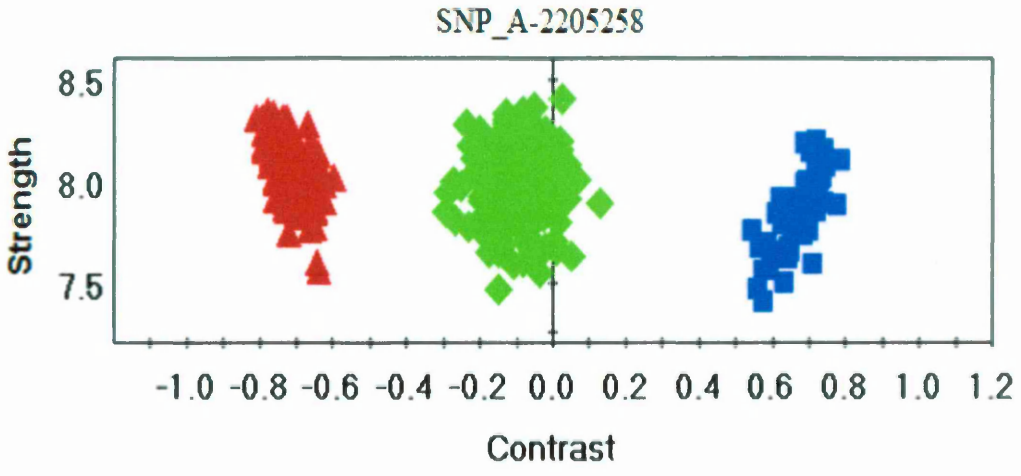



Figure 5.12: Cluster plots of SNP SNP_A-2205258 called by BRLMM algorithm. Each drawing consists of a set of points in "Contrast" vs. "Size" space. The points are colored according to the BRLMM genotype calls. Experiments with different genotypes for the SNP should cluster from left to right in contrast space: BB (red), AB (green) and AA (blue). Visual confidence can be assigned to a SNP whose experiments are resolved in Contrast space (the X axis), and have few No Calls (grey points).

### 5.4.3 Follow-up replication study in an independent cohort

Sequenom genotyping is a genotyping method that allows us to customise the SNPs to be genotyped, i.e. the SNPs of interest indentified by other methods can be genotyped with this method. It is a widely used method for the validation of GWAS data. We used Sequenom to validate the SNPs that were associated with susceptibility to dengue in the initial GWAS screening. As the purpose of this study was to identify SNPs associated with susceptibility or resistance to dengue, all dengue cases including severe and uncomplicated dengue patients were grouped into a "case" group while cord blood samples were used as population controls. Beside the top 179 SNPs that were chosen from GWAS screening, we decided to genotype an additional 28 SNPs that are mapped to DC-SIGN gene because some of these SNPs were demonstrated to be associated with susceptibility to dengue in

Thai children [137]. In total, we genotyped 207 SNPs in 2476 samples (1260 dengue cases and 1216 controls) using the Sequenom platform. The baseline characteristics of dengue cases used for this study is summarised in table 5.4. Of the 207 SNPs chosen for replication, we failed to genotype 26 SNPs because there was not primer available for the SNPs so that the genotyping reactions could not be designed. Data QC and association analyses were performed for 181 successfully genotyped SNPs using Helixtree software.

Table 5.4: Baseline characteristics of patients used in the validation study with the sequenom genotyping approach.

| **Variables** | **DSS** (N = 602) | **Uncomplicated dengue** (N = 658 ) | **Note** |
|---|---|---|---|
| **Age Median (range)** | 10 (1 – 15) | 12 (5 - 15) | |
| **Sex Male (%)** | 302 (50%) | 385 (59%) | |
| **Max Hct** | Missing N = 137 | Missing N = 54 | |
| **Median (range)** | 50 (37 - 60) | 44 (27.5 – 61) | |
| **Platelet nadir** | Missing N = 137 | Missing N = 54 | |
| **Median (range)** | 30800 (6390 – 174000) | 64250 (8510 – 379000) | |
| **Dengue confirmed cases** | 602 (100%) | 658 (100%) | |
| **By serology** | 590 (98.2%) | 637 (96.8%) | |
| **By clinical symptoms** | 12 (1.8%) | 3 (0.5%) | |
| **By PCR** | 0(%) | 18 (2.7%) | |

Table notes: The platelet nadir was the minimum platelet count from day 3 to day 8 of illness. Maximum Hct was the highest Hct between day 3 and day 8 of illness.

### 5.4.3.1 SNPs associated with susceptibility to dengue

We excluded 230 cases for being genotyped in GWAS leaving 1030 cases for further analysis. We compared frequencies of 181 SNPs (discovered in GWAS and DC-SIGN SNPs) in all dengue cases (n=1030, severe and mild dengue) with control samples (n=1216). Of the 181 SNPs, we removed two SNPs that had HWE P value in controls less than 2.8E-04 (Bonferoni multiple testing correction P value: 0.05/181 SNPs) and nine SNPs with call rate less than 0.93, leaving 170 SNPs for genetic association analysis. As the SNPs were chosen from the most significant SNPs from the trend test in GWAS, we also used the trend test to analyse them in this replication study. We identified 11 SNPs with trend test P value < 0.05 (without multiple testing corrections). We applied Fisher's test to combine P values from GWAS and from replication study and GWAviewer software to annotate the SNPs. The result of Fisher's combined values and annotation of the SNPs is summarised in table 5.5. The most significant SNPs of the combined data from GWAS (357 cases and 165 controls) with Sequenome (1030 cases and 1216 controls) was rs1366478 ($P = 6.3x10^{-5}$) which maps to intergenic region on chromosome 5 and rs17061643 ($P = 8.7x10^{-5}$) which maps to intronic region of the *KLF12* (Kruppel-like factor 12) gene on chromosome 13. The protein encoded by this gene is a member of the Kruppel-like zinc finger protein family. Other SNPs that also showed evidence for genome-wide association but with lower significance included rs10204649 ($P = 2.9x10^{-4}$) and rs298917 ($P = 6x10^{-4}$) which respectively maps to the intron and exon of *VWA3B* gene (von Willebrand factor A domain containing 3B) on chromosome 2. Other SNPs mapping to intergenic regions with less significance were rs7611871 ($P = 1.8x10^{-4}$) mapping to the *DPPA2* gene on chromosome 3, rs1863943 ($P = 1x10^{-4}$) on chromosome 5, rs12661760 ($P = 2.7x10^{-4}$) mapping to the *KIAA1244* gene on chromosome 6, rs1485306 ($P = 3.2x10^{-4}$) mapping to the *COLEC10* on chromosome 8, rs1885261 ($P = 3x10^{-4}$) mapping to the NPAS3 gene on chromosome 14 and rs9788917 ($P = 4x10^{-4}$) on chromosome 16. Functions of these genes in infectious diseases,

particularly in dengue are unknown and further investigations are needed in order to confirm the association of this gene with dengue.

Table 5.5: Summary of SNPs significantly associated with dengue

| SNP | Chr | Type | Gene | bp to gene | A1 | A2 | MAF | WGAS 357 cases vs. 165 controls | Replication 1030 cases vs. 1216 controls | Fisher's Combined P |
|---|---|---|---|---|---|---|---|---|---|---|
| **rs10204649** | 2 | Intronic | VWA3B | 0 | G | C | 0.05 | 1.25E-03 | 2.01E-02 | 2.90E-04 |
| **rs298917** | 2 | Synonymous | VWA3B | 0 | C | T | 0.049 | 1.63E-03 | 3.42E-02 | 6.02E-04 |
| **rs7611871** | 3 | Intronic | DPPA2 | 0 | G | A | 0.116 | 1.40E-03 | 1.09E-02 | 1.84E-04 |
| **rs1366478** | 5 | Intergenic | Unknown | -9 | A | G | 0.22 | 2.80E-04 | 1.70E-02 | 6.30E-05 |
| **rs1863943** | 5 | Intergenic | AC012607.8 | -25642 | T | C | 0.212 | 1.08E-03 | 7.87E-03 | 1.08E-04 |
| **rs12661760** | 6 | Intergenic | KIAA1244 | -23415 | A | G | 0.201 | 1.49E-03 | 1.56E-02 | 2.70E-04 |
| **rs1485306** | 8 | Intergenic | COLEC10 | -69185 | G | C | 0.366 | 6.16E-04 | 4.61E-02 | 3.26E-04 |
| **rs17061643** | 13 | Intronic | KLF12 | 0 | G | A | 0.06 | 1.12E-03 | 6.06E-03 | 8.75E-05 |
| **rs1885261** | 14 | Intronic | NPAS3 | 0 | G | C | 0.109 | 1.11E-03 | 2.64E-02 | 3.34E-04 |
| **rs9788917** | 16 | Intronic | AC092365.3-1 | -28509 | A | G | 0.235 | 1.06E-03 | 3.47E-02 | 4.12E-04 |

Note: chr: Chromosome; A1: allele 1, A2: allele 2; MAF: minor allele frequency.

### 5.4.3.2 The association of DC-SIGN SNPs with dengue

Previously, the SNPs rs4804803, rs8105483 and rs11465391 were demonstrated to be associated with protection for DF in Thai children [137]. We wanted to know if these DC-SIGN SNPs were also associated with disease status in Vietnamese Kinh with dengue. To answer this question, we investigated 28 SNPs in DC-SIGN in 1260 cases (uncomplicated and severe dengue) and 1216 controls. We successfully genotyped 16/28 SNPs in the total of 2476 samples. 12 SNPs failed to genotyped because we could not design the genotyping reaction. All of the 16 SNPs that were successfully typed were also typed in the previous study of Thai children [137]. To compare our results with previous published results, we applied a similar data analysis strategy. The frequency of these 16 SNPs in 1260 dengue cases and 1216 cord blood controls were also compared to identify SNPs associated with susceptibility to dengue. Statistical tests including Armitage's Trend test, basic allelic test, basic genotypic test, dominant test and recessive test were performed using Helixtree. The statistical results are summarised in Table 5.6. We identified four significant SNPs ($P < 0.05$, before multiple testing corrections) from comparison between dengue cases and controls. These SNPs which map to 3'-UTR region are rs4804802 (recessive $P = 0.028$), rs1544767 (recessive $P = 0.042$), rs1544766 (recessive $P = 0.025$) and rs11465413 (recessive $P = 0.028$). All of the significant SNPs were derived from recessive test with very marginal P values. This was not obtained from the study in Thai populations.

#### 5.4.3.3 DC-SIGN SNPs and disease status: Severe dengue and uncomplicated dengue comparison

The frequencies of 16 SNPs in 602 patients with DSS in 658 patients with uncomplicated dengue (DF, DHF grade I and II) were analysed using Helixtree software. We applied Armitage's Trend test, basic allelic test, basic genotypic test, dominant test and recessive test to analyse the data. Statistical results in comparison with previous study from Thai populations are summarised in table 5.7. From the tests performed, we did not identify any significant SNP (without multiple testing correction $P \leq 0.05$).

Table 5.6: Summary of association analysis of DC-SIGN SNPs between 1257 dengue cases (severe and mild dengue) and 1216 controls

| rs-id | Polymorphism | P value* | Marker statistics | | | Association test | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | MAF | Call Rate | HWE P | Armitage P | Allelic P | Genotypic P | Dominant P | Recessive P |
| **rs735240** | DCSIGN1-939 | | 0.196 | 0.982 | 0.530 | 0.405 | 0.402 | 0.332 | 0.232 | 0.610 |
| **rs735239** | DCSIGN1-871 | | 0.139 | 0.954 | 0.037 | 0.440 | 0.431 | 0.485 | 0.312 | 0.726 |
| **rs4804803** | DCSIGN1-336 | 0.52 | 0.080 | 0.973 | 0.327 | 0.152 | 0.156 | 0.356 | 0.165 | 0.552 |
| **rs2287886** | DCSIGN1-139 | 0.15 | 0.258 | 0.980 | 0.828 | 0.152 | 0.151 | 0.346 | 0.223 | 0.260 |
| **rs7252229** | DCSIGN1.in2+24 | | 0.073 | 0.977 | 0.956 | 0.397 | 0.397 | 0.351 | 0.548 | 0.160 |
| **rs11465384** | DCSIGN1.in4-28 | | 0.0002 | 0.942 | 0.992 | 0.325 | 0.325 | 0.325 | 0.325 | ? |
| **rs8105483** | DCSIGN1.in5-178 | 0.28 | 0.073 | 0.970 | 0.256 | 0.693 | 0.697 | 0.242 | 0.907 | 0.093 |
| **rs11465391** | DCSIGN1.in6-37 | 0.25 | 0.059 | 0.974 | 0.643 | 0.993 | 0.993 | 0.908 | 0.937 | 0.678 |
| **rs4804802** | DCSIGN1.1539 | | 0.379 | 0.974 | 0.043 | 0.423 | 0.432 | 0.046 | 0.667 | 0.028 |
| **rs10403018** | DCSIGN1.1599 | | 0.005 | 0.981 | 0.822 | 0.180 | 0.181 | 0.180 | 0.180 | ? |
| **rs4894801** | DCSIGN1.1667 | | 0.202 | 0.971 | 0.190 | 0.914 | 0.915 | 0.639 | 0.684 | 0.470 |
| **rs7248637** | DCSIGN1.2122 | | 0.375 | 0.983 | 0.162 | 0.547 | 0.553 | 0.079 | 0.608 | 0.054 |
| **rs1544767** | DCSIGN1.2281 | 0.15 | 0.374 | 0.983 | 0.264 | 0.466 | 0.471 | 0.069 | 0.681 | 0.042 |
| **rs1544766** | DCSIGN1.2426 | | 0.374 | 0.983 | 0.198 | 0.383 | 0.390 | 0.043 | 0.718 | 0.025 |
| **rs11465409** | DCSIGN1.3062 | | 0.072 | 0.979 | 0.084 | 0.710 | 0.715 | 0.517 | 0.836 | 0.251 |
| **rs11465413** | DCSIGN1.3197 | 0.5 | 0.102 | 0.979 | 0.072 | 0.383 | 0.392 | 0.089 | 0.646 | 0.028 |
| **rs4804800** | DCSIGN1.4020 | | 0.387 | 0.964 | 0.412 | 0.780 | 0.778 | 0.304 | 0.618 | 0.227 |
| **rs11260027** | DCSIGN1.43041 | | 0.220 | 0.976 | 0.993 | 0.621 | 0.621 | 0.385 | 0.992 | 0.184 |

*: P value determined in previous study in Thai patients

Table 5.7: Summary of association analysis of DC-SIGN SNPs between 600 severe dengue cases and 657 mild dengue controls.

| rs-id | Polymorphism | P value* | Marker statistics | | | Association test | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | MAF | Call Rate | HWE P | Armitage P | Allelic P | Genotypic P | Dominant P | Recessive P |
| **rs735240** | DCSIGN1-939 | | 0.191 | 0.976 | 0.185 | 0.470 | 0.462 | 0.316 | 0.822 | 0.133 |
| **rs4804803** | DCSIGN1-336 | 1.4E-07 | 0.075 | 0.962 | 0.461 | 0.905 | 0.906 | 0.915 | 0.851 | 0.732 |
| **rs2287886** | DCSIGN1-139 | 1.0E+00 | 0.249 | 0.969 | 0.954 | 0.644 | 0.644 | 0.545 | 0.958 | 0.286 |
| **rs7252229** | DCSIGN1.in2+24 | | 0.070 | 0.965 | 0.403 | 0.575 | 0.579 | 0.852 | 0.571 | 0.925 |
| **rs8105483** | DCSIGN1.in5-178 | 3.0E-05 | 0.071 | 0.957 | 0.072 | 0.523 | 0.534 | 0.812 | 0.519 | 0.962 |
| **rs11465391** | DCSIGN1.in6-37 | 2.0E-02 | 0.059 | 0.971 | 0.538 | 0.768 | 0.770 | 0.877 | 0.819 | 0.627 |
| **rs4804802** | DCSIGN1.1539 | | 0.384 | 0.962 | 0.852 | 0.920 | 0.919 | 0.593 | 0.683 | 0.452 |
| **rs10403018** | DCSIGN1.1599 | | 0.003 | 0.976 | 0.909 | 0.394 | 0.395 | 0.394 | 0.394 | ? |
| **rs4804801** | DCSIGN1.1667 | | 0.202 | 0.970 | 0.770 | 0.496 | 0.497 | 0.599 | 0.373 | 0.823 |
| **rs7248637** | DCSIGN1.2122 | | 0.379 | 0.975 | 0 606 | 0.875 | 0.874 | 0.785 | 0.849 | 0.570 |
| **rs1544767** | DCSIGN1.2281 | 1.4E-01 | 0.379 | 0.975 | 0.466 | 0.804 | 0.802 | 0.836 | 0.964 | 0.585 |
| **rs1544766** | DCSIGN1.2426 | | 0.380 | 0.980 | 0.480 | 0.715 | 0.712 | 0.725 | 0.979 | 0.455 |
| **rs11465409** | DCSIGN1.3062 | | 0.071 | 0.967 | 0.071 | 0.792 | 0.797 | 0.966 | 0.794 | 0.946 |
| **rs11465413** | DCSIGN1.3197 | 1.2E-01 | 0.098 | 0.965 | 0.013 | 0.443 | 0.460 | 0.492 | 0.527 | 0.282 |
| **rs11465421** | DCSIGN1.3852 | 6.4E-01 | -- | -- | -- | -- | -- | -- | -- | -- |
| **rs4804800** | DCSIGN1.4020 | | 0.389 | 0.952 | 0.099 | 0.277 | 0.266 | 0.447 | 0.531 | 0.210 |
| **rs11260027** | DCSIGN1.43041 | | 0.217 | 0.968 | 0.355 | 0.566 | 0.571 | 0.784 | 0.685 | 0.513 |

*: P value reported in previous paper; --: SNPs that were unable to genotype or filtered out by QC; MAF: Minor allele frequency; HWE: Hardy-Weinberg Equilibrium

### 5.4.4 GWAS: Severe dengue versus uncomplicated dengue

#### 5.4.4.1 Sample collection and QC

We selected 376 samples (187 DSS and 189 uncomplicated dengue controls) to genotype with the Affymetrix 500K platform. All the patients were children (less than 15 year olds) and Vietnamese Kinh ethnicity. All patients were diagnosed with DENV infection by serology, by PCR (see Chapter 2, section 2.2.1) or by clinical syndromes. All severe and uncomplicated dengue patients were classified based on newly updated WHO guidelines (see Chapter 2, section 2.1.2). The base-line characteristics of these patients are summarised in table 5.8. From the total of 376 samples, we removed 3 samples with BRLMM call rate of less than 93%. By RELPAIR (see section 5.3.6.6), we identified relationships between samples and removed another 26 samples. In the end, 29 samples were excluded, leaving 347 samples for further analysis (173 cases and 174 controls).

**Table 5.8: Base line characteristics of patients used in GWAS (affymetrix 500K genotyping)**

| Variables | DSS (N = 187) | Mild dengue (N = 189) |
|---|---|---|
| Age Mean (range) | 10.8 (7 – 14) | 11.2 (7 – 18) |
| Sex Male (%) | 50.8 (%) | 59.7 (%) |
| **Dengue confirmed cases (%)** | 100 % | 99.45 % |
| By serology (%) | 90.3 (%) | 94.7 (%) |
| By PCR (%) | 0 | 4.2 (%) |
| By clinical symptoms (%) | 9.7 (%) | 0.55 (%) |
| Unknown (%) | 0 | 0.55 (%) |

#### 5.4.4.2 Association analysis

To identify SNPs that were associated with severe dengue, we used Exampler to perform statistical tests comparing DSS cases with uncomplicated dengue controls. These tests including:

a. Allelic chi-square test (details in section 5.3.5.9)
b. Genotypic chi-square test (details in section 5.3.5.9)
c. Armitage's trend test (details in section 5.3.5.9)

From each test, the top 2000 most significant SNPs were chosen and pooled together. The SNPs were ranked by lowest P values and the 2000 SNPs with the lowest P values (the most significant) were selected. After choosing the SNPs, we performed SNP QC by looking at their call rate, HWE P value and MAF. All of the 2000 SNPs had a call rate over 93% by BRLMM and MAF >1%. We removed SNPs which deviated from HWE (HWE P $<1x10^{-7}$) in the controls and SNPs on the X chromosome. Finally, we chose 1536 SNPs for replication in an independent sample set using Illumina Golden Gate genotyping platform.

### 5.4.5 Nested analysis of severe dengue versus uncomplicated dengue cases

#### 5.4.5.1 Sample and SNP QC

The Illumina Golden Gate platform was used to genotype a set of 1536 SNPs in the 2$^{nd}$ sample set. From the total of 1,144 samples, DNA from 1107 samples was sufficient and qualified for genotyping by the new platform. Of the 1107 samples, we excluded 15 samples that could not be genotyped, 34 samples with sex discrepancies between genotyping results and clinical database, and 28 samples for being MZ or FS (most of which were duplication due to samples were collected twice) leaving 1030 samples (542 uncomplicated dengue controls and 488 DSS cases) for further analysis. The details of sample QC results are summarised in table 5.9. The base-line characteristics of these

samples are summarised in table 5.10. From the total of 1,536 SNPs discovered in GWAS, 81 SNPs were removed for having MAF ≤0.01, call rate ≤93%, and for deviation from Hardy-Weinberg equilibrium (P < 7e-4) in controls leaving 1455 SNPs for further analysis.

Table 5.9: QC results of samples used in replication study

| Filtering criteria | Sample filtered |
|---|---|
| Total samples | 1107 |
| Genotyping failed | 15 |
| Sex discrepancies | 34 |
| MZ or FS | 28 |
| # samples passed QC | 1030 |

Table 5.10: Base-line characteristics of 488 severe cases and 542 uncomplicated dengue controls used for replication with OPA golden gate technique.

| Variants | DSS (N = 488) | Mild dengue (N = 542) |
|---|---|---|
| Age Mean (range) | 9.4 (1 – 15) | 11 (5 – 15) |
| Sex Male (%) | 50.4% | 57.4% |
| **Dengue confirmed cases (%)** | 100% | 100% |
| By serology (%) | 99.4% | 98.9% |
| By PCR (%) | 0% | 0.7% |
| By clinical symptoms (%) | 0.6% | 0.4% |

### 5.4.5.2 SNPs associated with severe dengue

We genotyped and compared frequencies of 1455 SNPs in 542 uncomplicated dengue and 488 severe dengue patients. Statistical tests including trend test, genotypic test and allelic tests were performed using PLINK software. Figures 5.13 to 5.15 show Haploview plots of trend test P values of the 1455 SNPs against the chromosomal physical position. For consistency, P values of SNPs from the replication study were derived from the same statistical test performed in the GWAS. We identified 30 SNPs with P values < 0.05 (before multiple testing correction) in the replication study. By using Fisher's test, we combined P values in the GWAS and in the replication study. The list of these 30 SNPs and their annotation results are shown in table 5.11. The most significant SNP was rs12595534 ($P = 6.97x10^{-6}$, OR (95%CI) = 0.61 (0.5 – 0.76)) which maps to the intergenic region of the *KLF13* (Kruffel-like factor 13) gene on chromosome 15. Rs13294631 and rs16928238 were two significant SNPs that mapped to the intron of the *PTPRD* gene (protein tyrosine phosphatase, receptor type, D) on chromosome 9 with significance levels at $P = 2.3x10^{-5}$ and $P = 1.2x10^{-4}$ respectively. The distance between these two SNPs is 84kb apart indicating that they were independently identified. SNPs that map to the 3'-UTR and the intergenic region of the *MBIP* (*MAP3K12* binding inhibitory protein 1) on chromosome 14 were rs4900833 ($P = 3.5x10^{-5}$; OR (95CI) = 1.92 (1.35 -2.74)) and rs3844013 ($P = 2x10^{-4}$; OR (95CI) = 1.81 (1.25 – 2.62)), respectively. Rs4796751 ($P = 5.8x10^{-5}$, OR (95%CI) = 0.63 (0.49 – 0.84)) which maps to 3'-UTR region of the CNP (2',3'-cyclic nucleotide 3') gene on chromosome 17 was also identified.

### 5.4.5.3 IPA Gene network analysis

We used IPA to explore relationships between the genes in which significant SNPs were identified as associated with severe dengue. This analysis allows us to investigate the interactions of genes in which the SNPs themselves were found (n=14), or the closest

genes up to 25kb up or downstream of the candidate SNPs (n=7) and biological pathways based on its knowledge databases. Of the 21 genes, only 16 genes were mapped to the IPA knowledge databases which in turn, formed two networks. The first network contains 8/16 genes and was associated with cellular movement, cell signalling, nervous system development and function with the network score of 22 (equivalent to $P = 10^{-22)}$. This indicates that the network was highly unlikely formed by chance. The key molecules of the network were TNF (Tumour Necrosis Factor), *HPS90* (Heat Shock Protein 90kDa) and *HTT* (Huntingin). Genes in the network that contain or near by the associated SNPs are *PTPRD, CNP, OPCML, NKIRAS2, MBIP, PLA2G, DNAJC7* and *DNAH9*. The network is shown in figure 5.16. Collectively, the results suggest that there may be biologically plausible functional relationships exist between genes that contain or near by the association SNPs.

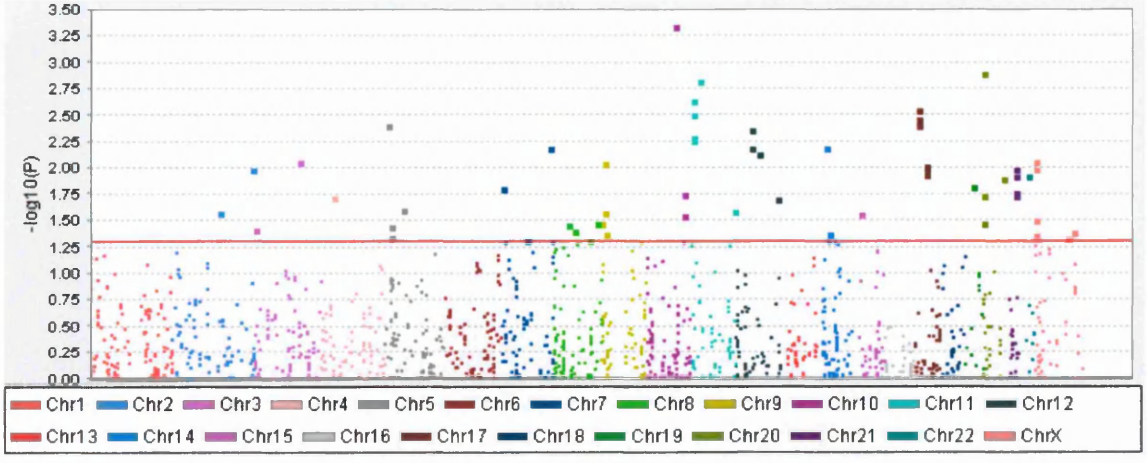



Figure 5.13: Trend test P values of 1455 SNPs plotted against physical chromosome position. The P values were derived from comparison between 488 severe dengue cases and 544 uncomplicated dengue controls. The Y axis is the logarithm scale of P value, the X axis is the chromosomal position indicated by different colours. Red lines is the cut-off at P = 0.05.

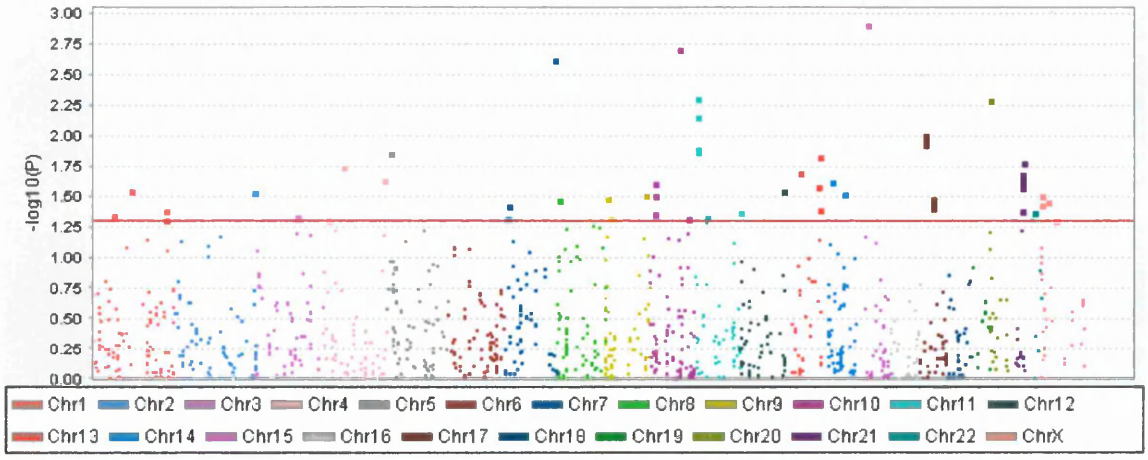



Figure 5.14: Genotypic test P values of 1455 SNPs plotted against physical chromosome positions by Haploview software. The P values were derived from comparison between 488 severe dengue cases and 544 uncomplicated dengue controls. The Y axis is the logarithm scale of P value, the X axis is the chromosomal position indicated by different colours. Red lines is the cut-off at P = 0.05.

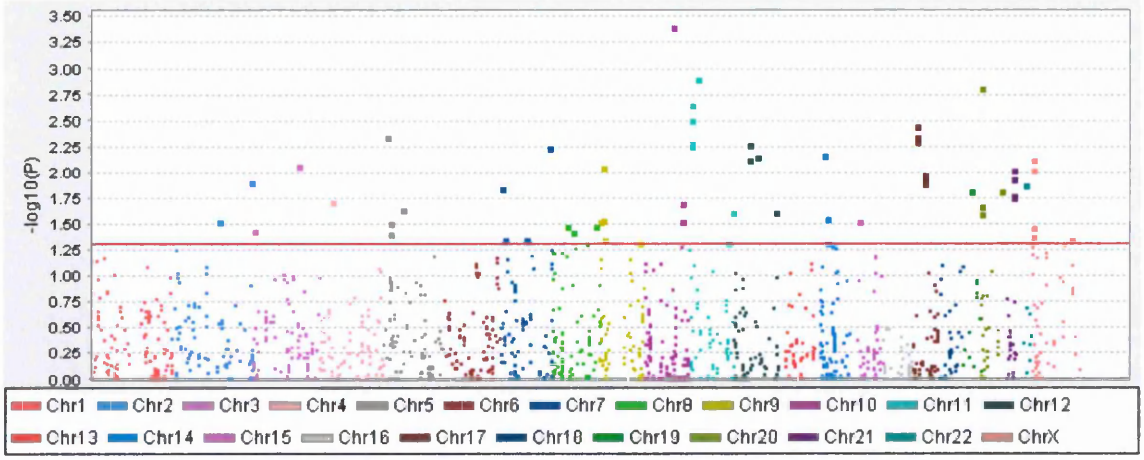



Figure 5.15: Allelic test P values of 1455 SNPs plotted against physical chromosome position. The P values were derived from comparison between 488 severe dengue cases and 544 uncomplicated dengue controls. The Y axis is the logarithm scale of P value, the X axis is the chromosomal position indicated by different colours. Red lines is the cut-off at P = 0.05.

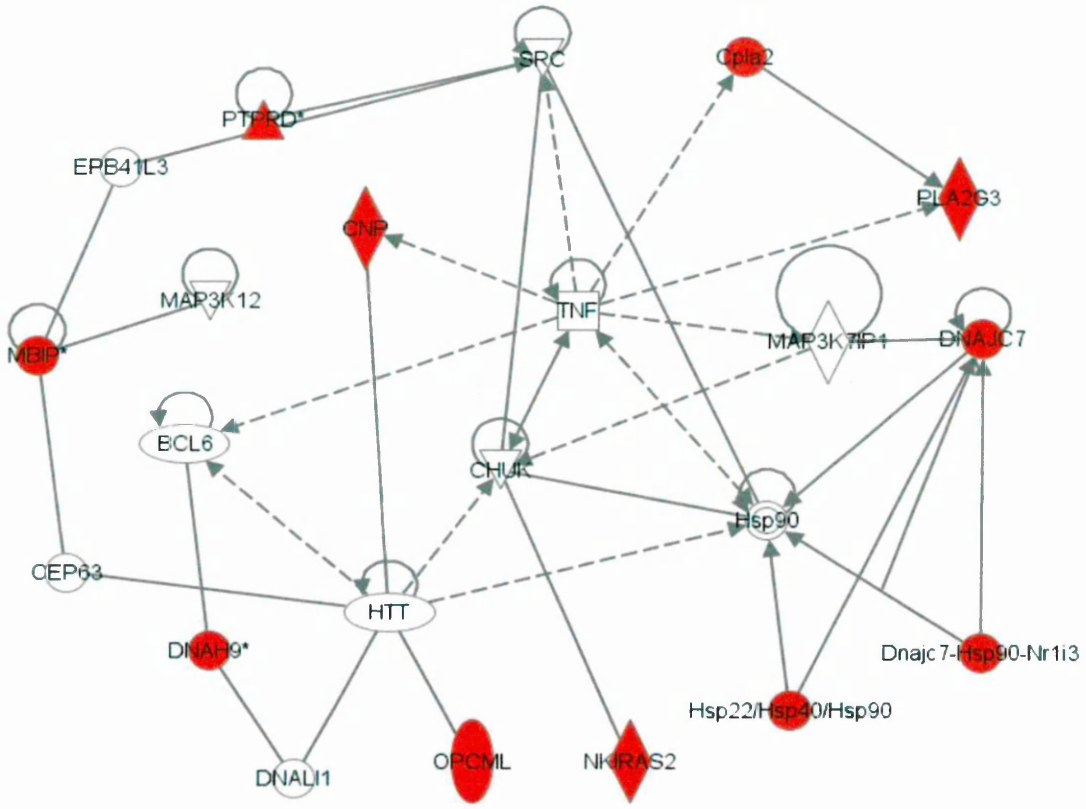



Figure 5.16: Gene network formed by genes that contain significant SNPs between severe dengue and uncomplicated dengue patients identified by IPA. From the list of 30 SNPs, we identified 21 genes which contain the SNPs or the closest genes up to 25kb up or downstream of the SNPs. Only 16 genes were recognised by IPA software. 11/16 genes directly or indirectly interact with each others to form the network. Direct interactions between genes were indicated in solid arrows. Indirect interactions were indicated by dash arrows. The network score given by IPA was 22 (equivalence to P value = $10^{-22}$, the network's Score = -log (Fisher's Exact test result) indicates the network was highly significant (not by chance). This is the number that ranks the network according to their degree of relevance to the Network Eligible Molecules in the dataset.

Table 5.11: Significant SNPs between severe dengue and uncomplicated dengue identified by WGAS, Illumina OPA validation and combined analysis

| SNP | Allele | MAF | Chr | Position | Type | Gene | Genome-Wide Association | | Illumnia OPA Validation | | Fisher's Combined Values | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | P value | OR (95% CI) | P value | OR (95% CI) | P value | OR (95% CI) |
| rs12595534 | A/G | 0.475 | 15 | 31458913 | Intergenic | KLF13 | 3.63E-04 | 0.46 (0.30 - 0.71) | 1.23E-03 | 0.67 (0.53 - 0.86) | 6.97E-06 | 0.61 (0.50 - 0.76) |
| rs1008405 | A/G | 0.232 | 20 | 2417085 | Downstream | AL049650.8 | 1.25E-03 | 1.79 (1.24 - 2.58) | 1.30E-03 | 1.37 (1.13 - 1.67) | 2.32E-05 | 1.46 (1.23 - 1.73) |
| rs13294631 | A/G | 0.150 | 9 | 8562080 | Intronic | PTPRD | 6.54E-05 | 2.82 (1.68 - 4.73) | 3.29E-02 | 1.36 (1.03 - 1.79) | 3.03E-05 | 1.60 (1.25 - 2.04) |
| rs4900833 | A/G | 0.067 | 14 | 36768276 | 3Prime_UTR | MBIP | 9.14E-05 | 4.04 (1.91 - 8.56) | 2.81E-02 | 1.56 (1.05 - 2.32) | 3.56E-05 | 1.92 (1.35 - 2.74) |
| rs4796751 | A/G | 0.127 | 17 | 40127537 | 3Prime_UTR | CNP | 1.33E-04 | 0.35 (0.20 - 0.61) | 3.30E-02 | 0.74 (0.56 - 0.99) | 5.84E-05 | 0.63 (0.49 - 0.82) |
| rs10839940 | A/G | 0.236 | 11 | 7953437 | Upstream | OR10A3 | 1.73E-03 | 2.01 (1.30 - 3.12) | 4.84E-03 | 2.08 (1.25 - 3.49) | 1.06E-04 | 2.04 (1.46 - 2.85 ) |
| rs2072193 | C/G | 0.457 | 22 | 31533796 | Non_synonymous | PLA2G3 | 2.04E-04 | 0.41 (0.25 - 0.66) | 4.23E-02 | 0.72 (0.52 - 0.98) | 1.09E-04 | 0.60 (0.46 - 0.79) |
| rs4239268 | A/G | 0.122 | 17 | 40154777 | Intronic | DNAJC7 | 2.23E-04 | 0.36 (0.20 - 0.63) | 3.89E-02 | 0.75 (0.56 - 1.00) | 1.10E-04 | 0.64 (0.50 - 0.83) |
| rs11078031 | A/C | 0.484 | 17 | 11661701 | Intronic | DNAH9 | 9.15E-04 | 2.37 (1.41 - 3.97) | 9.91E-03 | 1.36 (1.01 - 1.82) | 1.14E-04 | 1.56 (1.20 - 2.01) |
| rs11177202 | A/C | 0.047 | 12 | 68791570 | Intergenic | RAP1B | 1.33E-03 | 3.71 (1.58 - 8.69) | 7.00E-03 | 1.69 (1.15 - 2.48) | 1.17E-04 | 1.93 (1.36 - 2.74) |
| rs3744583 | A/G | 0.483 | 17 | 11656465 | Intronic | DNAH9 | 8.10E-04 | 2.41 (1.43 - 4.07) | 1.16E-02 | 1.32 (0.99 - 1.75) | 1.18E-04 | 1.51 (1.18 - 1.94) |
| rs16928238 | A/C | 0.255 | 9 | 8646183 | Intronic | PTPRD | 3.39E-04 | 1.90 (1.33 - 2.70) | 2.83E-02 | 1.25 (1.02 - 1.53) | 1.21E-04 | 1.39 (1.16 - 1.65) |
| rs824100 | A/G | 0.050 | 3 | 135172160 | Intergenic | Unknown | 1.25E-03 | 0.27 (0.11 - 0.63) | 8.49E-03 | 0.58 (0.39 - 0.87) | 1.32E-04 | 0.50 (0.35 - 0.73) |
| rs2236686 | A/G | 0.387 | 21 | 38888744 | Non_synonymous | AP001421.1 | 4.46E-04 | 2.20 (1.41 - 3.42) | 2.49E-02 | 1.17 (0.82 - 1.67) | 1.38E-04 | 1.50 (1.13 - 1.98) |
| rs3782999 | A/G | 0.378 | 13 | 99370485 | Intronic | SLC15A1 | 4.84E-04 | 0.47 (0.31 - 0.72) | 2.58E-02 | 0.84 (0.65 - 1.07) | 1.53E-04 | 0.73 (0.59 - 0.90) |
| rs4796759 | A/C | 0.124 | 17 | 40174841 | Intronic | NKIRAS2 | 3.82E-04 | 2.59 (1.52 - 4.41) | 3.56E-02 | 1.40 (1.06 - 1.85) | 1.66E-04 | 1.60 (1.25 - 2.05) |
| rs1482782 | C/G | 0.234 | 11 | 7941115 | Intergenic | OR10A6 | 1.04E-03 | 2.07 (1.34 - 3.19) | 1.33E-02 | 1.92 (1.14 - 3.24) | 1.69E-04 | 2.01 (1.43 - 2.80) |
| rs3844013 | A/G | 0.047 | 14 | 36745669 | Intergenic | MBIP | 3.34E-04 | 4.55 (1.85 - 11.21) | 4.85E-02 | 1.50 (1.00 - 2.25) | 1.95E-04 | 1.81 (1.25 - 2.62) |
| rs1448434 | C/G | 0.246 | 3 | 119043011 | Intronic | AC092981.3-1 | 3.93E-04 | 0.45 (0.28 - 0.70) | 4.68E-02 | 0.72 (0.56 - 0.93) | 2.19E-04 | 0.64 (0.51 - 0.80) |
| rs654036 | A/G | 0.301 | 10 | 113663573 | Intergenic | GPAM | 4.49E-04 | 0.46 (0.29 - 0.71) | 4.84E-02 | 0.92 (0.72 - 1.18) | 2.55E-04 | 0.78 (0.63 - 0.97) |
| rs2669501 | A/G | 0.435 | 13 | 104747823 | Intergenic | Unknown | 1.58E-03 | 2.10 (1.32 - 3.34) | 1.46E-02 | 1.37 (1.06 - 1.77) | 2.70E-04 | 1.52 (1.21 - 1.90) |
| rs4932739 | A/C | 0.024 | 19 | 24167222 | Intergenic | AC092279.2 | 2.09E-03 | 7.75 (1.76 - 34.14) | 1.54E-02 | 1.83 (1.12 - 3.01) | 3.65E-04 | 2.12 (1.32 - 3.40) |
| rs2105632 | A/G | 0.398 | 11 | 132336927 | Intronic | OPCML | 8.64E-04 | 2.88 (1.52 - 5.48) | 4.24E-02 | 1.39 (1.07 - 1.81) | 4.11E-04 | 1.54 (1.21 - 1.96) |
| rs6972511 | A/G | 0.175 | 7 | 78091663 | Intronic | AC007237.2 | 9.58E-04 | 0.51 (0.34 - 0.76) | 4.45E-02 | 0.79 (0.62 - 0.99) | 4.71E-04 | 0.71 (0.58 - 0.87) |
| rs17754112 | A/C | 0.117 | 14 | 66453174 | Intergenic | AL391261.3-2 | 1.51E-03 | 0.42 (0.25 - 0.73) | 2.98E-02 | 0.72 (0.55 - 0.95) | 4.95E-04 | 0.65 (0.51 - 0.82) |
| rs880056 | A/G | 0.366 | 8 | 59135137 | Intergenic | UBXN2B | 1.21E-03 | 0.60 (0.44 - 0.82) | 3.73E-02 | 0.82 (0.69 - 0.99) | 4.97E-04 | 0.76 (0.65 - 0.89) |
| rs10094349 | A/G | 0.375 | 8 | 4190312 | Intergenic | Unknown | 1.54E-03 | 0.33 (0.16 - 0.67) | 3.37E-02 | 0.95 (0.74 - 1.21) | 5.65E-04 | 0.84 (0.67 - 1.07) |
| rs10493264 | A/G | 0.370 | 1 | 59891290 | Intronic | FGGY | 1.30E-03 | 0.35 (0.18 - 0.68) | 4.49E-02 | 0.94 (0.73 - 1.20) | 6.28E-04 | 0.83 (0.66 - 1.05) |
| rs7993074 | A/G | 0.354 | 13 | 104625030 | Intergenic | Unknown | 1.74E-03 | 2.80 (1.44 - 5.42) | 4.08E-02 | 1.27 (0.89 - 1.80) | 7.49E-04 | 1.51 (1.10 - 2.06) |
| rs836137 | A/G | 0.397 | 11 | 34548424 | Intergenic | EHF | 1.74E-03 | 0.51 (0.33 - 0.78) | 4.67E-02 | 0.74 (0.58 - 0.94) | 8.46E-04 | 0.67 (0.54 - 0.83) |

## 5.5 Discussion

This is the first GWAS to look at host genetic susceptibility to dengue and dengue shock syndrome. We demonstrated for the first time that Vietnamese Kinh is distinct from other Asian populations such as the Chinese and Japanese. The most significant SNPs that were associated with susceptibility to dengue were rs1366478 ($P = 6.3x10^{-5}$), rs17061643 ($P = 8.7x10^{-5}$), rs10204649 ($P = 2.9x10^{-4}$) and rs298917 ($P = 6x10^{-4}$). The most significant SNPs that were associated with severe dengue are rs12595534 ($P = 6.97x10^{-6}$), rs1008405 ($p = 2.3x10^{-5}$), rs13294631 ($P = 3.03x10^{-5}$), rs4900833 ($P = 3.5x10^{-5}$), and rs4796751 ($P = 5.8x10^{-5}$). None of these SNPs reached a level that is currently regarded as significant in a GWAS ($10^{-7}$) [195]. We also identified a network that suggests biologically plausible functional relationships, which are extremely unlikely to have occurred by chance exist between genes that contain or are in the intermediate vicinity of the SNPs that are associated with severe dengue.

The limitation of this study was that power calculation was not performed. This was because the study was conducted in 2007 when the knowledge on the effect size of the disease was not available and that this was among the first GWAS in infectious disease when only very limited data on population structure and allele frequency in the Asian population was available. Infectious diseases are complex and an individual's susceptibility to the disease may be determined by multiple variants in multiple genes with small effect. It was shown recently that a sample size of 2000 cases and 2000 controls is needed to obtain sufficient power to detect associations with a modest effect size (O.R = 1.5) with a minor allele frequency greater than or equal to 5% [196]. This indicates clearly that our study with small sample size has insufficient power to detect any associated SNPs with an allele frequency of less than 5% and a minimum effect size of 1.5 . Nevertheless,

the data from this chapter provide insight into certain aspects of genetic susceptibility to dengue infection.

Understanding host genetic predisposition in other infectious diseases has lead to an improved understanding of pathogenesis of TB, HIV and malaria [197-203]. The genetic-based discovery that CCR5 is a co-receptor for HIV has lead to the development of therapeutic drugs that block CCR5. In dengue, an understanding of host genetic susceptibility could also assist in the development of antiviral drugs that prevent virus infection or modulate the host response. We identified several genes by GWAS that are implicated in susceptibility to dengue, but studies on the functions of the genes in the context of dengue are warranted.

Studies on HLA region s [121, 122, 129-133, 194], DC-SIGN, FcγR, Vitamin D receptor, and TAP [136-138] identified polymorphisms that were putatively associated with dengue. Unfortunately, none of these HLA polymorphisms have been replicated in independent cohorts. Moreover, many of these studies had small sample sizes with uncertain of case ascertainment and with unknown levels of population stratification. Surprisingly, we did not identify any MHC-related SNPs that were associated with dengue or with severe dengue. The inability to replicate previous cases-control studies of the HLA region in Vietnamese population [131] and Thai population [130, 132] might be because of the differences in genetic background between populations. Other possible explanations include the small sample size of previous studies and heterogeneous case mix.

Quality control is very important in GWAS because without stringent QC, spurious results could emerge given the huge number of tests required for analysis. We performed very stringent QC for every experimental step and also for data analysis. The genomic DNA was carefully QC'd to ensure that they were of good quality for the genotyping process. The genotyping data was assessed twice with DM algorithm and BRLMM algorithm to

assure that the SNP and sample call rates were high enough (≥93%) for downstream analysis. Population stratification in GWAS is a common problem but by Q-Q plots and by PCA, we showed that our cases and controls were homogenous. Sample mislabelling, sample duplication or samples from closely related relatives were also excluded through checking the IBD proportion that were shared between any pair of case-case, control – control or case - control. We also visually checked the potentially associated SNPs by looking at the clusters of genotype of SNPs (AA, AB and BB) called by the BRLMM algorithm. Spurious results could emerge in GWAS so that replication of GWAS results with an independent larger sample set is crucial. We performed a follow-up replication with a 2nd cohort of larger sample size to look at a smaller number of SNPs.

We successfully validated ten SNPs from the GWAS ($P < 0.05$) in the replication set with a larger sample size, with P values ranging from $6x10^{-5}$ to $1x10^{-4}$. The most significant SNP was rs1366478 ($P = 6.3x10^{-5}$) on the intergenic region of chromosome 5. The second most significant SNP was rs17061643 ($p = 8.7x10^{-5}$) which maps to the intron of the *KFL12* (Kruppel-like factor 12) gene on chromosome 13. The protein encoded by this gene is a member of Kruppel-like zinc finger protein family of which several members such as *KLF2*, *KLF4* and *KLF6* have been postulated to play an important role in mediating the function of endothelial cell. *KLF2* null mice died as a result of haemorrhage caused by incorrect vessel stabilization and defective tunica media formation (reviewed in [204, 205]). Furthermore, Kruppel-like factors are also expressed in endothelial cells during vasculogenesis and early angiogenesis (reviewed in [204]). Another SNPs that also showed evidence of genome-wide association but with lower significance included rs10204649 ($P = 2.9x10^{-4}$) and rs298917 ($P = 6x10^{-4}$) which mapped to the intron of the *VWA3B* gene (von Willebrand factor A domain containing 3B) on chromosome 2. Von Willebrand factor, primarily synthesised in vascular endothelial cells, is a multimeric plasma protein that mediates platelet adhesion and aggregation at the site of vascular injury which in turn

results in thrombocytopenia (reviewed in [206]). Further studies on the functions of *KLF12* and *VWA3B* are warranted. It is striking that 3/10 SNPs that were validated in the 2nd sample set were linked to genes involved in coagulation and endothelial cell function, both of which are central to the pathogenesis of dengue.

In a nested study, we also tried to indentify SNPs that were associated with severe dengue by comparing frequencies of SNPs between DSS and uncomplicated dengue patients. We identified 30 SNPs that showed strong evidence for association with severe dengue (uncorrected P value < $8x10^{-4}$) in both GWAS and replication study. The most significant SNP was rs12595534 (P = $6.9x10^{-6}$), which maps to the intergenic region of approximately 160kb upstream of the *KLF13* (Kruppel-like factor 13) gene on chromosome 15. This gene belongs to a family of transcriptional factors and is involved in B and T cell development [207]. Rs13294631 and rs16928238 were two significant SNPs that map to the intron of the *PTPRD* gene (protein tyrosine phosphatase, receptor type, D) on chromosome 9 with significance level at P = $2.3x10^{-5}$ and P = $1.2x10^{-4}$ respectively. The distance between these two SNPs was 84kb apart indicating that they were independently identified. The protein encoded by this gene is a member of the protein tyrosine phosphatase (PTP) family. PTPs are known to be signalling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. However, the functional effect of this protein in dengue has not been documented.

We used IPA in an unsupervised manner which allows identification of gene - gene relationships without a priori assumptions to investigate the interconnectedness of genes contain or near by the significant SNPs. This analysis linked eight of the 21 genetically associated genes that contains or near by the associated SNPs (25kb up or downstream) in a network that was associated with cellular movement, cell signalling, nervous system development and function. The key molecules of the network were *TNF* (Tumour Necrosis

Factor), hsp90 (Heat Shock Protein 90kDa) and *HTT* (Huntingin). TNF-α has been well documented in terms of its suspected contributions to dengue pathogenesis (reviewed in [208-210]). Hsp90 is an ATP dependent molecular chaperone which can regulate more than 100 proteins most of which are related to signal transduction. Many signal transduction systems use their interaction with hsp90 as an essential component of the signalling pathway. Several protein kinases, including the Src and Raf components of the MAP kinase system, are also bound to hsp90 (reviewed in [211]). *PLA2G3* belongs to the family of secreted phospholipase A2 which was postulated to play important roles in systemic inflammatory processes such as sepsis and acute lung injury [212]. *DNAJC7* (hsp40) is another heat shock protein which is demonstrated to have disease-suppressive activities. In PBMCs of patients with rheumatoid arthritis, hsp40 can induce secretion of IL-10 and reduce TNF-α secretion [213]. *NKIRAS2*, a member of NFκB pathway, was demonstrated to be regulated by mir-29B which negatively regulates expression of extracellular matrix related genes in chronic oxidative stress conditions [214]. *OPCML* (opioid binding protein/cell adhesion molecule-like) encodes a member of the IgLON subfamily in the immunoglobulin protein superfamily which was postulated to have tumor suppression function [215]. Investigations of the contribution of this gene network to dengue pathogenesis are warranted.

The DC-SIGN polymorphisms (rs4804803, rs8105483 and rs11465391) were demonstrated to confer protection to mild dengue against DHF in Thai patients [137]. However, our result for these DC-SIGN polymorphisms in Vietnamese dengue patients did not show any association with severe dengue. The reasons for these differences most likely reflect the different disease phenotype defined in our study versus the Thai study, where they defined "cases" as patients with DHF grade I and II and controls were DF, while our study defined "cases" was those with DSS and controls were DHF grade I and II and the different in genetic background of Thai and Vietnamese.

GWAS for infectious diseases are at an early stage with very limited of success because of the fact that infectious disease is polygenic and multi-factorial and limitation in the current approaches. Although several polymorphisms have been identified as susceptibility variants for infectious disease such as TB (*SLC11A1* and several others in *CCL1*) [197, 198], HIV-1 (*CCR5, HLA, HCP5, RNF39/ZNRD1* and *PROX1*) [199-202] and malaria in African (HbS) [203], many other genes and environmental factors as well as the gene-gene or gene-environment interactions that are potentially associated with the diseases are unidentified. To our knowledge, no GWAS study has been performed in dengue. For this reason, we did not have a chance to compare our data with data from other groups.

In conclusion, this is the first genome-wide association study in dengue in which all the cases were classified based on the most up-to-date WHO guideline. We also applied rigorous quality controls in every step of the genotyping experiment and also in data analysis. We are confident that our cases and controls were genetically homogenous. For potentially true associations, we visually checked their genotyping quality to assure that they were not genotyping errors. We identified several SNPs that showed evidence for significant association with dengue relative to controls or severe dengue relative to uncomplicated dengue. None of the SNPs was significant at suggestive genome-wide level but ontology analysis of genes implicated by associated SNPs suggests biologically plausible functional relationships between genes. Our study also suggested that much larger sample size is needed to be able to detect associations at genome-wide significant level but its preliminary findings provide basic knowledge of host genetic susceptibility in dengue and are valuable for future studies.

CHAPTER 6

# 6. GENERAL DISCUSSION

This thesis filled several significant gaps in our understanding of dengue pathogenesis. First, we defined the whole-blood transcriptional profile in acute dengue patients with mild and severe clinical outcomes. These are some of the first reports of the molecular response in dengue patients and shed light on the host response to systemic viral infections. An understanding of global gene expression in acute DENV infection could be useful for the evaluation of dengue vaccines as a successful vaccine could be expected to elicit a host response that mimics wild-type infection but without clinical signs. An improved understanding of the host response in acute infection should also be useful when considering clinical interventions in dengue. In particular, our data from chapter 3 and 4 suggested that the "host defense" transcriptional profile is observed well before the point of defervescence and cardiovascular decompensation in DSS patients. This may explain why adjuvant corticosteroid use in patients with established DSS [161] has little efficacy- in simple terms the inflammatory immune response "horse" has already "bolted" by the time cardiovascular decompensation occurs. A better approach is therefore to give adjuvant corticosteroid therapy early in the infection, at a time when the major elements of the host anti-viral and inflammatory immune response are most prominent and which possibly drive the capillary permeability syndrome that characterizes severe dengue. To this end, the OUCRU is currently conducting a randomized controlled trial of prednisolone, a corticosteroid, for the early treatment of dengue in children with <3 days of fever. This initial trial in 225 children trial will seek to determine if prednisolone is safe and can attenuate the host inflammatory response. In the future, it is possible that we will conduct larger trials with a view to determining if prednisolone prevents the development of important clinical complications such as DSS. In more general terms, it is plausible that insights into the host response to dengue as illuminated by the gene expression studies in this thesis will be relevant to other systemic viral infections, in particular flavivirus

infections with systemic phases such as Yellow Fever or Japanese encephalitis virus infection.

One of our major aims in characterising the transcriptional profile in children with dengue was to identify signatures that were associated with progression to DSS. Intriguingly, our findings identified several neutrophil-associated transcripts as more abundant in Vietnamese patients who progressed to DSS, suggesting a hitherto unrecognised association between neutrophil activation, pathogenesis and the development of DSS. We postulated that neutrophil adherence to endothelial cells and the secretion of soluble mediators of vascular permeability may represent a small step in the inflammatory cascade that synergises with other adaptive host responses to mediate capillary permeability in severe secondary infection. In infants with primary infection and severe dengue, an innate response that includes neutrophil activation, together with an inherently permeable vascular endothelium, might be sufficient to trigger clinically significant vascular leak. Our preliminary findings hinted that further studies on the contributions of neutrophils to dengue pathogenesis are required. Such studies of might be conducted in animal models of dengue, which have improved recently [216-218]. Alternatively, studies of DENV and neutrophil interactions the might also be performed on patient blood samples ex vivo using flow cytometry or in vitro using neutrophils in culture.

A lesson from our studies on global gene expression (Chapter 3 and Chapter 4) was that large patient sample sizes, with serial samples collected early on in the infection, offers the best way to gather further insights into the host transcriptional response. Such studies are difficult in children and complicated by the fact that outcomes such as DSS are relatively rare event. Nevertheless, studies similar to ours should be conducted in other populations of patients with different ethnic backgrounds and different circulating viruses. Alternative approaches to microarrays for gene profiling are also become available. Recently, a new high-throughput sequencing-based technique allowing whole transcriptome surveys from

subnanogram RNA quantities in an amplification-free manner was developed [219]. This technique could be a first choice for the coming years to replace gene expression microarrays. This approach would overcome some of the limitations of microarrays, e.g. this is an amplification-dependent method in which RNA has to be amplified before being hybridized onto the probes with possible introduction of bias.

Dengue, like other infectious diseases, is a complex disease in which outcome may not be decided by one factor but it may be the combination of multiple factors such as viral factors and host immunological and genetic factors. In Chapter 5 we conducted a genome-wide case – control association study to identify genetic variants that are associated with susceptibility to dengue or DSS by a GWAS approach. One of the limitations of GWAS is that it skewed toward common variants in the human genome with at least 1% so that if the disease is caused by one or combination of rare variants, it is impossible to detect them using current technology. Nonetheless, we demonstrated for the first time that Vietnamese Kinh ethnicity is homogenous and that they are not belong to other Asian populations such as Han Chinese and Japanese. This indicated also that cord blood samples could be used as population controls for GWAS with other infectious diseases. We identified ten and 30 SNPs that showed evidence for significant association with susceptibility to dengue and with DSS, respectively. However, none of the SNPs reached a genome-wide significant level (which is $10^{-7}$) [195], and the functional contributions of the related genes to dengue pathogenesis are unclear. Further replication studies focused on these variants are warranted, as are more functional studies of these genes during DENV infection. To this end, using the results from this thesis as a foundation, the OUCRU has recently commenced a GWAS with 2000 DSS cases and 2000 cord blood controls in collaboration with the Genome Institute of Singapore and funded by a Wellcome Trust project grant. It is possible then that some of the SNPs identified in this study will be replicated in this $2^{nd}$ samples set of patients. In general terms, future research into host genetic predisposition to

dengue and severe dengue could help not only to understand the disease pathogenesis but also to identify avenues to develop anti-DENV drugs. For example, the finding that individuals with functional deletions in their CCR5 receptor were highly resistant to HIV infection (reviewed in [220, 221]) subsequently led to the development and commercialisation of a small molecule drugs that block this virus-receptor interaction and thereby reduces HIV viral loads [222, 223]. A similar outcome is conceivable for dengue and would represent the ultimate justification for genomic research into the pathogenesis of this globally important disease.

## References


1. WHO/TDR, *DENGUE, Guidelines for diagnosis, treatment, prevention and control.* 2009.
2. Coffey, L.L., et al., *Human genetic determinants of dengue virus susceptibility.* Microbes Infect, 2009. **11**(2): p. 143-56.
3. Gubler, D.J. and G. Kuno, *Dengue and dengue hemorrhagic fever.* 1997, New York: CAB International.
4. Diallo, M., et al., *Potential role of sylvatic and domestic African mosquito species in dengue emergence.* Am J Trop Med Hyg, 2005. **73**(2): p. 445-9.
5. Laferl, H., et al., *Imported dengue fever in Austria 1990-2005.* Travel Med Infect Dis, 2006. **4**(6): p. 319-23.
6. Halstead, S.B., *Dengue haemorrhagic fever--a public health problem and a field for research.* Bull World Health Organ, 1980. **58**(1): p. 1-21.
7. Hayes, E.B. and D.J. Gubler, *Dengue and dengue hemorrhagic fever.* Pediatr Infect Dis J, 1992. **11**(4): p. 311-7.
8. Monath, T.P., *Dengue: the risk to developed and developing countries.* Proc Natl Acad Sci U S A, 1994. **91**(7): p. 2395-400.
9. Axel Kroeger, M.B.N., *Dengue: setting the global research agenda.* The Lancet, 2006. **368**: p. 2193-2195.
10. Gubler, D.J., *Dengue and dengue hemorrhagic fever: its history and resurgence as a global public health problem*, in *Dengue and dengue hemorrhagic fever*, D.J. Gubler and G. Kuno, Editors. 1997, CAB Internaltional: New York. p. p: 1-22.
11. Dai, B., *Dengue xuat huyet.* 2002.
12. Pasteur Institute, H.C.M.C., Vietnam, *Report of Dengue epidemiology, prevention and treatment activities in the south of Vietnam.* 2007(Jan 2008): p. 1-8.
13. Rodhain, F. and L. Rosen, *Mosquito vectors and dengue virus-vector relationships*, in *Dengue and dengue hemorrhagic fever*, D.J. Gubler and G. Kuno, Editors. 1997, CAB Internaltional: New York

p. p: 45-60.

14. Smith, *The history of dengue in tropical Asia and its probable relationship to the mosquito Aedes aegypti.* Tropical Medicine and Hygiene, 1956.
15. Gratz, N.G., *Critical review of the vector status of Aedes albopictus.* Med Vet Entomol, 2004. **18**(3): p. 215-27.
16. Trpis, M. and W. Hausermann, *Dispersal and other population parameters of Aedes aegypti in an African village and their possible significance in epidemiology of vectorborne diseases.* American Journal of Tropical Medicine and Hygiene, 1986. **35**: p. 143-152.
17. Jetten, T.H. and D.A. Focks, *Potential changes in the distribution of dengue transmission under climate warming.* Am J Trop Med Hyg, 1997. **57**(3): p. 285-97.
18. Blanc, G. and J. Caminopetros, *Recherches experimentales sur la dengue*. 1930, Archives de l'Institut Pasteur Hellenique. p. 2: 199-276.
19. Rosen, L., et al., *Transovarial transmission of dengue viruses by mosquitoes: Aedes albopictus and Aedes aegypti.* Am J Trop Med Hyg, 1983. **32**(5): p. 1108-19.
20. Gunther, J., et al., *Evidence of vertical transmission of dengue virus in two endemic localities in the state of Oaxaca, Mexico.* Intervirology, 2007. **50**(5): p. 347-52.
21. WHO, *Dengue haemorrhagic fever- Diagnosis, treatment, prevention and control.* 1997, World Health Organization: Geneva.
22. Tricou, V., et al., *A randomized controlled trial of chloroquine for the treatment of dengue in Vietnamese adults.* PLoS Negl Trop Dis. **4**(8): p. e785.
23. Shu, P.Y., et al., *Comparison of capture immunoglobulin M (IgM) and IgG enzyme-linked immunosorbent assay (ELISA) and nonstructural protein NS1 serotype-specific IgG ELISA for differentiation of primary and secondary dengue virus infections.* Clin Diagn Lab Immunol, 2003. **10**(4): p. 622-30.

24. Innis, B.L., et al., *An enzyme-linked immunosorbent assay to characterize dengue infections where dengue and Japanese encephalitis co-circulate.* Am J Trop Med Hyg, 1989. **40**(4): p. 418-27.
25. Srikiatkhachorn, A., et al., *Natural History of Plasma Leakage in Dengue Hemorrhagic Fever: A Serial Ultrasonographic Study.* Pediatr Infect Dis J, 2007. **26**(4): p. 283-290.
26. Hunsperger, E.A., et al., *Evaluation of commercially available anti-dengue virus immunoglobulin M tests.* Emerg Infect Dis, 2009. **15**(3): p. 436-40.
27. Deubel, V., et al., *Identification of dengue sequences by genomic amplification: rapid diagnosis of dengue virus serotypes in peripheral blood.* J Virol Methods, 1990. **30**(1): p. 41-54.
28. Lanciotti, R.S., et al., *Rapid detection and typing of dengue viruses from clinical samples by using reverse transcriptase-polymerase chain reaction.* J Clin Microbiol, 1992. **30**(3): p. 545-51.
29. Wang, W.K., et al., *Detection of dengue virus replication in peripheral blood mononuclear cells from dengue virus type 2-infected patients by a reverse transcription-real-time PCR assay.* J Clin Microbiol, 2002. **40**(12): p. 4472-8.
30. Chutinimitkul, S., et al., *Dengue typing assay based on real-time PCR using SYBR Green I.* J Virol Methods, 2005.
31. Shurtleff, A.C., et al., *Genetic variation in the 3' non-coding region of dengue viruses.* Virology, 2001. **281**(1): p. 75-87.
32. Chien, L.J., et al., *Development of real-time reverse transcriptase PCR assays to detect and serotype dengue viruses.* J Clin Microbiol, 2006. **44**(4): p. 1295-304.
33. Vaughn, D.W., et al., *Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity.* J Infect Dis, 2000. **181**(1): p. 2-9.
34. Alcon, S., et al., *Enzyme-linked immunosorbent assay specific to Dengue virus type 1 nonstructural protein NS1 reveals circulation of the antigen in the blood during the acute phase of disease in patients experiencing primary or secondary infections.* J Clin Microbiol, 2002. **40**(2): p. 376-81.
35. Chuansumrit, A., et al., *The use of dengue nonstructural protein 1 antigen for the early diagnosis during the febrile stage in patients with dengue infection.* Pediatr Infect Dis J, 2008. **27**(1): p. 43-8.
36. Hang, V.T., et al., *Diagnostic Accuracy of NS1 ELISA and Lateral Flow Rapid Tests for Dengue Sensitivity, Specificity and Relationship to Viraemia and Antibody Responses.* PLoS Negl Trop Dis, 2009. **3**(1): p. e360.
37. Kumarasamy, V., et al., *Evaluating the sensitivity of a commercial dengue NS1 antigen-capture ELISA for early diagnosis of acute dengue virus infection.* Singapore Med J, 2007. **48**(7): p. 669-73.
38. Blacksell, S.D., et al., *Evaluation of the Panbio dengue virus nonstructural 1 antigen detection and immunoglobulin M antibody enzyme-linked immunosorbent assays for the diagnosis of acute dengue infections in Laos.* Diagn Microbiol Infect Dis, 2008. **60**(1): p. 43-9.
39. Dussart, P., et al., *Evaluation of Two New Commercial Tests for the Diagnosis of Acute Dengue Virus Infection Using NS1 Antigen Detection in Human Serum.* PLoS Negl Trop Dis, 2008. **2**(8): p. e280.
40. Yamada, K., et al., *Virus isolation as one of the diagnostic methods for dengue virus infection.* J Clin Virol, 2002. **24**(3): p. 203-9.
41. Shu, P.Y., et al., *Dengue virus serotyping based on envelope and membrane and nonstructural protein NS1 serotype-specific capture immunoglobulin M enzyme-linked immunosorbent assays.* J Clin Microbiol, 2004. **42**(6): p. 2489-94.
42. Teles, F.R., D.M. Prazeres, and J.L. Lima-Filho, *Trends in dengue diagnosis.* Rev Med Virol, 2005.

43. Mackenzie, J.S., D.J. Gubler, and L.R. Petersen, *Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses.* Nat Med, 2004. **10**(12 Suppl): p. S98-109.
44. Cleaves, G.R. and D.T. Dubin, *Methylation status of intracellular dengue type 2 40 S RNA.* Virology, 1979. **96**(1): p. 159-65.
45. Henchal, E.A. and J.R. Putnak, *The dengue viruses.* Clin Microbiol Rev, 1990. **3**(4): p. 376-96.
46. Bartenschlager, R. and S. Miller, *Molecular aspects of Dengue virus replication.* Future Microbiol, 2008. **3**: p. 155-65.
47. Chiu, W.W., R.M. Kinney, and T.W. Dreher, *Control of translation by the 5'- and 3'-terminal regions of the dengue virus genome.* J Virol, 2005. **79**(13): p. 8303-15.
48. Yu, L. and L. Markoff, *The topology of bulges in the long stem of the flavivirus 3' stem-loop is a major determinant of RNA replication competence.* J Virol, 2005. **79**(4): p. 2309-24.
49. Alvarez, D.E., et al., *Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication.* Virology, 2005. **339**(2): p. 200-12.
50. Khromykh, A.A., et al., *Essential role of cyclization sequences in flavivirus RNA replication.* J Virol, 2001. **75**(14): p. 6719-28.
51. Tadano, M., et al., *Detection of dengue 4 virus core protein in the nucleus. I. A monoclonal antibody to dengue 4 virus reacts with the antigen in the nucleus and cytoplasm.* J Gen Virol, 1989. **70 ( Pt 6)**: p. 1409-15.
52. Stiasny, K. and F.X. Heinz, *Flavivirus membrane fusion.* J Gen Virol, 2006. **87**(Pt 10): p. 2755-66.
53. Chen, Y., T. Maguire, and R.M. Marks, *Demonstration of binding of dengue virus envelope protein to target cells.* J Virol, 1996. **70**(12): p. 8765-72.
54. Kuhn, R.J., et al., *Structure of dengue virus: implications for flavivirus organization, maturation, and fusion.* Cell, 2002. **108**(5): p. 717-25.
55. Tan, B.H., et al., *Recombinant dengue type 1 virus NS5 protein expressed in Escherichia coli exhibits RNA-dependent RNA polymerase activity.* Virology, 1996. **216**(2): p. 317-25.
56. Falgout, B., et al., *Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins.* J Virol, 1991. **65**(5): p. 2467-75.
57. Matusan, A.E., et al., *Mutagenesis of the Dengue virus type 2 NS3 protein within and outside helicase motifs: effects on enzyme activity and virus replication.* J Virol, 2001. **75**(20): p. 9633-43.
58. Winkler, G., et al., *Newly synthesized dengue-2 virus nonstructural protein NS1 is a soluble protein but becomes partially hydrophobic and membrane-associated after dimerization.* Virology, 1989. **171**(1): p. 302-5.
59. Flamand, M., et al., *Dengue virus type 1 nonstructural glycoprotein NS1 is secreted from mammalian cells as a soluble hexamer in a glycosylation-dependent fashion.* J Virol, 1999. **73**(7): p. 6104-10.
60. Libraty, D.H., et al., *High circulating levels of the dengue virus nonstructural protein NS1 early in dengue illness correlate with the development of dengue hemorrhagic fever.* J Infect Dis, 2002. **186**(8): p. 1165-8.
61. Avirutnan, P., et al., *Vascular leakage in severe dengue virus infections: a potential role for the nonstructural viral protein NS1 and complement.* J Infect Dis, 2006. **193**(8): p. 1078-88.
62. Avirutnan, P., et al., *Secreted NS1 of dengue virus attaches to the surface of cells via interactions with heparan sulfate and chondroitin sulfate E.* PLoS Pathog, 2007. **3**(11): p. e183.
63. Pierson, T.C. and M.S. Diamond, *Molecular mechanisms of antibody-mediated neutralisation of flavivirus infection.* Expert Rev Mol Med, 2008. **10**: p. e12.
64. Churdboonchart, V., et al., *Antibodies against dengue viral proteins in primary and secondary dengue hemorrhagic fever.* Am J Trop Med Hyg, 1991. **44**(5): p. 481-93.

65. Calvert, A.E., et al., *Non-structural proteins of dengue 2 virus offer limited protection to interferon-deficient mice after dengue 2 virus challenge.* J Gen Virol, 2006. **87**(Pt 2): p. 339-46.
66. Chung, K.M., et al., *Antibodies against West Nile Virus nonstructural protein NS1 prevent lethal infection through Fc gamma receptor-dependent and -independent mechanisms.* J Virol, 2006. **80**(3): p. 1340-51.
67. Wu, S.F., et al., *Evaluation of protective efficacy and immune mechanisms of using a non-structural protein NS1 in DNA vaccine against dengue 2 virus in mice.* Vaccine, 2003. **21**(25-26): p. 3919-29.
68. Garcia, G., et al., *Antibodies from patients with dengue viral infection mediate cellular cytotoxicity.* J Clin Virol, 2006. **37**(1): p. 53-7.
69. Laoprasopwattana, K., et al., *Antibody-Dependent Cellular Cytotoxicity Mediated by Plasma Obtained before Secondary Dengue Virus Infections: Potential Involvement in Early Control of Viral Replication.* J Infect Dis, 2007. **195**(8): p. 1108-16.
70. Green, S., et al., *Early CD69 expression on peripheral blood lymphocytes from children with dengue hemorrhagic fever.* J Infect Dis, 1999. **180**(5): p. 1429-35.
71. Shresta, S., et al., *Early activation of natural killer and B cells in response to primary dengue virus infection in A/J mice.* Virology, 2004. **319**(2): p. 262-73.
72. Azeredo, E.L., et al., *NK cells, displaying early activation, cytotoxicity and adhesion molecules, are associated with mild dengue disease.* Clin Exp Immunol, 2006. **143**(2): p. 345-56.
73. Mathew, A. and A.L. Rothman, *Understanding the contribution of cellular immunity to dengue disease pathogenesis.* Immunol Rev, 2008. **225**(1): p. 300-13.
74. Rothman, A.L., *Dengue: defining protective versus pathologic immunity.* J Clin Invest, 2004. **113**(7): p. 946-51.
75. Mongkolsapaya, J., et al., *Original antigenic sin and apoptosis in the pathogenesis of dengue hemorrhagic fever.* Nat Med, 2003. **9**(7): p. 921-7.
76. Rothman, A.L., I. Kurane, and F.A. Ennis, *Multiple specificities in the murine CD4+ and CD8+ T-cell response to dengue virus.* J Virol, 1996. **70**(10): p. 6540-6.
77. de Kruif, M.D., et al., *Differential gene expression changes in children with severe dengue virus infections.* PLoS Negl Trop Dis, 2008. **2**(4): p. e215.
78. Simmons, C.P., et al., *Patterns of host genome-wide gene transcript abundance in the peripheral blood of patients with acute dengue hemorrhagic fever.* J Infect Dis, 2007. **195**(8): p. 1097-107.
79. Azeredo, E.L., et al., *Characterisation of lymphocyte response and cytokine patterns in patients with dengue fever.* Immunobiology, 2001. **204**(4): p. 494-507.
80. Bosch, I., et al., *Increased production of interleukin-8 in primary human monocytes and in human epithelial and endothelial cell lines after dengue virus challenge.* J Virol, 2002. **76**(11): p. 5588-97.
81. Lee, Y.R., et al., *Dengue viruses can infect human primary lung epithelia as well as lung carcinoma cells, and can also induce the secretion of IL-6 and RANTES.* Virus Res, 2007. **126**(1-2): p. 216-25.
82. Nguyen, T.H., et al., *Dengue hemorrhagic fever in infants: a study of clinical and cytokine profiles.* J Infect Dis, 2004. **189**(2): p. 221-32.
83. Lee, Y.R., et al., *MCP-1, a highly expressed chemokine in dengue haemorrhagic fever/dengue shock syndrome patients, may cause permeability change, possibly through reduced tight junctions of vascular endothelium cells.* J Gen Virol, 2006. **87**(Pt 12): p. 3623-30.
84. Chen, H.C., et al., *Both virus and tumor necrosis factor alpha are critical for endothelium damage in a mouse model of dengue virus-induced hemorrhage.* J Virol, 2007. **81**(11): p. 5518-26.

85. Yen, Y.T., et al., *Enhancement by tumor necrosis factor alpha of dengue virus-induced endothelial cell production of reactive nitrogen and oxygen species is key to hemorrhage development.* J Virol, 2008. **82**(24): p. 12312-24.
86. Chen, J.P., et al., *Dengue virus induces expression of CXC chemokine ligand 10/IFN-gamma-inducible protein 10, which competitively inhibits viral binding to cell surface heparan sulfate.* J Immunol, 2006. **177**(5): p. 3185-92.
87. Hsieh, M.F., et al., *Both CXCR3 and CXCL10/IFN-Inducible Protein 10 Are Required for Resistance to Primary Infection by Dengue Virus.* J Immunol, 2006. **177**(3): p. 1855-63.
88. Libraty, D.H., et al., *Differing influences of virus burden and immune activation on disease severity in secondary dengue-3 virus infections.* J Infect Dis, 2002. **185**(9): p. 1213-21.
89. Wang, W.K., et al., *High levels of plasma dengue viral load during defervescence in patients with dengue hemorrhagic fever: implications for pathogenesis.* Virology, 2003. **305**(2): p. 330-8.
90. Rosen, L., M.T. Drouet, and V. Deubel, *Detection of dengue virus RNA by reverse transcription-polymerase chain reaction in the liver and lymphoid organs but not in the brain in fatal human infection.* Am J Trop Med Hyg, 1999. **61**(5): p. 720-4.
91. Jessie, K., et al., *Localization of dengue virus in naturally infected human tissues, by immunohistochemistry and in situ hybridization.* J Infect Dis, 2004. **189**(8): p. 1411-8.
92. Holmes, E.C., *The evolutionary biology of dengue virus.* Novartis Found Symp, 2006. **277**: p. 177-87; discussion 187-92, 251-3.
93. Thein, S., et al., *Risk factors in dengue shock syndrome.* Am J Trop Med Hyg, 1997. **56**(5): p. 566-72.
94. Guzman, M.G., et al., *Enhanced severity of secondary dengue-2 infections: death rates in 1981 and 1997 Cuban outbreaks.* Rev Panam Salud Publica, 2002. **11**(4): p. 223-7.
95. Balmaseda, A., et al., *Serotype-specific differences in clinical manifestations of dengue.* Am J Trop Med Hyg, 2006. **74**(3): p. 449-56.
96. Graham, R.R., et al., *A prospective seroepidemiologic study on dengue in children four to nine years of age in Yogyakarta, Indonesia I. studies in 1995-1996.* Am J Trop Med Hyg, 1999. **61**(3): p. 412-9.
97. Harris, V.K., et al., *Unusual association of Kikuchi's disease and dengue virus infection evolving into systemic lupus erythematosus.* J Indian Med Assoc, 2000. **98**(7): p. 391-3.
98. Messer, W.B., et al., *Epidemiology of dengue in Sri Lanka before and after the emergence of epidemic dengue hemorrhagic fever.* Am J Trop Med Hyg, 2002. **66**(6): p. 765-73.
99. Nisalak, A., et al., *Serotype-specific dengue virus circulation and dengue disease in Bangkok, Thailand from 1973 to 1999.* Am J Trop Med Hyg, 2003. **68**(2): p. 191-202.
100. Watts, D.M., et al., *Failure of secondary infection with American genotype dengue 2 to cause dengue haemorrhagic fever.* Lancet, 1999. **354**(9188): p. 1431-4.
101. Kouri, G.P., M.G. Guzman, and J.R. Bravo, *Why dengue haemorrhagic fever in Cuba? 2. An integral analysis.* Trans R Soc Trop Med Hyg, 1987. **81**(5): p. 821-3.
102. Bravo, J.R., M.G. Guzman, and G.P. Kouri, *Why dengue haemorrhagic fever in Cuba? 1. Individual risk factors for dengue haemorrhagic fever/dengue shock syndrome (DHF/DSS).* Trans R Soc Trop Med Hyg, 1987. **81**(5): p. 816-20.
103. Takhampunya, R., et al., *Phenotypic analysis of dengue virus isolates associated with dengue fever and dengue hemorrhagic fever for cellular attachment, replication and interferon signaling ability.* Virus Res, 2009. **145**(1): p. 31-8.
104. Cummings, D.A., et al., *Dynamic effects of antibody-dependent enhancement on the fitness of viruses.* Proc Natl Acad Sci U S A, 2005.
105. Twiddy, S.S., et al., *Phylogenetic relationships and differential selection pressures among genotypes of dengue-2 virus.* Virology, 2002. **298**(1): p. 63-72.
106. Halstead, S.B., et al., *Dengue hemorrhagic Fever in infants: research opportunities ignored.* Emerg Infect Dis, 2002. **8**(12): p. 1474-9.
107. Halstead, S.B. and E.J. O'Rourke, *Antibody-enhanced dengue virus infection in primate leukocytes.* Nature, 1977. **265**(5596): p. 739-41.

108. Halstead, S.B., J.S. Porterfield, and E.J. O'Rourke, *Enhancement of dengue virus infection in monocytes by flavivirus antisera.* Am J Trop Med Hyg, 1980. **29**(4): p. 638-42.
109. Goncalvez, A.P., et al., *Monoclonal antibody-mediated enhancement of dengue virus infection in vitro and in vivo and strategies for prevention.* Proc Natl Acad Sci U S A, 2007. **104**(22): p. 9422-7.
110. Halstead, S.B., *In vivo enhancement of dengue virus infection in rhesus monkeys by passively transferred antibody.* J Infect Dis, 1979. **140**(4): p. 527-33.
111. Kliks, S.C., et al., *Antibody-dependent enhancement of dengue virus growth in human monocytes as a risk factor for dengue hemorrhagic fever.* Am J Trop Med Hyg, 1989. **40**(4): p. 444-51.
112. Endy, T.P., et al., *Relationship of preexisting dengue virus (DV) neutralizing antibody levels to viremia and severity of disease in a prospective cohort study of DV infection in Thailand.* J Infect Dis, 2004. **189**(6): p. 990-1000.
113. Kliks, S.C., et al., *Evidence that maternal dengue antibodies are important in the development of dengue hemorrhagic fever in infants.* Am J Trop Med Hyg, 1988. **38**(2): p. 411-9.
114. Chau, T.N., et al., *Dengue in Vietnamese infants--results of infection-enhancement assays correlate with age-related disease epidemiology, and cellular immune responses correlate with disease severity.* J Infect Dis, 2008. **198**(4): p. 516-24.
115. Casanova, J.L. and L. Abel, *Human genetics of infectious diseases: a unified theory.* Embo J, 2007. **26**(4): p. 915-22.
116. Ge, D., et al., *Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance.* Nature, 2009. **461**(7262): p. 399-401.
117. Thomas, D.L., et al., *Genetic variation in IL28B and spontaneous clearance of hepatitis C virus.* Nature, 2009. **461**(7265): p. 798-801.
118. Tanaka, Y., et al., *Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C.* Nat Genet, 2009. **41**(10): p. 1105-9.
119. Guo, X., et al., *Strong influence of human leukocyte antigen (HLA)-DP gene variants on development of persistent chronic hepatitis B virus carriers in the Han Chinese population.* Hepatology. **53**(2): p. 422-8.
120. Guzman, M.G. and G. Kouri, *Dengue haemorrhagic fever integral hypothesis: confirming observations, 1987-2007.* Trans R Soc Trop Med Hyg, 2008. **102**(6): p. 522-3.
121. Guzman, M.G., et al., *Fatal dengue hemorrhagic fever in Cuba, 1997.* Int J Infect Dis, 1999. **3**(3): p. 130-5.
122. Guzman, M.G., *Global voices of science. Deciphering dengue: the Cuban experience.* Science, 2005. **309**(5740): p. 1495-7.
123. Sierra, B., et al., *HLA-A, -B, -C, and -DRB1 allele frequencies in Cuban individuals with antecedents of dengue 2 disease: advantages of the Cuban population for HLA studies of dengue virus infection.* Hum Immunol, 2007. **68**(6): p. 531-40.
124. Gonzalez, D., et al., *Classical dengue hemorrhagic fever resulting from two dengue infections spaced 20 years or more apart: Havana, Dengue 3 epidemic, 2001-2002.* Int J Infect Dis, 2005. **9**(5): p. 280-5.
125. Saluzzo, J.F., et al., *[Dengue 2 in eastern Senegal: serologic survey in simian and human populations. 1974-85].* Bull Soc Pathol Exot Filiales, 1986. **79**(3): p. 313-22.
126. Halstead, S.B., et al., *Haiti: absence of dengue hemorrhagic fever despite hyperendemic dengue virus transmission.* Am J Trop Med Hyg, 2001. **65**(3): p. 180-3.
127. Sharp, T.W., et al., *Dengue fever in U.S. troops during Operation Restore Hope, Somalia, 1992-1993.* Am J Trop Med Hyg, 1995. **53**(1): p. 89-94.
128. Gubler, D.J., et al., *Dengue 3 virus transmission in Africa.* Am J Trop Med Hyg, 1986. **35**(6): p. 1280-4.
129. Paradoa Perez, M.L., Y. Trujillo, and P. Basanta, *Association of dengue hemorrhagic fever with the HLA system.* Haematologia (Budap), 1987. **20**(2): p. 83-7.

130. Chiewsilp, P., R.M. Scott, and N. Bhamarapravati, *Histocompatibility antigens and dengue hemorrhagic fever.* Am J Trop Med Hyg, 1981. **30**(5): p. 1100-5.
131. Loke, H., et al., *Strong HLA class I--restricted T cell responses in dengue hemorrhagic fever: a double-edged sword?* J Infect Dis, 2001. **184**(11): p. 1369-73.
132. Stephens, H.A., et al., *HLA-A and -B allele associations with secondary dengue virus infections correlate with disease severity and the infecting viral serotype in ethnic Thais.* Tissue Antigens, 2002. **60**(4): p. 309-18.
133. Guzman, M.G. and G. Kouri, *Dengue and dengue hemorrhagic fever in the Americas: lessons and challenges.* J Clin Virol, 2003. **27**(1): p. 1-13.
134. Cooke, E.L., et al., *Functional analysis of the interleukin-1-receptor-associated kinase (IRAK-1) in interleukin-1 beta-stimulated nuclear factor kappa B (NF-kappa B) pathway activation: IRAK-1 associates with the NF-kappa B essential modulator (NEMO) upon receptor stimulation.* Biochem J, 2001. **359**(Pt 2): p. 403-10.
135. Fernandez-Mestre, M.T., et al., *TNF-alpha-308A allele, a possible severity risk factor of hemorrhagic manifestation in dengue fever patients.* Tissue Antigens, 2004. **64**(4): p. 469-72.
136. Loke, H., et al., *Susceptibility to dengue hemorrhagic fever in vietnam: evidence of an association with variation in the vitamin d receptor and Fc gamma receptor IIa genes.* Am J Trop Med Hyg, 2002. **67**(1): p. 102-6.
137. Sakuntabhai, A., et al., *A variant in the CD209 promoter is associated with severity of dengue disease.* Nat Genet, 2005. **37**(5): p. 507-13.
138. Soundravally, R. and S.L. Hoti, *Significance of transporter associated with antigen processing 2 (TAP2) gene polymorphisms in susceptibility to dengue viral infection.* J Clin Immunol, 2008. **28**(3): p. 256-62.
139. Liu, M., et al., *Early days: genomics and human responses to infection.* Curr Opin Microbiol, 2006. **9**(3): p. 312-9.
140. Chaussabel, D., V. Pascual, and J. Banchereau, *Assessing the human immune system through blood transcriptomics.* BMC Biol. **8**: p. 84.
141. Whitney, A.R., et al., *Individuality and variation in gene expression patterns in human blood.* Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1896-901.
142. Ubol, S., et al., *Differences in global gene expression in peripheral blood mononuclear cells indicate a significant role of the innate responses in progression of dengue fever but not dengue hemorrhagic fever.* J Infect Dis, 2008. **197**(10): p. 1459-67.
143. Nascimento, E.J., et al., *Gene expression profiling during early acute febrile stage of dengue infection can predict the disease outcome.* PLoS One, 2009. **4**(11): p. e7892.
144. Fink, J., et al., *Host gene expression profiling of dengue virus infection in cell lines and patients.* PLoS Negl Trop Dis, 2007. **1**(2): p. e86.
145. Hibberd, M.L., et al., *A genomics approach to understanding host response during dengue infection.* Novartis Found Symp, 2006. **277**: p. 206-14; discussion 214-7, 251-3.
146. R.Boom, C.J.A.S., M.M.M. Salimans, C.L.Jansen, P.M.E.Wertheim-Van Dillen and J. Van der Noordaa, *Rapid and simple method for purification of nucleic acids.* J. Clin. Microbiol, 1990: p. 495-503.
147. Laue, T., P. Emmerich, and H. Schmitz, *Detection of dengue virus RNA in patients after primary or secondary dengue infection by using the TaqMan automated amplification system.* J Clin Microbiol, 1999. **37**(8): p. 2543-7.
148. Cameron P. Simmons, S.P., Christiane Dolocek, Tran Nguyen Bich Chau, Michael Griffiths, Nguyen Thi Phuong Dung, Truong Hoang Long, Dang Minh Hoang, Nguyen Vinh Chau, Le Thi Thu Thao, Tran Tinh Hien, David A. Relman, Jeremy Farrar, *Patterns of host genome-wide gene transcript abundance in the peripheral blood of patients with acute dengue haemorrhagic fever.* Journal of Infectious Disease, 2007. **IN PRESS**.
149. Burke, D.S., et al., *A prospective study of dengue infections in Bangkok.* Am J Trop Med Hyg, 1988. **38**(1): p. 172-80.

150. Sangkawibha, N., et al., *Risk factors in dengue shock syndrome: a prospective epidemiologic study in Rayong, Thailand. I. The 1980 outbreak.* Am J Epidemiol, 1984. **120**(5): p. 653-69.
151. Mangada, M.M., et al., *Dengue-specific T cell responses in peripheral blood mononuclear cells obtained prior to secondary dengue virus infections in Thai schoolchildren.* J Infect Dis, 2002. **185**(12): p. 1697-703.
152. Green, S., et al., *Early immune activation in acute dengue illness is related to development of plasma leakage and disease severity.* J Infect Dis, 1999. **179**(4): p. 755-62.
153. Juffrie, M., et al., *Inflammatory mediators in dengue virus infection in children: interleukin-6 and its relation to C-reactive protein and secretory phospholipase A2.* Am J Trop Med Hyg, 2001. **65**(1): p. 70-5.
154. Shresta, S., et al., *Interferon-dependent immunity is essential for resistance to primary dengue virus infection in mice, whereas T- and B-cell-dependent immunity are less critical.* J Virol, 2004. **78**(6): p. 2701-10.
155. Diamond, M.S., et al., *Modulation of Dengue virus infection in human cells by alpha, beta, and gamma interferons.* J Virol, 2000. **74**(11): p. 4957-66.
156. Jones, M., et al., *Dengue virus inhibits alpha interferon signaling by reducing STAT2 expression.* J Virol, 2005. **79**(9): p. 5414-20.
157. Munoz-Jordan, J.L., et al., *Inhibition of alpha/beta interferon signaling by the NS4B protein of flaviviruses.* J Virol, 2005. **79**(13): p. 8004-13.
158. Mahalingam, S. and B.A. Lidbury, *Suppression of lipopolysaccharide-induced antiviral transcription factor (STAT-1 and NF-kappa B) complexes by antibody-dependent enhancement of macrophage infection by Ross River virus.* Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13819-24.
159. Sariol, C.A., et al., *Detection and genetic relationship of dengue virus sequences in seventeen-year-old paraffin-embedded samples from Cuba.* Am J Trop Med Hyg, 1999. **61**(6): p. 994-1000.
160. Kurane, I., et al., *High levels of interferon alpha in the sera of children with dengue virus infection.* Am J Trop Med Hyg, 1993. **48**(2): p. 222-9.
161. Panpanich, R., P. Sornchai, and K. Kanjanaratanakorn, *Corticosteroids for treating dengue shock syndrome.* Cochrane Database Syst Rev, 2006. **3**: p. CD003488.
162. PDVI, M.B.-P.D.V.I., *Global burden of Dengue.* 2009.
163. Bethell, D.B., et al., *Noninvasive measurement of microvascular leakage in patients with dengue hemorrhagic fever.* Clin Infect Dis, 2001. **32**(2): p. 243-53.
164. Cologna, R., P.M. Armstrong, and R. Rico-Hesse, *Selection for virulent dengue viruses occurs in humans and mosquitoes.* J Virol, 2005. **79**(2): p. 853-9.
165. Messer, W.B., et al., *Emergence and global spread of a dengue serotype 3, subtype III virus.* Emerg Infect Dis, 2003. **9**(7): p. 800-9.
166. Endy, T.P., et al., *Epidemiology of inapparent and symptomatic acute dengue virus infection: a prospective study of primary school children in Kamphaeng Phet, Thailand.* Am J Epidemiol, 2002. **156**(1): p. 40-51.
167. Anderson, K.B., et al., *Burden of symptomatic dengue infection in children at primary school in Thailand: a prospective study.* Lancet, 2007. **369**(9571): p. 1452-9.
168. Lynn, D.J., et al., *InnateDB: facilitating systems-level analyses of the mammalian innate immune response.* Mol Syst Biol, 2008. **4**: p. 218.
169. Kanehisa, M., et al., *KEGG for linking genomes to life and the environment.* Nucleic Acids Res, 2007. **36**: p. D480-4.
170. Joshi-Tope, G., et al., *Reactome: a knowledgebase of biological pathways.* Nucleic Acids Res, 2005. **33**(Database issue): p. D428-32.
171. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.* Journal of the Royal Statistical Society. Series B, 1995. **57**(1): p. 289-300.

172. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.
173. Samarajiwa, S.A., et al., *INTERFEROME: the database of interferon regulated genes.* Nucleic Acids Res, 2008.
174. Lin, C.Y., et al., *Hubba: hub objects analyzer—a framework of interactome hubs identification for network biology.* Nucleic Acids Res, 2008. **36**(2): p. W438-W443.
175. Yu, H., et al., *The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics.* PLoS Comput Biol, 2007. **3**(4): p. e59.
176. Long, H.T., et al., *Patterns of Gene Transcript Abundance in the Blood of Children with Severe or Uncomplicated Dengue Highlight Differences in Disease Evolution and Host Response to Dengue Virus Infection.* J Infect Dis, 2009.
177. Anderson, J.R. and R. Rico-Hesse, *Aedes aegypti vectorial capacity is determined by the infecting genotype of dengue virus.* Am J Trop Med Hyg, 2006. **75**(5): p. 886-92.
178. Chang, Y.J., M.J. Holtzman, and C.C. Chen, *Differential role of Janus family kinases (JAKs) in interferon-gamma-induced lung epithelial ICAM-1 expression: involving protein interactions between JAKs, phospholipase Cgamma, c-Src, and STAT1.* Mol Pharmacol, 2004. **65**(3): p. 589-98.
179. Somani, A.K., et al., *Src kinase activity is regulated by the SHP-1 protein-tyrosine phosphatase.* J Biol Chem, 1997. **272**(34): p. 21113-9.
180. An, H., et al., *Phosphatase SHP-1 promotes TLR- and RIG-I-activated production of type I interferon by inhibiting the kinase IRAK1.* Nat Immunol, 2008. **9**(5): p. 542-50.
181. Gough, D.J., et al., *IFNgamma signaling-does it mean JAK-STAT?* Cytokine Growth Factor Rev, 2008. **19**(5-6): p. 383-94.
182. Minoo, P., et al., *A novel SHP-1/Grb2-dependent mechanism of negative regulation of cytokine-receptor signaling: contribution of SHP-1 C-terminal tyrosines in cytokine signaling.* Blood, 2004. **103**(4): p. 1398-407.
183. Boekhoudt, G.H., M.R. Frazier-Jessen, and G.M. Feldman, *Immune complexes suppress IFN-gamma signaling by activation of the FcgammaRI pathway.* J Leukoc Biol, 2007. **81**(4): p. 1086-92.
184. Li, L., et al., *TRIP6 is a RIP2-associated common signaling component of multiple NF-kappaB activation pathways.* J Cell Sci, 2005. **118**(Pt 3): p. 555-63.
185. Lai, Y.J., W.C. Lin, and F.T. Lin, *PTPL1/FAP-1 negatively regulates TRIP6 function in lysophosphatidic acid-induced cell migration.* J Biol Chem, 2007. **282**(33): p. 24381-7.
186. Shresta, S., et al., *Critical roles for both STAT1-dependent and STAT1-independent pathways in the control of primary dengue virus infection in mice.* J Immunol, 2005. **175**(6): p. 3946-54.
187. Souza-Neto, J.A., S. Sim, and G. Dimopoulos, *An evolutionary conserved function of the JAK-STAT pathway in anti-dengue defense.* Proc Natl Acad Sci U S A, 2009. **106**(42): p. 17841-6.
188. Bethell, D.B., et al., *Pathophysiologic and prognostic role of cytokines in dengue hemorrhagic fever.* J Infect Dis, 1998. **177**(3): p. 778-82.
189. Hermant, B., et al., *Identification of proteases involved in the proteolysis of vascular endothelium cadherin during neutrophil transmigration.* J Biol Chem, 2003. **278**(16): p. 14002-12.
190. Juffrie, M., et al., *Inflammatory mediators in dengue virus infection in children: interleukin-8 and its relationship to neutrophil degranulation.* Infect Immun, 2000. **68**(2): p. 702-7.
191. Hartshorn, K.L., et al., *Innate defense against influenza A virus: activity of human neutrophil defensins and interactions of defensins with surfactant protein D.* J Immunol, 2006. **176**(11): p. 6962-72.
192. Furci, L., et al., *Alpha-defensins block the early steps of HIV-1 infection: interference with the binding of gp120 to CD4.* Blood, 2007. **109**(7): p. 2928-35.
193. Hazrati, E., et al., *Human alpha- and beta-defensins block multiple steps in herpes simplex virus infection.* J Immunol, 2006. **177**(12): p. 8658-66.

194. LaFleur, C., et al., *HLA-DR antigen frequencies in Mexican patients with dengue virus infection: HLA-DR4 as a possible genetic resistance factor for dengue hemorrhagic fever.* Hum Immunol, 2002. **63**(11): p. 1039-44.
195. *Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.* Nature, 2007. **447**(7145): p. 661-78.
196. Spencer, C.C., et al., *Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip.* PLoS Genet, 2009. **5**(5): p. e1000477.
197. Velez, D.R., et al., *Association of SLC11A1 with tuberculosis and interactions with NOS2A and TLR2 in African-Americans and Caucasians.* Int J Tuberc Lung Dis, 2009. **13**(9): p. 1068-76.
198. Thuong, N.T., et al., *Identification of tuberculosis susceptibility genes with human macrophage gene expression profiles.* PLoS Pathog, 2008. **4**(12): p. e1000229.
199. Pelak, K., et al., *Host determinants of HIV-1 control in African Americans.* J Infect Dis. **201**(8): p. 1141-9.
200. Herbeck, J.T., et al., *Multistage genomewide association study identifies a locus at 1q41 associated with rate of HIV-1 disease progression to clinical AIDS.* J Infect Dis. **201**(4): p. 618-26.
201. van Manen, D., et al., *Association of HLA-C and HCP5 gene regions with the clinical course of HIV-1 infection.* Aids, 2009. **23**(1): p. 19-28.
202. Hunt, P.W. and M. Carrington, *Host genetic determinants of HIV pathogenesis: an immunologic perspective.* Curr Opin HIV AIDS, 2008. **3**(3): p. 342-8.
203. Jallow, M., et al., *Genome-wide and fine-resolution association analysis of malaria in West Africa.* Nat Genet, 2009.
204. De Val, S. and B.L. Black, *Transcriptional control of endothelial cell development.* Dev Cell, 2009. **16**(2): p. 180-95.
205. Atkins, G.B. and M.K. Jain, *Role of Kruppel-like transcription factors in endothelial biology.* Circ Res, 2007. **100**(12): p. 1686-95.
206. Tsai, H.M., *Mechanisms of microvascular thrombosis in thrombotic thrombocytopenic purpura.* Kidney Int Suppl, 2009(112): p. S11-4.
207. Outram, S.V., et al., *KLF13 influences multiple stages of both B and T cell development.* Cell Cycle, 2008. **7**(13): p. 2047-55.
208. Chaturvedi, U.C., et al., *Cytokine cascade in dengue hemorrhagic fever: implications for pathogenesis.* FEMS Immunol Med Microbiol, 2000. **28**(3): p. 183-8.
209. Kurane, I., *Dengue hemorrhagic fever with special emphasis on immunopathogenesis.* Comp Immunol Microbiol Infect Dis, 2007. **30**(5-6): p. 329-40.
210. Leong, A.S., et al., *The pathology of dengue hemorrhagic fever.* Semin Diagn Pathol, 2007. **24**(4): p. 227-36.
211. Pratt, W.B., *The role of the hsp90-based chaperone system in signal transduction by nuclear receptors and receptors signaling via MAP kinase.* Annu Rev Pharmacol Toxicol, 1997. **37**: p. 297-326.
212. Vadas, P., et al., *Extracellular phospholipase A2 expression and inflammation: the relationship with associated disease states.* J Lipid Mediat, 1993. **8**(1): p. 1-30.
213. Tukaj, S., et al., *Hsp40 proteins modulate humoral and cellular immune response in rheumatoid arthritis patients.* Cell Stress Chaperones.
214. Luna, C., et al., *Role of miR-29b on the regulation of the extracellular matrix in human trabecular meshwork cells under chronic oxidative stress.* Mol Vis, 2009. **15**: p. 2488-97.
215. Fleming, J.S., et al., *Expression of ovarian tumour suppressor OPCML in the female CD-1 mouse reproductive tract.* Reproduction, 2009. **137**(4): p. 721-6.
216. Onlamoon, N., et al., *Dengue virus-induced hemorrhage in a nonhuman primate model.* Blood. **115**(9): p. 1823-34.
217. Yauch, L.E. and S. Shresta, *Mouse models of dengue virus infection and disease.* Antiviral Res, 2008. **80**(2): p. 87-93.

218. Martin, J., et al., *Viremia and the magnitude of the immune response upon infection of green monkeys with dengue virus type 2 are strain-dependent.* Curr Microbiol, 2009. **59**(6): p. 579-83.
219. Ozsolak, F., et al., *Digital transcriptome profiling from attomole-level RNA samples.* Genome Res.
220. Lederman, M.M., et al., *Biology of CCR5 and its role in HIV infection and treatment.* Jama, 2006. **296**(7): p. 815-26.
221. Westby, M. and E. van der Ryst, *CCR5 antagonists: host-targeted antivirals for the treatment of HIV infection.* Antivir Chem Chemother, 2005. **16**(6): p. 339-54.
222. Klibanov, O.M., *Vicriviroc, a CCR5 receptor antagonist for the potential treatment of HIV infection.* Curr Opin Investig Drugs, 2009. **10**(8): p. 845-59.
223. Sayana, S. and H. Khanlou, *Maraviroc: a new CCR5 antagonist.* Expert Rev Anti Infect Ther, 2009. **7**(1): p. 9-19.