MARISA SOFIA SOUSA OLIVEIRA

GENOMIC AND FUNCTIONAL INSIGHTS ON DENGUE INFECTION: THE ROLE OF HOST ANCESTRY

II

O trabalho apresentado nesta tese foi realizado no Instituto de Patologia e Imunologia Molecular da Universidade do Porto (Ipatimup), no Instituto de Investigação e Inovação em Saúde da Universidade do Porto (i3S) e no Institut Pasteur Paris.

Preceitos Legais

De acordo com a alínea d) do artigo 3.º do regulamento Geral dos Terceiros Ciclos de estudos da Universidade do Porto, nesta dissertação foram utilizados resultados dos trabalhos publicados ou em preparação abaixo indicados. No cumprimento do disposto no referido Decreto-Lei, a autora desta tese declara que interveio na conceção e na execução do trabalho experimental, na interpretação e discussão dos resultados e na redação dos manuscritos publicados, sob o nome Oliveira M:

– **Oliveira M,** Lert-itthiporn W, Cavadas B, Fernandes V, Chuaunsumrit A, Anunciação O, Casademont I, Koeth F, Tangnararatchakit K, Khor CC, Paul R, Malasit P, Matsuda F, Simon-Loriere E, Suriyaphol P, Pereira L, Sakuntabhai A. (2018) Joint ancestry and association test indicate two distinct pathogenic pathways involved in classical dengue fever and dengue shock syndrome. PLoS Negl Trop Dis, 12(2): e0006202. https://doi.org/10.1371/journal.pntd.0006202

– **Oliveira M**, Saraiva DP, Cavadas B, Fernandes V, Pedro N, Casademont I, Koeth F, Alshamali F, Harich N, Cherni L, Sierra B, Guzman MG, Sakuntabhai A and Pereira L. Population genetics-informed meta-analysis in seven genes associated with risk to dengue fever disease. (submitted)

– **Oliveira M**, Penova M, Prot M, Casademont I, Jacob Y, Pereira L, Simon-Loriere E and Sakuntabhai A. Functional evaluation of GRIP1 and PPP2R5E involvement in dengue fever. (in preparation)

Para a minha mãe

"Vous savez, la chance, c'est comme le Tour de France.
On l'attend longtemps et puis ça passe vite.
Alors quand le moment vient,
il faut sauter la barrière sans hésiter."

Le Fabuleux Destin d'Amélie Poulain (2001)

x

Acknowledgements

Arábias Verónica e Nicole, ao Bruno, não só pela contribuição visível no meu projecto mas por toda a importante partilha de ideias (e ideais) ao longo dos anos e, por fim, à Joana "Pequena", dada a empatia (quase) imediata e todo o apoio e amizade espontâneos.

Às meninas Ana Lu, Andreia S, Catarina, Joana, Marina, Natália, Patrícia, Sílvia, Sofia e Susana e também ao Miguel por toda a amizade, compreensão e companheirismo e em particular à Andreia B, que também tem sido a minha consciência em toda esta etapa.

Aos TugasPasteur pelo grupo incrível que se tem formado. Um obrigada muito especial por todo o carinho à Alexa, Alicia, Carolina, Filipe, Flávia, Freire, Jorge P, Jorge S e Marta...*We'll always have Euro2016*.

To all the great people I met in Paris and had the chance to know better through this time and mostly to Christiane, Erica, Coco, Antoine, Jovana, and Madhura for being always present and not let me down. E mais ainda à Sofia, companhia dos dias soalheiros.

A todos os meus amigos que já me acompanhavam antes desta aventura e têm sido presença constante. À Susana, Sara e Catarina pelos mais de 10 anos de amizade, ao SPAM, principalmente João e Filipa, que junta forças em todas as minhas visitas ao Porto, à Ana e Alex que são um pilar importante deste o mestrado e por todas as recordações e bons momentos tanto no Porto como em Paris. E claro, à Catarina S, que tem sido uma amiga excepcional e a quem admiro verdadeiramente a perserverança e força, entre tantas qualidades, e à Mónica, pelo afeto, altos e baixos, paciência, honestidade e tudo o mais que vem de uma amizade que já não se conta pelos dedos das mãos.

Um enorme e sentido obrigada ao Professor Zé Carlos que, mesmo não estando fisicamente presente, me continuará a inspirar ao longo da minha vida.

À minha família por todo o amor e carinho, ao Jorge por me mostrar que é possivel cuidar sem haver laços de sangue e em especial aos meus avós que lutaram para que eu aqui estivesse hoje e são sem dúvida um grande motivo de orgulho e respeito.

À minha mãe por tudo simplesmente... por lhe estar grata pelo que fez por mim, ser das pessoas mais correctas que conheço e ser excepcional sem facilitismos. E por isso, a ela dedico este trabalho.

Abstract

Dengue is the most widespread arthropod-borne disease, with two-thirds of the worldwide population living in risk areas for infection. Despite the major health and economic burden, the molecular mechanisms resulting in severe disease remain poorly understood. While dengue usually emerges in epidemic outbreaks in Latin America, it is endemic in Southeast Asia and Pacific regions where the four serotypes are present, causing recurrent outbreaks in short time frames. Dengue shock syndrome represents a major threat for early ages in Southeast Asia since most of the infections occur during childhood. To better understand the genetic factors underlying Southeast Asian susceptibility to dengue, in the present work we explored the human genetics influence on dengue outcomes combining three distinct strategies:

(1) Replication, by performing a population-genetics informed meta-analysis in seven genes known to be involved in dengue distinct phenotypes, *MICB*, *PLCE1*, *CD32*, *CD209*, *TNFA*, *OAS1* and *OAS3*. We genotyped and collected data from the literature for control/reference worldwide populations and dengue case-control cohorts (the three new cohorts were from Cambodia, Vietnam and Brazil) to refine the analysis;

(2) Discovery, by investigating the ancestry role in Southeast Asia through a coupled association-admixture analysis in new Thai cohorts exploring two phenotypes, life-threatening dengue shock syndrome (DSS) and classical dengue fever (DF). We aimed to unravel unique signatures and assess worldwide genetic risk for both outcomes;

(3) Validation, by performing functional evaluation of the antiviral activity of the most promising newly discovered genes. We established *in vitro* models for gene knockout and overexpression and conducted infection assays with DENV1.

The population-informed meta-analysis on the seven genes disclosed that risk or protection alleles presented intermediate to high minor allele frequencies across worldwide populations, and a high discriminatory power among the distinct population groups, although there was no evidence of selective pressure acting upon them. These alleles were specifically associated with a particular dengue phenotype: *TNFA*-rs1800629-A with DF (odds ratio (OR)= 0.67; p-value=0.0092), *CD32*-rs1801274-G with DHF (OR=0.82; p=0.038); *OAS3*-rs2285933-G (OR=0.55; p=0.016), *PLCE1*-rs2274223-G (OR=0.80; p=1.93e^{-10}) and *MICB*-rs3132468-C (OR=1.32; p=1.19e^{-12}) with DSS.

Five candidate genes were identified in the association-admixture genome-wide approach applied to Thai cohorts (252 DF, 159 DSS and 290 control). These genes are

involved in two independent pathogenic pathways: *PLCB4* (OR=0.58, p=1.3e10^{-2}), from phospholipase C family, was associated with susceptibility to DSS, possibly through involvement in blood vessels inflammation; and four genes, *AHRR* (OR=0.54; p=1.2e^{-4}), *CHST10* (OR=0.59; p=5.2e^{-4}), *GRIP1* (OR=0.56; p=1.6e^{-3}) and *PPP2R5E* (belonging to PP2A-B56 subunit; OR=0.62; p=2.0e^{-4}), in the xenobiotic metabolism pathway, were associated with DF, possible through link to viral proteins and immune system activation.

The immunofluorescence assays conducted to investigate the possible link between PPP2R5E-NS5 (non-structural protein 5) proteins at the *in silico* inferred LxxIxE motif, confirmed co-localization in a time-dependent manner with NS5 from DENV1 and DENV2. The only difference between the two strains was the entrance of both PPP2R5E and DENV2-NS5 into the host nucleus, with DENV2-NS5 remaining there while PPP2R5E returned to cytoplasm. PPP2R5E-NS5 from DENV1 always localized in the cytoplasm. The removal of the LxxIxE motif from NS5 led to no co-localization of PPP2R5E with DENV2-NS5. We successfully established *in vitro* models to further pursue functional experiments on *GRIP1* and *PPP2R5E* genes, using gene knockout by CRISPR-Cas9 and overexpression by constitutive mammalian expression. Infection tests with DENV1 in these transformed cells led to confirm that *GRIP1* and *PPP2R5E* overexpressions translated in considerable antiviral effects.

The worldwide genetic risks calculated for known and newly found genes independently pinpointed Africans as the most protected against DSS, contrasting with European and Asian populations that presented the highest predicted risk to develop severe dengue. These findings highlight the heterogeneity observed on human response to dengue infections and how ancestry is determinant in the risk of developing severe disease. Our findings of new ancestry-associated candidate genes, with proved functional impact, support research on personalized or at least population group based treatments to diminish dengue burden.

Resumo

A dengue é a doença transmitida por artrópodes mais dispersa a nível mundial, com dois terços da população a viver em áreas com risco de infecção. Apesar do grande impacto na economia e saúde, os mecanismos moleculares que levam à forma severa da doença ainda são pouco conhecidos. Enquanto que a dengue emerge mais frequentemente em surtos epidémicos na América Latina, esta é endémica nas regiões do Sudeste Asiático e Pacífico, onde são detectados os quatro serótipos, causando surtos recorrentes em curtos períodos de tempo. A síndrome do choque da dengue (DSS) representa uma grande ameaça no sudeste Asiático durante a infância, uma vez que a maioria das infecções ocorrem nessa faixa etária. Para compreender melhor os factores genéticos subjacentes à susceptibilidade à dengue no sudeste Asiático, no presente trabalho exploramos a influência da genética humana sobre o devenvolvimento da dengue, combinando três estratégias:

(1) Replicação através de uma meta-análise em sete genes associados a fenótipos distintos da dengue, *MICB*, *PLCE1*, *CD32*, *CD209*, *TNFA*, *OAS1* e *OAS3*. Procedeu-se à genotipagem e à recolha de dados da literatura para populações controlo/referência e coortes de indivíduos infectados com dengue (as três novas coortes eram do Cambodja, Vietname e Brasil) para refinar a análise;

(2) Descoberta de novos candidatos ligados ao papel da ancestralidade no Sudeste Asiático, aplicando a análise combinada de associação e mistura populacional em novos coortes tailandeses e explorando dois fenótipos, DSS, com risco de vida, e dengue clássica (DF). Os nossos objetivos consistiam em desvendar assinaturas únicas e avaliar o risco genético mundial para ambos os fenótipos;

(3) Validação através da avaliação funcional da atividade antiviral dos genes recém-descobertos mais promissores. Para tal, estabeleceram-se modelos *in vitro* para o *knockout* e sobre-expressão de genes e realizaram-se ensaios de infecção com DENV1.

A meta-análise informada pelos dados de genética populacional revelou que os alelos associados ao risco ou proteção apresentavam frequências alélicas intermédias a altas do alelo com frequência mínima e um alto poder discriminatório entre os distintos grupos populacionais embora, não tenha sido detectada evidência de pressão seletiva atuando sobre eles. Os alelos estudados estavam associados a fenótipos específicos da dengue: *TNFA*-rs1800629-A com DF (odds ratio (OR) = 0.67; p-value = 0.0092), *CD32*-rs1801274-G com DHF (OR = 0.82; p = 0.038) ; *OAS3*-rs2285933-G (OR =

0.55; p = 0.016), *PLCE1*–rs2274223–G (OR = 0.80; p = $1.93e^{-10}$) e *MICB*–rs3132468–C (OR = 1.32; p = $1.19e^{-12}$) com DSS.

A abordagem associação–ancestralidade combinada aplicada ao genoma dos grupos de casos e controlos tailandeses (252 DF, 159 DSS e 290 controlos) permitiu a identificação de cinco genes candidatos envolvidos em duas vias patogénicas independentes: *PLCB4* (OR = 0.58, p = $1.3e10^{-2}$), da família da fosfolipase C, associada à susceptibilidade à DSS, dado o potencial envolvimento na inflamação dos vasos sanguíneos; e quatro genes, *AHRR* (OR = 0.54; p = $1.2e^{-4}$), *CHST10* (OR = 0.59; p = $5.2e^{-4}$), *GRIP1* (OR = 0.56; p = $1.6e^{-3}$) e *PPP2R5E* (pertencente a subunidade PP2A-B56, OR = 0.62; p = $2.0e^{-4}$), da via metabólica dos xenobióticos, foram associados ao fenótipo DF, por possível interação com proteínas virais e ativação do sistema imune. Os ensaios de imunofluorescência realizados para investigar a possível interação entre as proteínas PPP2R5E–NS5 (proteína não estrutural 5) no local do motivo LxxIxE, inferido *in silico*, confirmou a co–localização com DENV1– e DENV2–NS5. A única diferença encontrada entre as proteínas das duas estirpes foi a entrada de PPP2R5E conjuntamente com a DENV2–NS5 no núcleo hospedeiro, com NS5 permanecendo lá enquanto o PPP2R5E retornava ao citoplasma. O complexo PPP2R5E–NS5 do DENV1 restringiu–se ao citoplasma. A remoção do motivo LxxIxE da NS5 levou a ausência de co–localização PPP2R5E–NS5. Conduzimos ainda testes funcionais com os genes *GRIP1* e *PPP2R5E*, tendo sido bem sucedidos no estabelecimento de modelos *in vitro* com *knockout* dos genes por CRISPR–Cas9 e sobreexpressão por expressão constitutiva. Os ensaios de infeção por DENV1 nas células transformadas revelaram que a sobreexpressão do *GRIP1* e *PPP2R5E* provocaram efeitos antivirais consideráveis.

Os riscos genéticos mundiais calculados para os genes conhecidos e os recém–descobertos indicaram independentemente os africanos como os mais protegidos contra DSS contrastando com as populações europeias e asiáticas que apresentaram o maior risco para dengue severa. Estas descobertas evidenciam a heterogeneidade observada na resposta humana à infeções por dengue e como a ancestralidade é determinante para a severidade da doença. As descobertas de novos genes candidatos associados à ancestralidade, com impacto funcional provado, justificam a investigação de tratamentos personalizados ou a nível do grupo populacional, para uma futura diminuição dos casos de dengue.

Table of Contents

List of Figures

Introduction

Paper I

Paper II

Paper III

Appendix A

Appendix B

Appendix C

List of Tables

Appendix C

Abbreviations

ADE	Antibody-dependent enhancement
AHR	Aryl Hydrocarbon Receptor
AHRR	Aryl Hydrocarbon Receptor Repressor
AIMs	ancestry informative markers
C protein	Capsid protein
CAR	Constitutive Androstane Receptor
Cas9	CRISPR associated protein 9
CD209	Cluster of Differentiation 209
cDNA	Complementary DNA
CEU	European ancestry from Utah
CHB	Han Chinese from Beijing
CHST10	Carbohydrate Sulfotransferase 10
CMV	Cytomegalovirus
CRISPR	Clustered regularly interspaced short palindromic repeats
CXCR4	C-X-C Motif Chemokine Receptor 4
CYD	Chimeric Yellow Fever Virus-DENV
DA	Dopaminergic
DAG1	Dystroglycan 1 gene
DARC	Duffy Blood Group, Chemokine Receptor
DC-SIGN	Dendritic cells-specific intercellular adhesion molecule-3-grabbing non-integrin
DDT	Dichlorodiphenyltrichloroethane
DENV	Dengue virus
DENV1-4	Dengue virus strain 1 to 4
DF	Dengue Fever
DHF	Dengue Hemorrhagic Fever
DNA	Deoxyribonucleic acid
DSB	Double strand break
DSS	Dengue Shock Syndrome
E protein	Envelop protein

eQTLs	Expression quantitative trait loci
ER	Endoplasmic reticulum
FCGRIIa	Low affinity immunoglobulin gamma Fc region receptor II a
F_{ST}	Fixation index
GRIP1	Glutamate receptor interacting protein
gRNA	Guide RNA
GSEA	Gene set enrichment analysis
GWAS	Genome-wide association studies
H	Amino acid histidine
HBB	β-globin
HDR	Homology-directed repair
HGDP	Human Genome Diversity Project
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
IFN-γ	Interferon gamma
IgG	Immunoglobulin G
iHS	integrated haplotype score
IL-10	Interleukin 10
IL-4	Interleukin 4
IL-6	Interleukin 6
Indels	Insertions/deletions
IRF3	Interferon regulatory factor 3
JPT	Japanese from Tokyo
Kb	Kilobase
LARGE	LARGE Xylosyl- And Glucuronyltransferase 1
LCT	Lactase
LD	Linkage disequilibrium
lncRNA	Long noncoding RNA
LXR	Liver X receptor
M protein	Membrane protein
MAF	Minor allele frequency

MCM6	Minichromosome Maintenance Complex Component 6
MHC	Major histocompatibility complex
MICB	MHC Class I Polypeptide-Related Sequence B
miRNA	MicroRNA
MOI	Multiplicity of infection
mRNA	Messenger RNA
NHEJ	Nonhomologous end-joining
NS	Non-structural protein
NS1	Non-structural protein 1
NS2A	Non-structural protein 2A
NS2B	Non-structural protein 2B
NS3	Non-structural protein 3
NS4A	Non-structural protein 4A
NS4B	Non-structural protein 4B
NS5	Non-structural protein 5
OAS	Oligoadenylate synthetase
OR	Odds ratio
ORP	OSBP-related protein
OSBPL10	Oxysterol Binding Protein Like 10
OTUB1	OTU Deubiquitinase, Ubiquitin Aldehyde Binding 1
PAHO	Pan American Health Organization
PCA	Principal component analysis
PLCB4	Phospholipase C Beta 4
PLCE1	Phospholipase C Epsilon 1
PP2A	Protein phosphatase 2A
PPAR	Peroxisome proliferator-activated receptor
PPP2R5E	Protein Phosphatase 2 Regulatory Subunit B'Epsilon
PrM	Pre-membrane protein
PVLR4	Poliovirus Receptor-Related Protein 4
R	Amino acid arginine
RdRp	RNA-dependent RNA polymerase

RHOA	Ras Homolog Family Member A
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNAse L	Endoribonuclease L
RNASeq	RNA Sequencing
RXRA	Retinoid X Receptor Alpha
shRNA	Short hairpin RNAs
siRNA	Small interfering RNAs
SNP	Single nucleotide polymorphism
TLR5	Toll Like Receptor 5
TNFA	Tumor Necrosis Factor alpha
TPST1	Tyrosylprotein Sulfotransferase 1
VCD	Virologically confirmed dengue
VDR	Vitamin D Receptor
WHO	World Health Organization
WHO/SEARO	World Health Organization, South-East Asia Regional Office
XP-EHH	Cross population extended haplotype homozygosity
YRI	Yoruba from Ibadan

I. General introduction

1. Human diversity

Human diversity has triggered the curiosity of researchers through centuries, who are trying to understand the traits that make our species distinctive. The diversity patterns observed in populations are variable and correlated with their past demographic events. Diversity is generated by mutation and rearranged in different backgrounds by recombination throughout generations. As mutations occur at random, time and effective size contribute to their accumulation. Thus, older and expanded populations will have higher genetic diversities than younger and constant (and bottlenecked) populations. Migration can contribute to increase the genetic diversity of the population receiving the migrants, while the population losing individuals will mostly lose genetic diversity. A new mutation will vary in frequency along time tending to disappear or to fixate. Most of the variations in allele frequencies will be governed by genetic drift (fluctuations due to random loss of individuals or no reproduction), but a rare proportion will be altered by selection if the genotype influences the phenotype (either positively, negatively or in a balancing way).

1.1. Population structure

Homo sapiens species appeared around 200 000 years ago in Africa (Cann et al., 1987). The most accepted model for the origin of modern humans is called "out-of-Africa" and states that the transition from early *Homo* species to *H. sapiens* occurred in Africa and, around 70 000 years ago, a small group of East Africans expanded to Southwest Asia/Arabian Peninsula and thence towards Europe, Asia and America (Figure 1) (Soares et al., 2012). There was a loss of heterozygosity associated with the sequential founder events in the world colonization, and the loss increases with the distance to East Africa. These bottlenecks have been identified when studying different genetic markers, such as single nucleotide polymorphisms (SNPs) and short tandem repeats (STRs) (Li et al., 2008; Underhill and Kivisild, 2007). Population structure is maintained by the greater difficulty in exchanging genes as geographic distance increases.
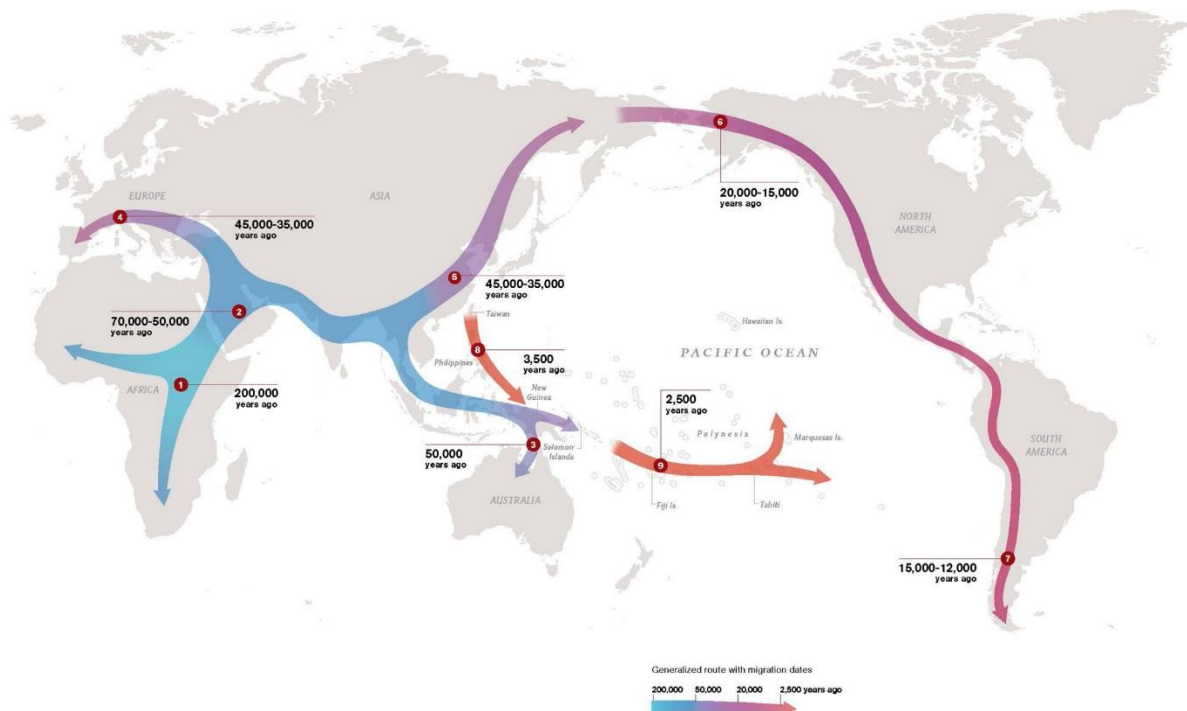
Figure 1. Out-of-Africa dispersal. Anatomically modern humans appeared in East Africa 200 000 years ago and started their dispersion through Arabian Peninsula, and from there to Europe, Asia and, at last, America. Adapted from https://www.nationalgeographic.org/

The event of out-of-Africa matched, most probably, the first noticeable expansion in the human population. But rapid expansions have been occurring especially since around 11,000 years ago (Atkinson et al., 2008), and the consequent extensive increment of the effective population size led to the emergence of a high proportion of new mutations since that date. In conformity, the analysis of the 1000 Genomes consortium (Altshuler et al., 2012) estimated that 17% of low-frequency (0.5–5%) variants were shared among populations belonging to the same ancestral group, while the rare variants (<0.5%) which occur within the populations amount to a proportion of 53%. Only high frequency variants (>10%) are observed in all worldwide populations. Considering their functional impact, the distribution of rare variants coding non-synonymous alterations, frequently deleterious (25–50%), varies with the ancestral population: the bottleneck in the origin of Europeans and East Asians renders them more prone to be homozygous for those variants in comparison with Africans (Altshuler et al., 2012; Fu et al., 2013).

1.2. Admixture of populations

A phenomenon that decreases the effects of population structure is migration, leading to the admixture of populations. Admixture brings together genetic diversity that emerged differently within population groups, since their divergence began in the out-of-Africa. Of course, when the migration event occurs, it leads to local structure in the population, which will decrease as interbreeding takes place.

Migration occurred throughout the world along time, and it is most noticeable in certain hotspots, as at the crossroads between continents (Figure 2B). This is notorious in Southwest Asia/Arabian Peninsula (Fernandes et al., 2012) where the three main human population backgrounds, sub-Saharan African, European and Asian, met. The encounter between these three ancestries occurred again, from the 15th century onwards, in America, where European and sub-Saharan African backgrounds were added to the Native American (of East Asian origin) ancestors, in variable proportions across the continent (Adhikari et al., 2017; Baharian et al., 2016). Europe displays some admixture with North African and sub-Saharan African ancestries in the south, in the Mediterranean region (Botigue et al., 2013). East Africa displays admixture with Southwest Asia/Arabian Peninsula background, and there are some signs of within-African migrations between different regions of the continent (Rito et al., 2013). East Asia also displays many signs of within-East Asia migrations (Abdulla et al., 2009), while the best examples of East Asian-African admixture are placed in Madagascar and Comoros Islands (Brucato et al., 2018; Pierron et al., 2017).
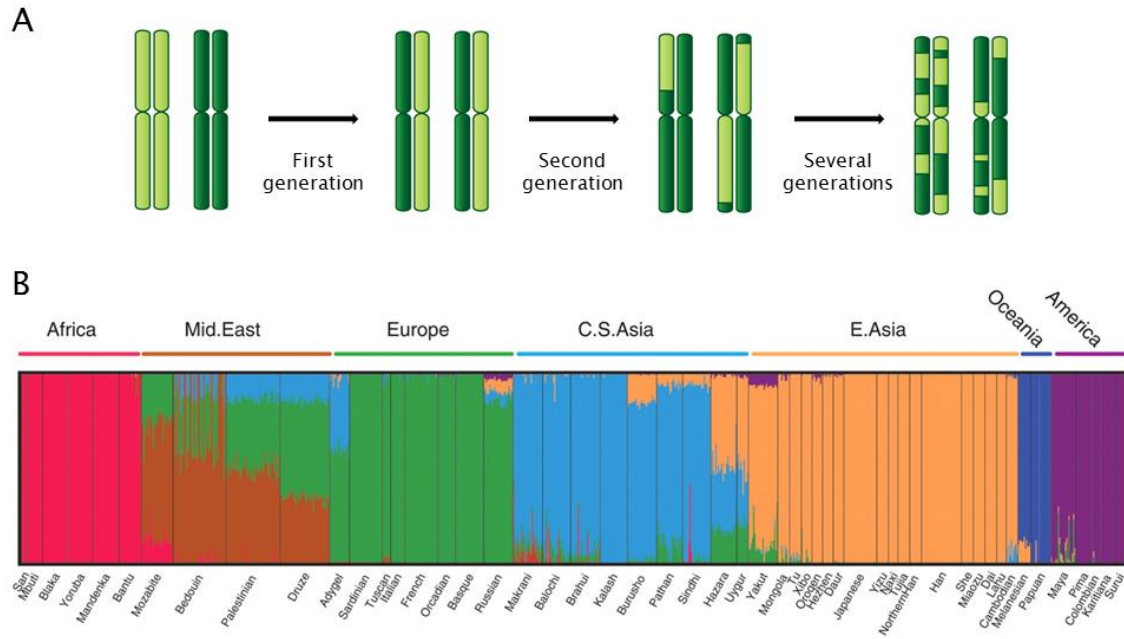
Figure 2. Admixture in human populations. A – Schematic representation of admixture along time in pairs of chromosomes of descendants from an admixture event between two parental ancestors. B – Admixture inference in worldwide representative populations, where each color represents one ancestral group and each vertical line an individual belonging to that population. Adapted from (Winkler et al., 2010) and (Li et al., 2008).

Genetically, admixture between ancestral groups is easier and securer to identify through haplotype diversity (Ardlie et al., 2002). Given the genome organization, the occurrence of mutations may be associated to other neighbor variants and transmitted together to future generations. This phenomenon is known as linkage disequilibrium (LD) and leads to a non-random loci distribution due to low recombination rate in close chromosome regions. Thus, LD tends to decay with the distance between the allele positions (McEvoy et al., 2011).

The extension of the LD blocks (Figure 2A) reflects the population age since younger populations present larger blocks than ancient ones (for example, African populations), due to the occurrence of less chromosomal recombination events during their history (McEvoy et al., 2011). The admixture of distinct ancestral populations freezes time to t=0, and the mixed ancestral blocks will decrease in size as recombination occurs since the admixture event (Nei and Li, 1973). Thus, the LD decay can give an accurate estimation of the time passed since the admixture event (Patterson et al., 2012).

1.3.   Adaptive pressures and selection

The *Homo sapiens* diaspora introduced the species to a wide range of climate variations, dietary shifts, pathogens and cultural changes that left an imprint in population diversity patterns (Akey et al., 2004). These genetic variants modelled by natural selection affect the population fitness and its reproductive success.

A mutation that confers phenotype advantage can be under the effect of positive selection. As consequence, its frequency will increase from generation to generation and eventually become fixed. A good example of positive selection is lactase persistence (Tishkoff et al., 2007). This derived trait is particularly frequent in Northern Europe where the mutation C/T-13910 in *MCM6* gene (near the *LCT* gene) reaches frequencies close to 100%. Other mutations in the same chromosome region were positively selected in pastoralist African populations, in a convergent adaptation (Tishkoff et al. 2007).

The opposite situation, where a mutation decreases the fitness, leads to purifying or negative selection, driving to the elimination of the deleterious polymorphism from the population. This is the most frequent type of selection in humans and can be identified in phylogenetic trees: a higher amount of non-synonymous mutations near the leaves than in older branches indicates that these deleterious mutations are eliminated along time (Kivisild et al., 2006).

Balancing selection occurs when heterozygosity introduces advantageous fitness and, therefore, favors the genetic diversification (Andrés et al., 2009). Malaria resistance is one of the best characterized cases of balancing selection. The allele Hb^s from β-globin gene (*HBB*) in homozygosity causes sickle-cell anemia, a condition that alters hemoglobin structure and results in the sickling form of red blood cells. This allele is observed in high frequency in areas where malaria is endemic and the heterozygous individuals are best protected against malaria than both homozygous for *HBB* (Hedrick, 2011). α-thalassemia is a group of diseases resulting from deletions in one of the two genes codifying α-globin. In resemblance with Hb^s allele, their distribution overlaps the malaria endemic areas and the heterozygous also seem to be protected against severe malaria (Williams et al., 2005).

Finding signatures of selection and adaptation in the human genome has been the focus of several genetic studies in recent years, when high-throughput databases begun to become available (Grossman et al., 2013). The most efficient algorithms to evaluate selection are haplotype-based (Figure 3), such as the integrated haplotype

41

score (iHS; (Voight et al., 2006)) and the cross population extended haplotype homozygosity (XP–EHH; (Sabeti et al., 2007)). iHS has good power to detect selective sweeps (consisting in the fixation of a beneficial mutation) at moderate frequency (50-80%), and XP–EHH is most powerful for selective sweeps above 80% frequency (Sabeti et al., 2007; Voight et al., 2006). The application (Grossman et al., 2013) of these algorithms to the 1000 Genomes database led to identify skin color, metabolism and infectious disease resistance as the main relevant selected traits in humans. The authors confirmed that many mutations fall in and around genes encoding the receptors or enzymes that modify the receptors for some of the most devastating pathogens in human history: *RHOA* and *OTUB1* (*Yersinia pestis*), *DAG1* (*Mycobacterium leprae*), *TLR5* (*Salmonella typhimurium* and others), *LARGE* (Lassa virus), *DARC* (*Plasmodium vivax* malaria), *PVRL4* (measles virus), *VDR* (*Mycobacterium tuberculosis*), *TPST1* (HIV), and *CXCR4* (HIV).
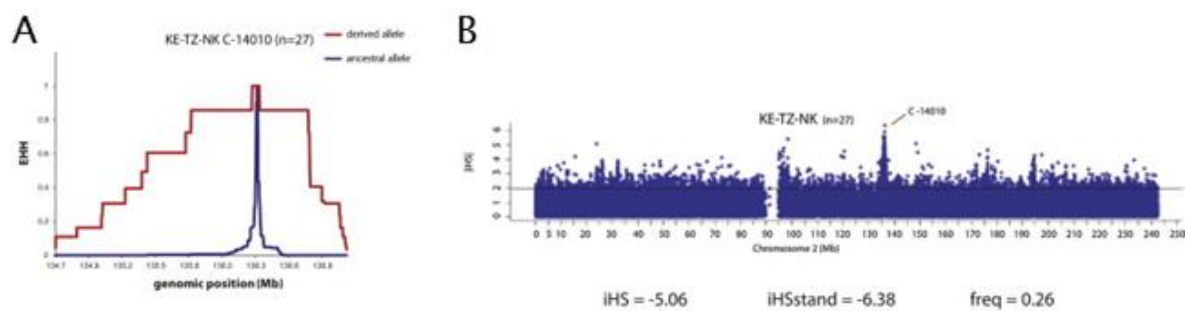
Figure 3. Example of positive selection detection through haplotype–based analyses in C–14010 polymorphism associated with lactase persistence trait in Kenyan and Tanzanian Niger–Kordofanian speakers. (A) Extended haplotype homozygosity (EHH). (B) Integrated haplotype score (iHS). Adapted from (Ranciaro et al., 2014).

1.4.   Public databases of human genetic variants

The technological development of recent decades allowed the characterization of large amounts of genomic data. Its processing provides a better understanding of human evolutionary history and genetic contribution to disease.

In 1991, the Human Genome Diversity Project (HGDP) was launched, aiming to obtain genetic information about the genome of indigenous populations around the world since they would provide more reliable insights about the human past and evolution (Cavalli–Sforza et al., 1991). The HGDP collection established 1064 lymphoblastoid

cell lines representative of the worldwide human diversity (Cann et al., 2002), that became the main reference panel for future studies and consortia.

The International HapMap Project arrived eleven years after the HGDP and proposed to find the common genetic patterns in the human genome of individuals from Africa, Asia and Europe. The plan included the genotyping of common SNPs, their analysis and correlation to construct LD maps to be used as support for genetic studies with biomedical application (Consortium, 2005). Data were fully released and open-access, reaching a total of 1184 characterized samples from 11 populations after Phase III in 2009, and covering up to 3.5 million SNPs (Altshuler et al., 2010; Frazer et al., 2007). These SNPs were used to construct chips than became an important tool in association studies. Despite the success of HapMap approach, its main focus was on common variants, excluding low frequency mutations due to technical limitations. The 1000 Genomes Project emerged in 2010 and proposed to overcome this limitation taking advantage of the recent high-throughput sequencing technologies. They aimed to collect the information for the full genome, covering bi-allelic variants with frequency lower than 1%. In the pilot phase, they used the extended sample collection from HapMap for three distinct approaches: high coverage whole-genome sequencing of two mother-father-child trios; high coverage exome sequencing corresponding to 906 genes in 697 individuals; and low coverage whole-genome sequencing of individuals from Yoruba from Ibadan (YRI), Japanese from Tokyo (JPT), Han Chinese from Beijing (CHB) and European ancestry from Utah (CEU) (Abecasis et al., 2010). In the latest phase, they applied whole-genome sequencing to 26 populations across Africa, South and East Asia, Europe and Americas, and extended the database to multi-allelic SNPs, structural variants and insertions/deletions (Auton et al., 2015). Up to date, this is the largest available database for human variation, providing numerous resources and tools that facilitate the analysis of candidate variants or genes from case-control cohorts across distinct populations. This population information also enables ethnic-informed imputation of SNPs not screened in a project, informs about allele frequencies and LD patterns across the globe and helps in design less-ascertained biased SNP arrays (Auton et al. 2015).

2. Complex diseases in the –*omics* era

Complex diseases result from a combination of several genetic and environmental factors. The first challenge is an accurate definition of phenotypes since for most of these diseases it is difficult to define discrete symptoms groups. Phenotypes depend on numerous genetic alleles distributed across several genes that can be variable between ethnic groups, and on environment factors that may vary at individual and population levels. The prevalence of the majority of complex diseases is not even across the world. For example, the presence of a certain infectious disease in a geographical region is dependent on a condition of factors that can range from temperature to the existence of a suitable vector.

The impact of complex diseases in the society is notable and includes some of the most common diseases: obesity, schizophrenia, type 2 diabetes, Crohn's disease, cardiomyopathy and infectious diseases. This led to a huge investment on association studies that could identify candidate genes. In order to achieve this goal, it is important to take advantage of the new high–throughput methods and the available genetic databases.

2.1. Genome-wide association studies

The technological improvements of the last two decades enabled the elaboration of SNP chips that allow to genotype thousands of variants at once. This large scale scanning method started to be applied to case–control cohorts to find new associations in an unbiased genome search, providing information about new candidate genes affecting the condition outcome (Visscher et al., 2017). These became known as genome–wide association studies (GWAS). The resource GWAS Catalog (www.ebi.ac.uk/gwas) includes all eligible GWAS since the first published one in 2005 (Klein et al., 2005), amounting by the 1st of September 2016 to 24,218 unique SNP–trait associations from 2,518 publications in 337 different journals (MacArthur et al., 2017).

In the classical GWAS, the statistical burden is very high (up to 10^{-8}) due to the multiple test correction (Bonferroni correction), which in consequence requires the genotyping of thousands of samples (Figure 4). In order to validate the findings and to decrease the report of false positive associations, the discovery set of SNPs that pass the significance threshold should be checked in a replication cohort composed by cases

and controls from the population used in the GWAS and, ideally, also in cohorts with a distinct ethnic origin (Manolio, 2010; Marigorta and Navarro, 2013). Most of the polymorphisms passing the significance threshold are tagging SNPs, not presenting a direct effect in the disease phenotype but being in LD with the true causative SNP. To disclose the real causative SNP, it is necessary to perform the fine mapping of the region in LD, and check for the strongest association, functional variants and presence of rare variants that had not been described. This approach has been applied with success to malaria (Jallow et al., 2009) and type 1 diabetes (Nejentsev et al., 2009).
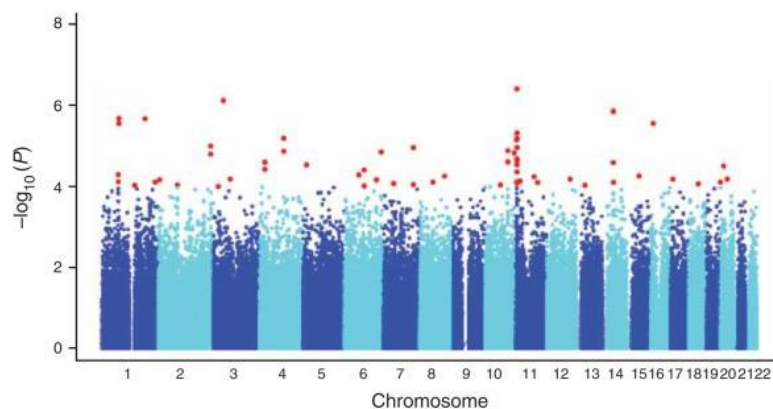
Figure 4. Manhattan plot of the $-\log^{10}$ p-values obtained in a control–case GWAS in severe malaria. The dots in red represent the SNPs passing the 10^{-4} significance threshold. Adapted from (Jallow et al., 2009).

The majority of the chips used in GWAS contain nonsynonymous and promoter SNPs, with a minor allele frequency (MAF) superior to 5% and not highly differentiated among populations. However, more than 80% of the candidate SNPs fall into intergenic and intronic regions, showing that noncoding SNPs play a major role in the outcomes of complex diseases, possibly by participating on gene regulation (Hindorff et al., 2009). Largely, the common variants being identified in the genome–wide approaches have a low to moderate individual contribution to the disease risk, with an odds ratio (OR) lower than 1.5 (Hindorff et al., 2009), as was already reported for, schizophrenia (Yue et al., 2011), malaria (Timmann et al., 2012) and asthma (Ferreira et al., 2011).

One common cause of false positives in traditional GWAS is population stratification, when cases and controls are not fully matched in terms of population group affiliation. To avoid this bias, the currently main used strategy is to apply a principal component analysis (PCA) to correct for stratification in the cohort (Price et al., 2006). These authors developed the EIGENSTRAT method, in three steps: first, principal components analysis is applied to the genotype data to infer continuous axes of

genetic variation (in data sets with ancestry differences between samples, axes of variation often have a geographic interpretation); second, genotypes and phenotypes are continuously adjusted by amounts attributable to ancestry along each axis (thus creating a virtual set of matched cases and controls); third, association statistics are computed using ancestry-adjusted genotypes and phenotypes.

This classical GWAS approach, with PCA-correction for population structure, loses the capacity to identify causative population-specific variants. Yet, it has been consistently shown, especially in admixed populations such as many American populations, that ancestry affects the susceptibility to several complex diseases (Farrer et al., 1997; Sierra et al., 2007; Takiar et al., 2015). Recently, new statistical methods have been developed, that take advantage of this admixed composition of populations, been known as admixture mapping or local ancestry affiliation (Li and Keating, 2014). These methods detect in the admixed individuals the mosaic of haplotypes inherited from the putative ancestral populations. The length of block will decrease along generations passed since the admixture event (Winkler et al., 2010), and the distribution in the descendants will be random in the overall (control) population. However, in cases, this distribution will not be random if a population-specific locus is associated with a disease (McKeigue, 1998; Patterson et al., 2004; Winkler et al., 2010): if an ancestry is associated with increased susceptibility, the disease locus will present a significantly higher proportion of that ancestry in cases versus controls (Figure 5); when an ethnicity is protective, the disease locus will present a significantly lower proportion of that ancestry in cases versus controls.
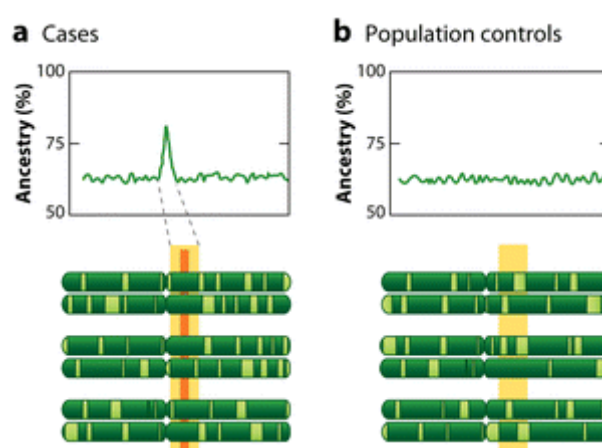
Figure 5. Disease locus identification based on the ancestral components. Adapted from (Winkler et al., 2010).

Since these methods are based in the detection of haplotype blocks (made of several SNPs), the statistical burden is lower than in the classical GWAS, leading also to lower sample sizes (Patterson et al., 2004; Winkler et al., 2010). This approach has been successfully applied to highly admixed populations in Latin American contributing with candidate genes to asthma (Mersha, 2015; Torgerson et al., 2012), blood lipids (Basu et al., 2009; Shetty et al., 2015), albuminuria (Brown et al., 2017), and type 2 diabetes (Martinez-Marignac et al., 2007).

2.2. Transcriptome studies

The human transcriptome studies are allowing to obtain data for all expressed RNAs, including the long noncoding- (lncRNA) and microRNA (miRNA). The two most common assays used are expression assays based on arrays and RNA Sequencing (RNASeq) or sequencing of cDNA (Casamassimi et al., 2017).

A useful contribution of transcriptome analysis to understand complex diseases is in the evaluation of GWAS-identified variants as controllers of the gene expression, the designated expression quantitative trait *loci* (eQTLs). Nicolae et al. (Nicolae et al., 2010) combined the data available in the GWAS catalog for 1598 SNPs with the information for HapMap lymphoblastoid cell lines (characterized for genetic and expression data), reporting that 625 of those trait-associated SNPs could be classified as eQTLs. This value shows that SNPs associated with complex traits are significantly more likely to be eQTLs than overall allele-frequency-matched SNPs.

Transcriptome-wide studies are also being used *per se* in association studies, especially in comparisons along disease progression (Sotiriou and Piccart, 2007) or between different phenotypes presented by the same disease (Mejias and Ramilo, 2014). Some studies focus on gene expression differential analyses between conditions, which has a high statistical burden (multiple testing for ~22,000 genes, or more if including miRNAs (Mootha et al., 2003)). However, methods denominated gene set enrichment analysis (GSEA) (Subramanian et al., 2005) take advantage of the fact that several of the genes that contribute to the complex disease act together in the same or similar metabolic pathways (Figure 6). So, by analyzing differences in gene expression between conditions within pathways containing at most hundreds of genes, they maximize statistical significance of tests.
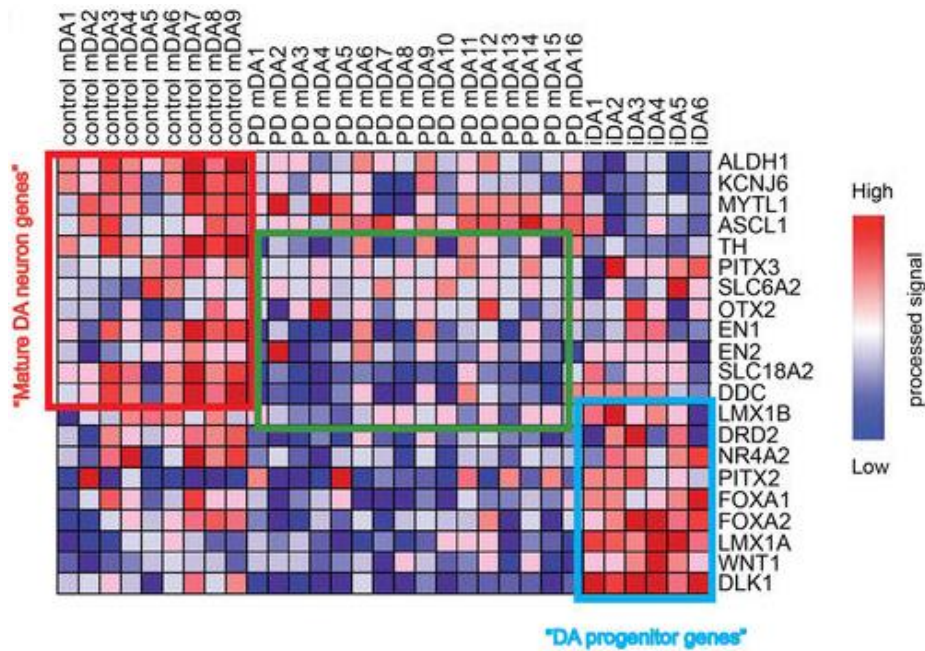
Figure 6. Heat map of dopaminergic (DA) genes generated in GSEA showing in blue the downregulated genes and in red the upregulated ones. The genes in the red rectangle are highly expressed in primary midbrain neurons; the blue rectangle contains genes highly expressed in induced dopaminergic neurons and, the green rectangle show genes with reduced expression levels in Parkinson's disease patients. Adapted from (Xia et al., 2016).

2.3. Genome-editing and phenotyping

The increasing number of candidate genes and polymorphisms being discovered in genome– and transcriptome–wide searches provide novel insights to understand the pathways and mechanisms behind common and complex diseases. These new hypotheses must be, nevertheless, functionally tested. Evaluating the candidates using *in vitro* or *in vivo* models is demanding, mainly if they are localized in noncoding regions and potentially have regulatory functions (Tak and Farnham, 2015).

Functional *in vitro* assays became easier when RNA interference (RNAi) was discovered in *Caenorhabditis elegans* (Fire et al., 1998), being widely used in the last 15 years for gene silencing. This gene silencing technique does not imply any alteration in the sequence since it only degrades the target mRNA after incorporation into the RNA–induced silencing complex (RISC) (Rana, 2007). The most used molecules are the small interfering RNAs (siRNAs) and the short hairpin RNAs (shRNAs). The siRNAs can be directly transfected into mammalian cells and lead to a transient gene knockdown or knockout. The shRNAs require the cloning into a plasmid that will be stably

transfected into the mammalian cells and provoke a stable knockdown of the gene of interest.

Recent developments in the genetic engineering field adapted the bacterial adaptive immune systems called clustered regularly interspaced short palindromic repeats (CRISPR)–Cas9. The Cas9 nuclease is used to cleave the genome when combined with a guide RNA (gRNA) specifically design to target the desired sequence (Ran et al., 2013). This method edits at DNA level, having repercussions at RNA and protein stages. Commonly, this system is used to create double strand breaks (DSB) in the target *locus* for modification (Figure 7). The reparation of the DSB can be performed in two ways, depending on the purpose of the work: to survey the impact of point mutations, the repair can occur by homology–directed repair (HDR) providing DNA templates that will be incorporated through recombination with the selected locus; for gene loss–of–function, nonhomologous end–joining (NHEJ) repair system introduces insertions/deletions (indels) that are expected to disrupt the open reading frame or promoter/enhancer activity depending on the target (Sander and Joung, 2014).
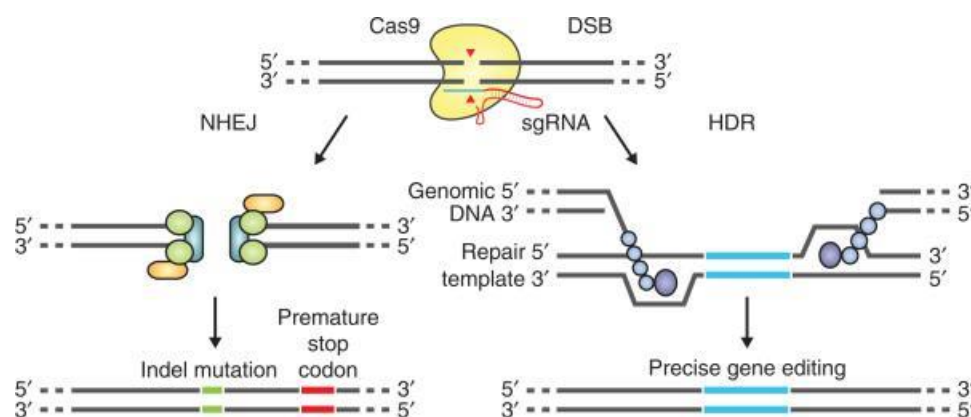
Figure 7. CRISPR/Cas9 system and DSB repair. The gRNA sequence directs Cas9 to cut and cause a double strand break in the target sequence. The repair mechanism can be through the NHEJ repair system, or include a template DNA that leads the HDR. Adapted from (Ran et al., 2013).

Thus, CRISPR has been successfully applied in the last years to knockout genes of interest and helped in disclosing their function and of other factors influencing the phenotype (Bressan et al., 2017; Cao et al., 2016; Chu et al., 2015; Puschnik et al., 2017). It is also possible to combine both genome engineering and RNAi techniques

to explore more complex phenotypes and validate the general function of disease-associated candidate genes (Mohr et al., 2014).

3. Dengue

Dengue virus origin is thought to be Africa, in common with other flaviviruses. During the 18th and 19th centuries, dengue spread to Asia and, after World War II, to the rest of the world. Its wide geographic distribution contributed to turn it in one of the most important neglected diseases worldwide, being endemic in several countries from Asia, Oceania and Americas, and representing a risk to 2/3 of the human population. Consequently, it has a considerable economic impact in the poorest regions of the world. Yet, dengue infection does not affect all the regions in an even manner, showing differences in outcome given different ethnicities.

3.1. Dengue virus

Dengue virus belongs to Flaviviridae family and *Flavivirus* genus, along with West Nile, Japanese Encephalitis, Yellow Fever and Zika viruses. Dengue infection can be caused by four phylogenetically distinct serotypes (DENV 1-4) that are constantly in circulation within the endemic areas. The serotype evolution occurred in non-human primates and it is hypothesized to have started circulation among humans around 1000 years ago (Wang et al., 2000).

3.1.1. Viral genome

The four dengue viruses present 65% genome similarity among them, contributing to their shared biological features, clinical symptoms and niche (Guzman et al., 2010; Halstead, 2008).

The virus genome is codified by a ~10.7 kb length positive-stranded RNA that works both as mRNA and polyprotein. This polyprotein comprises three structural proteins -capsid (C), premembrane (prM) and envelope (E) - and seven non-structural (NS) proteins - NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5 (Figure 8).
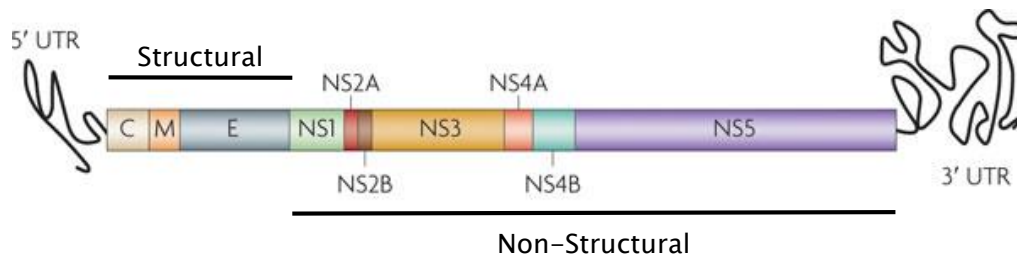
Figure 8. Dengue viral genome illustrating the ten coded proteins: capsid (C), premembrane (prM), envelope (E), and the non-structural (NS) proteins NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5. Adapted from (Guzman et al., 2016).

The ten viral proteins participate in processes required to produce efficient mature particles and increase the dengue virus fitness within the host. The structural proteins encapsulate (C protein) and coat (prM and E proteins) the viral genome and, after premembrane cleavage, originate the mature virion. The non-structural proteins participate in the hijack of host cell system and control of immune response to allow virus replication (Guzman and Harris, 2015).

3.1.2. Viral assembly and replication

To enable DENV entry into the human host cells, E proteins bind to receptors present in the host cell membrane allowing the virion receptor-mediated endocytosis (Acosta et al., 2014). Theses receptors can vary among species and cell types. After the virion uptake, the acidic late endosome membrane fuses with the virion envelope, leading to the nucleocapsid release into the cytoplasm and viral genome uncoating, through capsid dissociation, leaving the RNA free to be translated in the cytoplasm (Byk and Gamarnik, 2016; Yap et al., 2017).

The genomic RNA is translated into a single polyprotein in the ribosomes allocated in the endoplasmic reticulum (ER). Later, this polyprotein is cleaved by NS3 and other non-identified host proteases. To initiate the replication process, the virus hijacks the host machinery and, in combination with the non-structural proteins, promotes ER membrane invaginations connected to the cytoplasm (Acosta et al., 2014; Welsch et al., 2009). In a first stage, NS5 RNA-dependent RNA polymerase (RdRp) participates in the generation of a complementary negative strand that will function as template for the positive strands amplification. The synthetized RNAs link to the capsid forming nucleocapsids, that migrate to the ER lumen regions with prM and E enriched micro-

domains, given origin to imature virions (Perera and Kuhn, 2008). The prM cleavage is performed by host furin when the immature virion crosses the trans-Golgi network, leading to the pr peptide discard and the mature virion, that will be secreted in the extracellular matrix (Yu et al., 2008).

3.2. Vector

Dengue virus can be transmitted by two related species, *Aedes aegypti* and *Aedes albopictus*, in a similar process to other flaviviruses as Zika and Yellow Fever. The first vector is considered the most efficient one (Kraemer et al., 2015).

3.2.1. Geographical dispersion

Ae. aegypti has a tropical and subtropical distribution, being present in Asia, Africa and Americas. *Ae. albopictus* is considered less efficient for transmission, however, it has a wider distribution, being also present in Europe and North America, and was the cause of dengue autochthonous outbreaks in areas without *Ae. aegypti* (Guzman et al., 2016). The mosquitoes are highly adapted to the human environment, feeding preferably on humans and biting several individuals in short periods (Kraemer et al., 2015). Other factors that largely contributed for the overgrowing number of dengue cases in the last decades are overpopulated urban centers and uncontrolled urbanization, mostly in poor tropical areas, as well as increased travel flow and globalization (Wilder-Smith and Gubler, 2008) combined with the climate changes that allow the expansion and survival of the mosquito species (Figure 9), mostly *Ae. albopictus* since it is more robust than *Ae. aegypti* (Guzman et al., 2016; Wilder-Smith et al., 2017).
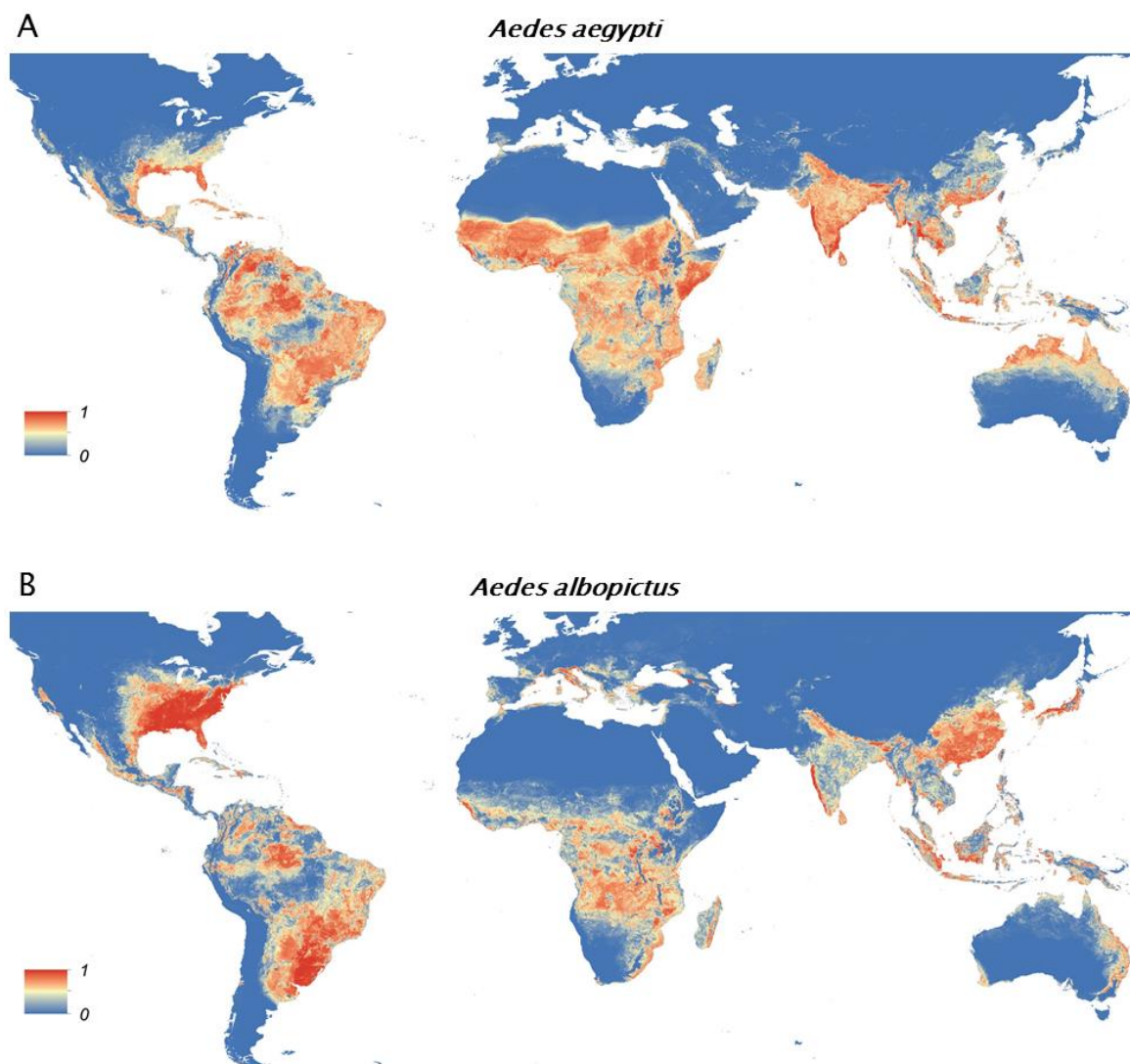
Figure 9. Predicted global distribution of Aedes aegypti (A) and Ae. albopictus (B). Darker blue indicates absence and red presence of the species. Adapted from (Kraemer et al., 2015).

3.2.2. Transmission cycle

It is believed that the ancient DENV forms were circulating in sylvatic cycles, infecting nonhuman primates and further evidence in favor of this supposition was found in Asia and Africa (Vasilakis et al., 2011). Currently, the four worldwide circulating dengue virus strains adapted to the referred *Aedes* species and human host.

The transmission cycle is carried on in female mosquitoes and starts after a blood meal from a dengue infected person during the acute phase of symptoms and viremia. In the mosquito, the virus begins the infection in the midgut, passing to other tissues and reaching the salivary glands 5 to 12 days later. Once here, the female becomes infective during its lifetime and able to pass the virus to non-infected humans. The

cycle repeats again around 4 to 7 days after the human infection, when the viremia reaches the highest values (Guzman et al., 2016).

3.2.3. Vector control

Several measures were tested to control the mosquito propagation, reproduction and viability. The first attempts occurred in the 1950s and 1960s in South America, by using dichlorodiphenyltrichloroethane (DDT) directly in the breeding sites, leading to a major decline of dengue incidences (cited in (Reiner et al., 2016)). Later controls using other compounds in Cuba (Kourí et al., 1998) and Singapore (Ooi et al., 2006), were less effective. More recently, an indoor insecticide was developed, which presented a good success rate (Wilder-Smith et al., 2017). Yet, all current procedures fail to work completely due to poor sustainability in the endemic areas, insecticide resistance, lack of monitoring and control to apply the strategies, insufficient economic funds, poor hygiene conditions and continuous travelling and vector spreading (Bowman et al., 2016).

Other approaches that are being explored lately include the mosquito infection with *Wolbachia*, or using genetic engineering to cause male sterility or transform vectors to lower efficiency for the virus (Ritchie and Devine, 2013). *Wolbachia* is an endosymbiotic gram negative bacterium that restricts transmission of human pathogenic flaviviruses and alphaviruses, by manipulating host biology to induce feminization, parthenogenesis, cytoplasmic incompatibility and male-killing, and by conferring antiviral activity to its host in a phenomenon not fully understood (Fraser et al., 2017). This antiviral activity can act through competition for resources (as cholesterol) or by priming the insect innate immune response pathways (Serbus et al., 2008; Werren et al., 2008).

3.3. Global burden and distribution

As previously referred, dengue virus is present around tropical and subtropical areas, where its incidence increased 30-fold in the last 50 years. According to the World Health Organization (WHO) 2009 guidelines (WHO, 2009), there are 50 million cases per year. Dengue cases are not evenly distributed, and the most affected area is undoubtedly Asia and the Pacific Islands, where around 70% of total population is considered to be at risk of dengue. Since the beginning of the 21st century, epidemic

dengue is reaching news locations, mostly in Southeast Asia and Western Pacific. WHO included Indonesia, Myanmar, Thailand, Sri Lanka and Timor-Leste in this risk group once dengue represents the major cause of hospitalizations and death during childhood. In the Western Pacific (WHO/SEARO, 2008), the largest outbreak occurred in 1998, leading to an increment of epidemic cases in the region. A combined strategy aiming to prevention and control – mainly focused on surveillance, case and vector management, outbreak response, social mobilization and communication and dengue research – is being applied to Southeast Asia and Pacific regions (WHO/SEARO, 2008).

The American continent was the first place to eradicate the main dengue vector during the 1960's and 1970's. Despite the success of the adopted measures, they revealed to be unsustainable due to constant mosquito re-infestations (PAHO, 2007). Since then, outbreaks occurred in Caribbean, Central and South Americas every 3-5 years. On the other side, the majority of the cases described in Canada and United States of America result from travelers to endemic areas such as Asia and epidemic Latin American countries. The only exceptions of autochthonous transmissions are Hawaii and the border between Mexico and Texas (Centers for Disease Control and Prevention, 2006).

Africa has long been reported as a non-endemic continent for dengue disease and the cases are not officially reported to WHO. Yet, dengue is present in Africa but, given the weak surveillance and lack of laboratorial confirmation, it may be misidentified due to similar symptoms with other tropical diseases that are much more common there (Stoler et al., 2014). Most of the testified dengue cases belong to travelers after returning to their countries or from sub-Saharan Africa (Neumayr et al., 2017). DENV1-3 serotypes were found in eastern and western Africa as a cause of acute fever, with higher incidence in the last 20 years, indicating the growing number of dengue infections in Africa, with higher occurrence in eastern Africa, as has been observed around the tropical and sub-tropical areas (Diallo et al., 2003; Gonzalez et al., 1985). More recently, there were reports of the first dengue outbreaks in Cabo Verde, during 2009 (Guedes et al., 2017), and in Angola in 2012 (Parreira et al., 2014).

Regarding Eastern Mediterranean, the outbreak frequency has been increasing since the 19th century, especially in the last 30 years. The most recent outbreaks happened in Sudan, Yemen, Saudi Arabia and Pakistan. In the case of Saudi Arabia, the number of cases and outbreaks is significantly enlarged in Jeddah that bears one the biggest airports and commercial port in the area and serves as a door to dengue endemic Asian countries (WHO/EMRO, 2005).

Europe does not belong to WHO classified dengue endemic regions, and most of the endemic cases come from travelers that returned from endemic countries. The main dengue vector was present in Europe until the middle of the 20th. Despite being almost vanished in continental Europe, the population from the Portuguese island Madeira reported bites during daylight and consequent pain provoked by *Ae. aegypti* activity in 2005. Later on, in 2012, two cases of uncomplicated dengue (DGS, 2013) were identified and, in the beginning of 2014, more than 2000 cases of confirmed dengue occurred there. All the detected cases were infected with a strain from serotype DENV1 related with viral strains present in South America (Alves et al., 2013). Another DENV1 strain circulated in southern France (La Ruche et al., 2010) and Germany (Schmidt-Chanasit et al., 2010) with origin in Martinique and, additionally, a DENV2 strain from Caribbean was found in the south of France in 2013 (Marchand et al., 2013). All the cases were transmitted by *Ae. albopictus* vector (Marchand et al., 2013). Given this, the risk of dengue epidemic in Europe is bigger than speculated due to adaptation of *Ae. albopictus*, the less efficient vector, that nowadays has a considerable presence in this continent (Parreira and Sousa, 2015).
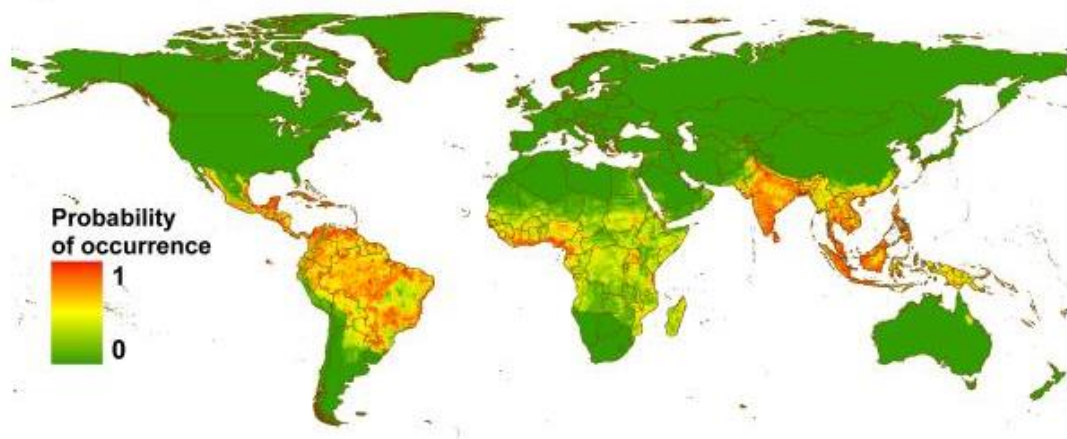
Figure 10. Probability of occurrence of dengue in 2010 based on evidence consensus. Green indicates low probability, yellow intermediate and red high probability. Adapted from (Bhatt et al., 2013).

These worldwide dengue cases that are reported to WHO do not reflect the correct picture. Bhatt et al. (Bhatt et al., 2013) undertook an exhaustive assembly of known records of dengue occurrence worldwide (8,309 geo-located records), and used a formal modelling framework to map the global distribution of dengue risk based on evidence (Figure 10). They got estimations higher than the reported ones: in 2010, there were 96 million cases worldwide. These cases were not equally distributed: 70%

occurred in Asia, predominantly in the most populated regions as India, followed by the Americas and Africa, each of them bearing around 15% of total infections. Aside from the identified cases as dengue, a concerning estimation of 294 million additional cases in the same year come from unapparent infections (including undiagnosed asymptomatic and mild infections). The underestimation of these cases results in lack of accuracy on predicting the virus population dynamics and the economic impact of the disease, affecting enormously the design of vaccination campaigns (Johansson et al., 2011) and control measures (Bhatt et al., 2013).

3.4. Clinical features

Dengue illness shares some clinical indicators with other infections caused by flaviviruses, which complicates its early diagnosis. It usually requires a close monitoring of the patients to follow up any alterations that can lead to severe complications. It is also important to check patient clinical histories, once it is more common to observe severe outcomes in people experiencing a secondary infection rather than in the primary one. Other factors such as age, viral serotype, ancestry and immune status play a role in the disease progress (Guzman et al., 2016).

3.4.1. Phenotype classification

After being bitten by an infective mosquito, the onset of symptoms in humans takes 4 to 7 days in average to manifest (Figure 11). During this period, named acute phase, the infected individuals exhibit high fever and high viremia peak. Then, depending on the development of warning signs, this is followed by the critical phase, associated with plasma leakage among others, that can be followed by the convalescent period or evolve to a more serious condition (Guzman et al., 2010).
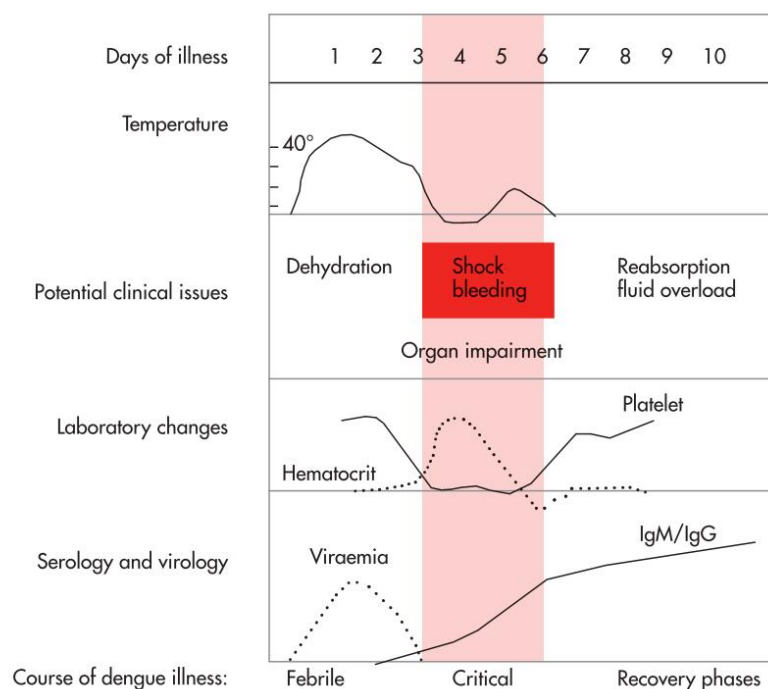
Figure 11. Dengue illness evolution. Adapted from (WHO, 2009).

In 2009, the WHO reviewed the classification of dengue symptoms to overcome the difficulties on classification and disease progression. The categories from the previous guidelines (WHO 1997), Dengue Fever (DF), Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS), were replaced by new designations – Dengue without warning signs, Dengue with warning signs and Severe Dengue – that better answered the wide range of dengue clinical manifestations and their correct structuring (WHO, 2009).

Dengue without warning signs includes mild symptoms such as nausea, vomiting, rash, aches, pains, tourniquet test positive and leukopenia. Dengue with warning signs, despite falling into the non–severe dengue category, already requires medical care and accompaniment due to the manifested symptoms – abdominal pain or tenderness, persistent vomiting, clinical fluid accumulation, mucosal bleeding, lethargy, restlessness, liver enlargement and increase in hematocrit concurrent with rapid decrease in platelet count. Severe dengue is characterized by severe plasma leakage, that leads to shock and fluid accumulation with respiratory distress, severe bleeding and organ involvement affecting the liver (serious alteration of the enzymatic levels), central nervous system (impaired consciousness), heart and other organs (such as kidneys) (WHO, 2009).

3.4.2. Immune condition

Human immune response to dengue infection is dependent on innumerous factors, being the most notable one the previous exposure to dengue virus. Once a person gets infected with one of the dengue virus serotypes, it acquires lifelong immunity to that specific serotype. Immunity against the remaining serotypes is also observed, yet, it only lasts for a few months after the primary infection (Montoya et al., 2013a). As result, these individuals carry two types of antibodies: the neutralizing, that allow the protection directed to the serotype from the primary infection, and the non-neutralizing, that are known to enhance DENV entry into cells carrying Fc receptors (Montoya et al., 2013b). This reaction is exclusive to dengue and it is called antibody-dependent enhancement (ADE). In individuals suffering a secondary infection, this cross-reactivity usually leads to an augmentation of severity and, in most cases, it evolves to severe dengue (Wahala and Silva, 2011). There are reports about infections with other serotypes after the secondary infection, yet the risk is considerably lower and is usually associated with mild or asymptomatic infections (Guzman et al., 2016; Olkowski et al., 2013).

The ADE can also be problematic to newborns and infants. It was already stated that mothers carrying neutralizing antibodies against one or two dengue serotypes pass them to the fetus through the bloodstream. This transference confers protection in the early months, yet, when the levels start to decrease it renders the infants more susceptibility to severe dengue when they are infected by another serotype (Chau et al., 2010; Clapham et al., 2015; Guzman et al., 2016).

3.4.3. Patient age, origin and social class

Dengue occurs mostly in Latin America and Southeast Asia and, currently, the four serotypes are circulating in both regions. Historically, the virus expansion occurred in the same century but in different periods: with World War II in the 1940s in Southeast Asia , and in the 1960s-1970s in the Americas (Halstead, 2006).

Despite the increasing number of cases in the Americas (Figure 12A), the highest death rate is observed in Asia (Figure 12B). In 2016, 2000 deaths were notified in Southeast Asia whereas the Americas reported only half number of deaths (Figure 12B).
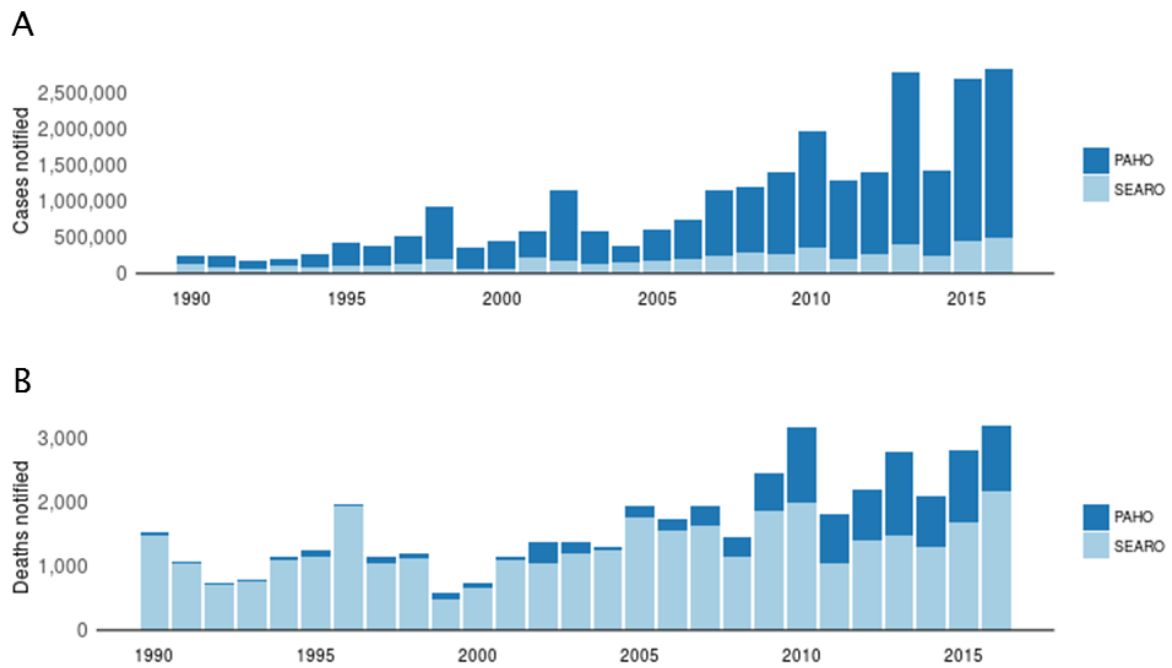
Figure 12. Number of dengue cases (A) and deaths (B) notified between 1990 and 2016 for two of the most affected WHO-recognized regions, coordinated by PAHO (Pan America Health Organization) and SEARO (South-East Asia Regional Office). Accessed on February 2018 and adapted from http://www.who.int/denguecontrol/epidemiology/dengue_data_application/en/.

An extensive study in pediatric cohorts from ten countries endemic for dengue fever from Southeast Asia (Indonesia, Malaysia, Vietnam, Philippines and Thailand) and Latin America (Brazil, Mexico, Colombia, Puerto Rico and Honduras), exhibited clear differences in the incidence of severity between the two regions (L'Azou et al., 2016). In all the comparisons, the Asian cohort presented the highest dengue burden, 4.6 (range: 2.2-6.6, per 100 person-years) virologically confirmed dengue (VCD) episodes versus 2.9 (range: 1.5-4.5, per 100 person-years) VDC episodes in Latin America. Also, the number of VCD cases requiring hospitalization was higher in Southeast Asia, with 19.1 % (range: 5.9-45.5%), than in Latin America, with 11.1% (range: 4.9-17.0%). Besides, dengue hemorrhagic fever cases were more frequently observed among Southeast Asians (0.3 episodes, range: 0.1-0.6, per 100 person-years), where all the countries registered severe cases, while, in Latin America, the average was 0.1 episodes (range: 0.0-0.1, per 100 person-years), with no episodes of DHF detected in Brazil and Mexico (L'Azou et al., 2016). In the general picture, the majority of hospitalizations due to severe dengue occurred in Southeast Asia, among children younger than 15 years old (Carlos et al., 2005; Kittigul et al., 2007), since they are more prone to manifest higher severe plasma leakage (Guzman et al., 2010)

61

that has been related with higher capillary permeability (Gamble et al., 2000). Despite the acquired immunity length of approximately two years, the hyperendemic status and four circulating serotypes represent a major threat to children public health in Southeast Asia (Gubler, 2011). Latin America presents more spaced outbreaks and has predominantly DF cases among adults. Brazil, that contributes with roughly 70% of total infections occurring in the Americas, has reported the same pattern, with a recent increment of children hospitalization due to DHF (Siqueira et al., 2005). During Cuba 1981 outbreak, vascular permeability were higher among younger individuals than in adults, whom exhibited mostly mild symptoms (Bravo et al., 1987) although both in Brazil and in Cuba ancestry can be a confounding factor (Sierra et al., 2007b).

Economic status is also a factor behind dengue susceptibility. The high occurrence of the mosquito in urban and suburban areas increases the risk for people living in places with unplanned urbanization and without capable water supplies (Chang et al., 2014). Human poverty conditions associated with slums also increase the risk of dengue fever due to water contamination and garbage accumulation, enabling the mosquito to breed in high population density areas (Knudsen and Slooff, 1992). On the other hand, wealth also brings risks associated with peri–domiciliary water storage containers and other surfaces as fountains, vases and pools, that act as mosquitoes breeding sites (Chang et al., 2014).

The increment of dengue cases justified estimation of costs, revealing that, in 2013, the total annual global cost of dengue illness was US$8.9 billion, without counting the cost related with persistent dengue cases (Shepard et al., 2016). This amount evidences the lack of good preventive measures and treatments that would not only decrease the disease burden but also the economic burden associated with it.

3.4.4. Antiviral strategies

Currently, there is no treatment available in case of dengue infection. Once a person gets infected, the subscribed care relies on resting and oral or intravenous fluid intake, according to dengue severity. Given the lack of preventive therapies or medication directed to treat dengue infection, there has been an effort in the last years to develop a reliable vaccine and discover antiviral drugs to block the evolution of symptoms in an early stage of the disease.

3.4.4.1. Antiviral drugs

Despite several candidate drugs appearing every year, the majority fails to cover the four serotypes, presents insufficient potency or even produces adverse effects in animals (Kaptein and Neyts, 2016).

Balapiravir has been tested for dengue infection due to its proved antiviral effect on the RNA-dependent RNA polymerase (RdRp) from Hepatitis C virus (Roberts et al., 2008) that is quite similar to the one from dengue viruses. It exhibited promising *in vitro* results, however, it failed to reduce viremia, viral load, cytokines expression, among other parameters in dengue patients, indicating that it may not be suitable for treating ongoing dengue infections (Nguyen et al., 2013).

Chloroquine was also evaluated due to its ability to alkalize the acidic intracellular environment, that favors the viral replication, and to decrease the levels of pro-inflammatory cytokines. In resemblance with the previous drug, it also showed some *in vitro* good results that did not compare when used in patients, in terms of viremia and cytokines expression. Yet, there is some evidence that chloroquine alleviates pain, improving life quality of patients during infection (Borges et al., 2013).

Lovastatin acts on the inhibition of cholesterol synthesis, which is required for efficient viral replication. However, no effect was observed in patients after its administration (Whitehorn et al., 2012). The treatment with Prednilosone, a known corticosteroid, achieved the same results as Lovastatin (Tam et al., 2012).

Celgosivir is an inhibitor of the endoplasmic-reticulum-resident α-glucosidase I enzyme and, *in vivo* assays in mice provided a promising antiviral effect on dengue virus (Watanabe et al., 2012). However, this prodrug was not able to produce any effect on patients and even in the *in vitro tests* it presented heterogeneous results depending on cell line or serotype (Kaptein and Neyts, 2016; Low et al., 2014, 2017).

3.4.4.2. Vaccines development

The burden of dengue has been increasing in the last decades, resulting in the exponential growth of the number of cases. The vaccine against dengue virus has been under investigation for several years due to the challenge in creating one vaccine that provides immunity against the four serotypes. Other desirable traits are lifetime protection and escape from the development of antibody-dependent enhancement. If the vaccine only offers protection against one of the serotypes, it is possible that

when the vaccinated individual is infected with a different serotype, the organism responds as if it was a secondary infection, what will increase the severity of symptoms (Guzman et al., 2016; Wilder-Smith and Macary, 2014). At the moment, there are three candidate vaccines. Two are at phase III of clinical trials, TDV from Takeda (Huang et al., 2013), and TV003 from Butantan, U.S. National Institutes of Health (Durbin et al., 2011). Dengvaxia (CYD-TDV) from Sanofi-Pasteur has the phase III completed and was licensed and recommended in 2016 by WHO (Halstead, 2017; Report, 2017).

Following WHO recommendations, Dengvaxia was introduced in countries with a high disease burden, data that should be supported by epidemiologic reports. This included several countries across Asia and South America, which are now being studied for surveillance, seroprevalence and other parameters to quantify the population response through the years. The present vaccine was designed in the yellow fever backbone, and includes the structural genes from dengue virus and the non-structural genes from yellow fever. In phase III, this vaccine showed to be more efficient in people that had a previous dengue infection than in seronegative individuals. It also increased the risk of hospitalization by dengue infection in children between two and five years old. Given these evidences, it was recommended to be given to individuals older than 9 years old. In countries that have cases of dengue among the adult age, it was established the maximum age of 45 years for vaccine intake. (Report, 2017; Vannice et al., 2017). Biologically, the 9 years old threshold makes no sense, and the obtained results are more likely related with a lower rate of seronegative children among the trial groups (Dans et al., 2017). Philippines and Brazil were the first two countries to introduce the vaccine in 2016, however, recent deaths in Philippine children after been vaccinated led WHO to withdraw the recommendation of vaccinating individuals that were not previously infected with wild dengue virus (WHO, 2017).

3.5. Human susceptibility to dengue

Dengue is a complex disease with several host factors interplaying for the outcome. The majority of the studies are concentrated on understanding its vector, virus evolution and the intricate complexity of human antibody response, mainly after the first infection, leading to a gap in knowledge regarding the human host factors that contribute to a higher risk of developing the disease and its severe form (Guzman et

al., 2016). In the last decades, technological improvements are allowing the generation of large amounts of genomic data, enabling studies on host genetic susceptibility to dengue.

3.5.1. Epidemiologic information on ethnicity influence in dengue

Host ancestry is known to affect dengue susceptibility or resistance to develop severe symptoms, as first indicated by epidemiological reports during outbreaks. Cuba has been systematically recording in several outbreaks the ethnic group distribution in dengue phenotypes. During three outbreaks – 1981, 1997 and 2001 – it was consistently observed that DHF and DSS frequency was lower among individuals with African ancestry evaluated through skin color (Sierra et al., 2007b) yet, increasing over time from 7% of the total cases in 1981 to 19% in the 2001 outbreak (Bravo et al., 1987; Gonzalez et al., 2005), and the number of fatalities followed the same tendency. Comparing the population composition with dengue cases frequency, there was an overrepresentation of Caucasian ethnicity individuals among the severe dengue cases, leading to speculate the higher susceptibility of this ethnic group (Sierra et al., 2007b). The Cuban samples from 1977 and 1981 outbreaks were additionally tested for T-cells response, revealing higher serotype-specific response and cross-reactivity in individuals classified with Caucasian in comparison with African ethnicity (Sierra et al., 2006). As severity of secondary dengue infections is associated with ADE, the greater specific response observed in people with predominant Caucasian origin corroborates their susceptibility to severe dengue in Cuba.

The data collected in Trinidad and Tobago during the outbreak of DENV1 in 1996 also indicates that people with African ancestry are less susceptible to dengue due to a higher frequency of south Asians among the detected cases than the estimated percentage in the general population (Brown et al., 2004). Haiti was also investigated in the context of dengue severity and ancestry, since it is predominantly of African descent. Although it is estimated that 30% of dengue infections occur per year in Haiti, and immunological tests performed in children indicated the occurrence of primary and secondary dengue infections, no cases of DHF or DSS have been reported as sporadic or as outbreaks in Haitians (Halstead et al., 2001), reinforcing the protective role of African ancestry.

Southeast Asia represents one of the most affected regions by dengue, however, the epidemiologic reports concerning severe dengue prevalence are still scarce. In 1992,

a report (Shekhar and Huat, 1992) based in more than 9000 cases of dengue fever and dengue hemorrhagic fever in Malaysia between 1973 and 1987, indicated the following frequencies per ethnic group: from the total number of informed cases, 56.5% were Chinese whereas 32.8% corresponded to Malays and 4.2% to Indians. The morbidity was also greater among Chinese (9 in 100 000) than in both Malay and Indian ethnicities (2.9 and 2.4 in 100 000, respectively) (Shekhar and Huat, 1992). Interestingly, dengue seroprevalence on 3293 Singapore adults indicated lower values amid the Malay group (50.2%) in comparison to the Chinese (57.0%) and Indians (62.0%) (Ang et al., 2015).

3.5.2. Genomic information on ethnicity influence in dengue

Epidemiologic records on ancestry in admixed populations are biased by personal identity perception and over-importance of characteristics like skin color. Genetic information is fundamental to accurately measure the ancestral proportions in an admixed population.

The first genetic study (Blanton et al., 2008) conducted to investigate ancestry in a dengue cohort was performed in the Brazilian state of Bahia in 49 confirmed DHF cases, 293 neighborhood cases of dengue fever and 294 asymptomatic controls, genotyping 30 ancestry informative markers (AIMs) and 282 SNPs unrelated with clinical manifestations of dengue. African (OR=0.13, p-value=0.02), and Afro-Brazilian (OR=0.28, p-value=0.02) genetic components were positively associated with protection against dengue hemorrhagic fever (Blanton et al., 2008).

A similar approach was applied to the Colombian population, through the genotyping of 30 AIMs in two main groups, dengue (mild cases with or without warning signs) and severe dengue, classified using both WHO 1997 and 2009 recommendations (Chacón-Duque et al., 2014). The results of this paper corroborated the findings in Brazil, and found a significant association between the protective effect of African ancestry against DHF and severe dengue, occurrence of hemorrhages and other warning signs. Moreover, severe outcomes occurrence depends on the African ancestry percentage, a decrease from 100% to 0% results in the exponential increment of severe dengue risk, from OR=0.023 (CI=0.001-0.493) to OR=43.860 (CI=2.029-947.929), respectively (Chacón-Duque et al., 2014).

Our group (Sierra et al., 2017) performed a global admixture analysis in Cuban dengue cohorts based on information from 2.5 million SNPs. The average African

ancestry was significantly lower in DHF (22.9%) when compared with DF, controls and especially asymptomatic groups (30.6%, p-value = 0.025; 30.0%, p-value = 0.041; 34.7%, p-value = 0.013, respectively). Nevertheless, the evidence of ancestry influence in dengue was not so straightforward when the samples were divided according to the city of origin: the African ancestry was more protective (OR=0.012) in Havana than in Guantanamo (OR=0.151). Our data show that there is an African protection conferred against DHF in Cuba, but other currently unknown confounding factors render it a complex relation, even in such a geographical restricted scenario as Cuba.

Until the moment, no study was conducted to evaluate ancestry admixture in Southeast Asian populations, urging the application of genetic approaches to enlighten the intrinsic factors leading to the observed disparities in infection observed worldwide.

3.6. Candidate genes associated with dengue

Dengue is a multifactorial disease, where genetics plays an important role on the determination of disease evolution and susceptibility. Some factors limit the application of genetic approaches to dengue cohorts, such as the requirement of patients follow-up for correct phenotype affiliation, an elevated sample size to increase the statistical power and the application of the same strategy to different geographic areas to validate the genetic associations across the Americas and Asia (Coffey et al., 2009).

3.6.1. Single gene approach

The first genetic studies regarding dengue disease were performed following the classical approach focusing on candidate genes and mostly, target SNPs. The use of this approach demands previous knowledge about the gene function and information on target SNPs from related diseases. The primary efforts were directed to genes known to be involved in immunity and signaling cascades, such as genes encoding cytokines and chemokines. Direct comparison between these works is biased by heterogeneity on ethnicity, age, circulating DENV strain, among others (Coffey et al., 2009).

3.6.1.1. Human Leukocyte Antigen system

The human leukocyte antigen (*HLA*) *loci* were considered the most appealing as starting point for association study to dengue infection. It is one of the biggest gene family, containing the most polymorphic genes and possessing thousands of SNPs with functional relevance (Rothman, 2011). Also, *HLA* expression increases when the cells get infected with DENV (Gan et al., 2015). HLA is the given name to the major histocompatibility complex (MHC) in humans and maps in chromosome six short arm (Figure 13). *HLA* genes are divided in three classes: class I, that encodes the classical genes *HLA-A*, *HLA-B* and *HLA-C*, expressed in the surface of the majority of nucleated cells; class II, encompassing three gene families, DR, DP and DQ, that are restrict to B and activated T lymphocytes and antigen–presenting cells; and class III that is responsible for coding complement components and *TNF* genes, among others, that not represent the classical HLA molecules (Choo, 2007).
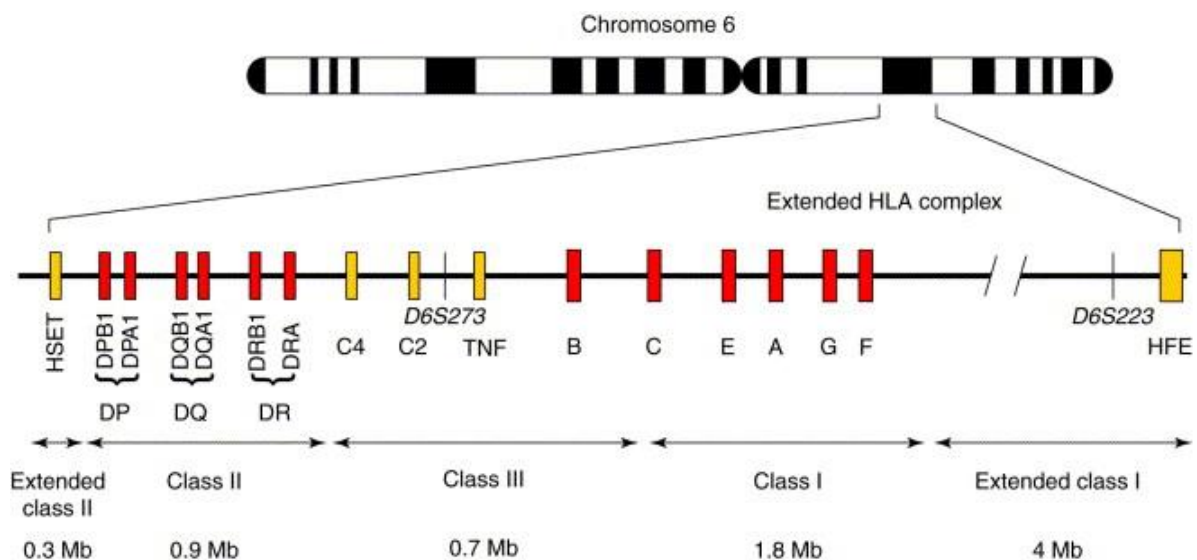
Figure 13. Classes and some genes belonging to HLA complex. Adapted from (Undlien et al., 2001).

Class I and class II *HLA* present characteristic LD patterns, leading to the detection of specific haplotypes that vary between ethnicities and that influence the immune response (Vina et al., 2012).

In 2002, a comprehensive *HLA-A* and *–B* allele genotyping centered on ethnic Thais (Stephens et al., 2002) found an association between peptide binding properties and secondary infection phenotype: *HLA-A**0203, more permissive in peptide binding, is linked to DF in opposition to *HLA-A**0207, more restrict and associated with dengue

hemorrhagic fever, but only in case of infection with DENV1 or DENV2. Independently, *HLA-B*52* was correlated with less severe DF in case of infection with DENV2 (Stephens et al., 2002). The work developed targeting a Vietnamese cohort (Nguyen et al., 2008) led to the discovery of additional susceptibility *HLA-A* alleles in the severe forms of dengue (DHF and DSS). In the same study, they also found an association between *HLA-DRB1*0901* and protection against dengue shock syndrome that was specific to Kinh ethnic group (Nguyen et al., 2008). In Malaysia, an HLA investigation in Malays, Chinese and Indian groups revealed ethnicity–specific HLA alleles associated with the disease, despite the groups being closely related (Appanna et al., 2010).

HLA association studies in dengue were also performed in Caribbean and Central America populations. In Cuba, the cohort from DENV2 1997 outbreak led to identify two new HLA class II DRB1 alleles, *DRB1*04* and *DRB1*07*, connected to resistance to all dengue categories (Sierra et al., 2007a). In Mexico, *DRB1*04* allele was also linked to resistance to dengue fever (LaFleur et al., 2002). In another Mexican cohort, DQB1 presented alleles for both resistance (*DQB1*0302* – OR=0.23, p–value=0.011) and susceptibility (*DQB1*0202* – OR=7.0, p–value= 0.012) to dengue fever phenotype (Falcon–Lezama et al., 2009).

More works were done in cohorts from India (K Alagarasu et al., 2013), Brazil (Cardozo et al., 2014), Philippines (Mercado et al., 2015) and Jamaica (Brown et al., 2011), disclosing a little more about HLA association and its dependence on serotype/strain, primary or secondary infection, ethnicity, among other leading factors to dengue development. Despite all these efforts, it is still unclear which alleles play a major role on susceptibility or resistance to dengue and how heterogeneous they can be between ethnicities.

3.6.1.2. Cytokine family

Cytokines are small segregated proteins responsible for cell signaling functions and modulation of cells behavior. Once the organism is infected with dengue, there is a direct correlation between the severity of the symptoms and the increase of cytokines production which is common known as "cytokine storm" (Rattanaburee et al., 2015).

TNF□ is a proinflammatory cytokine allocated in the HLA class III region. It is involved in inflammation and is a vasoactive immunomodulatory produced by monocytes (Vejbaesya et al., 2009). This cytokine is responsible for the enhancement of vascular permeability and, additionally, is upregulated in patients with severe dengue (K.

Alagarasu et al., 2013b). The best studied position within this gene is known as *TNF□-308A* and is allocated in the gene promoter 308 nucleotides upstream the starting transcriptional site (Coffey et al., 2009). Another SNP, *TNF□-238A*, which possibly locates close to a gene repressor site, was also reported to contribute to dengue phenotype. In a Thai cohort, the haplotype containing *TNF□-308A* was found in higher frequency in DHF phenotype during primary infection, yet, it did not reach significant values in comparison to healthy controls or DF patients (Vejbaesya et al., 2009). On the other hand, the haplotype defined with *TNF□-238A* showed a significant positive correlation with DHF patients with secondary infection (Vejbaesya et al., 2009). In 2013, the results from another Thai cohort corroborated the previous findings: despite the lack of significant p-values when comparing DHF against DF or healthy controls, there was an increased frequency of *TNF□-308A* in the individuals with bleeding (Chuansumrit et al., 2013). In the Vietnamese cohort tested for both mutations present on *TNF□*, no significant association was detected from all the comparisons (Loke et al., 2001). Regarding the panorama in the Americas, the analysis of *TNF□-308A* and -238A in a Mexican cohort only evidenced a minimal effect of these alleles in severity outcome (Sanchez-Leyva et al., 2017). The most significant association for *TNF□-308A* with dengue severity was discovered in the Venezuelan population where patients with hemorrhagic manifestations had a higher frequency of these allele in comparison with others, further supported by the higher expression of this cytokine in individuals bearing the -308A allele (Fernández-Mestre et al., 2004). The observations in the Cuban cohort composed by DHF cases and healthy controls were consistent with the data from Venezuela, evidencing the association of *TNF□-308A* high frequency with DHF incidence (Perez et al., 2010). These studies seem to point out that *TNF□* may be an indicator of severity for dengue infections, however data from Malaysia (Sam et al., 2015) and Sri Lanka (Fernando et al., 2015) show that results are not uniform in all populations and, further research should enlighten this relationship with severity.

IL-10 is another anti-inflammatory cytokine usually found in high levels in patients with DHF phenotype, being considered a severity marker. It is also an immunomodulation mediator segregated by monocytes, dendritic cells, T and B lymphocytes. A few studies tested three polymorphisms localized in the *IL-10* promoter (-1082, -819, -592) to access their prevalence across dengue phenotypes. No significant links were found for these SNPs in both Cuban and Venezuelan cohorts (Fernández-Mestre et al., 2004; Perez et al., 2010) when considered individually, but haplotype analyses led to the correlation of ACC/ATA haplotype with the DHF group

in Cuba (Perez et al., 2010). Also in this Cuban cohort, *IL-10* and *TNFα* polymorphisms joint analysis disclosed a predominant haplotype characterized by high *TNFα*/low *IL-10* expression in Cuban DHF patients (Perez et al., 2010). The genetic studies considering this cytokine are still scarce and not elucidative of *IL-10* overall contribution to susceptibility to dengue.

Few polymorphisms present in other cytokines, as *IFN-γ*, *IL-4* and *IL-6*, were genotyped in Cuba (Perez et al., 2010), Vietnam (Loke et al., 2002) and Venezuela (Fernández–Mestre et al., 2004), respectively, with no significant associations being detected. On the other hand, *TGFβ1*, explored in the Cuban cohort, evidenced the G allele from codon 25 as protective against DHF (Perez et al., 2010).

3.6.1.3. Other genes involved in virus entrance and replication in host cells

DC–SIGN (dendritic cells–specific intercellular adhesion molecule–3–grabbing non–integrin) is coded by *CD209* gene and is highly expressed in the surface of immature dendritic cells, acting as a partner recognition receptor (Liu et al., 2017). The involvement of this gene in dengue membrane attachment is a well–established fact given the internalization of DC–SIGN/DENV complexes (Liu et al., 2017). The work developed by Sakuntabhai and collaborators (Sakuntabhai et al., 2005) identified one SNP in this gene promoter, *DCSIGN1*–336, with a strong protective association against dengue fever. An opposite effect was shown by the genotypes GG and GA for the same SNP, which considerably increase the risk of dengue hemorrhagic fever. These evidences led the authors to hypothesize the action of distinct mechanisms behind both phenotypes (Sakuntabhai et al., 2005). The same observations were corroborated in a Taiwanese cohort (Wang et al., 2011), regarding DHF risk, but not for dengue fever, what may be explained by differences in frequencies or due to the system used to classify the clinical cases into dengue fever. Furthermore, they observed increased cell surface expression of DC–SIGN in AG heterozygous for *DCSIGN1*–336 (rs4804803) (Wang et al., 2011). In the Indian population, another SNP from this gene (rs2287886) was connected to DF, despite the lack of statistical significance (Alagarasu et al., 2013). DC–SIGN was also study in three Brazilian cohorts from distinct locations to check for any relationship with dengue phenotypes: it failed to reach statistical significance in Salvador and Belém (Oliveira et al., 2014; Silva et al., 2010), however, contrary to the observations in Thailand and Taiwan, in

Rio de Janeiro cohort (Xavier-Carvalho et al., 2013), G allele was associated with protection against severe dengue.

DENV antibody-dependent enhancement is known to contribute to the severity in secondary infections, due to the presence of non-neutralizing antibodies. It is hypothesized that this class of antibodies binds to the new dengue serotype infecting the organism and theses complexes enter the cells by Fc receptors (Guzman and Vazquez, 2010). *FcRIIa* is the most widespread Fc receptor and it is expressed in hematopoietic cells. It is also a low-affinity IgG receptor that can bind to a few immunoglobulin subtypes, being conditioned by a point mutation at position 131 (rs1801274) coding a nonsynonymous replacement of an arginine (R) for a histidine (H). And while R131 genotype interacts with IgG1 and IgG3, H131 genotype interacts only with IgG2 (García et al., 2010). This polymorphism has been studied to check its influence and relation with dengue outcome. In Pakistan (Mohsin et al., 2015), Cuba (García et al., 2010) and Vietnam (Loke et al., 2002), the results obtained were consistent among them, demonstrating that *FcRIIa*-H131 allele and HH/HR genotypes are associated with increased susceptibility to dengue, mostly to its severe outcome. The allele R131 is less competent for opsonization of IgG2 antibodies, which can lead to protection against severe dengue (Loke et al., 2002). The same was not observed analyzing the Mexican cohort; in fact, authors obtained the opposite results, evidencing a protection towards dengue infection when the individuals carried the same allele (Noecker et al., 2014).

The vitamin D receptor (*VDR*) is responsible for the immunoregulatory effects of 1,25-dihydroxyvitamin D_3 (1,25D_3), including monocytes activation, stimulation of immune response, lymphocyte proliferation and suppression of immunoglobulin production (Loke et al. 2002). *VDR* has been pointed out as a potential risk factor for dengue given its localization in monocytes, one of the cell types that the virus uses for replication (Coffey et al., 2009). Favoring that involvement, a polymorphism at position 352 (rs731236) on the gene was described to be involved in susceptibility to other infections as tuberculosis and hepatitis b virus (Bellamy et al., 1999). The first work to be published in dengue regarding this gene and polymorphism was performed by Loke and collaborators (Loke et al., 2002) whom successfully disclosed a link between T allele frequency and the risk for severe dengue however, these results failed to replicate in the Brazilian (Dettogni et al., 2015) and Indian (Alagarasu et al., 2012) populations. Despite this, another SNP in *VDR*, rs2228570, was found to be correlated with the risk of DHF in the Indian cohort (Alagarasu et al., 2012).

The oligoadenylate synthetase (OAS) proteins belong to interferon (IFN) stimulated genes, provoking an antiviral effect due to their ability in recognizing exogenous nucleic acids and activate antiviral pathways. These enzymes are essential to catalyze ATP into 2'–5' linked OAS, that later is responsible for endoribonuclease L (RNAse L) activation (Choi et al., 2015). This family comprises four genes, containing 10 isoforms, localized on chromosome 12: *OAS1*, *OAS2*, *OAS3*, with one, two and three OAS units respectively, and OAS-like (*OASL*), that lacks the 2'–5' linked OAS activity. OAS involvement in dengue infection is well established and Lin and coworkers (Lin et al., 2009) identified that the isoforms *OAS1*-p42, *OAS1*-p46 and *OAS3* (there is only one isoform in the latter) confer resistance against DENV2 and are determinant for the infection outcome. A recent study (Simon–Loriere et al., 2015) also concluded for a similar antiviral effect against DENV2 for both *OAS1* isoforms, while the experiments using DENV1 indicated that only *OAS1*-p42 had some antiviral action in this serotype. Considering OAS3, two nonsynonymous variants, *OAS3*_K18R (rs1859330) and *OAS3*_S381R (rs2285933), were evaluated functionally to access their antiviral power, revealing that *OAS3*–S381 was able to decrease viral production, especially when infection was caused by DENV2, reinforcing the serotype differences also found for *OAS1* (Simon–Loriere et al., 2015). The functional study was complemented by genotyping of the *OAS1* splicing variant (rs10774671) and both *OAS3* in Thai cohorts, revealing the following: *OAS1* splicing variant (generating p46 and not p48 and p52) was associated with an increased risk of developing plasma leakage and shock only in infections with DENV2; *OAS3*–R381 allele was associated with protection against shock also considering DENV2 infections. Another work (Alagarasu et al., 2013) focused on an Indian cohort and observed an increment of GG *OAS2* haplotype for all case groups and, the combination of six SNPs (two per gene) allowed the identification of two haplotypes significantly associated with DHF phenotype. Among the SNPs genotyped in India, only *OAS1* splice variant (rs10774671) was common with the work developed by Simon–Loriere et al. (2015), however, no significant association was detected for that SNP alone.

3.6.2. Classical genome-wide approaches applied to dengue

Currently, there is one classical GWAS published for dengue infection, focused in the dengue shock syndrome phenotype (Khor et al., 2011). It includes 2,008 children classified as DSS and 2,018 cord blood controls, whom were genotyped for more than

650 000 SNPs included in the Illumina 660W BeadChip. After PCA-correction for population structure, linked SNPs in two genes, *MICB* and *PLCE1* (Figure 14), passed the genome-wide p-value threshold ($>1*10^{-8}$), and were replicated in another Vietnamese cohort, composed by 1,737 DSS individuals and 2,934 controls. *MICB* locus is localized within the *HLA* complex, outside classes I and II, and is indirectly involved in cytokines pathways. On the other hand, *PLCE1* is connected with the nephrotic syndrome, that displays similar symptoms to severe dengue (Khor et al., 2011).
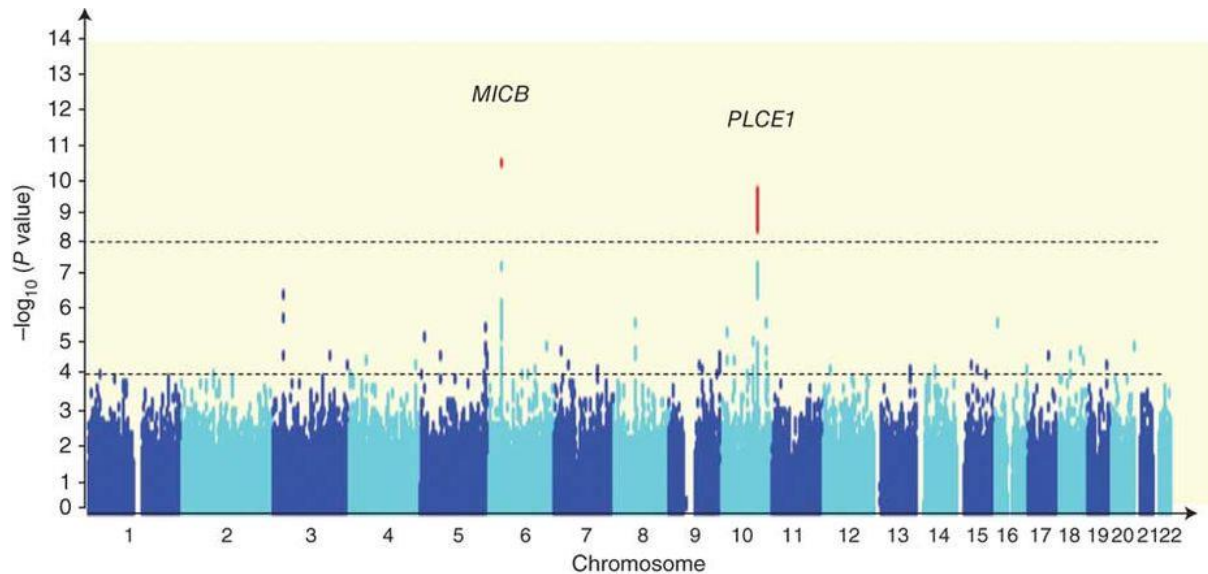
Figure 14. Manhattan plot containing the data from discovery and replication cohorts combined. The SNPs with p-value$<10^{-8}$ are identified in red and belong to *MICB* and *PLCE1* genes. Adapted from (Khor et al., 2011).

Whitehorn and coworkers (Whitehorn et al., 2013) explored two SNPs, rs3132468 and rs3740360 from *MICB* and *PLCE1*, respectively, and confirmed the significant association of these SNPs with DSS in children (*MICB*: OR=1.42, p-value=0.0014; *PLCE1*: OR=0.77, p-value=0.0094) and extended the comparison to non-DSS where the same association was observed (*MICB*: OR=1.15, p-value=0.0012; *PLCE1*: OR=0.92, p-value=0.018). Also in the Southeast Asian nexus, *MICB* (rs3132468) and *PLCE1* (rs3765524) were genotyped in Thai cohorts, showing to be associated with an increased risk of DSS (*MICB*: OR=1.58, p-value=0.0213; *PLCE1*: OR=1.49, p-value=0.0252), when compared versus non-DSS cases, in resemblance to Vietnam, indicating a role of these genes in severe dengue susceptibility (Dang et al., 2014).

3.6.3. Combining admixture mapping and association analysis

Early in 2017, Sierra and coworkers (Sierra et al., 2017) published the first study combining admixture mapping and association analysis in Cuban dengue cohorts genotyped for 2.5 million SNPs. Cuba has the ideal conditions for an ethnicity-based genome wide search due to the admixed nature of its population, with two predominant ancestral components (in addition to a residual Native American) – European and African from the colonizers and enslaved Africans, respectively, that began to mix there 500 years ago. As already described, the global association test revealed a significantly higher proportion of African ancestry in the asymptomatic/control group in comparison with the DHF, being a strong evidence of the protective role of African ethnicity against the development of severe dengue, as has been reported before based on epidemiologic data (Sierra et al., 2007b).

For mapping the candidate genes, Sierra et al. (2017) applied two strategies. First, an ancestry-controlled association test, where the case and control samples were carefully paired controlling for their African genetic proportion. The strongest sign was detected on chromosome 3, for six linked SNPs in Oxysterol Binding Protein Like 10 (*OSBPL10*) *locus*. This gene belongs to the OSBP-related protein (ORP) family, which participate in the non-vesicular lipid transport in ER and Golgi apparatus (Kentala et al., 2016). This protein family was already explored in the context of rhinovirus infection, where they participate in the cholesterol counter-flow involving ER and Golgi membranes (Roulin et al., 2014). Two haplotypes defined by the six alleles have opposing major frequencies in European and African populations, and, respectively, in Cuban DHF and asymptomatic/control groups. The African-protective haplotype seems to lead to a lower *OSBPL10* expression that matched the shRNA assay results of a significant decrease of dengue replication in cell lines with *OSBPL10* knockdown (Sierra et al., 2017).

Secondly, a complementary investigation of admixture blocks along the chromosomes in the Cuban cohorts identified Retinoid X Receptor Alpha (*RXRA*) gene-containing region has having a significantly higher frequency in African ancestry in asymptomatic/controls versus DHF. This gene is expressed in macrophages and liver cells, being a very important nuclear receptor that forms homodimers and heterodimers with other nuclear receptors, such as Liver X receptor (*LXR*), Peroxisome proliferator-activated receptor (*PPAR*), Constitutive Androstane receptor (*CAR*), and Vitamin D receptor (*VDR*), controlling several pathways and gene expression (Roszer et al., 2013). Ma and collaborators (Ma et al., 2014) characterized the interaction

between *RXRA* and type I *IFN*: *RXRA* downregulation is controlled during viral infections by the innate immune system, leading to *IFN* maximal performance. On other hand, the *RXRA* expression recovery occurs at the end of infection conducing to *IFN* suppression, demonstrating the important role of *RXRA* in the modulation of immune response (Ma et al., 2014). In the Cuban context, authors (Sierra et al., 2017) showed that *RXRA* expression varied over the different stages of infection in accordance to what has been reported (Ma et al., 2014). These two candidate genes identified in Cuba, *RXRA* and *OSBPL10*, , may be interplaying in the LXR/RXR pathway (Sierra et al., 2017), which controls lipid metabolism and cytokine production, but further investigation is required to clarify its fully involvement in dengue disease.

3.7. Transcriptome studies

Transcriptomic studies are being used to infer expression profiles during the various stages of infection or for distinct dengue phenotypes, allowing to identify genes that have a significant variation in expression levels. Most of the transcriptomic studies considering dengue were performed in whole-blood or some particular cell type present in the blood of patients. Notice that this kind of methodology only identifies genes that have changes in expression, and these changes can be secondarily controlled by other genes, which are the ones really associated with the phenotype and that will be identified in GWAS (such as *RXRA* discovered in Cuba). It is a complementary strategy, not replacing GWAS.

The first investigations were directed to uncover the pathways differentially expressed in DSS patients versus controls in Cambodia (Devignot et al., 2010) and Vietnam (Long et al., 2009; Simmons et al., 2007). The majority of the pathways and related genes identified in those works belonged to innate immunity mechanisms, as pro-inflammatory defense and IFN activation, apoptotic and host-lipid metabolism pathways.

Samples from dengue patients were collected in Venezuela on the early and late acute phases (Sun et al., 2013). In the early acute phase, the upregulated pathways were mostly related with innate immune response such as inflammatory response, cytokine signaling, interferon-gamma-mediated signaling and response to virus and double-stranded RNA binding. In the late acute phase, the upregulated pathways targeted aspects of cell cycle like cell division, mitosis, spindle organization and also DNA replication (Sun et al., 2013).

Another work in Thai, included DF, DHF, convalescent and control individuals (Kwissa et al., 2014) to look for pathways involved in the immune response to dengue phenotypes. Their findings evidenced distinct gene expression signatures associated with the disease progression: while genes encoding proinflammatory mediators and type I IFN related proteins exhibit higher expression during early acute illness (2 to 4 days after infection), later on the onset of symptoms (5 to 8 days), genes belonging to activated T and B cells provide the strongest signal. Comparing DF and DHF, they did not find any differentially expressed pathway between the two phenotypes.

A recent transcriptome study focused on the pathways differentially expressed in asymptomatic people and clinical dengue patients (Simon-Loriere et al., 2017a) revealed that antigen presentation and T cell and B cell activation related genes were the most upregulated pathways in asymptomatic, while clinical dengue was characterized by the increased secretion of anti-DENV antibodies and plasma cell development pathways.

II. Aims

The increasing number of dengue infections worldwide entails considerable economic costs and human suffering every year. The real disease burden is difficult to estimate, given the lack of follow-up of patients, misdiagnosis due to other tropical diseases with similar symptoms, and the uncertain number of asymptomatic that function as virus reservoirs. In addition, there is no vaccine or treatment available to efficiently control the disease. The knowledge of the mechanisms underlying the development of dengue in the human host is also limited. Genetic approaches have been applied to shed light on human susceptibility to dengue, especially given the evidence that dengue manifestations vary with host ancestry.

Southeast Asia is one of the most affected regions worldwide, with most countries having more than one serotype circulating in endemic-epidemic cycles. The main aim of this work is to disentangle the contribution of host ancestry in dengue outcomes, with a special focus on Southeast Asian nexus. We proposed to:

1.   **Investigate the worldwide risk of dengue using ten variants present in** *DCSIGN/CD209, TNFA, PLCE1, MICB, OAS1, OAS3* **and** *FCGR2A/CD32* **genes**. First, we obtained and analyzed the diversity patterns and selection evidence for the ten genetic markers by genotyping 15 populations and adding up published information to a total of 61 populations across the globe. Second, we performed meta-analyses incorporating new cohorts from Cambodia, Vietnam and Brazil and all the case-control cohorts available to evaluate the susceptibility or resistance conferred by the ten SNPs, pooling their ORs. Third, we combined both sets of information to obtain the worldwide genetic risk conferred by these variants to dengue disease.

2.   **Unravel the host ancestry influence in susceptibility/resistance to dengue infection in Southeast Asia.** We screened ~700,000 SNPs in dengue Thai cohorts and controls, the second Southeast Asian dengue set genotyped at a genome level after the published Vietnamese data. We started with the analysis of population structure of Thai and Vietnamese dengue samples to access the main ancestral components in these susceptible populations. Then, we performed the local ancestry mapping to evaluate the association of the identified ancestral components to dengue outcomes (dengue fever and dengue shock syndrome) and map candidate alleles/loci. The newly identified candidates in the Thai cohorts were further genotyped in a replication cohort and evaluated as potential eQTLs. The genetic risk to dengue fever and dengue shock syndrome conferred by these candidates were calculated for the various

worldwide regions. Co-localization assays were conducted for one of the candidate genes, for which *in silico* evidence pointed out a possible interaction with DENV NS5 protein;

3. Perform functional assays to access the influence of newly discovered candidate genes in DENV infection. We focused our functional validation on two candidate genes, *PPP2R5E* and *GRIP1*, pinpointed in the comparisons between dengue fever cases and healthy controls from the Thai cohorts. We established stable cell lines overexpressing individually the referred proteins using constitutive mammalian expression, and also knockout stable cell lines by taking advantage of the CRISPR/Cas9 system. We conducted infection assays in the transformed cell lines and evaluated the influence of the alterations by quantifying rate of infection and viral production.

III. Research work

Study of the host ancestry influence in susceptibility/resistance to dengue infection in Asia

Paper I – Population genetics–informed meta–analysis in seven genes associated with risk to dengue fever disease

Paper II – Joint ancestry and association test indicate two distinct pathogenic pathways involved in classical dengue fever and dengue shock syndrome

Paper I – Population genetics-informed meta-analysis in seven genes associated with risk to dengue fever disease

Population genetics-informed meta-analysis in seven genes associated with risk to dengue fever disease

Marisa Oliveira[1,2,3,4], Diana P. Saraiva[1,2,4*], Bruno Cavadas[1,2,3], Verónica Fernandes[1,2], Nicole Pedro[1,2], Isabelle Casademont[4,5], Fanny Koeth[4,5], Farida Alshamali[6], Nourdin Harich[7], Lotfi Cherni[8,9], Beatriz Sierra[10], Maria G. Guzman[10], Anavaj Sakuntabhai[4,5,11] and Luisa Pereira[1,2,12**]

[1] i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200–135 Porto, Portugal

[2] Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), 4200–135 Porto, Portugal

[3] Instituto de Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto, 4050–313 Porto, Portugal

[4] Institut Pasteur, Functional Genetics of Infectious Diseases Unit, 75724 Paris Cedex 15, France

[5] Pasteur Kyoto International Joint Research Unit for Integrative Vaccinomics, Kyoto, Japan

[6] General Department of Forensic Sciences and Criminology, Dubai Police General Headquarters, PO Box 1493, Dubai, United Arab Emirates

[7] Laboratoire des Sciences Anthropogénétiques et Biotechnologies, Départment de Biologie, Université Chouaïb Doukkali, El Jadida 24000, Morocco

[8] Laboratory of Genetics, Immunology and Human Pathology, Faculté de Sciences de Tunis, Université de Tunis El Manar, Tunis 2092, Tunisia

[9] Tunis and High Institute of Biotechnology, University of Monastir, 5000 Monastir, Tunisia

[10] Virology Department, PAHO/WHO Collaborating Center for the Study of Dengue and its Vector, Pedro Kourí Institute of Tropical Medicine (IPK), 601 Havana, Cuba

[11] CNRS UMR2000: Génomique évolutive, modélisation et santé (GEMS), Paris, France

[12] Faculdade de Medicina da Universidade do Porto, Portugal

* Present address: CEDOC, Nova Medical School, Faculdade de Ciências Médicas da Universidade Nova de Lisboa, Lisbon, Portugal

Short title: Meta-analysis in seven candidate genes for dengue disease

**Corresponding author: Luisa Pereira (luisap@ipatimup.pt)

Abstract

Population genetics theory predicted that rare frequent markers would be the main contributors for heritability of complex diseases, but meta-analyses of genome-wide association studies are revealing otherwise common markers, present in all population groups, as the identified candidate genes. In this work, we applied a population-genetics informed meta-analysis to 10 markers located in seven genes said to be associated with dengue fever disease. Seven markers (in *PLCE1*, *CD32*, *CD209*, *OAS1* and *OAS3* genes) have high-frequency and the other three (in *MICB* and *TNFA* genes) have intermediate frequency. Most of these markers have high discriminatory power between population groups, but their frequencies follow the rules of genetic drift, and seem to have not been under strong selective pressure. There was a good agreement in directional consistency across trans-ethnic association signals, in East Asian and Latin American cohorts, with heterogeneity generated by randomness between studies and especially by low sample sizes. This led to confirm the following significant associations: with DF, odds ratio of 0.67 for *TNFA*-rs1800629-A; with DHF, 0.82 for *CD32*-rs1801274-G; with DSS, 0.55 for *OAS3*-rs2285933-G, 0.80 for *PLCE1*-rs2274223-G and 1.32 for *MICB*-rs3132468-C. The overall genetic risks confirmed sub-Saharan African populations and descendants as the best protected against the severer forms of the disease, while Southeast and Northeast Asians are the least protected ones. European and close neighbours are the best protected against dengue fever, while, again, Southeast and Northeast Asians are the least protected ones. These risk scores provide important predictive information for the largely naïve European and North American regions, as well as for Africa where misdiagnosis with other hemorrhagic diseases is of concern.

Keywords: Dengue disease; meta-analysis; East Asian and Latin American cohorts; worldwide diversity; global risk

Abbreviations:

DENV – dengue virus; DF – Dengue Fever; DHF – Dengue Hemorrhagic Fever; DSS – Dengue Shock Syndrome; EHH – Extended Haplotype Homozygosity; GWAS – genome-wide association studies; MAF – minimum allele frequency; SNP – Single Nucleotide Polymorphism; WHO – World Health Organization

Declarations of interest: none

1. Introduction

Dengue fever (DF) is amongst the 17 neglected diseases prioritized by the World Health Organization (WHO) in the year 2012 (WHO, 2012). The list will for sure enlarge as new epidemic agents emerge, such as chikungunya and Zika viruses (Gautret and Simon, 2016). Less than 10% of the global health stipends are destined to the treatment of these diseases, which infect around 90% of sick individuals in the world (Morel, 2003), contributing immensely to the global health burden.

Dengue is an acute systemic disease caused by four virus types (DENV1–4) which are transmitted to humans by an *Aedes* mosquito vector (mainly *Ae. aegypti* and *Ae. albopictus*). Most DENV infections are asymptomatic, but in some individuals, it can lead to a wide range of clinical symptoms, from mild fever to life threatening dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) (Guzman and Harris, 2015). Usually, but not exclusively, the severe disease appears when the patient is infected by a heterologous DENV, in a non-completely understood cross-reaction with antibodies against the previous DENV type (Vaughn et al., 2000). Control efforts have failed in counterbalance the effects of human international travel and global warming, which are introducing new vectors and pathogens into novel geographic areas (Messina et al., 2014). *Ae. aegypti* is well adapted to urban environments, breeding in any container of water. This species was able to re-establish in countries of South and Central America, from where it had been eliminated in the mid-20th century, and it also dispersed throughout Southeast Asia during and after World War 2 (Messina et al., 2014). The geographical range of the secondary dengue vector, *Ae. albopictus*, has also expanded substantially over the past 30 years, being the major potential vector of DENV in Europe (Kraemer et al., 2015). This increased movement of people, vector and virus will also raise the co-circulation of serotypes, leading to a higher risk of sequential infections and severe disease.

In this high dispersion scenario, it is relevant to estimate the worldwide risk of dengue fever. Bhatt et al. (Bhatt et al., 2013) undertook an exhaustive assembly of known records of dengue occurrence worldwide (8,309 geo-located records), and used a formal modelling framework to map the global distribution of dengue risk based on evidence. Dengue was predicted to be ubiquitous throughout the tropics, with the highest risk zones in the Americas (14%) and Asia (70%; India alone contributed 34%), varying locally by the influence of rainfall, temperature and the degree of urbanization. Global values of 390 million (95% credible interval [284–528]) dengue infections per year, of which 96 million [67–136] are symptomatic, were estimated. These values surpass more than thrice the dengue burden estimates by WHO, of 50–

100 million infections per year (WHO, 2009). Predicted risk in Africa, although unevenly distributed, was much more widespread than previously suggested, 16 [11-22] million infections, or 16% of the global total, nearly equivalent to that of the Americas. This disparity seems to support the fact that dengue disease can be masked by symptomatically similar illnesses in Africa (Stoler et al., 2014). Oceania contributed less than 0.2% of global apparent infections. The estimated 294 [217-392] million unapparent worldwide infections, although having no immediate implications for clinical management, is a huge potential viral reservoir, as asymptomatic individuals can also transmit the virus (Duong et al., 2015), having implications in designing vaccination campaigns.

A very important factor to consider in the worldwide risk evaluation to dengue disease is the host genetic component (reviewed recently in (Xavier-Carvalho et al., 2017)). A predictive map based on genetic risk information could provide useful insights into the still largely naïve regions of Europe and North America, and regions where registries/diagnosis mix-up predominates as in Africa. *Homo sapiens* is a young species dating around 200,000 years (McEvoy et al., 2011; Rito et al., 2013), with a single unique origin in Africa. The structure of the modern human population was initiated at 60,000 years ago, in the out-of-Africa event when a small group of Eastern Africans migrated towards the Arabian Peninsula/Near East, and thence, first to Asia (arriving in Australia at around 50,000 years ago) and then (from 45,000 years ago) to Europe (Soares et al., 2009). The bottleneck in the origin of all Asians and Europeans explains their high homozygosity when compared with the high heterozygosity observed in Africans, so that the former have proportionally more deleterious genetic variation than the latter (Lohmueller et al., 2008). The out-of-Africa event was probably fuelled by the first noticeable population expansion (Soares et al., 2012), but continuous expansion of the human population begun in the Holocene, around 12,000 years ago, when climatic conditions became similar to present-day, also promoting the cultural Neolithic revolution of agriculture and domestication (Atkinson et al., 2008). Population expansions lead to increased amounts of rare (low MAF, minimum allele frequency) and population-restricted variants. Thus, 73% of protein-coding variants and 86% of predicted-deleterious variants in humans arose only in the past 5,000-10,000 years, and are rare and population specific (Fu et al., 2013). A considerable lower proportion of variants are shared between continents, presenting higher MAF (>1%), and being older (Abecasis et al., 2012). Through genetic drift, variants will be randomly lost or fixed in the population along time, or, if they confer susceptibility/resistance to infectious

diseases, they can be under the effect of selection (Karlsson et al., 2014). Negative selection reduces frequencies of SNPs, which are maintained as rare variants, while positive selection increases MAF of protective alleles, generating heterogeneity in selected polymorphisms among populations along time, as pathogens emerge and become important agents in specific geographical regions. The factors contributing to population divergence, in terms of (1) variants (especially the rare ones), (2) MAF of common variants, (3) linkage disequilibrium (LD) patterns (tag SNPs vary in LD with the causative SNP), (4) biological adaptations, (5) phenotypic prevalence, and (6) effective sizes, have important implications in the power and the reproducibility of association studies.

Association studies in dengue disease began by pinpointing potential candidate genes involved in the immune response to viral infection in Asian and Latin American populations (Coffey et al., 2009). One of the first identified candidates was DC-SIGN (dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin 1) protein, coded by *CD209* gene. DC-SIGN is an essential molecule for the interaction of virus particles with a cellular receptor responsible for its internalization in cells otherwise refractory to infection by the four DENV serotypes (Lozach et al., 2005). In accordance, cell surface DC-SIGN expression correlates with DENV infection rates in vitro, while antibodies against DC-SIGN can block infection (Navarro-Sanchez et al., 2003; Tassaneetrithep et al., 2003). At the genetic level, the G allele at position 336 (rs4804803) of DC-SIGN was associated with a dominant protection against DF (odds ratio, OR=4.90; p=2×10^{-6}) but not against DHF, while among individuals with dengue, genotypes GG and GA strongly increased the risk of contracting DHF versus DF (OR=5.84; p=1.4×10^{-7}) in three Thai dengue cohorts (Sakuntabhai et al., 2005). Further functional studies showed that dendritic cells from AG-heterozygous have a significantly higher cell surface DC-SIGN expression than AA-homozygous and a higher production of TNF-α, IL-12p40, and IP-10 in response to DENV infection (Sakuntabhai et al., 2005). However, against expectations, DENV replication was significantly lower in AG individuals, probably due to the enhanced production of IP-10 (Sakuntabhai et al., 2005). Replication was positive in Taiwanese cohorts (Wang et al., 2011), for GG/GA genotypes when comparing DHF with DF, other febrile diseases and controls (p=0.003; 3×10^{-5} and 0.001), with ORs in the range 2-5. Replication failed in India ((Alagarasu et al., 2013a); although other markers around this SNP presented significant association values) and Mexico (Noecker et al., 2014; Vargas-Castillo et al., 2017). More difficult to explain was the closeness to significant association of GG genotype in Rio de Janeiro, Brazil, but in the other direction, being

protective against severe dengue (Xavier-Carvalho et al., 2013) (another Brazilian cohort from Salvador da Bahia showed no association (Silva et al., 2010)).

Another set of genes associated with dengue, and for which there is already functional evidence, is the oligoadenylate synthetase family (*OAS1*, *OAS3* and *OAS2* genes, in this order in chromosome 12), which consists in interferon-induced cytoplasmic double-stranded RNA (dsRNA) sensors. Upon activation by dsRNA, OAS converts ATP to oligoadenylate, which in turn activates the latent ribonuclease RNAseL that catalyses the degradation of viral and host RNA, thus reducing protein synthesis in the cell (Peisley and Hur, 2013). DENV infection leads to an upregulation of OAS expression in endothelium HUVEC cells (Warke et al., 2003), and the stable silencing of RNAseL in A549 lung cells increases DENV titers by 10- to 40-fold in agreement with its antiviral activity (Acosta et al., 2014). Some *OAS* polymorphisms (*OAS1* – rs1131454 and rs10774671, *OAS3* – rs2285932 and rs2072136, and *OAS2* – rs15895 and rs1732778) were correlated with clinical severity of dengue in India (Alagarasu et al., 2013b), when associated in haplotypes (*OAS2* G-G all patients vs. controls, p-corrected=0.012, OR=1.73 95% CI 1.16-2.59; *OAS3-OAS2* C-G-A-G all patients vs. controls, p-corrected =0.0486, OR=0.09, 95% CI 0.00-0.64). Even more interesting, a recent study in Thai dengue cohorts showed that the influence of the *OAS* polymorphism is probably dengue virus strain-dependent: OAS3-G allele (Ser381Arg; rs2285933) conferred protection against shock (OR=0.37; p<0.001) when the infection was caused by DENV2 (OR=0.13; p=0.007) but not when due to DENV1 (Simon-Loriere et al., 2015). Also, by conducting a multivariate analysis, there was a serotype-specific effect of the OAS1-G splicing variant (rs10774671) on the risk of plasma leakage (p=0.019) in DENV-2 infections (OR=3.17; 95% CI, 1.31-7.68) but not in DENV-1 infections (OR=0.81; 95% CI, .37-1.75), while no statistical significant associations were found for the *OAS3*-rs1859330 missense marker (Arg18Lys) in the Thai population.

CD32 or *FCGR2A* (Fc fragment of IgC receptor IIa) gene, together with other Fcγ receptors expressed in monocytes, macrophages and dendritic cells, mediate antibody dependent enhancement (ADE). This phenomenon occurs during secondary infection with a heterologous serotype, when antibodies formed against primary infecting serotype may enhance the entry of a different serotype and may contribute to disease severity (Alagarasu et al., 2015). A study from Cuba on DENV4 infected patients and FCGR2A p.R131H polymorphism reported that the R allele and R/R genotype are associated with protection to symptomatic dengue while the H allele and H/H genotype are associated with both DF and DHF (Garcia et al., 2010). Genotype

frequencies of FCGR2A p.R131H were not statistically different between DF, DHF and controls in general in India, but the RR genotype was observed significantly increased in DENV cases with thrombocytopenia (Alagarasu et al., 2015).

TNFA (tumor necrosis factor alpha) gene codes a proinflammatory cytokine, mainly secreted by macrophages in response to bacterial or viral infections. Functional assays have shown that transient suppression of TNF-α production during the early period of ADE infection in THP-1 cells promoted the initiation of DENV replication (Chareonsirisuthigul et al., 2007). These results were contradicted in another work, where TNF-α at high and medium concentration inhibited DENV replication in human dendritic cells, a primary DENV target cells (cited in (Sam et al., 2015)). The promoter −308A allele has been related to enhanced TNF-α gene transcriptional activity (Kroeger et al., 1997), and different studies have associated it with diverse dengue phenotypes or failed to do so: protective against development of DHF/DSS in Malaysia (Sam et al., 2015); not associated in Mexico (Garcia-Trejo et al., 2011); dominant model marginally protective against DF in Brazil (Santos et al., 2017); risk factor for DHF in Venezuela (Fernandez-Mestre et al., 2004) and Cuba (Perez et al., 2010).

As technological improvements enabled genotyping of thousands of SNPs at once in a chip, association surveys upgraded to unbiased genome-wide association studies (GWAS). The thousands of SNPs screened in GWAS elevate the statistical burden, pushing the significance threshold to the level of 10^{-8}, implying the genotyping of thousands of cases and controls. A common variant identified in the discovery population, if rare in a replication population, will imply a greater sample size to achieve comparable statistical power to confirm the significant association (Rosenberg et al., 2010). And, definitely, trans-ethnic studies are needed, as it is evident that no single population is sufficient to fully uncover the variants underlying a certain disease in all populations (Rosenberg et al., 2010). The only traditional GWAS published so far in dengue context was performed in Vietnamese children (2,008 dengue shock cases and 2,018 controls; replicated in 1,737 cases and 2,934 controls) and showed significant association of polymorphisms within *MICB* (MHC class I polypeptide-related sequence B) and *PLCE1* (phospholipase C, epsilon 1) genes with dengue shock syndrome (Khor et al., 2011). The authors hypothesise that the possible protective role of *PLCE1* gene is through maintenance of normal endothelial barrier function, as mutations in this gene are associated with nephrotic syndrome, a kidney disorder characterised by dysfunction of the glomerular basement membrane resulting in proteinuria and hypoproteinemia that, when severe, leads to reduced vascular oncotic pressure and edema. Still, functional assays are needed to fully

ascertain the mechanism by which this gene may be involved in dengue disease. Several linked *PLCE1* SNPs presented significant association values with DSS in Vietnamese children (Khor et al., 2011), and further works have confirmed association of some of these SNPs in other dengue cohorts (such as rs3740360 with dengue without shock in Vietnamese children and adults (Whitehorn et al., 2013); rs3765524 with DSS in Thai children (Dang et al., 2014)). An *in silico* assay predicted that rs2274223 (H1619R) is deleterious as it induces structural changes in the C2 domain of PLCE protein (Taqi et al., 2016). As occurs in most GWAS-discovered candidate markers, the ORs for these SNPs have a limited impact, varying around the value of 0.76–0.85. *MICB* gene is located on the highly gene dense HLA region in chromosome 6, and encodes a stress-inducible activating ligand for the NKG2D type II receptor on natural killer and CD8+ T cells. The ligation of NKG2D by MICB protein stimulates antiviral effector functions in natural killer cells, including cytokine expression and the cytolytic response (Muntasell et al., 2010). The rs3132468-C risk allele is associated with lower mRNA expression (Whitehorn et al., 2013) and may lead to dysfunctional natural killer and/or CD8+ T cell activation early in infection, resulting in a higher viral burden *in vivo* (Khor et al., 2011). However, given the location of this gene in the rich HLA-A region, it is still possible that its association with dengue fever is due to hitchhiking.

Ethnicity is traditionally considered a confounding factor in association studies, leading to false-positives when cases and controls are not correctly paired for it, and a principal component-based correction is usually applied in GWAS (Price et al., 2006). Nevertheless, ethnicity is an important source of variability to be accounted for, and not just discarded as noise. In this work, we applied population genetics-informed meta-analyses to evaluate the genetic risk/protection conferred by 10 SNPs across seven genes mostly associated with dengue in the literature (*PLCE1*, *TNFA*, DC-SIGN/*CD209*, *OAS1*, *OAS3*, FCGR2A/*CD32* and *MICB*; Table 1). We surveyed the literature for case-control cohorts (Alagarasu et al., 2013a; Alagarasu et al., 2013b; Alagarasu et al., 2013c; Dang et al., 2014; Dang et al., 2016; Fernandez-Mestre et al., 2004; Fernando et al., 2015; Garcia-Trejo et al., 2011; Garcia et al., 2010; Loke et al., 2002; Oliveira et al., 2014; Sakuntabhai et al., 2005; Sam et al., 2015; Sanchez-Leyva et al., 2017; Santos et al., 2017; Simon-Loriere et al., 2015; Vargas-Castillo et al., 2017; Wang et al., 2011; Whitehorn et al., 2013; Xavier-Carvalho et al., 2013), took advantage of available datasets chip-genotyped (Khor et al., 2011; Oliveira et al., 2018; Sierra et al., 2017), and added a few case-control cohorts of our own

(mostly Cambodia, Vietnam and few Brazilian individuals). Then this information was used in genetic risk evaluation at the worldwide scale.

2. Material and Methods

2.1. Genotyping and datasets

The 10 SNPs were included in an OpenArray™ ordered from Life Technologies™. Master mix and DNA with concentration higher than 10 ng/µl were automatically transferred from a 384-well plate to the array in AccuFill™ (Life Tecnhologies™) by capillarity. The arrays were submitted to amplification on a QuantStudio™ 12K Flex (Life Technologies™). The results were analyzed in TaqMan® Genotyper software (Life Technologies™).

We screened the allele frequencies in our collection of samples from several worldwide populations. These population samples were collected from individuals at random, healthy in the moment of collection, in which the only selection criteria was the affiliation in that population for at least three generations back in time. Sampled individuals gave informed consent and the Ethics Committee from University of Porto guaranteed accordance to global ethical guidelines (17/CEUP/2012). Around 1,312 individuals from 15 countries were genotyped in this study. Other publically available population data were used for extracting the allele frequency information: 1000 Genomes database, phase 3 (Auton et al., 2015), composed of 2,504 complete genomes from 26 populations; 35 Koreans (Zhang et al., 2014); 100 Southeast Asian Malays (Wong et al., 2013); 498 Dutch (Consortium, 2014). A few genome-wide SNP chip based population studies were also consulted (Hellenthal et al., 2014; Henn et al., 2012; Li et al., 2008) as well as Maasai data (n=90) from HapMap (Consortium, 2005). Thus, the number of individuals successfully genotyped for each SNP was between 4,055 and 4,830, and of populations between 43 and 61. The maximum number of populations per population group was: 11 European; 3 North African; 11 Sub-Saharan African; 11 Southwest Asian; 8 South Asian; 5 Northeast Asian; 4 Southeast Asian; 2 North American; 1 Caribe; and 5 Latin American. Data are presented in Supplemental Tables A.1–A.10.

A few control-case samples from Cambodia, Vietnam and Brazil were also genotyped for these markers, and data are presented in Table A.11. Written informed consent was obtained from all subjects or, in case of individuals under 18 years of age, from their parents or tutors. Local ethics committees approved the protocol.

2.2. Population genetics and selection analyses

Hardy-Weinberg equilibrium at each locus for each population and F_{ST} genetic distances between populations were calculated in the software Arlequin v 3.0

(Excoffier et al., 2005). Heat maps, done in R package fields, represent the average F_{ST} genetic distances between and within (except same population comparison) population groups and average MAFs for the 10 SNPs in each population group. To visualize the geographical distribution of allele frequencies, we constructed interpolation maps using the "Spatial Analyst Extension" of ArcMap version 10.3.1 for Desktop (http://desktop.arcgis.com/en/arcmap/). We used the "Inverse Distance Weighted" (IDW) option with a power of two for the interpolation of the surface. IDW assumes that each input point has a local influence that decreases with distance. The geographic location used is the centre of the distribution area from which the individual samples of each population were collected.

Mantel tests, to evaluate the correlation between F_{ST} genetic distance matrices and great-circle geographic distance matrices were conducted in R using package Ape (Paradis et al., 2004) and Mantel correlograms were obtained in mpmcorrelogram, by using 10,000 permutations. W Kendall's coefficient of concordance and Friedman' chi-square statistic tests were registered, while in the correlograms the rM referring to the vector with the computed Mantel correlations for each of 12 distance classes (and information on significant p-values after a progressive Bonferroni correction) were plot.

We used the package REHH (Gautier and Vitalis, 2012) to evaluate natural or artificial selection in the candidate SNPs, through the extended haplotype homozygosity (EHH) measure described by Sabeti et al. (Sabeti et al., 2002). This measure was calculated for each SNP in all 1000 Genomes (Abecasis et al., 2012) populations belonging to the three main human population groups: sub-Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan). 5Mb phased data around the focal SNP from 1000 Genomes database were inputted in REHH. For comparison purposes, we also obtained this measure for *MCM6* SNP rs4988235, which is known to be under strong positive selection in Northern European populations (reviewed in (Gerbault, 2013)).

By using the 1000 Genomes project populations of Great Britain, China-Han and Yoruba, as proxies for European, East Asian and African population groups, we checked the linkage disequilibrium (LD) for the two SNPs in *PLCE1*, two SNPs in *MICB* and three SNPs in *OAS1/3* in Haploview v 4.2 (Barrett et al., 2005).

2.3. Meta-analyses

Meta-analyses were conducted in the R package Metafor (Viechtbauer, 2010). 2x2 contingency tables for MAF, and recessive and dominant genotypes were used as inputs for OR calculations in the following comparisons (whenever possible due to sample size/number of studies): DF *vs* controls; DHF *vs* controls; DSS *vs* controls; symptomatic *vs* controls. In a few studies, the control groups included asymptomatic individuals. Funnel and radial plots were obtained and checked to test for publication biases, and heterogeneity across studies was also assessed through Cochran's Q and I^2 statistics. Pooled ORs were obtained by using three models: an inverse–variance weighted fixed–effects model; Der Simonian–Laird estimator as a random–effects model; a restricted maximum–likelihood estimator as another random–effects model. A fixed–effect meta–analysis assumes that the true effect of intervention is equal in all studies (that is, fixed across studies), implying that the observed differences among study results are due solely to chance, i.e. that there is no statistical heterogeneity. A random–effects meta–analysis assumes that the effects being estimated in the different studies are not identical, but are random following some distribution.

2.4. Genetic risk

The dengue disease genetic risk scores for DF and DHF+DSS were calculated by multiplying each individual's risk allele count for each locus by the reported beta coefficient for that polymorphism and summing the product for all corresponding loci (as in (Marden et al., 2014)). The beta parameter for each SNP corresponds with the natural log of the ORs.

3. Results and Discussion

3.1. Worldwide diversity for the 10 SNPs

The 10 SNPs associated with dengue fever can be divided into two groups in terms of frequencies and distribution across the globe (Fig. 1). The SNPs *PLCE1*–rs2274223-G, *PLCE1*–rs3765524-T, *CD209*–rs4804803-G, *OAS1*–rs10774671-G, *OAS3*–rs1859330-G, *OAS3*–rs2285933-G and *CD32*–rs1801274-G are high frequent, rounding $0.20 > MAF > 0.50$ (Fig. 1A). In some cases, the minimum allele becomes even the maximum allele in Africa, and the amplitude in the frequency between the two alleles can be very high, being 0.41 in *OAS1* and 0.38 in *CD209*. For all these seven SNPs, the highest MAF frequencies are observed in sub-Saharan African populations, decreasing towards Europe and being lowest in Southeast/Northeast Asians (Fig. 1B and Fig. A.2–A.6). This distribution of allele frequencies leads to the highest average F_{ST} genetic distances (Fig. A1) being observed between sub-Saharan African populations or their descendants (Caribe, and less so Latin/North America that are mixes between populations having variable African input) and Southeast/Northeast Asians, reaching extreme values of 0.32 and 0.29 for *CD209* and *OAS1* SNPs. When applying the Mantel test to evaluate correlations between the genetic and the geographic distances for each SNP (Table 1; Fig. A.10), it is possible to verify the existence of overall significant positive correlations, indicating that there is no deviation from the rule of lower exchange of genes as geographical distance increases. Also, checking for the possible effect of natural positive selection upon these SNPs (Fig. 2 and Fig. A.11–A.17), the measure EHH indicates that no selective sweeps occurred in these makers. For comparison purposes, Fig. 2 (and Fig. A.18) includes the profiles for the highly positively selected marker rs4988235 in *MCM6* gene, which in European populations confers lactose tolerance in human adult life (an extended haplotype homozygosity is observed for the derived allele, especially in northern Europeans; the derived allele is absent in sub-Saharan Africans and East Asians) (Gerbault, 2013). As can be observed in the figure, the extended haplotype in this case is considerable larger than the ones observed in the genes associated with dengue.

The remaining three SNPs, *TNFA*–rs1800629-A, *MICB*–rs3132468-C and *MICB*–rs3134899-C, are intermediately frequent ($\sim 0.05 > MAF > 0.20$) across the globe (Fig. 1A), presenting higher MAF in Europeans and North Africans/Southwest Asians, decreasing towards Africa and Asia (Fig. 1C; Fig. A.7–A.9). Distinctly, SNPs on *MICB* and *TNFA* genes are the ones that display lower inter-population group divergence, with values of F_{ST} attaining at most 0.06–0.08 and especially between Asian and

European ancestries (Fig. A.1). These four markers also display a positive correlation in the Mantel test (Table 1; Fig. A.10), except for *MICB*–rs3132468–C allele, where values are not significant. The three markers equally do not show signals of positive selection (Fig. A.19–A.21).

Table 1. **General information about the 10 SNPs screened in this work** and results for the Mantel test (W= Kendall's coefficient of concordance; p-value for Friedman' chi-square statistic tests of W; significant values are indicated on bold). MAF refers to the minimum allele frequency in the global human population.

Gene	SNP	Position in build GRCh37	Function/Location	MAF	Mantel test W	p-value
PLCE1	rs2274223	10:96066341	Missense (His1927Arg)	G	0.677	**0.0001**
PLCE1	rs3765524	10:96058298	Missense (Thr1777Ile)	T	0.659	**0.0001**
TNFA	rs1800629	6:31543031	Upstream gene variant	A	0.557	**0.0219**
DC–SIGN (*CD209*)	rs4804803	19:7812733	Upstream gene variant	G	0.616	**0.0004**
*OAS1*_splicing	rs10774671	12:113357193	Splice acceptor variant	G	0.603	**0.0010**
*OAS3*_K18R	rs1859330	12:113376388	Missense (Arg18Lys)	G	0.607	**0.0010**
*OAS3*_S381R	rs2285933	12:113386779	Missense (Ser381Arg)	G	0.724	**0.0001**
FCgRIIa (*CD32*)	rs1801274	1:161479745	Missense (His166Arg)	G	0.646	**0.0001**
MICB	rs3132468	6:31475486	Intron variant	C	0.486	0.6919
MICB	rs3134899	6:31473286	Intron variant	C	0.585	**0.0029**
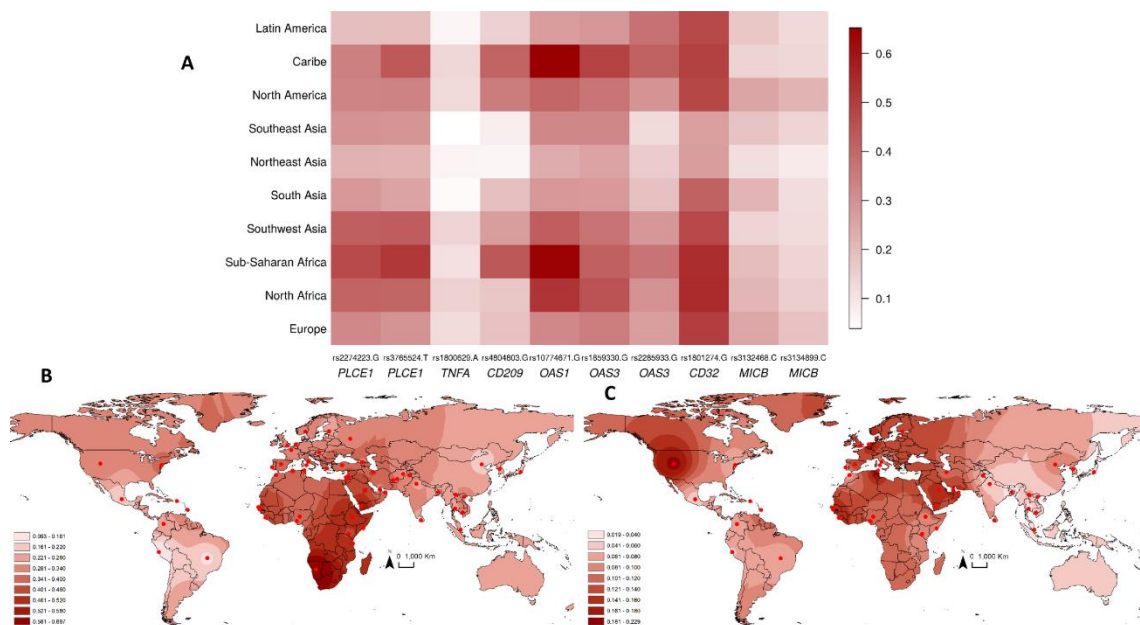
Figure 1. Worldwide diversity for the 10 SNPs. (A) Heatmap for the minimum allele frequency (MAF) in the 10 SNPs. (B) Map of allele frequency distribution for *PLCE1*–rs2274223–G, which follows a pattern shared with *PLCE1*–rs3765524–T, *CD209*–rs4804803–G, *OAS1*–rs10774671–G, *OAS3*–rs1859330–G, *OAS3*–rs2285933–G and *CD32*–rs1801274–G (maps for these SNPs are displayed in Fig A.2–A.7). (C) Map of allele frequency distribution for *TNFA*–rs1800629–A, which follows a pattern shared with *MICB*–rs3132468–C and *MICB*–rs3134899–C (maps for these are displayed in Fig A.8–A.9).
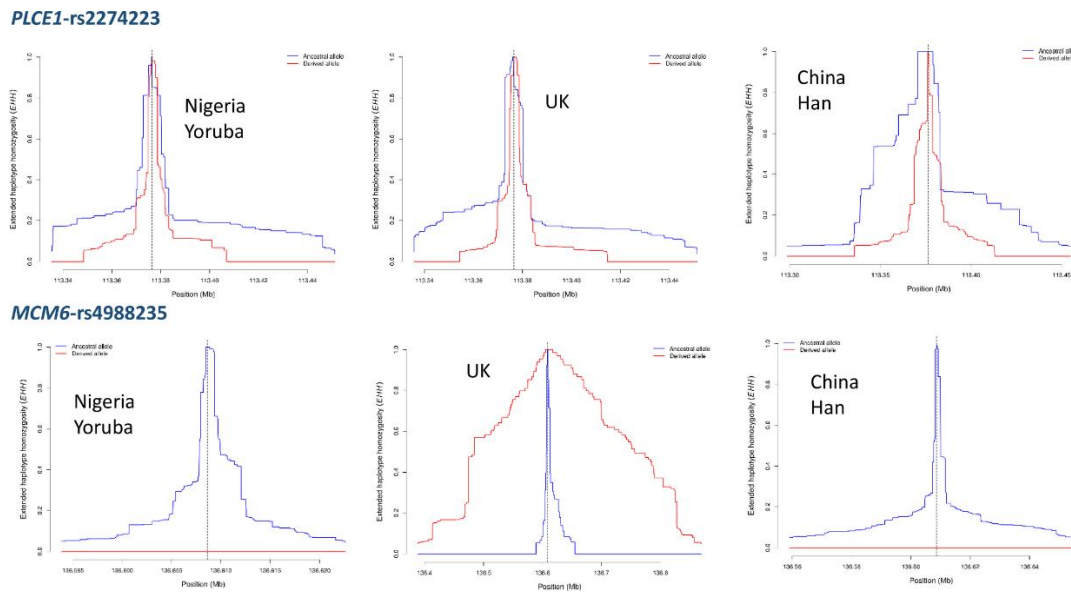
Figure 2. Extended Haplotype Homozygosity (EHH) for *PLCE1*-rs2274223-G and the well-known positively selected European marker for lactase persistence *MCM6*-rs4988235. Data for one population representative from the three main human population groups (Nigerian Yorubans from sub–Saharan Africa, Great Britain from Europe and Chinese Han from East Asia; 1000 Genomes database).

3.2. Meta-analysis

The 2009 WHO classification for dengue disease phenotypes (WHO, 2009) introduced important changes in relation to the 1997 classification (WHO, 1997), aiming at a better clinical structuring of dengue patients. However, it led to addition of noise when joining cohorts from different studies, with considerable overlap between phenotypic classes. Despite these biases, we decided it was worth to compare the information gain from considering (a) stricter phenotypic classes and (b) a broader symptomatic class with higher sample size. Thus, we performed meta–analyses in the following phenotypic classes *vs* healthy controls: DSS (dengue shock syndrome); DHF (dengue hemorrhagic fever); DF (dengue fever); and Symptomatic (all the other classes

together and a few cases that were identified as presenting symptoms but not allowing to classify patients in the previous classes).

The SNPs with larger cohorts are the ones discovered in the Vietnamese GWAS (Khor et al., 2011), rounding thousands of patients and controls (Table A.11), although they have been mainly screened in East Asian populations. Beginning by the SNPs in *MICB* gene, *MICB*–rs3132468–C was genotyped in 9493 patients (4360 DSS, 224 DHF, 3971 DF and 938 with symptoms) *vs* 6785 controls, while *MICB*–rs3134899–C was screened in 4781 patients (3987 DSS, 225 DHF, 472 DF and 97 with symptoms) *vs* 5630 controls. The most significant results are for the comparison DSS *vs* controls (Fig. 3), although values are still significant for the symptomatic *vs* controls comparison (Fig. A.24), while comparisons for the less severe phenotypes are not significant (Fig. A.22– A.23). In the DSS *vs* controls, all the studies showed a high concordance between them, as can be confirmed in the funnel and radial plots (Fig 3B,C,E,F), and also through the non–significant Cochran's Q measure and nil I^2 values. The pooled OR values in DSS were of 1.32 [1.22–1.42] for *MICB*–rs3132468–C and 1.27 [1.16–1.39] for *MICB*–rs3134899–C, being highly statistically significant ($1.19e^{-12}$ and $3.13e^{-7}$, respectively). The concordance between these OR values reflect the fact that the two markers are linked (Fig. A.33) and may be tagging the not yet identified causative marker in this chromosome region.
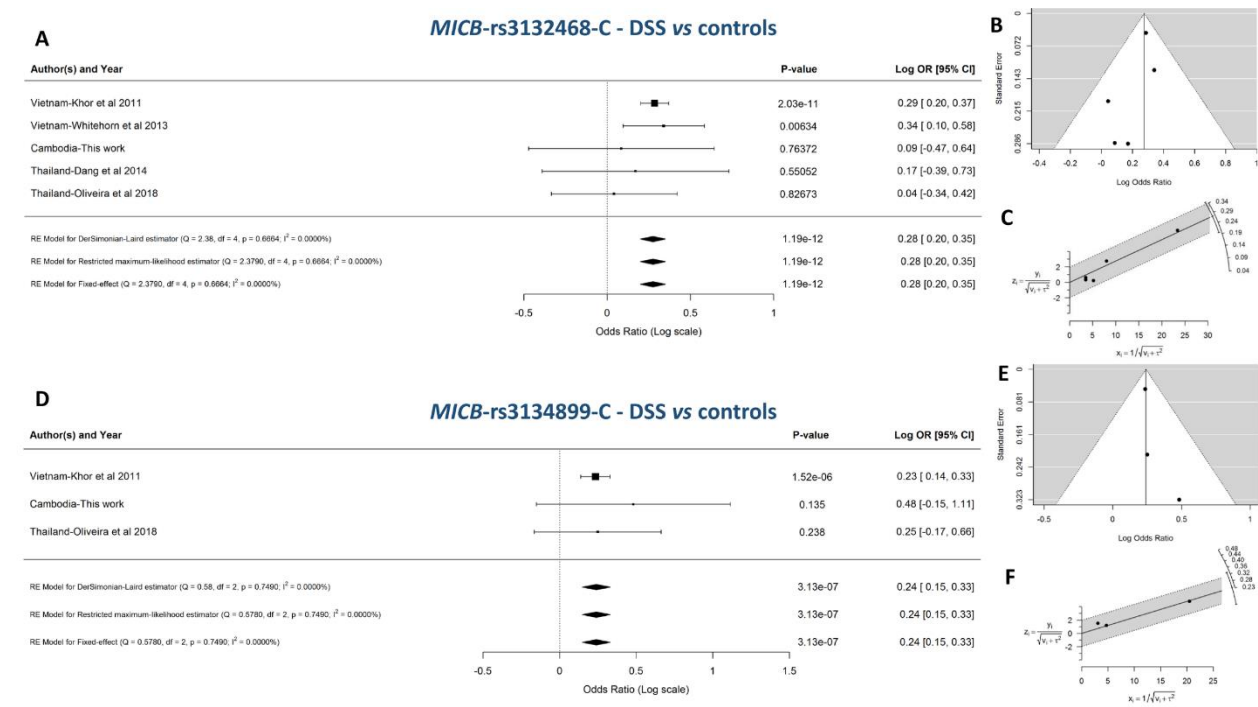
Figure 3. Meta-analysis in *MICB*-rs3132468-C (A-C) and *MICB*-rs3134899-C (D-F) for DSS *vs* controls comparison. (A and D) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's

Q measure (and p–values) and I^2 values for the tested three models (weighted Mantel–Haenszel fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B and E) Funnel plots. (C and F) Radial plots.

For *PLCE1* markers, also thousands of patients and controls (Table A.11), mainly (97%) Eastern Asians, have been genotyped. *PLCE1*–rs2274223–G was genotyped in 4763 patients (3987 DSS, 217 DHF, 192 DF and 90 with symptoms) versus 5614 controls, while *PLCE1*–rs3765524–T was screened in 5133 patients (3904 DSS, 157 DHF, 141 DF and 931 with symptoms) versus 5281 controls. The small Brazilian cohorts and the Thai cohort from Oliveira et al. (which has a higher frequency of heterozygous in DSS class than neighbour cohorts; (Oliveira et al., 2018)) introduced heterogeneity in the meta–analyses for *PLCE1*–rs2274223–G (Fig. A.25A–C). If these samples are removed from the analyses, the other studies give a homogeneous signal, leading to the detection of a significant protective association for this allele and DSS (Fig. 4A), with an OR of 0.80 [0.76–0.86] (p=1.93e^{-10}). The significant value is still detectable in the Symptomatic vs control comparison (p=2.81e^{-10}), but not in the other phenotypes (Fig. A.25D–L). Identical results are obtained for the linked (Fig. A.33) *PLCE1*–rs3765524–T allele: OR=0.79 [0.75–0.85] (p=1.79e^{-11}) in DSS (Fig. 4B; Fig A.26).
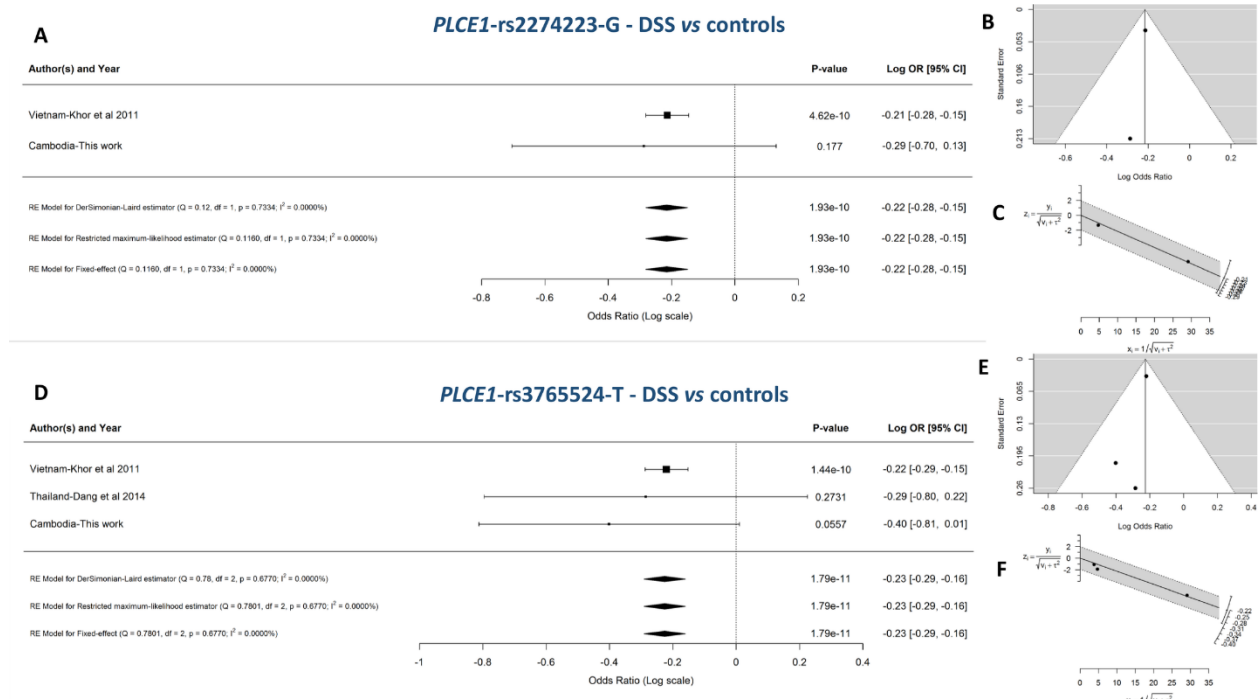
Figure 4. Meta-analysis in *PLCE1*-rs2274223-G (A-C) and *PLCE1*-rs3765524-T (D-F) for DSS *vs* controls comparison. (A and D) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's Q measure (and p–values) and I^2 values for the tested three models (weighted

Mantel–Haenszel fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B and E) Funnel plots. (C and F) Radial plots.

CD32–rs1801274–G meta–analysis departed from 5343 patients (3988 DSS, 588 DHF, 673 DF and 94 with symptoms) versus 6195 controls, with 8% of the individuals being from Latin America. This SNP was contained in the chip screened in Khor et al. (Khor et al., 2011), and that is the reason for this large dataset. Notice that this SNP is not significantly associated with DSS in this Vietnamese cohort alone (see Fig. 5). A first test (Fig. A.27) showed that our data from Cambodia and Cuban cohort from Garcia et al. (Garcia et al., 2010) were introducing biases ($I^2 \approx 65\%$), in opposite directions. When removing these two studies, the other studies are homogeneous, as can be verified in the funnel and radial plots, as well as in Cochran's Q and I^2 measures. Association for this marker (Fig. 5) is close to association in DHF (p=0.053), with a protective OR (0.82 [0.67–1.00]), which becomes significant (p=0.038) in the larger symptomatic comparison with OR (0.94 [0.89–1.00]). Curiously, despite the large sample size in DSS comparison (~7x larger than DHF), the association is not significant in this phenotype class. It thus seems that CD32 is more closely associated with DHF.
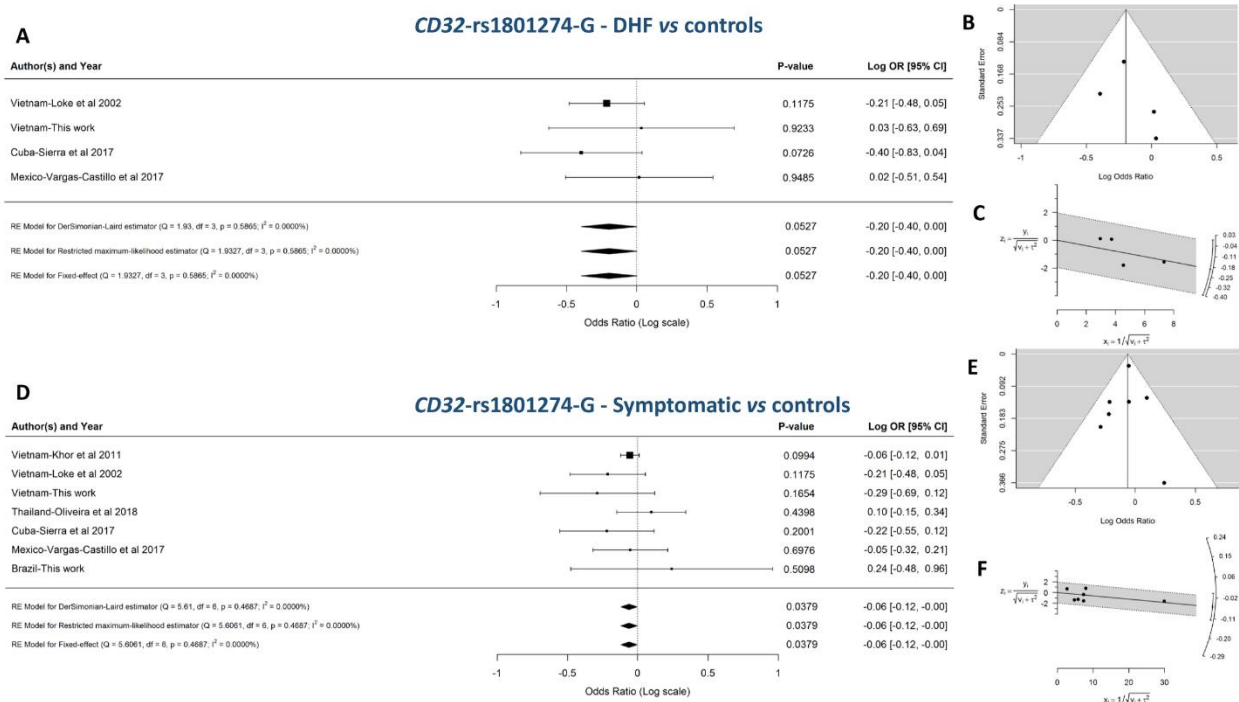
Figure 5. Meta-analysis in *CD32*-rs1801274-G for DHF *vs* controls (A-C) and Symptomatic *vs* controls (D-F) comparisons. (A and D) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's Q measure (and p–values) and I^2 values for the tested three models (weighted

Mantel–Haenszel fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B and E) Funnel plots. (C and F) Radial plots.

The total sample sizes for the remaining SNPs are smaller given their absence from the chip screened in Vietnam (Khor et al., 2011). *CD209*–rs4804803–G, one of the first dengue–associated genetic marker, amounted to 3177 patients (123 DSS, 1262 DHF, 1682 DF and 110 with symptoms) versus 1616 controls. Around 10% of patients and 26% of controls are from Latin America, while the rest are of East/South Asian ancestry. This disparity in the sample sizes between ethnicities does not allow to test the hypothesis if the effect of this SNP could be opposite between East Asian and Latin American backgrounds, as advanced previously (Xavier–Carvalho et al., 2013). The strong significant causative effect of *CD209*–rs4804803–G in DF in the Thai cohort reported in Sakuntabhai et al. (Sakuntabhai et al., 2005) was not obtained in any other work available so far. Even after removing that Thai cohort from the test, the remaining samples input a significant heterogeneity into the meta–analysis in DF *vs* control comparison (Fig. 6A), so that overall there is no significant effect in this phenotype in the three tested models. The same being true for the phenotype DHF (Fig. 6B), and when considering all Symptomatic (Fig. A.28).
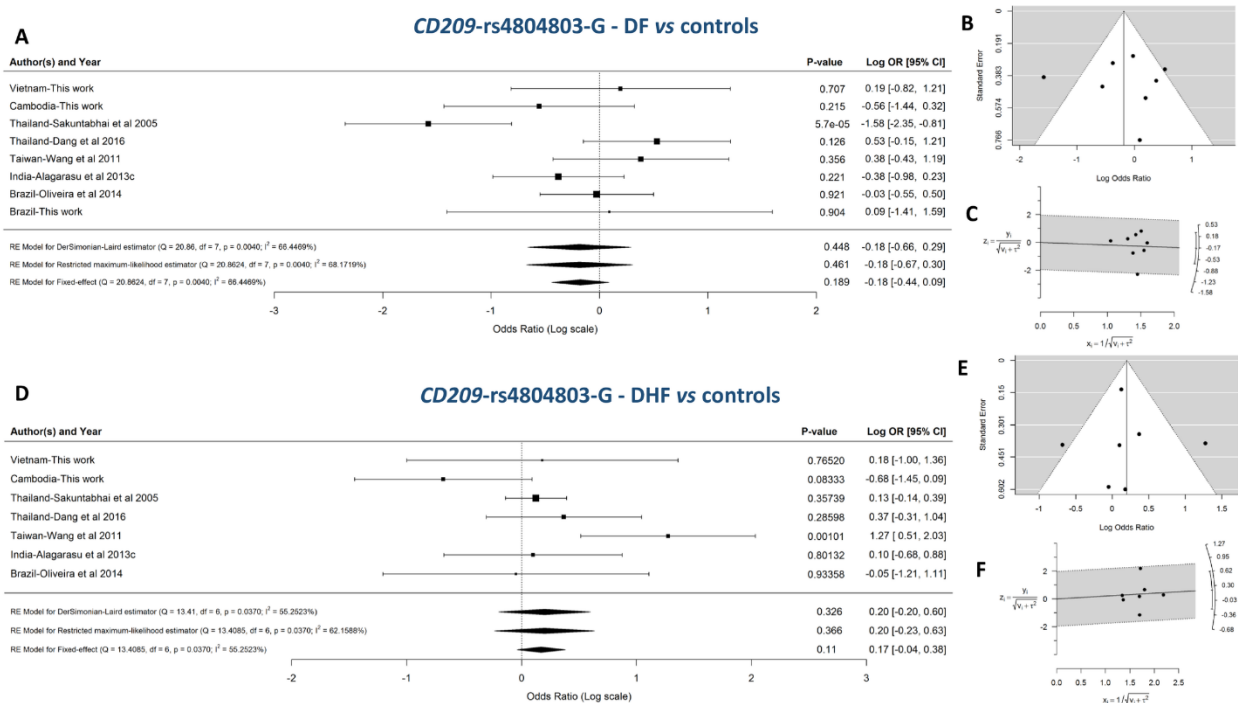
Figure 6. Meta-analysis in *CD209*-rs4804803-G for DF *vs* controls (A-C) and DHF *vs* controls (D-F) comparisons. (A and D) Forest plots, pooled log OR ([95% CI] and p-values), Cochran's *Q* measure (and p–values) and *I²* values for the tested three models (weighted

Mantel–Haenszel fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B and E) Funnel plots. (C and F) Radial plots.

TNFA–rs1800629–A dataset sums up 1943 patients (165 DSS, 776 DHF, 906 DF and 96 with symptoms) versus 1934 controls, of which 51% patients and 74% controls are from Latin American ancestry. When samples are considered together, only the larger Symptomatic *vs* control comparison presents a significant protective effect (OR=0.81 [0.67–0.98]; p=0.022 to p=0.046 depending on the model), although heterogeneity is still around 20% for I^2 measure (Fig. 7A). Limiting the analysis to the Latin American samples, a significant protective effect (OR=0.67 [0.50–0.90]; p=0.0092) was observed for the DF *vs* control comparison (Fig. 7B), and not for the severer phenotypes (Fig. A.29).
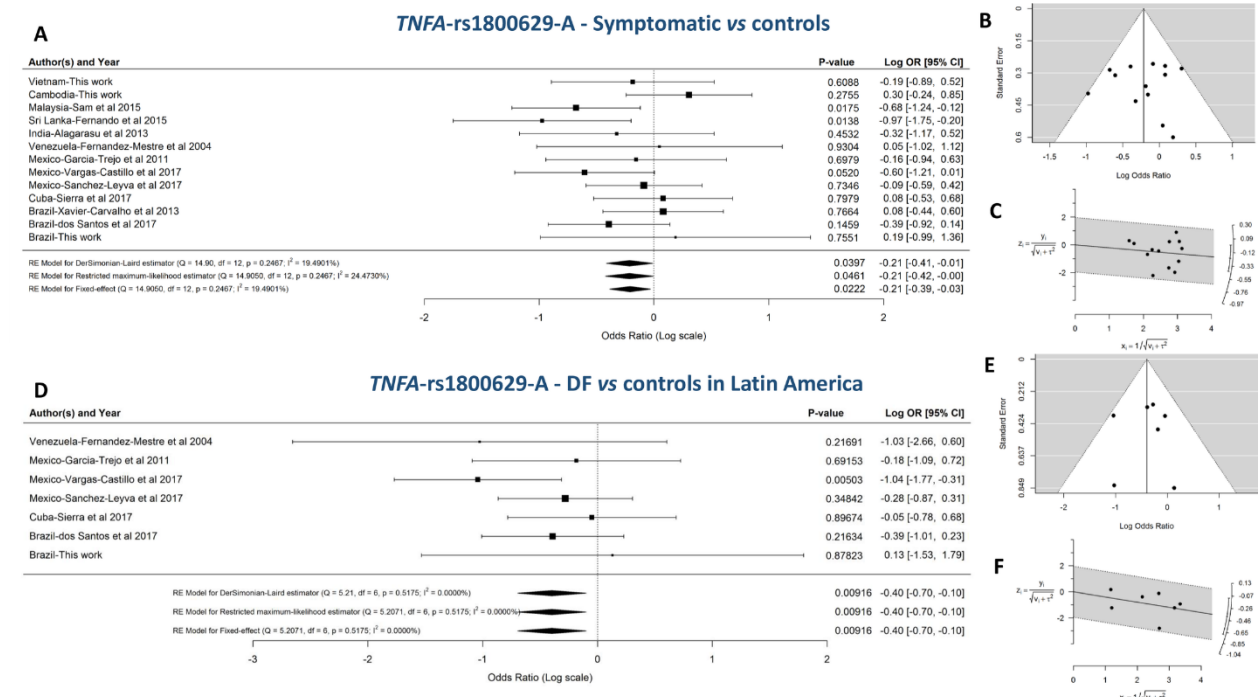
Figure 7. Meta-analysis in *TNFA*-rs1800629-A for symptomatic *vs* controls (A-C) and DF *vs* controls in Latin America (D-F) comparisons. (A and D) Forest plots, pooled log OR ([95% CI] and p-values), Cochran's *Q* measure (and p-values) and I^2 values for the tested three models (weighted Mantel–Haenszel fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B and E) Funnel plots. (C and F) Radial plots.

Lastly, the OAS family presented the lower sample sizes screened so far, and with the exceptions of a few individuals from India (Alagarasu et al., 2013b) and Cuba (Sierra

et al., 2017) typed for *OAS1*–rs10774671–G, the Thai cohorts from Simon–Loriere et al. (Simon–Loriere et al., 2015) dominate these meta–analyses. Notice that Simon–Loriere et al. (Simon–Loriere et al., 2015) did not provide controls, and in order to be consistent in our tests we included our Thai population screening as controls to compare with these Thai patient cohorts. We also decided to perform an overall evaluation, disregarding DENV strain, which was not accounted for in all other studies besides Simon–Loriere et al. (Simon–Loriere et al., 2015). Numbers amounted to 1125 patients (185 DSS, 505 DHF, 407 DF and 28 with symptoms) *vs* 402 controls for *OAS1*–rs10774671–G, 931 patients (187 DSS, 412 DHF, 304 DF and 28 with symptoms) *vs* 242 controls for *OAS3*–rs1859330–G, and 941 patients (187 DSS, 419 DHF, 307 DF and 28 with symptoms) *vs* 236 controls for *OAS3*–rs2285933–G. Values were non–significant for all *OAS1*–rs10774671–G and linked (Fig. A.33) *OAS3*–rs1859330–G comparisons (Fig. A.30–A.31). While *OAS3*–rs2285933–G confers significant protection in DSS, with OR of 0.55 [0.34–0.90], p=0.016 considering the fixed–effect model (Fig. 8; Fig. A.32).
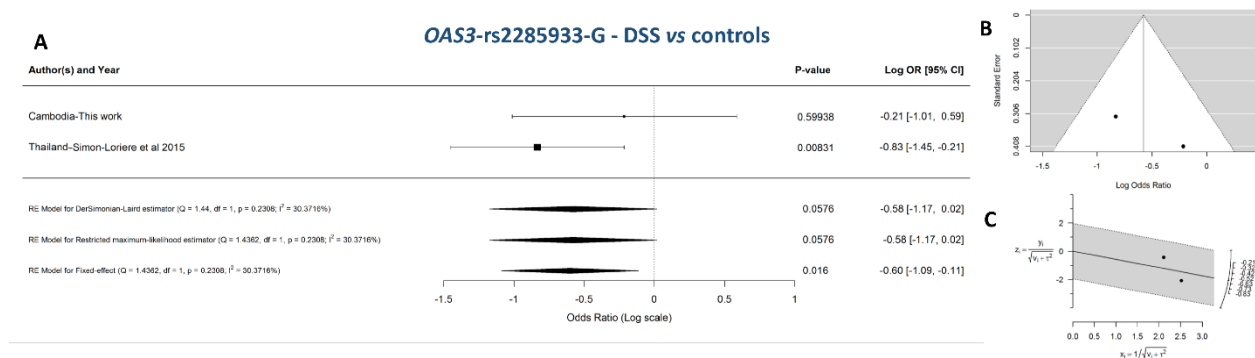
Figure 8. Meta-analysis in *OAS3*-rs2285933-G for DSS *vs* controls comparison (A-C). (A) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and *I²* values for the tested three models (weighted Mantel–Haenszel fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B) Funnel plots. (C) Radial plots.

3.3. Genetic risk scores

For markers confirmed here to be associated with the phenotype DSS, the protective are highly frequent in Africa (*PLCE1* SNPs and *OAS3*–rs2285933–G), while the causative are low frequent in Africa (*MICB* SNPs). The same is true for the *CD32*–rs1801274–G, which seems to protect from DHF. Thus, when summing up the risk scores conferred by *PLCE1*–rs2274223–G (linked *PLCE1*–rs3765524–T was left out), *MICB*–rs3132468–C (linked *MICB*–rs3134899–C was also left out), *OAS3*–rs2285933–

G and *CD32*–rs1801274–G (Fig. 9A), sub–Saharan populations and their descendants are the best protected against the severer dengue phenotypes, in contrast with both Northeast Asian and Southeast populations, which are least protected. The difference in the means of genetic risk between those populations groups are highly significant (sub–Saharan Africa *vs* Southeast Asia, $p=2.96e^{-63}$; sub–Saharan Africa *vs* Northeast Asia, $p=7.66e^{-56}$). European populations and neighbor North African and Southwest Asian populations are in between those genetic risk scores, being significantly different from both (sub–Saharan Africa vs Europe, $p=2.02e^{-16}$; Europe *vs* Southeast Asia, $p=2.63e^{-8}$).

The protection conferred by *TNFA*–rs1800629–A to DF is higher in European, North African and Southwest Asian populations, and decreases towards Africa and especially East Asia (Fig 9B), being statistically significantly different for the latter (sub–Saharan Africa *vs* Europe, $p=0.85$; Europe *vs* Southeast Asia, $p=2.22e^{-16}$).
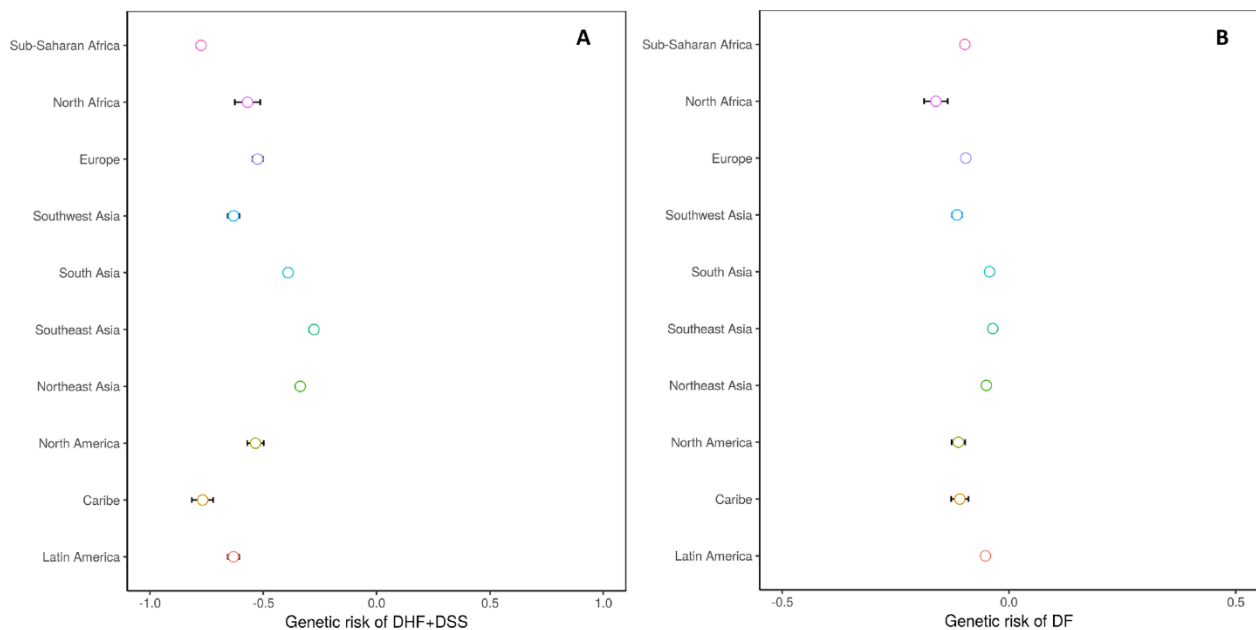
Figure 9. Genetic risk (median and standard deviation) for DHF+DSS (A) and DF (B) in population samples reported in Tables A.1-A.10. (A) Values were calculated taking into account the significant association found here for *PLCE1*–rs2274223–G, *MICB*–rs3132468–C, *OAS3*–rs2285933–G and *CD32*–rs1801274–G. (B) Values were calculated for the significant association with *TNFA*–rs1800629–A.

4. Conclusions

Population genetic models predict that a large fraction of missing heritability for complex traits could be explained by *loci* that contain classes of rare (MAF <1%) susceptibility variants (Pritchard, 2001). However, GWAS have been identifying mainly common variants, already present in the ancestral African population that migrated out-of-Africa (Rosenberg et al., 2010). In fact, by analysing 400 susceptibility-conferring SNPs across a spectrum of qualitative and quantitative traits, Park et al. (Park et al., 2011) observed that both their number and their collective contribution to genetic variance were highest for more common MAF categories (30-40% or 40-50%), and dropped substantially for lower allele-frequency categories (5-10% or 10-20%). The authors confirmed that these results are not caused by lower statistical power for detection of association for low-frequency SNPs. Another simulation analysis (Simons et al., 2014) showed that variance due to rare alleles is only expected for diseases caused primarily by strong deleterious mutations (diseases tightly coupled to fitness). The 10 markers surveyed here for association with dengue disease follow these tendencies of SNPs uncovered by GWAS for non-strong uncoupling diseases. Seven markers (in *PLCE1*, *CD32*, *CD209*, *OAS1* and *OAS3* genes) have high-frequency MAF and the other three (in *MICB* and *TNFA* genes) have intermediate frequency MAF. Most of these markers have high discriminatory power between population groups. This is especially so for the SNPs in *CD209* and *OAS1* genes, which attain F_{ST} values between 0.20-0.32 in comparisons involving sub-Saharan African and Northeast/Southeast Asian ancestries. This level of values is attained by the so-called ancestry informative markers (our own calculations in SNPs referred in (Marcheco-Teruel et al., 2014)), while reference values taking into account multiple SNPs based on whole genome sequencing are at the level of 0.106 for European-East Asian and 0.139 for European-West African comparisons (Bhatia et al., 2013). Despite this high population discrimination, frequencies of these alleles follow the rules of genetic drift and seem not to have been under strong selective pressure.

If a large proportion of causal variants are common and shared across diverse populations, they will be replicated in trans-ethnic GWAS, as demonstrated in a large-scale review (Marigorta and Navarro, 2013) for 28 diseases in European, East Asian and African ancestries. Even when power is insufficient to achieve statistically independent significance, there are high rates of directional consistency across trans-ethnic GWAS signals (Ntzani et al., 2012). We confirmed this tendency to directional consistency between studies performed in East Asia and Latin America in dengue, with

heterogeneity generated by randomness between studies and especially by low sample sizes. Also, the patterns of LD for the *PLCE1*, *MICB* and *OAS* genes are quite consistent in the three main population groups, and, so, the potential of replication in East Asia or Latin America for these markers is identical.

Another important result from Park et al. (Park et al., 2011) study was that the genetic variance due to individual susceptibility SNPs, on average, remains fairly constant over different ranges of allele frequency, as for example in Crohn's disease, the average regression effects corresponded to OR of 1.08 for MAF=0.45, 1.13 for MAF=0.15, and 1.16 for MAF=0.05. Our results also corroborate this evidence, with close protective OR values, between 0.55 for *OAS3*–rs2285933-G, 0.67 for *TNFA*–rs1800629-A, 0.80 for *PLCE1*–rs2274223-G and 0.82 for *CD32*–rs1801274-G, despite differences in MAFs. Even the causative OR was of similar order of magnitude, 1.32 for *MICB*–rs3132468-C.

Our meta-analysis was not informative enough to replicate overall the significant association between *CD209*–rs4804803-G and DF/DHF reported by Sakuntabhai et al. (Sakuntabhai et al., 2005) in Thai, neither the opposite-direction association reported in Brazilians (Xavier-Carvalho et al., 2013). This result could be interpreted as surprising given the collected functional evidence that this polymorphism interferes with response to DENV infection (Sakuntabhai et al., 2005), but follows many failed replication studies, even of associations that turned out to be genuine. The meta-analysis conducted by Lohmueller et al. (Lohmueller et al., 2003) of 301 association studies of 25 disease loci, showed that only 11 loci were confirmed in the replications, and, even more striking, 24 of the 25 loci presented lower ORs than the initial study. This result is a consequence of the data set used to identify the variant of interest being a non-random population sample (Zollner and Pritchard, 2007). In fact, tests oversample affected individuals relative to their frequency in the population (an unreal similar number of cases and controls are tested), and there is a major ascertainment effect that occurs when a variant is of interest, specifically because it was significant for association. As most of the associations have modest effects, when there is a significant result, it may imply that the genotype counts of cases and controls are more different from each other than expected. Consequently, the estimates of effect size are biased upwards, an effect known as the "winner's curse". Gorroochurn et al. (Gorroochurn et al., 2007) developed a formula for the replication power of a second association study based on the p-value of an initial study, concluding that: (1) a p-value only slightly lower than the nominal α results in only

approximately 50% replication power; (2) very low p-values are required to achieve a replication power of at least 80% (e.g., at α=0.05, a p<0.005 is required). The authors state that because many initially significant findings result in low replication power, replication failure should not be surprising or be interpreted as necessarily refuting the initial findings.

Our results also support the structuring of dengue patients into specific phenotype classes, despite the lowering in sample size for the statistical evaluation. It seems that particular polymorphisms can confer susceptibility or protection against particular molecular phenomena, which play a different role along the disease course, as we have recently (Oliveira et al., 2018) demonstrated in a GWAS conducted in Thai cohorts. In this study we were able to identify four genes involved in xenobiotic metabolism that are protecting against dengue fever phenotype, possibly through linking to the virus and conditioning of its localization in host cells. While a new gene (*PLCB4*) of the *PLCE1*-family protects against dengue shock syndrome and may be related with cytokine dynamics, inflammation and activation of vascular endothelium cells. In the current study, *TNFA* seems to protect for DF, *CD32* protects for DHF (the large sample size for DSS in the meta-analysis shows that the non-association with this phenotype class is not a sample size issue), while *PLCE1*, *MICB* and *OAS3* play a role at the level of DSS.

We were not able to evaluate the influence of dengue virus strain in the associations detected in these markers. Simon-Loriere et al. (Simon-Loriere et al., 2015) showed statistically and functionally that this is an important factor when ascertaining the association for OAS family. Most studies do not provide information about the circulating strain when cohorts were collected, rendering it impossible to conduct meta-analyses incorporating the strain information. Groups working in the field should henceforth make an effort to provide this information for proper future evaluation.

As failures in replications accumulate in GWAS, and sample sizes in the order of thousands are difficult to cope with (implying huge efforts in sample collection and financing), the net result was a general disinterest in performing GWAS, and specifically in neglected tropical diseases. Fortunately, some improvements in statistical methods are allowing to take advantage of ethnicity information in conducting more powerful association studies in lower sample size cohorts (Maples et al., 2013; Shriner et al., 2011). These methods analyse blocks of SNPs that can be affiliated to one of the ancestral populations. The blocks are randomly distributed by

recombination across the chromosomes of the admixed descendant individuals, reflecting the frequency contributions of the parental populations (expectation for the control group), except in genomic locations where a candidate gene confers susceptibility to a certain disease (significant frequency increase in the parental population with higher MAF for the susceptibility marker) in cases. As there are only hundreds of ancestral blocks, these tests have considerable lower statistical burdens than the traditional GWAS with thousands of SNPs. We took advantage of joined admixture-association studies and conducted work in Cuban (Sierra et al., 2017) and Thai (Oliveira et al., 2018) cohorts, mapping new candidate genes: lipid metabolism-related *OSBPL10* and *RXRA*; and the already mentioned four genes involved in xenobiotic metabolism (*CHST10*, *AHRR*, *GRIP1* and *PPP2R5E*) and *PLCB4* of the *PLCE1-*family. The genetic risks conferred by the markers reported in these publications add precisely to the genetic risks displayed by the markers studied here. Sub-Saharan African populations and descendants are the best protected against the severer forms of the disease, while Southeast and Northeast Asians are the least protected ones. European and close neighbours are the best protected against dengue fever, while, again, Southeast and Northeast Asians are the least protected ones.

Acknowledgments

References

Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., McVean, G.A., 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491, 56-65.

Acosta, E.G., Kumar, A., Bartenschlager, R., 2014. Revisiting dengue virus-host cell interaction: new insights into molecular and cellular virology. Advances in virus research 88, 1-109.

Alagarasu, K., Bachal, R.V., Damle, I., Shah, P.S., Cecilia, D., 2015. Association of FCGR2A p.R131H and CCL2 c.-2518 A>G gene variants with thrombocytopenia in patients with dengue virus infection. Human immunology 76, 819-822.

Alagarasu, K., Damle, I.M., Bachal, R.V., Mulay, A.P., Shah, P.S., Dayaraj, C., 2013a. Association of promoter region polymorphisms of CD209 gene with clinical outcomes of dengue virus infection in Western India. Infect Genet Evol 17, 239-242.

Alagarasu, K., Honap, T., Damle, I.M., Mulay, A.P., Shah, P.S., Cecilia, D., 2013b. Polymorphisms in the oligoadenylate synthetase gene cluster and its association with clinical outcomes of dengue virus infection. Infect Genet Evol 14, 390-395.

Alagarasu, K., Mulay, A.P., Singh, R., Gavade, V.B., Shah, P.S., Cecilia, D., 2013c. Association of HLA-DRB1 and TNF genotypes with dengue hemorrhagic fever. Human immunology 74, 610-617.

Atkinson, Q.D., Gray, R.D., Drummond, A.J., 2008. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. Mol Biol Evol 25, 468-474.

Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R., 2015. A global reference for human genetic variation. Nature 526, 68-74.

Barrett, J.C., Fry, B., Maller, J., Daly, M.J., 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21, 263-265.

Bhatia, G., Patterson, N., Sankararaman, S., Price, A.L., 2013. Estimating and interpreting FST: the impact of rare variants. Genome Res 23, 1514-1521.

Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., Drake, J.M., Brownstein, J.S., Hoen, A.G., Sankoh, O., Myers, M.F., George, D.B., Jaenisch, T., Wint, G.R., Simmons, C.P., Scott, T.W., Farrar, J.J., Hay, S.I., 2013. The global distribution and burden of dengue. Nature 496, 504-507.

Chareonsirisuthigul, T., Kalayanarooj, S., Ubol, S., 2007. Dengue virus (DENV) antibody-dependent enhancement of infection upregulates the production of anti-inflammatory cytokines, but suppresses anti-DENV free radical and pro-inflammatory cytokine production, in THP-1 cells. The Journal of general virology 88, 365-375.

Coffey, L.L., Mertens, E., Brehin, A.C., Fernandez-Garcia, M.D., Amara, A., Despres, P., Sakuntabhai, A., 2009. Human genetic determinants of dengue virus susceptibility. Microbes Infect 11, 143-156.

Consortium, T.G.o.t.N., 2014. Whole-genome sequence variation, population structure and demographic history of the Dutch population. Nat Genet 46, 818-825.

Consortium, T.I.H., 2005. A haplotype map of the human genome. Nature 437, 1299-1320.

Dang, T.N., Naka, I., Sa-Ngasang, A., Anantapreecha, S., Chanama, S., Wichukchinda, N., Sawanpanyalert, P., Patarapotikul, J., Tsuchiya, N., Ohashi, J., 2014. A replication study confirms the association of GWAS-identified SNPs at MICB and PLCE1 in Thai patients with dengue shock syndrome. BMC Med Genet 15, 58.

Dang, T.N., Naka, I., Sa-Ngasang, A., Anantapreecha, S., Wichukchinda, N., Sawanpanyalert, P., Patarapotikul, J., Tsuchiya, N., Ohashi, J., 2016. Association of BAK1 single nucleotide polymorphism with a risk for dengue hemorrhagic fever. BMC Med Genet 17, 43.

Duong, V., Lambrechts, L., Paul, R.E., Ly, S., Lay, R.S., Long, K.C., Huy, R., Tarantola, A., Scott, T.W., Sakuntabhai, A., Buchy, P., 2015. Asymptomatic humans transmit dengue virus to mosquitoes. Proc Natl Acad Sci U S A 112, 14688-14693.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1, 47-50.

Fernandez-Mestre, M.T., Gendzekhadze, K., Rivas-Vetencourt, P., Layrisse, Z., 2004. TNF-alpha-308A allele, a possible severity risk factor of hemorrhagic manifestation in dengue fever patients. Tissue antigens 64, 469-472.

Fernando, A.N., Malavige, G.N., Perera, K.L., Premawansa, S., Ogg, G.S., De Silva, A.D., 2015. Polymorphisms of Transporter Associated with Antigen Presentation, Tumor Necrosis Factor-alpha and Interleukin-10 and their Implications for Protection and Susceptibility to Severe Forms of Dengue Fever in Patients in Sri Lanka. Journal of global infectious diseases 7, 157-164.

Fu, W., O'Connor, T.D., Jun, G., Kang, H.M., Abecasis, G., Leal, S.M., Gabriel, S., Rieder, M.J., Altshuler, D., Shendure, J., Nickerson, D.A., Bamshad, M.J., Akey, J.M., 2013. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature 493, 216-220.

Garcia-Trejo, A.R., Falcon-Lezama, J.A., Juarez-Palma, L., Granados, J., Zuniga-Ramos, J., Rangel, H., Barquera, R., Vargas-Alarcon, G., Ramos, C., 2011. Tumor necrosis factor alpha promoter polymorphisms in Mexican patients with dengue fever. Acta tropica 120, 67-71.

Garcia, G., Sierra, B., Perez, A.B., Aguirre, E., Rosado, I., Gonzalez, N., Izquierdo, A., Pupo, M., Danay Diaz, D.R., Sanchez, L., Marcheco, B., Hirayama, K., Guzman, M.G., 2010. Asymptomatic dengue infection in a Cuban population confirms the protective role of the RR variant of the FcgammaRIIa polymorphism. The American journal of tropical medicine and hygiene 82, 1153-1156.

Gautier, M., Vitalis, R., 2012. rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics 28, 1176-1177.

Gautret, P., Simon, F., 2016. Dengue, chikungunya and Zika and mass gatherings: What happened in Brazil, 2014. Travel medicine and infectious disease 14, 7-8.

Gerbault, P., 2013. The onset of lactase persistence in Europe. Human heredity 76, 154-161.

Gorroochurn, P., Hodge, S.E., Heiman, G.A., Durner, M., Greenberg, D.A., 2007. Non-replication of association studies: "pseudo-failures" to replicate? Genetics in medicine : official journal of the American College of Medical Genetics 9, 325-331.

Guzman, M.G., Harris, E., 2015. Dengue. Lancet (London, England) 385, 453-465.

Hellenthal, G., Busby, G.B., Band, G., Wilson, J.F., Capelli, C., Falush, D., Myers, S., 2014. A genetic atlas of human admixture history. Science 343, 747-751.

Henn, B.M., Botigue, L.R., Gravel, S., Wang, W., Brisbin, A., Byrnes, J.K., Fadhlaoui-Zid, K., Zalloua, P.A., Moreno-Estrada, A., Bertranpetit, J., Bustamante, C.D., Comas, D., 2012. Genomic ancestry of North Africans supports back-to-Africa migrations. PLoS Genet 8, e1002397.

Karlsson, E.K., Kwiatkowski, D.P., Sabeti, P.C., 2014. Natural selection and infectious disease in human populations. Nat Rev Genet 15, 379-393.

Khor, C.C., Chau, T.N., Pang, J., Davila, S., Long, H.T., Ong, R.T., Dunstan, S.J., Wills, B., Farrar, J., Van Tram, T., Gan, T.T., Binh, N.T., Tri le, T., Lien le, B., Tuan, N.M., Tham, N.T., Lanh, M.N., Nguyet, N.M., Hieu, N.T., Van, N.V.C.N., Thuy, T.T., Tan, D.E., Sakuntabhai, A., Teo, Y.Y., Hibberd, M.L., Simmons, C.P., 2011. Genome-wide association study identifies susceptibility loci for dengue shock syndrome at MICB and PLCE1. Nat Genet 43, 1139-1141.

Kraemer, M.U., Sinka, M.E., Duda, K.A., Mylne, A.Q., Shearer, F.M., Barker, C.M., 2015. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. Elife 4, e08347.

Kroeger, K.M., Carville, K.S., Abraham, L.J., 1997. The -308 tumor necrosis factor-alpha promoter polymorphism effects transcription. Molecular immunology 34, 391-399.

Li, J.Z., Absher, D.M., Tang, H., Southwick, A.M., Casto, A.M., Ramachandran, S., Cann, H.M., Barsh, G.S., Feldman, M., Cavalli-Sforza, L.L., Myers, R.M., 2008. Worldwide human relationships inferred from genome-wide patterns of variation. Science 319, 1100-1104.

Lohmueller, K.E., Indap, A.R., Schmidt, S., Boyko, A.R., Hernandez, R.D., Hubisz, M.J., Sninsky, J.J., White, T.J., Sunyaev, S.R., Nielsen, R., Clark, A.G., Bustamante, C.D., 2008. Proportionally more deleterious genetic variation in European than in African populations. Nature 451, 994-997.

Lohmueller, K.E., Pearce, C.L., Pike, M., Lander, E.S., Hirschhorn, J.N., 2003. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 33, 177-182.

Loke, H., Bethell, D., Phuong, C.X., Day, N., White, N., Farrar, J., Hill, A., 2002. Susceptibility to dengue hemorrhagic fever in vietnam: evidence of an association with variation in the vitamin d receptor and Fc gamma receptor IIa genes. The American journal of tropical medicine and hygiene 67, 102-106.

Lozach, P.Y., Burleigh, L., Staropoli, I., Navarro-Sanchez, E., Harriague, J., Virelizier, J.L., Rey, F.A., Despres, P., Arenzana-Seisdedos, F., Amara, A., 2005. Dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin (DC-SIGN)-mediated enhancement of dengue virus infection is independent of DC-SIGN internalization signals. J Biol Chem 280, 23698-23708.

Maples, B.K., Gravel, S., Kenny, E.E., Bustamante, C.D., 2013. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. Am J Hum Genet 93, 278-288.

Marcheco-Teruel, B., Parra, E.J., Fuentes-Smith, E., Salas, A., Buttenschon, H.N., Demontis, D., Torres-Espanol, M., Marin-Padron, L.C., Gomez-Cabezas, E.J., Alvarez-Iglesias, V., Mosquera-Miguel, A., Martinez-Fuentes, A., Carracedo, A., Borglum, A.D., Mors, O., 2014. Cuba: exploring the history of admixture and the genetic basis of pigmentation using autosomal and uniparental markers. PLoS Genet 10, e1004488.

Marden, J.R., Walter, S., Tchetgen Tchetgen, E.J., Kawachi, I., Glymour, M.M., 2014. Validation of a polygenic risk score for dementia in black and white individuals. Brain and behavior 4, 687-697.

Marigorta, U.M., Navarro, A., 2013. High trans-ethnic replicability of GWAS results implies common causal variants. PLoS Genet 9, e1003566.

McEvoy, B.P., Powell, J.E., Goddard, M.E., Visscher, P.M., 2011. Human population dispersal "Out of Africa" estimated from linkage disequilibrium and allele frequencies of SNPs. Genome Res 21, 821-829.

Messina, J.P., Brady, O.J., Scott, T.W., Zou, C., Pigott, D.M., Duda, K.A., Bhatt, S., Katzelnick, L., Howes, R.E., Battle, K.E., Simmons, C.P., Hay, S.I., 2014. Global spread of dengue virus types: mapping the 70 year history. Trends in microbiology 22, 138-146.

Morel, C.M., 2003. Neglected diseases: under-funded research and inadequate health interventions. Can we change this reality? EMBO reports 4 Spec No, S35-38.

Muntasell, A., Magri, G., Pende, D., Angulo, A., Lopez-Botet, M., 2010. Inhibition of NKG2D expression in NK cells by cytokines secreted in response to human cytomegalovirus infection. Blood 115, 5170-5179.

Navarro-Sanchez, E., Altmeyer, R., Amara, A., Schwartz, O., Fieschi, F., Virelizier, J.L., Arenzana-Seisdedos, F., Despres, P., 2003. Dendritic-cell-specific ICAM3-grabbing non-integrin is essential for the productive infection of human dendritic cells by mosquito-cell-derived dengue viruses. EMBO reports 4, 723-728.

Noecker, C.A., Amaya-Larios, I.Y., Galeana-Hernandez, M., Ramos-Castaneda, J., Martinez-Vega, R.A., 2014. Contrasting associations of polymorphisms in FcgammaRIIa and DC-SIGN with the clinical presentation of dengue infection in a Mexican population. Acta tropica 138, 15-22.

Ntzani, E.E., Liberopoulos, G., Manolio, T.A., Ioannidis, J.P., 2012. Consistency of genome-wide associations across major ancestral groups. Human genetics 131, 1057-1071.

Oliveira, L.F., Lima, C.P., Azevedo Rdo, S., Mendonca, D.S., Rodrigues, S.G., Carvalho, V.L., Pinto, E.V., Maia, A.L., Maia, M.H., Vasconcelos, J.M., Silva, A.L., Nunes, M.R., Sena, L., Vasconcelos, P.F., Santos, E.J., 2014. Polymorphism of DC-SIGN (CD209) promoter in association with clinical symptoms of dengue fever. Viral immunology 27, 245-249.

Oliveira, M., Lert-Itthiporn, W., Cavadas, B., Fernandes, V., Chuansumrit, A., Anunciação, O., Casademont, I., Koeth, F., Penova, M., Tangnararatchakit, K., Khor, C.C., Paul, R., Malasit, P., Matsuda, F., Simon-Loriere, E., Suriyaphol, P., Pereira, L., Sakuntabhai, A., 2018. Joint ancestry and association test indicate two distinct pathogenic pathways involved in classical dengue fever and dengue shock syndrome. PLoS Negl Trop Dis 12, e0006202.

Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20, 289-290.

Park, J.H., Gail, M.H., Weinberg, C.R., Carroll, R.J., Chung, C.C., Wang, Z., Chanock, S.J., Fraumeni, J.F., Jr., Chatterjee, N., 2011. Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. Proc Natl Acad Sci U S A 108, 18026-18031.

Peisley, A., Hur, S., 2013. Multi-level regulation of cellular recognition of viral dsRNA. Cellular and molecular life sciences : CMLS 70, 1949-1963.

Perez, A.B., Sierra, B., Garcia, G., Aguirre, E., Babel, N., Alvarez, M., Sanchez, L., Valdes, L., Volk, H.D., Guzman, M.G., 2010. Tumor necrosis factor-alpha, transforming growth factor-beta1, and interleukin-10 gene polymorphisms: implication in protection or susceptibility to dengue hemorrhagic fever. Human immunology 71, 1135-1140.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38, 904-909.

Pritchard, J.K., 2001. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet 69, 124-137.

Rito, T., Richards, M.B., Fernandes, V., Alshamali, F., Cerny, V., Pereira, L., Soares, P., 2013. The first modern human dispersals across Africa. PLoS One 8, e80031.

Rosenberg, N.A., Huang, L., Jewett, E.M., Szpiech, Z.A., Jankovic, I., Boehnke, M., 2010. Genome-wide association studies in diverse populations. Nat Rev Genet 11, 356–366.

Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J., Ackerman, H.C., Campbell, S.J., Altshuler, D., Cooper, R., Kwiatkowski, D., Ward, R., Lander, E.S., 2002. Detecting recent positive selection in the human genome from haplotype structure. Nature 419, 832–837.

Sakuntabhai, A., Turbpaiboon, C., Casademont, I., Chuansumrit, A., Lowhnoo, T., Kajaste-Rudnitski, A., Kalayanarooj, S.M., Tangnararatchakit, K., Tangthawornchaikul, N., Vasanawathana, S., Chaiyaratana, W., Yenchitsomanus, P.T., Suriyaphol, P., Avirutnan, P., Chokephaibulkit, K., Matsuda, F., Yoksan, S., Jacob, Y., Lathrop, G.M., Malasit, P., Despres, P., Julier, C., 2005. A variant in the CD209 promoter is associated with severity of dengue disease. Nat Genet 37, 507–513.

Sam, S.S., Teoh, B.T., Chinna, K., AbuBakar, S., 2015. High producing tumor necrosis factor alpha gene alleles in protection against severe manifestations of dengue. International journal of medical sciences 12, 177–186.

Sanchez-Leyva, M., Sanchez-Zazueta, J.G., Osuna-Ramos, J.F., Rendon-Aguilar, H., Felix-Espinoza, R., Becerra-Loaiza, D.S., Sanchez-Garcia, D.C., Romero-Quintana, J.G., Castillo Ureta, H., Velarde-Rodriguez, I., Velarde-Felix, J.S., 2017. Genetic Polymorphisms of Tumor Necrosis Factor Alpha and Susceptibility to Dengue Virus Infection in a Mexican Population. Viral immunology 30, 615–621.

Santos, A., de Moura, E.L., Ferreira, J.M., de Moura, A.W.A., Ferreira, A.D.S., Bezerra, R.P., de Siqueira Figueiredo, D., de Farias, K.F., de Andrade, T.G., de Souza Figueiredo, E.V.M., 2017. Association of TNFA (-308G/A), IFNG (+874 A/T) and IL-10 (-819 C/T) polymorphisms with protection and susceptibility to dengue in Brazilian population. Asian Pacific journal of tropical medicine 10, 1065–1071.

Shriner, D., Adeyemo, A., Rotimi, C.N., 2011. Joint ancestry and association testing in admixed individuals. PLoS Comput Biol 7, e1002325.

Sierra, B., Triska, P., Soares, P., Garcia, G., Perez, A.B., Aguirre, E., Oliveira, M., 2017. OSBPL10, RXRA and lipid metabolism confer African-ancestry protection against dengue haemorrhagic fever in admixed Cubans. PLoS Pathog 13, e1006220.

Silva, L.K., Blanton, R.E., Parrado, A.R., Melo, P.S., Morato, V.G., Reis, E.A., Dias, J.P., Castro, J.M., Vasconcelos, P.F., Goddard, K.A., Barreto, M.L., Reis, M.G., Teixeira,

M.G., 2010. Dengue hemorrhagic fever is associated with polymorphisms in JAK1. Eur J Hum Genet 18, 1221-1227.

Simon-Loriere, E., Lin, R.J., Kalayanarooj, S.M., Chuansumrit, A., Casademont, I., Lin, S.Y., Yu, H.P., Lert-Itthiporn, W., Chaiyaratana, W., Tangthawornchaikul, N., Tangnararatchakit, K., Vasanawathana, S., Chang, B.L., Suriyaphol, P., Yoksan, S., Malasit, P., Despres, P., Paul, R., Lin, Y.L., Sakuntabhai, A., 2015. High Anti-Dengue Virus Activity of the OAS Gene Family Is Associated With Increased Severity of Dengue. J Infect Dis 212, 2011-2020.

Simons, Y.B., Turchin, M.C., Pritchard, J.K., Sella, G., 2014. The deleterious mutation load is insensitive to recent population history. Nat Genet 46, 220-224.

Soares, P., Alshamali, F., Pereira, J.B., Fernandes, V., Silva, N.M., Afonso, C., Costa, M.D., Musilova, E., Macaulay, V., Richards, M.B., Cerny, V., Pereira, L., 2012. The Expansion of mtDNA Haplogroup L3 within and out of Africa. Mol Biol Evol 29, 915-927.

Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Rohl, A., Salas, A., Oppenheimer, S., Macaulay, V., Richards, M.B., 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. Am J Hum Genet 84, 740-759.

Stoler, J., Al Dashti, R., Anto, F., Fobil, J.N., Awandare, G.A., 2014. Deconstructing "malaria": West Africa as the next front for dengue fever surveillance and control. Acta tropica 134, 58-65.

Taqi, M.M., Waseem, D., Ismatullah, H., Haider, S.A., Faisal, M., 2016. In silico transcriptional regulation and functional analysis of dengue shock syndrome associated SNPs in PLCE1 and MICB genes. 16, 335-345.

Tassaneetrithep, B., Burgess, T.H., Granelli-Piperno, A., Trumpfheller, C., Finke, J., Sun, W., Eller, M.A., Pattanapanyasat, K., Sarasombath, S., Birx, D.L., Steinman, R.M., Schlesinger, S., Marovich, M.A., 2003. DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells. The Journal of experimental medicine 197, 823-829.

Vargas-Castillo, A.B., Ruiz-Tovar, K., Vivanco-Cid, H., Quiroz-Cruz, S., Escobar-Gutierrez, A., Cerna-Cortes, J.F., Vaughan, G., Fonseca-Coronado, S., 2017. Association of Single-Nucleotide Polymorphisms in Immune-Related Genes with Development of Dengue Hemorrhagic Fever in a Mexican Population. Viral immunology.

Vaughn, D.W., Green, S., Kalayanarooj, S., Innis, B.L., Nimmannitya, S., Suntayakorn, S., Endy, T.P., Raengsakulrach, B., Rothman, A.L., Ennis, F.A., Nisalak, A., 2000.

Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. J Infect Dis 181, 2–9.

Viechtbauer, W., 2010. Conducting Meta-Analyses in R with the metafor Package. J Stat Softw 36, 1–48.

Wang, L., Chen, R.F., Liu, J.W., Lee, I.K., Lee, C.P., Kuo, H.C., Huang, S.K., Yang, K.D., 2011. DC-SIGN (CD209) Promoter -336 A/G polymorphism is associated with dengue hemorrhagic fever and correlated to DC-SIGN expression and immune augmentation. PLoS Negl Trop Dis 5, e934.

Warke, R.V., Xhaja, K., Martin, K.J., Fournier, M.F., Shaw, S.K., Brizuela, N., de Bosch, N., Lapointe, D., Ennis, F.A., Rothman, A.L., Bosch, I., 2003. Dengue virus induces novel changes in gene expression of human umbilical vein endothelial cells. Journal of virology 77, 11822–11832.

Whitehorn, J., Chau, T.N., Nguyet, N.M., Kien, D.T., Quyen, N.T., Trung, D.T., Pang, J., Wills, B., Van Vinh Chau, N., Farrar, J., Hibberd, M.L., Khor, C.C., Simmons, C.P., 2013. Genetic variants of MICB and PLCE1 and associations with non-severe dengue. PLoS One 8, e59067.

WHO, 1997. Dengue haemorrhagic fever: diagnosis, treatment, prevention and control., 2nd ed. World Health Organization, Geneva.

WHO, 2009. Dengue: guidelines for diagnosis, treatment, prevention and control. World Health Organization, Geneve, p. 160.

WHO, 2012. Accelerating work to overcome the global impact of neglected tropical diseases. A roadmap for implementation. World Health Organization, Geneva.

Wong, L.P., Ong, R.T., Poh, W.T., Liu, X., Chen, P., Li, R., Lam, K.K., Pillai, N.E., Sim, K.S., Xu, H., Sim, N.L., Teo, S.M., Foo, J.N., Tan, L.W., Lim, Y., Koo, S.H., Gan, L.S., Cheng, C.Y., Wee, S., Yap, E.P., Ng, P.C., Lim, W.Y., Soong, R., Wenk, M.R., Aung, T., Wong, T.Y., Khor, C.C., Little, P., Chia, K.S., Teo, Y.Y., 2013. Deep whole-genome sequencing of 100 southeast Asian Malays. Am J Hum Genet 92, 52–66.

Xavier-Carvalho, C., Cardoso, C.C., de Souza Kehdy, F., Pacheco, A.G., Moraes, M.O., 2017. Host genetics and dengue fever. Infect Genet Evol 56, 99–110.

Xavier-Carvalho, C., Gibson, G., Brasil, P., Ferreira, R.X., de Souza Santos, R., Goncalves Cruz, O., de Oliveira, S.A., de Sa Carvalho, M., Pacheco, A.G., Kubelka, C.F., Moraes, M.O., 2013. Single nucleotide polymorphisms in candidate genes and dengue severity in children: a case-control, functional and meta-analysis study. Infect Genet Evol 20, 197–205.

Zhang, W., Meehan, J., Su, Z., Ng, H.W., Shu, M., Luo, H., Ge, W., Perkins, R., Tong, W., Hong, H., 2014. Whole genome sequencing of 35 individuals provides insights into the genetic architecture of Korean population. BMC Bioinformatics 15 Suppl 11, S6.

Zollner, S., Pritchard, J.K., 2007. Overcoming the winner's curse: estimating penetrance parameters from case-control data. Am J Hum Genet 80, 605–615.

Paper II – Joint ancestry and association test indicate two distinct pathogenic pathways involved in classical dengue fever and dengue shock syndrome

PLOS NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Joint ancestry and association test indicate two distinct pathogenic pathways involved in classical dengue fever and dengue shock syndrome

Marisa Oliveira[1,2,3,4], Worachart Lert-itthiporn[5], Bruno Cavadas[1,2,3], Verónica Fernandes[1,2], Ampaiwan Chuansumrit[6], Orlando Anunciação[2], Isabelle Casademont[4,7], Fanny Koeth[4,7], Marina Penova[4,7,8], Kanchana Tangnararatchakit[6], Chiea Chuen Khor[9,10], Richard Paul[4,7,11], Prida Malasit[12,13], Fumihiko Matsuda[7,8], Etienne Simon-Lorière[4,7,11], Prapat Suriyaphol[5], Luisa Pereira[1,2,14‡]*, Anavaj Sakuntabhai[4,7,11‡]*

1 i3S—Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal, 2 Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), Porto, Portugal, 3 Instituto de Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto, Porto, Portugal, 4 Functional Genetics of Infectious Diseases Unit, Institut Pasteur, Paris, France, 5 Bioinformatics and Data Management for Research, Office for Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, 6 Department of Pediatrics, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand, 7 Pasteur Kyoto International Joint Research Unit for Integrative Vaccinomics, Kyoto, Japan, 8 Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, 9 Genome Institute of Singapore, A-STAR, Singapore, Singapore, 10 Department of Biochemistry, National University of Singapore, Singapore, Singapore, 11 CNRS, Unité de Recherche Associée 3012, Paris, France, 12 Dengue Hemorrhagic Fever Research Unit, Office for Research and Development, Siriraj Hospital, Faculty of Medicine, Mahidol University, Bangkok, Thailand, 13 Medical Biotechnology Unit, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Pathumthani, Thailand, 14 Department of Pathology, Faculty of Medicine, University of Porto, Porto, Portugal

‡ LP and AS jointly directed this work.
* luisap@ipatimup.pt (LP); anavaj.sakuntabhai@pasteur.fr (AS)

Abstract

Ethnic diversity has been long considered as one of the factors explaining why the severe forms of dengue are more prevalent in Southeast Asia than anywhere else. Here we take advantage of the admixed profile of Southeast Asians to perform coupled association-admixture analyses in Thai cohorts. For dengue shock syndrome (DSS), the significant haplotypes are located in genes coding for phospholipase C members (*PLCB4* added to previously reported *PLCE1*), related to inflammation of blood vessels. For dengue fever (DF), we found evidence of significant association with *CHST10*, *AHRR*, *PPP2R5E* and *GRIP1* genes, which participate in the xenobiotic metabolism signaling pathway. We conducted functional analyses for *PPP2R5E*, revealing by immunofluorescence imaging that the coded protein co-localizes with both DENV1 and DENV2 NS5 proteins. Interestingly, only DENV2-NS5 migrated to the nucleus, and a deletion of the predicted top-linking motif in NS5 abolished the nuclear transfer. These observations support the existence of differences between serotypes in their cellular dynamics, which may contribute to differential infection outcome risk. The contribution of the identified genes to the genetic risk render Southeast and North-east Asian populations more susceptible to both phenotypes, while African populations are

129

best protected against DSS and intermediately protected against DF, and Europeans the best protected against DF but the most susceptible against DSS.

Author summary

Dengue fever is endemic in tropical and subtropical areas of East Asia and America, but globalization and climate changes are introducing vector and virus to the naïve regions of Europe and North America. In this work we conducted a statistically robust, coupled association-admixture test in two dengue cohorts from Thailand (classical dengue fever, DF, and dengue shock syndrome, DSS) and a published Vietnamese (DSS only) cohort. We identified new candidate genes associated with DF risk and confirmed known gene family association with DSS risk. In DF, phosphatase control is crucial, including through binding to viral proteins, as we showed for PPP2R5E protein co-localization with DENV1 and DENV2-NS5 proteins within liver cells and differential cellular localizations along time. In DSS, cytokine dynamics, inflammation and activation of vascular endothelium cells are dominant features. The particular genetic risk conferred by these genes indicates that Southeast and Northeast Asians are highly susceptible to both phenotypes, while Africans are best protected against DSS, and Europeans best protected against DF but the most susceptible against DSS.

Introduction

Dengue virus (DENV) is the most common mosquito-borne viral infection, infecting approximately 390 million people per year worldwide with one quarter developing dengue disease (MIM: 614371) [1]. Symptoms range from undifferentiated fever, classical dengue fever (DF) to shock syndrome (DSS; hemorrhage, plasma leakage and vital organ impairment) [2].

Recent -omic approaches provide unbiased genomic insights into mechanisms associated with dengue disease. There has been only one publication on classical genome wide association study (GWAS) of dengue [3] compared to a considerable number of transcriptomic studies [4–7]. The reason for this discrepancy is that cohorts of thousands of individuals are required for GWAS to reach genome wide significance. The GWAS work conducted on a cohort of Vietnamese children [3] included 2,008 DSS samples versus 2,018 controls, replicated in 1,737 versus 2,934, and found SNPs in genes *MICB* and *PLCE1* associated with DSS phenotype. Lately, analytical improvements based on admixture mapping have reduced the sample size requirement from thousands to hundreds of individuals or even fewer [8]. Most human populations have some degree of ancestry admixture, which brings together haplotypes that occur at different frequencies in parental populations. Admixture mapping analyses these blocks across the mosaic descendant chromosomes and allows to compare their distribution between case and control cohorts. The lower number of blocks compared with individual SNPs reduces considerably the statistical burden. We have successfully conducted such an admixture study in dengue cohorts from Cuba [9], and identified two genes involved in lipid metabolism which showed to be protective against the risk of dengue hemorrhagic fever, a protection conferred by the African inherited ancestry. Whereas for *RXRA* gene there was already functional evidence of its involvement in infection [10], we also demonstrated functionally by shRNA that the knockdown of *OSBPL10* gene had a significant negative impact in DENV replication rate [9].

130

Epidemiologic reports have shown the existence of ethnic differences in susceptibility to dengue fever not only in Cuba [11] but also in Malaysia [12] where the incidence rate by ethnic group was 3.7:1:1.3 for Chinese, Malays and Indians, respectively, in the years 1970's and 1980's, although no cross-evaluation was performed with other socio-demographic factors. In the present study, we take advantage of the admixed profile of Southeast Asians (in the nexus between South, Northeast and Southeast Asia) to perform coupled association-admixture analyses (BMIX; [13]) of case/control cohorts of dengue patients: Thai dengue patients who developed DF (n = 252) or DSS (n = 159), and a control blood donor group (n = 290); and the published Vietnamese dataset (2018 controls and 2008 DSS patients;[3]). Although the admixture in the region has been taking place along time, since the first arrival of modern human after the out-of-Africa migration, a considerable migration from south China began in the XV century and increased in the XIX and XX centuries, mainly towards Thailand where about 40% of the population has some Chinese admixture and 14% are identifiable Thai Chinese [14]. This is a similar scenario to the admixture that took place in the Americas, where these local admixture inference tools have been successfully applied [9, 15, 16]. We were able to identify distinct candidate genes conferring susceptibility/resistance to the risk of DF and DSS, arguing in favor of independent pathogenic mechanisms for the establishment of the two phenotypes. We further confirmed that one DF candidate gene codes for a human protein that co-localizes with the DENV1 and DENV2-NS5 proteins, and, in the latter case, transiently relocated from the cytoplasm to the nucleus. We also inferred the relative worldwide genetic risks contributed by the detected candidate genes based on their frequencies for the susceptible/resistant haplotypes.

Results

Ancestry of Thai and Vietnamese cohorts

All analyzed individuals have some degree of admixture (Fig 1; S1 Fig). The Northeast Asian background is dominant in Vietnam (77.3%) and decreases in Thailand (56.4%), in contrast to the Southeast Asian component, which increases from 20.7% in Vietnam to 35.1% in Thailand. The South Asian influence is 8.5% in Thailand and 2.0% in Vietnam. Within the dengue cohorts, we observed a statistically significant increase in the Southeast Asian background in Thailand for both DF (4.1% increase; p-value = 1.25×10^{-7}) and DSS (4.8% increase; p-value = 5.90×10^{-8}) compared to Thai control.

DSS cohorts analyses

We began by checking if the BMIX results on the published Vietnamese cohort [3] are in accordance with the results from the classical association mapping, a test of the robustness of the algorithm. BMIX indicates also the association of DSS with *MICB* and *PLCE1* genes (Table 1, Fig 2A, S1 and S4 Tables). The identified region surrounding *MICB* encompasses seven significant SNPs, placed along 165,080 bp, from the downstream *MICA* to the upstream *LTB* gene, a region highly rich in genes. Three linked (S2 Fig) SNPs in *MICB* have the most significant p-values, forming the protective haplotype GTT (OR = 0.77; p-value<0.0001), which is the most frequent haplotype in worldwide populations (Fig 3C). The susceptible *MICB* haplotype ACC (OR = 1.39; p-value<0.0001) is more frequent in Europeans and South Asians (0.18 to 0.34). The two SNPs found for *PLCE1* reached significant p-values and are almost in complete linkage (S5 Fig). The DSS protective *PLCE1* haplotype (CG; OR = 0.75; p-value< 0.0001) is more frequent (Fig 3B) in Northeast Asia (0.12–0.28) and Southeast Asia (0.19), followed by Europe (0.04–0.14) and absent in Africa.

131
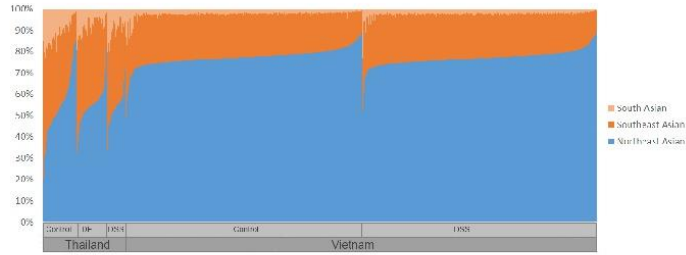
Fig 1. Global ancestry inferred through RFMix when using three parental ancestries (South, Northeast and Southeast Asian) for the global dataset. Each vertical line represents an individual, and the three colours represent the proportion of the three parental populations in each genome (light orange for South Asian, dark orange for Southeast Asian and blue for Northeast Asian).

We further analyzed the Thai DSS *vs*. control cohort (Table 1, Fig 2B, S2 and S5 Tables), and obtained a reliable signal of six linked (S3 Fig) significant SNPs for *PLCB4* (phospholipase C, beta 4; S4A Fig), a gene in the same family as *PLCE1*, and participating in many common pathways, such as dendritic cell maturation, PI3K signaling in B lymphocytes and PPARA/RXRA activation. The DSS protective *PLCB4* haplotype (GAGAGG; OR = 0.58; p-value =

Table 1. Odds ratios (ORs), 95% confidence intervals and Yates p-values (corrected for continuity) of the $\chi 2$ test for the significant haplotypes/SNPs in the phenotype and populations for which association was detected.

Gene	Protective haplotype				Susceptible haplotype			
	Frequency Control	Frequency Case	OR (95 CI)	P value	Frequency Control	Frequency Case	OR (95 CI)	P value
	Vietnamese DSS study							
MICB	rs2534666-rs2855807-rs3132468							
	G-T-T				A-C-C			
	0.78	0.73	0.77 (0.70–0.85)	<1.0 x 10^{-4}	0.13	0.18	1.39 (1.23–1.57)	<1.0 x 10^{-4}
PLCE1	rs3740360-rs2274223							
	C-G				A-A			
	0.27	0.22	0.75 (0.68–0.84)	<1.0 x 10^{-4}	0.69	0.75	1.30 (1.18–1.43)	<1.0 x 10^{-4}
	Thai DSS study							
PLCB4	rs16995800-rs2299676-rs7269910-rs1997696-rs6133707-rs6056595							
	G-A-G-A-G-G				A-G-A-C-A-A			
	0.17	0.10	0.58 (039–0.88)	1.3 x 10^{-2}	0.61	0.70	1.48 (1.10–1.98)	1.1 x 10^{-2}
	Thai DF study							
CHST10	rs4850931-rs1030902-rs2241811-rs2241810-rs4149518-rs2241809-rs4149510-rs4851313-rs3828193							
	C-T-C-T-A-C-G-G-G				T-G-T-C-G-T-A-A-T			
	0.26	0.17	0.59 (0.44–0.79)	5.2 x 10^{-4}	0.62	0.74	1.78 (1.37–2.31)	<1.0 x 10^{-4}
AHRR	rs6555205-rs2721020							
	T-C				C-T			
	0.25	0.16	0.54 (0.40–0.74)	1.2 x 10^{-4}	0.67	0.79	1.89 (1.43–2.48)	<1.0 x 10^{-4}
PPP2R5E	rs3829766-rs6573513-rs743221-rs7144210							
	A-C-G-G				GTAA			
	0.43	0.32	0.62 (0.48–0.79)	2.0 x 10^{-4}	0.39	0.50	1.59 (1.25–2.03)	2.0 x 10^{-4}
GRIP1	rs1480010							
	T				C			
	0.17	0.11	0.56 (0.39–0.80)	1.6 x 10^{-3}	0.83	0.89	1.79 (1.26–2.56)	1.6 x 10^{-3}

132

Fig 2. Manhattan plots of BMIX analysis in Vietnamese DSS vs Control (A), Thai DSS vs Control (B) and Thai DF vs Control (C) for Northeast and Southeast Asian ancestries. The red line represents the significance threshold. The protein coding genes with significantly associated SNPs are identified.

0.013) is rare in most worldwide populations (Fig 3A), reaching the highest frequencies in Africa (0.21–0.28). Only one *PLCE1* SNP (rs2274223) was present in the chip used in the Thai dengue cohort and it did not reach significance.

Individually, the conventional association study with PCA correction for population stratification in Thai DSS *vs.* control could not identify any candidate gene when correcting for multiple test (S6A Fig–the two singled out significant SNPs are spurious signals as linked SNPs do not display significant p-values). We also tested 10 runs of pseudo datasets, permutating case and control labels (S1 Text). No SNP is significant in the association tests, and the BMIX

133
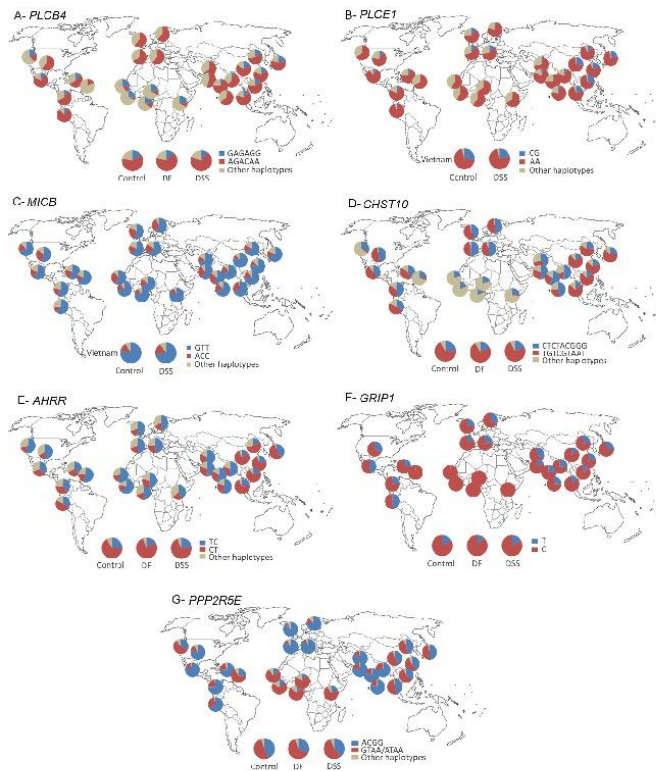
Fig 3. Worldwide (from the 1000 Genomes database) and Thai dengue cohorts (control, DF and DSS) frequencies for significantly associated haplotypes in the various genes. A- *PLCB4*; B- *PLCE1*; C- *MICB*; D- *CHST10*; E- *AHRR*; F- *GRIP1*; G- *PPP2R5E*. The protective and causative haplotypes are highlighted.

algorithm identified spurious significant SNPs (mostly isolated in different chromosomes) that do not replicate between runs and that are different from the case-control comparison. Overall posterior p-values were also lower in the pseudo datasets. The higher spurious detections in BMIX than in the association test agree with the fact that the statistical burden of the local ancestry test is considerably lower than the one for the association test, which raises the possibility of detecting a positive signal. The randomness between runs reflects the high variability between individuals in admixture percentages and in distribution of ancestry blocks along the genomes. This argues for a double-careful interpretation of BMIX results in the context of the disease. For the Thai DSS *vs.* control, the fact that the *PLCB4* gene belongs to the same family of the previously independently identified *PLCE1* gene is an important additional evidence for considering that gene a strong candidate in DSS phenotype.

Calculating the genetic risk of DSS according to the worldwide population frequency of the phospholipase C and *MICB* protective and susceptible haplotypes (Fig 4A), it can be observed that African and descendent Caribbean populations are best protected, while European, Asian and Latin American populations are more susceptible to DSS.

134
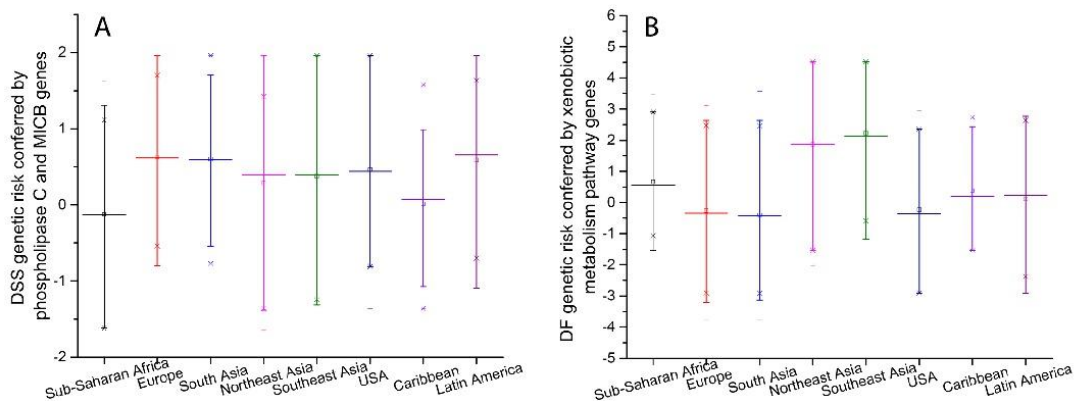
Fig 4. Genetic risk for the various worldwide regions by considering an additive model of protective and causative haplotypes/SNPs for DSS (A) and DF (B). Median (middle line), mean (little square), 95% confidence interval (whiskers) and extreme values (crosses) are indicated.

We genotyped two of the BMIX-identified significantly associated *PLCB4* SNPs in further Thai control (n = 244) and case (n = 20) samples (S7 Table), and the rs1997696 SNP presents a p-value over the significance threshold of a traditional GWAS (p = 4.7×10^{-8}).

Dengue fever cohort analysis

When comparing Thai DF *vs.* control, a distinctive genetic signature was obtained. Three genes located on different chromosomes had at least two SNPs above the BMIX significant posterior probability threshold of 0.5 (Fig 2C), forming haplotypes (S7, S8 and S9 Figs). *CHST10* codes for carbohydrate sulfotransferase 10 (S4C Fig), has nine significant SNPs (Table 1, S3 and S6 Tables), forming the protective haplotype CTCTACGGG (OR = 0.59; p-value = 0.0005), whereas the haplotype TGTCGTAAT increased risk of DF (OR = 1.78; p-value<0.0001). The protective haplotype is frequent in South Asian populations (0.38–0.57), whereas the susceptible haplotype is frequent in Northeast Asia (0.61–0.74) and very rare in the African populations (Fig 3D). *AHRR* (S4D Fig) codes for aryl-hydrocarbon receptor (*AHR*) repressor, has two significant SNPs, and similarly to *CHST10*, the protective *AHRR* haplotype (TC—OR = 0.54; p-value = 0.0001) is more frequent in South Asian and African populations (between 0.40–0.60) and the opposite haplotype (CT—OR = 1.89; p-value<0.0001) is more frequent in Northeast Asian populations (0.57–0.78) (Fig 3E). *PPP2R5E* (S4F Fig) codes for protein phosphatase 2 (PP2A), regulatory subunit B', epsilon isoform (also known as PP2A-B56), has four significant SNPs, whose protective haplotype (ACGG—OR = 0.62; p-value = 0.0002) showed high frequency in South Asian populations (0.76–0.86), while African populations have the lowest frequency of this haplotype (0.09–0.20) (Fig 3G). Interestingly, the proteins coded by these three genes, and by another gene, *GRIP1* (S4E Fig) that codes for glutamate receptor interacting protein 1, with one significant SNP (T; OR = 0.56; p-value = 0.0016), are involved in the xenobiotic metabolism signaling pathway (Fig 5D). The *GRIP1* protective allele is more frequent in South Asian populations (0.19–0.32) and absent in Africa (Fig 3F).

Again, individually, the conventional association study with PCA correction for population stratification could not identify any candidate gene when correcting for multiple test (S6B Fig).

135
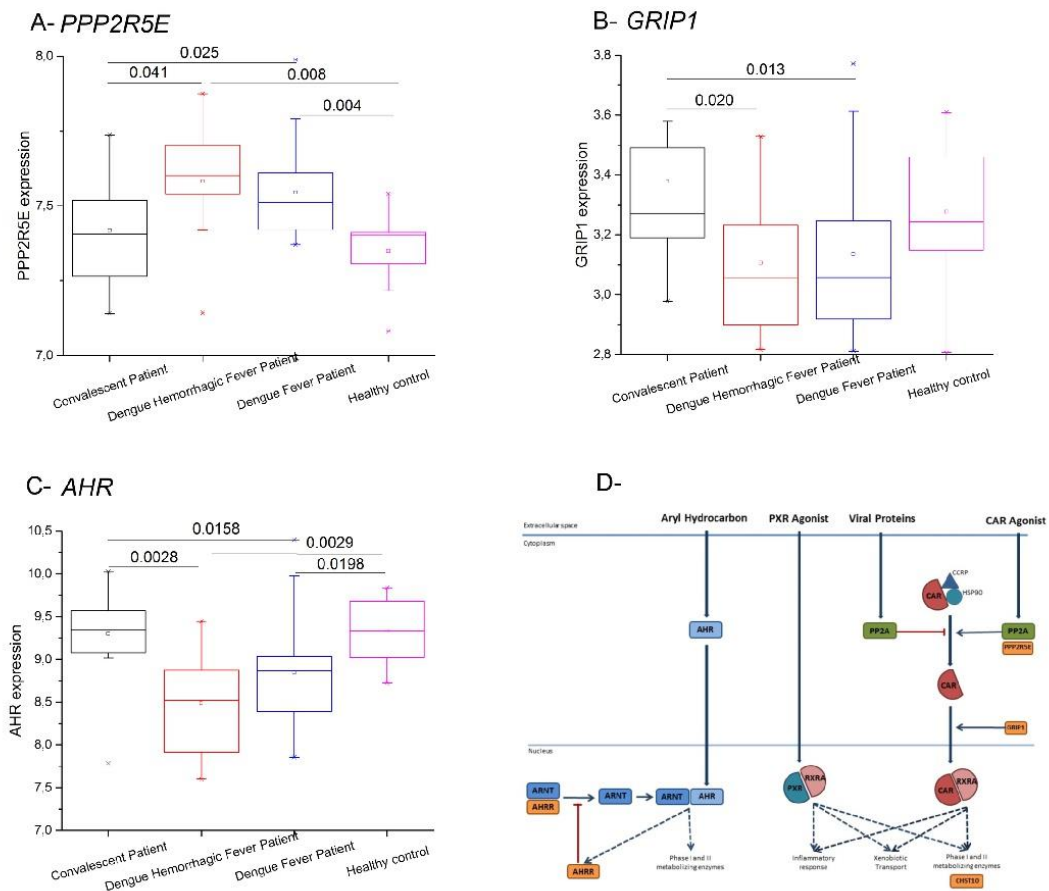
Fig 5. **Significantly altered gene expression for *PPP2R5E* (A), *GRIP1* (B) and *AHR* (C) in Thai dengue cohort along the course of disease from a transcriptome dataset for whole blood[17].** Significant p-values are indicated. D) Scheme of the xenobiotic metabolism signalling pathway (based on Ingenuity database information), highlighting the three nuclear transcription factors: the constitutive active receptor (CAR); the pregnane X receptor (PXR); and the aryl hydrocarbon receptor (AHR).

The contribution to the genetic risk to DF, inferred from an additive model combining the protective and susceptible haplotypes of the four xenobiotic-related genes (Fig 4B), indicates highest protection in European and South Asian populations and highest risk in Northeast and Southeast Asians. African and Latin American/Caribbean populations have an intermediate risk conferred by these genes to DF.

The genotyping of six SNPs in these four genes in additional Thai controls (n = 245) and cases (n = 55) improves p-values of a traditional association test in the total cohort to levels of 10^{-5} in four SNPs and 10^{-4} in two SNPs (S7 Table). These values are significant after Bonferroni correction for the set-test of six SNPs.

136

We analyzed the expression of these genes in the xenobiotic pathway in a transcriptome dataset including patients sampled during acute phase of DF, DHF and convalescence compared with controls [17]. *CHST10* and *AHRR* expressions did not significantly change during dengue infection (S10 Fig), however, there was a significant increase in *PPP2R5E* expression and a significant decrease in *AHR* (negatively regulated by *AHRR*) and *GRIP1* expressions during acute dengue infection (Fig 5A, 5B and 5C). These findings are further evidence that *PPP2R5E*, *GRIP1* and *AHR* can be involved in dengue infection and development of dengue disease. We further checked in the GTEx database if the DF candidate SNPs act as eQTLs. All candidate protective alleles in *PPP2R5E* and *AHRR* genes significantly reduce the expression of the respective proteins (S11 Fig). The candidate SNP in *GRIP1* gene is not an eQTL in the GTEx cohort, and the two eQTLs (rs11176317 and rs12322014) close to the candidate rs1480010 are not in LD with it. As the GTEx cohort is mainly of European ancestry, we cannot ascertain if this *GRIP1* SNP or other linked SNPs can be eQTLs in Asian populations.

Immunofluorescence co-localization imaging of PPP2R5E and NS5 protein from DENV1 and 2

The recent identification of conserved motifs that provide binding specificity to the PP2A-B56 phosphatase [18] led us to further test the hypothesis of the potential binding of this regulatory region of PP2A protein to DENV proteins. We began by performing an *in silico* search [19] for the high-affinity LxxIxE motif as well as the intermediate- and low-affinity motifs in the protein reference sequences of the four DENV serotypes (S8 Table and Fig 6A). NS5 presents between three and six motifs in all four DENV serotypes, and at least two of these motifs (LxxIxE and LxxVxE) are highly conserved. Other viral proteins also bear motifs, but are more heterogeneous between DENV serotypes.

We then tested the hypothesis that PP2A-B56 can interact with NS5 by conducting confocal immunofluorescence co-localization tests. We transfected Huh7 cells with a mammalian expression plasmid containing DENV2-NS5 tagged with an orange fluorescence protein. We fixed and stained with antibody against PPP2R5E at 24h, 48h and 72h after transfection. In non-transfected cells, PPP2R5E is localized in the cytoplasm (Fig 6B). At 24h of post-transfection, both PPP2R5E and NS5 are localized in the cytoplasm, but by 48h they both co-localize in the nucleus, and at 72h PPP2R5E returns to the cytoplasm while NS5 remains in the nucleus (Fig 6C). We then deleted the xLxxIxE motif in our DENV2-NS5 vector (Fig 6D) and transfected cells in the same way. The deletion of this motif prevented the translocation of the viral NS5 protein to the nucleus (Fig 6D).

Testifying to the existence of differences between serotypes, the immunofluorescence co-localization test between PPP2R5E and DENV1-NS5 (Fig 6E) showed that the two proteins co-localize in the cytoplasm, but the entrance in the nucleus is almost negligible, and little accumulation of NS5 can be detected in the nucleus at 72h.

Discussion

Our successful association-admixture analyses in Thai population have provided evidence that different genes/pathways contribute to the genetic susceptibility or resistance to different outcome of dengue infection. We suggest that xenobiotics and lipid metabolism, as well as interaction of viral proteins to these molecules and to its phosphatases, are critical in the development of classical DF, whereas more severe forms of dengue are caused by over reactive immunity leading to cytokine storm and/or defect in endothelial cell dysfunction and coagulation system.

137

A

Motif sequence	Start	Stop	Disorder score	Serotype	Domain
LKKVTEVKG	2587	2595	0.30	DENV1	NS5 Methyltransferase
LCDIGESSP	2637	2645	0.40	DENV1	NS5 Methyltransferase
ITDIMEPEH	3045	3053	0.31	DENV1	NS5 RNA Dependent RNA Pclymerase
LKKVTEVKG	2585	2593	0.34	DENV2	NS5 Methyltransferase
LCDIGESSP	2635	2643	0.43	DENV2	NS5 Methyltransferase
LGFLNEDHW	2980	2988	0.14	DENV2	NS5
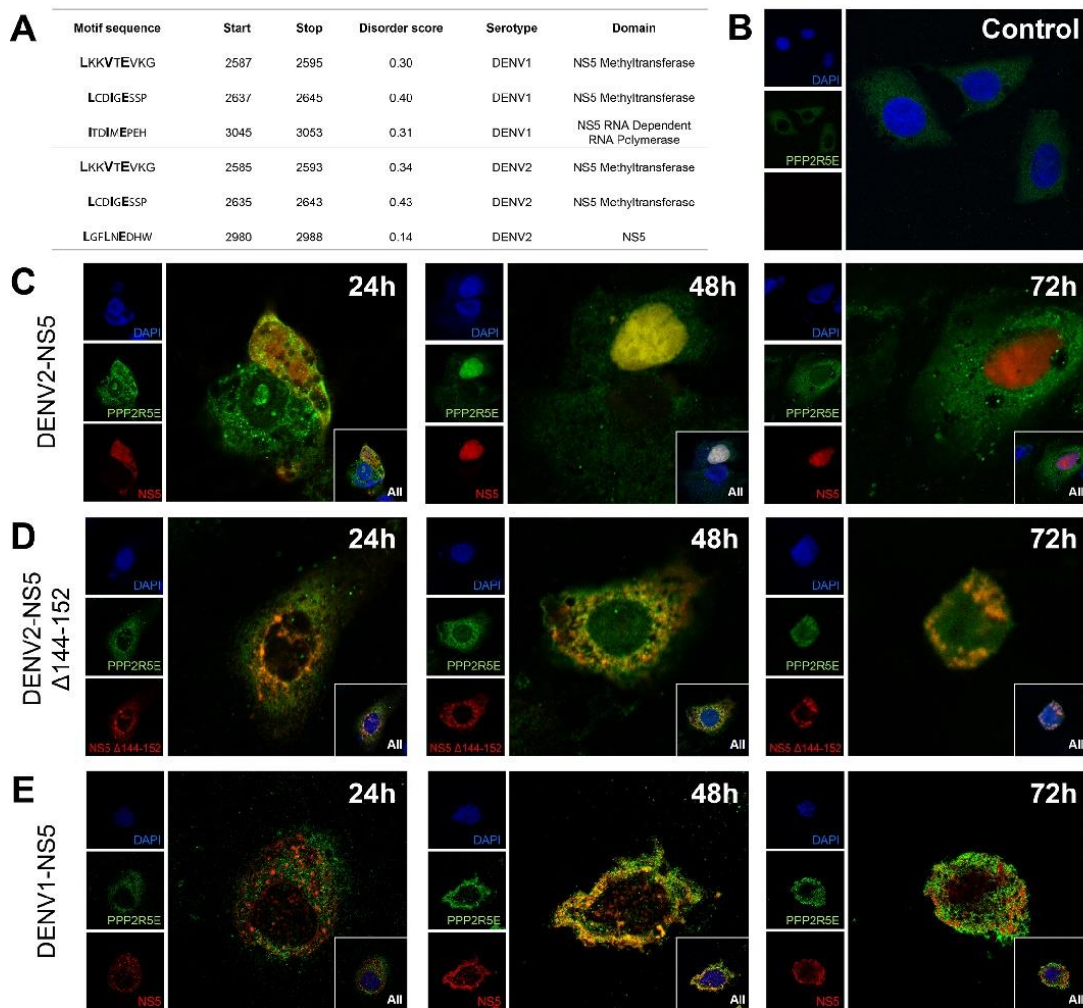
Fig 6. Confocal imaging of PPP2R5E and NS5 from DENV1 and DENV2. A. The main link motifs in NS5 proteins from DENV1 and DENV2. B. Subcellular localization of PPP2R5E in Huh7 control cells. C. Subcellular localization of PPP2R5E in Huh7 control and after 24h, 48h and 72h of transfection with DENV2-NS5 protein. D. Subcellular localization of mutated PPP2R5E in Huh7 control and after 24h, 48h and 72h of transfection with DENV2-NS5 protein. E. Subcellular localization of PPP2R5E in Huh7 control and after 24h, 48h and 72h of transfection with DENV1-NS5 protein. Green immunofluorescence indicates PPP2R5E, red indicates NS5, blue flags nucleus. Yellow signals indicate co-localization of NS5 and PPP2R5E, and was obtained by overlapping the two panels. Original magnification, ×630.

https://doi.org/10.1371/journal.pntd.0006202.g006

We reinforced the association of phospholipase C gene family with the DSS phenotype in Thai patients, as had been found in Vietnamese DSS patients [3], but this time the gene detected was *PLCB4* instead of *PLCE1*. These enzymes have been implicated in a high number of signal transduction pathways [19] and on immune regulation [20]. More recently, Lin et al.

138

[21] found *PLCB4/PLCB1* susceptibility loci for coronary artery aneurysm in Kawasaki disease when analyzing a Han Chinese cohort. During the acute stage of this disease, inflammation occurs by the infiltration of T cells and macrophages and the activation of vascular endothelium cells (ECs) with increased serum proinflammatory cytokines and predominant damage of small-, medium-sized vessels and the coronary artery. The injured vascular tissues show subendothelial edema, vascular damage, gap formation, and fenestration of ECs. It thus seems that phospholipase C genes are involved in several diseases presenting the phenotype of inflammation of the blood vessels, as it is the case in dengue shock syndrome.

The association of the four genes (*CHST10*, *AHRR*, *PPP2R5E* and *GRIP1*) from the xenobiotic metabolism signaling pathway with Thai DF patients is the first genetic evidence for its implication in dengue pathogenesis, although functional studies have already indicated this association before. Xenobiotics are toxic non-endogenous compounds that together with other toxic endogenous compounds must be eliminated from the body by drug/xenobiotic metabolizing enzymes (DME/XME) and transporters. The DMEs are induced by their own substrates, through signaling cascades involving three specific receptors: the constitutive active receptor (CAR), the pregnane X receptor (PXR) and the aryl hydrocarbon receptor (AHR) (Fig 5D) [22]. In particular PP2A regulates the CAR:HSP90 complex, allowing CAR release and its eventual translocation to the nucleus [23], an event that could be similar/parallel to PP2A involvement in the nuclear translocation of viral proteins. Several lines of evidence show that some of PP2A regulatory (such as PP2A-B56) and scaffold subunits can bind viral proteins [24–26]. In addition, the capsid of West Nile virus, a flavivirus related to DENV, binds to the inhibitor of PP2A proteins (I_2^{PP2A} or SET) in the precise site where I_2^{PP2A} binds to PP2A, causing an increase of PP2A activity in several cell types [27]. These evidences led us to perform immunofluorescence assays that revealed that PPP2R5E co-localizes with the NS5 protein of DENV2, first in the cytoplasm and then in the nucleus, and that this nuclear translocation does not take place when the specific link motif is deleted from the NS5 protein. NS5 is a crucial viral protein responsible for the virus replication at the endoplasmic reticulum. NS5 accumulation in the nucleus seems to occur late in DENV infection as a hyperphosphorylated form unable to bind NS3. Adding a layer of complexity to these events, the test of the NS5 nuclear translocation in the four DENV serotypes showed that it only occurs in DENV2 and DENV3 [28], leading the authors to hypothesize that the NS5 nuclear localization is not strictly required for virus replication but that it is more likely to have an auxiliary function in the life cycle of specific DENV serotypes. Our *in silico* analysis indicates that all NS5 proteins from the four DENV strains contain the specific link-motifs, but our immunofluorescence assay with DENV1-NS5 confirmed the observation of the authors, that the protein of this serotype does not enter the nucleus. Extra factors, such as importin proteins, must contribute to this differential import into the nucleus. Thus, the dynamics of NS5 seems to play a major role in dengue infection, potentially impacting differential strain virulence, and the PPP2R5E-NS5 interaction must be taken into consideration in future studies.

Increased PP2A activity could favor viral infection not only through binding of viral proteins but also through regulation of regulatory T cells (Treg) [29]. Suppression and impairment of anti-viral activity of interferon α (IFNα) has been shown in hepatitis C infection through inhibition of Jak1/Tyk2/STAT1 phosphorylation [30] and upregulation of PP2A dependent upon NS5A protein [24]. Both Treg and Jak1/Tyk2/STAT1 pathway have been shown to be important in DENV infection [31–34].

The significant association detected for *AHRR* and lower expression of AhR during acute dengue infection suggests another possible key factor in dengue infection outcome. AhR has been shown to be involved in mediating the biotransformation and carcinogenic/teratogenic effects of environmental toxins. Recently, its role in innate and adaptive immunity has been

139

demonstrated, through its involvement in regulation of CD4, CD8 and Treg after viral infection [35]. We hypothesize that the outcome of dengue infection depends on a fine-tuning of the xenobiotic metabolism signaling pathway between the AhR and PP2A/CAR pathway. While increased PP2A/CAR activity promotes viral infection, decreased AhR activity results in uncontrolled immune homeostasis leading to dengue disease. The confirmation that our candidate protective alleles in *PPP2R5E* and *AHRR* genes are eQTLs leading to lower expression of the respective proteins supports this hypothesis. When cells are invaded by DENV, there will be an increase in the expression of *PPP2R5E*, possibly due to the interaction with NS5. So Asian individuals that are genetically *PPP2R5E*-low-expressing are protected against this link. For the *AHRR* gene, the protective phenotype also has a lower expression profile, and as this protein will inhibit AHR protein, these individuals will have a higher expression of AHR, which is opposite to the expression pattern observed in the dengue patients transcriptome.

For both DF and DSS phenotypes, Northeast and Southeast Asian populations have a higher ancestral prone risk when compared with other geographical regions, considering the particular genes identified in this work. Specifically, Southeast Asian ancestry has a slightly higher risk for DF than Northeast Asian ancestry (Fig 4B). These genetic predictions agree with observations that almost 75% of the global population exposed to dengue live in Asia-Pacific, with rates of severe dengue being 18 times higher in this region compared with the Americas [36]. African and its descendant populations are the most protected ones against DSS, and displaying an intermediate protection against DF, adding genetic evidence to previous claims that this ancestry is protected against worse dengue phenotypes [37, 38]. Our inferred genetic risk for DF in Africa, slightly higher that the risk in America, agrees quite well with the risk predictions inferred by Bhatt et al. [1] of 16% and 14%, respectively, of the global burden. Climatic change and globalization are enlarging the spread of dengue vector and virus to northern latitudes, putting Europe and North America at risk of autochthonous infections [39]. The considerable number of autochthonous infections that occurred in Madeira Island, Portugal, in 2012/2013 [40] is the first example of a reality that can take place in a near future in continental Europe. The genetic risk calculated here, for the newly and confirmed susceptible/resistant haplotypes, shows that European populations (as well as South Asian and USA) present an even higher risk than Southeast Asian populations to DSS, while they are the best protected ones against DF.

Materials and methods

Samples and genotyping

We enrolled 411 patients (age range, 1–31 years; male to female ratio, 0.984) with symptomatic DENV infection during 2000–2003 from two hospitals in Bangkok (Ramathibodi, and Siriraj) and one in Khonkaen, Thailand (S1 Fig). These patients were first admitted to the hospitals with suspicions of dengue infection based on clinical features, and DENV infection was later confirmed by either detection of viral genome or a comparable immunoglobulin G (IgG) and immunoglobulin M (IgM) titers, measured by an enzyme-linked immunosorbent assay, in late acute and/or convalescent sera. Dengue severity was defined according to 1997 World Health Organization (WHO) criteria [2], and we ended up with the following case cohorts: 252 patients with DF with no evidence of plasma leakage but incapacitating dengue fever; 159 with severe plasma leakage and/or bleeding, leading to shock or profound shock (grades 3 and 4). Information about the DENV serotype and primary/secondary infection are provided in S9 Table. These case samples were genotyped with the Illumina Human 660W Quad BeadChip (Illumina, San Diego, CA, USA). The control cohort was collected in the same hospitals, geographically matching the cases, consisting in 290 healthy individuals with no fever and being

140

treated for minor injuries. Control individuals were genotyped with the Illumina HumanOm-niExpress BeadChip.

Quality control was performed in PLINK [41], and SNPs with more than 5% missing geno-types, minor allele frequency (MAF) below 5%, and Hardy-Weinberg equilibrium (HWE) deviation p-values of less than 0.001 were filtered out from downstream analyses. We also checked visually for outliers in principal-component analysis (PCA; S12 Fig), and excluded samples that had evidence of being a second-degree relative or closer to another sample in the study (identity by descent >30%; or identity by state >90%). All studied samples passed these criteria. SNPs located in X and Y chromosomes and in mitochondrial DNA were removed from the analyses, leading to a final account of common 261,660 autosomal SNPs.

The Vietnamese cohort consisted in 2018 controls and 2008 DSS patients typed with the Illumina Human 660W Quad BeadChip, amounting in 479,905 SNPs after QC.

Eight SNPs detected in the BMIX analyses were genotyped in additional Thai cohort samples (61 DF; 20 DSS; and 250 controls), through TaqMan assays (Life Technologies, Carlsbad, CA, USA) with commercial probes for the SNPs. The screen was conducted on a QuantStudio 12K Flex (Life Technologies, Carlsbad, CA, USA) and the results were analyzed in TaqMan-Genotyper software (Life Technologies, Carlsbad, CA, USA).

Ethics statement

Written informed consent was obtained from all subjects or, in case of individuals under 18 years of age, from their parents or tutors. The protocol was approved by the ethics committees from the Faculty of Medicine, Ramathibodi Hospital, Mahidol University; the Faculty of Medicine, Siriraj Hospital, Mahidol University; the Khon Kaen Hospital; and the Thailand Ministry of Public Health.

Association test, admixture mapping and BMIX

The tests were conducted in the following comparisons: Thai DF versus controls (DF test); Thai DSS versus controls (DSS test); and Vietnamese DSS versus controls (DSS test). Samples were phased in SHAPEIT v.2 [42] using HapMap reference panel and fine-scale genetic map.

For admixture mapping, we applied the RFMix algorithm [43] and used three ancestral data sets: the phased data from the 1,000 Genomes Database [44] for the Chinese Dai in Xishuang-banna (CDX); and the Indian Telugu from the UK (ITU) representing the Northeast and South Asian ancestries, respectively; and the Malaysian complete genomes from Singapore [45] representing the Southeast Asian ancestry. We checked if the global admixture profile inferred by RFMix would be reproduced when using ADMIXTURE analysis [46] for K = 3 running together the parental populations and Thai and Vietnamese cohorts. As can be verified in S13 Fig, the three-background admixture profiles for the Thai and Vietnamese individuals are iden-tical between RFMix and ADMIXTURE analysis. ITU population is quite homogeneous, while CDX and Malaysians have themselves around 20–30% admixture, but this does not affect the overall proportion inference of the three components. Information on the three ancestry back-grounds was obtained for each locus along chromosomes for every individual, and these values were averaged in each cohort. Two-tailed Mann-Whitney test (non-parametric test, not requir-ing normal distribution) was applied to assess the significance between the global ancestry pro-portions inferred for the three ancestral backgrounds within the Thai and Vietnamese cohorts.

BMIX [13] was implemented on the Thai and Vietnamese groups, by coupling the admix-ture mapping inferred through the RFMix algorithm with the association data. SNPs with a posterior probability of a joint ancestry and association effect equal or higher to 0.5 were con-sidered significant. The annotation of the significant SNPs was inferred by using the Variant

141

Effect Predictor (VEP) tool from Ensembl. The involvement of the significant genes in pathways was checked in the Ingenuity Pathway Database (https://targetexplorer.ingenuity.com/index.htm).

Calculation of genetic risk

Odds ratios (ORs), 95% confidence intervals, beta parameters and Yates p-values (corrected for continuity) for the significant haplotypes/SNPs in DF and DSS phenotypes were calculated in http://vassarstats.net/odds2x2.html. The dengue phenotype (DF or DSS) genetic risk score was calculated by multiplying each individual's significant haplotype/allele (either protective or causative) count for each locus by the respective beta coefficient and summing the product for all loci (as in [47]). For DF, the four genes (*CHST10*, *AHRR*, *PPP2R5E* and *GRIP1*) participating in the xenobiotic metabolism signaling pathway were considered. For DSS, both phospholipase C (*PLCB4* and *PLCE1*) and *MICB* genes were included in the calculation.

Expression data

Expression of candidate genes was checked in a Thai whole blood transcriptome cohort [17] measured with the Human U133 Plus 2.0 Arrays (Affymetrix, Santa Clara, CA, USA), including: nine healthy controls samples; 28 samples collected between days 2 and 9 after onset of symptoms (acute illness) from secondarily infected patients (18 DF and 10 DHF); and 19 samples collected at convalescence, four weeks or later after discharge.

Cell culture, transfection and co-localization assay

Huh7 liver cell line was cultured in DMEM (Life Technologies, Carlsbad, CA, USA) with 10% bovine serum and penicillin-streptomycin and maintained at 37˚C in 5% CO_2 and used for three transfection assays: wild type DENV2-NS5; LxxIxE-deleted DENV2-NS5; and wild type DENV1-NS5. The DENV2-NS5 was fused with an orange fluorescent protein and cloned into pCMV3-C-His vector (DENV-2 (strain New Guinea C–GenBank: AF038403) NS5 open reading frame (ORF) (2700bp) mammalian expression plasmid; SinoBiological, Beijing, China). The wild-type DENV2-NS5 clones were mutagenized with the Q5 Site-Directed Mutagenesis Kit (New England Biolabs, Ipswich, MA, USA), following the manufacturers' protocol. DENV1-NS5 (isolate KDH0026A, Kamphaeng Phet Provincial Hospital, Thailand–GenBank: HG316481) was also fused with OFP and cloned into the same vector using overlapping primers and the HiFi DNA Assembly Protocol (New England Biolabs, Ipswich, MA, USA), according to the manufacturers' recommendations. The designed primers are described on S10 Table, and the LxxIxE-deletion and DENV1-NS5 assembly were confirmed by Sanger sequencing performed in a 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Transient transfections were performed using Lipofectamine 3000 reagent (Invitrogen) according to the manufacturer's protocol. Wildtype and mutated DENV2-NS5 proteins expressions were confirmed in ZOE Fluorescent Cell Imager (Bio-Rad, Hercules, CA, USA). Transfected cells were harvested 24h, 48h and 72h after transfection and fixed with 4% paraformaldehyde. PPP2R5E was tagged with a primary rabbit anti-PPP2R5E antibody (Atlas Antibodies, Bromma, Sweden) and revealed with a secondary goat anti-rabbit Alexa Fluor 488 antibody (Thermo Fisher Scientific, Waltham, MA, USA). Imaging was obtained on a TCS SP5 II (Leica, Wetzlar, Germany) Laser Scanning Confocal microscope. Images (green and red channel) were aligned with Huygens Software (Chromatic Aberration Corrector module), by using 0.2 μm TetraSpeck Microspheres (Thermo Fisher Scientific, Waltham, MA, USA) embedded in the same conditions as the samples. The images were merged and adjusted for brightness and contrast with Fiji software [48].

142

Supporting information

S1 Fig. Global ancestry inferred through RFMix when using three parental ancestries (South, Northeast and Southeast Asian) for the three Thai hospital dengue cohorts. Each vertical line represents an individual, and the three colours represent the proportion of the three parental populations in each genome (light orange for South Asian, dark orange for Southeast Asian and blue for Northeast Asian).
(DOCX)

S2 Fig. LD (D') values for the *MICB* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.
(DOCX)

S3 Fig. LD (D') values for the *PLCB4* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.
(DOCX)

S4 Fig. Locus zoom of the chromosomal region around genes (A–*PLCB4*, B–*PLCE1*, C–*CHST10*, D–*AHRR*, E–*GRIP1*, F–*PPP2R5E*) with significant p-values obtained for DSS and DF tests. The Asian recombination map was used.
(DOCX)

S5 Fig. LD (D') values for the *PLCE1* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.
(DOCX)

S6 Fig. Manhattan plots for the conventional association tests with PCA correction for population structure. A–DSS test (p-values and D' for SNPs surrounding the two spurious SNPs are highlighted). B—DF test. The red line indicates the significance threshold.
(DOCX)

S7 Fig. LD (D') values for the *CHST10* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.
(DOCX)

S8 Fig. LD (D') values for the *AHRR* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.
(DOCX)

S9 Fig. LD (D') values for the *PPP2R5E* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.
(DOCX)

S10 Fig. Gene expression for *CHST10* (A) and *AHRR* (B) in Thai dengue cohort along the course of disease from a transcriptome dataset for whole blood.[17] No significant differences in expression were observed.
(DOCX)

143

S11 Fig. mRNA expression profiles for the eQTLs in *PPP2R5E* and *AHRR* genes (information from GTEx database). The protective alleles are indicated in green while the causative alleles are in red.
(DOCX)

S12 Fig. PCA of the Thai samples. Plot of PC1 versus PC2 in the Thai control and patient cohorts.
(DOCX)

S13 Fig. ADMIXTURE plot for K = 3 for Thai and Vietnamese cohorts and the parental populations used in this work (CDX–Chinese Dai in Xishuangbanna; ITU–Indian Telugu from the UK (ITU); and MAL–Malaysian).
(DOCX)

S1 Table. Significant SNPs in BMIX analysis for Vietnam DSS test. The base position refers to GRCh37 genome assembly.
(DOCX)

S2 Table. Significant SNPs in BMIX analysis for Thai DSS vs control test. The base position refers to GRCh37 genome assembly.
(DOCX)

S3 Table. Significant SNPs in BMIX analysis for Thai DF test. The base position refers to GRCh37 genome assembly.
(DOCX)

S4 Table. Annotation of the significant SNPs in BMIX analysis for Vietnam DSS test, inferred by using the Variant Effect Predictor (VEP) tool from Ensemble.
(DOCX)

S5 Table. Annotation of the significant SNPs in BMIX analysis for Thai DSS test, inferred by using the Variant Effect Predictor (VEP) tool from Ensemble.
(DOCX)

S6 Table. Annotation of the significant SNPs in BMIX analysis for Thai DF test, inferred by using the Variant Effect Predictor (VEP) tool from Ensemble.
(DOCX)

S7 Table. Association p-values in the entire Thai cohort in six and two SNPs selected from the sets of BMIX-associated SNPs with DF and DSS phenotypes, respectively.
(DOCX)

S8 Table. Detailed identification of motif sequences across DENV serotypes.
(DOCX)

S9 Table. Information on DENV serotype and primary/secondary in the Thai cohort.
(DOCX)

S10 Table. Primers used for mutagenesis and DNA assembly protocols.
(DOCX)

S1 Text. Association and BMIX results for the 10 runs of pseudo datasets, permutating case and control labels.
(DOCX)

144

Acknowledgments

Author Contributions

Conceptualization: Prapat Suriyaphol, Luisa Pereira, Anavaj Sakuntabhai.

Data curation: Marisa Oliveira, Worachart Lert-itthiporn, Bruno Cavadas, Verónica Fernandes, Ampaiwan Chuansumrit, Orlando Anunciação, Fanny Koeth, Kanchana Tangnararatchakit, Chiea Chuen Khor, Prida Malasit, Fumihiko Matsuda, Prapat Suriyaphol.

Formal analysis: Marisa Oliveira, Worachart Lert-itthiporn, Bruno Cavadas, Verónica Fernandes, Orlando Anunciação, Isabelle Casademont, Etienne Simon-Lorière, Luisa Pereira.

Funding acquisition: Luisa Pereira, Anavaj Sakuntabhai.

Investigation: Marisa Oliveira, Isabelle Casademont, Fanny Koeth, Marina Penova, Luisa Pereira.

Methodology: Marisa Oliveira, Worachart Lert-itthiporn, Bruno Cavadas, Verónica Fernandes, Orlando Anunciação, Prapat Suriyaphol.

Project administration: Luisa Pereira, Anavaj Sakuntabhai.

Resources: Ampaiwan Chuansumrit, Kanchana Tangnararatchakit, Chiea Chuen Khor, Prida Malasit, Fumihiko Matsuda, Prapat Suriyaphol, Anavaj Sakuntabhai.

Software: Marisa Oliveira, Worachart Lert-itthiporn, Bruno Cavadas, Verónica Fernandes, Orlando Anunciação.

Supervision: Prapat Suriyaphol, Luisa Pereira, Anavaj Sakuntabhai.

Writing – original draft: Marisa Oliveira, Richard Paul, Luisa Pereira, Anavaj Sakuntabhai.

Writing – review & editing: Marisa Oliveira, Worachart Lert-itthiporn, Richard Paul, Luisa Pereira, Anavaj Sakuntabhai.

References

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. Nature. 2013; 496(7446):504–7. Epub 2013/04/09. https://doi.org/10.1038/nature12060 PMID: 23563266; PubMed Central PMCID: PMCPmc3651993.

2. WHO. Dengue haemorrhagic fever: diagnosis, treatment, prevention and control. 2nd ed. Geneva: World Health Organization; 1997. 84 p.

3. Khor CC, Chau TN, Pang J, Davila S, Long HT, Ong RT, et al. Genome-wide association study identifies susceptibility loci for dengue shock syndrome at MICB and PLCE1. Nat Genet. 2011; 43(11):1139–41. Epub 2011/10/18. https://doi.org/10.1038/ng.960 PMID: 22001756; PubMed Central PMCID: PMCPmc3223402.

4. Long HT, Hibberd ML, Hien TT, Dung NM, Van Ngoc T, Farrar J, 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; 199(4):537–46. Epub 2009/01/14. https://doi.org/10.1086/596507 PMID: 19138155; PubMed Central PMCID: PMCPmc4333209.

145

5. Hoang LT, Lynn DJ, Henn M, Birren BW, Lennon NJ, Le PT, et al. The early whole-blood transcriptional signature of dengue virus and features associated with progression to dengue shock syndrome in Vietnamese children and young adults. Journal of virology. 2010; 84(24):12982–94. Epub 2010/10/15. https://doi.org/10.1128/JVI.01224-10 PMID: 20943967; PubMed Central PMCID: PMCPmc3004338.

6. Simmons CP, Popper S, Dolocek C, Chau TN, Griffiths M, Dung NT, 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):1097–107. Epub 2007/03/16. https://doi.org/10.1086/512162 PMID: 17357045; PubMed Central PMCID: PMCPmc4042601.

7. Devignot S, Sapet C, Duong V, Bergon A, Rihet P, Ong S, et al. Genome-wide expression profiling deciphers host responses altered during dengue shock syndrome and reveals the role of innate immunity in severe dengue. PLoS One. 2010; 5(7):e11671. Epub 2010/07/24. https://doi.org/10.1371/journal.pone.0011671 PMID: 20652028; PubMed Central PMCID: PMCPmc2907396.

8. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, et al. Ancient admixture in human history. Genetics. 2012; 192(3):1065–93. Epub 2012/09/11. https://doi.org/10.1534/genetics.112.145037 PMID: 22960212; PubMed Central PMCID: PMCPmc3522152.

9. Sierra B, Triska P, Soares P, Garcia G, Perez AB, Aguirre E, et al. OSBPL10, RXRA and lipid metabolism confer African-ancestry protection against dengue haemorrhagic fever in admixed Cubans. PLoS Pathog. 2017; 13(2):e1006220. https://doi.org/10.1371/journal.ppat.1006220 PMID: 28241052.

10. Ma F, Liu SY, Razani B, Arora N, Li B, Kagechika H, et al. Retinoid X receptor alpha attenuates host antiviral response by suppressing type I interferon. Nat Commun. 2014; 5:5494. Epub 2014/11/25. https://doi.org/10.1038/ncomms6494 PMID: 25417649; PubMed Central PMCID: PMCPmc4380327.

11. de la Sierra B, Kouri G, Guzman MG. Race: a risk factor for dengue hemorrhagic fever. Archives of virology. 2007; 152(3):533–42. Epub 2006/11/16. https://doi.org/10.1007/s00705-006-0869-x PMID: 17106622.

12. Shekhar KC, Huat OL. Epidemiology of dengue/dengue hemorrhagic fever in Malaysia—a retrospective epidemiological study. 1973–1987. Part II: Dengue fever (DF). Asia-Pacific journal of public health / Asia-Pacific Academic Consortium for Public Health. 1992; 6(3):126–33. Epub 1992/01/01. https://doi.org/10.1177/101053959200600302 PMID: 1342799.

13. Shriner D, Adeyemo A, Rotimi CN. Joint ancestry and association testing in admixed individuals. PLoS Comput Biol. 2011; 7(12):e1002325. Epub 2012/01/05. https://doi.org/10.1371/journal.pcbi.1002325 PMID: 22216000; PubMed Central PMCID: PMCPmc3245293.

14. Barbie J. Chinese culture and demographic trends in Thailand and Malaysia. Chinese journal of population science. 1992; 4(2):139–48. Epub 1992/01/01. PMID: 12317919.

15. Galanter JM, Gignoux CR, Torgerson DG, Roth LA, Eng C, Oh SS, et al. Genome-wide association study and admixture mapping identify different asthma-associated loci in Latinos: the Genes-environments & Admixture in Latino Americans study. J Allergy Clin Immunol. 2014; 134(2):295–305. Epub 2014/01/11. https://doi.org/10.1016/j.jaci.2013.08.055 PMID: 24406073; PubMed Central PMCID: PMC4085159.

16. Jeff JM, Armstrong LL, Ritchie MD, Denny JC, Kho AN, Basford MA, et al. Admixture mapping and subsequent fine-mapping suggests a biologically relevant and novel association on chromosome 11 for type 2 diabetes in African Americans. PLoS One. 2014; 9(3):e86931. Epub 2014/03/07. https://doi.org/10.1371/journal.pone.0086931 PMID: 24595071; PubMed Central PMCID: PMC3940426.

17. Kwissa M, Nakaya HI, Onlamoon N, Wrammert J, Villinger F, Perng GC, et al. Dengue virus infection induces expansion of a CD14(+)CD16(+) monocyte population that stimulates plasmablast differentiation. Cell Host Microbe. 2014; 16(1):115–27. Epub 2014/07/02. https://doi.org/10.1016/j.chom.2014.06.001 PMID: 24981333; PubMed Central PMCID: PMC4116428.

18. Hertz EP, Kruse T, Davey NE, Lopez-Mendez B, Sigurethsson JO, Montoya G, et al. A Conserved Motif Provides Binding Specificity to the PP2A-B56 Phosphatase. Molecular cell. 2016; 63(4):686–95. Epub 2016/07/28. https://doi.org/10.1016/j.molcel.2016.06.024 PMID: 27453045.

19. Cocco L, Follo MY, Manzoli L, Suh PG. Phosphoinositide-specific phospholipase C in health and disease. Journal of lipid research. 2015; 56(10):1853–60. Epub 2015/03/31. https://doi.org/10.1194/jlr.R057984 PMID: 25821234; PubMed Central PMCID: PMCPmc4583093.

20. Kawakami T, Xiao W. Phospholipase C-beta in immune cells. Advances in biological regulation. 2013; 53(3):249–57. Epub 2013/08/29. https://doi.org/10.1016/j.jbior.2013.08.001 PMID: 23981313; PubMed Central PMCID: PMCPmc4324595.

21. Lin YJ, Chang JS, Liu X, Tsang H, Chien WK, Chen JH, et al. Genetic variants in PLCB4/PLCB1 as susceptibility loci for coronary artery aneurysm formation in Kawasaki disease in Han Chinese in Taiwan. Sci Rep. 2015; 5:14762. Epub 2015/10/06. https://doi.org/10.1038/srep14762 PMID: 26434682; PubMed Central PMCID: PMCPmc4593004.

146

22. Maglich JM, Stoltz CM, Goodwin B, Hawkins-Brown D, Moore JT, Kliewer SA. Nuclear pregnane x receptor and constitutive androstane receptor regulate overlapping but distinct sets of genes involved in xenobiotic detoxification. Molecular pharmacology. 2002; 62(3):638–46. Epub 2002/08/16. PMID: 12181440.

23. Yoshinari K, Kobayashi K, Moore R, Kawamoto T, Negishi M. Identification of the nuclear receptor CAR:HSP90 complex in mouse liver and recruitment of protein phosphatase 2A in response to pheno-barbital. FEBS letters. 2003; 548(1–3):17–20. Epub 2003/07/30. PMID: 12885400.

24. Georgopoulou U, Tsitoura P, Kalamvoki M, Mavromara P. The protein phosphatase 2A represents a novel cellular target for hepatitis C virus NS5A protein. Biochimie. 2006; 88(6):651–62. Epub 2006/02/08. https://doi.org/10.1016/j.biochi.2005.12.003 PMID: 16460864.

25. Maertens GN. B'-protein phosphatase 2A is a functional binding partner of delta-retroviral integrase. Nucleic Acids Res. 2016; 44(1):364–76. Epub 2015/12/15. https://doi.org/10.1093/nar/gkv1347 PMID: 26657642; PubMed Central PMCID: PMCPmc4705670.

26. Kwun HJ, Shuda M, Camacho CJ, Gamper AM, Thant M, Chang Y, et al. Restricted protein phospha-tase 2A targeting by Merkel cell polyomavirus small T antigen. 2015; 89(8):4191–200. https://doi.org/10.1128/JVI.00157-15 PMID: 25631078.

27. Hunt TA, Urbanowski MD, Kakani K, Law LM, Brinton MA, Hobman TC. Interactions between the West Nile virus capsid protein and the host cell-encoded phosphatase inhibitor, I2PP2A. Cellular microbiol-ogy. 2007; 9(11):2756–66. Epub 2007/09/18. https://doi.org/10.1111/j.1462-5822.2007.01046.x PMID: 17868381.

28. Hannemann H, Sung PY, Chiu HC, Yousuf A, Bird J, Lim SP, et al. Serotype-specific differences in den-gue virus non-structural protein 5 nuclear localization. J Biol Chem. 2013; 288(31):22621–35. Epub 2013/06/19. https://doi.org/10.1074/jbc.M113.481382 PMID: 23770669; PubMed Central PMCID: PMCPmc3829348.

29. Apostolidis SA, Rodriguez-Rodriguez N, Suarez-Fueyo A, Dioufa N, Ozcan E, Crispin JC, et al. Phos-phatase PP2A is requisite for the function of regulatory T cells. Nature immunology. 2016; 17(5):556–64. Epub 2016/03/15. https://doi.org/10.1038/ni.3390 PMID: 26974206; PubMed Central PMCID: PMCPmc4837024.

30. Shanker V, Trincucci G, Heim HM, Duong HT. Protein phosphatase 2A impairs IFNalpha-induced anti-viral activity against the hepatitis C virus through the inhibition of STAT1 tyrosine phosphorylation. Jour-nal of viral hepatitis. 2013; 20(9):612–21. Epub 2013/08/06. https://doi.org/10.1111/jvh.12083 PMID: 23910645.

31. Luhn K, Simmons CP, Moran E, Dung NT, Chau TN, Quyen NT, et al. Increased frequencies of CD4+ CD25(high) regulatory T cells in acute dengue infection. The Journal of experimental medicine. 2007; 204(5):979–85. Epub 2007/04/25. https://doi.org/10.1084/jem.20061381 PMID: 17452519; PubMed Central PMCID: PMCPmc2118571.

32. Munoz-Jordan JL, Sanchez-Burgos GG, Laurent-Rolle M, Garcia-Sastre A. Inhibition of interferon sig-naling by dengue virus. Proc Natl Acad Sci U S A. 2003; 100(24):14333–8. Epub 2003/11/13. https://doi.org/10.1073/pnas.2335168100 PMID: 14612562; PubMed Central PMCID: PMCPmc283592.

33. Ho LJ, Hung LF, Weng CY, Wu WL, Chou P, Lin YL, et al. Dengue virus type 2 antagonizes IFN-alpha but not IFN-gamma antiviral effect via down-regulating Tyk2-STAT signaling in the human dendritic cell. J Immunol. 2005; 174(12):8163–72. Epub 2005/06/10. PMID: 15944325.

34. Shresta S, Sharar KL, Prigozhin DM, Snider HM, Beatty PR, Harris E. 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):3946–54. Epub 2005/09/09. PMID: 16148142.

35. Kreitinger JM, Beamer CA. Environmental Immunology: Lessons Learned from Exposure to a Select Panel of Immunotoxicants. 2016; 196(8):3217–25. https://doi.org/10.4049/jimmunol.1502149 PMID: 27044635.

36. Murray NE, Quam MB, Wilder-Smith A. Epidemiology of dengue: past, present and future prospects. Clinical epidemiology. 2013; 5:299–309. Epub 2013/08/31. https://doi.org/10.2147/CLEP.S34440 PMID: 23990732; PubMed Central PMCID: PMCPmc3753061.

37. Chacon-Duque JC, Adhikari K, Avendano E, Campo O, Ramirez R, Rojas W, et al. African genetic ancestry is associated with a protective effect on Dengue severity in colombian populations. Infect Genet Evol. 2014; 27:89–95. Epub 2014/07/16. https://doi.org/10.1016/j.meegid.2014.07.003 PMID: 25017656.

38. Jaenisch T, Junghanss T, Wills B, Brady OJ, Eckerle I, Farlow A, et al. Dengue expansion in Africa-not recognized or not happening? Emerg Infect Dis. 2014; 20(10). Epub 2014/10/02. https://doi.org/10.3201/eid2010.140487 PMID: 25271370; PubMed Central PMCID: PMC4193177.

39. Jaenisch T, Sakuntabhai A, Wilder-Smith A. Dengue research funded by the European Commission-scientific strategies of three European dengue research consortia. PLoS Negl Trop Dis. 2013; 7(12):

147

e2320. Epub 2013/12/19. https://doi.org/10.1371/journal.pntd.0002320 PMID: 24349584; PubMed Central PMCID: PMCPmc3861113.

40. Lourenco J, Recker M. The 2012 Madeira dengue outbreak: epidemiological determinants and future epidemic potential. PLoS Negl Trop Dis. 2014; 8(8):e3083. Epub 2014/08/22. https://doi.org/10.1371/journal.pntd.0003083 PMID: 25144749; PubMed Central PMCID: PMCPmc4140668.

41. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3):559–75. Epub 2007/08/19. https://doi.org/10.1086/519795 PMID: 17701901; PubMed Central PMCID: PMC1950838.

42. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods. 2012; 9(2):179–81. Epub 2011/12/06. https://doi.org/10.1038/nmeth.1785 PMID: 22138821.

43. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. Am J Hum Genet. 2013; 93(2):278–88. Epub 2013/08/06. https://doi.org/10.1016/j.ajhg.2013.06.020 PMID: 23910464; PubMed Central PMCID: PMC3738819.

44. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491(7422):56–65. Epub 2012/11/07. https://doi.org/10.1038/nature11632 PMID: 23128226; PubMed Central PMCID: PMC3498066.

45. Wong LP, Ong RT, Poh WT, Liu X, Chen P, Li R, et al. Deep whole-genome sequencing of 100 south-east Asian Malays. Am J Hum Genet. 2013; 92(1):52–66. Epub 2013/01/08. https://doi.org/10.1016/j.ajhg.2012.12.005 PMID: 23290073; PubMed Central PMCID: PMCPmc3542459.

46. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009; 19(9):1655–64. Epub 2009/08/04. https://doi.org/10.1101/gr.094052.109 PMID: 19648217; PubMed Central PMCID: PMCPmc2752134.

47. Marden JR, Walter S, Tchetgen Tchetgen EJ, Kawachi I, Glymour MM. Validation of a polygenic risk score for dementia in black and white individuals. Brain and behavior. 2014; 4(5):687–97. Epub 2014/10/21. https://doi.org/10.1002/brb3.248 PMID: 25328845; PubMed Central PMCID: PMCPmc4107377.

48. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012; 9(7):676–82. Epub 2012/06/30. https://doi.org/10.1038/nmeth.2019 PMID: 22743772; PubMed Central PMCID: PMCPmc3855844.

148

Functional studies to assess the influence of the newly discovered candidate genes in DENV infection

Paper III – Functional evaluation of GRIP1 and PPP2R5E involvement in dengue fever

Paper III – Functional evaluation of GRIP1 and PPP2R5E involvement in dengue fever

Functional evaluation of *GRIP1* and *PPP2R5E* involvement in dengue fever

Marisa Oliveira[1,2,3,4], Marina Penova[4,5,6], Matthieu Prot[4,5], Isabelle Casademont[4,5], Yves Jacob[4,7], Luisa Pereira[1,2,8,9], Etienne Simon-Loriere[4,5,7,9], Anavaj Sakuntabhai[4,5,7,9]

[1] i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135 Porto, Portugal

[2] Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), 4200-135 Porto, Portugal

[3] Instituto de Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto, 4050-313 Porto, Portugal

[4] Institut Pasteur, Functional Genetics of Infectious Diseases Unit, 75724 Paris Cedex 15, France

[5] Pasteur Kyoto International Joint Research Unit for Integrative Vaccinomics, Kyoto, Japan

[6] Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

[7] CNRS URA3012, Paris 75015, France

[8] Faculdade de Medicina da Universidade do Porto, Portugal

[9] These authors jointly directed this work.

Abstract

Genome-wide association studies (GWAS) have been extensively used in complex diseases to identify phenotype-associated genes. However, the functional validation of candidate genes and/or genes variants is mostly ignored due to lack of appropriate *in vitro* and *in vivo* models. Recently, a dengue fever ancestry-informed GWAS applied to Thai case-control cohorts identified two new candidate genes, *GRIP1* and *PPP2R5E*, found in association with classical dengue fever. Here we established stable cell lines with gene knockout by CRISPR-Cas9 and constitutive gene overexpression using a mammalian expression vector to evaluate the effects of these genes on virus replication and infection capacity. The infection assays using DENV1 showed that both *GRIP1* and *PPP2R5E* overexpression (OE) exhibit antiviral effect against DENV1, with a stronger impact 24h after infection: less than 5% (p-value<0.0001) and 40% (p-value<0.0001) of infection at MOI 0.1 and 0.5, respectively, while the wild type cell line presented 15% and 75% of infection for the same MOIs. At 24h, the same effect was observed at virus titer level, with a significant reduction for *GRIP1* overexpression (p-value<0.0001; MOI 0.5: p-value=0.0002), and an even more accentuated decrease for *PPP2R5E* overexpression (both MOIs: p-value<0.0001). No evident signs were observed in the knockout cell lines, possibly due to low GRIP1 and PPP2R5E protein level in the cell line used. *PPP2R5E* has been previous reported to interact with DENV-NS5 and be targeted for degradation by HIV-1, so it might exert an antiviral action by direct interaction with DENV viral proteins, whereas *GRIP1* seems to modulate IRF3 and, consequently, the innate immunity activation. Our established stable cell lines are efficient *in vitro* models to continue investigating the molecular mechanisms by which these genes play a role in DENV infection.

Introduction

Dengue is the most common arthropod-borne disease worldwide (Bhatt et al., 2013). It is well spread through the tropics and surroundings, occurring in endemic-epidemic cycles (Vasilakis and Weaver, 2008). Its burden increased exponentially in the last decades and 390 million dengue infections are estimated to occur per year, with only 96 million presenting clinical manifestations (Bhatt et al., 2013). Dengue virus comprises four serotypes (DENV 1-4) and is transmitted by *Aedes* species, especially by *Ae. aegypti* and, less efficiently, by *Ae. Albopictus* (Guzman et al., 2010). Once an individual is infected with one of the four serotypes, the manifestations can go from classical dengue fever (DF), with mild symptoms, to life-threatening severe dengue (Guzman et al., 2016). Despite the efforts, there is no preventive treatment available and Dengvaxia, the only vaccine that passed all the trials, was recently recommend to be only used in seropositive individuals by World Health Organization (WHO) (WHO, 2017).

Genetic susceptibility has been indicated as one of the main factors behind the observed differences among ethnic groups regarding the disease outcomes and severity (Ang et al., 2015; Brown et al., 2004; Shekhar and Huat, 1992; Sierra et al., 2007). The inclusion of ancestry information in recent genome-wide association studies (GWAS) led to the discovery of new candidate genes (Oliveira et al., 2018; Sierra et al., 2017), which conferred genetic risk matches quite well the high susceptibility of Southeast and Northeast Asian populations and the lower susceptibility associated with higher African input in Latin Americans. Southeast Asia also presented a high severe dengue risk applying a careful population genetic-based meta-analyses of traditional dengue-associated genetic markers, such as *TNFA*, *CD209*, and *CD32* (Oliveira et al. submitted). In fact, in epidemiological terms, 70% of total dengue infections in 2010 occurred in the Asian continent, mostly in the Western Pacific and Southeast Asian regions, where is detected the presence of several serotypes, with also an estimate of around 200 million of unapparent infections (Bhatt et al., 2013). This high incidence per year leads to a high economical cost, including hospital admissions, patients follow-up and treatments, urging investigation aiming to understand the molecular mechanisms underlying the disease. Transcriptomic studies have been conducted in dengue fever cohorts, to unravel pathways involved in the disease, commonly evidencing several immune system pathways altered (Kwissa et al., 2014; Simon-Loriere et al., 2017b). The functional follow-up of GWAS and transcriptome studies have the potential to reveal the molecular mechanisms

working as natural defenses against dengue fever, besides cleaning up the literature from spurious association signals.

The first GWAS published for dengue infection analyzed children with dengue shock syndrome (DSS) against cord blood donors from Vietnam and found two new candidate genes, *MICB*, and *PLCE1*, for dengue severity yet, no functional evidence was reported (Khor et al., 2011). Our recent ancestry-based genome-wide search on the Thai population (Oliveira et al., 2018) indicated a significant association of the related *PLCB4* gene with dengue shock syndrome, adding support to phospholipase C members role with this outcome due to potential involvement in the inflammation of blood vessels (Lin et al., 2015). The phenotype of knockout mice for these genes can give valuable information. Mice with *PLCE1* knockout exhibited lower pro-inflammatory mediators induced by LPS and microvascular leakage, whereas siRNA-mediated PLCE1 depletion seemed to modulate endothelial cells inflammation response by inhibiting NF-kB feedback to pro-inflammatory action (Bijli et al., 2016). *PLCB4* was detected in the blood vessels and heart atrium of mice embryos (Visel, 2004), reinforcing the participation of both genes detected in association with dengue shock syndrome in processes involving immune response and blood vessels phenotype.

Our study in Thai dengue cohorts (Oliveira et al., 2018) identified, from the comparison between dengue fever and healthy controls, four statistically associated genes with participation in the xenobiotics metabolism pathway – *AHRR*, *CHST10*, *GRIP1* and *PPP2R5E*. Potential expression quantitative trait *loci* (eQTLs) were searched in the European database GTEx (Lonsdale et al., 2013) where all four candidate protective *PPP2R5E* alleles significantly reduce the expression of the gene, while the candidate *GRIP1* allele is not an eQTL, serving as an indication since the extrapolation to Asian populations cannot be direct as some ethnicity differences in eQTLs were observed (Quach et al., 2016). In addition, *PPP2R5E* and *GRIP1* expression changed significantly among the different dengue phenotypes in a distinct Thai cohort (Kwissa et al., 2014), increasing in patients when comparing to controls and convalescents for the first, and decreasing for the second. In our paper (Oliveira et al., 2018), we also provided *in silico* and immune-localization evidence that there is interaction between DENV non-structural protein 5 (NS5) and PPP2R5E, with variability between NS5 proteins from serotypes 1 and 2 in relation to entrance into the host nucleus. This potential role of PPP2R5E protein in binding to viral proteins, and co-localization within the host cell, should be further explored in the context of protection of this gene against dengue fever.

Other works indicated the possible involvement of these genes in infectious diseases. *PPP2R5E* belongs to the phosphatase 2A regulatory subunit B family from the PP2A holoenzymes family. PP2A is a heterotrimeric serine/threonine phosphatase with a catalytic C subunit, a scaffold A subunit, and a changeable regulatory B subunit. The interaction of PP2A holoenzyme with viral proteins has been described in the literature: it is targeted by Merkel cell polyomavirus (MCV) small T protein (sT) to allow optimal replication (Kwun et al., 2009); West Nile Virus (WNV) capsid protein interacts with SET/I2PP2A, resulting in the increment of PP2A activity (Hunt et al., 2007); and Hepatitis C virus NS5A protein increased PP2A activity and interacted with subunits A and C (Georgopoulou et al., 2006). Later evidences reported that PP2A–B56 regulatory subunit recognizes a conversed motif (LxxIxE) (Hertz et al., 2016) in the interacting proteins including the Ebola virus nucleoprotein (Kruse et al., 2017). PP2A activity was tested upon dengue infection with DENV2 and DENV4 and, its activity did not suffer alteration with infection, yet PP2A inhibition with okadaic acid had an antiviral effect on both serotypes (Soto–Acosta et al., 2017).

GRIP1 (Glutamate Receptor Interacting Protein 1) is a scaffolding protein and participates in the traffic of transmembrane proteins such as Constitutive Androstane Receptor (CAR) (Min et al., 2002). Besides, Liver X Receptor (LXR), another nuclear receptor from the same family as CAR, competes with Interferon Regulatory Factor (IRF3) to bind GRIP1 modulating the inflammatory response by LPS (Miao et al., 2016). GRIP1 interaction with IRF3 has been documented before, also in a competitive manner with glucocorticoid receptors (GR) and *GRIP1* overexpression rescued IRF3–dependent gene expression in macrophages (Reily et al., 2006). Furthermore, *GRIP1* seems to regulate the inflammatory genes during the transcription cycle by modulating and complexing with GR, supporting the potential role of *GRIP1* in innate immunity (Gupte et al., 2013).

Given these evidences we decided in this work to perform classical infection functional assays to investigate the impact of *PPP2R5E* and *GRIP1* genes in dengue disease. For that we established stable cell lines to overexpress (OE) and knockout (KO) *PPP2R5E* and *GRIP1* genes, that could be used as *in vitro* models to conduct the infection assays using DENV1 serotype and investigate their potential role as human host proteins in the development of infection.

Material and Methods

Cell Culture

HEK 293 cells were maintained in DMEM–Dulbecco's Modified Eagle Medium (Thermo Fisher Scientific, Waltham, MA, USA) with 10% fetal bovine serum (FBS) and penicillin–streptomycin (P/S) and maintained at 37°C in 5% CO_2.

CRISPR/Cas9-Mediated PPP2R5E e GRIP1 Knockout

Sequences corresponding to the fourth exon of *PPP2R5E* and the third exon of *GRIP1* (Table C.1) were cloned into the BbsI (New England Biolabs Inc., Ipswich, MA, USA) site of the pSpCas9(BB)–2A–Puro plasmid. pSpCas9(BB)–2A–Puro (PX459) V2.0 was a gift from Feng Zhang(Ran et al., 2013) (Addgene plasmid # 62988). The guide RNAs correct insertion were confirmed by Sanger Sequencing. Cells were transfected with 500 ng plasmid DNA using Lipofectamine LTX (Thermo Fisher Scientific, Waltham, MA, USA), and 48 h later were treated with 1 mg/mL puromycin (Thermo Fisher Scientific, Waltham, MA, USA) until all non–transfected cells, used as control, had died (~72 h). The double strand break and nonhomologous end joining were confirmed by Sanger sequencing (Table C.2; Figure C.3).

ORF Cloning

Primers were designed to include NheI and XhoI restriction enzyme sites and targeting *GRIP1* and *PPP2R5E* (Table C.1). The fragments were PCR–amplified from a donor vector and cloned into the pcDNA3.1 plasmid (Thermo Fisher Scientific, Waltham, MA, USA), previously digested with NheI and XhoI. The construct was confirmed by Sanger Sequencing prior to transfection. Cells were transfected with 500 ng of linearized plasmid using Lipofectamine LTX (Thermo Fisher Scientific, Waltham, MA, USA) and selected with 0.5 mg/mL Geneticin (G418) (Thermo Fisher Scientific, Waltham, MA, USA) 48 h later. Prior transfection, the Geneticin concentration was optimized considering cells viability and morphology. Cells were maintained under antibiotic selection and were assayed for protein expression via western blot.

Western Blots

Surviving cells were cultured and assayed for PPP2R5E and GRIP1 protein levels via western blot. Samples diluted in Laemmli buffer were run on polyacrylamide gels and transferred using the iBlot 2 Dry Blotting System (Thermo Fisher Scientific, Waltham, MA, USA). Membranes were probed with primary antibody for PPP2R5E (1:1000, Abcam, Cambridge, UK), GRIP1 (1:1000, Santa Cruz Biotechnology, Dallas, TX, USA (H-4)) and beta-Actin loading control (1:1000, Thermo Fisher Scientific, Waltham, MA, USA) at 4°C overnight and were incubated with the corresponding secondary antibodies, Goat anti-Rabbit IgG (H+L), (Thermo Fisher Scientific, Waltham, MA, USA) and Pierce Goat Anti.Mouse IgG (H+L), (Thermo Fisher Scientific, Waltham, MA, USA), both peroxidase conjugated. Membranes were treated with SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific, Waltham, MA, USA), exposed to Amersham hyperfilm (GE Healthcare, Little Chalfont, UK), and developed.

RNA Purification and cDNA synthesis

Total RNA was extracted using the RNeasy Mini kit (Qiagen, Hilden, Germany). The RNA concentration was measured on NanoDrop. and used for cDNA synthesis using the Maxima H Minus reverse transcriptase (Life Technologies, Carlsbad, CA, USA), with 500 ng RNA and random hexamer primers, following manufacturer's instructions.

DENV infection and Virus Titration

DENV-1 strain (isolate KDH0026A, Kamphaeng Phet Provincial Hospital, Thailand; GenBank: HG316481) was propagated in the mosquito C6/36 cell line grown in Leibovitz's L-15 Medium (Thermo Fisher Scientific, Waltham, MA, USA) containing 2% fetal bovine serum (FBS) and complemented with 1% Non-Essential Amino Acids (NEAA) and 10% Tryptose Phosphate Broth solution. HEK 293 and stably transformed cell lines were used in the infection assays. Dengue viruses were diluted in serum-free DMEM at a multiplicity of infection (MOI) of 0.1 and 0.5. Viral inoculum was mixed in the cells suspension. The virus was neutralized 2 h later by adding the same volume of DMEM complemented with 10% FBS and P/S, resulting in a final concentration of 5% FBS in the supernatant, and kept at 37°C in 5% CO_2. Infected cells and supernatants were harvested at 24 h, 48 h and 72 h post infection.

Ten-fold dilutions of cell supernatants were prepared in serum-free DMEM and used to inoculate confluent monolayers of Vero-E6 cells for 2 h at 37°C. Cells were then overlaid with DMEM with 10% FBS and P/S: Carboxymethyl cellulose (CMC), 1/1 ratio,

performing a final concentration of 5% FBS and 16g/L of CMC. DENV samples were incubated for 72 h. All cells were fixed with 4% Paraformaldehyde for 20 min.

Infected cells collected at 24 h, 48 h and 72 h post infection, were fixed and permabilized with BD Cytofix/Cytoperm and washed with Perm/Wash solutions (BD Biosciences, San Jose, CA), to enable intracellular staining. Cells were stained with the Mouse anti-Flavivirus E protein (4G2) primary antibody, following by labeling with a Goat anti-mouse Alexa Fluor 488 antibody (Thermo Fisher Scientific, Waltham, MA, USA) to detect infected cells by flow cytometry (MACSQuant Analyzer 10, Miltenyi Biotec, Bergisch Gladbach, Germany). Samples were analyzed using Kaluza Analysis 1.5a software (Beckman Coulter, Brea, CA, USA) to determine the percentage of infected cells. Wild type HEK 293 were used as control in all comparisons.

Gene expression analysis

The RNA from infected HEK 293 wild type and all stable transfected cell was collected and purified 24 h and 48 h after infection. The level of gene expression was measured by reverse transcription-quantitative PCR (RT-qPCR). Briefly, total cellular RNA was isolated and assayed using TaqMan primer/probe sets for *GRIP1* (Hs00901672_m1) and *PPP2R5E* (Hs00952135_m1), using Human *GUSB* (Beta Glucuronidase) Endogenous Control (VIC™/MGB probe, primer limited) as an internal control (Applied Biosystems, Foster City, CA, USA). RT-qPCR amplification and data analysis were performed on a QuantStudio™ 12K Flex (Life Technologies, Carlsbad, CA, USA). The results were analyzed in TaqMan® Genotyper software (Life Technologies, Carlsbad, CA, USA). Relative gene expression was calculated using the DCT method, comparing the gene to *GUSB*.

Statistical Analysis

Prism 7.0 (GraphPad) was used to generate graphs and perform statistical analysis. For virus titer analyses, one-tailed unpaired t-test was used to compare groups and the null hypothesis was rejected when p-value<0.05.

Results

Establishment of stable transfected cell lines

RNA expressions of *GRIP1* and *PPP2R5E* genes were checked in The Human Protein Atlas Database to choose a suitable cell line to proceed with gene editing and constitutive expression of both genes of interest (Figure C.1). Human embryonic kidney (HEK 293) cell line was used as a model due to the following characteristics: (1) reasonable gene expression levels (6.8 and 10.8 transcripts per million (TPM) for *GRIP1* and *PPP2R5E*, respectively); (2) permissive handling to construct stably edited cell lines using CRISPR–Cas9 for gene knockout (Li et al., 2017; Zhang et al., 2017); (3) constitutive expression using pcDNA3.1 system(Cortes–Gonzalez et al., 2010; Esmail et al., 2016); (4) and desirable infection rates for dengue infection (Rattanaburee et al., 2015).

Wild type HEK 293 cells were independently transfected with the four constructed plasmids, in order to establish stable cell lines with OE or KO of *PPP2R5E* or *GRIP1* genes. The protein level in the transfected cell lines were accessed via Western Blot assay (Figure 1A, B), and mRNA expression was quantified through RT–qPCR (Figure 1C). For *GRIP1*–OE, protein quantification indicated a fold change of 27, while the mean increase in RNA level was lower (19.88). The same pattern was observed for *PPP2R5E*–OE, with 5.99 and 2.39 fold changes in protein and mRNA levels, respectively. No expression is observed at protein level for CRISPR–Cas9 mediated *PPP2R5E*–KO and *GRIP1*–KO and, the RNA fold changes were considerably low (0.25 and 0.48, respectively).

Cell proliferation rate was investigated (Figure C.2) at 24 h, 48 h and 72 h to check if the established cell lines suffered any perturbation due the introduced gene expression changes. The cell counts at 24 h after plating were used as reference (equal to 1) for comparison with the other time–points. At 48 h, all cell lines presented around the double of initial cells, dividing again at 72 h.
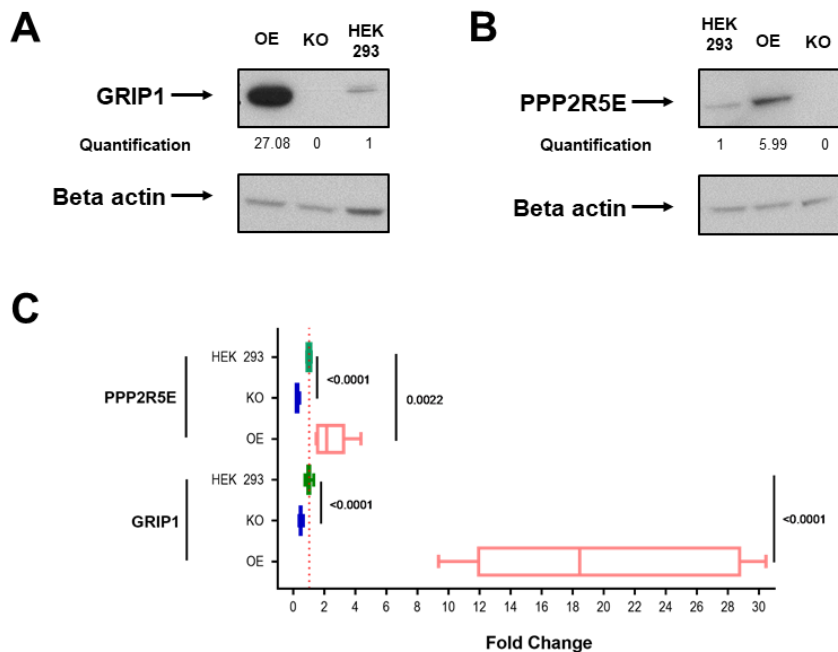
Figure 1. Protein and mRNA expression levels in the established stable cell lines: A and B - Western blots for GRIP1 and PPP2R5E proteins in the wild type HEK 293 and respective OE and KO cell lines; C – RT-PCR results for *GRIP1* and *PPP2R5E* mRNA expressions in the wild type HEK 293 and respective OE and KO cell lines.

Infection of wild type HEK 293 and PPP2R5E and GRIP1 expression profiles

Wild type HEK 293 were infected with DENV1 at MOI 0.5 to access the viral titer at 24 h and 48 h post infection and to investigate potential changes in mRNA expression of *PPP2R5E* and *GRIP1* genes. HEK 293 reached desirable high DENV1 titers ($>1*10^6$ ffu/mL) and a significant increase in virus replication (p–value=0.0011) was observed between 24 h to 48 h after infection (Figure 2A). The mRNA levels of both genes were also accessed during the course of infection (Figure 2B, C). At 24 h, a slight expression decrease was observed in both genes relatively to basal expression values (GRIP: p–value=0.5990; PPP2R5E: p–value=0.0601) and, at 48 h, the tendency inverted, with an increase of *GRIP1* expression (p–value=0.1316), while *PPP2R5E* expression (p–value=0.8509) returned to normal values. In none of the cases, statistical significance was reached.
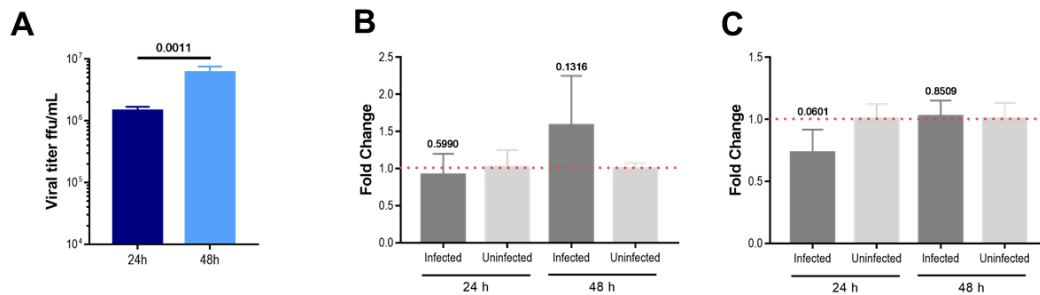
Figure 2. Efficiency of infection of wild type HEK 293 cell line. A – Viral titer (in ffu/mL) at MOI 0.5, 24 h and 48 h after infection. B – *GRIP1* and C – *PPP2R5E* mRNA expression profiles in infected wild type HEK 293 cell line, at MOI 0.5, 24h and 48h after infection versus non-infected line.

DENV1 infection rates vary in the established cell lines

To evaluate the antiviral effect possibly associated with the expression alteration of GRIP1 or PPP2R5E in our established cell lines, we performed DENV1 infections testing two MOIs, 0.1 and 0.5, and measured the percentage of infection obtained at 24 h and 48 h after infection.

As expected, the MOI was determinant in the infection success at 24 h: cells infected with MOI of 0.1 (Figure 3A) reached a maximum of 15% of infection, while MOI 0.5 (Figure 3B) infection rate was 80%. The highest infection rates were observed for wild type control and KO modified cell lines, with *GRIP1*-KO reaching a borderline significant value (p-value= 0.02436). The OE cell lines presented the lowest infection rates, less than 5% (a 2/3 decrease) for MOI 0.1 (p-value<0.0001), and between 35% and 45% (around 1/2 decrease) for MOI 0.5 (p-value<0.0001).

At 48 h, the difference in infection efficiency was considerable for MOI 0.1 (Figure 3A), keeping the tendency observed at 24 h: wild type and KO presented a total of 90% infection, contrasting with the OE cell lines that reached statistically significant lower infection values (60% for *GRIP1* and 50% for *PPP2R5E*, p-value<0.0001). For MOI 0.5 at 48 h (Figure 3B), all cell lines had infection above 90%, and here the increases in infection for KO were statistically significant in both cell lines, while the decreases in infection for OE were not.

At 72 h, the infections rates decreased for both MOIs possibly due to cell death caused by virus saturation (data not shown). We ended up not pursuing this time-point.
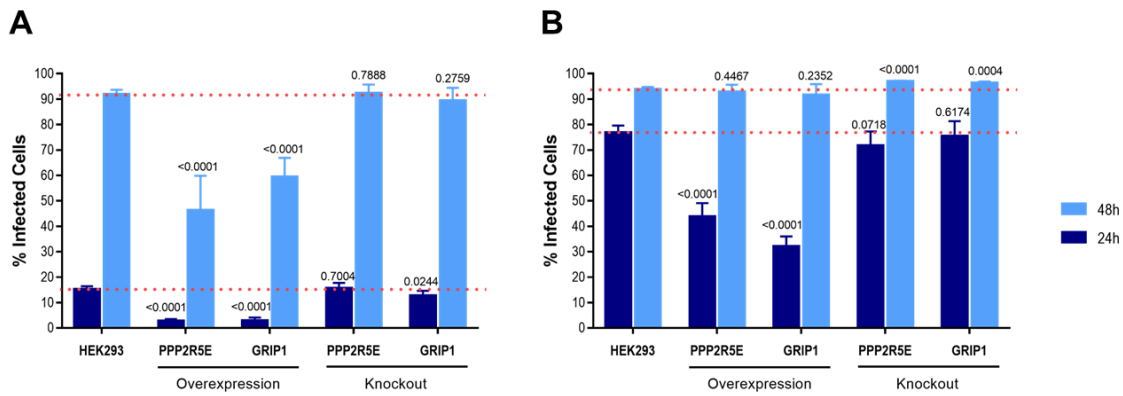
Figure 3. Infection rates (in % of infected cells) in *GRIP1* and *PPP2R5E* OE and KO cell lines in comparison with wild type HEK 293, for MOI 0.1 (A) and MOI 0.5 (B) at 24 h and 48 h after infection.

Virus replication is influenced by these genes

At 24 h and for both MOIs, the pattern mirrored the infection rate results (Figure 4A, B). Wild type HEK 293 and KO cell lines had the highest titers ($1*10^4$ ffu/mL for MOI 0.1 and near $1*10^6$ ffu/mL for MOI 0.5). *PPP2R5E*–OE and *GRIP1*–OE cell lines had the most significant antiviral effects, being approximately one log lower than the control (for both: p–value<0.0001). The differences found at MOI 0.5 were in the same range and, again, *PPP2R5E*–OE (p–value<0.0001) and *GRIP1*–OE (p–value=0.0002) lines presented the strongest antiviral effect.

The titers of the samples collected 48 h after infection still presented significant differences for both MOIs in contrast with the infection rate. The results for MOI of 0.1 (Figure 4A) mirrored the 24 h experiment with *PPP2R5E*–OE (p–value<0.0001) and *GRIP1*–OE (p–value= 0.0003) cell lines, presenting statistically significant decreases in DENV1 antiviral activity, while the KO cell lines did not display significant antiviral activity. With MOI of 0.5, a slightly higher viral concentration for *PPP2R5E*–KO cell line was observed but it was not statistically significant (p–value= 0.6079). The stable *GRIP1*–OE (p–value= 0.4614) and *PPP2R5E*–OE (p–value= 0.0581) cell lines kept the lowest viral titer, yet values did not reach the significance level as was observed in the corresponding infection rate results (Figure 4B).
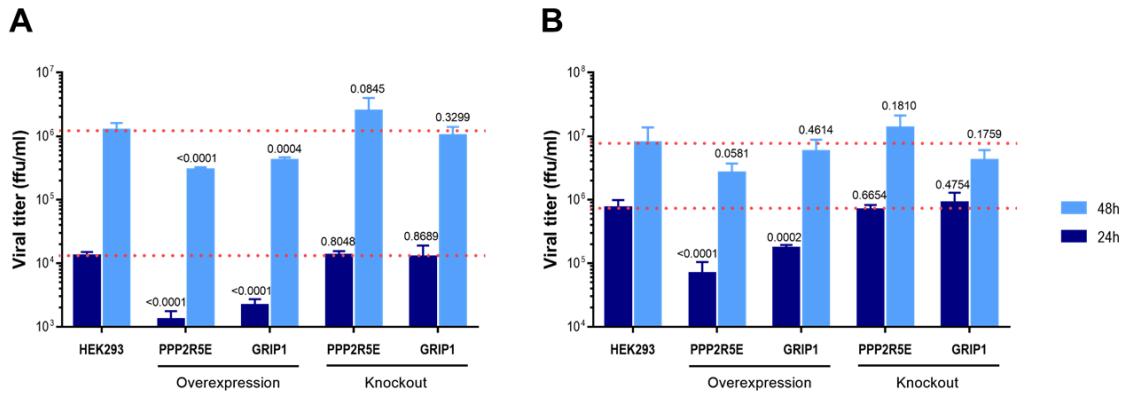
Figure 4. Virus replication (viral titer in ffu/mL) in *GRIP1* and *PPP2R5E* OE and KO cell lines in comparison with wild type HEK 293, for MOI 0.1 (A) and MOI 0.5 (B) at 24 h and 48 h after infection.

Discussion and conclusions

In this study, we took in consideration our newly discovered candidate genes for dengue fever susceptibility (Oliveira et al., 2018) and explored the genes *PPP2R5E* and *GRIP1*. To investigate their effect upon dengue infection, we successfully established stable cell lines, using HEK 293, to overexpress and knockout the selected genes and performed infection assays with DENV1.

HEK 293 is one of the most interesting cell lines to study the function of recombinant proteins, shRNA, CRISPR–Cas9, and other similar technologies. They have been used for the transient or stable overexpression systems of recombinant proteins, using simple methods as polyethyleneimine and calcium phosphate for plasmid delivery into mammalian cells. The efficiency of overexpression is highly dependent of the promoter present in the expression plasmid. The cytomegalovirus (CMV) promoter is a stronger promoter often used to achieve high protein yields when tolerated by the cell line (Khan, 2013). The expression plasmid chosen in this work, pcDNA3.1, possesses a CMV promoter and was successfully integrated and led to elevated yields of PPP2R5E and especially GRIP1 proteins. The observed difference in expression efficiency between both proteins may be related with their cell localization. In fact, the yields of cytosolic recombinant proteins are usually considerably lower and hard to obtain (Büssow, 2015) and, PPP2R5E is uniformly classified as a cytosolic protein. GRIP1 is classified as a non–cytosolic protein in The Human Protein Atlas and UniProt databases, being localized in the plasma membrane, endoplasmic reticulum and cytoplasmic vesicles. The CRISPR–Cas9 system was applied for the first time to *GRIP1* and *PPP2R5E* genes and ablated completely these proteins from HEK 293, albeit they present vestigial RNA expression. This CRISPR–Cas9 protocol was more efficient to completely eliminate the proteins, when compared with previous tests on GRIP1–KO by using siRNA or shRNA methods (Kato et al., 2010),and on PPP2R5E–KO by shRNA (Liu et al., 2014).

The infection assays performed using a low–passage DENV1 strain from Thailand evidence antiviral activity when we independently overexpress *GRIP1* and *PPP2R5E* that is lost over time due to virus saturation. The independent knockout of the genes seemed to lead to an increase in infection, but this increase was very limited and did not reach significant values. This result may be related to the basal low protein expression observed in the wild–type cell line. These stable transformed cell lines will allow us to continue pursuing functional tests to try to explain the molecular mechanisms behind the significant decrease in DENV1 infection rate when overexpressing these genes. The next step will be to obtain the complete

transcriptome in transformed cell lines infected at MOI 0.5, 24 h after infection to investigate the pathways that were altered by the protein changes and by DENV1 infection.

GRIP1 is a known interactor of Interferon regulatory factor 3 (IRF3) (Reily et al., 2006), that has a crucial role in the innate immune response against DNA and RNA viruses by regulating the transcription of type I IFN genes (IFN–alpha and IFN–beta) and IFN–stimulated genes (ISG). The transient overexpression of *GRIP1* (by using a pcDNA3 vector) was able to rescue *IRF3* gene-dependent expression, due to IRF3 and glucocorticoids competition for GRIP1, and the opposite effect was observed with *GRIP1* knockdown. DENV1 non–structural proteins NS2A and NS4B antagonized the IRF3–dependent immune response by blocking its activation by phosphorylation, which is essential for the innate immune response in the early stages of dengue infection (Dalrymple et al., 2015). We suggest that overexpression of *GRIP1* in the early time-points leads to an increased interaction with IRF3 and consequent activation of innate immune response that is still present 48 h after infection at lower MOIs, due to lower virus saturation.

In an earlier work (Oliveira et al., 2018), we reported the interaction assessed by immunofluorescence assay between PPP2R5E and NS5 of both DENV1 and DENV2 serotypes, yet with distinct dynamics: while PPP2R5E transiently entered the nucleus following DENV2–NS5 expression, it remained in the cytoplasm when DENV1–NS5 was present (Oliveira et al., 2018). Different NS5 localizations were previously described (Hannemann et al., 2013). In addition to our observations, recently, the PP2A–B56 subunit was associated with other viruses. Ebola virus (EBOV) nucleoprotein also recruits PP2A–B56 phosphatase, in order to dephosphorylate the transcription factor VP30, allowing the transcription of EBOV genome (Kruse et al., 2017). Plus, when PP2A–B56 was inhibited, EBOV genome transcription was reduced. On the other hand, HIV–1 accessory protein Vif targets the five elements belonging to PP2A–B56 regulatory subunit for degradation (Greenwood et al., 2016), a required step for *in vivo* pathogenesis. Regarding DENV, there is only one additional study considering PP2A (Soto-Acosta et al., 2017). In this work, no changes in PP2A activity were observed when infecting with DENV2 or DENV4, although PP2A activity decrease with its antagonist okadaic acid in a dose dependent-manner was accompanied by a reduction in the viral titer, most accentuated in DENV2. In our system, the overexpression of *PPP2R5E* provoked a substantial decrease of DENV1 viral titer at both MOIs, and *PPP2R5E* expression was downregulated in wild type HEK 293 cells at 24 h after infection, recovering at 48 h after infection. Our data suggest that *PPP2R5E*

may be involved in some antiviral regulation in the early phase of infection. However, our KO cell line did not register a significant increase in infection, which may be due to function redundancy among the five components of PP2A–B56, since they are highly conserved (Sommer et al., 2015). Given that PP2A activity is regulated by PP2A–B56 subunits, an overexpression of PPP2R5E could lead to the increment of PP2A activity; however, the effect observed in the DENV1 replication did not follow the previously reported for DENV2 and DENV4, where the reduction of PP2A activity conducted to a decrease in the virus yield. It also contradicts eQTL data from GTEx, where the protective SNPs caused a lower *PPP2R5E* mRNA expression (see (Oliveira et al., 2018)). We have still to ascertain if the current data obtained for DENV1 virus replication and infectivity can be replicated using other serotypes as DENV2. This assay will help understanding if there a strain–dependent phenomenon, to join to the transient translocation to the nucleus of PPP2R5E–DENV2–NS5 proteins and not of PPP2R5E–DENV1–NS5. We are now initiating these assays with DENV2, which was not available so far in our lab.

Acknowledgements

References

Ang, L.W., Cutter, J., James, L., 2015. Seroepidemiology of dengue virus infection in the adult population in tropical Singapore. Epidemiol. Infect. 1585-1593.

Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., Drake, J.M., Brownstein, J.S., Hoen, A.G., Sankoh, O., Myers, M.F., George, D.B., Jaenisch, T., Wint, G.R.W., Simmons, C.P., Scott, T.W., Farrar, J.J., Hay, S.I., 2013. The global distribution and burden of dengue. Nature 496(7446), 504-507.

Bijli, K.M., Fazal, F., Slavin, S.A., Leonard, A., Grose, V., Alexander, W.B., Smrcka, A. V., Rahman, A., 2016. Phospholipase C-є signaling mediates endothelial cell inflammation and barrier disruption in acute lung injury. Am. J. Physiol. – Lung Cell. Mol. Physiol. 311(2), L517-L524.

Brown, T.U., Babb, K., Nimrod, M., Carrington, C.V.F., Salas, R.A., 2004. A Retrospective Study of the 1996 DEN-1 Epidemic in Trinidad : Demographic and Clinical Features 28, 7-19.

Büssow, K., 2015. Stable mammalian producer cell lines for structural biology. Curr. Opin. Struct. Biol. 32(Table 1), 81-90.

Cortes-Gonzalez, C., Barrera-Chimal, J., Ibarra-Sanchez, M., Gilbert, M., Gamba, G., Zentella, A., Flores, M.E., Bobadilla, N.A., 2010. Opposite effect of Hsp90alpha and Hsp90beta on eNOS ability to produce nitric oxide or superoxide anion in human embryonic kidney cells. Cell. Physiol. Biochem. 26(15), 657-668.

Dalrymple, N.A., Cimica, V., Mackow, E.R., 2015. Dengue Virus NS Proteins Inhibit RIG-I / MAVS Signaling by Blocking TBK1 / IRF3 Phosphorylation : Dengue Virus Serotype 1 NS4A Is a Unique Interferon-Regulating Virulence Determinant. MBio 6(3), 1-12.

Esmail, S., Yao, Y., Kartner, N., Li, J., Reithmeier, R.A.F., Manolson, M.F., 2016. N-Linked Glycosylation Is Required for Vacuolar H+-ATPase (V-ATPase) a4 Subunit Stability, Assembly, and Cell Surface Expression. J. Cell. Biochem. 2768(April),

2757-2768.

Georgopoulou, U., Tsitoura, P., Kalamvoki, M., Mavromara, P., 2006. The protein phosphatase 2A represents a novel cellular target for hepatitis C virus NS5A protein. Biochimie 88(6), 651-62.

Greenwood, E.J.D., Matheson, N.J., Wals, K., van den Boomen, D.J.H., Antrobus, R., Williamson, J.C., Lehner, P.J., 2016. Temporal proteomic analysis of HIV infection reveals remodelling of the host phosphoproteome by lentiviral Vif variants. Elife 5(September2016), 1-30.

Gupte, R., Muse, G.W., Chinenov, Y., Adelman, K., Rogatsky, I., 2013. Glucocorticoid receptor represses proinflammatory genes at distinct steps of the transcription cycle. Proc. Natl. Acad. Sci. U. S. A. 110(36), 14616-14621.

Guzman, M.G., Gubler, D.J., Izquierdo, A., Martinez, E., Halstead, S.B., 2016. Dengue infection. Nat. Rev. Dis. Prim. 2, 16055.

Guzman, M.G., Halstead, S.B., Artsob, H., Buchy, P., Farrar, J., Gubler, D.J., Hunsperger, E., Kroeger, A., Margolis, H.S., Martínez, E., Nathan, M.B., Pelegrino, J.L., Simmons, C., Yoksan, S., Peeling, R.W., 2010. Dengue: a continuing global threat. Nat. Rev. Microbiol. 8(12), S7-S16.

Hannemann, H., Sung, P.-Y., Chiu, H.-C., Yousuf, A., Bird, J., Lim, S.P., Davidson, A.D., 2013. Serotype-specific differences in dengue virus non-structural protein 5 nuclear localization. J. Biol. Chem. 288(31), 22621-35.

Hertz, E.P.T., Kruse, T., Davey, N.E., López-Méndez, B., Sigurðsson, J.O., Montoya, G., Olsen, J. V., Nilsson, J., 2016. A Conserved Motif Provides Binding Specificity to the PP2A-B56 Phosphatase. Mol. Cell 63(4), 686-695.

Hunt, T.A., Urbanowski, M.D., Kakani, K., Law, L.M.J., Brinton, M.A., Hobman, T.C., 2007. Interactions between the West Nile virus capsid protein and the host cell-encoded phosphatase inhibitor, I2PP2A. Cell. Microbiol. 9(11), 2756-2766.

Kato, M., Takaishi, H., Yoda, M., Tohmonda, T., Takito, J., Fujita, N., Hosogane, N., Horiuchi, K., Kimura, T., Okada, Y., Saito, T., Kawaguchi, H., Kikuchi, T., Matsumoto, M., Toyama, Y., Chiba, K., 2010. GRIP1 enhances estrogen receptor α-dependent extracellular matrix gene expression in chondrogenic cells. Osteoarthr. Cartil. 18(7), 934-941.

Khan, K.H., 2013. Gene expression in mammalian cells and its applications. Adv. Pharm. Bull. 3(2), 257-263.

Khor, C.C., Chau, T.N.B., Pang, J., Davila, S., Long, H.T., Ong, R.T.H., Dunstan, S.J., Wills, B., Farrar, J., Van Tram, T., Gan, T.T., Binh, N.T.N., Tri, L.T., Lien, L.B., Tuan,

N.M., Tham, N.T.H., Lanh, M.N., Nguyet, N.M., Hieu, N.T., Van N Vinh Chau, N., Thuy, T.T., Tan, D.E.K., Sakuntabhai, A., Teo, Y.-Y., Hibberd, M.L., Simmons, C.P., 2011. Genome-wide association study identifies susceptibility loci for dengue shock syndrome at MICB and PLCE1. Nat. Genet. 43(11), 1139-1141.

Kruse, T., Biedenkopf, N., Hertz, E.P.T., Dietzel, E., Stalmann, G., López-Méndez, B., Davey, N.E., Nilsson, J., Becker, S., 2017. The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30. Mol. Cell 1-10.

Kwissa, M., Nakaya, H.I., Onlamoon, N., Wrammert, J., Villinger, F., Perng, G.C., Yoksan, S., Pattanapanyasat, K., Chokephaibulkit, K., Ahmed, R., Pulendran, B., 2014. Dengue virus infection induces expansion of a CD14(+)CD16(+) monocyte population that stimulates plasmablast differentiation. Cell Host Microbe 16(1), 115-127.

Kwun, H.J., Guastafierro, A., Shuda, M., Meinke, G., Bohm, A., Moore, P.S., Chang, Y., 2009. The Minimum Replication Origin of Merkel Cell Polyomavirus Has a Unique Large T-Antigen Loading Architecture and Requires Small T-Antigen Expression for Optimal Replication. J. Virol. 83(23), 12118-12128.

Li, J., Summerlin, M., Nitiss, K.C., Nitiss, J.L., Hanakahi, L.A., 2017. TDP1 is required for efficient non-homologous end joining in human cells. DNA Repair (Amst). 60(October), 40-49.

Lin, Y.J., Chang, J.S., Liu, X., Tsang, H., Chien, W.K., Chen, J.H., Hsieh, H.Y., Hsueh, K.C., Shiao, Y.T., Li, J.P., Lin, C.W., Lai, C.H., Wu, J.Y., Chen, C.H., Lin, J.G., Lin, T.H., Liao, C.C., Huang, S.M., Lan, Y.C., Ho, T.J., Liang, W.M., Yeh, Y.C., Lin, J.C., Tsai, F.J., 2015. Genetic variants in PLCB4/PLCB1 as susceptibility loci for coronary artery aneurysm formation in Kawasaki disease in Han Chinese in Taiwan. Sci. Rep. 5(October), 1-12.

Liu, X., Liu, Q., Fan, Y., Wang, S., Liu, X., Zhu, L., Liu, M., Tang, H., 2014. Downregulation of PPP2R5E expression by miR-23a suppresses apoptosis to facilitate the growth of gastric cancer cells. FEBS Lett. 588(17), 3160-3169.

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron, J., Fleming, J., Siminoff, L., Traino, H., Mosavel, M., Barker, L., Jewell, S., Rohrer, D., Maxim, D., Filkins, D., Harbach, P., Cortadillo, E., Berghuis, B., Turner, L., Hudson, E., Feenstra, K., Sobin, L., Robb, J., Branton, P., Korzeniewski, G., Shive, C., Tabor, D., Qi, L., Groch, K., Nampally, S.,

Buia, S., Zimmerman, A., Smith, A., Burges, R., Robinson, K., Valentino, K., Bradbury, D., Cosentino, M., Diaz-Mayoral, N., Kennedy, M., Engel, T., Williams, P., Erickson, K., Ardlie, K., Winckler, W., Getz, G., DeLuca, D., Daniel MacArthur, Kellis, M., Thomson, A., Young, T., Gelfand, E., Donovan, M., Meng, Y., Grant, G., Mash, D., Marcus, Y., Basile, M., Liu, J., Zhu, J., Tu, Z., Cox, N.J., Nicolae, D.L., Gamazon, E.R., Im, H.K., Konkashbaev, A., Pritchard, J., Stevens, M., Flutre, T., Wen, X., Dermitzakis, E.T., Lappalainen, T., Guigo, R., Monlong, J., Sammeth, M., Koller, D., Battle, A., Mostafavi, S., McCarthy, M., Rivas, M., Maller, J., Rusyn, I., Nobel, A., Wright, F., Shabalin, A., Feolo, M., Sharopova, N., Sturcke, A., Paschal, J., Anderson, J.M., Wilder, E.L., Derr, L.K., Green, E.D., Struewing, J.P., Temple, G., Volpi, S., Boyer, J.T., Thomson, E.J., Guyer, M.S., Ng, C., Abdallah, A., Colantuoni, D., Insel, T.R., Koester, S.E., A Roger Little, Bender, P.K., Lehner, T., Yao, Y., Compton, C.C., Vaught, J.B., Sawyer, S., Lockhart, N.C., Demchok, J., Moore, H.F., 2013. The Genotype-Tissue Expression (GTEx) project. Nat. Genet. 45(6), 580-585.

Miao, C.M., He, K., Li, P.Z., Liu, Z.J., Zhu, X.W., Ou, Z.B., Ruan, X.Z., Gong, J.P., Liu, C.A., 2016. LXRalpha represses LPS-induced inflammatory responses by competing with IRF3 for GRIP1 in Kupffer cells. Int. Immunopharmacol. 35, 272-279.

Min, G., Kim Kemper, J., Kemper, B., 2002. Glucocorticoid receptor-interacting protein 1 mediates ligand-independent nuclear translocation and activation of constitutive androstane receptor in vivo. J. Biol. Chem. 277(29), 26356-26363.

Oliveira, M., Lert-itthiporn, W., Cavadas, B., Chuansumrit, A., Anunciac, O., Casademont, I., Koeth, F., Penova, M., Tangnararatchakit, K., Khor, C.C., Paul, R., Malasit, P., Matsuda, F., Simon-Iori, E., 2018. Joint ancestry and association test indicate two distinct pathogenic pathways involved in classical dengue fever and dengue shock syndrome. PLoS Negl. Trop. Dis. 1-20.

Quach, H., Rotival, M., Pothlichet, J., Loh, Y.H.E., Dannemann, M., Zidane, N., Laval, G., Patin, E., Harmant, C., Lopez, M., Deschamps, M., Naffakh, N., Duffy, D., Coen, A., Leroux-Roels, G., Clément, F., Boland, A., Deleuze, J.F., Kelso, J., Albert, M.L., Quintana-Murci, L., 2016. Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. Cell 167(3), 643-656.e17.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8(11), 2281-2308.

Rattanaburee, T., Junking, M., Panya, A., Sawasdee, N., Songprakhon, P., Suttitheptumrong, A., Limjindaporn, T., Haegeman, G., Yenchitsomanus, P.T., 2015. Inhibition of dengue virus production and cytokine/chemokine expression

by ribavirin and compound A. Antiviral Res. 124, 83-92.

Reily, M.M., Pantoja, C., Hu, X., Chinenov, Y., Rogatsky, I., 2006. The GRIP1:IRF3 interaction as a target for glucocorticoid receptor-mediated immunosuppression. EMBO J. 25(1), 108-17.

Shekhar, K.C., Huat, O.L., 1992. Epidemiology of dengue/dengue hemorrhagic fever in Malaysia--a retrospective epidemiological study. 1973-1987. Part II: Dengue fever (DF). Asia-Pacific J. public Heal. 6(3), 126-133.

Sierra, B., Kouri, G., Guzman, M.G., 2007. Race: a risk factor for dengue hemorrhagic fever. Arch. Virol. 152(3), 533-542.

Sierra, B., Triska, P., Soares, P., Garcia, G., Perez, A.B., Aguirre, E., Oliveira, M., Cavadas, B., Alvarez, M., Ruiz, D., Samuels, D.C., Sakuntabhai, A., Pereira, L., Guzman, M.G., 2017. OSBPL10 , RXRA and lipid metabolism confer African-ancestry protection against dengue haemorrhagic fever in admixed Cubans. PLoS Pathog. 13(2), 1-21.

Simon-Loriere, E., Duong, V., Tawfik, A., Ung, S., Ly, S., Casadémont, I., Prot, M., Courtejoie, N., Bleakley, K., Buchy, P., Tarantola, A., Dussart, P., Cantaert, T., Sakuntabhai, A., 2017. Increased adaptive immune responses and proper feedback regulation protect against clinical dengue. Sci. Transl. Med. 9(405).

Sommer, L.M., Cho, H., Choudhary, M., Seeling, J.M., 2015. Evolutionary analysis of the B56 gene family of PP2A regulatory subunits. Int. J. Mol. Sci. 16(5), 10134-10157.

Soto-Acosta, R., Bautista-Carbajal, P., Cervantes-Salazar, M., Angel-Ambrocio, A.H., del Angel, R.M., 2017. DENV up-regulates the HMG-CoA reductase activity through the impairment of AMPK phosphorylation: A potential antiviral target. PLOS Pathog. 13(4), e1006257.

Vasilakis, N., Weaver, S.C., 2008. The History and Evolution of Human Dengue Emergence, Advances in Virus Research. Elsevier Inc.

Visel, A., 2004. GenePaint.org: an atlas of gene expression patterns in the mouse embryo. Nucleic Acids Res. 32(90001), 552D-556.

WHO, 2017. GACVS Statement on Dengvaxia ® (CYD-TDV) [WWW Document]. December 7, 2017. URL http://www.who.int/vaccine_safety/committee/GACVS-StatementonDengvaxia-CYD-TDV/en/

Zhang, W., Xiao, D., Shan, L., Zhao, J., Mao, Q., Xia, H., 2017. Generation of apoptosis-resistant HEK293 cells with CRISPR/Cas mediated quadruple gene knockout for improved protein and virus production. Biotechnol. Bioeng. 114(11), 2539-2549.

Supplements

Table C.1 – Primers used for pcDNA3.1 cloning and CRISPR/Cas9 guideRNAs.

Table C.2 – PCR Primers designed for CRISPR/Cas9 cell lines editing confirmation.

Figure C.1– mRNA levels (in transcripts per million) for *GRIP1* and *PPP2R5E* in four different cell lines. Data extracted from The Human Protein Atlas (https://www.proteinatlas.org/).

Figure C.2– Proliferation of wild type HEK 293 and transformed cell lines. (A) Fold Change at 48h and 72h after plating (normalize to 24h values); (B) Fold Change at 72h after plating (normalize to 48h values).

Figure C.3– CRISPR/Cas9 editing confirmation by Sanger sequencing. (A) Detection of a 4 bp deletion in *GRIP1* knockout cell line; (B) Detection of a 1 bp insertion in *PPP2R5E* knockout cell line.

IV. Concluding remarks

Literature in dengue fever is populated by papers dealing with one or a couple of immune-related candidate markers, genotyped in one case-control sample from Southeast Asia or Latin America (e.g. (K Alagarasu et al., 2013b, 2013c; Kalichamy Alagarasu et al., 2013a; Dang et al., 2016, 2014; Fernández-Mestre et al., 2004; Fernando et al., 2015; García-Trejo et al., 2011; García et al., 2010; Loke et al., 2002; Oliveira et al., 2014; Sakuntabhai et al., 2005; Sam et al., 2015; Sanchez-Leyva et al., 2017; Wang et al., 2011; Whitehorn et al., 2013; Xavier-Carvalho et al., 2013). Usually those studied samples have very limited sample sizes. There is a high heterogeneity in obtained results, from non-significant associations to opposite significant associations, either positive of negative, for the same marker in sometimes related populations. In this thesis, he began by performing a careful population genetics investigation of ten of these markers from seven genes, followed by meta-analyses in all available and our own cohorts from Southeast Asia or Latin America. These markers were selected amongst the most-studied ones, which allowed to gather information for thousands of individuals (the largest genotyped cohorts were for *MICB*-rs3132468-C, genotyped in 9493 patients (4360 DSS, 224 DHF, 3971 DF and 938 with symptoms) *vs* 6785 controls; the smallest were for *OAS3*-rs1859330-G, with 931 patients (187 DSS, 412 DHF, 304 DF and 28 with symptoms) *vs* 242 controls). This approach contributed with important insights into host genetic susceptibility to dengue fever, allowing us to begin establishing a comprehensive framework on this issue.

In agreement with results from meta-analysis of GWAS for non-strong uncoupling diseases (Park et al., 2011; Rosenberg et al., 2010), we confirmed that seven of these markers (in *PLCE1*, *CD32*, *CD209*, *OAS1* and *OAS3* genes) have high-frequency MAF (20-50%), and the other three (in *MICB* and *TNFA* genes) have intermediate-frequency MAF (5-20%). Most of these markers (highest for *CD209* and *OAS1* genes) have high discriminatory power between population groups, attaining F_{ST} values between 0.20-0.32 in comparisons involving sub-Saharan African and Northeast/Southeast Asian ancestries. Genetic distances for these markers are positively related with geographic distances, following the rules of genetic drift. None of the markers showed signs of strong selective pressure.

A very important insight from our meta-analyses is that it supports the structuring of dengue patients into specific phenotype classes, even at the loss of sample size for the statistical evaluation. It seems that particular polymorphisms can confer susceptibility or protection against particular molecular phenomena, which play a different role along the disease course. *TNFA* seems to protect for DF, *CD32* protects

for DHF, while *PLCE1*, *MICB* and *OAS3* play a role at the level of DSS. This fact is supported by our second work (Oliveira et al., 2018), in which we demonstrated in a GWAS conducted in Thai cohorts that four genes involved in xenobiotic metabolism are protecting against DF, while a new gene (*PLCB4*) of the *PLCE1*-family protects against DSS. There was a directional (causative or protective) consistency between studies performed in East Asia and Latin America in dengue, a shared trend with other high-frequency markers discovered in trans-ethnic GWAS for other diseases (Marigorta and Navarro, 2013; Ntzani et al., 2012). Significant OR values obtained in the meta-analyses were close to each other, between 0.55 for *OAS3*-rs2285933-G, 0.67 for *TNFA*-rs1800629-A, 0.80 for *PLCE1*-rs2274223-G and 0.82 for *CD32*-rs1801274-G, despite differences in MAFs between those markers. Even the causative OR was of similar order of magnitude, 1.32 for *MICB*-rs3132468-C. Also for the markers discovered in the Thai GWAS (Oliveira et al., 2018), OR values varied between 0.54-0.62 for the protective alleles/haplotypes and between 1.48-1.89 for the causative ones. Our meta-analysis was not informative enough to replicate overall the significant association between *CD209*-rs4804803-G and DF/DHF (Sakuntabhai et al., 2005) in Thai, not supporting collected functional evidence that this polymorphism interferes with response to DENV infection (Sakuntabhai et al., 2005). This result is just an example amongst several failed replication studies, even of associations that turned out to be genuine (Lohmueller et al., 2003). Further screening of this SNP can help in clarifying this issue. Current reporting status on genotyping in dengue cohorts does not allow to evaluate the influence of dengue virus strain in the associations detected in these markers, a factor that Simon-Loriere et al. (Simon-Loriere et al., 2015) showed to be of importance in markers from *OAS* family. As it is becoming increasingly evident that different dengue virus strains behave differently, including at molecular level (Hannemann et al., 2013), this issue merits a call of attention to the dengue research community, whom should begin to include this information in their publications even if in that specific cohort there is no sufficient sample size to evaluate the effect of this variable. A proper inclusion of strain variability can be evaluated in meta-analyses.

The results obtained in the meta-analyses for the ten markers enabled us to provide the first genetic risk estimation conferred by these markers, at a worldwide scale. Adding up the risk scores conferred by *PLCE1*, *MICB*, *OAS3* and *CD32* in the severest forms of dengue DHF and DSS, sub-Saharan populations and their descendants are the best protected against the severer dengue phenotypes, in contrast with both Northeast Asian and Southeast populations, which are least protected. European

populations and neighbour North African and Southwest Asian populations are in between those genetic risk scores. The protection conferred by *TNFA* to DF is higher in European, North African and Southwest Asian populations, and decreases towards Africa and especially East Asia. These results fit nicely epidemiological registries in endemic regions (Brown et al., 2004b; Shekhar and Huat, 1992) and provide first insights into still largely naïve regions (Europe and North America) as well as regions with poor health registries (Africa). We also verified in the second paper presented in this thesis that the genetic risk conferred by the four genes involved in xenobiotic metabolism (*CHST10*, *AHRR*, *GRIP1* and *PPP2R5E*) to DF, and of *PLCB4, PLCE1* and *MICB* add precisely to the genetic risks displayed by the previous ten markers, and, also to *OSBPL10* haplotype risk described by our group in a Cuban cohort (Sierra et al. 2017).

GWAS have the advantage of provide unbiased genome-wide evaluations when compared with the pin-point old strategy of test obvious immune-related genes. In dengue context, the needed sample sizes in the order of thousands were difficult to cope with, implying huge efforts in sample collection and financing, especially in Latin American countries where epidemic episodes are sparse by several years. The net result was only one traditional GWAS in a Vietnamese DSS cohort. Our group (Sierra et al. 2017) was the first to apply improved statistical methods that allow to take advantage of ethnicity information (Maples et al., 2013; Shriner et al., 2011). These methods analyse blocks of SNPs that can be affiliated to one of the ancestral populations. The blocks are randomly distributed by recombination across the chromosomes of the admixed descendant individuals, reflecting the frequency contributions of the parental populations (expectation for the control group), except in genomic locations where a candidate gene confers susceptibility to a certain disease (significant frequency increase in the parental population with higher MAF for the susceptibility marker) in cases. As there are only hundreds of ancestral blocks, these tests have considerable lower statistical burdens than the traditional GWAS with thousands of SNPs. Our work in Cuban cohorts (Sierra et al., 2017) allowed to map new candidate genes associated with DHF, the lipid metabolism-related *OSBPL10* and *RXRA.* In this thesis, as already mentioned, we conducted the same strategy in Thai cohorts for DF and DSS phenotypes. We reinforced the association of phospholipase C gene family with the DSS phenotype, by identifying *PLCB4* gene. Both *PLCB4* and *PLCE1* have been implicated in a high number of signal transduction pathways (Cocco et al., 2015) and on immune regulation (Kawakami and Xiao, 2013), and also in diseases presenting similar phenotype to the severest forms of dengue (Lin et al.,

2015). Four genes (*CHST10*, *AHRR*, *PPP2R5E* and *GRIP1*) from the xenobiotic metabolism signalling pathway were identified as protective against DF in the Thai patients. Several independent studies had already related these genes with infectious disease, and even dengue (Audas et al., 2016; Georgopoulou et al., 2006; Maertens, 2016). Ours was the first genomic independent observation.

Again, our genetic results from the two first papers in this thesis have provided evidence that different genes/pathways contribute to the genetic susceptibility or resistance to different outcome of dengue infection. We suggest that xenobiotics and lipid metabolism, as well as interaction of viral proteins to these molecules and to its phosphatases, are critical in the development of classical DF, whereas more severe forms of dengue are caused by over reactive immunity leading to cytokine storm and/or defect in endothelial cell dysfunction and coagulation system. We further tried to conduct some functional assays to follow-up on our GWAS results, as these have the potential to reveal the molecular mechanisms that evolution explored as natural defences against dengue disease, besides cleaning up the literature from spurious association signals.

In order to evaluate the involvement of these genes in interaction with viral proteins, we performed immunofluorescence assays in PPP2R5E protein, for which a recent publication (Hertz et al., 2016) attributed a specific link motif to several proteins, including DENV proteins as we confirmed by *in silico* aligning. We selected NS5 from DENV1 and DENV2 to perform the tests, as this is a crucial viral protein responsible for the virus replication at the endoplasmic reticulum (Perera and Kuhn, 2008). The results revealed that PPP2R5E protein co-localizes with the NS5 protein of DENV2, first in the cytoplasm and then in the nucleus. We also showed that this nuclear translocation does not take place when the specific link motif is deleted from the NS5 protein. The immunofluorescence assay with DENV1-NS5 showed that both proteins co-localize, but there was no entrance of the complex of this serotype into the nucleus. This different molecular behaviour agrees with a previous investigation of the NS5 nuclear translocation in the four DENV serotypes, which showed that it only occurs in DENV2 and DENV3 (Hannemann et al., 2013), leading the authors to hypothesize that the NS5 nuclear localization is not strictly required for virus replication but that it is more likely to have an auxiliary function in the life cycle of specific DENV serotypes. Extra factors, such as importin proteins, must contribute to this differential import into the nucleus. Thus, the dynamics of NS5 seems to play a major role in dengue infection, potentially impacting differential strain virulence, and the PPP2R5E-NS5 interaction must be taken into consideration in future studies.

We were further successful in establishing stable transformed cell lines, using HEK 293, to overexpress and knockout *PPP2R5E* and *GRIP1* genes, and already performed infection assays with DENV1. The induced overexpression with a CMV promoter led to significantly elevated yields of PPP2R5E and, especially, GRIP1 proteins. The observed difference in expression efficiency between both proteins may reflect lower yields of recombinant proteins if it allocated in the cytosol (Khan, 2013), and PPP2R5E is uniformly classified as a cytosolic protein, while GRIP1 is localized in the plasma membrane, endoplasmic reticulum and cytoplasmic vesicles. The CRISPR–Cas9 system was applied for the first time to *GRIP1* and *PPP2R5E* genes, being able to ablate completely these proteins in HEK 293 transformed cell lines, albeit presenting vestigial RNA expression.

The infection assays performed using a DENV1 strain from Thailand evidence antiviral activity when we independently overexpress *GRIP1* and *PPP2R5E*. The independent knockout of the genes seemed to lead to an increase in infection, but this increase was vestigial, not reaching significant values. Most probably this is due to the departing low expression protein values observed in the wild type cell line. These stable transformed cell lines will allow us to continue pursuing functional tests to try to explain the molecular mechanisms behind the significant decrease in DENV1 infection rate when overexpressing these genes. The next step will be to obtain the complete transcriptome in transformed cell lines infected at MOI 0.5, 24 h after infection. Literature provides hypotheses that we will investigate: *GRIP1* overexpression in the early time–points may lead to an increased interaction with IRF3 and consequent activation of innate immune response; expectations that an overexpression of *PPP2R5E* would lead to an increment of PP2A activity were opposed in the DENV1 infection test performed by us, which may reflect again differences between strains. We are now initiating infection assays with DENV2, which was not available so far in our lab.

It is quite interesting that all protective genes identified in GWAS (Khor et al., 2011; Oliveira et al., 2018; Sierra et al., 2017) play a role in pathways centred in the key nuclear receptors RXRA, CAR, LXR, PPAR and VDR that act as transcription factors of many important genes. We had already called attention to this fact in Sierra et al. paper, but the Thai GWAS results obtained in this thesis strengthen that claim. RXRA forms heterodimers with VDR, previously identified as dengue protective in Vietnamese (Loke et al., 2002). The RXRA–VDR heterodimers negatively control the expression of several immune function genes. Also, *PLCE1*, protective against DSS in Vietnamese children (Khor et al., 2011), interacts with RXRA in the PPARA/RXRA

activation pathway, positively controlling the expression of genes again related with lipid metabolism. *OSBPL10* discovered in Cuban cohorts acts in the RXRA/LXR activation pathway that controls lipid metabolism in liver cells and in lipid metabolism and cytokines production in macrophages (Sierra et al. 2017). We added now *PLCB4* to *PLCE1* that act both in PPARA/RXRA activation pathway, and the four *CHST10*, *AHRR*, *PPP2R5E* and *GRIP1* genes which interact in several steps of the xenobiotic metabolism where CAR/RXRA and PXR/RXRA dimers control inflammatory response, xenobiotic transport and phase I and II metabolizing enzymes. These multiple association hits in related pathways deserve a closer functional investigation in the future.

The results gathered in this thesis contributed substantial knowledge to the contribution of host ancestry in dengue outcomes, with a special focus on Southeast Asian nexus. We provided new candidates that improved the proportion of known heritability on dengue phenotypes, and began establishing stable in vitro models for its functional evaluation, having already confirmed significant implications in infection rates and viral replication. We clarified associations for previously indicated candidates through meta-analyses and built up genetic risk evaluation at a worldwide scale. Despite candidate genes have been identified in different ancestry backgrounds, they are intermediate- to highly-frequent, shared by all population groups and all agreeing in that Northeast Asian and Southeast Asian populations are the ones in higher risk to DF and DHF/DSS. Definitely, including ancestry information in the evaluation of association is valuable, improving statistical power of smaller cohorts.

References

Abdulla, M.A., Ahmed, I., Assawamakin, A., Bhak, J., Brahmachari, S.K., Calacal, G.C., Chaurasia, A., Chen, C.-H., Chen, J., Chen, Y.-T., Chu, J., Cutiongco-de la Paz, E.M.C., De Ungria, M.C.A., Delfin, F.C., Edo, J., Fuchareon, S., Ghang, H., Gojobori, T., Han, J., Ho, S.-F., Hoh, B.P., Huang, W., Inoko, H., Jha, P., Jinam, T.A., Jin, L., Jung, J., Kangwanpong, D., Kampuansai, J., Kennedy, G.C., Khurana, P., Kim, H.-L., Kim, K., Kim, S., Kim, W.-Y., Kimm, K., Kimura, R., Koike, T., Kulawonganunchai, S., Kumar, V., Lai, P.S., Lee, J.-Y., Lee, S., Liu, E.T., Majumder, P.P., Mandapati, K.K., Marzuki, S., Mitchell, W., Mukerji, M., Naritomi, K., Ngamphiw, C., Niikawa, N., Nishida, N., Oh, B., Oh, S., Ohashi, J., Oka, A., Ong, R., Padilla, C.D., Palittapongarnpim, P., Perdigon, H.B., Phipps, M.E., Png, E., Sakaki, Y., Salvador, J.M., Sandraling, Y., Scaria, V., Seielstad, M., Sidek, M.R., Sinha, A., Srikummool, M., Sudoyo, H., Sugano, S., Suryadi, H., Suzuki, Y., Tabbada, K.A., Tan, A., Tokunaga, K., Tongsima, S., Villamor, L.P., Wang, E., Wang, Y., Wang, H., Wu, J.-Y., Xiao, H., Xu, S., Yang, J.O., Shugart, Y.Y., Yoo, H.-S., Yuan, W., Zhao, G., Zilfalil, B.A., 2009. Mapping human genetic diversity in Asia. Science 326(5959), 1541–1545.

Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., McVean, G.A., 2010. A map of human genome variation from population-scale sequencing. Nature 467(7319), 1061–1073.

Acosta, E.G., Kumar, A., Bartenschlager, R., 2014. Revisiting dengue virus–host cell interaction: New insights into molecular and cellular virology, 1st ed, Advances in Virus Research. Elsevier Inc.

Adhikari, K., Chacon-Duque, J.C., Mendoza-Revilla, J., Fuentes-Guajardo, M., Ruiz-Linares, A., 2017. The Genetic Diversity of the Americas. Annu. Rev. Genomics Hum. Genet. 18, 277–296.

Akey, J.M., Eberle, M.A., Rieder, M.J., Carlson, C.S., Shriver, M.D., Nickerson, D.A., Kruglyak, L., 2004. Population history and natural selection shape patterns of genetic variation in 132 genes. PLoS Biol. 2(10).

Alagarasu, K., Damle, I.M., Bachal, R. V., Mulay, A.P., Shah, P.S., Dayaraj, C., 2013a. Association of promoter region polymorphisms of CD209 gene with clinical outcomes of dengue virus infection in Western India. Infect. Genet. Evol. 17, 239–242.

Alagarasu, K., Honap, T., Damle, I.M., Mulay, A.P., Shah, P.S., Cecilia, D., 2013a. Polymorphisms in the oligoadenylate synthetase gene cluster and its association with clinical outcomes of dengue virus infection. Infect. Genet. Evol. 14, 390–395.

Alagarasu, K., Honap, T., Mulay, A.P., Bachal, R. V, Shah, P.S., Cecilia, D., 2012. Association of vitamin D receptor gene polymorphisms with clinical outcomes of dengue virus infection. Hum. Immunol. 73(11), 1194–1199.

Alagarasu, K., Mulay, A.P., Sarikhani, M., Rashmika, D., Shah, P.S., Cecilia, D., 2013c. Profile of human leukocyte antigen class I alleles in patients with dengue infection from Western India. Hum. Immunol. 74(12), 1624–1628.

Alagarasu, K., Mulay, A.P., Singh, R., Gavade, V.B., Shah, P.S., Cecilia, D., 2013b. Association of HLA-DRB1 and TNF genotypes with dengue hemorrhagic fever. Hum. Immunol. 74(5), 610–617.

Altshuler, D.M., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., Flicek, P., Gabriel, S.B., Gibbs, R.A., Green, E.D., Hurles, M.E., Knoppers, B.M., Korbel, J.O., Lander, E.S., Lee, C., Lehrach, H., Mardis, E.R., Marth, G.T., McVean, G.A., Nickerson, D.A., Schmidt, J.P., Sherry, S.T., Wang, J., Wilson, R.K., Dinh, H., Kovar, C., Lee, S., Lewis, L., Muzny, D., Reid, J., Wang, M., Fang, X., Guo, X., Jian, M., Jiang, H., Jin, X., Li, G., Li, J., Li, Y., Li, Z., Liu, X., Lu, Y., Ma, X., Su, Z., Tai, S., Tang, M., Wang, B., Wang, G., Wu, H., Wu, R., Yin, Y., Zhang, W., Zhao, J., Zhao, M., Zheng, X., Zhou, Y., Gupta, N., Clarke, L., Leinonen, R., Smith, R.E., Zheng-Bradley, X., Grocock, R., Humphray, S., James, T., Kingsbury, Z., Sudbrak, R., Albrecht, M.W., Amstislavskiy, V.S., Borodina, T.A., Lienhard, M., Mertes, F., Sultan, M., Timmermann, B., Yaspo, M.L., Fulton, L., Fulton, R., Weinstock, G.M., Balasubramaniam, S., Burton, J., Danecek, P., Keane, T.M., Kolb-Kokocinski, A., McCarthy, S., Stalker, J., Quail, M., Davies, C.J., Gollub, J., Webster, T., Wong, B., Zhan, Y., Auton, A., Yu, F., Bainbridge, M., Challis, D., Evani, U.S., Lu, J., Nagaswamy, U., Sabo, A., Wang, Y., Yu, J., Coin, L.J.M., Fang, L., Li, Q., Li, Z., Lin, H., Liu, B., Luo, R., Qin, N., Shao, H., Wang, B., Xie, Y., Ye, C., Yu, C., Zhang, F., Zheng, H., Zhu, H., Garrison, E.P., Kural, D., Lee, W.P., Fung Leong, W., Ward, A.N., Wu, J., Zhang, M., Griffin, L., Hsieh, C.H., Mills, R.E., Shi, X., Von Grotthuss, M., Zhang, C., Daly, M.J., Depristo, M.A., Banks, E., Bhatia, G., Carneiro, M.O., Del Angel, G., Genovese, G., Handsaker, R.E., Hartl, C., McCarroll, S.A., Nemesh, J.C., Poplin, R.E., Schaffner, S.F., Shakir, K., Yoon, S.C., Lihm, J., Makarov, V., Jin, H., Kim, W., Cheol Kim, K., Rausch, T., Beal, K., Cunningham, F., Herrero, J., McLaren, W.M., Ritchie, G.R.S., Gottipati, S., Keinan, A., Rodriguez-Flores, J.L., Sabeti, P.C., Grossman, S.R., Tabrizi, S., Tariyal, R., Cooper, D.N., Ball, E. V, Stenson, P.D., Barnes, B., Bauer, M., Keira Cheetham, R., Cox, T., Eberle, M., Kahn, S., Murray, L., Peden, J., Shaw, R., Ye, K., Batzer, M.A., Konkel, M.K., Walker, J.A., MacArthur, D.G., Lek, M., Herwig, R., Shriver, M.D., Bustamante, C.D., Byrnes,

J.K., De La Vega, F.M., Gravel, S., Kenny, E.E., Kidd, J.M., Maples, B.K., Moreno-Estrada, A., Zakharia, F., Halperin, E., Baran, Y., Craig, D.W., Christoforides, A., Homer, N., Izatt, T., Kurdoglu, A.A., Sinari, S.A., Squire, K., Xiao, C., Sebat, J., Bafna, V., Ye, K., Burchard, E.G., Hernandez, R.D., Gignoux, C.R., Haussler, D., Katzman, S.J., James Kent, W., Howie, B., Ruiz-Linares, A., Dermitzakis, E.T., Lappalainen, T., Devine, S.E., Liu, X., Maroo, A., Tallon, L.J., Rosenfeld, J.A., Michelson, L.P., Min Kang, H., Anderson, P., Angius, A., Bigham, A., Blackwell, T., Busonero, F., Cucca, F., Fuchsberger, C., Jones, C., Jun, G., Li, Y., Lyons, R., Maschio, A., Porcu, E., Reinier, F., Sanna, S., Schlessinger, D., Sidore, C., Tan, A., Kate Trost, M., Awadalla, P., Hodgkinson, A., Lunter, G., Marchini, J.L., Myers, S., Churchhouse, C., Delaneau, O., Gupta-Hinch, A., Iqbal, Z., Mathieson, I., Rimmer, A., Xifara, D.K., Oleksyk, T.K., Fu, Y., Liu, X., Xiong, M., Jorde, L., Witherspoon, D., Xing, J., Browning, B.L., Alkan, C., Hajirasouliha, I., Hormozdiari, F., Ko, A., Sudmant, P.H., Chen, K., Chinwalla, A., Ding, L., Dooling, D., Koboldt, D.C., McLellan, M.D., Wallis, J.W., Wendl, M.C., Zhang, Q., Tyler-Smith, C., Albers, C.A., Ayub, Q., Chen, Y., Coffey, A.J., Colonna, V., Huang, N., Jostins, L., Li, H., Scally, A., Walter, K., Xue, Y., Zhang, Y., Gerstein, M.B., Abyzov, A., Balasubramanian, S., Chen, J., Clarke, D., Fu, Y., Habegger, L., Harmanci, A.O., Jin, M., Khurana, E., Jasmine Mu, X., Sisu, C., Degenhardt, J., Stütz, A.M., Keira Cheetham, R., Church, D., Michaelson, J.J., Blackburne, B., Lindsay, S.J., Ning, Z., Frankish, A., Harrow, J., Mu, X.J., Fowler, G., Hale, W., Kalra, D., Barker, J., Kelman, G., Kulesha, E., Radhakrishnan, R., Roa, A., Smirnov, D., Streeter, I., Toneva, I., Vaughan, B., Ananiev, V., Belaia, Z., Beloslyudtsev, D., Bouk, N., Chen, C., Cohen, R., Cook, C., Garner, J., Hefferon, T., Kimelman, M., Liu, C., Lopez, J., Meric, P., O'Sullivan, C., Ostapchuk, Y., Phan, L., Ponomarov, S., Schneider, V., Shekhtman, E., Sirotkin, K., Slotta, D., Zhang, H., Barnes, K.C., Beiswanger, C., Cai, H., Cao, H., Gharani, N., Henn, B., Jones, D., Kaye, J.S., Kent, A., Kerasidou, A., Mathias, R., Ossorio, P.N., Parker, M., Reich, D., Rotimi, C.N., Royal, C.D., Sandoval, K., Su, Y., Tian, Z., Tishkoff, S., Toji, L.H., Via, M., Wang, Y., Yang, H., Yang, L., Zhu, J., Bodmer, W., Bedoya, G., Ming, C.Z., Yang, G., Jia You, C., Peltonen, L., Garcia-Montero, A., Orfao, A., Dutil, J., Martinez-Cruzado, J.C., Brooks, L.D., Felsenfeld, A.L., McEwen, J.E., Clemm, N.C., Duncanson, A., Dunn, M., Guyer, M.S., Peterson, J.L., Lacroute, P., 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491(7422), 56-65.

Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Peltonen, L., Dermitzakis, E., Bonnen, P.E.,

Altshuler, D.M., Gibbs, R.A., de Bakker, P.I.W., Deloukas, P., Gabriel, S.B., Gwilliam, R., Hunt, S., Inouye, M., Jia, X., Palotie, A., Parkin, M., Whittaker, P., Yu, F., Chang, K., Hawes, A., Lewis, L.R., Ren, Y., Wheeler, D., Gibbs, R.A., Muzny, D.M., Barnes, C., Darvishi, K., Hurles, M., Korn, J.M., Kristiansson, K., Lee, C., McCarrol, S.A., Nemesh, J., Dermitzakis, E., Keinan, A., Montgomery, S.B., Pollack, S., Price, A.L., Soranzo, N., Bonnen, P.E., Gibbs, R.A., Gonzaga-Jauregui, C., Keinan, A., Price, A.L., Yu, F., Anttila, V., Brodeur, W., Daly, M.J., Leslie, S., McVean, G., Moutsianas, L., Nguyen, H., Schaffner, S.F., Zhang, Q., Ghori, M.J.R., McGinnis, R., McLaren, W., Pollack, S., Price, A.L., Schaffner, S.F., Takeuchi, F., Grossman, S.R., Shlyakhter, I., Hostetter, E.B., Sabeti, P.C., Adebamowo, C.A., Foster, M.W., Gordon, D.R., Licinio, J., Manca, M.C., Marshall, P.A., Matsuda, I., Ngare, D., Wang, V.O., Reddy, D., Rotimi, C.N., Royal, C.D., Sharp, R.R., Zeng, C., Brooks, L.D., McEwen, J.E., 2010. Integrating common and rare genetic variation in diverse human populations. Nature 467(7311), 52-58.

Alves, M.J., Fernandes, P.L., Amaro, F., Osorio, H., Luz, T., Parreira, P., Andrade, G., Ze-Ze, L., Zeller, H., 2013. Clinical presentation and laboratory findings for the first autochthonous cases of dengue fever in Madeira island, Portugal, October 2012. Euro Surveill. Bull. Eur. sur les Mal. Transm. = Eur. Commun. Dis. Bull. 18(6).

Andrés, A.M., Hubisz, M.J., Indap, A., Torgerson, D.G., Degenhardt, J.D., Boyko, A.R., Gutenkunst, R.N., White, T.J., Green, E.D., Bustamante, C.D., Clark, A.G., Nielsen, R., 2009. Targets of balancing selection in the human genome. Mol. Biol. Evol. 26(12), 2755-2764.

Ang, L.W., Cutter, J., James, L., 2015. Seroepidemiology of dengue virus infection in the adult population in tropical Singapore. Epidemiol. Infect. 1585-1593.

Appanna, R., Ponnampalavanar, S., Lum, L., See, C., Sekaran, S.D., 2010. Susceptible and Protective HLA Class 1 Alleles against Dengue Fever and Dengue Hemorrhagic Fever Patients in a Malaysian Population. PLoS Negl. Trop. Dis. 5(9), 1-10.

Ardlie, K.G., Kruglyak, L., Seielstad, M., 2002. Patterns of linkage disequilibrium in the human genome. Nat. Rev. Genet. 3(4), 299-309.

Atkinson, Q.D., Gray, R.D., Drummond, A.J., 2008. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. Mol. Biol. Evol. 25(2), 468-474.

Audas, T.E., Hardy-Smith, P.W., Penney, J., Taylor, T., Lu, R., 2016. Characterization of nuclear foci-targeting of Luman/CREB3 recruitment factor (LRF/CREBRF) and

its potential role in inhibition of herpes simplex virus-1 replication. Eur. J. Cell Biol. 95(12), 611–622.

Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R., 2015. A global reference for human genetic variation. Nature 526(7571), 68–74.

Baharian, S., Barakatt, M., Gignoux, C.R., Shringarpure, S., Errington, J., Blot, W.J., Bustamante, C.D., Kenny, E.E., Williams, S.M., Aldrich, M.C., Gravel, S., 2016. The Great Migration and African-American Genomic Diversity. PLoS Genet. 12(5), e1006059.

Basu, A., Tang, H., Lewis, C.E., North, K., Curb, J.D., Quertermous, T., Mosley, T.H., Boerwinkle, E., Zhu, X., Risch, N.J., 2009. Admixture mapping of quantitative trait loci for blood lipids in African-Americans. Hum. Mol. Genet. 18(11), 2091–2098.

Bellamy, R., Ruwende, C., Corrah, T., McAdam, K.P., Thursz, M., Whittle, H.C., Hill, A. V, 1999. Tuberculosis and chronic hepatitis B virus infection in Africans and variation in the vitamin D receptor gene. J. Infect. Dis. 179(3), 721–724.

Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., Drake, J.M., Brownstein, J.S., Hoen, A.G., Sankoh, O., Myers, M.F., George, D.B., Jaenisch, T., Wint, G.R.W., Simmons, C.P., Scott, T.W., Farrar, J.J., Hay, S.I., 2013. The global distribution and burden of dengue. Nature 496(7446), 504–507.

Bijli, K.M., Fazal, F., Slavin, S.A., Leonard, A., Grose, V., Alexander, W.B., Smrcka, A. V., Rahman, A., 2016. Phospholipase C-ε signaling mediates endothelial cell inflammation and barrier disruption in acute lung injury. Am. J. Physiol. – Lung Cell. Mol. Physiol. 311(2), L517–L524.

Blanton, R.E., Silva, L.K., Morato, V.G., Parrado, A.R., Nunes, R.T., Dias, J.P., Melo, P.R.S., Reis, E.A.G., Goddard, K.A.B., Rodrigues, S.G., Vasconcelos, P.F.C., Castro, J.M., Reis, M.G., 2008. Genetic ancestry and income are associated with dengue hemorrhagic fever in a highly admixed population. Eur. J. Hum. Genet. 16, 762–765.

Borges, M.C., Castro, L.A., Fonseca, B.A.L. de, 2013. Chloroquine use improves dengue-related symptoms. Mem. Inst. Oswaldo Cruz 108(5), 596–599.

Botigue, L.R., Henn, B.M., Gravel, S., Maples, B.K., Gignoux, C.R., Corona, E., Atzmon, G., Burns, E., Ostrer, H., Flores, C., Bertranpetit, J., Comas, D., Bustamante, C.D., 2013. Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. Proc. Natl. Acad. Sci. 110(29), 11791–11796.

Bowman, L.R., Donegan, S., McCall, P.J., 2016. Is Dengue Vector Control Deficient in Effectiveness or Evidence?: Systematic Review and Meta-analysis. PLoS Negl.

Trop. Dis. 10(3), e0004551.

Bravo, J.R., Guzman, M.G., Kouri, G.P., 1987. 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. 81(5), 816–820.

Bressan, R.B., Dewari, P.S., Kalantzaki, M., Gangoso, E., Matjusaitis, M., Garcia-Diaz, C., Blin, C., Grant, V., Bulstrode, H., Gogolok, S., Skarnes, W.C., Pollard, S.M., 2017. Efficient CRISPR/Cas9-assisted gene targeting enables rapid and precise genetic manipulation of mammalian neural stem cells. Development 144(4), 635–648.

Brown, L.A., Sofer, T., Stilp, A.M., Baier, L.J., Kramer, H.J., Masindova, I., Levy, D., Hanson, R.L., Moncrieft, A.E., Redline, S., Rosas, S.E., Lash, J.P., Cai, J., Laurie, C.C., Browning, S., Thornton, T., Franceschini, N., 2017. Admixture Mapping Identifies an Amerindian Ancestry Locus Associated with Albuminuria in Hispanics in the United States. J. Am. Soc. Nephrol. ASN.2016091010.

Brown, M.G., Salas, R.A., Vickers, I.E., Heslop, O.D., Smikle, M.F., 2011. Dengue HLA associations in Jamaicans. West Indian Med. J. 60(2), 126–131.

Brown, T.U., Babb, K., Nimrod, M., Carrington, C.V.F., Salas, R.A., 2004. A Retrospective Study of the 1996 DEN-1 Epidemic in Trinidad : Demographic and Clinical Features. Dengue Bull. 28, 7–19.

Brucato, N., Fernandes, V., Mazières, S., Kusuma, P., Cox, M.P., Ng'ang'a, J.W., Omar, M., Simeone-Senelle, M.C., Frassati, C., Alshamali, F., Fin, B., Boland, A., Deleuze, J.F., Stoneking, M., Adelaar, A., Crowther, A., Boivin, N., Pereira, L., Bailly, P., Chiaroni, J., Ricaut, F.X., 2018. The Comoros Show the Earliest Austronesian Gene Flow into the Swahili Corridor. Am. J. Hum. Genet. 102(1), 58–68.

Büssow, K., 2015. Stable mammalian producer cell lines for structural biology. Curr. Opin. Struct. Biol. 32(Table 1), 81–90.

Byk, L.A., Gamarnik, A. V., 2016. Properties and Functions of the Dengue Virus Capsid Protein. Annu. Rev. Virol. 3(1), 263–281.

Cann, H.M., de Toma, C., Cazes, L., Legrand, M.-F., Morel, V., Piouffre, L., Bodmer, J., Bodmer, W.F., Bonne-Tamir, B., Cambon-Thomsen, A., Chen, Z., Chu, J., Carcassi, C., Contu, L., Du, R., Excoffier, L., Ferrara, G.B., Friedlaender, J.S., Groot, H., Gurwitz, D., Jenkins, T., Herrera, R.J., Huang, X., Kidd, J., Kidd, K.K., Langaney, A., Lin, A.A., Mehdi, S.Q., Parham, P., Piazza, A., Pistillo, M.P., Qian, Y., Shu, Q., Xu, J., Zhu, S., Weber, J.L., Greely, H.T., Feldman, M.W., Thomas, G., Dausset, J., Cavalli-Sforza, L.L., 2002. A Human Genome Diversity Cell Line Panel. Science (80-.). 296(5566), 261–262.

Cann, R.L., Stoneking, M., Wilson, A.C., 1987. Mitochondrial DNA and human evolution. Nature 325(6099), 31-36.

Cao, J., Wu, L., Zhang, S.M., Lu, M., Cheung, W.K.C., Cai, W., Gale, M., Xu, Q., Yan, Q., 2016. An easy and efficient inducible CRISPR/Cas9 platform with improved specificity for multiple gene targeting. Nucleic Acids Res. 44(19), 1-10.

Cardozo, D.M., Moliterno, R.A., Sell, A.M., Andréia, G., Guelsin, S., Beltrame, L.M., Clementino, S.L., Reis, P.G., Alves, H.V., Mazini, P.S., Eliete, J., Visentainer, L., 2014. Evidence of HLA-DQB1 Contribution to Susceptibility of Dengue Serotype 3 in Dengue Patients in Southern Brazil. J. Trop. Med. 2014(968262).

Carlos, C.C., Oishi, K., Cinco, M.T., Mapua, C.A., Inoue, S., Cruz, D.J., Pancho, M.A., Tanig, C.Z., Matias, R.R., Morita, K., Natividad, F.F., Igarashi, A., Nagatake, T., 2005. Comparison of clinical features and hematologic abnormalities between dengue fever and dengue hemorrhagic fever among children in the Philippines. Am. J. Trop. Med. Hyg. 73(2), 435-440.

Casamassimi, A., Federico, A., Rienzo, M., Esposito, S., Ciccodicola, A., 2017. Transcriptome profiling in human diseases: New advances and perspectives. Int. J. Mol. Sci. 18(8), 1652.

Cavalli-Sforza, L.L., Wilson, A.C., Cantor, C.R., Cook-Deegan, R.M., King, M.C., 1991. Call for a worldwide survey of human genetic diversity: a vanishing opportunity for the Human Genome Project. Genomics 11(2), 490-491.

Centers for Disease Control and Prevention, 2006. Travel-associated dengue - United States, 2005. Morb. Mortal. Wkly. Rep. 55(25), 700-702.

Chacón-Duque, J.C., Adhikari, K., Avendaño, E., Campo, O., Ramirez, R., Rojas, W., Ruiz-Linares, A., Restrepo, B.N., Bedoya, G., 2014. African genetic ancestry is associated with a protective effect on Dengue severity in colombian populations. Infect. Genet. Evol. 27, 89-95.

Chang, A.Y., Fuller, D.O., Carrasquillo, O., Beier, J.C., 2014. Social justice, climate change, and dengue. Health Hum. Rights 16(1), 93-104.

Chau, T.N.B., Anders, K.L., Lien, L.B., Hung, N.T., Hieu, L.T.M., Tuan, N.M., Thuy, T.T., Phuong, L.T., Tham, N.T.H., Lanh, M.N., Farrar, J.J., Whitehead, S.S., Simmons, C.P., 2010. Clinical and virological features of Dengue in Vietnamese infants. PLoS Negl. Trop. Dis. 4(4), e657.

Choi, U.Y., Kang, J., Hwang, Y.S., Kim, Y., 2015. Oligoadenylate synthase-like (OASL) proteins : dual functions and associations with diseases. Exp. Mol. Med. 47(3), e144-6.

Choo, S.Y., 2007. The HLA system: genetics, immunology, clinical testing, and clinical

implications. Yonsei Med. J. 48(1), 11-23.

Chu, H.W., Rios, C., Huang, C., Wesolowska-Andersen, A., Burchard, E.G., O'Connor, B.P., Fingerlin, T.E., Nichols, D., Reynolds, S.D., Seibold, M.A., 2015. CRISPR-Cas9-mediated gene knockout in primary human airway epithelial cells reveals a proinflammatory role for MUC18. Gene Ther. 22(10), 822-829.

Chuansumrit, A., Anantasit, N., Sasanakul, W., Chaiyaratana, W., Tangnararatchakit, K., Butthep, P., Chunhakan, S., Yoksan, S., 2013. Tumour necrosis factor gene polymorphism in dengue infection: association with risk of bleeding. Paediatr. Int. Child Health 33(2), 97-101.

Clapham, H., Cummings, D.A.T., Nisalak, A., Kalayanarooj, S., 2015. Epidemiology of Infant Dengue Cases Illuminates Serotype-Specificity in the Interaction between Immunity and Disease , and Changes in Transmission Dynamics. PLoS Negl Trop Dis 9(12), 1-10.

Cocco, L., Follo, M.Y., Manzoli, L., Suh, P.-G., 2015. Phosphoinositide-specific phospholipase C in health and disease. J. Lipid Res. 56(10), 1853-1860.

Coffey, L.L., Mertens, E., Brehin, A.C., Fernandez-Garcia, M.D., Amara, A., Després, P., Sakuntabhai, A., 2009. Human genetic determinants of dengue virus susceptibility. Microbes Infect. 11(2), 143-156.

Consortium, T.I.H., 2005. A haplotype map of the human genome. Nature 437(7063), 1299-1320.

Cortes-Gonzalez, C., Barrera-Chimal, J., Ibarra-Sanchez, M., Gilbert, M., Gamba, G., Zentella, A., Flores, M.E., Bobadilla, N.A., 2010. Opposite effect of Hsp90alpha and Hsp90beta on eNOS ability to produce nitric oxide or superoxide anion in human embryonic kidney cells. Cell. Physiol. Biochem. 26(15), 657-668.

Dalrymple, N.A., Cimica, V., Mackow, E.R., 2015. Dengue Virus NS Proteins Inhibit RIG-I / MAVS Signaling by Blocking TBK1 / IRF3 Phosphorylation : Dengue Virus Serotype 1 NS4A Is a Unique Interferon-Regulating Virulence Determinant. MBio 6(3), 1-12.

Dang, T.N., Naka, I., Sa-Ngasang, A., Anantapreecha, S., Chanama, S., Wichukchinda, N., Sawanpanyalert, P., Patarapotikul, J., Tsuchiya, N., Ohashi, J., 2014. A replication study confirms the association of GWAS-identified SNPs at MICB and PLCE1in Thai patients with dengue shock syndrome. BMC Med. Genet. 15(1), 58.

Dang, T.N., Naka, I., Sa-ngasang, A., Anantapreecha, S., Wichukchinda, N., 2016. Association of BAK1 single nucleotide polymorphism with a risk for dengue hemorrhagic fever. BMC Med. Genet. 11-15.

Dans, A.L., Dans, L.F., Lansang, M.A.D., Silvestre, M.A.A., Guyatt, G.H., 2017. Review

of a licensed dengue vaccine: Inappropriate subgroup analyses and selective reporting may cause harm in mass vaccination programs. J. Clin. Epidemiol. 1-3.

Dettogni, R.S., Tristão-Sá, R., dos Santos, M., da Silva, F.F., Louro, I.D., 2015. Single nucleotide polymorphisms in immune system genes and their association with clinical symptoms persistence in dengue-infected persons. Hum. Immunol. 76(10), 717-723.

Devignot, S., Sapet, C., Duong, V., Bergon, A., Rihet, P., Ong, S., Lorn, P.T., Chroeung, N., Ngeav, S., Tolou, H.J., Buchy, P., Couissinier-Paris, P., 2010. Genome-wide expression profiling deciphers host responses altered during dengue shock syndrome and reveals the role of innate immunity in severe dengue. PLoS One 5(7).

DGS, 2013. Dengue na Ilha da Madeira – situação em 01 de setembro de 2013. Informação Mens. sobre dengue na Ilha da Madeira 1084, 1.

Diallo, M., Ba, Y., Sall, A.A., Diop, O.M., Ndione, J.A., Mondo, M., Girault, L., Mathiot, C., 2003. Amplification of the sylvatic cycle of dengue virus type 2, Senegal, 1999-2000: entomologic findings and epidemiologic considerations. Emerg. Infect. Dis. 9(3), 362-367.

Durbin, A.P., Kirkpatrick, B.D., Pierce, K.K., Schmidt, A.C., Whitehead, S.S., 2011. Development and clinical evaluation of multiple investigational monovalent DENV vaccines to identify components for inclusion in a live attenuated tetravalent DENV vaccine. Vaccine 29(42), 7242-7250.

Esmail, S., Yao, Y., Kartner, N., Li, J., Reithmeier, R.A.F., Manolson, M.F., 2016. N-Linked Glycosylation Is Required for Vacuolar H+-ATPase (V-ATPase) a4 Subunit Stability, Assembly, and Cell Surface Expression. J. Cell. Biochem. 2768(April), 2757-2768.

Falcon-Lezama, J.A., Ramos, C., Zuniga, J., Juarez-Palma, L., Rangel-Flores, H., Garcia-Trejo, A.R., Acunha-Alonzo, V., Granados, J., Vargas-Alarcon, G., 2009. HLA class I and II polymorphisms in Mexican Mestizo patients with dengue fever. Acta Trop. 112(2), 193-197.

Farrer, L.A., Cupples, L.A., Haines, J.L., Hyman, B., Kukull, W.A., Mayeux, R., Myers, R.H., Pericak-Vance, M.A., Risch, N., van Duijn, C.M., 1997. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. JAMA 278(16), 1349-1356.

Fernandes, V., Alshamali, F., Alves, M., Costa, M.D., Pereira, J.B., Silva, N.M., Cherni, L., Harich, N., Cerny, V., Soares, P., Richards, M.B., Pereira, L., 2012. The Arabian

cradle: Mitochondrial relicts of the first steps along the Southern route out of Africa. Am. J. Hum. Genet. 90(2), 347–355.

Fernández-Mestre, M.T., Gendzekhadze, K., Rivas-Vetencourt, P., Layrisse, Z., 2004. TNF-α-308A allele, a possible severity risk factor of hemorrhagic manifestation in dengue fever patients. Tissue Antigens 64(4), 469–472.

Fernando, A.N., Malavige, G.N., Perera, K.L.N., Premawansa, S., Ogg, G.S., De Silva, A.D., 2015. Polymorphisms of Transporter Associated with Antigen Presentation, Tumor Necrosis Factor-alpha and Interleukin-10 and their Implications for Protection and Susceptibility to Severe Forms of Dengue Fever in Patients in Sri Lanka. J. Glob. Infect. Dis. 7(4), 157–164.

Ferreira, M.A.R., Matheson, M.C., Duffy, D.L., Marks, G.B., Hui, J., Le Souëf, P., Danoy, P., Baltic, S., Nyholt, D.R., Jenkins, M., Hayden, C., Willemsen, G., Ang, W., Kuokkanen, M., Beilby, J., Cheah, F., De Geus, E.J.C., Ramasamy, A., Vedantam, S., Salomaa, V., Madden, P.A., Heath, A.C., Hopper, J.L., Visscher, P.M., Musk, B., Leeder, S.R., Jarvelin, M.R., Pennell, C., Boomsma, D.I., Hirschhorn, J.N., Walters, H., Martin, N.G., James, A., Jones, G., Abramson, M.J., Robertson, C.F., Dharmage, S.C., Brown, M.A., Montgomery, G.W., Thompson, P.J., 2011. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. Lancet 378(9795), 1006–1014.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391(6669), 806–811.

Fraser, J.E., De Bruyne, J.T., Iturbe-Ormaetxe, I., Stepnell, J., Burns, R.L., Flores, H.A., O'Neill, S.L., 2017. Novel Wolbachia-transinfected Aedes aegypti mosquitoes possess diverse fitness and vector competence phenotypes. PLoS Pathog. 13(12), 1–19.

Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., Pasternak, S., Wheeler, D.A., Willis, T.D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., Hu, W., Li, C., Lin, W., Liu, S., Pan, H., Tang, X., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q., Zhao, H., Zhao, H., Zhou, J., Gabriel, S.B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R.C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Altshuler, D., Shen, Y., Yao, Z., Huang, W., Chu, X., He, Y., Jin, L., Liu, Y., Shen, Y., Sun, W., Wang, H., Wang, Y., Wang, Y., Xiong, X., Xu, L., Waye, M.M.Y., Tsui, S.K.W., Xue, H., Wong, J.T.-F., Galver, L.M., Fan, J.-B., Gunderson, K., Murray, S.S., Oliphant, A.R., Chee, M.S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J.-F., Phillips, M.S.,

Roumy, S., Sallee, C., Verner, A., Hudson, T.J., Kwok, P.-Y., Cai, D., Koboldt, D.C., Miller, R.D., Pawlikowska, L., Taillon-Miller, P., Xiao, M., Tsui, L.-C., Mak, W., Song, Y.Q., Tam, P.K.H., Nakamura, Y., Kawaguchi, T., Kitamoto, T., Morizono, T., Nagashima, A., Ohnishi, Y., Sekine, A., Tanaka, T., Tsunoda, T., Deloukas, P., Bird, C.P., Delgado, M., Dermitzakis, E.T., Gwilliam, R., Hunt, S., Morrison, J., Powell, D., Stranger, B.E., Whittaker, P., Bentley, D.R., Daly, M.J., de Bakker, P.I.W., Barrett, J., Chretien, Y.R., Maller, J., McCarroll, S., Patterson, N., Pe'er, I., Price, A., Purcell, S., Richter, D.J., Sabeti, P., Saxena, R., Schaffner, S.F., Sham, P.C., Varilly, P., Altshuler, D., Stein, L.D., Krishnan, L., Smith, A.V., Tello-Ruiz, M.K., Thorisson, G.A., Chakravarti, A., Chen, P.E., Cutler, D.J., Kashuk, C.S., Lin, S., Abecasis, G.R., Guan, W., Li, Y., Munro, H.M., Qin, Z.S., Thomas, D.J., McVean, G., Auton, A., Bottolo, L., Cardin, N., Eyheramendy, S., Freeman, C., Marchini, J., Myers, S., Spencer, C., Stephens, M., Donnelly, P., Cardon, L.R., Clarke, G., Evans, D.M., Morris, A.P., Weir, B.S., Tsunoda, T., Mullikin, J.C., Sherry, S.T., Feolo, M., Skol, A., Zhang, H., Zeng, C., Zhao, H., Matsuda, I., Fukushima, Y., Macer, D.R., Suda, E., Rotimi, C.N., Adebamowo, C.A., Ajayi, I., Aniagwu, T., Marshall, P.A., Nkwodimmah, C., Royal, C.D.M., Leppert, M.F., Dixon, M., Peiffer, A., Qiu, R., Kent, A., Kato, K., Niikawa, N., Adewole, I.F., Knoppers, B.M., Foster, M.W., Clayton, E.W., Watkin, J., Gibbs, R.A., Belmont, J.W., Muzny, D., Nazareth, L., Sodergren, E., Weinstock, G.M., Wheeler, D.A., Yakub, I., Gabriel, S.B., Onofrio, R.C., Richter, D.J., Ziaugra, L., Birren, B.W., Daly, M.J., Altshuler, D., Wilson, R.K., Fulton, L.L., Rogers, J., Burton, J., Carter, N.P., Clee, C.M., Griffiths, M., Jones, M.C., McLay, K., Plumb, R.W., Ross, M.T., Sims, S.K., Willey, D.L., Chen, Z., Han, H., Kang, L., Godbout, M., Wallenburg, J.C., L'Archeveque, P., Bellemare, G., Saeki, K., Wang, H., An, D., Fu, H., Li, Q., Wang, Z., Wang, R., Holden, A.L., Brooks, L.D., McEwen, J.E., Guyer, M.S., Wang, V.O., Peterson, J.L., Shi, M., Spiegel, J., Sung, L.M., Zacharia, L.F., Collins, F.S., Kennedy, K., Jamieson, R., Stewart, J., 2007. A second generation human haplotype map of over 3.1 million SNPs. Nature 449(7164), 851–861.

Fu, W., O'Connor, T.D., Jun, G., Kang, H.M., Abecasis, G., Leal, S.M., Gabriel, S., Altshuler, D., Shendure, J., Nickerson, D.A., Bamshad, M.J., Akey, J.M., 2013. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature 493(7431), 216–220.

Gamble, J., Bethell, D., Day, N.P.J., Loc, P.P., Phu, N.H., Gartside, I.B., Farrar, J.F., White, N.J., 2000. Age-related changes in microvascular permeability: a significant factor in the susceptibility of children to shock? Clin. Sci. 98(2), 211 LP–216.

Gan, C.S., Yusof, R., Othman, S., 2015. Different serotypes of dengue viruses differently regulate the expression of the host cell antigen processing machinery. Acta Trop. 149, 8–14.

García-Trejo, A.R., Falcón-Lezama, J.A., Juárez-Palma, L., Granados, J., Zúñiga-Ramos, J., Rangel, H., Barquera, R., Vargas-Alarcón, G., Ramos, C., 2011. Tumor necrosis factor alpha promoter polymorphisms in Mexican patients with dengue fever. Acta Trop. 120(1-2), 67–71.

García, G., Sierra, B., Pérez, A.B., Aguirre, E., Rosado, I., Gonzalez, N., Izquierdo, A., Pupo, M., Ruiz, D., Díaz, D., Sánchez, L., Marcheco, B., Hirayama, K., Guzmán, M.G., 2010. Asymptomatic dengue infection in a Cuban population confirms the protective role of the RR variant of the FcgammaRIIa polymorphism. Am. J. Trop. Med. Hyg. 82(6), 1153–1156.

Georgopoulou, U., Tsitoura, P., Kalamvoki, M., Mavromara, P., 2006. The protein phosphatase 2A represents a novel cellular target for hepatitis C virus NS5A protein. Biochimie 88(6), 651–62.

Gonzalez, D., Castro, O.E., Kouri, G., Perez, J., Martinez, E., Vazquez, S., Rosario, D., Cancio, R., Guzman, M.G., 2005. 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. 9(5), 280–285.

Gonzalez, J.P., Du Saussay, C., Gautun, J.C., McCormick, J.B., Mouchet, J., 1985. [Dengue in Burkina Faso (ex-Upper Volta): seasonal epidemics in the urban area of Ouagadougou]. Bull. Soc. Pathol. Exot. Filiales 78(1), 7–14.

Greenwood, E.J.D., Matheson, N.J., Wals, K., van den Boomen, D.J.H., Antrobus, R., Williamson, J.C., Lehner, P.J., 2016. Temporal proteomic analysis of HIV infection reveals remodelling of the host phosphoproteome by lentiviral Vif variants. Elife 5(September2016), 1–30.

Grossman, S.R., Andersen, K.G., Shlyakhter, I., Tabrizi, S., Winnicki, S., Yen, A., Park, D.J., Griesemer, D., Karlsson, E.K., Wong, S.H., Cabili, M., Adegbola, R.A., Bamezai, R.N.K., Hill, A.V.S., Vannberg, F.O., Rinn, J.L., Lander, E.S., Schaffner, S.F., Sabeti, P.C., 2013. Identifying recent adaptations in large-scale genomic data. Cell 152(4), 703–713.

Gubler, D.J., 2011. Dengue, Urbanization and Globalization: The Unholy Trinity of the 21st Century. Trop. Med. Health 39(4SUPPLEMENT), S3–S11.

Guedes, D.R.D., Gomes, E.T.B., Paiva, M.H.S., Melo-Santos, M.A.V. d., Alves, J., Gómez, L.F., Ayres, C.F.J., 2017. Circulation of DENV2 and DENV4 in Aedes aegypti (Diptera: Culicidae) mosquitoes from Praia, Santiago Island, Cabo Verde. J. Insect

Sci. 17(4, 1 July 2017), 86.

Gupte, R., Muse, G.W., Chinenov, Y., Adelman, K., Rogatsky, I., 2013. Glucocorticoid receptor represses proinflammatory genes at distinct steps of the transcription cycle. Proc. Natl. Acad. Sci. U. S. A. 110(36), 14616–14621.

Guzman, M.G., Gubler, D.J., Izquierdo, A., Martinez, E., Halstead, S.B., 2016. Dengue infection. Nat. Rev. Dis. Prim. 2, 16055.

Guzman, M.G., Halstead, S.B., Artsob, H., Buchy, P., Farrar, J., Gubler, D.J., Hunsperger, E., Kroeger, A., Margolis, H.S., Martínez, E., Nathan, M.B., Pelegrino, J.L., Simmons, C., Yoksan, S., Peeling, R.W., 2010. Dengue: a continuing global threat. Nat. Rev. Microbiol. 8(12), S7–S16.

Guzman, M.G., Harris, E., 2015. Dengue. Lancet 385(9966), 453–465.

Guzman, M.G., Vazquez, S., 2010. The complexity of antibody-dependent enhancement of dengue virus infection. Viruses 2(12), 2649–2662.

Halstead, S.B., 2017. Which Dengue Vaccine Approach Is the Most Promising, and Should We Be Concerned about Enhanced Disease after Vaccination? There Is Only One True Winner. Cold Spring Harb. Perspect. Biol.

Halstead, S.B., 2008. Dengue virus-mosquito interactions. Annu. Rev. Entomol. 53, 273–291.

Halstead, S.B., 2006. Dengue in the Americas and Southeast Asia: do they differ? Rev. Panam. Salud Pública 20(6), 407–415.

Halstead, S.B., Streit, T.G., Lafontant, J.G., Putvatana, R., Russell, K., Sun, W., Kanesa-Thasan, N., Hayes, C.G., Watts, D.M., 2001. Haiti: absence of dengue hemorrhagic fever despite hyperendemic dengue virus transmission. Am. J. Trop. Med. Hyg. 65(3), 180–183.

Hannemann, H., Sung, P.-Y., Chiu, H.-C., Yousuf, A., Bird, J., Lim, S.P., Davidson, A.D., 2013. Serotype-specific differences in dengue virus non-structural protein 5 nuclear localization. J. Biol. Chem. 288(31), 22621–35.

Hedrick, P.W., 2011. Population genetics of malaria resistance in humans. Heredity (Edinb). 107(4), 283–304.

Hertz, E.P.T., Kruse, T., Davey, N.E., López-Méndez, B., Sigurðsson, J.O., Montoya, G., Olsen, J. V., Nilsson, J., 2016. A Conserved Motif Provides Binding Specificity to the PP2A-B56 Phosphatase. Mol. Cell 63(4), 686–695.

Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., Manolio, T.A., 2009. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. U. S. A. 106(23), 9362-9367.

Huang, C.Y.-H., Kinney, R.M., Livengood, J.A., Bolling, B., Arguello, J.J., Luy, B.E., Silengo, S.J., Boroughs, K.L., Stovall, J.L., Kalanidhi, A.P., Brault, A.C., Osorio, J.E., Stinchcomb, D.T., 2013. Genetic and phenotypic characterization of manufacturing seeds for a tetravalent dengue vaccine (DENVax). PLoS Negl. Trop. Dis. 7(5), e2243.

Hunt, T.A., Urbanowski, M.D., Kakani, K., Law, L.M.J., Brinton, M.A., Hobman, T.C., 2007. Interactions between the West Nile virus capsid protein and the host cell-encoded phosphatase inhibitor, I2PP2A. Cell. Microbiol. 9(11), 2756-2766.

Jallow, M., Teo, Y.Y., Small, K.S., Rockett, K.A., Deloukas, P., Clark, T.G., Kivinen, K., Bojang, K.A., Conway, D.J., Pinder, M., Sirugo, G., Sisay-joof, F., Usen, S., Auburn, S., Bumpstead, S.J., Campino, S., Coffey, A., Dunham, A., Fry, A.E., Green, A., Gwilliam, R., Hunt, S.E., Inouye, M., Jeffreys, A.E., Mendy, A., Palotie, A., Potter, S., Ragoussis, J., Rogers, J., Rowlands, K., Somaskantharajah, E., Whittaker, P., Widden, C., Donnelly, P., Howie, B., Marchini, J., Morris, A., Sanjoaquin, M., Achidi, E.A., Agbenyega, T., Allen, A., Amodu, O., Corran, P., Ibrahim, M., Karunaweera, N., Kokwaro, G., Koram, K.A., Lemnge, M., Makani, J., Marsh, K., Michon, P., Modiano, D., Molyneux, M.E., Mueller, I., Parker, M., Peshu, N., Plowe, C. V, Puijalon, O., Reeder, J., Tall, A., Taylor, T.E., Thera, M., Troye-blomberg, M., Williams, T.N., 2009. Genome-wide and fine-resolution association analysis of malaria in West Africa. Nat. Genet. 41(6), 657-665.

Johansson, M.A., Hombach, J., Cummings, D.A.T., 2011. Models of the impact of dengue vaccines: a review of current research and potential approaches. Vaccine 29(35), 5860-5868.

Kaptein, S.J.F., Neyts, J., 2016. ScienceDirect Towards antiviral therapies for treating dengue virus infections. Curr. Opin. Pharmacol. 30, 1-7.

Kato, M., Takaishi, H., Yoda, M., Tohmonda, T., Takito, J., Fujita, N., Hosogane, N., Horiuchi, K., Kimura, T., Okada, Y., Saito, T., Kawaguchi, H., Kikuchi, T., Matsumoto, M., Toyama, Y., Chiba, K., 2010. GRIP1 enhances estrogen receptor α-dependent extracellular matrix gene expression in chondrogenic cells. Osteoarthr. Cartil. 18(7), 934-941.

Kawakami, T., Xiao, W., 2013. Phospholipase C-beta in immune cells. Adv. Biol. Regul. 53(3), 249-257.

Kentala, H., Weber-Boyvat, M., Olkkonen, V.M., 2016. OSBP-Related Protein Family: Mediators of Lipid Transport and Signaling at Membrane Contact Sites, International Review of Cell and Molecular Biology. Elsevier Inc.

Khan, K.H., 2013. Gene expression in mammalian cells and its applications. Adv.

Pharm. Bull. 3(2), 257–263.

Khor, C.C., Chau, T.N.B., Pang, J., Davila, S., Long, H.T., Ong, R.T.H., Dunstan, S.J., Wills, B., Farrar, J., Van Tram, T., Gan, T.T., Binh, N.T.N., Tri, L.T., Lien, L.B., Tuan, N.M., Tham, N.T.H., Lanh, M.N., Nguyet, N.M., Hieu, N.T., Van N Vinh Chau, N., Thuy, T.T., Tan, D.E.K., Sakuntabhai, A., Teo, Y.-Y., Hibberd, M.L., Simmons, C.P., 2011. Genome-wide association study identifies susceptibility loci for dengue shock syndrome at MICB and PLCE1. Nat. Genet. 43(11), 1139–1141.

Kittigul, L., Pitakarnjanakul, P., Sujirarat, D., Siripanichgon, K., 2007. The differences of clinical manifestations and laboratory findings in children and adults with dengue virus infection. J. Clin. Virol. 39(2), 76–81.

Kivisild, T., Shen, P., Wall, D.P., Do, B., Sung, R., Davis, K., Passarino, G., Underhill, P.A., Scharfe, C., Torroni, A., Scozzari, R., Modiano, D., Coppa, A., De Knijff, P., Feldman, M., Cavalli-Sforza, L.L., Oefner, P.J., 2006. The role of selection in the evolution of human mitochondrial genomes. Genetics 172(1), 373–387.

Klein, R.J., Zeiss, C., Chew, E.Y., Tsai, J.-Y., Sackler, R.S., Haynes, C., Henning, A.K., SanGiovanni, J.P., Mane, S.M., Mayne, S.T., Bracken, M.B., Ferris, F.L., Ott, J., Barnstable, C., Hoh, J., 2005. Complement factor H polymorphism in age-related macular degeneration. Science 308(5720), 385–389.

Knudsen, A.B., Slooff, R., 1992. Vector-borne disease problems in rapid urbanization: new approaches to vector control. Bull. World Health Organ. 70(1), 1–6.

Kourí, G., Guzmán, M.G., Valdés, L., Carbonel, I., Del Rosario, D., Vazquez, S., Laferté, J., Delgado, J., Cabrera, M. V., 1998. Reemergence of dengue in Cuba: A 1997 epidemic in Santiago de Cuba. Emerg. Infect. Dis. 4(1), 89–92.

Kraemer, M.U.G., Sinka, M.E., Duda, K.A., Mylne, A.Q.N., Shearer, F.M., Barker, C.M., Moore, C.G., Carvalho, R.G., Coelho, G.E., Van Bortel, W., Hendrickx, G., Schaffner, F., Elyazar, I.R.F., Teng, H.-J., Brady, O.J., Messina, J.P., Pigott, D.M., Scott, T.W., Smith, D.L., Wint, G.R.W., Golding, N., Hay, S.I., 2015. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. Elife 4, e08347.

Kruse, T., Biedenkopf, N., Hertz, E.P.T., Dietzel, E., Stalmann, G., López-Méndez, B., Davey, N.E., Nilsson, J., Becker, S., 2017. The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30. Mol. Cell 1–10.

Kwissa, M., Nakaya, H.I., Onlamoon, N., Wrammert, J., Villinger, F., Perng, G.C., Yoksan, S., Pattanapanyasat, K., Chokephaibulkit, K., Ahmed, R., Pulendran, B., 2014. Dengue virus infection induces expansion of a CD14(+)CD16(+) monocyte

population that stimulates plasmablast differentiation. Cell Host Microbe 16(1), 115–127.

Kwun, H.J., Guastafierro, A., Shuda, M., Meinke, G., Bohm, A., Moore, P.S., Chang, Y., 2009. The Minimum Replication Origin of Merkel Cell Polyomavirus Has a Unique Large T-Antigen Loading Architecture and Requires Small T-Antigen Expression for Optimal Replication. J. Virol. 83(23), 12118–12128.

L'Azou, M., Moureau, A., Sarti, E., Nealon, J., Zambrano, B., Wartel, T.A., Villar, L., Capeding, M.R.Z., Ochiai, R.L., 2016. Symptomatic Dengue in Children in 10 Asian and Latin American Countries. N. Engl. J. Med. 374(12), 1155–1166.

La Ruche, G., Souares, Y., Armengaud, A., Peloux-Petiot, F., Delaunay, P., Despres, P., Lenglet, A., Jourdain, F., Leparc-Goffart, I., Charlet, F., Ollier, L., Mantey, K., Mollet, T., Fournier, J.P., Torrents, R., Leitmeyer, K., Hilairet, P., Zeller, H., Van Bortel, W., Dejour-Salamanca, D., Grandadam, M., Gastellu-Etchegorry, M., 2010. First two autochthonous dengue virus infections in metropolitan France, September 2010. Euro Surveill. Bull. Eur. sur les Mal. Transm. = Eur. Commun. Dis. Bull. 15(39), 19676.

LaFleur, C., Granados, J., Vargas-Alarcon, G., Ruiz-Morales, J., Villarreal-Garza, C., Higuera, L., Hernandez-Pacheco, G., Cutino-Moguel, T., Rangel, H., Figueroa, R., Acosta, M., Lazcano, E., Ramos, C., 2002. 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. 63(11), 1039–1044.

Li, J., Summerlin, M., Nitiss, K.C., Nitiss, J.L., Hanakahi, L.A., 2017. TDP1 is required for efficient non-homologous end joining in human cells. DNA Repair (Amst). 60(October), 40–49.

Li, J.Z., Absher, D.M., Tang, H., Southwick, A.M., Casto, A.M., Ramachandran, S., Cann, H.M., Barsh, G.S., Feldman, M., Cavalli-sforza, L.L., Myers, R.M., 2008. Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation 25(February), 1100–1105.

Li, Y.R., Keating, B.J., 2014. Trans-ethnic genome-wide association studies: Advantages and challenges of mapping in diverse populations. Genome Med. 6(10), 1–14.

Lin, R.-J., Yu, H.-P., Chang, B.-L., Tang, W.-C., Liao, C.-L., Lin, Y.-L., 2009. Distinct Antiviral Roles for Human 2',5'-Oligoadenylate Synthetase Family Members against Dengue Virus Infection. J. Immunol. 183(12), 8035–8043.

Lin, Y.J., Chang, J.S., Liu, X., Tsang, H., Chien, W.K., Chen, J.H., Hsieh, H.Y., Hsueh, K.C., Shiao, Y.T., Li, J.P., Lin, C.W., Lai, C.H., Wu, J.Y., Chen, C.H., Lin, J.G., Lin,

T.H., Liao, C.C., Huang, S.M., Lan, Y.C., Ho, T.J., Liang, W.M., Yeh, Y.C., Lin, J.C., Tsai, F.J., 2015. Genetic variants in PLCB4/PLCB1 as susceptibility loci for coronary artery aneurysm formation in Kawasaki disease in Han Chinese in Taiwan. Sci. Rep. 5(October), 1–12.

Liu, P., Ridilla, M., Patel, P., Betts, L., Gallichotte, E., Shahidi, L., Thompson, N.L., Jacobson, K., 2017. Beyond attachment: Roles of DC-SIGN in dengue virus infection. Traffic 18(4), 218–231.

Liu, X., Liu, Q., Fan, Y., Wang, S., Liu, X., Zhu, L., Liu, M., Tang, H., 2014. Downregulation of PPP2R5E expression by miR-23a suppresses apoptosis to facilitate the growth of gastric cancer cells. FEBS Lett. 588(17), 3160–3169.

Lohmueller, K.E., Pearce, C.L., Pike, M., Lander, E.S., Hirschhorn, J.N., 2003. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat. Genet. 33(2), 177–182.

Loke, H., Bethell, D., Phuong, C.X.T., Day, N., White, N., Farrar, J., Hill, A., 2002. Susceptibility to dengue hemorrhagic fever in Vietnam: Evidence of an association with variation in the vitamin D receptor and FCγ receptor IIA genes. Am. J. Trop. Med. Hyg. 67(1), 102–106.

Loke, H., Bethell, D.B., Phuong, C.X., Dung, M., Schneider, J., White, N.J., Day, N.P., Farrar, J., Hill, A. V, 2001. Strong HLA class I–restricted T cell responses in dengue hemorrhagic fever: a double-edged sword? J. Infect. Dis. 184(11), 1369–1373.

Long, H.T., Hibberd, M.L., Hien, T.T., Dung, N.M., Van, T., 2009. Europe PMC Funders Group 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 199(4), 537–546.

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron, J., Fleming, J., Siminoff, L., Traino, H., Mosavel, M., Barker, L., Jewell, S., Rohrer, D., Maxim, D., Filkins, D., Harbach, P., Cortadillo, E., Berghuis, B., Turner, L., Hudson, E., Feenstra, K., Sobin, L., Robb, J., Branton, P., Korzeniewski, G., Shive, C., Tabor, D., Qi, L., Groch, K., Nampally, S., Buia, S., Zimmerman, A., Smith, A., Burges, R., Robinson, K., Valentino, K., Bradbury, D., Cosentino, M., Diaz-Mayoral, N., Kennedy, M., Engel, T., Williams, P., Erickson, K., Ardlie, K., Winckler, W., Getz, G., DeLuca, D., Daniel MacArthur, Kellis, M., Thomson, A., Young, T., Gelfand, E., Donovan, M., Meng, Y., Grant, G., Mash, D., Marcus, Y., Basile, M., Liu, J., Zhu, J., Tu, Z., Cox, N.J., Nicolae, D.L., Gamazon, E.R., Im, H.K., Konkashbaev, A., Pritchard, J., Stevens, M., Flutre, T.,

Wen, X., Dermitzakis, E.T., Lappalainen, T., Guigo, R., Monlong, J., Sammeth, M., Koller, D., Battle, A., Mostafavi, S., McCarthy, M., Rivas, M., Maller, J., Rusyn, I., Nobel, A., Wright, F., Shabalin, A., Feolo, M., Sharopova, N., Sturcke, A., Paschal, J., Anderson, J.M., Wilder, E.L., Derr, L.K., Green, E.D., Struewing, J.P., Temple, G., Volpi, S., Boyer, J.T., Thomson, E.J., Guyer, M.S., Ng, C., Abdallah, A., Colantuoni, D., Insel, T.R., Koester, S.E., A Roger Little, Bender, P.K., Lehner, T., Yao, Y., Compton, C.C., Vaught, J.B., Sawyer, S., Lockhart, N.C., Demchok, J., Moore, H.F., 2013.The Genotype–Tissue Expression (GTEx) project. Nat.Genet. 45(6),580–585.

Low, J.G., Sung, C., Wijaya, L., Wei, Y., Rathore, A.P.S., Watanabe, S., Tan, B.H., Toh, L., Chua, L.T., Hou, Y., Chow, A., Howe, S., Chan, W.K., Tan, K.H., Chung, J.S., Cherng, B.P., Lye, D.C., Tambayah, P.A., Ng, L.C., Connolly, J., Hibberd, M.L., Leo, Y.S., Cheung, Y.B., Ooi, E.E., Vasudevan, S.G., 2014. Efficacy and safety of celgosivir in patients with dengue fever (CELADEN): a phase 1b, randomised, double-blind, placebo-controlled, proof-of-concept trial. Lancet. Infect. Dis. 14(8), 706-715.

Low, J.G.H., Ooi, E.E., Vasudevan, S.G., 2017. Current Status of Dengue Therapeutics Research and Development. J. Infect. Dis. 215(Suppl 2), 1-7.

Ma, F., Liu, S.-Y., Razani, B., Arora, N., Li, B., Kagechika, H., Tontonoz, P., Nunez, V., Ricote, M., Cheng, G., 2014. Retinoid X receptor alpha attenuates host antiviral response by suppressing type I interferon. Nat. Commun. 5, 5494.

MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J., MayPendlington, Z., Welter, D., Burdett, T., Hindorff, L., Flicek, P., Cunningham, F., Parkinson, H., 2017. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 45(D1), D896-D901.

Maertens, G.N., 2016. B'-protein phosphatase 2A is a functional binding partner of delta-retroviral integrase. Nucleic Acids Res. 44(1), 364-376.

Manolio, T.A., 2010. Genomewide association studies and assessment of the risk of disease. N. Engl. J. Med. 363(2), 166-176.

Maples, B.K., Gravel, S., Kenny, E.E., Bustamante, C.D., 2013. RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. Am. J. Hum. Genet. 93(2), 278-288.

Marchand, E., Prat, C., Jeannin, C., Lafont, E., Bergmann, T., Flusin, O., Rizzi, J., Roux, N., Busso, V., Deniau, J., Noel, H., Vaillant, V., Leparc-Goffart, I., Six, C., Paty, M.C., 2013. Autochthonous case of dengue in France, October 2013. Euro Surveill. Bull. Eur. sur les Mal. Transm. = Eur. Commun. Dis. Bull. 18(50), 20661.

Marigorta, U.M., Navarro, A., 2013. High Trans-ethnic Replicability of GWAS Results Implies Common Causal Variants. PLoS Genet. 9(6).

Martinez-Marignac, V.L., Valladares, A., Cameron, E., Chan, A., Perera, A., Globus-Goldberg, R., Wacher, N., Kumate, J., McKeigue, P., O'Donnell, D., Shriver, M.D., Cruz, M., Parra, E.J., 2007. Admixture in Mexico City: Implications for admixture mapping of Type 2 diabetes genetic risk factors. Hum. Genet. 120(6), 807-819.

McEvoy, B.P., Powell, J.E., Goddard, M.E., Visscher, P.M., 2011. Human population dispersal "'Out of Africa'" estimated from linkage disequilibrium and allele frequencies of SNPs. Genome Res. 21(6), 821-829.

McKeigue, P.M., 1998. Mapping Genes That Underlie Ethnic Differences in Disease Risk: Methods for Detecting Linkage in Admixed Populations, by Conditioning on Parental Admixture. Am. J. Hum. Genet. 63(1), 241-251.

Mejias, A., Ramilo, O., 2014. Transcriptional profiling in infectious diseases: Ready for prime time? J. Infect. 68(SUPPL1), S94-S99.

Mercado, E.S., Espino, F.E., Perez, M.L.M., Bilar, J.M., Bajaro, J.D.P., Huy, N.T., Baello, B.Q., Kikuchi, M., Hirayama, K., 2015. HLA-A*33:01 as protective allele for severe dengue in a population of Filipino children. PLoS One 10(2), e0115619.

Mersha, T.B., 2015. Mapping asthma-associated variants in admixed populations. Front. Genet. 6(SEP), 1-21.

Miao, C.M., He, K., Li, P.Z., Liu, Z.J., Zhu, X.W., Ou, Z.B., Ruan, X.Z., Gong, J.P., Liu, C.A., 2016. LXRalpha represses LPS-induced inflammatory responses by competing with IRF3 for GRIP1 in Kupffer cells. Int. Immunopharmacol. 35, 272-279.

Min, G., Kim Kemper, J., Kemper, B., 2002. Glucocorticoid receptor-interacting protein 1 mediates ligand-independent nuclear translocation and activation of constitutive androstane receptor in vivo. J. Biol. Chem. 277(29), 26356-26363.

Mohr, S.E., Smith, J.A., Shamu, C.E., Neumüller, R.A., Perrimon, N., 2014. RNAi screening comes of age: Improved techniques and complementary approaches. Nat. Rev. Mol. Cell Biol. 15(9), 591-600.

Mohsin, S.N., Mahmood, S., Amar, A., Ghafoor, F., Raza, S.M., Saleem, M., 2015. Association of FcgammaRIIa Polymorphism with Clinical Outcome of Dengue Infection: First Insight from Pakistan. Am. J. Trop. Med. Hyg. 93(4), 691-696.

Montoya, M., Gresh, L., Mercado, J.C., Williams, K.L., Jose, M., Gutierrez, G., Kuan, G., Gordon, A., Balmaseda, A., Harris, E., 2013a. Symptomatic Versus Inapparent Outcome in Repeat Dengue Virus Infections Is Influenced by the Time Interval between Infections and Study Year. PLoS Negl Trop Dis 7(8), 1-10.

Montoya, M., Gresh, L., Mercado, J.C., Williams, K.L., Jose, M., Gutierrez, G., Kuan, G., Gordon, A., Balmaseda, A., Harris, E., 2013b. Symptomatic Versus Inapparent Outcome in Repeat Dengue Virus Infections Is Influenced by the Time Interval between Infections and Study Year 7(8), 1-10.

Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., Houstis, N., Daly, M.J., Patterson, N., Mesirov, J.P., Golub, T.R., Tamayo, P., Spiegelman, B., Lander, E.S., Hirschhorn, J.N., Altshuler, D., Groop, L.C., 2003. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. 34(3), 267.

Nei, M., Li, W.H., 1973. Linkage disequilibrium in subdivided populations. Genetics 75(1), 213-219.

Nejentsev, S., Walker, N., Riches, D., Egholm, M., Todd, J. a, 2009. Rare Variants of IFIH1, a Gene Implicated in Antiviral Responses, Protect Against Type 1 Diabetes. Science (80-.). 333(April), 387-389.

Neumayr, A., Munoz, J., Schunk, M., Bottieau, E., Cramer, J., Calleri, G., Lopez-Velez, R., Angheben, A., Zoller, T., Visser, L., Serre-Delcor, N., Genton, B., Castelli, F., Van Esbroeck, M., Matteelli, A., Rochat, L., Sulleiro, E., Kurth, F., Gobbi, F., Norman, F., Torta, I., Clerinx, J., Poluda, D., Martinez, M., Calvo-Cano, A., Sanchez-Seco, M.P., Wilder-Smith, A., Hatz, C., Franco, L., 2017. Sentinel surveillance of imported dengue via travellers to Europe 2012 to 2014: TropNet data from the DengueTools Research Initiative. Euro Surveill. Bull. Eur. sur les Mal. Transm. = Eur. Commun. Dis. Bull. 22(1).

Nguyen, N.M., Tran, C.N.B., Phung, L.K., Duong, K.T.H., Huynh, H. le A., Farrar, J., Nguyen, Q.T.H., Tran, H.T., Nguyen, C.V.V., Merson, L., Hoang, L.T., Hibberd, M.L., Aw, P.P.K., Wilm, A., Nagarajan, N., Nguyen, D.T., Pham, M.P., Nguyen, T.T., Javanbakht, H., Klumpp, K., Hammond, J., Petric, R., Wolbers, M., Nguyen, C.T., Simmons, C.P., 2013. A randomized, double-blind placebo controlled trial of balapiravir, a polymerase inhibitor, in adult dengue patients. J. Infect. Dis. 207(9), 1442-1450.

Nguyen, T.P.L., Kikuchi, M., Vu, T.Q.H., Do, Q.H., Tran, T.T., Vo, D.T., Ha, M.T., Vo, V.T., Cao, T.P.N., Tran, V.D., Oyama, T., Morita, K., Yasunami, M., Hirayama, K., 2008. Protective and enhancing HLA alleles, HLA-DRB1*0901 and HLA-A*24, for severe forms of dengue virus infection, dengue hemorrhagic fever and dengue shock syndrome. PLoS Negl. Trop. Dis. 2(10), e304.

Nicolae, D.L., Gamazon, E., Zhang, W., Duan, S., Eileen Dolan, M., Cox, N.J., 2010.

Trait-associated SNPs are more likely to be eQTLs: Annotation to enhance discovery from GWAS. PLoS Genet. 6(4).

Noecker, C.A., Amaya-Larios, I.Y., Galeana-Hernandez, M., Ramos-Castaneda, J., Martinez-Vega, R.A., 2014. Contrasting associations of polymorphisms in FcgammaRIIa and DC-SIGN with the clinical presentation of dengue infection in a Mexican population. Acta Trop. 138, 15-22.

Ntzani, E.E., Liberopoulos, G., Manolio, T.A., Ioannidis, J.P.A., 2012. Consistency of genome-wide associations across major ancestral groups. Hum. Genet. 131(7), 1057-1071.

Oliveira, M., Lert-itthiporn, W., Cavadas, B., Chuansumrit, A., Anunciac, O., Casademont, I., Koeth, F., Penova, M., Tangnararatchakit, K., Khor, C.C., Paul, R., Malasit, P., Matsuda, F., Simon-lori, E., 2018. Joint ancestry and association test indicate two distinct pathogenic pathways involved in classical dengue fever and dengue shock syndrome. PLoS Negl. Trop. Dis. 1-20.

Oliveira, L.F. de, Lima, C.P.S. de, Azevedo, R. do S.S., Mendonça, D.S.F. de, Rodrigues, S.G., Carvalho, V.L., Pinto, E.V., Maia, A.L., Maia, M.H.T., Vasconcelos, J.M., Silva, A.L.S. da, Nunes, M.R.T., Sena, L., Vasconcelos, P.F., Santos, E.J.M. dos, 2014. Polymorphism of DC-SIGN (*CD209*) Promoter in Association with Clinical Symptoms of Dengue Fever. Viral Immunol. 27(5), 245-249.

Olkowski, S., Forshey, B.M., Morrison, A.C., Rocha, C., Vilcarromero, S., Halsey, E.S., Kochel, T.J., Scott, T.W., Stoddard, S.T., 2013. Reduced Risk of Disease During Postsecondary Dengue Virus Infections. J. Infect. Dis. 208(6), 1026-1033.

Ooi, E.E., Goh, K.T., Gubler, D.J., 2006. Dengue prevention and 35 years of vector control in Singapore. Emerg. Infect. Dis. 12(6), 887-893.

PAHO, 2007. Pan American Health Organization. Prevención y control del dengue en las Américas. Resolución CSP27.R15. 27. a Conf. Sanit. Panam. CSP27 R15 (Esp.)(1-5 de octubre).

Park, J.-H., Gail, M.H., Weinberg, C.R., Carroll, R.J., Chung, C.C., Wang, Z., Chanock, S.J., Fraumeni, J.F., Chatterjee, N., 2011. Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. Proc. Natl. Acad. Sci. 108(44), 18026-18031.

Parreira, R., Conceição, C., Centeno-Lima, S., Marques, N., Saraiva, J. da C., Abreu, C., Sá, L., Sarmento, A., Atouguia, J., Moneti, V., Azevedo, T., Nina, J., Mansinho, K., Antunes, A., Teodósio, R., Nazareth, T., Seixas, J., 2014. Angola's 2013 dengue outbreak: Clinical, laboratory and molecular analyses of cases from four Portuguese institutions. J. Infect. Dev. Ctries. 8(9), 1210-1215.

Parreira, R., Sousa, C.A., 2015. Dengue fever in Europe : could there be an epidemic in the future? Expert Rev. Anti. Infect. Ther. 13(1), 29–40.

Patterson, N., Hattangadi, N., Lane, B., Lohmueller, K.E., Hafler, D.A., Oksenberg, J.R., Hauser, S.L., Smith, M.W., O'Brien, S.J., Altshuler, D., Daly, M.J., Reich, D., 2004. Methods for High-Density Admixture Mapping of Disease Genes. Am. J. Hum. Genet. 74(5), 979–1000.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., Reich, D., 2012. Ancient admixture in human history. Genetics 192(3), 1065–1093.

Perera, R., Kuhn, R.J., 2008. Structural proteomics of dengue virus. Curr. Opin. Microbiol. 11(4), 369–377.

Perez, A.B., Sierra, B., Garcia, G., Aguirre, E., Babel, N., Alvarez, M., Sanchez, L., Valdes, L., Volk, H.D., Guzman, M.G., 2010. Tumor necrosis factor-alpha, transforming growth factor-beta1, and interleukin-10 gene polymorphisms: implication in protection or susceptibility to dengue hemorrhagic fever. Hum. Immunol. 71(11), 1135–1140.

Pierron, D., Heiske, M., Razafindrazaka, H., Rakoto, I., Rabetokotany, N., Ravololomanga, B., Rakotozafy, L.M.-A., Rakotomalala, M.M., Razafiarivony, M., Rasoarifetra, B., Raharijesy, M.A., Razafindralambo, L., Ramilisonina, Fanony, F., Lejamble, S., Thomas, O., Mohamed Abdallah, A., Rocher, C., Arachiche, A., Tonaso, L., Pereda-loth, V., Schiavinato, S., Brucato, N., Ricaut, F.-X., Kusuma, P., Sudoyo, H., Ni, S., Boland, A., Deleuze, J.-F., Beaujard, P., Grange, P., Adelaar, S., Stoneking, M., Rakotoarisoa, J.-A., Radimilahy, C., Letellier, T., 2017. Genomic landscape of human diversity across Madagascar. Proc. Natl. Acad. Sci. 114(32), E6498–E6506.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38(8), 904–909.

Puschnik, A.S., Majzoub, K., Ooi, Y.S., Carette, J.E., 2017. A CRISPR toolbox to study virus-host interactions. Nat. Rev. Microbiol. 15(6), 351–364.

Quach, H., Rotival, M., Pothlichet, J., Loh, Y.H.E., Dannemann, M., Zidane, N., Laval, G., Patin, E., Harmant, C., Lopez, M., Deschamps, M., Naffakh, N., Duffy, D., Coen, A., Leroux-Roels, G., Clément, F., Boland, A., Deleuze, J.F., Kelso, J., Albert, M.L., Quintana-Murci, L., 2016. Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. Cell 167(3), 643–656.e17.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome

engineering using the CRISPR-Cas9 system. Nat. Protoc. 8(11), 2281-2308.

Rana, T.M., 2007. Illuminating the silence: Understanding the structure and function of small RNAs. Nat. Rev. Mol. Cell Biol. 8(1), 23-36.

Ranciaro, A., Campbell, M.C., Hirbo, J.B., Ko, W.Y., Froment, A., Anagnostou, P., Kotze, M.J., Ibrahim, M., Nyambo, T., Omar, S.A., Tishkoff, S.A., 2014. Genetic origins of lactase persistence and the spread of pastoralism in Africa. Am. J. Hum. Genet. 94(4), 496-510.

Rattanaburee, T., Junking, M., Panya, A., Sawasdee, N., Songprakhon, P., Suttitheptumrong, A., Limjindaporn, T., Haegeman, G., Yenchitsomanus, P.T., 2015. Inhibition of dengue virus production and cytokine/chemokine expression by ribavirin and compound A. Antiviral Res. 124, 83-92.

Reily, M.M., Pantoja, C., Hu, X., Chinenov, Y., Rogatsky, I., 2006. The GRIP1:IRF3 interaction as a target for glucocorticoid receptor-mediated immunosuppression. EMBO J. 25(1), 108-17.

Reiner, R.C., Achee, N., Barrera, R., Burkot, T.R., Chadee, D.D., Devine, G.J., Endy, T., Gubler, D., Hombach, J., Kleinschmidt, I., Lenhart, A., Lindsay, S.W., Longini, I., Mondy, M., Morrison, A.C., Perkins, T.A., Vazquez-Prokopec, G., Reiter, P., Ritchie, S.A., Smith, D.L., Strickman, D., Scott, T.W., 2016. Quantifying the Epidemiological Impact of Vector Control on Dengue. PLoS Negl. Trop. Dis. 10(5), 1-11.

Report, W.H.O., 2017. Dengue vaccine: WHO position paper, July 2016 - recommendations 6-7.

Ritchie, S.A., Devine, G.J., 2013. Confusion, knock-down and kill of Aedes aegypti using metofluthrin in domestic settings: a powerful tool to prevent dengue transmission? Parasit. Vectors 6(1), 262.

Rito, T., Richards, M.B., Fernandes, V., Alshamali, F., Cerny, V., Pereira, L., Soares, P., 2013. The first modern human dispersals across Africa. PLoS One 8(11), 1-16.

Roberts, S.K., Cooksley, G., Dore, G.J., Robson, R., Shaw, D., Berns, H., Hill, G., Klumpp, K., Najera, I., Washington, C., 2008. Robust antiviral activity of R1626, a novel nucleoside analog: a randomized, placebo-controlled study in patients with chronic hepatitis C. Hepatology 48(2), 398-406.

Rosenberg, N.A., Huang, L., Jewett, E.M., Szpiech, Z.A., Jankovic, I., Boehnke, M., 2010. Genome-wide association studies in diverse populations. Nat. Rev. Genet. 11(5), 356-366.

Roszer, T., Menendez-Gutierrez, M.P., Cedenilla, M., Ricote, M., 2013. Retinoid X receptors in macrophage biology. Trends Endocrinol. Metab. 24(9), 460-468.

Rothman, A.L., 2011. Immunity to dengue virus : a tale of original antigenic sin and tropical cytokine storms. Nat. Rev. Immunol. 11(8), 532–543.

Roulin, P.S., Lotzerich, M., Torta, F., Tanner, L.B., van Kuppeveld, F.J.M., Wenk, M.R., Greber, U.F., 2014. Rhinovirus uses a phosphatidylinositol 4-phosphate/cholesterol counter-current for the formation of replication compartments at the ER–Golgi interface. Cell Host Microbe 16(5), 677–690.

Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E.H., McCarroll, S.A., Gaudet, R., Schaffner, S.F., Lander, E.S., Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., Pasternak, S., Wheeler, D.A., Willis, T.D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., Hu, W., Li, C., Lin, W., Liu, S., Pan, H., Tang, X., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q., Zhao, H., Zhao, H., Zhou, J., Gabriel, S.B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R.C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Altshuler, D., Shen, Y., Yao, Z., Huang, W., Chu, X., He, Y., Jin, L., Liu, Y., Shen, Y., Sun, W., Wang, H., Wang, Y., Wang, Y., Xiong, X., Xu, L., Waye, M.M.Y., Tsui, S.K.W., Xue, H., Wong, J.T.F., Galver, L.M., Fan, J.B., Gunderson, K., Murray, S.S., Oliphant, A.R., Chee, M.S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J.F., Phillips, M.S., Roumy, S., Sallée, C., Verner, A., Hudson, T.J., Kwok, P.Y., Cai, D., Koboldt, D.C., Miller, R.D., Pawlikowska, L., Taillon-Miller, P., Xiao, M., Tsui, L.C., Mak, W., You, Q.S., Tam, P.K.H., Nakamura, Y., Kawaguchi, T., Kitamoto, T., Morizono, T., Nagashima, A., Ohnishi, Y., Sekine, A., Tanaka, T., Tsunoda, T., Deloukas, P., Bird, C.P., Delgado, M., Dermitzakis, E.T., Gwilliam, R., Hunt, S., Morrison, J., Powell, D., Stranger, B.E., Whittaker, P., Bentley, D.R., Daly, M.J., De Bakker, P.I.W., Barrett, J., Chretien, Y.R., Maller, J., McCarroll, S., Patterson, N., Pe'Er, I., Price, A., Purcell, S., Richter, D.J., Saxena, R., Sham, P.C., Stein, L.D., Krishnan, L., Smith, A.V., Tello-Ruiz, M.K., Thorisson, G.A., Chakravarti, A., Chen, P.E., Cutler, D.J., Kashuk, C.S., Lin, S., Abecasis, G.R., Guan, W., Li, Y., Munro, H.M., Qin, Z.S., Thomas, D.J., McVean, G., Auton, A., Bottolo, L., Cardin, N., Eyheramendy, S., Freeman, C., Marchini, J., Myers, S., Spencer, C., Stephens, M., Donnelly, P., Cardon, L.R., Clarke, G., Evans, D.M., Morris, A.P., Weir, B.S., Johnson, T.A., Mullikin, J.C., Sherry, S.T., Feolo, M., Skol, A., Zhang, H., Matsuda, I., Fukushima, Y., MacEr, D.R., Suda, E., Rotimi, C.N., Adebamowo, C.A., Ajayi, I., Aniagwu, T., Marshall, P.A., Nkwodimmah, C., Royal, C.D.M., Leppert, M.F., Dixon, M., Peiffer, A., Qiu, R., Kent, A., Kato, K., Niikawa, N., Adewole, I.F., Knoppers, B.M., Foster, M.W.,

Clayton, E.W., Watkin, J., Muzny, D., Nazareth, L., Sodergren, E., Weinstock, G.M., Yakub, I., Birren, B.W., Wilson, R.K., Fulton, L.L., Rogers, J., Burton, J., Carter, N.P., Clee, C.M., Griffiths, M., Jones, M.C., McLay, K., Plumb, R.W., Ross, M.T., Sims, S.K., Willey, D.L., Chen, Z., Han, H., Kang, L., Godbout, M., Wallenburg, J.C., L'Archevêque, P., Bellemare, G., Saeki, K., Wang, H., An, D., Fu, H., Li, Q., Wang, Z., Wang, R., Holden, A.L., Brooks, L.D., McEwen, J.E., Guyer, M.S., Wang, V.O., Peterson, J.L., Shi, M., Spiegel, J., Sung, L.M., Zacharia, L.F., Collins, F.S., Kennedy, K., Jamieson, R., Stewart, J., 2007. Genome-wide detection and characterization of positive selection in human populations. Nature 449(7164), 913–918.

Sakuntabhai, A., Turbpaiboon, C., Casadémont, I., Chuansumrit, A., Lowhnoo, T., Kajaste-Rudnitski, A., Kalayanarooj, S.M., Tangnararatchakit, K., Tangthawornchaikul, N., Vasanawathana, S., Chaiyaratana, W., Yenchitsomanus, P., Suriyaphol, P., Avirutnan, P., Chokephaibulkit, K., Matsuda, F., Yoksan, S., Jacob, Y., Lathrop, G.M., Malasit, P., Desprès, P., Julier, C., 2005. A variant in the CD209 promoter is associated with severity of dengue disease. Nat. Genet. 37(5), 507–513.

Sam, S., Teoh, B., Chinna, K., Abubakar, S., 2015. High Producing Tumor Necrosis Factor Alpha Gene Alleles in Protection against Severe Manifestations of Dengue. Int. J. Med. Sci. 12(2), 177–186.

Sanchez-Leyva, M., Sanchez-Zazueta, J.G., Osuna-Ramos, J.F., Rendon-Aguilar, H., Felix-Espinoza, R., Becerra-Loaiza, D.S., Sanchez-Garcia, D.C., Romero-Quintana, J.G., Castillo Ureta, H., Velarde-Rodriguez, I., Velarde-Felix, J.S., 2017. Genetic Polymorphisms of Tumor Necrosis Factor Alpha and Susceptibility to Dengue Virus Infection in a Mexican Population. Viral Immunol. 30(8), 615–621.

Sander, J.D., Joung, J.K., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32(4), 347–350.

Schmidt-Chanasit, J., Haditsch, M., Schöneberg, I., Günther, S., Stark, K., Frank, C., 2010. Dengue virus infection in a traveller returning from croatia to Germany. Eurosurveillance 15(40), 2–3.

Serbus, L.R., Casper-Lindley, C., Landmann, F., Sullivan, W., 2008. The Genetics and Cell Biology of *Wolbachia* –Host Interactions. Annu. Rev. Genet. 42(1), 683–707.

Shekhar, K.C., Huat, O.L., 1992. Epidemiology of dengue/dengue hemorrhagic fever in Malaysia--a retrospective epidemiological study. 1973–1987. Part II: Dengue fever (DF). Asia-Pacific J. public Heal. 6(3), 126–133.

Shepard, D.S., Undurraga, E.A., Halasa, Y.A., Stanaway, J.D., 2016. The global economic burden of dengue: a systematic analysis. Lancet Infect. Dis. 16(8), 935–

941.

Shetty, P.B., Tang, H., Feng, T., Tayo, B., Morrison, A.C., Kardia, S.L.R., Hanis, C.L., Arnett, D.K., Hunt, S.C., Boerwinkle, E., Rao, D.C., Cooper, R.S., Risch, N., Zhu, X., 2015. Variants for HDL-C, LDL-C, and triglycerides identified from admixture mapping and fine-mapping analysis in African American families. Circ. Cardiovasc. Genet. 8(1), 106-113.

Shriner, D., Adeyemo, A., Ramos, E., Chen, G., Rotimi, C.N., 2011. Mapping of disease-associated variants in admixed populations. Genome Biol. 12(5), 1-8.

Sierra, B., Alegre, R., Perez, A.B., Garcia, G., Sturn-Ramirez, K., Obasanjo, O., Aguirre, E., Alvarez, M., Rodriguez-Roche, R., Valdes, L., Kanki, P., Guzman, M.G., 2007a. 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. 68(6), 531-540.

Sierra, B., Garcia, G., Perez, A.B., Morier, L., Alvarez, M., Kouri, G., Guzman, M.G., 2006. Ethnicity and difference in dengue virus-specific memory T cell responses in Cuban individuals. Viral Immunol. 19(4), 662-668.

Sierra, B., Kouri, G., Guzman, M.G., 2007b. Race: a risk factor for dengue hemorrhagic fever. Arch. Virol. 152(3), 533-542.

Sierra, B., Triska, P., Soares, P., Garcia, G., Perez, A.B., Aguirre, E., Oliveira, M., Cavadas, B., Alvarez, M., Ruiz, D., Samuels, D.C., Sakuntabhai, A., Pereira, L., Guzman, M.G., 2017. OSBPL10 , RXRA and lipid metabolism confer African-ancestry protection against dengue haemorrhagic fever in admixed Cubans. PLoS Pathog. 13(2), 1-21.

Silva, L.K., Blanton, R.E., Parrado, A.R., Melo, P.S., Morato, V.G., Reis, E.A.G., Dias, J.P., Castro, J.M., Vasconcelos, P.F.C., Goddard, K.A.B., Barreto, L., Reis, M.G., Glo, M., 2010. Dengue hemorrhagic fever is associated with polymorphisms in JAK1. Eur. J. Hum. Genet. 18(September 2009), 1221-1227.

Simmons, C.P., Popper, S., Dolocek, C., Nguyen, T., Chau, B., Griffiths, M., Thi, N., Dung, P., Long, T.H., Hoang, D.M., Chau, N.V., Thi, L., Thao, T., Hien, T.T., Relman, D.A., Farrar, J., 2007. Patterns of Host Genome – Wide Gene Transcript Abundance in the Peripheral Blood of Patients with Acute Dengue Hemorrhagic Fever. J. Infect. Dis. 195(8), 1097-1107.

Simon-Loriere, E., Duong, V., Tawfik, A., Ung, S., Ly, S., Casadémont, I., Prot, M., Courtejoie, N., Bleakley, K., Buchy, P., Tarantola, A., Dussart, P., Cantaert, T., Sakuntabhai, A., 2017a. Increased adaptive immune responses and proper feedback regulation protect against clinical dengue. Sci. Transl. Med. 9(405).

Simon-Loriere, E., Duong, V., Tawfik, A., Ung, S., Ly, S., Casadémont, I., Prot, M., Courtejoie, N., Bleakley, K., Buchy, P., Tarantola, A., Dussart, P., Cantaert, T., Sakuntabhai, A., 2017b. Increased adaptive immune responses and proper feedback regulation protect against clinical dengue. Sci. Transl. Med. 9(405).

Simon-Loriere, E., Lin, R.J., Kalayanarooj, S.M., Chuansumrit, A., Casademont, I., Lin, S.Y., Yu, H.P., Lert-Itthiporn, W., Chaiyaratana, W., Tangthawornchaikul, N., Tangnararatchakit, K., Vasanawathana, S., Chang, B.L., Suriyaphol, P., Yoksan, S., Malasit, P., Despres, P., Paul, R., Lin, Y.L., Sakuntabhai, A., 2015. High anti-dengue virus activity of the OAS gene family is associated with increased severity of dengue. J. Infect. Dis. 212(12), 2011–2020.

Siqueira, J.B., Martelli, C.M.T., Coelho, G.E., Simplicio, A.C.D.R., Hatch, D.L., 2005. Dengue and dengue hemorrhagic fever, Brazil, 1981–2002. Emerg. Infect. Dis. 11(1), 48–53.

Soares, P., Alshamali, F., Pereira, J.B., Fernandes, V., Silva, N.M., Afonso, C., Costa, M.D., Musilová, E., MacAulay, V., Richards, M.B., Černý, V., Pereira, L., 2012. The expansion of mtDNA haplogroup L3 within and out of Africa. Mol. Biol. Evol. 29(3), 915–927.

Sommer, L.M., Cho, H., Choudhary, M., Seeling, J.M., 2015. Evolutionary analysis of the B56 gene family of PP2A regulatory subunits. Int. J. Mol. Sci. 16(5), 10134–10157.

Sotiriou, C., Piccart, M.J., 2007. Taking gene-expression profiling to the clinic: When will molecular signatures become relevant to patient care? Nat. Rev. Cancer 7(7), 545–553.

Soto-Acosta, R., Bautista-Carbajal, P., Cervantes-Salazar, M., Angel-Ambrocio, A.H., del Angel, R.M., 2017. DENV up-regulates the HMG-CoA reductase activity through the impairment of AMPK phosphorylation: A potential antiviral target. PLOS Pathog. 13(4), e1006257.

Stephens, H.A.F., Klaythong, R., Sirikong, M., Vaughn, D.W., Green, S., Kalayanarooj, S., Endy, T.P., Libraty, D.H., Nisalak, A., Innis, B.L., Rothman, A.L., Ennis, F.A., Chandanayingyong, D., 2002. 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 60(4), 309–318.

Stoler, J., al Dashti, R., Anto, F., Fobil, J.N., Awandare, G.A., 2014. Deconstructing "malaria": West Africa as the next front for dengue fever surveillance and control. Acta Trop. 134(1), 58–65.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A.,

Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., Mesirov, J.P., 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. 102(43), 15545–15550.

Sun, P., Garcia, J., Comach, G., Vahey, M.T., Wang, Z., Forshey, B.M., Morrison, A.C., Sierra, G., Bazan, I., Rocha, C., Vilcarromero, S., Blair, P.J., Scott, T.W., Camacho, D.E., Ockenhouse, C.F., Halsey, E.S., Kochel, T.J., 2013. Sequential waves of gene expression in patients with clinically defined dengue illnesses reveal subtle disease phases and predict disease severity. PLoS Negl. Trop. Dis. 7(7), e2298.

Tak, Y.G., Farnham, P.J., 2015. Making sense of GWAS: Using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. Epigenetics and Chromatin 8(1), 1–18.

Takiar, R., Lutsey, P.L., Zhao, D., Guallar, E., Schneider, A.L.C., Grams, M.E., Appel, L.J., Selvin, E., Michos, E.D., 2015. The associations of 25-hydroxyvitamin D levels, vitamin D binding protein gene polymorphisms, and race with risk of incident fracture-related hospitalization: Twenty-year follow-up in a bi-ethnic cohort (the ARIC Study). Bone 78, 94–101.

Tam, D.T.H., Ngoc, T. V, Tien, N.T.H., Kieu, N.T.T., Thuy, T.T.T., Thanh, L.T.C., Tam, C.T., Truong, N.T., Dung, N.T., Qui, P.T., Hien, T.T., Farrar, J.J., Simmons, C.P., Wolbers, M., Wills, B.A., 2012. Effects of short-course oral corticosteroid therapy in early dengue infection in Vietnamese patients: a randomized, placebo-controlled trial. Clin. Infect. Dis. 55(9), 1216–1224.

Timmann, C., Thye, T., Vens, M., Evans, J., May, J., Ehmen, C., Sievertsen, J., Muntau, B., Ruge, G., Loag, W., Ansong, D., Antwi, S., Asafo-Adjei, E., Nguah, S.B., Kwakye, K.O., Akoto, A.O.Y., Sylverken, J., Brendel, M., Schuldt, K., Loley, C., Franke, A., Meyer, C.G., Agbenyega, T., Ziegler, A., Horstmann, R.D., 2012. Genome-wide association study indicates two novel resistance loci for severe malaria. Nature 489(7416), 443–446.

Tishkoff, S.A., Reed, F.A., Ranciaro, A., Voight, B.F., Babbitt, C.C., Silverman, J.S., Powell, K., Mortensen, H.M., Hirbo, J.B., Osman, M., Ibrahim, M., Omar, S.A., Lema, G., Nyambo, T.B., Ghori, J., Bumpstead, S., Pritchard, J.K., Wray, G.A., Deloukas, P., 2007. Convergent adaptation of human lactase persistence in Africa and Europe. Nat. Genet. 39(1), 31–40.

Torgerson, D.G., Gignoux, C.R., Galanter, J.M., Drake, K.A., Roth, L.A., Eng, C., Huntsman, S., Torres, R., Avila, P.C., Chapela, R., Ford, J.G., Rodríguez-Santana, J.R., Rodríguez-Cintrón, W., Hernandez, R.D., Burchard, E.G., 2012. Case-control admixture mapping in Latino populations enriches for known asthma-associated

genes. J. Allergy Clin. Immunol. 130(1), 76-82.

Underhill, P.A., Kivisild, T., 2007. Use of Y Chromosome and Mitochondrial DNA Population Structure in Tracing Human Migrations. Annu. Rev. Genet. 41(1), 539–564.

Vannice, K., Chocarro, L., Pfleiderer, M., Bellah, A., Ward, M., Yoon, I., Hombach, J., 2017. The value of multi-country joint regulatory reviews : The experience of a WHO joint technical consultation on the CYD-TDV (Dengvaxia) dossier. Vaccine 35(43), 5731-5733.

Vasilakis, N., Cardosa, J., Hanley, K.A., Holmes, E.C., Weaver, S.C., 2011. Fever from the forest : prospects for the continued emergence of sylvatic dengue virus and its impact on public health. Nat. Rev. Microbiol. 9(7), 532-541.

Vasilakis, N., Weaver, S.C., 2008. The History and Evolution of Human Dengue Emergence, Advances in Virus Research. Elsevier Inc.

Vejbaesya, S., Luangtrakool, P., Luangtrakool, K., Kalayanarooj, S., Vaughn, D.W., Endy, T.P., Mammen, M.P., Green, S., Libraty, D.H., Ennis, F.A., Rothman, A.L., Stephens, H.A.F., 2009. *TNF* and *LTA* Gene, Allele, and Extended HLA Haplotype Associations with Severe Dengue Virus Infection in Ethnic Thais. J. Infect. Dis. 199(10), 1442-1448.

Vina, M.A.F., Hollenbach, J.A., Lyke, K.E., Sztein, M.B., Maiers, M., Klitz, W., Cano, P., Mack, S., Single, R., Brautbar, C., Israel, S., Raimondi, E., Khoriaty, E., 2012. Tracking human migrations by the analysis of the distribution of HLA alleles , lineages and haplotypes in closed and open populations. Phil. Trans. R. Soc. B 367, 820-829.

Visel, A., 2004. GenePaint.org: an atlas of gene expression patterns in the mouse embryo. Nucleic Acids Res. 32(90001), 552D-556.

Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., Mccarthy, M.I., Brown, M.A., Yang, J., 2017. 10 Years of GWAS Discovery : Biology , Function , and Translation. Am. J. Hum. Genet. 101(1), 5-22.

Voight, B.F., Kudaravalli, S., Wen, X., Pritchard, J.K., 2006. A map of recent positive selection in the human genome. PLoS Biol. 4(3), 0446-0458.

Wahala, W.M.P.B., Silva, A.M. de, 2011. The human antibody response to dengue virus infection. Viruses 3(12), 2374-2395.

Wang, E., Ni, H., Xu, R., Barrett, A.D.T., Watowich, S.J., Gubler, D.J., Weaver, S.C., 2000. Evolutionary Relationships of Endemic / Epidemic and Sylvatic Dengue Viruses. J. Virol. 74(7), 3227-3234.

Wang, L., Chen, R.F., Liu, J.W., Lee, I.K., Lee, C.P., Kuo, H.C., Huang, S.K., Yang, K.D.,

2011. DC-SIGN (CD209) promoter -336 A/G polymorphism is associated with dengue hemorrhagic fever and correlated to DC-SIGN expression and immune augmentation. PLoS Negl. Trop. Dis. 5(1), e934.

Watanabe, S., Rathore, A.P.S., Sung, C., Lu, F., Khoo, Y.M., Connolly, J., Low, J., Ooi, E.E., Lee, H.S., Vasudevan, S.G., 2012. Dose- and schedule-dependent protective efficacy of celgosivir in a lethal mouse model for dengue virus infection informs dosing regimen for a proof of concept clinical trial. Antiviral Res. 96(1), 32-35.

Welsch, S., Miller, S., Romero-brey, I., Merz, A., Bleck, C.K.E., Walther, P., Fuller, S.D., Antony, C., Krijnse-locker, J., Bartenschlager, R., 2009. Article Composition and Three-Dimensional Architecture of the Dengue Virus Replication and Assembly Sites. Cell Host Microbe 5(4), 365-375.

Werren, J.H., Baldo, L., Clark, M.E., 2008. Wolbachia: Master manipulators of invertebrate biology. Nat. Rev. Microbiol. 6(10), 741-751.

Whitehorn, J., Chau, T.N.B., Nguyet, N.M., Kien, D.T.H., Quyen, N.T.H., Trung, D.T., Pang, J., Wills, B., Van Vinh Chau, N., Farrar, J., Hibberd, M.L., Khor, C.C., Simmons, C.P., 2013. Genetic Variants of MICB and PLCE1 and Associations with Non-Severe Dengue. PLoS One 8(3), 1-7.

Whitehorn, J., Van Vinh Chau, N., Truong, N.T., Tai, L.T.H., Van Hao, N., Hien, T.T., Wolbers, M., Merson, L., Dung, N.T.P., Peeling, R., Simmons, C., Wills, B., Farrar, J., 2012. Lovastatin for adult patients with dengue: protocol for a randomised controlled trial. Trials 13, 203.

WHO, 2017. GACVS Statement on Dengvaxia ® (CYD-TDV) [WWW Document]. December 7, 2017. URL http://www.who.int/vaccine_safety/committee/GACVS-StatementonDengvaxia-CYD-TDV/en/

WHO, 2009. Dengue guidelines for diagnosis, treatment, prevention and control. 3rd editio(3rd edition), 1-147.

WHO/EMRO, 2005. World Health Organization, Regional Office for the Eastern Mediterranean, Division of Communicable Disease Control. Newsletter 6:7-8.

WHO/SEARO, 2008. Concrete measure key in controlling dengue in South East Asia. Press Release SEA/PR/1479.

Wilder-Smith, A., Gubler, D.J., 2008. Geographic expansion of dengue: the impact of international travel. Med. Clin. North Am. 92(6), 1377-90, x.

Wilder-Smith, A., Gubler, D.J., Weaver, S.C., Monath, T.P., Heymann, D.L., Scott, T.W., 2017. Epidemic arboviral diseases: priorities for research and public health. Lancet Infect. Dis. 17(3), e101-e106.

Wilder-Smith, A., Macary, P., 2014. Dengue: challenges for policy makers and vaccine

developers. Curr. Infect. Dis. Rep. 16(5), 404.

Williams, T.N., Wambua, S., Uyoga, S., Macharia, A., Mwacharo, J.K., Newton, C.R.J.C., Maitland, K., 2005. Both heterozygous and homozygous alpha+ thalassemias protect against severe and fatal Plasmodium falciparum malaria on the coast of Kenya. Blood 106(1), 368–371.

Winkler, C.A., Nelson, G.W., Smith, M.W., 2010. Admixture Mapping Comes of Age. Annu. Rev. Genomics Hum. Genet. 11(1), 65–89.

Xavier-Carvalho, C., Gibson, G., Brasil, P., Ferreira, R.X., De Souza Santos, R., Gonçalves Cruz, O., De Oliveira, S.A., De Sá Carvalho, M., Pacheco, A.G., Kubelka, C.F., Moraes, M.O., 2013. Single nucleotide polymorphisms in candidate genes and dengue severity in children: A case-control, functional and meta-analysis study. Infect. Genet. Evol. 20, 197–205.

Xia, N., Zhang, P., Fang, F., Wang, Z., Rothstein, M., Angulo, B., Chiang, R., Taylor, J., Reijo Pera, R.A., 2016. Transcriptional comparison of human induced and primary midbrain dopaminergic neurons. Sci. Rep. 6(September 2014), 1–9.

Yap, S.S.L., Nguyen-khuong, T., Rudd, P.M., Alonso, S., 2017. Dengue Virus Glycosylation: What Do We Know? Front. Microbiol. 8(July), 1–16.

Yu, I.-M., Zhang, W., Holdaway, H.A., Li, L., Kostyuchenko, V.A., Chipman, P.R., Kuhn, R.J., Rossmann, M.G., Chen, J., 2008. Structure of the immature dengue virus at low pH primes proteolytic maturation. Science 319(5871), 1834–1837.

Yue, W.-H., Wang, H.-F., Sun, L.-D., Tang, F.-L., Liu, Z.-H., Zhang, H.-X., Li, W.-Q., Zhang, Y.-L., Zhang, Y., Ma, C.-C., Du, B., Wang, L.-F., Ren, Y.-Q., Yang, Y.-F., Hu, X.-F., Wang, Y., Deng, W., Tan, L.-W., Tan, Y.-L., Chen, Q., Xu, G.-M., Yang, G.-G., Zuo, X., Yan, H., Ruan, Y.-Y., Lu, T.-L., Han, X., Ma, X.-H., Wang, Y., Cai, L.-W., Jin, C., Zhang, H.-Y., Yan, J., Mi, W.-F., Yin, X.-Y., Ma, W.-B., Liu, Q., Kang, L., Sun, W., Pan, C.-Y., Shuang, M., Yang, F.-D., Wang, C.-Y., Yang, J.-L., Li, K.-Q., Ma, X., Li, L.-J., Yu, X., Li, Q.-Z., Huang, X., Lv, L.-X., Li, T., Zhao, G.-P., Huang, W., Zhang, X.-J., Zhang, D., 2011. Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2. Nat. Genet. 43(12), 1228–1231.

Zhang, W., Xiao, D., Shan, L., Zhao, J., Mao, Q., Xia, H., 2017. Generation of apoptosis-resistant HEK293 cells with CRISPR/Cas mediated quadruple gene knockout for improved protein and virus production. Biotechnol. Bioeng. 114(11), 2539–2549.

Appendices

Appendix A – Supplementary information from paper I

Population genetics-informed meta-analyses in seven genes associated with risk to dengue fever disease

Marisa Oliveira, Diana P. Saraiva, Bruno Cavadas, Verónica Fernandes, Nicole Pedro, Isabelle Casademont, Fanny Koeth, Farida Alshamali, Nourdin Harich, Lotfi Cherni, Beatriz Sierra, Maria G. Guzman, Anavaj Sakuntabhai and Luisa Pereira

Table A.1 Raw population data for *PLCE1*–rs2274223 SNP.

Region	Population	Genotype			N	Frequency genotype			Allele count		Freq Alelle A	Freq Allele G	Hardy-Weinberg p - value
		AA	AG	GG		AA	AG	GG	Allele A	Allele G			
Europe	Portugal	21	16	6	43	0.488	0.372	0.140	58	28	0.674	0.326	0.312
Europe	Spain	38	53	16	107	0.355	0.495	0.150	129	85	0.603	0.397	0.841
Europe	Italy	42	47	19	108	0.389	0.435	0.176	131	85	0.606	0.394	0.420
Europe	France	15	11	2	28	0.536	0.393	0.071	41	15	0.732	0.268	1.000
Europe	UK	42	38	12	92	0.457	0.413	0.130	122	62	0.663	0.337	0.486
Europe	Netherland	241	199	57	497	0.485	0.400	0.115	681	313	0.685	0.315	0.114
Europe	Finland	57	37	5	99	0.576	0.374	0.051	151	47	0.763	0.237	1.000
Europe	Norway	10	6	2	18	0.556	0.333	0.111	26	10	0.722	0.278	0.560
Europe	Greece	7	10	3	20	0.350	0.500	0.150	24	16	0.600	0.400	1.000
Europe	Hungary	8	11	1	20	0.400	0.550	0.050	27	13	0.675	0.325	0.607
Europe	Russia	14	8	3	25	0.560	0.320	0.120	36	14	0.720	0.280	0.330
Europe	Georgia	9	9	2	20	0.450	0.450	0.100	27	13	0.675	0.325	1.000
North Africa	Morocco	21	18	9	48	0.438	0.375	0.188	60	36	0.625	0.375	0.217
North Africa	Algeria–Mozabite	10	14	5	29	0.345	0.483	0.172	34	24	0.586	0.414	1.000
North Africa	Tunisia	29	36	20	85	0.341	0.424	0.235	94	76	0.553	0.447	0.192
Sub–Saharan Africa	Ethiopia	18	27	21	66	0.273	0.409	0.318	63	69	0.477	0.523	0.145
Sub–Saharan Africa	Somalia	29	34	22	85	0.341	0.400	0.259	92	78	0.541	0.459	0.083
Sub–Saharan Africa	Nigeria – Yoruba	45	50	14	109	0.413	0.459	0.128	140	78	0.642	0.358	1.000
Sub–Saharan Africa	Nigeria – Esan	37	47	15	99	0.374	0.475	0.152	121	77	0.611	0.389	1.000
Sub–Saharan Africa	The Gambia	34	58	21	113	0.301	0.513	0.186	126	100	0.558	0.442	0.850
Sub–Saharan Africa	Kenya – Luhya	44	45	12	101	0.436	0.446	0.119	133	69	0.658	0.342	1.000
Sub–Saharan Africa	Sierra Leone	34	33	18	85	0.400	0.388	0.212	101	69	0.594	0.406	0.074
Sub–Saharan Africa	Senegal	13	35	15	63	0.206	0.556	0.238	61	65	0.484	0.516	0.456
Sub–Saharan Africa	Tanzania–Sandawe	6	15	7	28	0.214	0.536	0.250	27	29	0.482	0.518	1.000
Sub–Saharan Africa	South Africa– San	2	14	14	30	0.067	0.467	0.467	18	42	0.300	0.700	0.688
Sub–Saharan Africa	Maasai – Kenya	26	75	42	143	0.182	0.524	0.294	127	159	0.444	0.556	0.504
Southwest Asia	Yemen	10	34	22	66	0.152	0.515	0.333	54	78	0.409	0.591	0.798
Southwest Asia	Saudi Arabia	28	43	22	93	0.301	0.462	0.237	99	87	0.532	0.468	0.532

219

Southwest Asia	Oman	26	37	13	76	0.342	0.487	0.171	89	63	0.586	0.414	1.000
Southwest Asia	UAE	47	35	7	89	0.528	0.393	0.079	129	49	0.725	0.275	1.000
Southwest Asia	Druze	15	21	6	42	0.357	0.500	0.143	51	33	0.607	0.393	1.000
Southwest Asia	Jordania	7	10	3	20	0.350	0.500	0.150	24	16	0.600	0.400	1.000
Southwest Asia	Palestina	13	25	8	46	0.283	0.543	0.174	51	41	0.554	0.446	0.565
Southwest Asia	Syria	6	6	4	16	0.375	0.375	0.250	18	14	0.563	0.438	0.347
Southwest Asia	Turkey	7	7	3	17	0.412	0.412	0.176	21	13	0.618	0.382	0.623
Southwest Asia	Iran	32	47	15	94	0.340	0.500	0.160	111	77	0.590	0.410	0.832
South Asia	Afghanistan – Hazara	12	9	1	22	0.545	0.409	0.045	33	11	0.750	0.250	1.000
South Asia	Afghanistan – Brahui in Balochistan	12	10	3	25	0.480	0.400	0.120	34	16	0.680	0.320	0.664
South Asia	Pakistan – Balochi in Balochistan	9	12	3	24	0.375	0.500	0.125	30	18	0.625	0.375	1.000
South Asia	Pakistan – Punjavi in Lahore	47	36	13	96	0.490	0.375	0.135	130	62	0.677	0.323	0.166
South Asia	India – Gujarati in Texas	60	41	5	106	0.566	0.387	0.047	161	51	0.759	0.241	0.789
South Asia	India – Telegu in UK	54	43	6	103	0.524	0.417	0.058	151	55	0.733	0.267	0.618
South Asia	Sri Lanka – Tamil in UK	60	32	11	103	0.583	0.311	0.107	152	54	0.738	0.262	0.071
South Asia	Bangladesh – Bengali	50	28	8	86	0.581	0.326	0.093	128	44	0.744	0.256	0.167
Northeast Asia	China – Dai in Xishuangbanna	52	39	8	99	0.525	0.394	0.081	143	55	0.722	0.278	0.807
Northeast Asia	China – Han in Beijing	68	31	4	103	0.660	0.301	0.039	167	39	0.811	0.189	0.755
Northeast Asia	China – Han in South	66	37	5	108	0.611	0.343	0.046	169	47	0.782	0.218	1.000
Northeast Asia	Japan	66	33	5	104	0.635	0.317	0.048	165	43	0.793	0.207	0.765
Northeast Asia	Korea	21	12	2	35	0.600	0.343	0.057	54	16	0.771	0.229	1.000
Southeast Asia	Vietnam	50	41	10	101	0.495	0.406	0.099	141	61	0.698	0.302	0.812
Southeast Asia	Thailand	39	25	7	71	0.549	0.352	0.099	103	39	0.725	0.275	0.372
Southeast Asia	Cambodia	40	34	7	81	0.494	0.420	0.086	114	48	0.704	0.296	1.000
Southeast Asia	Singapore	56	33	7	96	0.583	0.344	0.073	145	47	0.755	0.245	0.579
North America	African ancestry – Southwest USA	27	31	8	66	0.409	0.470	0.121	85	47	0.644	0.356	1.000
North America	Utah – European ancestry	44	48	7	99	0.444	0.485	0.071	136	62	0.687	0.313	0.250
Caribe	African Caribbean in Barbados	43	40	13	96	0.448	0.417	0.135	126	66	0.656	0.344	0.500
Latin America	Colombia – Medelin	52	37	5	94	0.553	0.394	0.053	141	47	0.750	0.250	0.785
Latin America	Mexico – LA, USA	46	19	2	67	0.687	0.284	0.030	111	23	0.828	0.172	1.000
Latin America	Peru – Lima	71	14	1	86	0.826	0.163	0.012	156	16	0.907	0.093	0.536
Latin America	Puerto Rico	54	38	13	105	0.514	0.362	0.124	146	64	0.695	0.305	0.164
Latin America	Brazil	24	10	0	34	0.706	0.294	0.000	58	10	0.853	0.147	1.000

Table A.2– Raw population data for *PLCE1*–rs3765524 SNP.

		Genotype				Frequency genotype			Allele count				
Region	Population	CC	CT	TT	N	CC	CT	TT	Allele C	Allele T	Freq Alelle C	Freq Allele T	Hardy-Weinberg p - value
Europe	Portugal	36	33	8	77	0.468	0.429	0.104	105	49	0.682	0.318	1.000
Europe	Spain	42	51	14	107	0.393	0.477	0.131	135	79	0.631	0.369	1.000

Europe	Italy	44	49	15	108	0.407	0.454	0.139	137	79	0.634	0.366	0.837
Europe	France	15	12	1	28	0.536	0.429	0.036	42	14	0.750	0.250	0.647
Europe	UK	46	35	11	92	0.500	0.380	0.120	127	57	0.690	0.310	0.328
Europe	Netherlands	268	178	52	498	0.538	0.357	0.104	714	282	0.717	0.283	0.008
Europe	Finland	61	33	5	99	0.616	0.333	0.051	155	43	0.783	0.217	0.773
Europe	Norway	12	5	1	18	0.667	0.278	0.056	29	7	0.806	0.194	0.510
Europe	Greece	7	11	2	20	0.350	0.550	0.100	25	15	0.625	0.375	0.647
Europe	Hungary	10	9	1	20	0.500	0.450	0.050	29	11	0.725	0.275	1.000
Europe	Russia	14	8	3	25	0.560	0.320	0.120	36	14	0.720	0.280	0.329
Europe	Georgia	9	9	2	20	0.450	0.450	0.100	27	13	0.675	0.325	1.000
North Africa	Morocco	22	14	12	48	0.458	0.292	0.250	58	38	0.604	0.396	0.008
North Africa	Algeria–Mozabite	11	13	5	29	0.379	0.448	0.172	35	23	0.603	0.397	0.711
North Africa	Tunisia	33	34	22	89	0.371	0.382	0.247	100	78	0.562	0.438	0.034
Sub–Saharan Africa	Ethiopia	21	22	26	69	0.304	0.319	0.377	64	74	0.464	0.536	0.000
Sub–Saharan Africa	Somalia	28	43	21	92	0.304	0.467	0.228	99	85	0.538	0.462	0.674
Sub–Saharan Africa	Nigeria – Yoruba	33	53	23	109	0.303	0.486	0.211	119	99	0.546	0.454	0.849
Sub–Saharan Africa	Nigeria – Esan	31	48	20	99	0.313	0.485	0.202	110	88	0.556	0.444	0.841
Sub–Saharan Africa	The Gambia	28	54	31	113	0.248	0.478	0.274	110	116	0.487	0.513	0.704
Sub–Saharan Africa	Kenya – Luhya	34	51	16	101	0.337	0.505	0.158	119	83	0.589	0.411	0.837
Sub–Saharan Africa	Sierra Leone	24	38	23	85	0.282	0.447	0.271	86	84	0.506	0.494	0.386
Sub–Saharan Africa	Senegal	12	36	23	71	0.169	0.507	0.324	60	82	0.423	0.577	0.812
Sub–Saharan Africa	Tanzania–Sandawe	7	16	5	28	0.250	0.571	0.179	30	26	0.536	0.464	0.703
Sub–Saharan Africa	Namibia– San	1	13	16	30	0.033	0.433	0.533	15	45	0.250	0.750	0.640
Sub–Saharan Africa	Maasai – Kenya	20	80	43	143	0.140	0.559	0.301	120	166	0.420	0.580	0.089
Southwest Asia	Yemen	10	36	27	73	0.137	0.493	0.370	56	90	0.384	0.616	0.807
Southwest Asia	Saudi Arabia	25	47	22	94	0.266	0.500	0.234	97	91	0.516	0.484	1.000
Southwest Asia	Oman	27	41	17	85	0.318	0.482	0.200	95	75	0.559	0.441	0.829
Southwest Asia	UAE	47	35	9	91	0.516	0.385	0.099	129	53	0.709	0.291	0.610
Southwest Asia	Druze	15	20	7	42	0.357	0.476	0.167	50	34	0.595	0.405	1.000
Southwest Asia	Jordania	7	10	3	20	0.350	0.500	0.150	24	16	0.600	0.400	1.000
Southwest Asia	Palestina	13	25	8	46	0.283	0.543	0.174	51	41	0.554	0.446	0.566
Southwest Asia	Syria	6	6	4	16	0.375	0.375	0.250	18	14	0.563	0.438	0.347
Southwest Asia	Turkey	7	7	3	17	0.412	0.412	0.176	21	13	0.618	0.382	0.625
Southwest Asia	Iran	41	39	15	95	0.432	0.411	0.158	121	69	0.637	0.363	0.272
South Asia	Afghanistan – Hazara	14	7	1	22	0.636	0.318	0.045	35	9	0.795	0.205	1.000

Region	Population	AA	AG	GG	N	AA	AG	GG	Allele A	Allele G	Freq A	Freq G	HW p-value
South Asia	Afghanistan – Brahui in Balochistan	12	11	2	25	0.480	0.440	0.080	35	15	0.700	0.300	1.000
South Asia	Pakistan – Balochi in Balochistan	10	12	2	24	0.417	0.500	0.083	32	16	0.667	0.333	1.000
South Asia	Pakistan – Punjavi in Lahore	53	34	9	96	0.552	0.354	0.094	140	52	0.729	0.271	0.309
South Asia	India – Gujarati in Texas	63	38	5	106	0.594	0.358	0.047	164	48	0.774	0.226	1.000
South Asia	India – Telegu in UK	57	41	5	103	0.553	0.398	0.049	155	51	0.752	0.248	0.602
South Asia	Sri Lanka – Tamil in UK	61	31	11	103	0.592	0.301	0.107	153	53	0.743	0.257	0.038
South Asia	Bangladesh – Bengali	54	24	8	86	0.628	0.279	0.093	132	40	0.767	0.233	0.064
Northeast Asia	China – Dai in Xishuangbanna	52	39	8	99	0.525	0.394	0.081	143	55	0.722	0.278	0.807
Northeast Asia	China – Han in Beijing	68	31	4	103	0.660	0.301	0.039	167	39	0.811	0.189	0.755
Northeast Asia	China – Han in South	67	36	5	108	0.620	0.333	0.046	170	46	0.787	0.213	1.000
Northeast Asia	Japan	66	33	5	104	0.635	0.317	0.048	165	43	0.793	0.207	0.765
Northeast Asia	Korea	21	12	2	35	0.600	0.343	0.057	54	16	0.771	0.229	1.000
Southeast Asia	Vietnam	50	41	10	101	0.495	0.406	0.099	141	61	0.698	0.302	0.814
Southeast Asia	Thailand	42	25	10	77	0.545	0.325	0.130	109	45	0.708	0.292	0.058
Southeast Asia	Cambodia	41	39	11	91	0.451	0.429	0.121	121	61	0.665	0.335	0.814
Southeast Asia	Singapore	55	34	7	96	0.573	0.354	0.073	144	48	0.750	0.250	0.589
North America	African ancestry – Southwest USA	22	35	9	66	0.333	0.530	0.136	79	53	0.598	0.402	0.455
North America	Utah – European ancestry	51	42	6	99	0.515	0.424	0.061	144	54	0.727	0.273	0.617
Caribe	African Caribbean in Barbados	34	40	22	96	0.354	0.417	0.229	108	84	0.563	0.438	0.146
Latin America	Colombia – Medelin	52	37	5	94	0.553	0.394	0.053	141	47	0.750	0.250	0.786
Latin America	Mexico – LA, USA	47	19	1	67	0.701	0.284	0.015	113	21	0.843	0.157	1.000
Latin America	Peru – Lima	71	14	1	86	0.826	0.163	0.012	156	16	0.907	0.093	0.536
Latin America	Puerto Rico	55	36	14	105	0.524	0.343	0.133	146	64	0.695	0.305	0.064
Latin America	Brazil	28	12	1	41	0.683	0.293	0.024	68	14	0.829	0.171	1.000

Table A.3– Raw population data for *TNFA–*rs1800629 SNP.

		Genotype				Frequency genotype			Allele count				
Region	Population	AA	AG	GG	N	AA	AG	GG	Allele A	Allele G	Freq Alelle A	Freq Allele G	Hardy-Weinberg p - value
Europe	Portugal	0	10	50	60	0.000	0.167	0.833	10	110	0.083	0.917	1.000
Europe	Spain	0	31	76	107	0.000	0.290	0.710	31	183	0.145	0.855	0.120
Europe	Italy	0	20	88	108	0.000	0.185	0.815	20	196	0.093	0.907	0.595
Europe	UK	3	16	73	92	0.033	0.174	0.793	22	162	0.120	0.880	0.112
Europe	Netherlands	24	143	330	497	0.048	0.288	0.664	191	803	0.192	0.808	0.113
Europe	Finland	1	23	75	99	0.010	0.232	0.758	25	173	0.126	0.874	1.000
North Africa	Morocco	0	3	20	23	0.000	0.130	0.870	3	43	0.065	0.935	1.000
North Africa	Tunisia	4	27	45	76	0.053	0.355	0.592	35	117	0.230	0.770	1.000
Sub–Saharan Africa	Ethiopia	2	5	42	49	0.041	0.102	0.857	9	89	0.092	0.908	0.038

Sub–Saharan Africa	Somalia	1	13	77	91	0.011	0.143	0.846	15	167	0.082	0.918	0.466
Sub–Saharan Africa	Nigeria – Yoruba	1	20	88	109	0.009	0.183	0.807	22	196	0.101	0.899	1.000
Sub–Saharan Africa	Nigeria – Esan	2	21	76	99	0.020	0.212	0.768	25	173	0.126	0.874	0.647
Sub–Saharan Africa	The Gambia	2	28	83	113	0.018	0.248	0.735	32	194	0.142	0.858	1.000
Sub–Saharan Africa	Kenya – Luhya	0	18	83	101	0.000	0.178	0.822	18	184	0.089	0.911	1.000
Sub–Saharan Africa	Sierra Leone	2	23	60	85	0.024	0.271	0.706	27	143	0.159	0.841	1.000
Sub–Saharan Africa	Senegal	4	15	54	73	0.055	0.205	0.740	23	123	0.158	0.842	0.066
Sub–Saharan Africa	Maasai – Kenya	1	21	121	143	0.007	0.147	0.846	23	263	0.080	0.920	1.000
Southwest Asia	Yemen	2	16	48	66	0.030	0.242	0.727	20	112	0.152	0.848	0.630
Southwest Asia	Saudi Arabia	2	26	66	94	0.021	0.277	0.702	30	158	0.160	0.840	1.000
Southwest Asia	Oman	4	19	62	85	0.047	0.224	0.729	27	143	0.159	0.841	0.208
Southwest Asia	UAE	2	18	58	78	0.026	0.231	0.744	22	134	0.141	0.859	0.640
Southwest Asia	Iran	1	14	60	75	0.013	0.187	0.800	16	134	0.107	0.893	1.000
South Asia	Pakistan – Punjavi in Lahore	0	11	85	96	0.000	0.115	0.885	11	181	0.057	0.943	1.000
South Asia	India – Gujarati in Texas	1	9	96	106	0.009	0.085	0.906	11	201	0.052	0.948	0.238
South Asia	India – Telegu in UK	1	6	96	103	0.010	0.058	0.932	8	198	0.039	0.961	0.131
South Asia	Sri Lanka – Tamil in UK	0	19	84	103	0.000	0.184	0.816	19	187	0.092	0.908	0.597
South Asia	Bangladesh – Bengali	0	4	82	86	0.000	0.047	0.953	4	168	0.023	0.977	1.000
Northeast Asia	China – Dai in Xishuangbanna	0	13	86	99	0.000	0.131	0.869	13	185	0.066	0.934	1.000
Northeast Asia	China – Han in Beijing	1	17	85	103	0.010	0.165	0.825	19	187	0.092	0.908	1.000
Northeast Asia	China – Han in South	0	13	95	108	0.000	0.120	0.880	13	203	0.060	0.940	1.000
Northeast Asia	Japan	0	4	100	104	0.000	0.038	0.962	4	204	0.019	0.981	1.000
Northeast Asia	Korea	1	5	29	35	0.029	0.143	0.829	7	63	0.100	0.900	0.283
Southeast Asia	Vietnam	0	11	90	101	0.000	0.109	0.891	11	191	0.054	0.946	1.000
Southeast Asia	Thailand	0	4	77	81	0.000	0.049	0.951	4	158	0.025	0.975	1.000
Southeast Asia	Cambodia	0	10	81	91	0.000	0.110	0.890	10	172	0.055	0.945	1.000
Southeast Asia	Singapore	0	5	91	96	0.000	0.052	0.948	5	187	0.026	0.974	1.000
North America	African ancestry – Southwest USA	0	9	57	66	0.000	0.136	0.864	9	123	0.068	0.932	1.000
North America	Utah – European ancestry	2	33	64	99	0.020	0.333	0.646	37	161	0.187	0.813	0.511
Caribe	African Caribbean in Barbados	1	24	71	96	0.010	0.250	0.740	26	166	0.135	0.865	1.000
Latin America	Colombia – Medelin	0	13	81	94	0.000	0.138	0.862	13	175	0.069	0.931	1.000
Latin America	Mexico – LA, USA	0	7	60	67	0.000	0.104	0.896	7	127	0.052	0.948	1.000
Latin America	Peru – Lima	0	10	76	86	0.000	0.116	0.884	10	162	0.058	0.942	1.000
Latin America	Puerto Rico	2	14	89	105	0.019	0.133	0.848	18	192	0.086	0.914	0.156
Latin America	Brazil	0	5	36	41	0.000	0.122	0.878	5	77	0.061	0.939	1.000

223

Table A.4– Raw population data for *CD209*–rs4804803 SNP.

Region	Population	Genotype				Frequency genotype			Allele count		Freq Alelle A	Freq Alelle G	Hardy-Weinberg p - value
		AA	AG	GG	N	AA	AG	GG	Allele A	Allele G			
Europe	Portugal	48	8	2	58	0.828	0.138	0.034	104	12	0.897	0.103	0.096
Europe	Spain	67	32	8	107	0.626	0.299	0.075	166	48	0.776	0.224	0.163
Europe	Italy	73	31	4	108	0.676	0.287	0.037	177	39	0.819	0.181	0.746
Europe	UK	56	30	6	92	0.609	0.326	0.065	142	42	0.772	0.228	0.550
Europe	Netherlands	337	142	19	498	0.677	0.285	0.038	816	180	0.819	0.181	0.451
Europe	Finland	63	31	5	99	0.636	0.313	0.051	157	41	0.793	0.207	0.758
North Africa	Morocco	18	4	0	22	0.818	0.182	0.000	40	4	0.909	0.091	1.000
North Africa	Tunisia	39	26	5	70	0.557	0.371	0.071	104	36	0.743	0.257	0.762
Sub–Saharan Africa	Ethiopia	13	10	8	31	0.419	0.323	0.258	36	26	0.581	0.419	0.069
Sub–Saharan Africa	Somalia	35	38	12	85	0.412	0.447	0.141	108	62	0.635	0.365	0.815
Sub–Saharan Africa	Nigeria – Yoruba	37	46	26	109	0.339	0.422	0.239	120	98	0.550	0.450	0.124
Sub–Saharan Africa	Nigeria – Esan	33	46	20	99	0.333	0.465	0.202	112	86	0.566	0.434	0.683
Sub–Saharan Africa	The Gambia	24	62	27	113	0.212	0.549	0.239	110	116	0.487	0.513	0.349
Sub–Saharan Africa	Kenya – Luhya	46	42	13	101	0.455	0.416	0.129	134	68	0.663	0.337	0.506
Sub–Saharan Africa	Sierra Leone	19	42	24	85	0.224	0.494	0.282	80	90	0.471	0.529	1.000
Sub–Saharan Africa	Senegal	17	29	14	60	0.283	0.483	0.233	63	57	0.525	0.475	0.800
Southwest Asia	Yemen	33	27	5	65	0.508	0.415	0.077	93	37	0.715	0.285	1.000
Southwest Asia	Saudi Arabia	47	35	12	94	0.500	0.372	0.128	129	59	0.686	0.314	0.227
Southwest Asia	Oman	35	27	11	73	0.479	0.370	0.151	97	49	0.664	0.336	0.190
Southwest Asia	UAE	41	29	4	74	0.554	0.392	0.054	111	37	0.750	0.250	1.000
Southwest Asia	Iran	51	15	5	71	0.718	0.211	0.070	117	25	0.824	0.176	0.030
South Asia	Pakistan – Punjavi in Lahore	64	24	8	96	0.667	0.250	0.083	152	40	0.792	0.208	0.026
South Asia	India – Gujarati in Texas	69	34	3	106	0.651	0.321	0.028	172	40	0.811	0.189	0.760
South Asia	India – Telegu in UK	76	24	3	103	0.738	0.233	0.029	176	30	0.854	0.146	0.442
South Asia	Sri Lanka – Tamil in UK	65	31	7	103	0.631	0.301	0.068	161	45	0.782	0.218	0.247
South Asia	Bangladesh – Bengali	54	31	1	86	0.628	0.360	0.012	139	33	0.808	0.192	0.180
Northeast Asia	China – Dai in Xishuangbanna	84	15	0	99	0.848	0.152	0.000	183	15	0.924	0.076	1.000
Northeast Asia	China – Han in Beijing	90	13	0	103	0.874	0.126	0.000	193	13	0.937	0.063	1.000
Northeast Asia	China – Han in South	90	18	0	108	0.833	0.167	0.000	198	18	0.917	0.083	1.000
Northeast Asia	Japan	96	8	0	104	0.923	0.077	0.000	200	8	0.962	0.038	1.000
Northeast Asia	Korea	32	3	0	35	0.914	0.086	0.000	67	3	0.957	0.043	1.000
Southeast Asia	Vietnam	86	15	0	101	0.851	0.149	0.000	187	15	0.926	0.074	1.000

224

Southeast Asia	Thailand	62	8	1	71	0.873	0.113	0.014	132	10	0.930	0.070	0.288
Southeast Asia	Cambodia	74	11	1	86	0.860	0.128	0.012	159	13	0.924	0.076	0.387
Southeast Asia	Singapore	78	17	1	96	0.813	0.177	0.010	173	19	0.901	0.099	1.000
North America	African ancestry – Southwest USA	17	38	11	66	0.258	0.576	0.167	72	60	0.545	0.455	0.225
North America	Utah – European ancestry	55	39	5	99	0.556	0.394	0.051	149	49	0.753	0.247	0.788
Caribe	African Caribbean in Barbados	36	41	19	96	0.375	0.427	0.198	113	79	0.589	0.411	0.291
Latin America	Colombia – Medelin	67	22	5	94	0.713	0.234	0.053	156	32	0.830	0.170	0.132
Latin America	Mexico – LA, USA	46	16	5	67	0.687	0.239	0.075	108	26	0.806	0.194	0.057
Latin America	Peru – Lima	73	13	0	86	0.849	0.151	0.000	159	13	0.924	0.076	1.000
Latin America	Puerto Rico	68	33	4	105	0.648	0.314	0.038	169	41	0.805	0.195	1.000
Latin America	Brazil	24	7	0	31	0.774	0.226	0.000	55	7	0.887	0.113	1.000

Table A.5– Raw population data for *OAS1*–rs10774671 SNP.

		Genotype				Frequency genotype			Allele count				Hardy-Weinberg p - value
Region	Population	AA	AG	GG	N	AA	AG	GG	Allele A	Allele G	Freq Alelle A	Freq Alelle G	
Europe	Portugal	37	24	2	63	0.587	0.381	0.032	98	28	0.778	0.222	0.714
Europe	Spain	44	55	8	107	0.411	0.514	0.075	143	71	0.668	0.332	0.128
Europe	Italy	36	53	19	108	0.333	0.491	0.176	125	91	0.579	0.421	1.000
Europe	UK	41	36	15	92	0.446	0.391	0.163	118	66	0.641	0.359	0.173
Europe	Netherlands	243	193	62	498	0.488	0.388	0.124	679	317	0.682	0.318	0.018
Europe	Finland	51	40	8	99	0.515	0.404	0.081	142	56	0.717	0.283	1.000
North Africa	Morocco	6	10	6	22	0.273	0.455	0.273	22	22	0.500	0.500	0.683
North Africa	Tunisia	14	35	21	70	0.200	0.500	0.300	63	77	0.450	0.550	1.000
Sub–Saharan Africa	Ethiopia	9	19	18	46	0.196	0.413	0.391	37	55	0.402	0.598	0.091
Sub–Saharan Africa	Somalia	16	42	30	88	0.182	0.477	0.341	74	102	0.420	0.580	0.829
Sub–Saharan Africa	Nigeria – Yoruba	12	48	49	109	0.110	0.440	0.450	72	146	0.330	0.670	1.000
Sub–Saharan Africa	Nigeria – Esan	9	41	49	99	0.091	0.414	0.495	59	139	0.298	0.702	1.000
Sub–Saharan Africa	The Gambia	16	43	54	113	0.142	0.381	0.478	75	151	0.332	0.668	0.139
Sub–Saharan Africa	Kenya – Luhya	15	50	36	101	0.149	0.495	0.356	80	122	0.396	0.604	0.837
Sub–Saharan Africa	Sierra Leone	7	46	32	85	0.082	0.541	0.376	60	110	0.353	0.647	0.153
Sub–Saharan Africa	Senegal	13	24	30	67	0.194	0.358	0.448	50	84	0.373	0.627	0.067

Region	Population	AA	AG	GG	N	AA	AG	GG	Allele A	Allele G	Freq Allele A	Freq Allele G	Hardy-Weinberg p-value
Sub–Saharan Africa	Maasai – Kenya	11	64	68	143	0.077	0.448	0.476	86	200	0.301	0.699	0.551
Southwest Asia	Yemen	22	31	12	65	0.338	0.477	0.185	75	55	0.577	0.423	1.000
Southwest Asia	Saudi Arabia	31	38	25	94	0.330	0.404	0.266	100	88	0.532	0.468	0.066
Southwest Asia	Oman	23	45	12	80	0.288	0.563	0.150	91	69	0.569	0.431	0.638
Southwest Asia	UAE	28	33	15	76	0.368	0.434	0.197	89	63	0.586	0.414	0.353
Southwest Asia	Iran	32	28	15	75	0.427	0.373	0.200	92	58	0.613	0.387	0.087
South Asia	Pakistan – Punjavi in Lahore	42	40	14	96	0.438	0.417	0.146	124	68	0.646	0.354	0.379
South Asia	India – Gujarati in Texas	53	41	12	106	0.500	0.387	0.113	147	65	0.693	0.307	0.362
South Asia	India – Telegu in UK	53	40	10	103	0.515	0.388	0.097	146	60	0.709	0.291	0.631
South Asia	Sri Lanka – Tamil in UK	55	36	12	103	0.534	0.350	0.117	146	60	0.709	0.291	0.151
South Asia	Bangladesh – Bengali	54	31	1	86	0.628	0.360	0.012	139	33	0.808	0.192	0.181
Northeast Asia	China – Dai in Xishuangbanna	54	39	6	99	0.545	0.394	0.061	147	51	0.742	0.258	1.000
Northeast Asia	China – Han in Beijing	52	39	12	103	0.505	0.379	0.117	143	63	0.694	0.306	0.254
Northeast Asia	China – Han in South	68	35	5	108	0.630	0.324	0.046	171	45	0.792	0.208	0.776
Northeast Asia	Japan	73	28	3	104	0.702	0.269	0.029	174	34	0.837	0.163	0.733
Northeast Asia	Korea	20	12	3	35	0.571	0.343	0.086	52	18	0.743	0.257	0.656
Southeast Asia	Vietnam	45	40	16	101	0.446	0.396	0.158	130	72	0.644	0.356	0.192
Southeast Asia	Thailand	43	29	5	77	0.558	0.377	0.065	115	39	0.747	0.253	1.000
Southeast Asia	Cambodia	44	32	14	90	0.489	0.356	0.156	120	60	0.667	0.333	0.060
Southeast Asia	Singapore	44	36	16	96	0.458	0.375	0.167	124	68	0.646	0.354	0.077
North America	African ancestry – Southwest USA	23	27	16	66	0.348	0.409	0.242	73	59	0.553	0.447	0.211
North America	Utah – European ancestry	38	50	11	99	0.384	0.505	0.111	126	72	0.636	0.364	0.512
Caribe	African Caribbean in Barbados	8	51	37	96	0.083	0.531	0.385	67	125	0.349	0.651	0.122
Latin America	Colombia – Medelin	49	32	13	94	0.521	0.340	0.138	130	58	0.691	0.309	0.054
Latin America	Mexico – LA, USA	37	25	5	67	0.552	0.373	0.075	99	35	0.739	0.261	0.756
Latin America	Peru – Lima	68	17	1	86	0.791	0.198	0.012	153	19	0.890	0.110	1.000
Latin America	Puerto Rico	39	53	13	105	0.371	0.505	0.124	131	79	0.624	0.376	0.536
Latin America	Brazil	15	18	2	35	0.429	0.514	0.057	48	22	0.686	0.314	0.436

Table A.6- Raw population data for *OAS3*–rs1859330 SNP.

		Genotype				Frequency genotype			Allele count		Freq Alelle A	Freq Alelle G	Hardy-Weinberg p-value
Region	Population	AA	AG	GG	N	AA	AG	GG	Allele A	Allele G			
Europe	Portugal	31	39	3	73	0.425	0.534	0.041	101	45	0.692	0.308	0.145
Europe	Spain	41	53	13	107	0.383	0.495	0.121	135	79	0.631	0.369	0.678
Europe	Italy	33	54	21	108	0.306	0.500	0.194	120	96	0.556	0.444	1.000
Europe	UK	39	38	15	92	0.424	0.413	0.163	116	68	0.630	0.370	0.269
Europe	Netherlands	244	178	73	495	0.493	0.360	0.147	666	324	0.673	0.327	0.000

Europe	Finland	52	39	8	99	0.525	0.394	0.081	143	55	0.722	0.278	0.805
North Africa	Morocco	7	10	5	22	0.318	0.455	0.227	24	20	0.545	0.455	0.455
North Africa	Tunisia	24	35	16	75	0.320	0.467	0.213	83	67	0.553	0.447	0.643
Sub–Saharan Africa	Ethiopia	27	13	4	44	0.614	0.295	0.091	67	21	0.761	0.239	0.220
Sub–Saharan Africa	Somalia	32	47	10	89	0.360	0.528	0.112	111	67	0.624	0.376	0.366
Sub–Saharan Africa	Nigeria – Yoruba	32	54	23	109	0.294	0.495	0.211	118	100	0.541	0.459	1.000
Sub–Saharan Africa	Nigeria – Esan	28	43	28	99	0.283	0.434	0.283	99	99	0.500	0.500	0.225
Sub–Saharan Africa	The Gambia	24	61	28	113	0.212	0.540	0.248	109	117	0.482	0.518	0.454
Sub–Saharan Africa	Kenya – Luhya	53	35	13	101	0.525	0.347	0.129	141	61	0.698	0.302	0.095
Sub–Saharan Africa	Sierra Leone	28	40	17	85	0.329	0.471	0.200	96	74	0.565	0.435	0.665
Sub–Saharan Africa	Senegal	17	33	20	70	0.243	0.471	0.286	67	73	0.479	0.521	0.809
Southwest Asia	Yemen	26	30	10	66	0.394	0.455	0.152	82	50	0.621	0.379	0.796
Southwest Asia	Saudi Arabia	40	35	19	94	0.426	0.372	0.202	115	73	0.612	0.388	0.050
Southwest Asia	Oman	31	43	11	85	0.365	0.506	0.129	105	65	0.618	0.382	0.651
Southwest Asia	UAE	34	31	12	77	0.442	0.403	0.156	99	55	0.643	0.357	0.320
Southwest Asia	Iran	34	27	13	74	0.459	0.365	0.176	95	53	0.642	0.358	0.081
South Asia	Pakistan – Punjavi in Lahore	43	39	14	96	0.448	0.406	0.146	125	67	0.651	0.349	0.367
South Asia	India – Gujarati in Texas	54	40	12	106	0.509	0.377	0.113	148	64	0.698	0.302	0.354
South Asia	India – Telegu in UK	53	40	10	103	0.515	0.388	0.097	146	60	0.709	0.291	0.633
South Asia	Sri Lanka – Tamil in UK	58	33	12	103	0.563	0.320	0.117	149	57	0.723	0.277	0.049
South Asia	Bangladesh – Bengali	54	31	1	86	0.628	0.360	0.012	139	33	0.808	0.192	0.180
Northeast Asia	China – Dai in Xishuangbanna	54	39	6	99	0.545	0.394	0.061	147	51	0.742	0.258	1.000
Northeast Asia	China – Han in Beijing	53	38	12	103	0.515	0.369	0.117	144	62	0.699	0.301	0.239
Northeast Asia	China – Han in South	68	35	5	108	0.630	0.324	0.046	171	45	0.792	0.208	0.777
Northeast Asia	Japan	74	27	3	104	0.712	0.260	0.029	175	33	0.841	0.159	0.717
Northeast Asia	Korea	18	8	9	35	0.514	0.229	0.257	44	26	0.629	0.371	0.003
Southeast Asia	Vietnam	47	38	16	101	0.465	0.376	0.158	132	70	0.653	0.347	0.122
Southeast Asia	Thailand	44	32	5	81	0.543	0.395	0.062	120	42	0.741	0.259	1.000
Southeast Asia	Cambodia	37	40	14	91	0.407	0.440	0.154	114	68	0.626	0.374	0.653
Southeast Asia	Singapore	45	40	11	96	0.469	0.417	0.115	130	62	0.677	0.323	0.644
North America	African ancestry – Southwest USA	28	29	9	66	0.424	0.439	0.136	85	47	0.644	0.356	0.787
North America	Utah – European ancestry	34	54	11	99	0.343	0.545	0.111	122	76	0.616	0.384	0.200
Caribe	African Caribbean in Barbados	23	52	21	96	0.240	0.542	0.219	98	94	0.510	0.490	0.538
Latin America	Colombia – Medelin	47	39	8	94	0.500	0.415	0.085	133	55	0.707	0.293	1.000
Latin America	Mexico – LA, USA	38	24	5	67	0.567	0.358	0.075	100	34	0.746	0.254	0.746

Latin America	Peru – Lima	69	16	1	86	0.802	0.186	0.012	154	18	0.895	0.105	1.000
Latin America	Puerto Rico	34	55	16	105	0.324	0.524	0.152	123	87	0.586	0.414	0.549
Latin America	Brazil	13	22	4	39	0.333	0.564	0.103	48	30	0.615	0.385	0.323

Table A.7– Raw population data for *OAS3*–rs2285933 SNP.

		Genotype				Frequency genotype			Allele count				
Region	Population	CC	CG	GG	N	CC	CG	GG	Allele C	Allele G	Freq Alelle C	Freq Allele G	Hardy-Weinberg p - value
Europe	Portugal	36	23	7	66	0.545	0.348	0.106	95	37	0.720	0.280	0.434
Europe	Spain	58	45	4	107	0.542	0.421	0.037	161	53	0.752	0.248	0.294
Europe	Italy	60	41	7	108	0.556	0.380	0.065	161	55	0.745	0.255	1.000
Europe	UK	42	40	10	92	0.457	0.435	0.109	124	60	0.674	0.326	1.000
Europe	Netherlands	260	184	54	498	0.522	0.369	0.108	704	292	0.707	0.293	0.017
Europe	Finland	60	36	3	99	0.606	0.364	0.030	156	42	0.788	0.212	0.550
North Africa	Morocco	9	10	3	22	0.409	0.455	0.136	28	16	0.636	0.364	1.000
North Africa	Tunisia	43	18	7	68	0.632	0.265	0.103	104	32	0.765	0.235	1.000
Sub–Saharan Africa	Ethiopia	28	7	5	40	0.700	0.175	0.125	63	17	0.788	0.213	0.042
Sub–Saharan Africa	Somalia	40	38	10	88	0.455	0.432	0.114	118	58	0.670	0.330	0.813
Sub–Saharan Africa	Nigeria – Yoruba	35	49	25	109	0.321	0.450	0.229	119	99	0.546	0.454	0.342
Sub–Saharan Africa	Nigeria – Esan	26	48	25	99	0.263	0.485	0.253	100	98	0.505	0.495	0.841
Sub–Saharan Africa	The Gambia	42	56	15	113	0.372	0.496	0.133	140	86	0.619	0.381	0.693
Sub–Saharan Africa	Kenya – Luhya	38	49	14	101	0.376	0.485	0.139	125	77	0.619	0.381	0.836
Sub–Saharan Africa	Sierra Leone	37	37	11	85	0.435	0.435	0.129	111	59	0.653	0.347	0.811
Sub–Saharan Africa	Senegal	23	33	8	64	0.359	0.516	0.125	79	49	0.617	0.383	0.600
Southwest Asia	Yemen	35	21	9	65	0.538	0.323	0.138	91	39	0.700	0.300	0.075
Southwest Asia	Saudi Arabia	53	33	7	93	0.570	0.355	0.075	139	47	0.747	0.253	0.583
Southwest Asia	Oman	35	31	8	74	0.473	0.419	0.108	101	47	0.682	0.318	0.787
Southwest Asia	UAE	38	31	9	78	0.487	0.397	0.115	107	49	0.686	0.314	0.788
Southwest Asia	Iran	43	26	6	75	0.573	0.347	0.080	112	38	0.747	0.253	0.540
South Asia	Pakistan – Punjavi in Lahore	77	18	1	96	0.802	0.188	0.010	172	20	0.896	0.104	1.000
South Asia	India – Gujarati in Texas	56	41	9	106	0.528	0.387	0.085	153	59	0.722	0.278	0.807
South Asia	India – Telegu in UK	67	31	5	103	0.650	0.301	0.049	165	41	0.801	0.199	0.541
South Asia	Sri Lanka – Tamil in UK	72	29	2	103	0.699	0.282	0.019	173	33	0.840	0.160	1.000
South Asia	Bangladesh – Bengali	54	29	3	86	0.628	0.337	0.035	137	35	0.797	0.203	1.000

Northeast Asia	China – Dai in Xishuangbanna	76	22	1	99	0.768	0.222	0.010	174	24	0.879	0.121	1.000
Northeast Asia	China – Han in Beijing	75	24	4	103	0.728	0.233	0.039	174	32	0.845	0.155	0.259
Northeast Asia	China – Han in South	68	34	6	108	0.630	0.315	0.056	170	46	0.787	0.213	0.565
Northeast Asia	Japan	72	32	0	104	0.692	0.308	0.000	176	32	0.846	0.154	0.121
Northeast Asia	Korea	25	8	2	35	0.714	0.229	0.057	58	12	0.829	0.171	0.238
Southeast Asia	Vietnam	75	24	2	101	0.743	0.238	0.020	174	28	0.861	0.139	1.000
Southeast Asia	Thailand	53	23	1	77	0.688	0.299	0.013	129	25	0.838	0.162	0.679
Southeast Asia	Cambodia	74	14	1	89	0.831	0.157	0.011	162	16	0.910	0.090	0.524
Southeast Asia	Singapore	73	22	1	96	0.760	0.229	0.010	168	24	0.875	0.125	1.000
North America	African ancestry – Southwest USA	27	33	6	66	0.409	0.500	0.091	87	45	0.659	0.341	0.425
North America	Utah – European ancestry	51	46	2	99	0.515	0.465	0.020	148	50	0.747	0.253	0.031
Caribe	African Caribbean in Barbados	31	50	15	96	0.323	0.521	0.156	112	80	0.583	0.417	0.536
Latin America	Colombia – Medelin	42	43	9	94	0.447	0.457	0.096	127	61	0.676	0.324	0.815
Latin America	Mexico – LA, USA	30	34	3	67	0.448	0.507	0.045	94	40	0.701	0.299	0.142
Latin America	Peru – Lima	20	41	25	86	0.233	0.477	0.291	81	91	0.471	0.529	0.668
Latin America	Puerto Rico	44	47	14	105	0.419	0.448	0.133	135	75	0.643	0.357	0.832
Latin America	Brazil	13	14	5	32	0.406	0.438	0.156	40	24	0.625	0.375	1.000

Table A.8- Raw population data for *CD32*–rs1801274 SNP.

		Genotype				Frequency genotype			Allele count		Freq Alelle A	Freq Alelle G	Hardy-Weinberg p - value
Region	**Population**	**AA**	**AG**	**GG**	**N**	**AA**	**AG**	**GG**	**Allele A**	**Allele G**			
Europe	Portugal	18	27	11	56	0.321	0.482	0.196	63	49	0.563	0.438	1.000
Europe	Spain	23	55	29	107	0.215	0.514	0.271	101	113	0.472	0.528	0.846
Europe	Italy	39	49	20	108	0.361	0.454	0.185	127	89	0.588	0.412	0.553
Europe	UK	18	37	37	92	0.196	0.402	0.402	73	111	0.397	0.603	0.131
Europe	Netherlands	145	236	117	498	0.291	0.474	0.235	526	470	0.528	0.472	0.274
Europe	Finland	23	45	31	99	0.232	0.455	0.313	91	107	0.460	0.540	0.421
North Africa	Morocco	4	9	8	21	0.190	0.429	0.381	17	25	0.405	0.595	0.660
North Africa	Tunisia	16	41	16	73	0.219	0.562	0.219	73	73	0.500	0.500	0.355
Sub–Saharan Africa	Ethiopia	10	14	21	45	0.222	0.311	0.467	34	56	0.378	0.622	0.027
Sub–Saharan Africa	Somalia	18	42	28	88	0.205	0.477	0.318	78	98	0.443	0.557	0.828
Sub–Saharan Africa	Nigeria – Yoruba	24	57	28	109	0.220	0.523	0.257	105	113	0.482	0.518	0.703
Sub–Saharan Africa	Nigeria – Esan	12	56	31	99	0.121	0.566	0.313	80	118	0.404	0.596	0.100
Sub–Saharan Africa	The Gambia	28	61	24	113	0.248	0.540	0.212	117	109	0.518	0.482	0.456

229

Region	Population	CC	CT	TT	N	CC	CT	TT	Allele C	Allele T	Freq Alelle C	Freq Allele T	Hardy-Weinberg p-value
Sub–Saharan Africa	Kenya – Luhya	30	46	25	101	0.297	0.455	0.248	106	96	0.525	0.475	0.424
Sub–Saharan Africa	Sierra Leone	13	40	32	85	0.153	0.471	0.376	66	104	0.388	0.612	1.000
Sub–Saharan Africa	Senegal	10	42	15	67	0.149	0.627	0.224	62	72	0.463	0.537	0.051
Sub–Saharan Africa	Maasai – Kenya	33	77	32	142	0.232	0.542	0.225	143	141	0.504	0.496	0.403
Southwest Asia	Yemen	18	32	15	65	0.277	0.492	0.231	68	62	0.523	0.477	1.000
Southwest Asia	Saudi Arabia	20	50	24	94	0.213	0.532	0.255	90	98	0.479	0.521	0.678
Southwest Asia	Oman	25	36	23	84	0.298	0.429	0.274	86	82	0.512	0.488	0.187
Southwest Asia	UAE	19	36	22	77	0.247	0.468	0.286	74	80	0.481	0.519	0.648
Southwest Asia	Iran	30	34	11	75	0.400	0.453	0.147	94	56	0.627	0.373	0.807
South Asia	Pakistan – Punjavi in Lahore	34	49	13	96	0.354	0.510	0.135	117	75	0.609	0.391	0.530
South Asia	India – Gujarati in Texas	39	48	19	106	0.368	0.453	0.179	126	86	0.594	0.406	0.546
South Asia	India – Telegu in UK	32	49	22	103	0.311	0.476	0.214	113	93	0.549	0.451	0.694
South Asia	Sri Lanka – Tamil in UK	35	39	29	103	0.340	0.379	0.282	109	97	0.529	0.471	0.017
South Asia	Bangladesh – Bengali	37	36	13	86	0.430	0.419	0.151	110	62	0.640	0.360	0.481
Northeast Asia	China – Dai in Xishuangbanna	49	45	5	99	0.495	0.455	0.051	143	55	0.722	0.278	0.314
Northeast Asia	China – Han in Beijing	43	51	9	103	0.417	0.495	0.087	137	69	0.665	0.335	0.374
Northeast Asia	China – Han in South	49	53	6	108	0.454	0.491	0.056	151	65	0.699	0.301	0.110
Northeast Asia	Japan	67	34	3	104	0.644	0.327	0.029	168	40	0.808	0.192	0.757
Northeast Asia	Korea	18	15	2	35	0.514	0.429	0.057	51	19	0.729	0.271	1.000
Southeast Asia	Vietnam	53	40	8	101	0.525	0.396	0.079	146	56	0.723	0.277	1.000
Southeast Asia	Thailand	55	23	1	79	0.696	0.291	0.013	133	25	0.842	0.158	0.679
Southeast Asia	Cambodia	50	31	7	88	0.568	0.352	0.080	131	45	0.744	0.256	0.571
Southeast Asia	Singapore	39	42	15	96	0.406	0.438	0.156	120	72	0.625	0.375	0.516
North America	African ancestry – Southwest USA	16	37	13	66	0.242	0.561	0.197	69	63	0.523	0.477	0.458
North America	Utah – European ancestry	27	48	24	99	0.273	0.485	0.242	102	96	0.515	0.485	0.841
Caribe	African Caribbean in Barbados	20	57	19	96	0.208	0.594	0.198	97	95	0.505	0.495	0.102
Latin America	Colombia – Medelin	34	46	14	94	0.362	0.489	0.149	114	74	0.606	0.394	1.000
Latin America	Mexico – LA, USA	21	26	20	67	0.313	0.388	0.299	68	66	0.507	0.493	0.084
Latin America	Peru – Lima	23	46	17	86	0.267	0.535	0.198	92	80	0.535	0.465	0.665
Latin America	Puerto Rico	28	59	18	105	0.267	0.562	0.171	115	95	0.548	0.452	0.235
Latin America	Brazil	7	16	12	35	0.200	0.457	0.343	30	40	0.429	0.571	0.735

Table A.9– Raw population data for *MICB*–rs3132468 SNP.

		Genotype				Frequency genotype			Allele count		Freq Alelle C	Freq Allele T	Hardy-Weinberg p - value
Region	Population	CC	CT	TT	N	CC	CT	TT	Allele C	Allele T			
Europe	Portugal	4	18	55	77	0.052	0.234	0.714	26	128	0.169	0.831	0.209

Europe	Spain	4	35	68	107	0.037	0.327	0.636	43	171	0.201	0.799	1.000
Europe	Italy	6	34	68	108	0.056	0.315	0.630	46	170	0.213	0.787	0.564
Europe	UK	12	33	47	92	0.130	0.359	0.511	57	127	0.310	0.690	0.142
Europe	Netherlands	26	211	261	498	0.052	0.424	0.524	263	733	0.264	0.736	0.049
Europe	Finland	12	44	43	99	0.121	0.444	0.434	68	130	0.343	0.657	1.000
North Africa	Morocco	3	6	17	26	0.115	0.231	0.654	12	40	0.231	0.769	0.090
North Africa	Tunisia	3	24	48	75	0.040	0.320	0.640	30	120	0.200	0.800	1.000
Sub-Saharan Africa	Ethiopia	6	9	27	42	0.143	0.214	0.643	21	63	0.250	0.750	0.009
Sub-Saharan Africa	Somalia	4	26	60	90	0.044	0.289	0.667	34	146	0.189	0.811	0.508
Sub-Saharan Africa	Nigeria – Yoruba	2	34	73	109	0.018	0.312	0.670	38	180	0.174	0.826	0.520
Sub-Saharan Africa	Nigeria – Esan	1	22	76	99	0.010	0.222	0.768	24	174	0.121	0.879	1.000
Sub-Saharan Africa	The Gambia	12	45	56	113	0.106	0.398	0.496	69	157	0.305	0.695	0.510
Sub-Saharan Africa	Kenya – Luhya	1	20	80	101	0.010	0.198	0.792	22	180	0.109	0.891	1.000
Sub-Saharan Africa	Sierra Leone	2	26	57	85	0.024	0.306	0.671	30	140	0.176	0.824	1.000
Sub-Saharan Africa	Senegal	5	25	41	71	0.070	0.352	0.577	35	107	0.246	0.754	0.748
Sub-Saharan Africa	Maasai – Kenya	6	47	90	143	0.042	0.329	0.629	59	227	0.206	0.794	1.000
Southwest Asia	Yemen	1	12	52	65	0.015	0.185	0.800	14	116	0.108	0.892	0.544
Southwest Asia	Saudi Arabia	3	28	63	94	0.032	0.298	0.670	34	154	0.181	0.819	1.000
Southwest Asia	Oman	1	14	70	85	0.012	0.165	0.824	16	154	0.094	0.906	0.541
Southwest Asia	UAE	5	25	46	76	0.066	0.329	0.605	35	117	0.230	0.770	0.522
Southwest Asia	Iran	2	10	58	70	0.029	0.143	0.829	14	126	0.100	0.900	0.125
South Asia	Pakistan – Punjavi in Lahore	1	24	71	96	0.010	0.250	0.740	26	166	0.135	0.865	1.000
South Asia	India – Gujarati in Texas	8	46	52	106	0.075	0.434	0.491	62	150	0.292	0.708	0.814
South Asia	India – Telegu in UK	8	41	54	103	0.078	0.398	0.524	57	149	0.277	0.723	1.000
South Asia	Sri Lanka – Tamil in UK	4	34	65	103	0.039	0.330	0.631	42	164	0.204	0.796	1.000
South Asia	Bangladesh – Bengali	3	25	58	86	0.035	0.291	0.674	31	141	0.180	0.820	1.000
Northeast Asia	China – Dai in Xishuangbanna	0	12	87	99	0.000	0.121	0.879	12	186	0.061	0.939	1.000
Northeast Asia	China – Han in Beijing	2	28	73	103	0.019	0.272	0.709	32	174	0.155	0.845	1.000
Northeast Asia	China – Han in South	1	21	86	108	0.009	0.194	0.796	23	193	0.106	0.894	1.000
Northeast Asia	Japan	2	28	74	104	0.019	0.269	0.712	32	176	0.154	0.846	1.000
Northeast Asia	Korea	1	6	28	35	0.029	0.171	0.800	8	62	0.114	0.886	0.362
Southeast Asia	Vietnam	3	22	76	101	0.030	0.218	0.752	28	174	0.139	0.861	0.396
Southeast Asia	Thailand	3	23	53	79	0.038	0.291	0.671	29	129	0.184	0.816	0.715
Southeast Asia	Cambodia	2	21	68	91	0.022	0.231	0.747	25	157	0.137	0.863	0.668
Southeast Asia	Singapore	9	33	54	96	0.094	0.344	0.563	51	141	0.266	0.734	0.293
North America	African ancestry – Southwest USA	1	25	40	66	0.015	0.379	0.606	27	105	0.205	0.795	0.272

North America	Utah – European ancestry	10	41	48	99	0.101	0.414	0.485	61	137	0.308	0.692	0.814
Caribe	African Caribbean in Barbados	2	24	70	96	0.021	0.250	0.729	28	164	0.146	0.854	1.000
Latin America	Colombia – Medelin	5	26	63	94	0.053	0.277	0.670	36	152	0.191	0.809	0.316
Latin America	Mexico – LA, USA	1	26	40	67	0.015	0.388	0.597	28	106	0.209	0.791	0.268
Latin America	Peru – Lima	2	17	67	86	0.023	0.198	0.779	21	151	0.122	0.878	0.604
Latin America	Puerto Rico	3	38	64	105	0.029	0.362	0.610	44	166	0.210	0.790	0.554
Latin America	Brazil	1	9	31	41	0.024	0.220	0.756	11	71	0.134	0.866	1.000

Table A.10– Raw population data for *MICB*–rs3134899 SNP.

		Genotype				Frequency genotype			Allele count		Freq Alelle C	Freq Alelle T	Hardy-Weinberg p - value
Region	Population	CC	CT	TT	N	CC	CT	TT	Allele C	Allele T			
Europe	Portugal	1	18	43	62	0.016	0.290	0.694	20	104	0.161	0.839	1.000
Europe	Spain	3	33	71	107	0.028	0.308	0.664	39	175	0.182	0.818	1.000
Europe	Italy	5	27	76	108	0.046	0.250	0.704	37	179	0.171	0.829	0.301
Europe	France	1	10	17	28	0.036	0.357	0.607	12	44	0.214	0.786	1.000
Europe	UK	8	30	54	92	0.087	0.326	0.587	46	138	0.250	0.750	0.262
Europe	Netherlands	19	203	276	498	0.038	0.408	0.554	241	755	0.242	0.758	0.015
Europe	Finland	12	36	51	99	0.121	0.364	0.515	60	138	0.303	0.697	0.160
Europe	Norway	2	3	13	18	0.111	0.167	0.722	7	29	0.194	0.806	0.086
Europe	Greece	1	2	17	20	0.050	0.100	0.850	4	36	0.100	0.900	0.152
Europe	Hungary	0	4	16	20	0.000	0.200	0.800	4	36	0.100	0.900	1.000
Europe	Russia	1	9	15	25	0.040	0.360	0.600	11	39	0.220	0.780	1.000
Western Asia	Georgia	0	4	16	20	0.000	0.200	0.800	4	36	0.100	0.900	1.000
North Africa	Morocco	0	8	37	45	0.000	0.178	0.822	8	82	0.089	0.911	1.000
North Africa	Algeria–Mozabite	1	11	17	29	0.034	0.379	0.586	13	45	0.224	0.776	1.000
North Africa	Tunisia	0	28	59	87	0.000	0.322	0.678	28	146	0.161	0.839	0.113
Sub–Saharan Africa	Ethiopia	2	12	39	53	0.038	0.226	0.736	16	90	0.151	0.849	0.319
Sub–Saharan Africa	Somalia	0	21	70	91	0.000	0.231	0.769	21	161	0.115	0.885	0.611
Sub–Saharan Africa	Nigeria – Yoruba	2	34	73	109	0.018	0.312	0.670	38	180	0.174	0.826	0.517
Sub–Saharan Africa	Nigeria – Esan	1	21	77	99	0.010	0.212	0.778	23	175	0.116	0.884	1.000
Sub–Saharan Africa	The Gambia	8	39	66	113	0.071	0.345	0.584	55	171	0.243	0.757	0.458
Sub–Saharan Africa	Kenya – Luhya	1	15	85	101	0.010	0.149	0.842	17	185	0.084	0.916	0.520
Sub–Saharan Africa	Sierra Leone	1	22	62	85	0.012	0.259	0.729	24	146	0.141	0.859	1.000

|													
Sub–Saharan Africa	Senegal	1	20	44	65	0.015	0.308	0.677	22	108	0.169	0.831	0.675
Sub–Saharan Africa	Tanzania–Sandawe	1	5	22	28	0.036	0.179	0.786	7	49	0.125	0.875	0.348
Sub–Saharan Africa	Maasai – Kenya	1	21	121	143	0.007	0.147	0.846	23	263	0.080	0.920	1.000
Southwest Asia	Yemen	1	13	58	72	0.014	0.181	0.806	15	129	0.104	0.896	0.557
Southwest Asia	Saudi Arabia	0	31	63	94	0.000	0.330	0.670	31	157	0.165	0.835	0.068
Southwest Asia	Oman	0	17	66	83	0.000	0.205	0.795	17	149	0.102	0.898	0.592
Southwest Asia	UAE	1	31	60	92	0.011	0.337	0.652	33	151	0.179	0.821	0.288
Southwest Asia	Druze	0	7	35	42	0.000	0.167	0.833	7	77	0.083	0.917	1.000
Southwest Asia	Jordania	0	6	14	20	0.000	0.300	0.700	6	34	0.150	0.850	1.000
Southwest Asia	Palestina	2	10	34	46	0.043	0.217	0.739	14	78	0.152	0.848	0.258
Southwest Asia	Syria	0	5	11	16	0.000	0.313	0.688	5	27	0.156	0.844	1.000
Southwest Asia	Turkey	0	3	14	17	0.000	0.176	0.824	3	31	0.088	0.912	1.000
Southwest Asia	Iran	1	14	77	92	0.011	0.152	0.837	16	168	0.087	0.913	0.510
South Asia	Pakistan – Punjavi in Lahore	1	15	80	96	0.010	0.156	0.833	17	175	0.089	0.911	0.540
South Asia	India – Gujarati in Texas	1	25	80	106	0.009	0.236	0.755	27	185	0.127	0.873	1.000
South Asia	India – Telegu in UK	5	21	77	103	0.049	0.204	0.748	31	175	0.150	0.850	0.048
South Asia	Sri Lanka – Tamil in UK	2	25	76	103	0.019	0.243	0.738	29	177	0.141	0.859	1.000
South Asia	Bangladesh – Bengali	1	15	70	86	0.012	0.174	0.814	17	155	0.099	0.901	0.583
Northeast Asia	China – Dai in Xishuangbanna	0	11	88	99	0.000	0.111	0.889	11	187	0.056	0.944	1.000
Northeast Asia	China – Han in Beijing	1	23	79	103	0.010	0.223	0.767	25	181	0.121	0.879	1.000
Northeast Asia	China – Han in South	0	15	93	108	0.000	0.139	0.861	15	201	0.069	0.931	1.000
Northeast Asia	Japan	2	25	77	104	0.019	0.240	0.740	29	179	0.139	0.861	1.000
Northeast Asia	Korea	0	5	30	35	0.000	0.143	0.857	5	65	0.071	0.929	1.000
Southeast Asia	Vietnam	0	17	84	101	0.000	0.168	0.832	17	185	0.084	0.916	1.000
Southeast Asia	Thailand	0	24	55	79	0.000	0.304	0.696	24	134	0.152	0.848	0.715
Southeast Asia	Cambodia	0	13	78	91	0.000	0.143	0.857	13	169	0.071	0.929	1.000
Southeast Asia	Singapore	8	34	54	96	0.083	0.354	0.563	50	142	0.260	0.740	0.432
North America	African ancestry – Southwest USA	1	23	42	66	0.015	0.348	0.636	25	107	0.189	0.811	0.436
North America	Utah – European ancestry	9	33	57	99	0.091	0.333	0.576	51	147	0.258	0.742	0.196
Caribe	African Caribbean in Barbados	2	22	72	96	0.021	0.229	0.750	26	166	0.135	0.865	0.677
Latin America	Colombia – Medelin	5	24	65	94	0.053	0.255	0.691	34	154	0.181	0.819	0.172
Latin America	Mexico – LA, USA	1	17	49	67	0.015	0.254	0.731	19	115	0.142	0.858	1.000
Latin America	Peru – Lima	0	6	80	86	0.000	0.070	0.930	6	166	0.035	0.965	1.000
Latin America	Puerto Rico	3	31	71	105	0.029	0.295	0.676	37	173	0.176	0.824	1.000
Latin America	Brazil	0	8	25	33	0.000	0.242	0.758	8	58	0.121	0.879	1.000
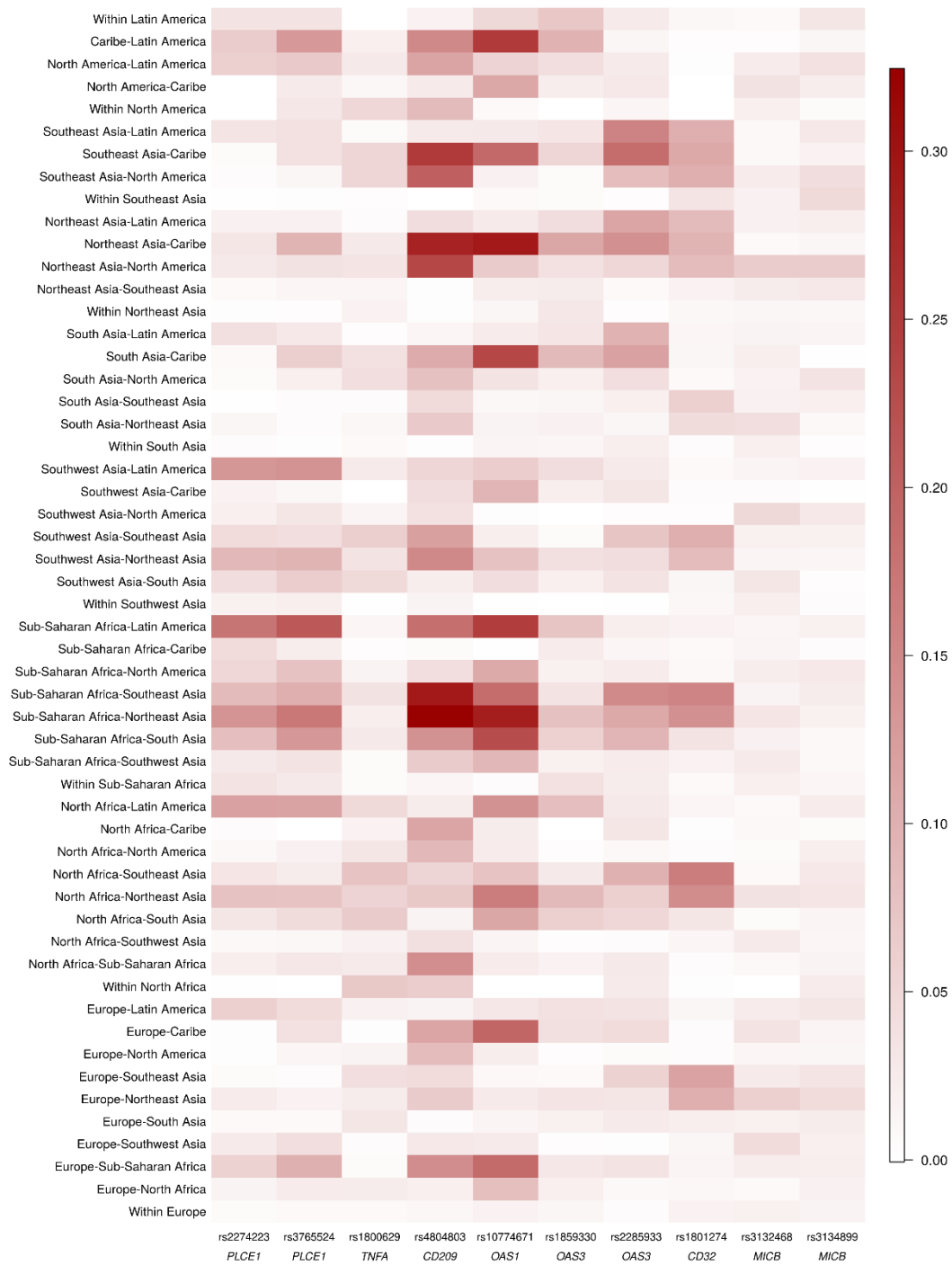
Figure A.1– Heatmap for the average F_{ST} genetic distances within and between population groups (Europe, North Africa, Sub–Saharan Africa, Southwest Asia, South Asia, Northeast Asia, Southeast Asia, North America, Caribe and Latin America) for each of the 10 SNPs.
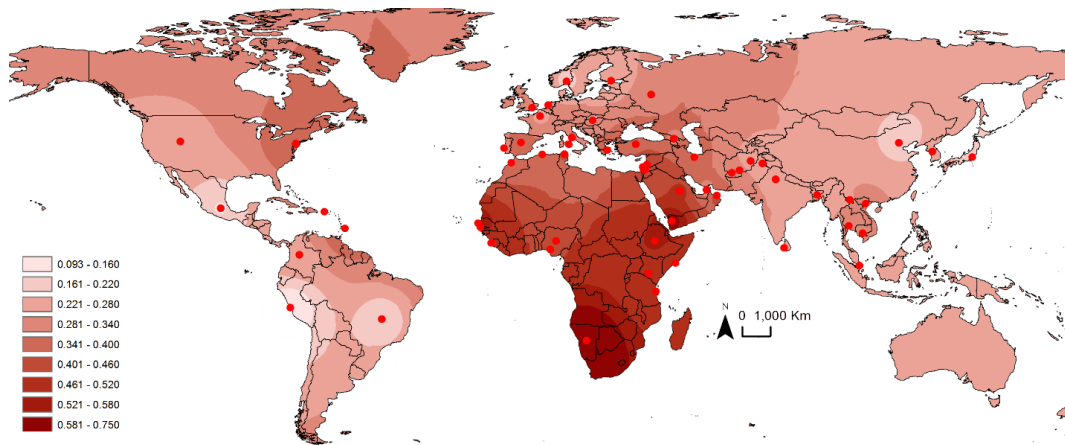
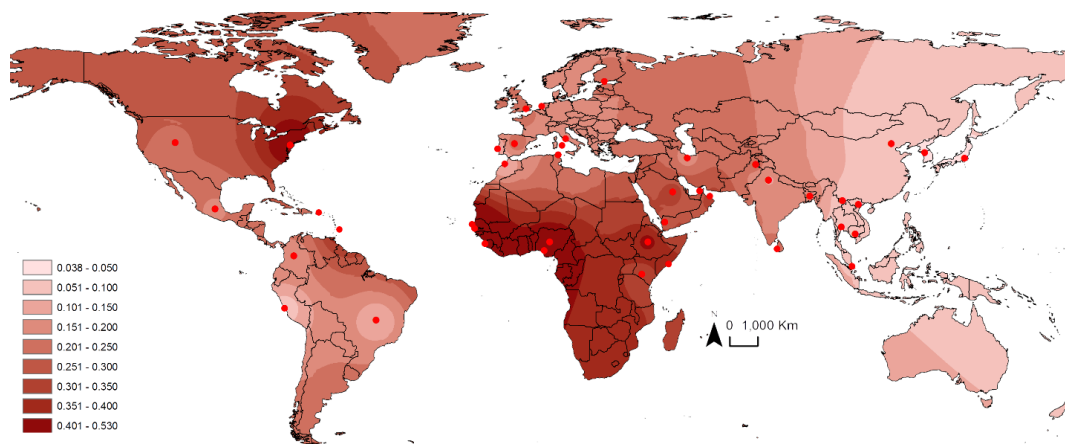**Figure A.2**– Map of allele frequency distribution for *PLCE1*–rs3765524–T.

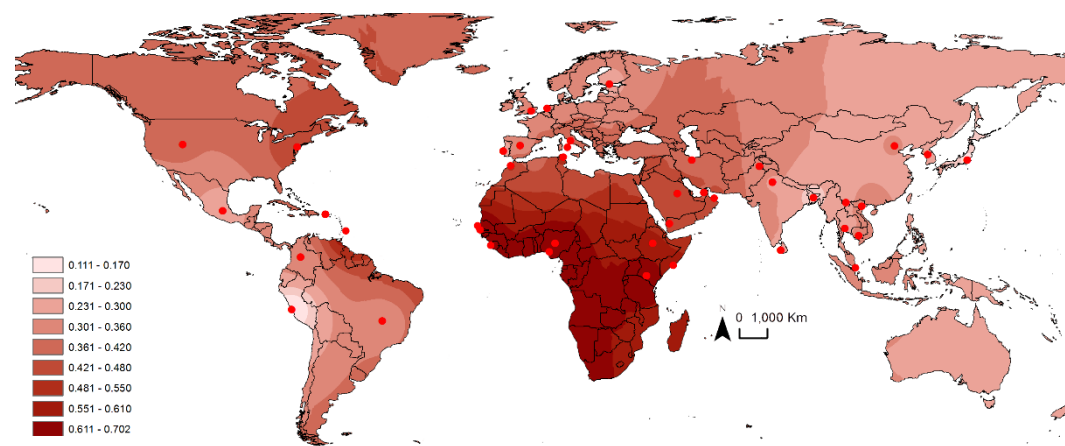**Figure A.3**– Map of allele frequency distribution for *CD209*–rs4804803–G.

Figure A.4– Map of allele frequency distribution for *OAS1*–rs10774671–G.

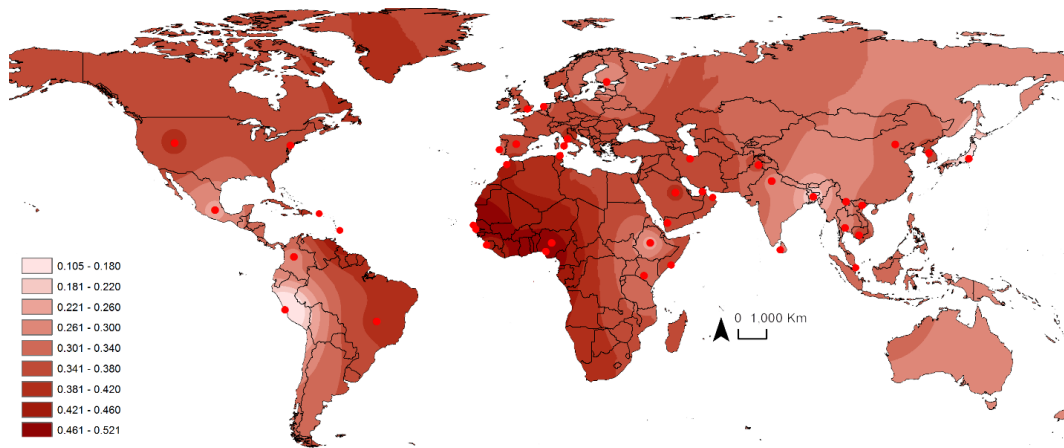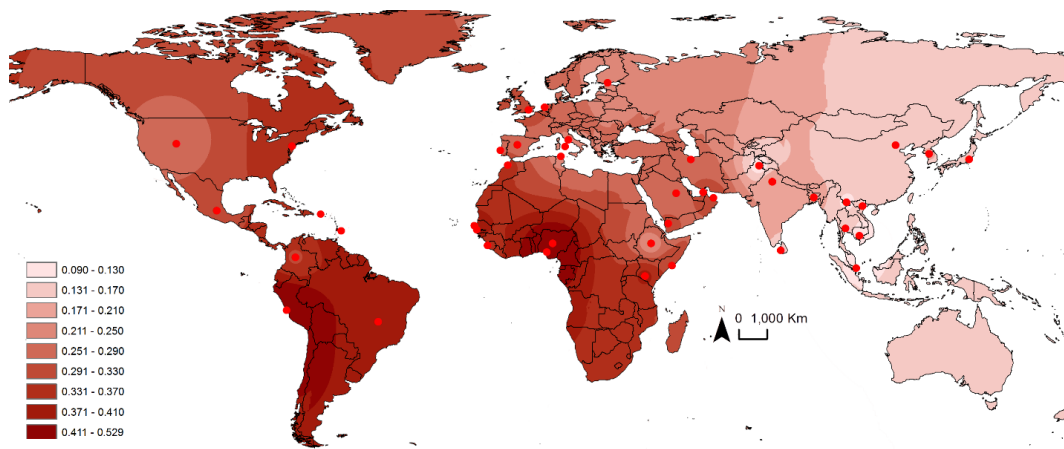**Figure A.5**– Map of allele frequency distribution for, *OAS3*–rs1859330–G.

Figure A.6– Map of allele frequency distribution for *OAS3*–rs2285933–G.
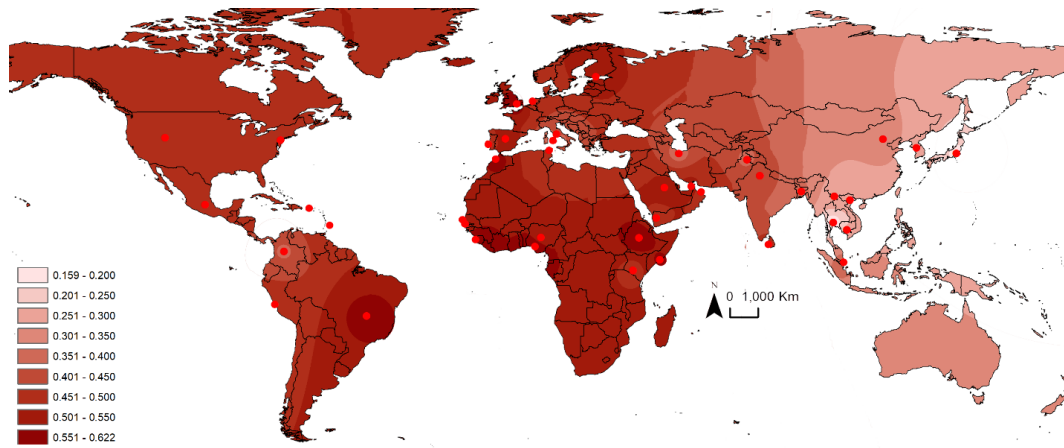

Figure A.7– Map of allele frequency distribution for *CD32*–rs1801274–G.
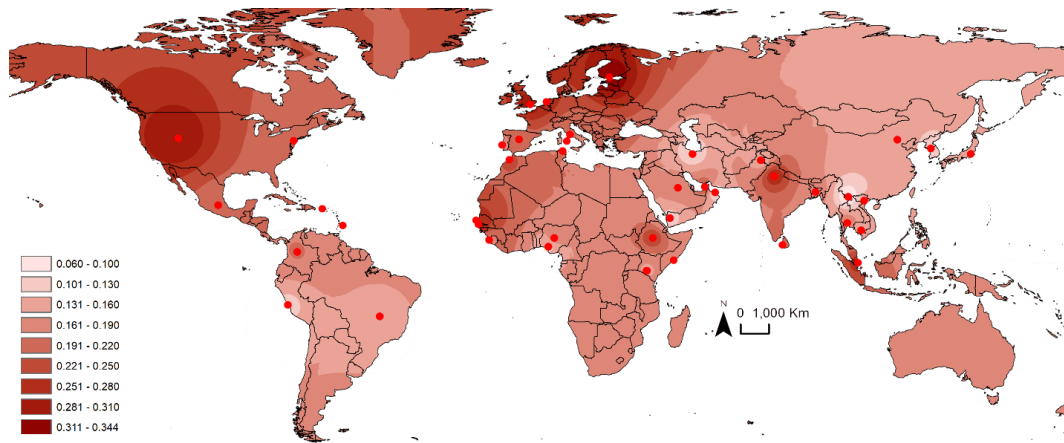
Figure A.8– Map of allele frequency distribution for *MICB*–rs3132468–C.
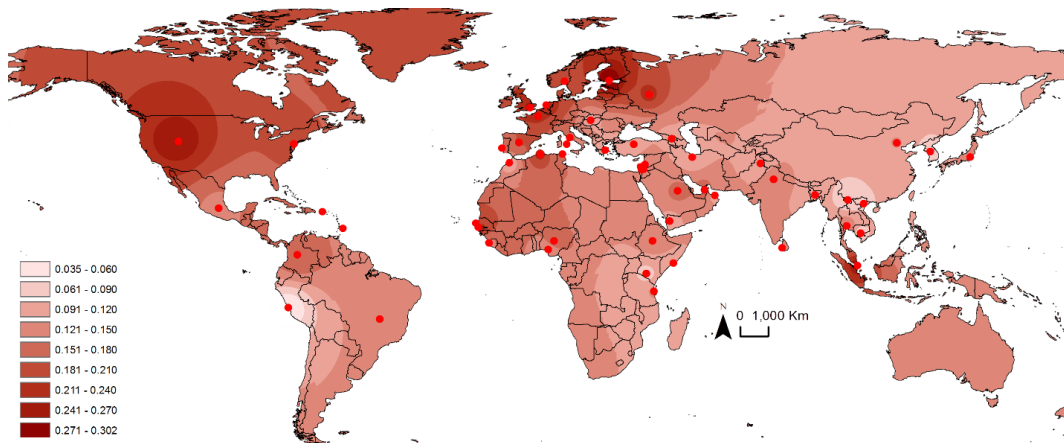
Figure A.9– Map of allele frequency distribution for *MICB*–rs3134899–C.
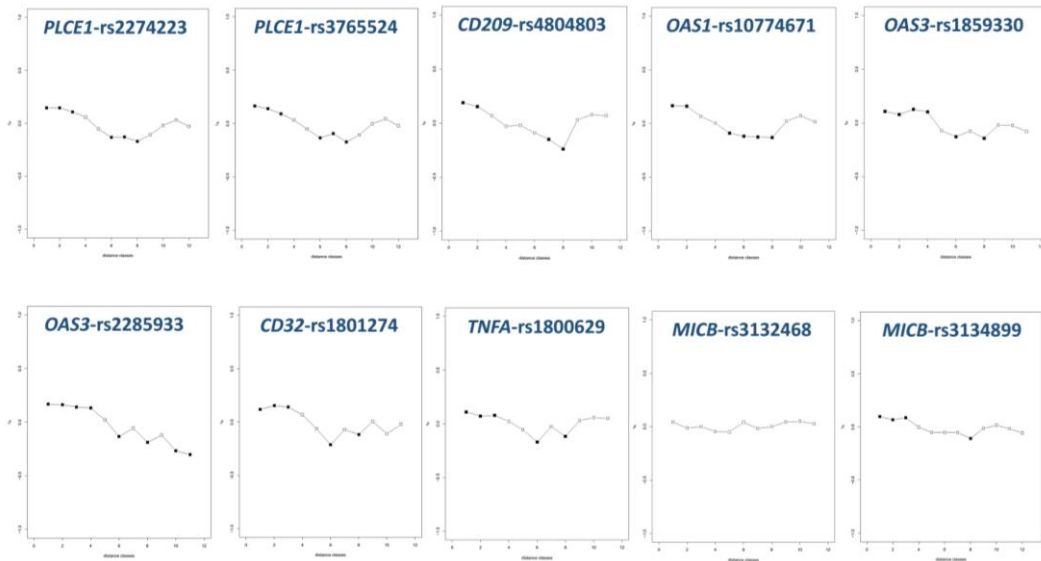
Figure A.10– Mantel correlograms for the 10 studied SNPs. Distance classes are great–circle geographic distances in thousands of km. Filled symbols represent significant p–values while unfilled symbols are non–significant p–values after a progressive Bonferroni correction. rM refers to the vector with the computed Mantel correlations for each distance class.

Figure A.11– Extended Haplotype Homozygosity (EHH) for *PLCE1_*rs2274223. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).

Figure A.12– Extended Haplotype Homozygosity (EHH) for *PLCE1_*rs3765524. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).

Figure A.13– Extended Haplotype Homozygosity (EHH) for *CD209_* rs4804803. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).

Figure A.14– Extended Haplotype Homozygosity (EHH) for *OAS1*–rs10774671. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).
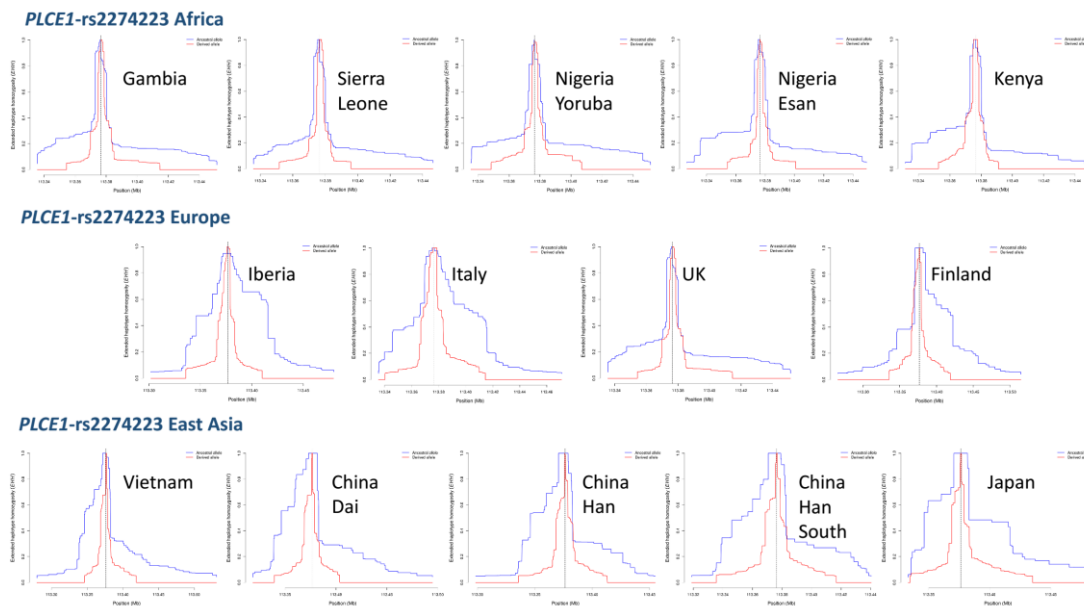

Figure A.15– Extended Haplotype Homozygosity (EHH) for *OAS3*–rs1859330. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).
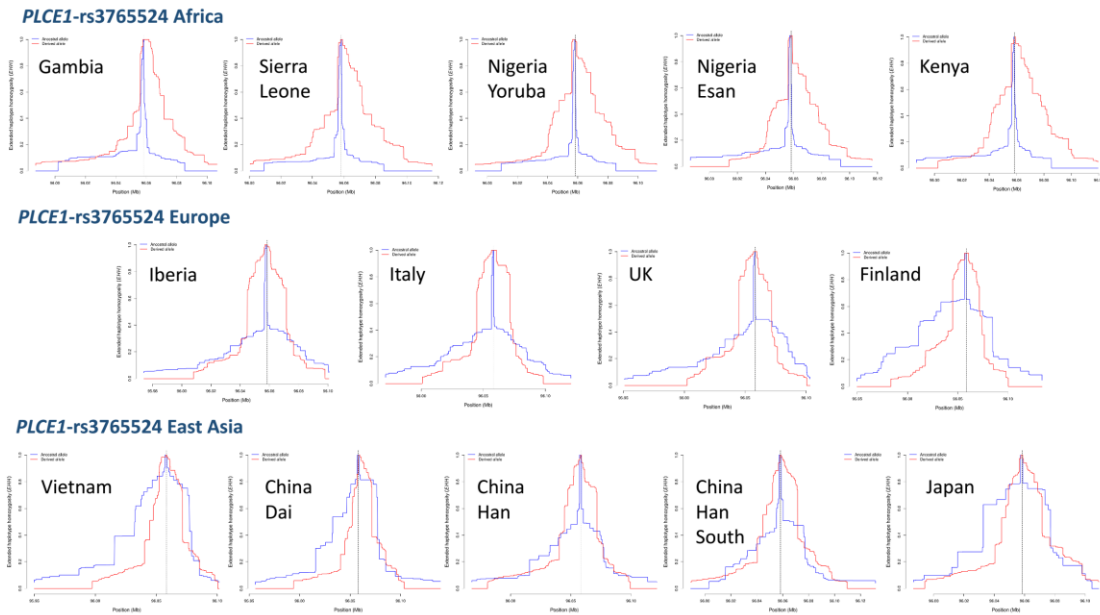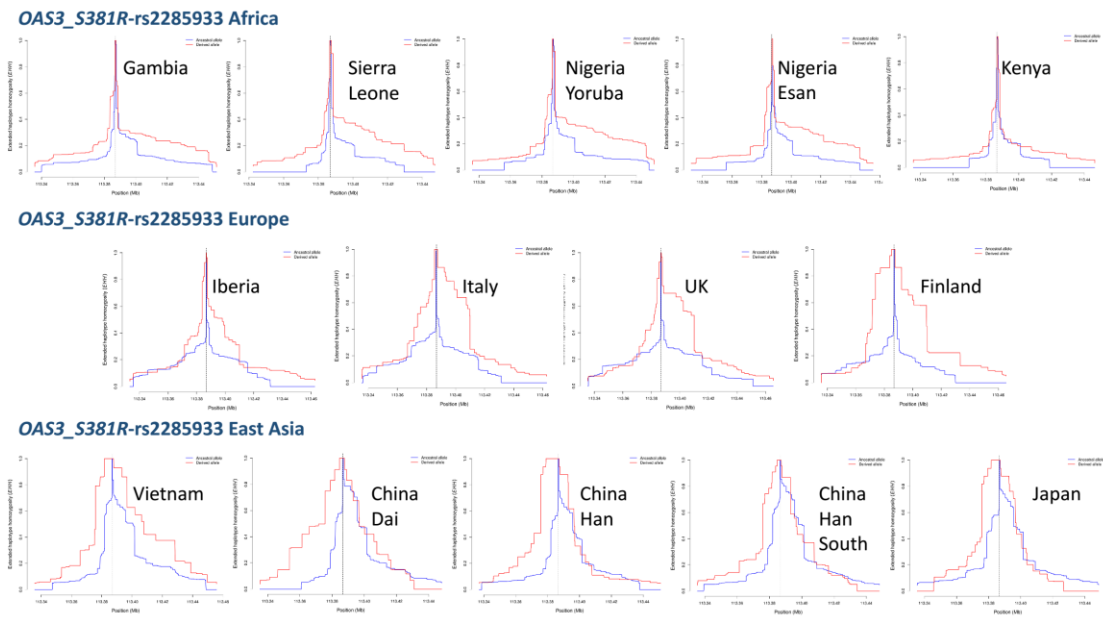
Figure A.16– Extended Haplotype Homozygosity (EHH) for *OAS3*–rs2285933. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).
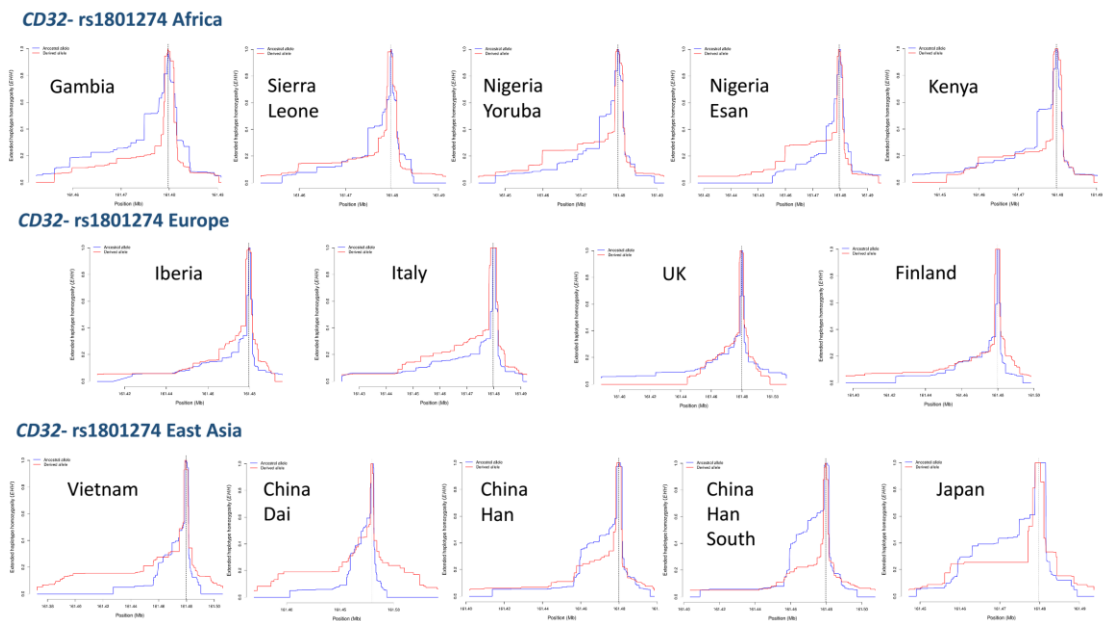
Figure A.17– Extended Haplotype Homozygosity (EHH) for *CD32*–rs1801274. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).
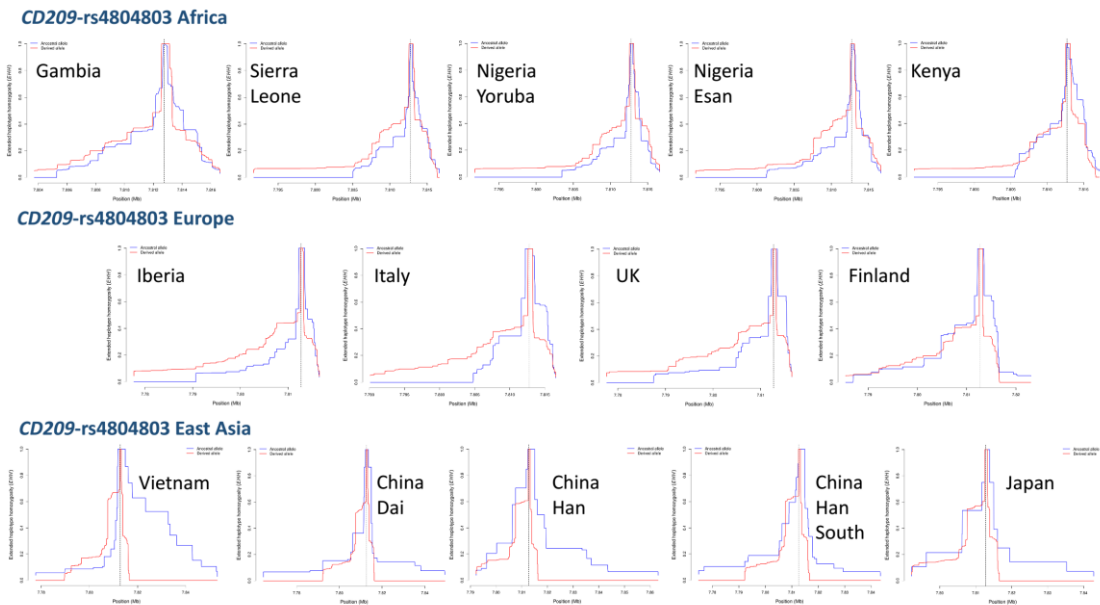
241
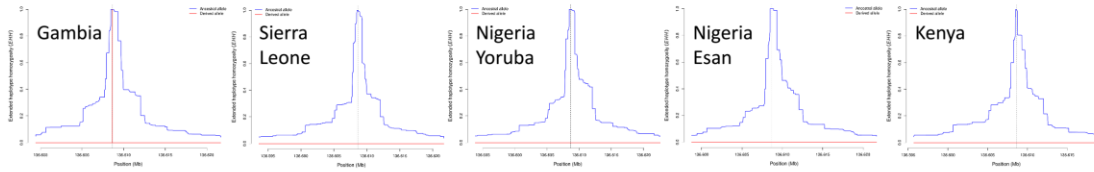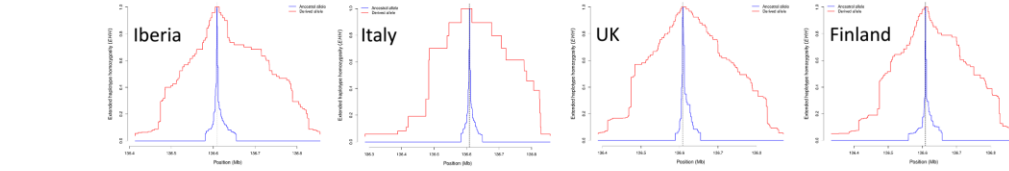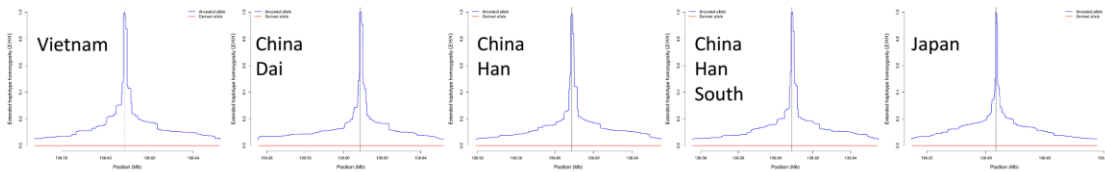
Figure A.18– Extended Haplotype Homozygosity (EHH) for *MCM6*–rs4988235. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).
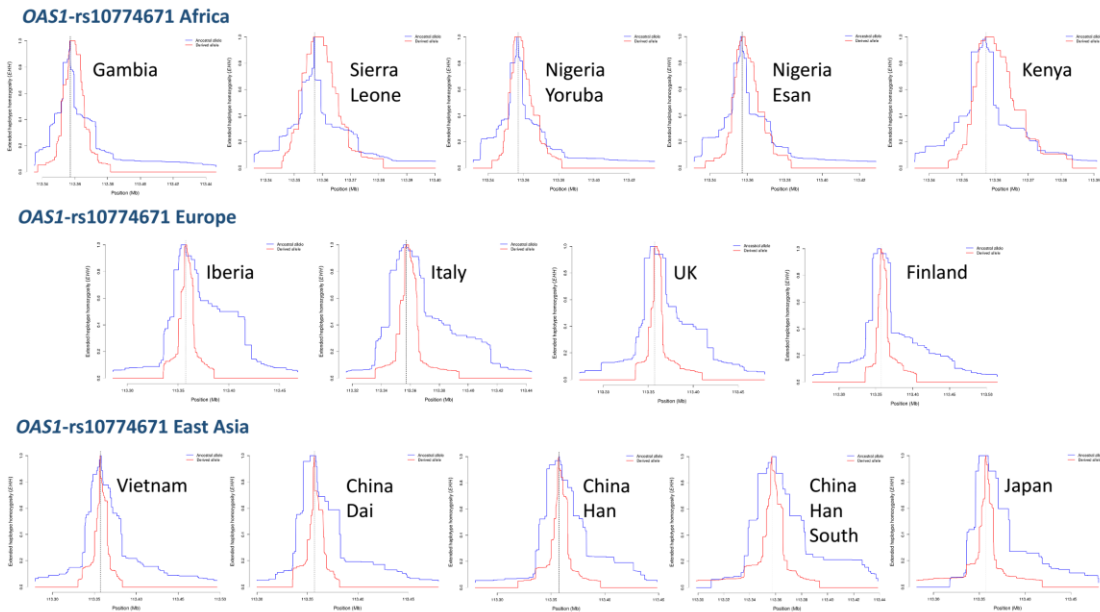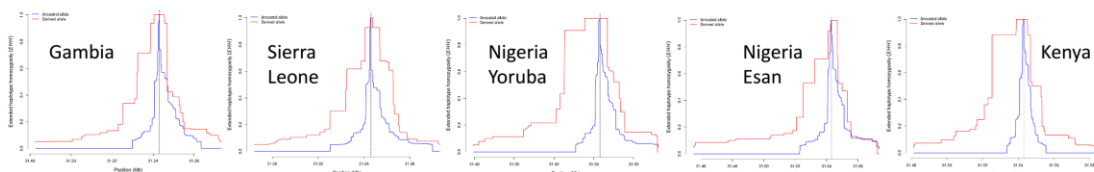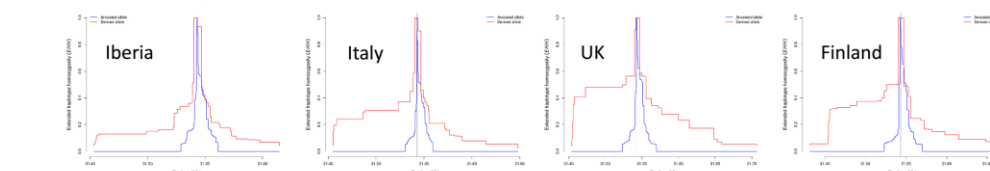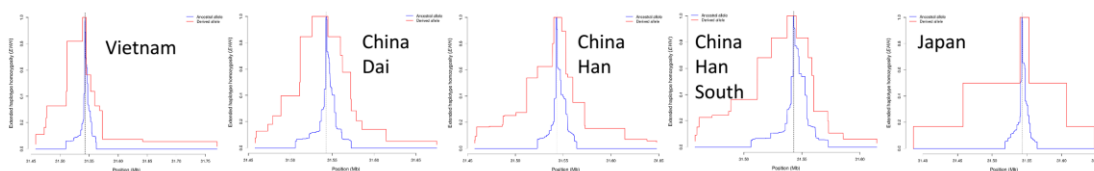

Figure A.19– Extended Haplotype Homozygosity (EHH) for *TNFA*–rs1800629. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).
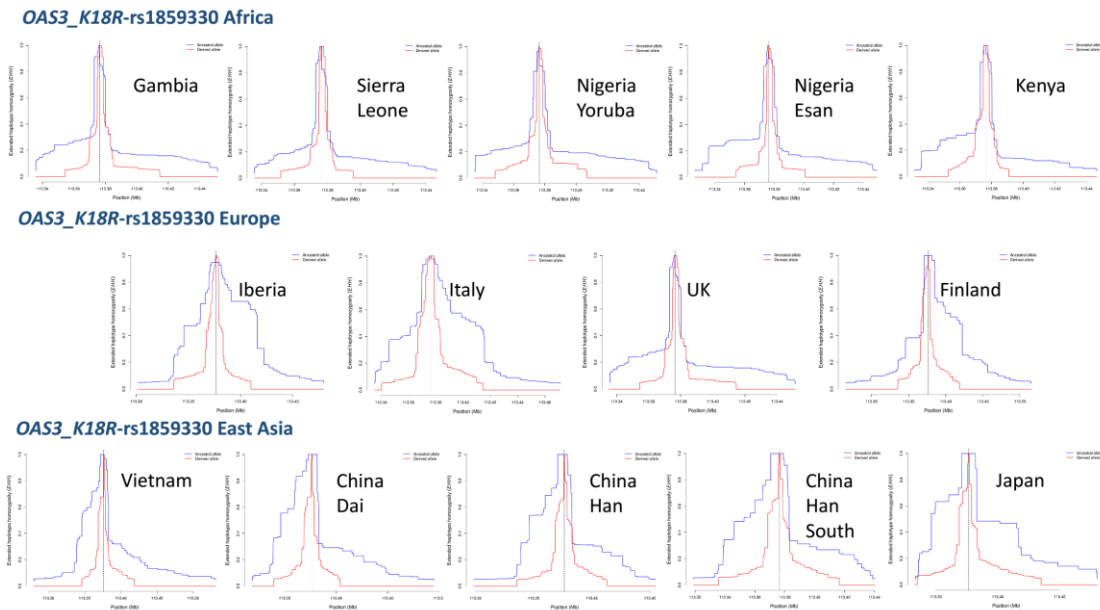
242
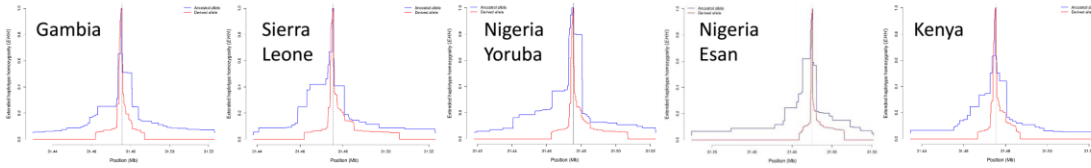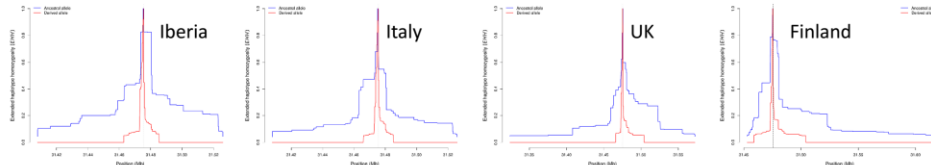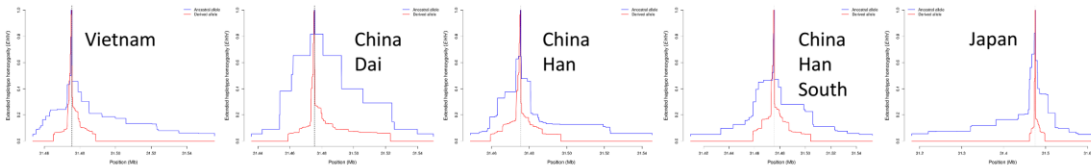
Figure A.20– Extended Haplotype Homozygosity (EHH) for *MICB*–rs3132468. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan).
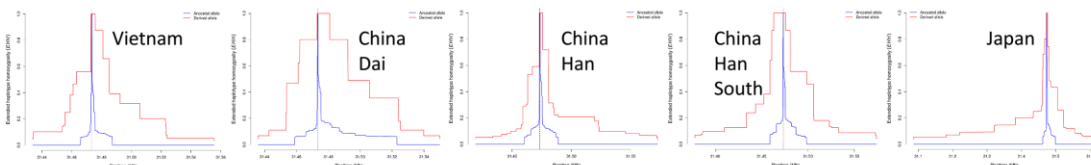
Figure A.21– Extended Haplotype Homozygosity (EHH) for *MICB*–rs3134899. Data for 1000 Genomes database representative populations from the three main human population groups: sub–Saharan Africa (Gambia, Sierra Leone, Nigerian Yoruba, Nigerian Esan and Kenya), Europe (Iberia, Italy, Great Britain, Finland) and East Asia (Vietnam, Chinese Dai, Chinese Han, South Chinese Han and Japan)
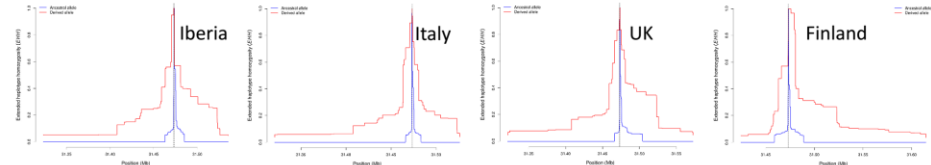
243

Table A.11– Genotypes for the 10 studied SNPs in the dengue and control cohorts obtained in this work and already published.

Gene	SNP ID	Alleles	MAF	Phenotype	Sample size	Population	Reference	Genotypes	Alleles
PLCE1	rs2274223	G/A	G	DSS	3745	Vietnam	Khor et al. 2011	250GG 1436AG 2059AA	1936G 5554A
PLCE1	rs2274223	G/A	G	Control	4952	Vietnam	Khor et al. 2011	450GG 2086AG 2416AA	2986G 6918A
PLCE1	rs2274223	G/A	G	DHF	20	Vietnam	This work	3GG 7AG 10AA	13G 27A
PLCE1	rs2274223	G/A	G	DF	25	Vietnam	This work	3GG 6AG 16AA	12G 38A
PLCE1	rs2274223	G/A	G	Symptomatic	68	Vietnam	This work	2GG 35AG 31AA	39G 97A
PLCE1	rs2274223	G/A	G	Control	76	Vietnam	This work	11GG 28AG 37AA	50G 102A
PLCE1	rs2274223	G/A	G	DSS	159	Thailand	Oliveira et al. 2018	14GG 76AG 69AA	104G 214A
PLCE1	rs2274223	G/A	G	DF	252	Thailand	Oliveira et al. 2018	31GG 110AG 111AA	172G 332A
PLCE1	rs2274223	G/A	G	Control	290	Thailand	Oliveira et al. 2018	30GG 110AG 150AA	170G 410A
PLCE1	rs2274223	G/A	G	DSS	83	Cambodia	This work	7GG 38AG 38AA	52G 114A
PLCE1	rs2274223	G/A	G	DHF	137	Cambodia	This work	17GG 65AG 55AA	99G 175A
PLCE1	rs2274223	G/A	G	DF	103	Cambodia	This work	14GG 51AG 38AA	79G 127A
PLCE1	rs2274223	G/A	G	Control	123	Cambodia	This work	20GG 53AG 50AA	93G 153A
PLCE1	rs2274223	G/A	G	DHF	60	Cuba	Sierra et al. 2017	4GG 29AG 27AA	37G 83A
PLCE1	rs2274223	G/A	G	DF	76	Cuba	Sierra et al. 2017	4GG 38AG 34AA	46G 106A
PLCE1	rs2274223	G/A	G	Control	137	Cuba	Sierra et al. 2017	15GG 61AG 61AA	91G 183A
PLCE1	rs2274223	G/A	G	DF	13	Brazil	This work	1GG 5AG 7AA	7G 19A
PLCE1	rs2274223	G/A	G	Symptomatic	22	Brazil	This work	1GG 6AG 15AA	8G 36A
PLCE1	rs2274223	G/A	G	Control	36	Brazil	This work	10AG 26AA	10G 62A
PLCE1	rs3765524	T/C	T	DSS	3745	Vietnam	Khor et al. 2011	247TT 1430CT 2068CC	1924T 5566C
PLCE1	rs3765524	T/C	T	Control	4952	Vietnam	Khor et al. 2011	449TT 2084CT 2419CC	2982T 6922C
PLCE1	rs3765524	T/C	T	DHF	20	Vietnam	This work	3TT 7CT 10CC	13T 27C
PLCE1	rs3765524	T/C	T	DF	25	Vietnam	This work	3TT 5CT 17CC	11T 39C
PLCE1	rs3765524	T/C	T	Symptomatic	68	Vietnam	This work	2TT 34CT 32CC	38T 98C
PLCE1	rs3765524	T/C	T	Control	76	Vietnam	This work	11TT 28CT 37CC	50T 102C
PLCE1	rs3765524	T/C	T	DSS	76	Thailand	Dang et al. 2014	4TT 28CT 44CC	36T 116C
PLCE1	rs3765524	T/C	T	non–DSS (DF+DHF)	841	Thailand	Dang et al. 2014	84TT 364CT 393CC	532T 1150C
PLCE1	rs3765524	T/C	T	Control	77	Thailand	This work	10TT 25CT 42CC	45T 109C
PLCE1	rs3765524	T/C	T	DSS	83	Cambodia	This work	7TT 37CT 39CC	51T 115C
PLCE1	rs3765524	T/C	T	DHF	137	Cambodia	This work	17TT 65CT 55CC	99T 175C
PLCE1	rs3765524	T/C	T	DF	103	Cambodia	This work	14TT 51CT 38CC	79T 127C
PLCE1	rs3765524	T/C	T	Control	133	Cambodia	This work	24TT 58CT 51CC	106T 160C
PLCE1	rs3765524	T/C	T	DF	13	Brazil	This work	1TT 5CT 7CC	7T 19C

244

PLCE1	rs3765524	T/C	T	Symptomatic	22	Brazil	This work	1TT 7CT 14CC	9T 35C
PLCE1	rs3765524	T/C	T	Control	43	Brazil	This work	1TT 12CT 30CC	14T 72C
PLCE1	rs753724	T/G	T	DSS	3745	Vietnam	Khor et al. 2011	199TT 1328TG 2218GG	1726T 5764G
PLCE1	rs753724	T/G	T	Control	4952	Vietnam	Khor et al. 2011	362TT 1955TG 2635GG	2679T 7225G
PLCE1	rs753724	T/G	T	DHF	20	Vietnam	This work	3TT 7TG 10GG	13T 27G
PLCE1	rs753724	T/G	T	DF	25	Vietnam	This work	3TT 5TG 17GG	11T 39G
PLCE1	rs753724	T/G	T	Symptomatic	68	Vietnam	This work	2TT 33TG 33GG	37T 99G
PLCE1	rs753724	T/G	T	Control	76	Vietnam	This work	10TT 26TG 40GG	46T 106G
PLCE1	rs753724	T/G	T	DSS	83	Cambodia	This work	4TT 33TG 46GG	41T 125G
PLCE1	rs753724	T/G	T	DHF	137	Cambodia	This work	11TT 58TG 68GG	80T 194G
PLCE1	rs753724	T/G	T	DF	103	Cambodia	This work	11TT 45TG 47GG	67T 139G
PLCE1	rs753724	T/G	T	Control	136	Cambodia	This work	5TT 38TG 93GG	48T 224G
PLCE1	rs753724	T/G	T	DF	13	Brazil	This work	1TG 12GG	1T 25G
PLCE1	rs753724	T/G	T	Symptomatic	22	Brazil	This work	1TT 1TG 20GG	3T 41G
PLCE1	rs753724	T/G	T	Control	33	Brazil	This work	33GG	66G
TNFA	rs1800629	G/A	A	DHF	23	Vietnam	This work	19GG 3AG 1AA	5A 41G
TNFA	rs1800629	G/A	A	DF	25	Vietnam	This work	21GG 2AG 2AA	6A 44G
TNFA	rs1800629	G/A	A	Symptomatic	72	Vietnam	This work	65GG 6AG 1AA	8A 136G
TNFA	rs1800629	G/A	A	Control	80	Vietnam	This work	68GG 9AG 3AA	15A 145G
TNFA	rs1800629	G/A	A	DSS	83	Cambodia	This work	66GG 17AG	17A 149G
TNFA	rs1800629	G/A	A	DHF	145	Cambodia	This work	128GG 17AG	17A 273G
TNFA	rs1800629	G/A	A	DF	106	Cambodia	This work	82GG 23AG 1AA	25A 187G
TNFA	rs1800629	G/A	A	Control	135	Cambodia	This work	117GG 18AG	18A 252G
TNFA	rs1800629	G/A	A	DHF	195	Malaysia	Sam et al. 2015	178GG 17AG	17A 373G
TNFA	rs1800629	G/A	A	DF	86	Malaysia	Sam et al. 2015	74GG 17AG 1AA	13A 159G
TNFA	rs1800629	G/A	A	Controls	120	Malaysia	Sam et al. 2015	97GG 22AG 1AA	24A 216G
TNFA	rs1800629	G/A	A	DHF	107	Sri Lanka	Fernando et al. 2015	95GG 10AG 2AA	12A 200G
TNFA	rs1800629	G/A	A	Controls	62	Sri Lanka	Fernando et al. 2015	47GG 13AG 2AA	17A 107G
TNFA	rs1800629	G/A	A	DHF	29	Marathi – Western India	Alagarasu et al. 2013	24GG 5AG	5A 53G
TNFA	rs1800629	G/A	A	DF	85	Marathi – Western India	Alagarasu et al. 2013	80GG 5AG	5A 165G
TNFA	rs1800629	G/A	A	Controls	110	Marathi – Western India	Alagarasu et al. 2013	96GG 13AG	13A 205G
TNFA	rs1800629	G/A	A	DHF	25	Venezuela	Fernandez–Mestre et al. 2004	18GG 7AG	7A 43G
TNFA	rs1800629	G/A	A	DF	41	Venezuela	Fernandez–Mestre et al. 2004	39GG 2AG	2A 80G
TNFA	rs1800629	G/A	A	Control	46	Venezuela	Fernandez–Mestre et al. 2004	40GG 6AG	6A 86G
TNFA	rs1800629	G/A	A	DHF	45	Mexico	García–Trejo et al. 2011	41GG 4AG	4A 86G
TNFA	rs1800629	G/A	A	DF	85	Mexico	García–Trejo et al. 2011	78GG 7AG	7A 163G

TNFA	rs1800629	G/A	A	Controls	163	Mexico	García–Trejo et al. 2011	148GG 14AG 1AA	16A 310G
TNFA	rs1800629	G/A	A	DHF	31	Mexico – Veracruz	Vargas–Castillo et al. 2017	25GG 5AG 1AA	7A 55G
TNFA	rs1800629	G/A	A	DF	138	Mexico – Veracruz	Vargas–Castillo et al. 2017	128GG 10AG	10A 566G
TNFA	rs1800629	G/A	A	Controls	304	Mexico – Veracruz	Vargas–Castillo et al. 2017	275GG 29AG	29A 579G
TNFA	rs1800629	G/A	A	DHF	67	Mexico – Sinaloa	Sanchez–Leyva et al. 2017	58GG 6AG 3AA	12A 122G
TNFA	rs1800629	G/A	A	DF	172	Mexico – Sinaloa	Sanchez–Leyva et al. 2017	155GG 16AG 1AA	18A 326G
TNFA	rs1800629	G/A	A	Controls	257	Mexico – Sinaloa	Sanchez–Leyva et al. 2017	224GG 31AG 2AA	35A 479G
TNFA	rs1800629	G/A	A	DHF	60	Cuba	Sierra et al. 2017	49GG 10AG 1AA	12A 108G
TNFA	rs1800629	G/A	A	DF	77	Cuba	Sierra et al. 2017	65GG 12AG	12A 142G
TNFA	rs1800629	G/A	A	Control	135	Cuba	Sierra et al. 2017	115GG 18AG 2AA	22A 248G
TNFA	rs1800629	G/A	A	DSS	82	Brazil	Xavier–Carvalho et al. 2013	64GG 16AG 2AA	20A 144G
TNFA	rs1800629	G/A	A	Controls (many asymptomatic)	343	Brazil	Xavier–Carvalho et al. 2013	271GG 66AG 6AA	78A 608G
TNFA	rs1800629	G/A	A	DHF	49	Brazil – Arapiraca, Northeast Brazil	dos Santos et al. 2017	39GG 10AG	10A 88G
TNFA	rs1800629	G/A	A	DF	78	Brazil – Arapiraca, Northeast Brazil	dos Santos et al. 2017	63GG 14AG 1AA	16A 140G
TNFA	rs1800629	G/A	A	Controls	135	Brazil – Arapiraca, Northeast Brazil	dos Santos et al. 2017	99GG 33AG 3AA	39A 231G
TNFA	rs1800629	G/A	A	DF	13	Brazil	This work	11GG 2AG	2A 24G
TNFA	rs1800629	G/A	A	Symptomatic	24	Brazil	This work	20GG 4AG	4A 44G
TNFA	rs1800629	G/A	A	Control	44	Brazil	This work	38GG 6AG	6A 82G
DC–SIGN (CD209)	rs4804803	A/G	G	DHF	24	Vietnam	This work	4AG 20AA	4G 44A
DC–SIGN (CD209)	rs4804803	A/G	G	DF	27	Vietnam	This work	6AG 21AA	6G 48G
DC–SIGN (CD209)	rs4804803	A/G	G	Symptomatic	75	Vietnam	This work	1GG 12AG 62AA	14G 136A
DC–SIGN (CD209)	rs4804803	A/G	G	Control	85	Vietnam	This work	1GG 10AG 74AA	12G 158A
DC–SIGN (CD209)	rs4804803	A/G	G	DSS	38	Cambodia	This work	9AG 29AA	9G 67A
DC–SIGN (CD209)	rs4804803	A/G	G	DHF	103	Cambodia	This work	10AG 93AA	10G 196A
DC–SIGN (CD209)	rs4804803	A/G	G	DF	64	Cambodia	This work	7AG 57AA	7G 121A
DC–SIGN (CD209)	rs4804803	A/G	G	Control	120	Cambodia	This work	1GG 20AG 99AA	22G 218A

DC–SIGN (CD209)	rs4804803	A/G	G	DHF	452	Thailand	Sakuntabhai et al. 2005	4GG 97AG 351AA	105G 799A
DC–SIGN (CD209)	rs4804803	A/G	G	DF	752	Thailand	Sakuntabhai et al. 2005	7AG 143AA	7G 293A
DC–SIGN (CD209)	rs4804803	A/G	G	Controls	693	Thailand	Sakuntabhai et al. 2005	9GG 126AG 558AA	144G 1242A
DC–SIGN (CD209)	rs4804803	A/G	G	DHF	507	Thailand	Dang et al. 2016	4GG 92AG 411AA	100G 914A
DC–SIGN (CD209)	rs4804803	A/G	G	DF	408	Thailand	Dang et al. 2016	8GG 77AG 323AA	93G 723A
DC–SIGN (CD209)	rs4804803	A/G	G	Controls	71	Thailand	This work	1GG 8AG 62AA	10G 132A
DC–SIGN (CD209)	rs4804803	A/G	G	DHF	135	Taiwan	Wang et al. 2011	2GG 29AG 104AA	33G 237A
DC–SIGN (CD209)	rs4804803	A/G	G	DF	176	Taiwan	Wang et al. 2011	19AG 157AA	19G 333A
DC–SIGN (CD209)	rs4804803	A/G	G	Controls	120	Taiwan	Wang et al. 2011	9AG 111AA	9G 231A
DC–SIGN (CD209)	rs4804803	A/G	G	DHF	29	Western Indian	Alagarasu et al. 2013c	10AG 19AA	10G 48A
DC–SIGN (CD209)	rs4804803	A/G	G	DF	83	Western Indian	Alagarasu et al. 2013c	19AG 64AA	19G 147A
DC–SIGN (CD209)	rs4804803	A/G	G	Controls	104	Western Indian	Alagarasu et al. 2013c	2GG 29AG 73AA	33G 175A
DC–SIGN (CD209)	rs4804803	A/G	G	DSS	85	Brazil	Xavier–Carvalho et al. 2013	1GG 33AG 51AA	35G 135A
DC–SIGN (CD209)	rs4804803	A/G	G	Controls (many asymptomatic)	322	Brazil	Xavier–Carvalho et al. 2013	28GG 122AG 172AA	178G 466A
DC–SIGN (CD209)	rs4804803	A/G	G	DHF	12	Pará, Brazil	Oliveira et al. 2014	4AG 8AA	4G 20A
DC–SIGN (CD209)	rs4804803	A/G	G	DF	156	Pará, Brazil	Oliveira et al. 2014	6GG 41AG 109AA	53G 259A
DC–SIGN (CD209)	rs4804803	A/G	G	Controls	72	Pará, Brazil	Oliveira et al. 2014	3GG 19AG 50AA	25G 119A
DC–SIGN (CD209)	rs4804803	A/G	G	DF	16	Brazil	This work	3AG 13AA	3G 29A
DC–SIGN (CD209)	rs4804803	A/G	G	Symptomatic	35	Brazil	This work	4GG 4AG 27AA	12G 58A
DC–SIGN (CD209)	rs4804803	A/G	G	Control	29	Brazil	This work	5AG 24AA	5G 53A
OAS1_splicing	rs10774671	A/G	G	DHF	7	Vietnam	This work	5AA 1AG 1GG	3G 11A
OAS1_splicing	rs10774671	A/G	G	DF	14	Vietnam	This work	6AA 8AG	8G 20A
OAS1_splicing	rs10774671	A/G	G	Symptomatic	28	Vietnam	This work	20AA 7AG 1GG	9G 47A
OAS1_splicing	rs10774671	A/G	G	Control	31	Vietnam	This work	15AA 15AG 1GG	17G 45A
OAS1_splicing	rs10774671	A/G	G	DSS	53	Cambodia	This work	25AA 22AG 6GG	34G 72A

OAS1_splicing	rs10774671	A/G	G	DHF	49	Cambodia	This work	18AA 25AG 6GG	37G 61A
OAS1_splicing	rs10774671	A/G	G	DF	53	Cambodia	This work	26AA 24AG 3GG	30G 76A
OAS1_splicing	rs10774671	A/G	G	Control	129	Cambodia	This work	57AA 49AG 23GG	95G 163A
OAS1_splicing	rs10774671	A/G	G	DSS	132	Thailand	Simon–Loriere et al. 2015	58AA 61AG 13GG	87G 177A
OAS1_splicing	rs10774671	A/G	G	DHF	361	Thailand	Simon–Loriere et al. 2015	196AA 142AG 23GG	188G 534A
OAS1_splicing	rs10774671	A/G	G	DF	234	Thailand	Simon–Loriere et al. 2015	132AA 87AG 15GG	117G 351A
OAS1_splicing	rs10774671	A/G	G	Control	77	Thailand	This work	43AA 29AG 5GG	39G 115A
OAS1_splicing	rs10774671	A/G	G	DHF	29	Pune, Maharashtra, India	Alagarasu et al. 2013	10AA 15AG 4GG	24G 35A
OAS1_splicing	rs10774671	A/G	G	DF	29	Pune, Maharashtra, India	Alagarasu et al. 2013	27AA 41AG 12GG	65G 95A
OAS1_splicing	rs10774671	A/G	G	Control	29	Pune, Maharashtra, India	Alagarasu et al. 2013	45AA 47AG 13GG	73G 137A
OAS1_splicing	rs10774671	A/G	G	DHF	59	Cuba	Sierra et al. 2017	18AA 31AG 10GG	51G 67A
OAS1_splicing	rs10774671	A/G	G	DF	77	Cuba	Sierra et al. 2017	28AA 33AG 16GG	65G 89A
OAS1_splicing	rs10774671	A/G	G	Control	136	Cuba	Sierra et al. 2017	50AA 59AG 27GG	113G 159A
OAS3_K18R	rs1859330	G/A	G	DHF	7	Vietnam	This work	1GG 1AG 5AA	3G 11A
OAS3_K18R	rs1859330	G/A	G	DF	14	Vietnam	This work	9AG 5AA	9G 19A
OAS3_K18R	rs1859330	G/A	G	Symptomatic	28	Vietnam	This work	1GG 7AG 20AA	9G 47A
OAS3_K18R	rs1859330	G/A	G	Control	31	Vietnam	This work	1GG 15AG 15AA	17G 45A
OAS3_K18R	rs1859330	G/A	G	DSS	52	Cambodia	This work	5GG 21AG 26AA	31G 73A
OAS3_K18R	rs1859330	G/A	G	DHF	50	Cambodia	This work	6GG 23AG 21AA	35G 65A
OAS3_K18R	rs1859330	G/A	G	DF	55	Cambodia	This work	4GG 24AG 27AA	32G 78A
OAS3_K18R	rs1859330	G/A	G	Control	130	Cambodia	This work	23GG 58AG 49AA	104G 156A
OAS3_K18R	rs1859330	G/A	G	DSS	135	Thailand	Simon–Loriere et al. 2015	14GG 64AG 57AA	92G 178A
OAS3_K18R	rs1859330	G/A	G	DHF	355	Thailand	Simon–Loriere et al. 2015	29GG 144AG 182AA	202G 508A
OAS3_K18R	rs1859330	G/A	G	DF	235	Thailand	Simon–Loriere et al. 2015	24GG 91AG 120AA	139G 331A
OAS3_K18R	rs1859330	G/A	G	Controls	81	Thailand	This work	5GG 32AG 44AA	42G 120A
OAS3_S381R	rs2285933	G/C	G	DHF	7	Vietnam	This work	2CG 5CC	2G 12C
OAS3_S381R	rs2285933	G/C	G	DF	14	Vietnam	This work	4CG 10CC	4G 24C
OAS3_S381R	rs2285933	G/C	G	Symptomatic	28	Vietnam	This work	8CG 20CC	8G 48C
OAS3_S381R	rs2285933	G/C	G	Control	31	Vietnam	This work	10CG 21CC	10G 52C
OAS3_S381R	rs2285933	G/C	G	DSS	52	Cambodia	This work	9CG 43CC	9G 95C
OAS3_S381R	rs2285933	G/C	G	DHF	48	Cambodia	This work	1GG 13CG 34CC	15G 81C
OAS3_S381R	rs2285933	G/C	G	DF	55	Cambodia	This work	11CG 44CC	11G 99C
OAS3_S381R	rs2285933	G/C	G	Control	128	Cambodia	This work	2GG 21CG 105CC	25G 231C
OAS3_S381R	rs2285933	G/C	G	DSS	135	Thailand	Simon–Loriere et al. 2015	2GG 17CG 116CC	21G 249C
OAS3_S381R	rs2285933	G/C	G	DHF	364	Thailand	Simon–Loriere et al. 2015	7GG 101CG 256CC	115G 613C
OAS3_S381R	rs2285933	G/C	G	DF	238	Thailand	Simon–Loriere et al. 2015	7GG 71CG 160CC	85G 391C
OAS3_S381R	rs2285933	G/C	G	Controls	77	Thailand	This work	1GG 23CG 53CC	25G 129C

Gene	SNP	Alleles	Ancestral	Group	N	Country	Reference	Genotype counts	Allele counts
FCgRIIa (CD32)	rs1801274	A/G	G	DSS	3745	Vietnam	Khor et al. 2011	327GG 1559AG 1859AA	2213G 5277A
FCgRIIa (CD32)	rs1801274	A/G	G	Control	4952	Vietnam	Khor et al. 2011	467GG 2107AG 2378AA	3041G 6863A
FCgRIIa (CD32)	rs1801274	A/G	G	DHF	302	Vietnam	Loke et al 2002	19GG 118AG 165AA	157G 448A
FCgRIIa (CD32)	rs1801274	A/G	G	Control	238	Vietnam	Loke et al 2002	25GG 94AG 119AA	144G 332A
FCgRIIa (CD32)	rs1801274	A/G	G	DHF	23	Vietnam	This work	4GG 12AG 7AA	20G 26A
FCgRIIa (CD32)	rs1801274	A/G	G	DF	25	Vietnam	This work	2GG 13AG 10AA	17G 33A
FCgRIIa (CD32)	rs1801274	A/G	G	Symptomatic	72	Vietnam	This work	7GG 35AG 30AA	49G 95A
FCgRIIa (CD32)	rs1801274	A/G	G	Control	82	Vietnam	This work	15GG 40AG 27AA	70G 94A
FCgRIIa (CD32)	rs1801274	A/G	G	DSS	84	Cambodia	This work	6GG 39AG 39AA	51G 117A
FCgRIIa (CD32)	rs1801274	A/G	G	DHF	143	Cambodia	This work	11GG 63AG 69AA	85G 201A
FCgRIIa (CD32)	rs1801274	A/G	G	DF	103	Cambodia	This work	6GG 39AG 58AA	51G 155A
FCgRIIa (CD32)	rs1801274	A/G	G	Control	122	Cambodia	This work	7GG 46AG 69AA	60G 184A
FCgRIIa (CD32)	rs1801274	A/G	G	DSS	159	Thailand	Oliveira et al. 2018	11GG 59AG 89AA	81G 237A
FCgRIIa (CD32)	rs1801274	A/G	G	DF	252	Thailand	Oliveira et al. 2018	17GG 102AG 133AA	136G 368A
FCgRIIa (CD32)	rs1801274	A/G	G	Control	289	Thailand	Oliveira et al. 2018	15GG 112AG 162AA	142G 436A
FCgRIIa (CD32)	rs1801274	A/G	G	DHF	29	Cuba	Garcia et al. 2010	6GG 10AG 13AA	22G 36A
FCgRIIa (CD32)	rs1801274	A/G	G	DF	68	Cuba	Garcia et al. 2010	18GG 33AG 17AA	69G 67A
FCgRIIa (CD32)	rs1801274	A/G	G	Control (Subclinical)	42	Cuba	Garcia et al. 2010	16GG 23AG 3AA	55G 29A
FCgRIIa (CD32)	rs1801274	A/G	G	DHF	60	Cuba	Sierra et al. 2017	14GG 25AG 21AA	53G 67A
FCgRIIa (CD32)	rs1801274	A/G	G	DF	77	Cuba	Sierra et al. 2017	19GG 42AG 16AA	80G 74A
FCgRIIa (CD32)	rs1801274	A/G	G	Control	137	Cuba	Sierra et al. 2017	37GG 74AG 26AA	148G 126A
FCgRIIa (CD32)	rs1801274	A/G	G	DHF	31	Mexico – Veracruz	Vargas–Castillo et al. 2017	10GG 13AG 8AA	33G 29A
FCgRIIa (CD32)	rs1801274	A/G	G	DF	138	Mexico – Veracruz	Vargas–Castillo et al. 2017	34GG 73AG 31AA	141G 135A
FCgRIIa (CD32)	rs1801274	A/G	G	Controls	304	Mexico – Veracruz	Vargas–Castillo et al. 2017	88GG 145AG 71AA	321G 287A
FCgRIIa (CD32)	rs1801274	A/G	G	DF	10	Brazil	This work	4GG 5AG 1AA	13G 7A
FCgRIIa (CD32)	rs1801274	A/G	G	Symptomatic	22	Brazil	This work	6GG 13AG 3AA	25G 19A
FCgRIIa (CD32)	rs1801274	A/G	G	Control	29	Brazil	This work	7GG 17AG 5AA	31G 27A
MICB	rs3132468	C/T	C	DSS	3745	Vietnam	Khor et al. 2011	108CC 1055CT 2582TT	1271C 6219T
MICB	rs3132468	C/T	C	Control	4952	Vietnam	Khor et al. 2011	88CC 1142CT 3722TT	1318C 8586T
MICB	rs3132468	C/T	C	DSS	297	Vietnam	Whitehorn et al. 2013	9CC 88CT 200TT	106C 488T
MICB	rs3132468	C/T	C	DF (All mild cases)	3500	Vietnam	Whitehorn et al. 2013	80CC 897CT 2523TT	1057C 5943T
MICB	rs3132468	C/T	C	Control	1068	Vietnam	Whitehorn et al. 2013	19CC 248CT 801TT	286C 1850T
MICB	rs3132468	C/T	C	DHF	23	Vietnam	This work	1CT 22TT	1C 45T
MICB	rs3132468	C/T	C	DF	25	Vietnam	This work	1CC 6CT 18TT	8C 42T
MICB	rs3132468	C/T	C	Symptomatic	74	Vietnam	This work	6CC 19CT 49TT	31C 117T

| MICB	rs3132468	C/T	C	Control	81	Vietnam	This work	3CC 15CT 63TT	21C 141T
MICB	rs3132468	C/T	C	DSS	83	Cambodia	This work	3CC 18CT 62TT	24C 142T
MICB	rs3132468	C/T	C	DHF	142	Cambodia	This work	4CC 31CT 107TT	39C 245T
MICB	rs3132468	C/T	C	DF	105	Cambodia	This work	1CC 25CT 79TT	27C 183T
MICB	rs3132468	C/T	C	Control	134	Cambodia	This work	3CC 30CT 101TT	36C 232T
MICB	rs3132468	C/T	C	DSS	76	Thailand	Dang et al. 2014	3CC 26CT 47TT	32C 120T
MICB	rs3132468	C/T	C	non-DSS (DF+DHF)	841	Thailand	Dang et al. 2014	19CC 213CT 609TT	251C 1431T
MICB	rs3132468	C/T	C	Control	79	Thailand	This work	3CC 23CT 53TT	29C 129T
MICB	rs3132468	C/T	C	DSS	159	Thailand	Oliveira et al. 2018	3CC 44CT 112TT	50C 268T
MICB	rs3132468	C/T	C	DF	252	Thailand	Oliveira et al. 2018	2CC 78CT 172TT	82C 422T
MICB	rs3132468	C/T	C	Control	290	Thailand	Oliveira et al. 2018	6CC 76CT 208TT	88C 492T
MICB	rs3132468	C/T	C	DHF	59	Cuba	Sierra et al. 2017	3CC 21CT 35TT	27C 81T
MICB	rs3132468	C/T	C	DF	76	Cuba	Sierra et al. 2017	4CC 20CT 52TT	28C 124T
MICB	rs3132468	C/T	C	Control	137	Cuba	Sierra et al. 2017	2CC 48CT 87TT	52C 222T
MICB	rs3132468	C/T	C	DF	13	Brazil	This work	3CT 10TT	3C 23T
MICB	rs3132468	C/T	C	Symptomatic	23	Brazil	This work	1CC 3CT 19TT	5C 41T
MICB	rs3132468	C/T	C	Control	44	Brazil	This work	1CC 9CT 34TT	11C 77T
MICB	rs3134899	C/T	C	DSS	3745	Vietnam	Khor et al. 2011	56CC 802CT 2887TT	914C 6576T
MICB	rs3134899	C/T	C	Control	4952	Vietnam	Khor et al. 2011	49CC 883CT 4020TT	981C 8923T
MICB	rs3134899	C/T	C	DHF	23	Vietnam	This work	1CT 22TT	1C 45T
MICB	rs3134899	C/T	C	DF	25	Vietnam	This work	5CT 20TT	5C 45T
MICB	rs3134899	C/T	C	Symptomatic	74	Vietnam	This work	17CT 57TT	17C 131T
MICB	rs3134899	C/T	C	Control	81	Vietnam	This work	1CC 13CT 67TT	15C 147T
MICB	rs3134899	C/T	C	DSS	83	Cambodia	This work	4CC 13CT 66TT	21C 145T
MICB	rs3134899	C/T	C	DHF	142	Cambodia	This work	2CC 26CT 114TT	30C 254T
MICB	rs3134899	C/T	C	DF	105	Cambodia	This work	15CT 90TT	15C 195T
MICB	rs3134899	C/T	C	Control	134	Cambodia	This work	1CC 20CT 113TT	22C 246T
MICB	rs3134899	C/T	C	DSS	159	Thailand	Oliveira et al. 2018	2CC 39CT 118TT	43C 275T
MICB	rs3134899	C/T	C	DF	252	Thailand	Oliveira et al. 2018	3CC 61CT 188TT	67C 437T
MICB	rs3134899	C/T	C	Control	290	Thailand	Oliveira et al. 2018	2CC 59CT 229TT	63C 517T
MICB	rs3134899	C/T	C	DHF	60	Cuba	Sierra et al. 2017	2CC 18CT 40TT	22C 98T
MICB	rs3134899	C/T	C	DF	77	Cuba	Sierra et al. 2017	2CC 22CT 53TT	26C 128T
MICB	rs3134899	C/T	C	Control	137	Cuba	Sierra et al. 2017	2CC 35CT 100TT	39C 235T
MICB	rs3134899	C/T	C	DF	13	Brazil	This work	3CT 10TT	3C 23T
MICB	rs3134899	C/T	C	Symptomatic	23	Brazil	This work	1CC 3CT 19TT	5C 41T
MICB	rs3134899	C/T	C	Control	36	Brazil	This work	8CT 28TT	8C 64T
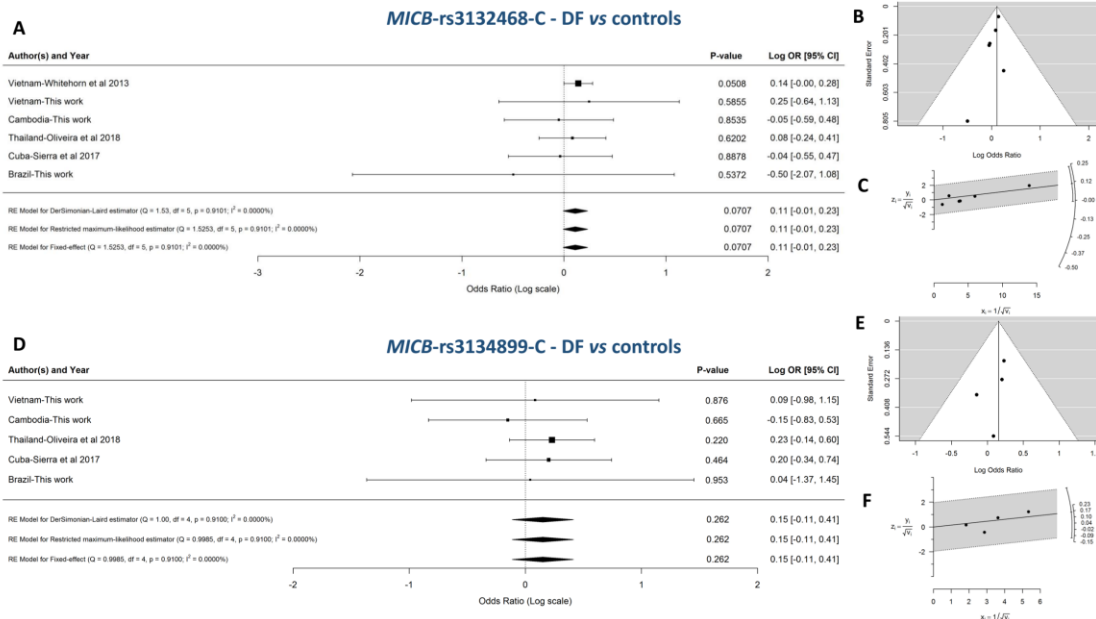
Figure A.22– Meta–analysis in *MICB*–rs3132468–C (A–C) and *MICB*–rs3134899–C (D–F) for DF *vs* controls comparison. (A,D) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and *I²* values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E) Funnel plots. (C,F) Radial plots.
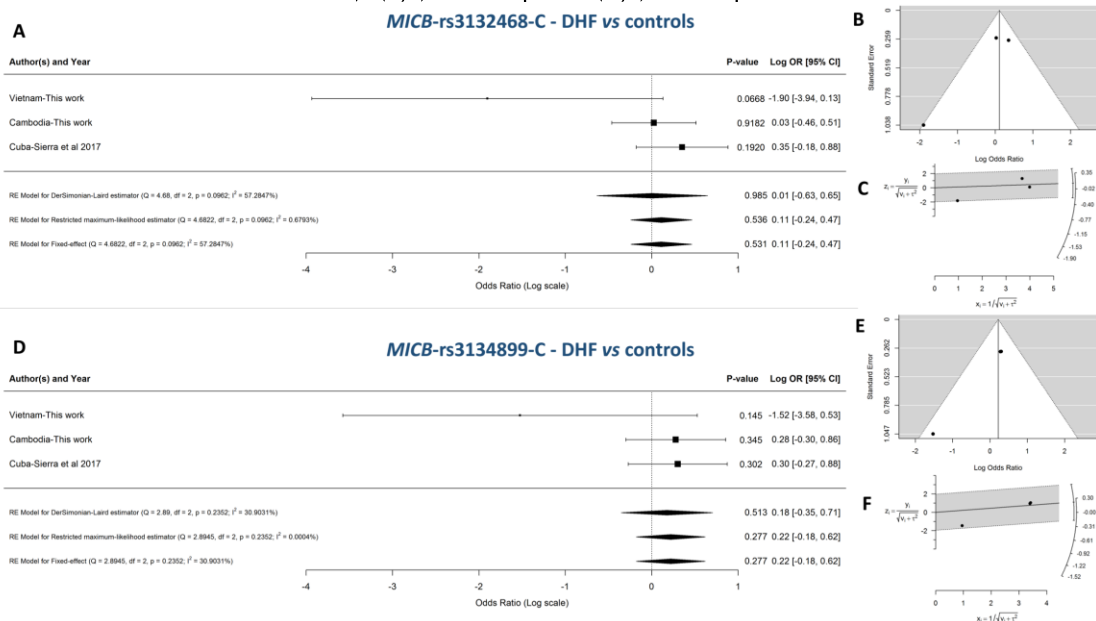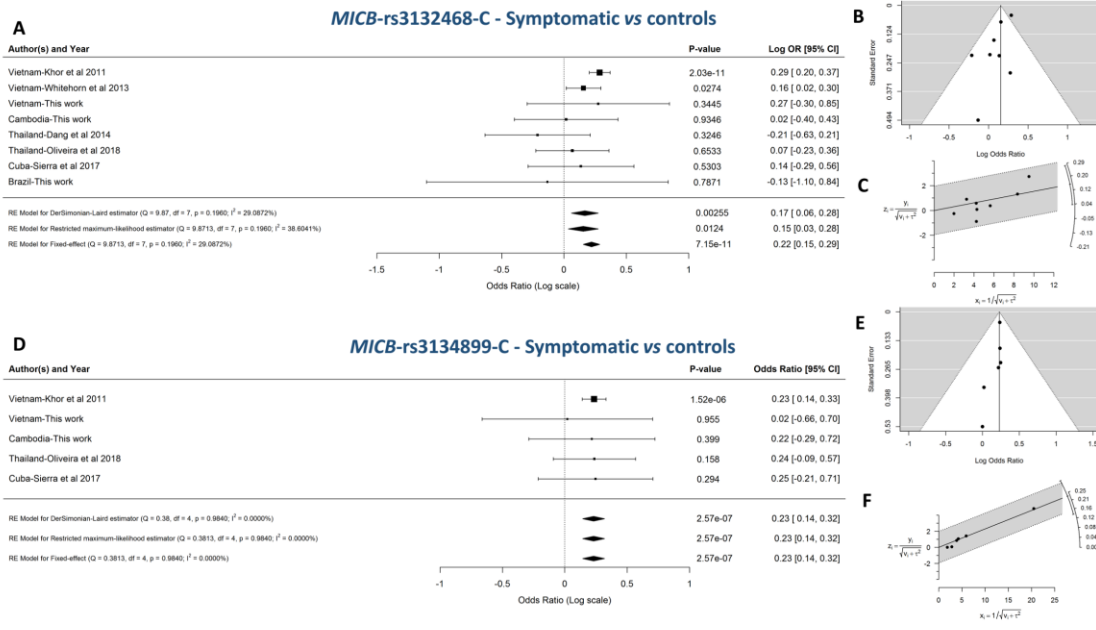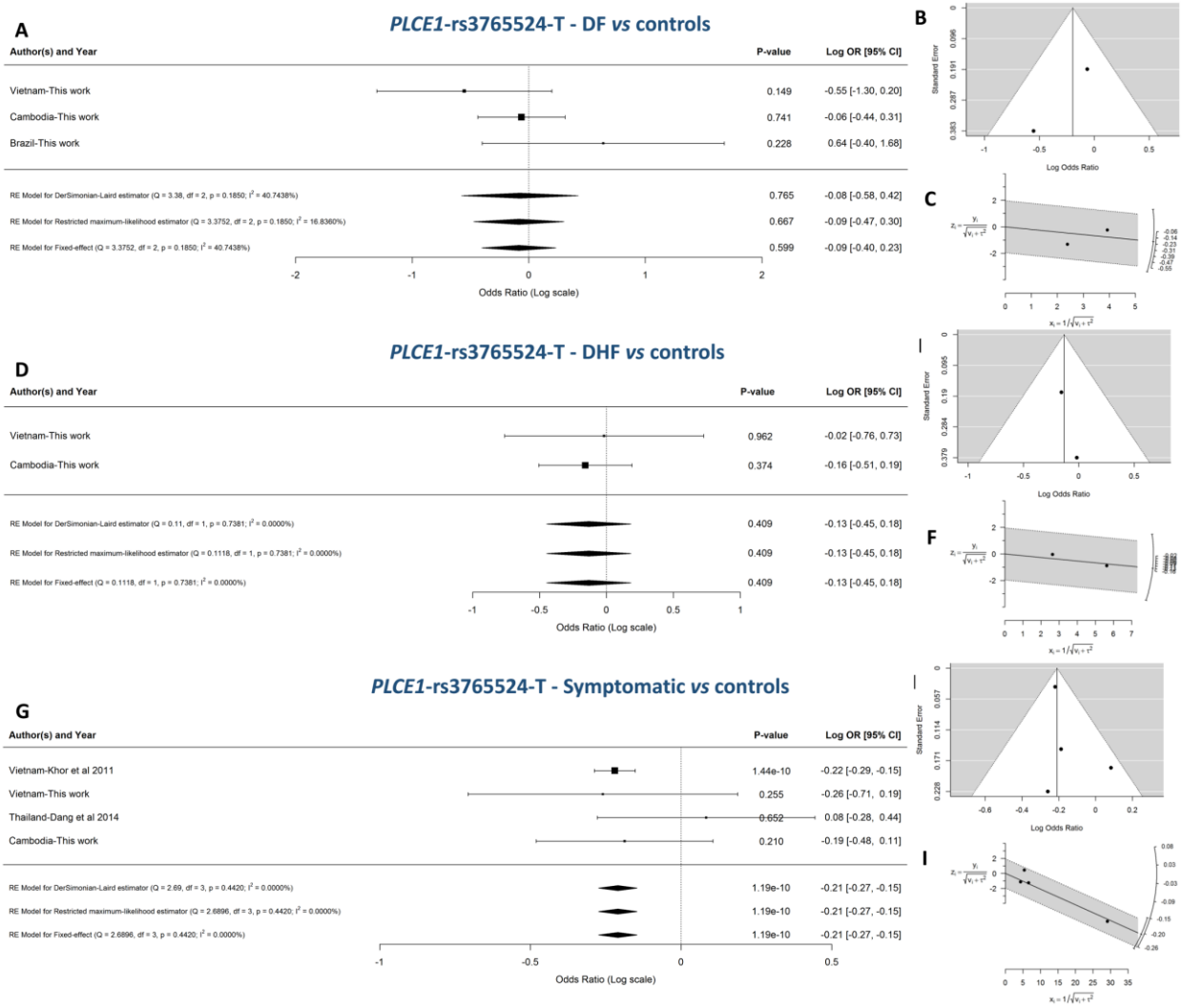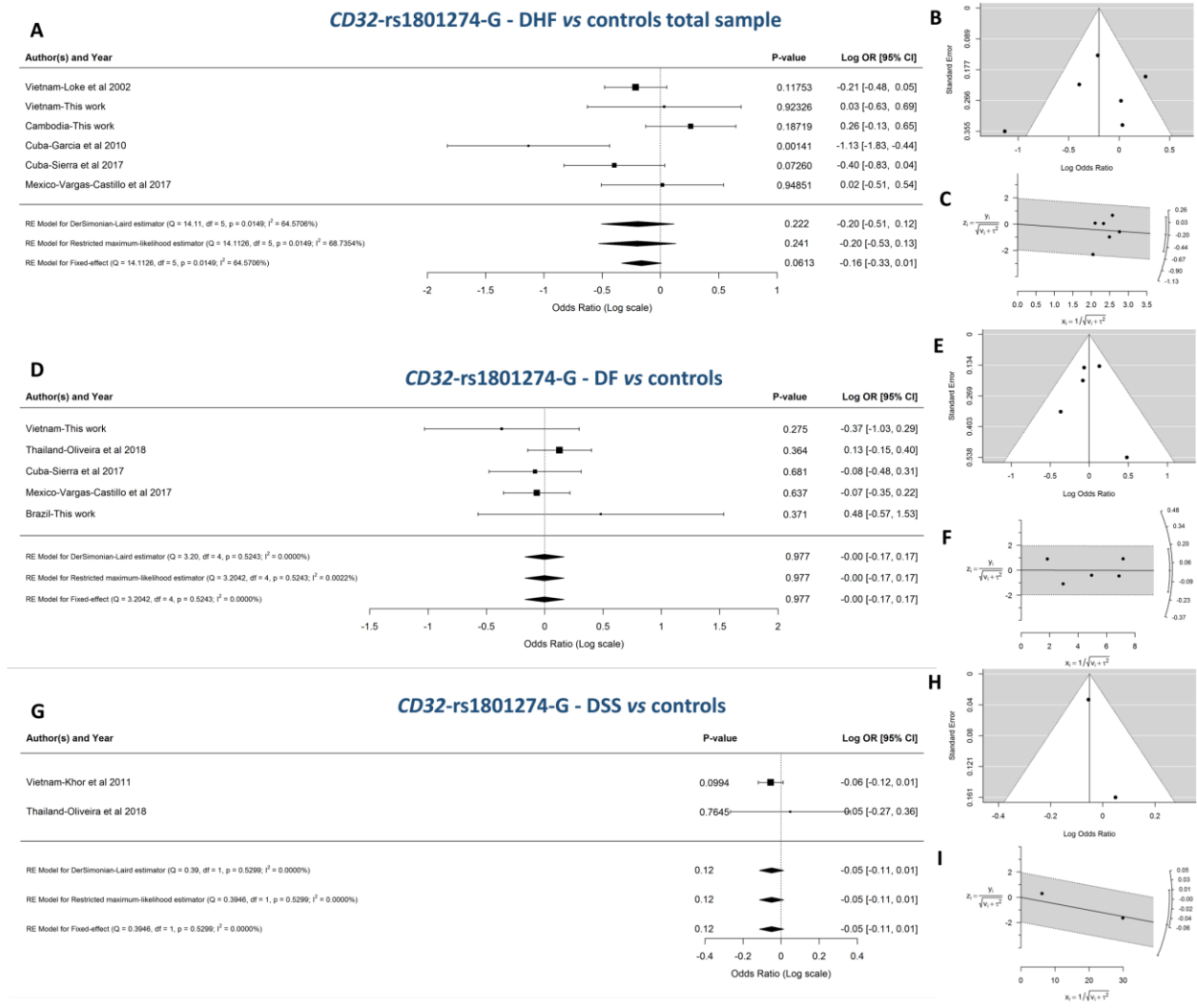

Figure A.23– Meta–analysis in *MICB*–rs3132468–C (A–C) and *MICB*–rs3134899–C (D–F) for DHF *vs* controls comparison. (A,D) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and *I²* values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E) Funnel plots. (C,F) Radial plots.

Figure A.24– Meta–analysis in *MICB*–rs3132468–C (A–C) and *MICB*–rs3134899–C (D–F) for Symptomatic *vs* controls comparison. (A,D) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and I^2 values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E) Funnel plots. (C,F) Radial plots.
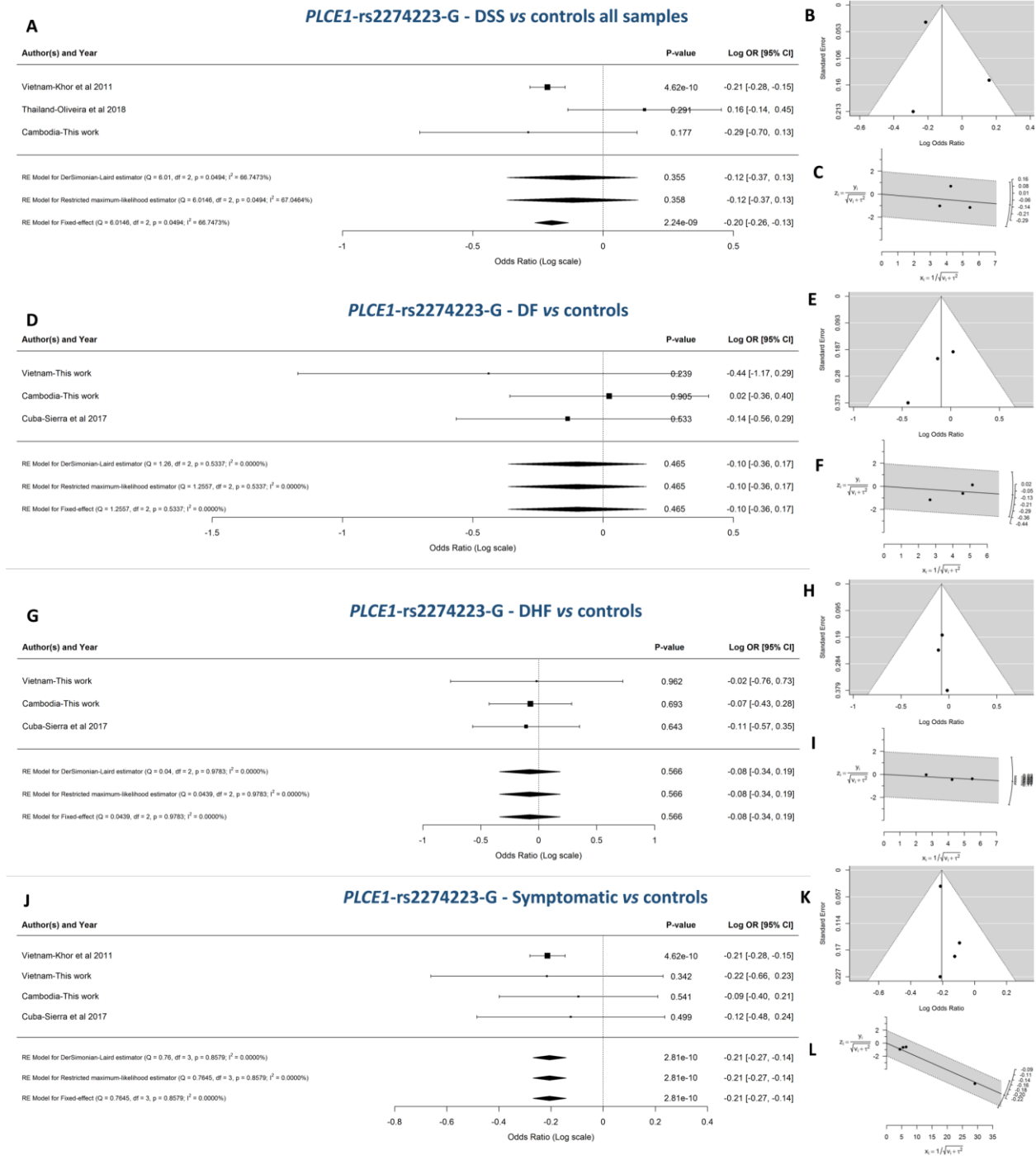
Figure A.25– Meta-analysis in *PLCE1*–rs2274223–G for DSS vs controls in all samples (A–C), and removing Brazil and Cambodia for DF vs controls (D–F), DHF vs controls (G–I) and Symptomatic *vs* controls (J–L) comparisons. (A,D,G,J) Forest plots, pooled log OR ([95% CI] and p-values), Cochran's *Q* measure (and p-values) and I^2 values for the tested three models (weighted fixed–effect, Der Simonian– Laird, and restricted maximum–likelihood). (B,E,H,K) Funnel plots. (C,F,I,L) Radial plots.

Figure A.26– Meta–analysis in *PLCE1*–rs3765524–T for DF vs controls (A–C), for DHF vs controls (D–F), and Symptomatic vs controls (G–I) comparisons. (A,D,G) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's Q measure (and p–values) and I^2 values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E,H) Funnel plots. (C,F,I) Radial plots.

Figure A.27– Meta–analysis in *CD32*–rs1801274–G for DHF vs controls considering all studies (A–C), for DF vs controls (D–F) and for Dss vs controls (G–I) comparison. (A,D,G) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and *I²* values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E,H) Funnel plots. (C,F,I) Radial plots.
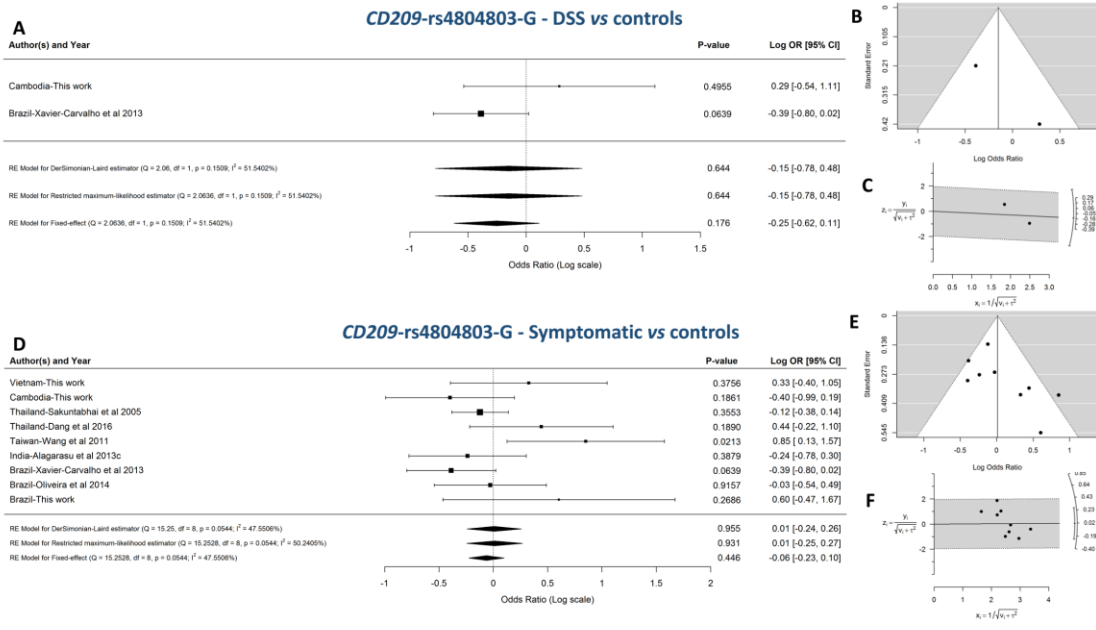
Figure A.28– Meta–analysis in *CD209*–rs4804803–G for DSS vs controls (A–C), and for Symptomatic vs controls (D–F) comparisons. (A,D) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and I^2 values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E) Funnel plots. (C,F) Radial plots.
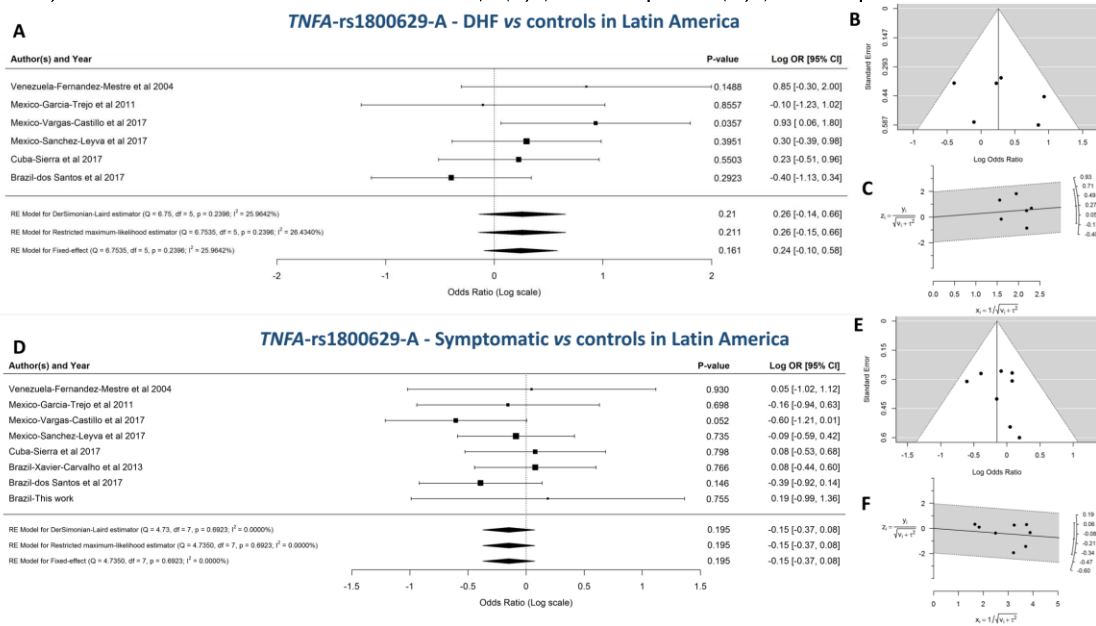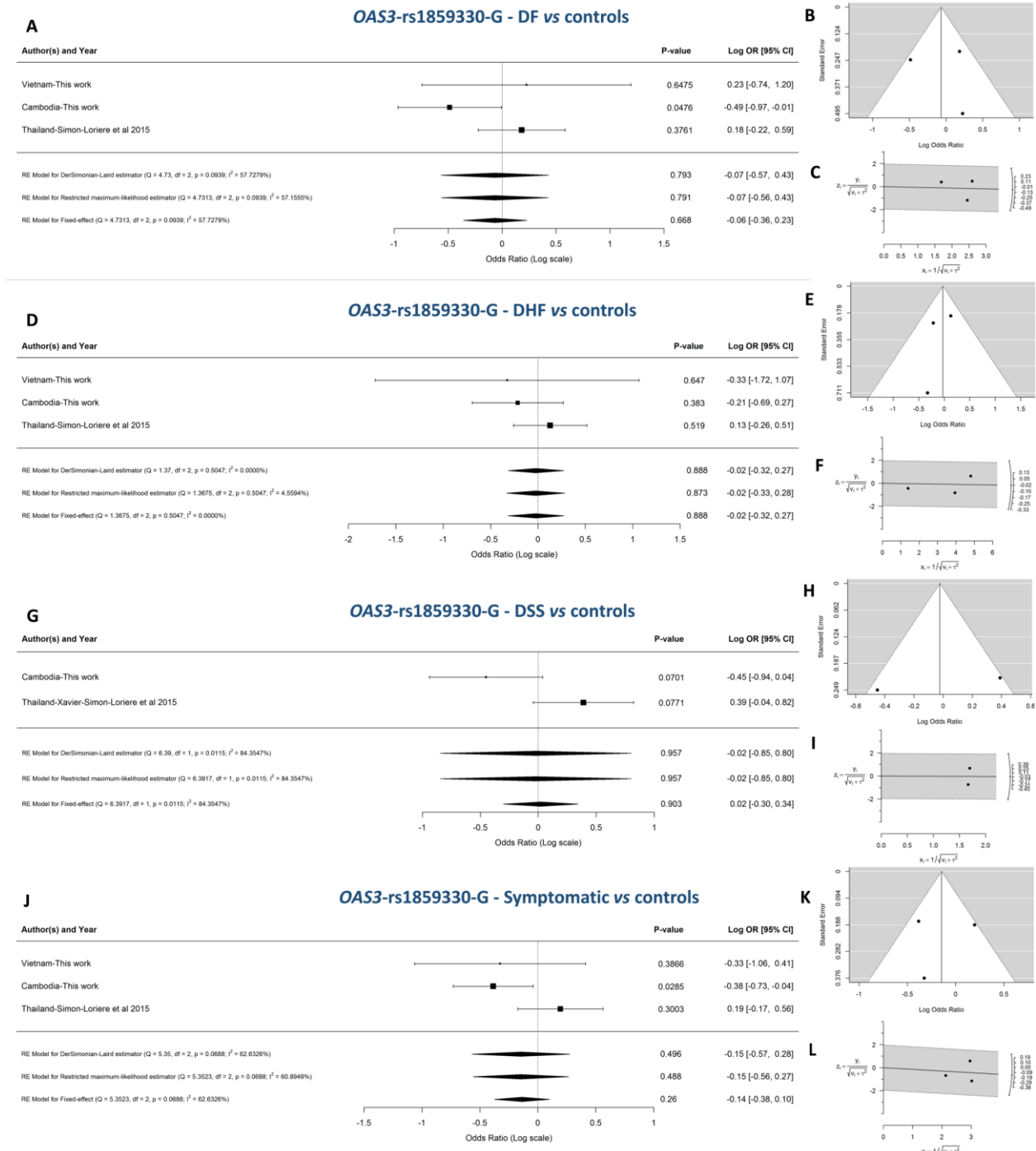

Figure A.29– Meta–analysis in *TNFA*–rs1800629–A for DHF vs controls (A–C), and for Symptomatic vs controls (D–F) in Latin America comparisons. (A,D) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and I^2 values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E) Funnel plots. (C,F) Radial plots.
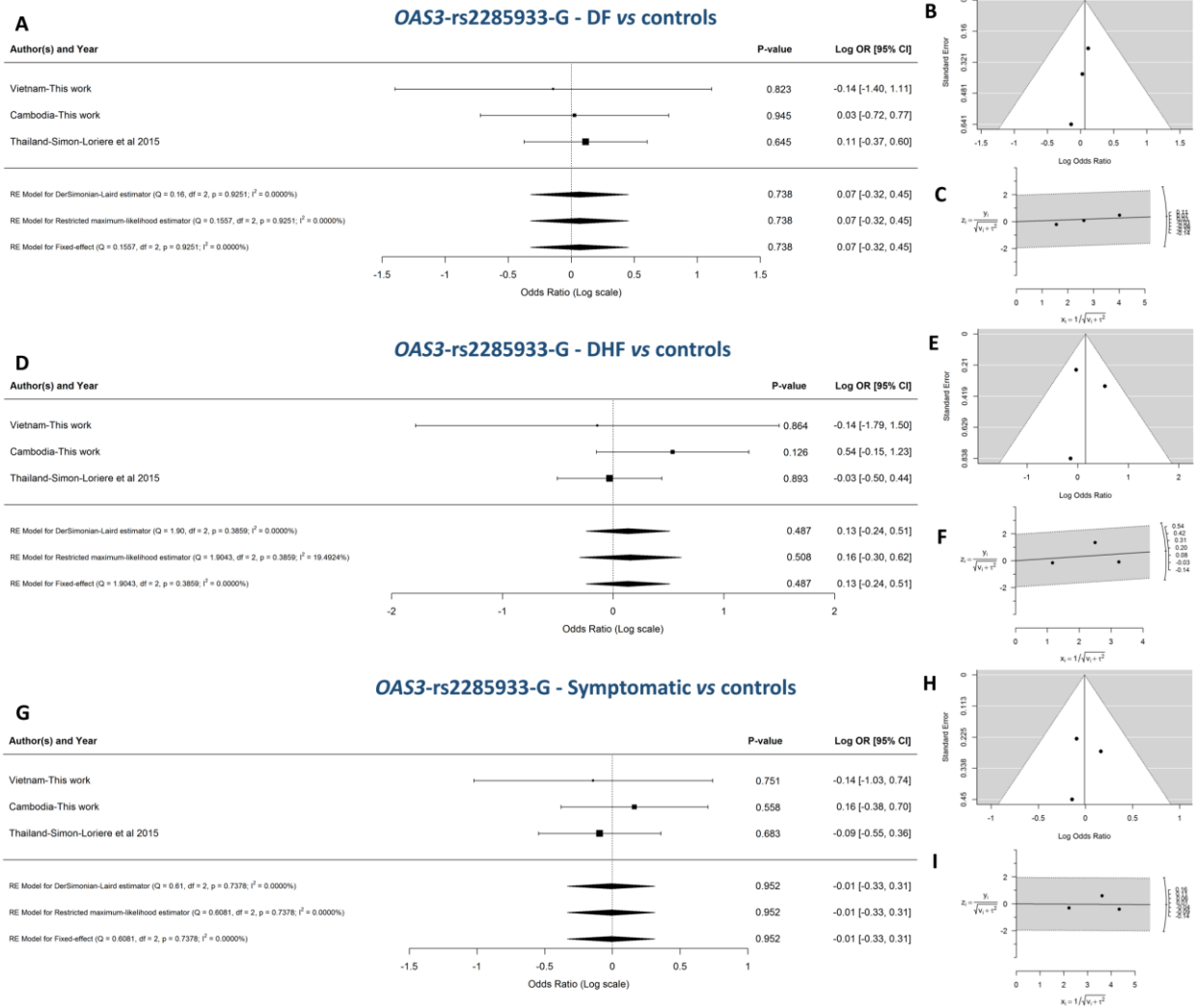
256
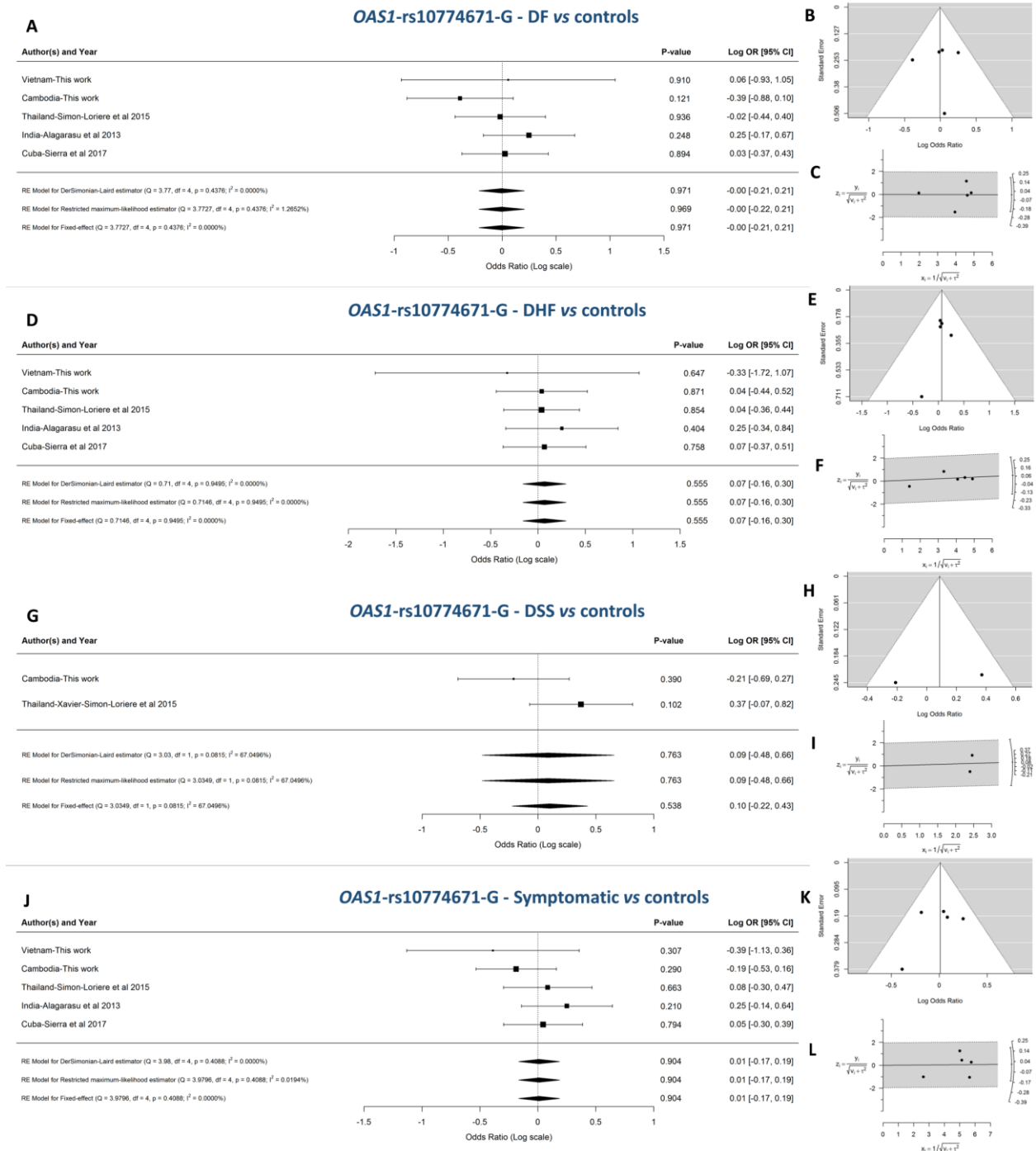
Figure A.30– Meta–analysis in *OAS1*–rs10774671–G for DF vs controls (A–C), DHF vs controls (D–F), DSS vs controls (G–I), and for Symptomatic vs controls (J–L) comparisons. (A,D,G,J) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and *I²* values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E,H,K) Funnel plots. (C,F,I,L) Radial plots.

Figure A.31– Meta–analysis in *OAS3*–rs1859330–G for DF vs controls (A–C), DHF vs controls (D–F), DSS vs controls (G–I), and for Symptomatic vs controls (J–L) comparisons. (A,D,G,J) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and I^2 values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E,H,K) Funnel plots. (C,F,I,L) Radial plots.

Figure A.32– Meta–analysis in *OAS3*–rs2285933–G for DF vs controls (A–C), DHF vs controls (D–F), and for Symptomatic vs controls (G–I) comparisons. (A,D,G) Forest plots, pooled log OR ([95% CI] and p–values), Cochran's *Q* measure (and p–values) and *I²* values for the tested three models (weighted fixed–effect, Der Simonian–Laird, and restricted maximum–likelihood). (B,E,H) Funnel plots. (C,F,I) Radial plots.
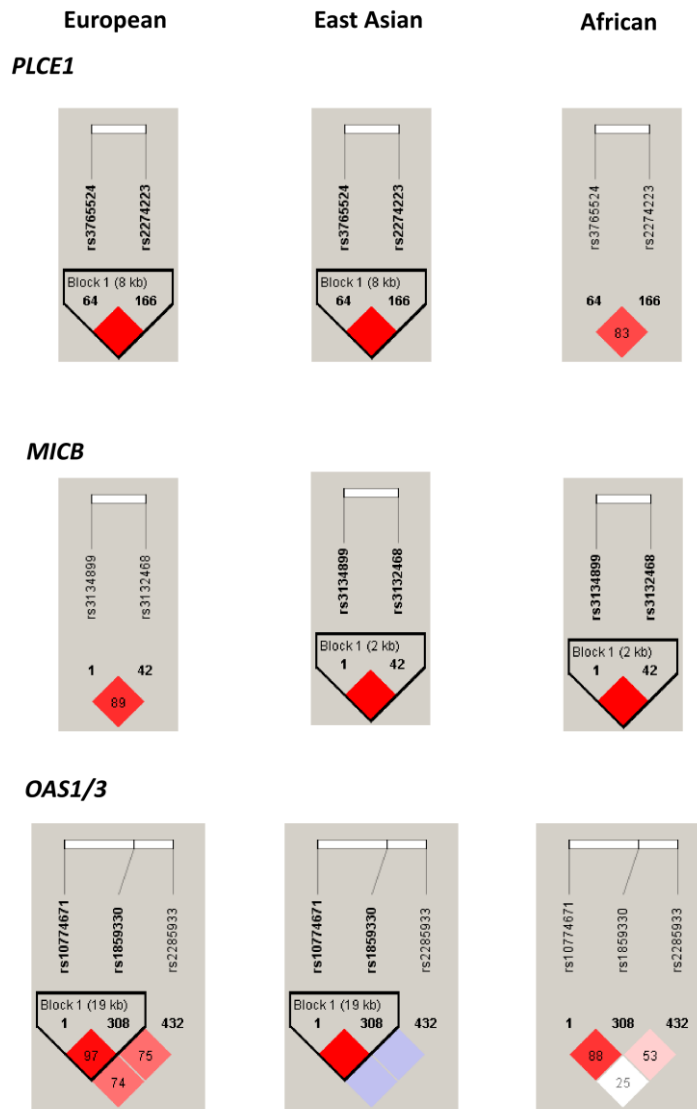
Figure A.33– Linkage disequilibrium (LD) for the two SNPs in *PLCE1*, two SNPs in *MICB* and three SNPs in *OAS1/3* in the 1000 Genomes project populations of Great Britain, China–Han and Yoruba, as proxies for European, East Asian and African population groups. Graphs obtained in Haploview v 4.2.
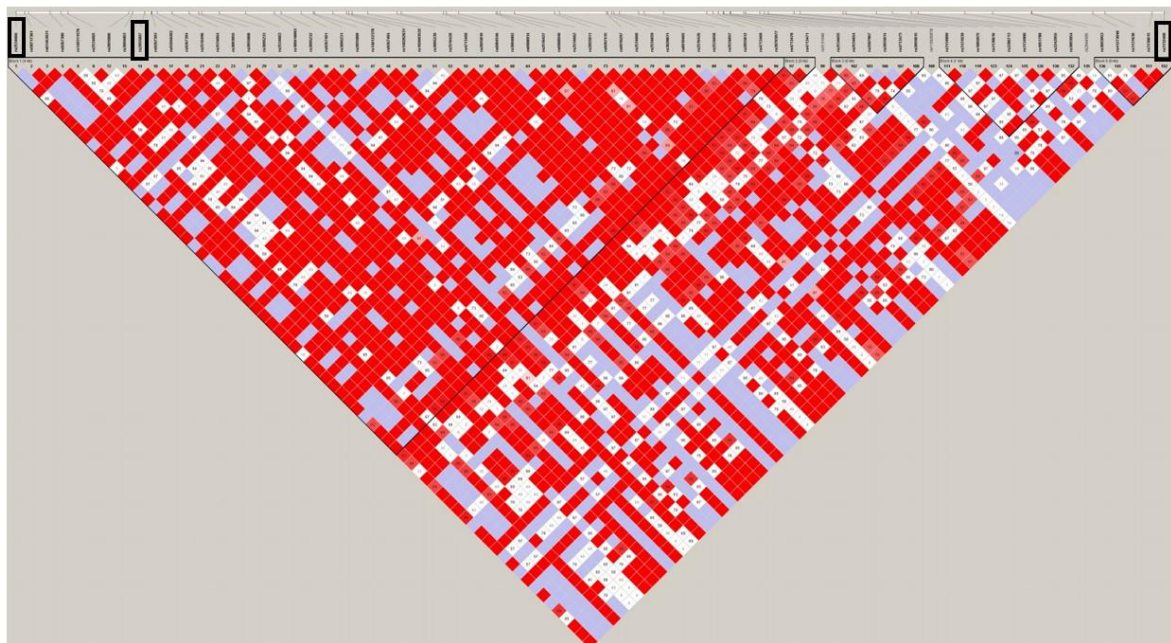
Appendix B – Supplementary information from paper II

Joint ancestry and association test indicate two distinct pathogenic pathways involved in classical dengue fever and dengue shock syndrome
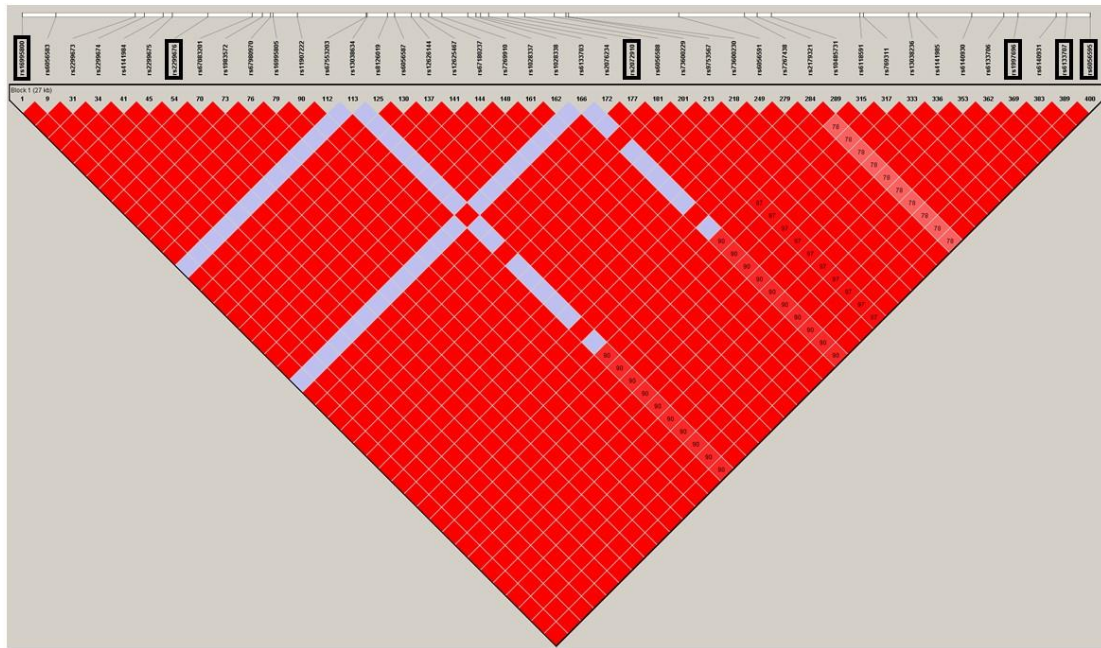
Marisa Oliveira, Worachart Lert-itthiporn, Bruno Cavadas, Verónica Fernandes, Ampaiwan Chuansumrit, Orlando Anunciação, Isabelle Casademont, Fanny Koeth, Marina Penova, Kanchana Tangnararatchakit, Chiea Chuen Khor, Richard Paul, Prida Malasit, Fumihiko Matsuda, Etienne Simon-Lorière, Prapat Suriyaphol, Luisa Pereira and Anavaj Sakuntabhai
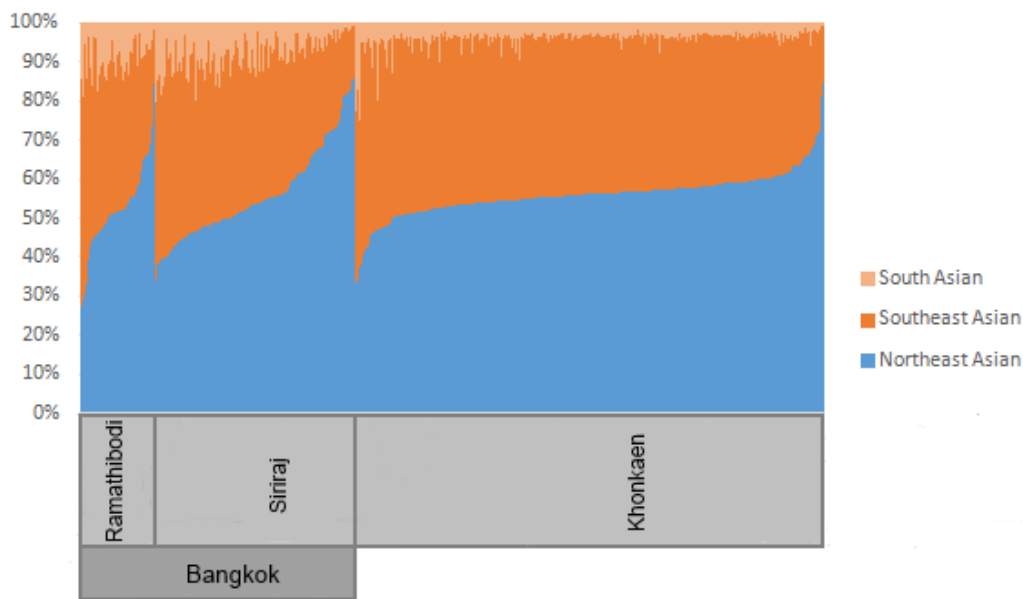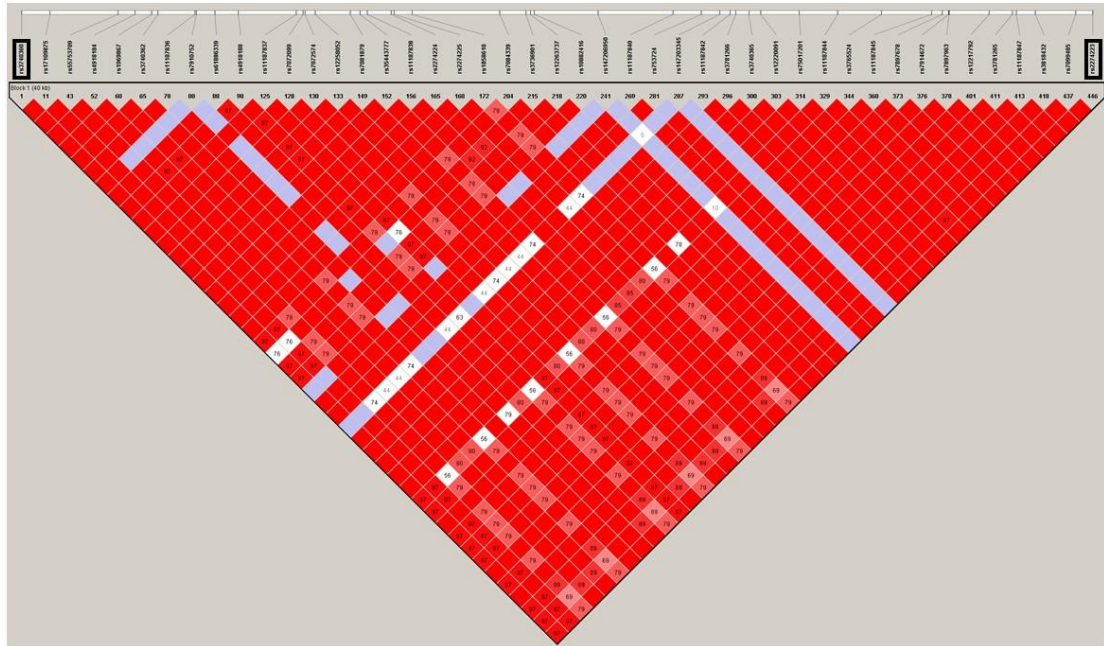
S1 Figure. Global ancestry inferred through RFMix when using three parental ancestries (South, Northeast and Southeast Asian) for the three Thai hospital dengue cohorts. Each vertical line represents an individual, and the three colours represent the proportion of the three parental populations in each genome (light orange for South Asian, dark orange for Southeast Asian and blue for Northeast Asian).
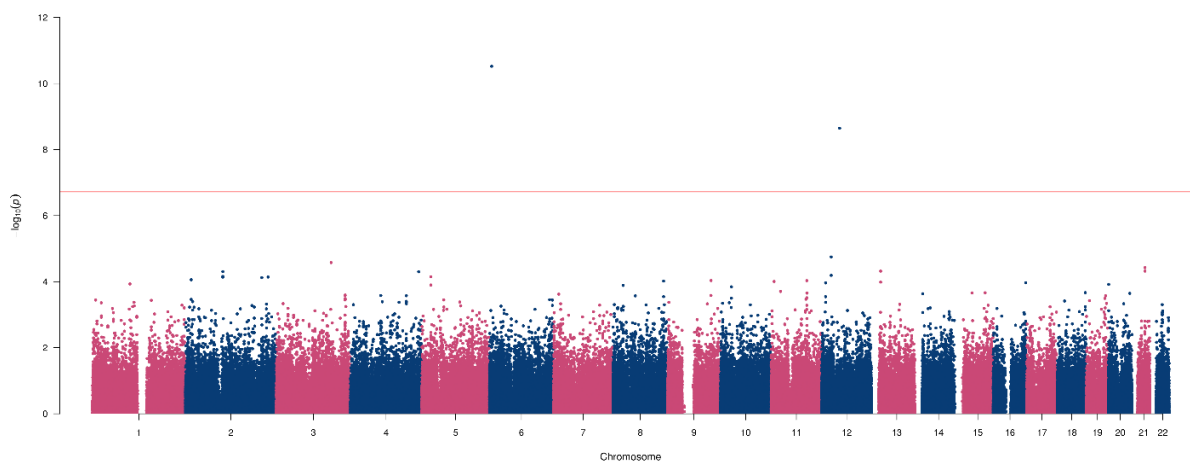

S2 Figure. LD (D') values for the *MICB* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.

263

S3 Figure. LD (D') values for the *PLCB4* **region in the Chinese population (CDX) from 1000 Genomes database.** BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.

S4 Figure. Locus zoom of the chromosomal region around genes (A – *PLCB4*, B – *PLCE1*, C – *CHST10*, D – *AHRR*, E – *GRIP1*, F – *PPP2R5E*) with significant p-values obtained for DSS and DF tests. The Asian recombination map was used.

S5 Figure. LD (D') values for the *PLCE1* **region in the Chinese population (CDX) from 1000 Genomes database.** BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.

| Tag SNP	Chr	SNP	BP	Allele	OR	P-value	D' with tag SNP
	6	rs6939701	5934718	C	1.027	0.871	0.462
	6	rs3018	5938368	T	0.9635	0.8772	0.745
Tag	**6**	**rs4959364**	**5947139**	**A**	**0.3191**	**3.00E-11**	
	6	rs11759185	5958615	T	0.977	0.8939	0.964
	6	rs1555536	5961601	A	0.9602	0.8128	0.830
	12	rs11168353	48408212	T	0.9491	0.8367	1.000
	12	rs12819124	48409054	A	0.68	0.1146	1.000
Tag	**12**	**rs6580649**	**48410517**	**C**	**0.3528**	**2.26E-09**	
	12	rs11168357	48412138	A	1.167	0.3924	1.000
	12	rs7299271	48414058	C	1.074	0.6797	1.000

A
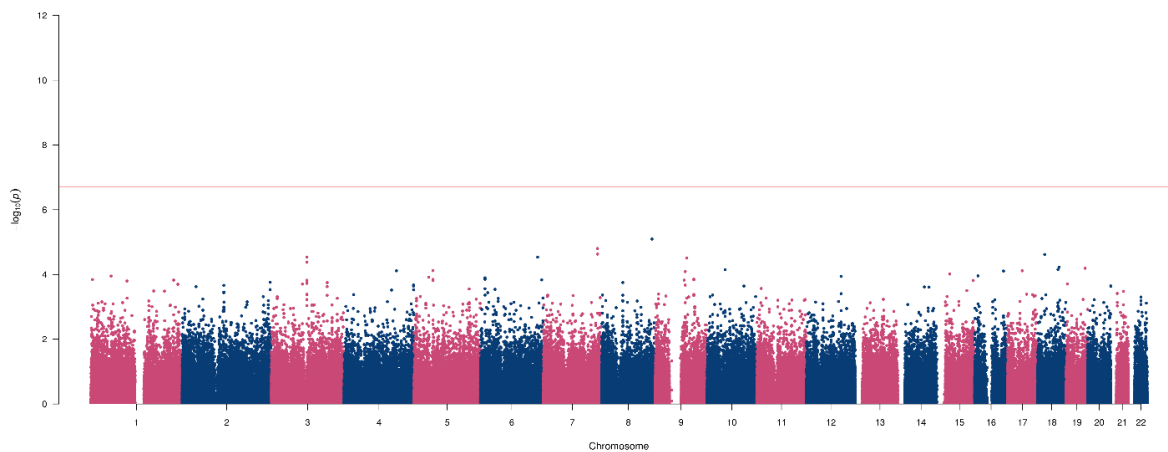
B

S6 Figure Manhattan plots for the conventional association tests with PCA correction for population structure. A – DSS test (p-values and D' for SNPs surrounding the two spurious SNPs are highlighted). B – DF test. The red line indicates the significance threshold. (DOC)

267
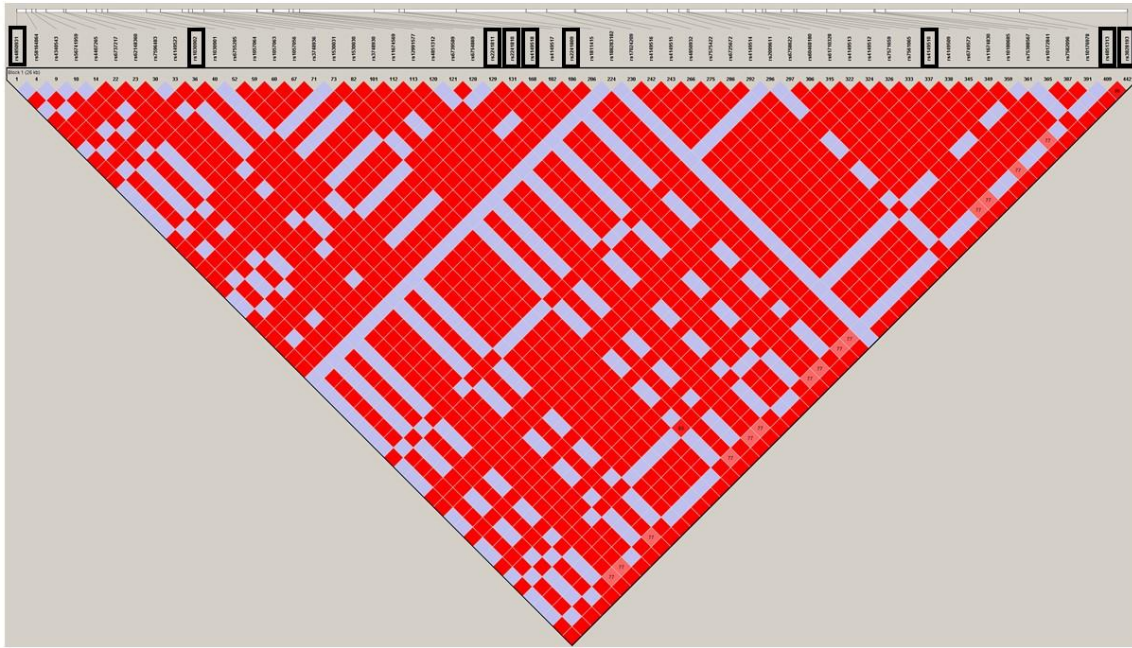
S7 Figure. LD (D') values for the *CHST10* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.

S8 Figure. LD (D') values for the *AHRR* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.
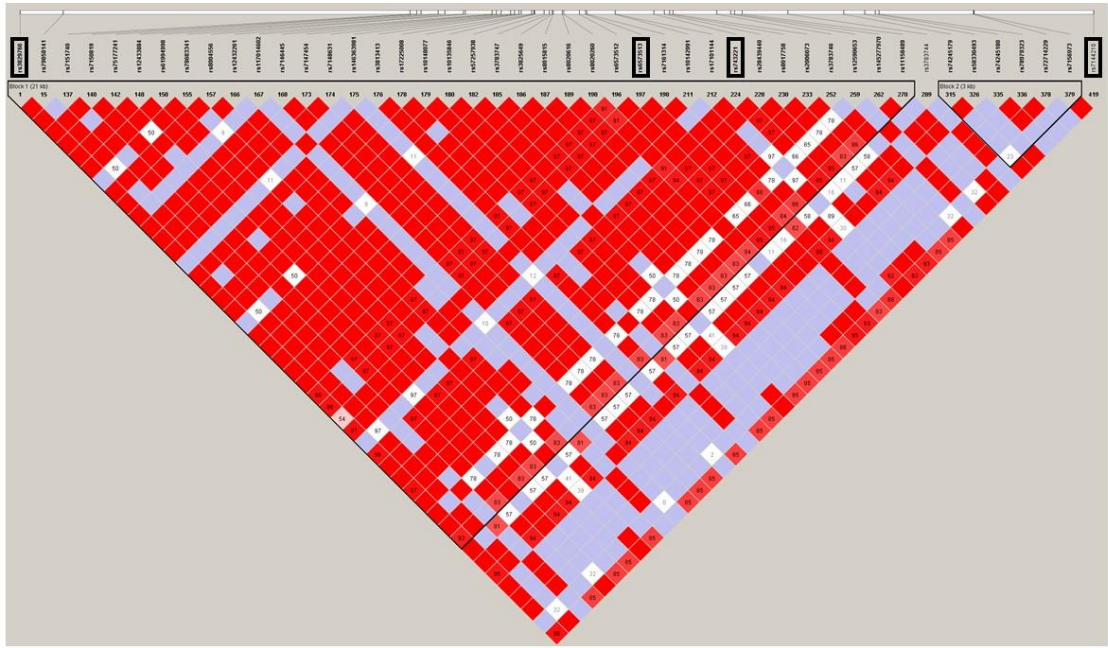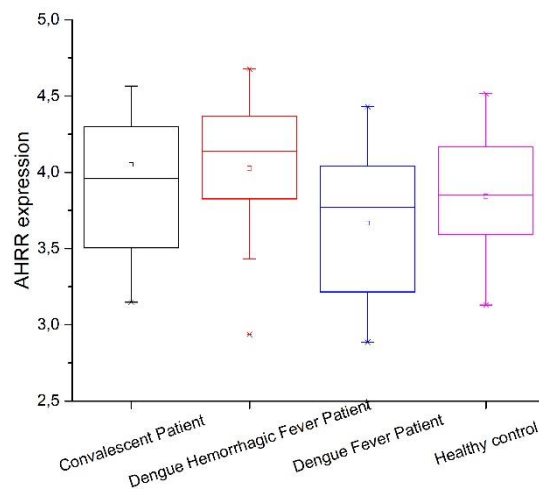
S9 Figure. LD (D') values for the *PPP2R5E* region in the Chinese population (CDX) from 1000 Genomes database. BMIX identified significant SNPs are indicated by a box. All SNPs have at least 5% minimum allele frequency in the population analysed.
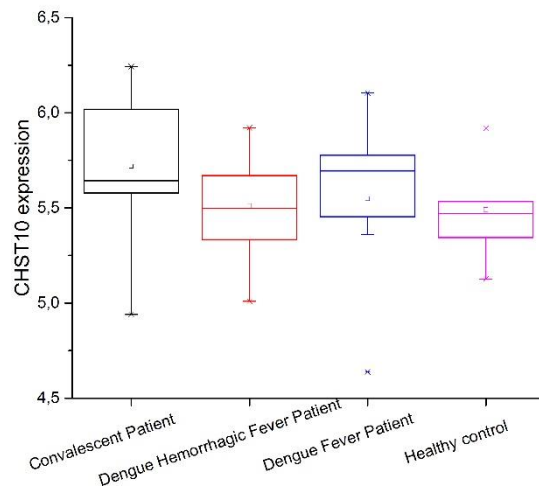
A

B

S10 Figure. Gene expression for *CHST10* (A) and *AHRR* (B) in Thai dengue cohort along the course of disease from a transcriptome dataset for whole blood.[1] No significant differences in expression were observed.

References

1. Kwissa, M. *et al.* Dengue virus infection induces expansion of a CD14(+)CD16(+) monocyte population that stimulates plasmablast differentiation. *Cell Host Microbe* **16**, 115–27 (2014).
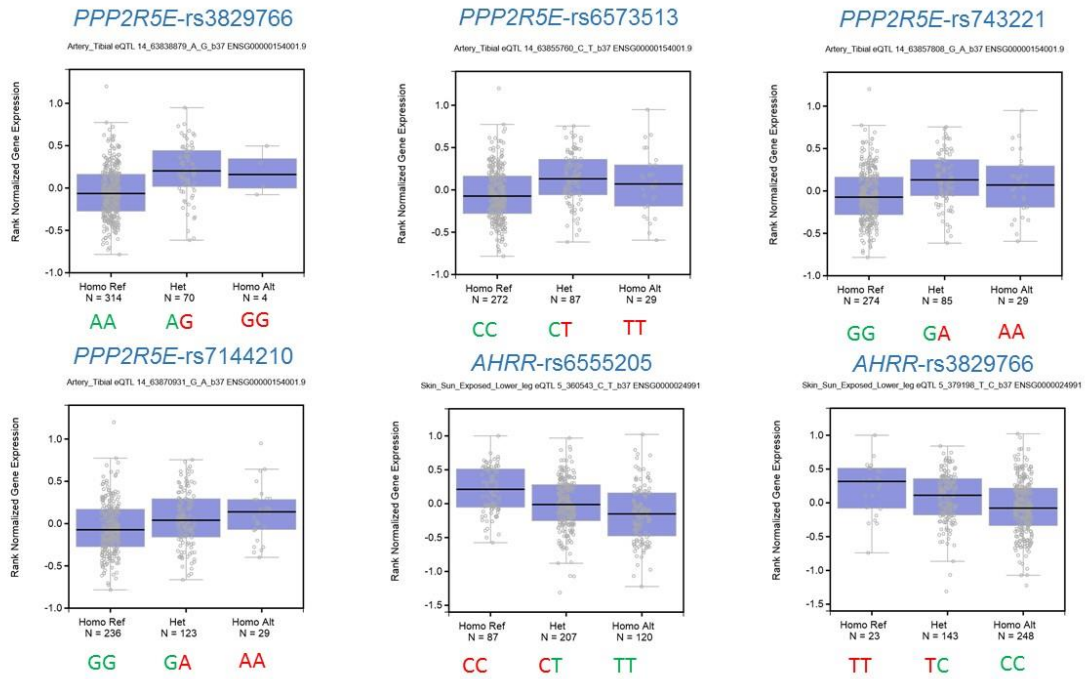
S11 Figure mRNA expression profiles for the eQTLs in *PPP2R5E* and *AHRR* genes (information from GTEx database). The protective alleles are indicated in green while the causative alleles are in red.
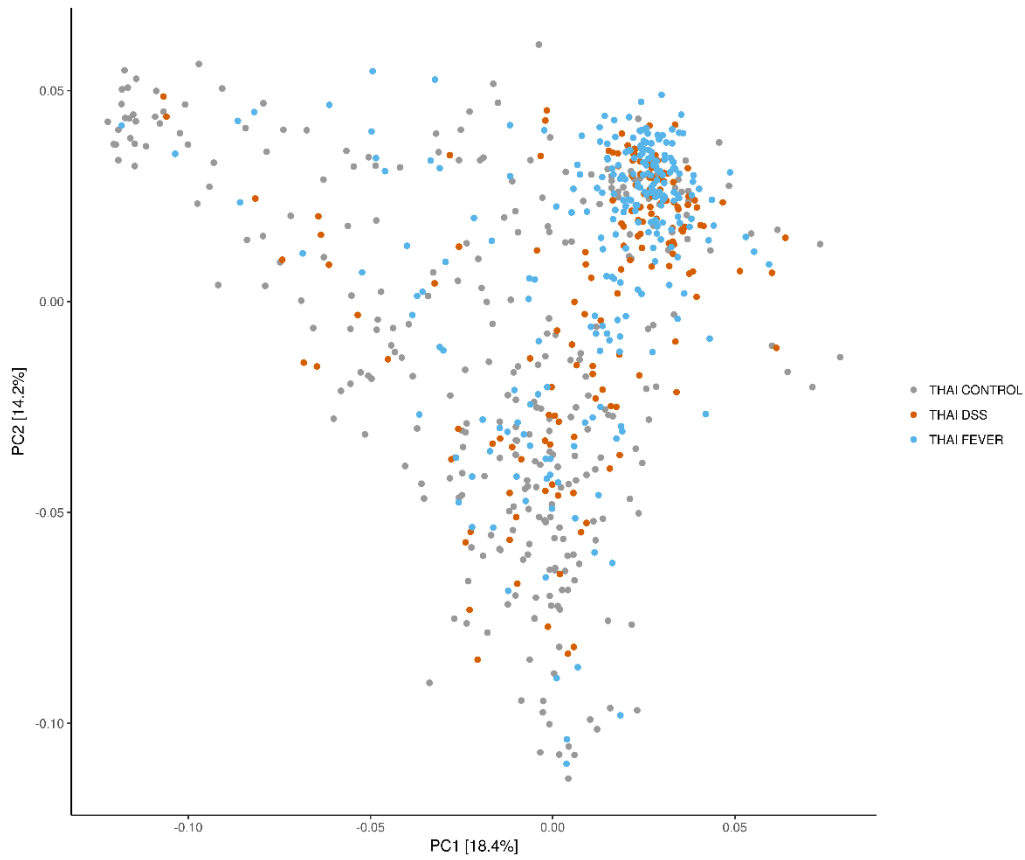
S12 Fig. PCA of the Thai samples. Plot of PC1 versus PC2 in the Thai control and patient cohorts.
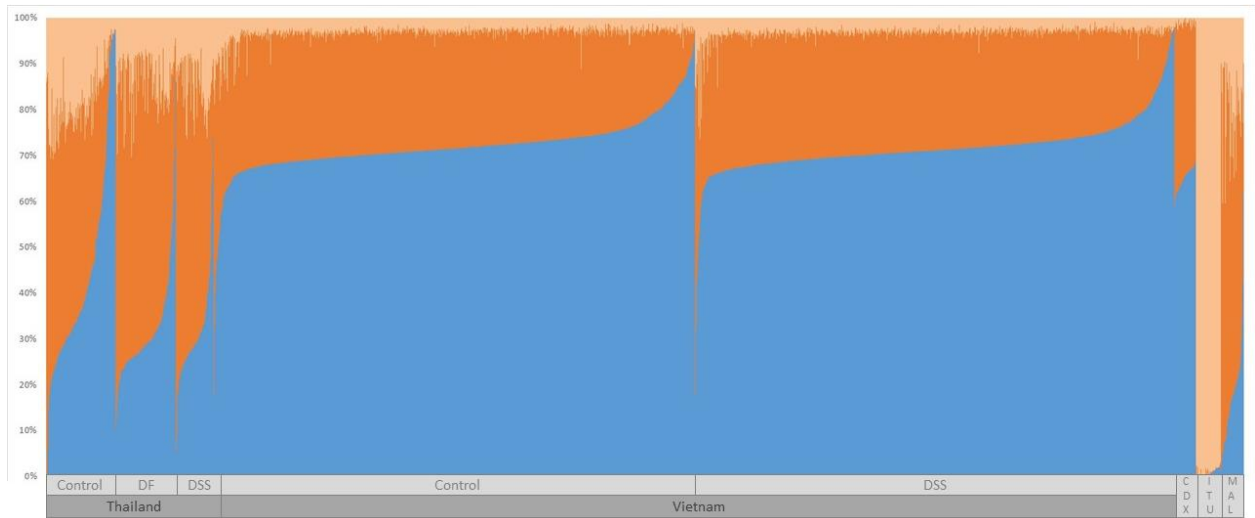
S13 Fig. ADMIXTURE plot for K=3 for Thai and Vietnamese cohorts and the parental populations used in this work (CDX - Chinese Dai in Xishuangbanna; ITU- Indian Telugu from the UK (ITU); and MAL - Malaysian).

S1 Table. Significant SNPs in BMIX analysis for Vietnam DSS test. The base position refers to GRCh37 genome assembly.

Chr	SNP	BP	Allele	Association p-value	OR	BMIX posterior p-value in Northeast Asian ancestry	BMIX posterior p-value in Southeast Asian ancestry	Gene
4	rs17603961	37386498	A	0.0007608	1.210	0.503	0.589	NWD2
6	rs1051794	31379109	G	0.04197	1.195	0.896	0.762	MICA
6	rs1131904	31383071	G	0.0002064	1.192	0.883	0.737	MICA/HCP5
6	rs2534666	31468546	A	1.46E-08	1.343	0.993	0.999	MICB/Y_RNA
6	rs2855807	31469323	C	9.38E-07	1.294	0.993	0.999	MICB/Y_RNA
6	rs3132468	31475486	C	7.04E-08	1.398	0.999	1.000	MICB
6	rs9267487	31511350	G	0.00121	1.442		0.637	DDX39B/DDX39B-AS1/ATP6V1G2/SNORD84/NFKBIL1
6	rs3093662	31544189	G	0.001019	1.439		0.705	LTB/LTA/TNF
9	rs2417485	106867106	C	0.0019776	1.193	0.821	0.837	SMC2
9	rs2122576	106870187	C	0.0023773	1.191	0.854	0.842	SMC2
10	rs3740360	96025491	C	7.05E-08	0.756	0.529	0.649	PLCE1
10	rs2274223	96066341	G	1.26E-07	0.768		0.504	PLCE1
10	rs2421027	124203648	G	5.16E-05	0.496		0.550	
12	rs1231794	31884832	A	0.0013009	1.224	0.523		AMN1
20	rs6074355	11786106	A	0.0001216	1.506	0.776	0.887	LINC00687/AL080274.1
20	rs6074356	11786432	A	0.0001268	1.512	0.773	0.885	LINC00687/AL080274.1

S2 Table. Significant SNPs in BMIX analysis for Thai DSS vs control test. The base position refers to GRCh37 genome assembly.

Chr	SNP	BP	Allele	Association p–value	OR	BMIX posterior p–value in Northeast Asian ancestry	BMIX posterior p–value in Southeast Asian ancestry	Gene
1	rs705731	202298894	C	2.09E–05	0.497	0.585		*UBE2T*
2	rs1884725	31571786	A	0.001065	1.778	0.628		*XDH*
2	rs7566302	33595197	A	0.002199	0.576		0.583	*LTBP1*
2	rs11679130	35730792	C	0.0008383	1.605		0.569	
2	rs11682759	204192201	C	5.15E–05	1.767	0.541		*ABI2/RP11-363J17.1*
2	rs13383306	230179536	A	0.0006001	2.159	0.522		
4	rs17256627	11873033	T	9.81E–05	0.401	0.924	0.852	
4	rs11937407	81413618	C	3.00E–05	0.538	0.866		*C4orf22*
4	rs13109014	85260011	T	2.01E–06	2.446	0.961		
5	rs1501938	24059579	G	4.08E–05	2.004	0.833	0.643	*C5orf17*
5	rs6452189	24163109	T	1.41E–05	1.974	0.936	0.655	*C5orf17*
6	rs4959364	5947139	A	7.22E–17	0.285	1.000	1.000	
8	rs10105057	19614339	C	0.0007059	1.622		0.679	*CSGALNACT1*
8	rs7837390	19626120	C	0.00057	1.632		0.734	
8	rs6983707	91133678	G	0.0001822	1.706	0.688		
9	rs16922639	106004562	T	2.56E–05	2.476		0.958	*RP11-341A22.2*
12	rs6580649	48410517	C	2.26E–11	0.357	1.000	1.000	*RP1-228P16.4*
20	rs16995800	9352562	G	0.006618	0.565	0.703		*PLCB4*
20	rs2299676	9357437	A	0.006497	0.582	0.542		*PLCB4*
20	rs7269910	9363565	G	0.003453	0.553	0.917		*PLCB4*
20	rs1997696	9378671	A	0.007033	0.667	0.691		*PLCB4*
20	rs6133707	9379949	G	0.008858	0.675	0.679		*PLCB4*
20	rs6056595	9380556	G	0.008289	0.672	0.691		*PLCB4*

S3 Table. Significant SNPs in BMIX analysis for Thai DF test. The base position refers to GRCh37 genome assembly.

Chr	SNP	BP	Allele	Association p–value	OR	BMIX posterior p–value in Northeast Asian ancestry	BMIX posterior p–value in Southeast Asian ancestry	Gene
1	rs12028426	171927696	T	0.0003996	1.579		0.938	DNM3
2	rs394874	17655480	T	3.17E–05	0.387		0.891	
2	rs2309798	100956949	G	0.0001066	0.594		0.826	
2	rs4850931	101005145	C	0.0002867	0.583		0.556	CHST10
2	rs1030902	101007178	T	4.40E–05	0.577		0.940	CHST10
2	rs2241811	101011724	C	0.0003	0.582		0.565	CHST10
2	rs2241810	101011877	T	3.19E–05	0.571		0.957	CHST10
2	rs4149518	101013649	A	3.19E–05	0.571		0.957	CHST10
2	rs2241809	101014363	C	3.54E–05	0.573		0.854	CHST10
2	rs4149510	101023635	G	3.19E–05	0.571		0.957	CHST10
2	rs4851313	101029002	G	8.99E–05	0.561		0.835	CHST10
2	rs3828193	101031561	G	5.57E–06	0.545	0.794	0.987	CHST10
5	rs6555205	360543	T	1.03E–05	0.510		0.851	AHRR
5	rs2721020	379198	C	3.61E–05	0.558		0.505	AHRR
5	rs1994929	23507631	T	0.0008452	1.666	0.780		PRDM9
5	rs7708103	177410416	G	0.001826	0.664		0.648	RP11-1252I4.2
6	rs532098	32578052	A	8.60E–05	0.612	0.597		
6	rs9397270	156163798	T	6.82E–08	2.019	0.864		
8	rs2255522	137523980	G	5.23E–06	0.566	0.910	0.976	RP11-431D12.1
12	rs1480010	67076016	T	0.001176	0.557	0.929		GRIP1

275

Chr	SNP	Position	Allele	P	OR			Gene
12	rs2717418	70973596	G	0.001935	1.524	0.845		*PTPRB*
14	rs9323435	63780460	C	2.49E-05	1.696		0.853	*GPHB5*
14	rs3829766	63838879	G	0.0003975	1.547		0.527	*PPP2R5E*
14	rs6573513	63855760	C	0.0003278	0.638		0.662	*PPP2R5E*
14	rs743221	63857808	G	0.0002449	0.632		0.718	*PPP2R5E*
14	rs7144210	63870931	G	0.000268	0.633		0.843	*PPP2R5E*
16	rs7184164	64281609	C	0.002642	1.847		0.576	*AC012322.1*
21	rs2212870	20941913	T	0.01392	1.355	0.545		
21	rs2825968	21415693	C	0.004167	1.421	0.887		
21	rs2825993	21443895	G	0.001717	0.622	0.573		
21	rs2826059	21543190	T	0.0003941	0.459	0.738	0.549	

S4 Table. Annotation of the significant SNPs in BMIX analysis for Vietnam DSS test, inferred by using the Variant Effect Predictor (VEP) tool from Ensemble.

Chr	SNP	Impact in VEP	Gene	Consequence
4	rs17603961	Modifier	*NWD2*	intron variant
6	rs1051794	Moderate/Modifier	*MICA*	missense, intron, non-coding transcript variant
6	rs1131904	Modifier	*MICA/HCP5*	3 prime UTR, intron, non-coding transcript variant
6	rs2534666	Modifier	*MICB/Y_RNA*	downstream gene, intron, upstream gene, regulatory region (enhancer) variant
6	rs2855807	Modifier	*MICB/Y_RNA*	downstream gene, intron, upstream gene variant
6	rs3132468	Modifier	*MICB*	downstream gene, intron variant
6	rs9267487	Modifier	*DDX39B/DDX39B-AS1/ATP6V1G2/SNORD84/NFKBIL1*	upstream gene, downstream gene, intron, NMD transcript, regulatory region (promoter) variant

276

6	rs3093662	Modifier	*LTB/LTA/TNF*	downstream gene, intron, regulatory region (promoter) variant
9	rs2417485	Modifier	*SMC2*	intron, downstream gene variant
9	rs2122576	Modifier	*SMC2*	intron variant
10	rs3740360	Modifier	*PLCE1*	intron variant
10	rs2274223	Moderate	*PLCE1*	missense variant
10	rs2421027	Modifier		intergenic variant
12	rs12317948	Modifier	*AMN1*	upstream gene variant
20	rs6074355	Modifier	*LINC00687/AL080274.1*	downstream gene, intron, non-coding transcript variant
20	rs6074356	Modifier	*LINC00687/AL080274.1*	downstream gene, intron, non-coding transcript variant

S5 Table. Annotation of the significant SNPs in BMIX analysis for Thai DSS test, inferred by using the Variant Effect Predictor (VEP) tool from Ensemble.

Chr	SNP	Impact in VEP	Gene	Consequence
1	rs705731	Modifier	*UBE2T*	downstream gene variant
2	rs1884725	Low	*XDH*	synonymous variant
2	rs7566302	Modifier	*LTBP1*	intron variant
2	rs11679130	Modifier		intergenic variant
2	rs11682759	Modifier	*ABI2/RP11-363J17.1*	upstream gene, downstream gene, regulatory region (promoter) variant
2	rs13383306	Modifier		intergenic variant
4	rs17256627	Modifier		intergenic variant
4	rs11937407	Modifier	*C4orf22*	intron, non–coding transcript, NMD transcript variant
4	rs13109014	Modifier		intergenic variant
5	rs1501938	Modifier	*C5orf17*	intron variant, NMD transcript variant
5	rs6452189	Modifier	*C5orf17*	intron, NMD transcript, regulatory region (promoter flanking region) variant
6	rs4959364	Modifier		intergenic variant
8	rs10105057	Modifier	*CSGALNACT1*	intron, non–coding transcript, regulatory region (promoter) variant
8	rs7837390	Modifier		intergenic variant
8	rs6983707	Modifier		intergenic variant
9	rs16922639	Modifier	*RP11-341A22.2*	intron, non–coding transcript variant
12	rs6580649	Modifier	*RP1-228P16.4*	intron, non–coding, downstream gene variant
20	rs16995800	Modifier	*PLCB4*	intron, non–coding transcript variant
20	rs2299676	Modifier	*PLCB4*	intron, non–coding transcript variant
20	rs7269910	Modifier	*PLCB4*	intron, non–coding transcript variant
20	rs1997696	Modifier	*PLCB4*	intron, non–coding transcript, regulatory region (promoter flanking region) variant
20	rs6133707	Modifier	*PLCB4*	intron, non–coding transcript, regulatory region (promoter flanking region) variant
20	rs6056595	Modifier	*PLCB4*	intron, non–coding transcript, regulatory region (promoter flanking region) variant

S6 Table. Annotation of the significant SNPs in BMIX analysis for Thai DF test, inferred by using the Variant Effect Predictor (VEP) tool from Ensemble.

Chr	SNP	Impact in VEP	Gene	Consequence
1	rs12028426	Modifier	DNM3	intron variant
2	rs394874	Modifier		intergenic variant
2	rs2309798	Modifier		intergenic variant
2	rs4850931	Modifier	CHST10	downstream gene variant
2	rs1030902	Modifier	CHST10	downstream gene variant
2	rs2241811	Modifier	CHST10	downstream gene, intron, regulatory region (TF binding site) variant
2	rs2241810	Modifier	CHST10	downstream gene, intron, regulatory region (TF binding site) variant
2	rs4149518	Modifier	CHST10	downstream gene, intron variant
2	rs2241809	Modifier	CHST10	downstream gene, splice region, intron variant
2	rs4149510	Modifier	CHST10	intron, downstream gene, non–coding transcript variant
2	rs4851313	Modifier	CHST10	intron, downstream gene, non–coding transcript variant
2	rs3828193	Modifier	CHST10	5 prime UTR, non–coding transcript exon, non–coding transcript variant
5	rs6555205	Modifier	AHRR	intron, regulatory region (open chromatin region) variant
5	rs2721020	Modifier	AHRR	intron, regulatory region (promoter flanking region) variant
5	rs1994929	Modifier	PRDM9	intron, upstream gene variant
5	rs7708103	Modifier	RP11–1252I4.2	downstream gene variant
6	rs532098	Modifier		intergenic variant
6	rs9397270	Modifier		intergenic variant
8	rs2255522	Modifier	RP11–431D12.1	intron, non–coding transcript variant
12	rs1480010	Modifier	GRIP1	intron, upstream gene variant
12	rs2717418	Modifier	PTPRB	intron variant
14	rs9323435	Modifier	GPHB5	intron, non–coding transcript variant
14	rs3829766	Modifier	PPP2R5E	downstream gene, 3 prime UTR variant
14	rs6573513	Modifier	PPP2R5E	intron, non–coding transcript, regulatory region (CTCF and TF binding sites) variant
14	rs743221	Modifier	PPP2R5E	intron, non–coding transcript variant
14	rs7144210	Modifier	PPP2R5E	intron, non–coding transcript variant
16	rs7184164	Modifier	AC012322.1	intron, non–coding transcript variant
21	rs2212870	Modifier		intergenic variant
21	rs2825968	Modifier		intergenic variant

21	rs2825993	Modifier		intergenic variant
21	rs2826059	Modifier		intergenic variant

S7 Table. Association p-values in the entire Thai cohort in six and two SNPs selected from the sets of BMIX-associated SNPs with DF and DSS phenotypes, respectively.

				GWAS Discovery				Total cohort			
Phenotype	SNP (alleles)	Chromosome (position) HG37	Candidate gene	MAF cases	MAF controls	OR	P–value	MAF cases	MAF controls	OR	P–value
DF	rs2241811 (G/A)	2 (101011724)	CHST10	0.175	0.268	0.582	0.0003	0.181	0.256	0.643	0.0004132
DF	rs2241809 (G/A)	2 (101014363)	CHST10	0.248	0.365	0.573	3.54E–05	0.248	0.343	0.629	3.84E–05
DF	rs6555205 (A/G)	5 (360543)	AHRR	0.157	0.265	0.510	1.03E–05	0.157	0.249	0.562	1.02E–05
DF	rs1480010 (A/G)	12 (67076016)	GRIP1	0.105	0.174	0.557	0.001176	0.098	0.176	0.508	1.35E–05
DF	rs6573513 (G/A)	14 (63855760)	PPP2R5E	0.349	0.459	0.638	0.0003278	0.359	0.447	0.691	0.0004071
DF	rs7144210 (G/A)	14 (63870931)	PPP2R5E	0.347	0.457	0.633	0.000268	0.354	0.453	0.661	8.21E–05
DSS	rs7269910 (G/A)	20 (9363565)	PLCB4	0.116	0.195	0.553	0.003453	0.123	0.193	0.585	0.002431
DSS	rs1997696 (A/C)	20 (9378671)	PLCB4	0.289	0.381	0.667	0.007033	0.327	0.493	0.500	4.74E–08

S8 Table. Detailed identification of motif sequences across DENV serotypes.

DENV serotype	Motif type	Motif sequence	Start	Stop	DENV protein
DENV1	LxxLxE	LGELCEDTM	155	163	ER anchor for the protein C
DENV1	MxxIxE	MFLIAENKI	1333	1341	
DENV1	LxxLxE	LTTLWEGSP	2446	2454	NS4B
DENV1	LxxVxE	LKKVTEVKG	2587	2595	NS5
DENV1	LxxIxE	LCDIGESSP	2637	2645	NS5
DENV1	IxxIxE	ITDIMEPEH	3045	3053	NS5
DENV2	LxxLxE	LGELCEDTI	155	163	ER anchor for the protein C
DENV2	LxxIxE	LNLITEMGR	2097	2105	NS4A
DENV2	LxxLxE	LSELPETLE	2138	2146	NS4A
DENV2	LxxVxE	LKNVREVKG	2585	2593	NS5
DENV2	LxxIxE	LCDIGESSP	2635	2643	NS5
DENV2	LxxLxE	LGFLNEDHW	2980	2988	NS5
DENV3	CxxIxE	CPHITEVEP	167	175	ER anchor for the protein C
DENV3	IxxIxE	IRPINEKEE	1108	1116	NS1
DENV3	LxxLxE	LAILFEEVM	1140	1148	NS2A
DENV3	LxxVxE	LDLVTEIGR	2096	2104	NS4A
DENV3	LxxVxE	LKKVTEVRG	2584	2592	NS5
DENV3	LxxIxE	LCDIGESSP	2634	2642	NS5
DENV3	MxxIxE	MDVIGERIK	2763	2771	NS5
DENV3	IxxIxE	IKRIKEEHN	2770	2778	NS5
DENV3	LxxLxE	LGFLNEDHW	2978	2986	NS5
DENV4	LxxLxE	LRFLGEDGC	1095	1103	NS1
DENV4	IxxIxE	IRDVEETNM	1430	1438	NS2B
DENV4	LxxLxE	LDILTEIAS	2096	2104	NS4A
DENV4	LxxLxE	LNELPESLE	2137	2145	NS4A
DENV4	LxxVxE	LLWVAEIQP	2186	2194	NS4A
DENV4	IxxIxE	IRWIVERGM	2549	2557	NS5
DENV4	LxxVxE	LKNVTEVKG	2581	2589	NS5
DENV4	LxxIxE	LCDIGESSS	2631	2639	NS5
DENV4	LxxLxE	LQRLQEEHK	2768	2776	NS5
DENV4	LxxLxE	LGFLNEDHW	2976	2984	NS5
DENV4	LxxIxE	LGYILEEID	3002	3010	NS5

S9 Table. Information on DENV serotype and primary/secondary infection in the Thai cohort.

	DF	DSS
Serotype		
DENV1	85	21
DENV2	38	36
DENV3	15	3
DENV4	41	9
Not available	73	90
Immune Status		
Primary	32	6
Secondary	194	135
Not available	26	18

S10 Table. Primers used for mutagenesis and DNA assembly protocols.

| Gene	Description	Forward Primer (5'-3')	Reverse Primer (5'-3')
DENV2–NS5	Mutagenesis	AATCCCACGGTAGAAGCA	TGTGTCACACTTTTCTGG
DENV1–NS5	Assembly insert	– GCTTGGTACCATGGGCACGG GAGCCCAA	CTCCACCCCCCCAGAGTGCCC CTTCGGG
DENV1–NS5	Assembly plasmid	– GGCACTCTGGGGGGGTGGAG GCTCTGATAG	CCGTGCCCATGGTACCAAGCT TGGTGGCGG

S1 Text

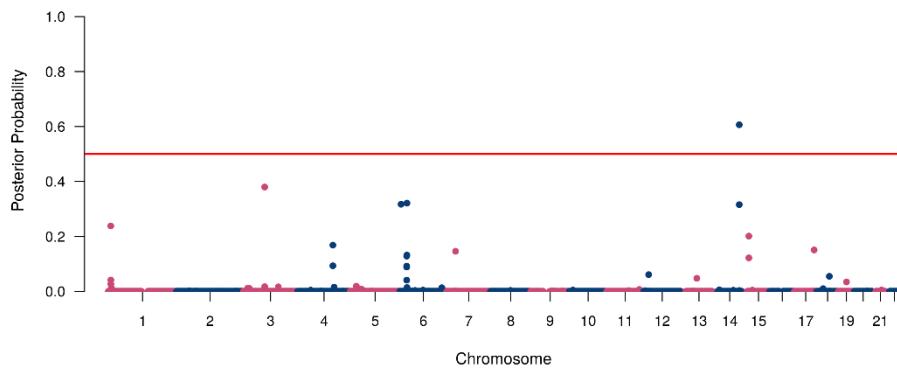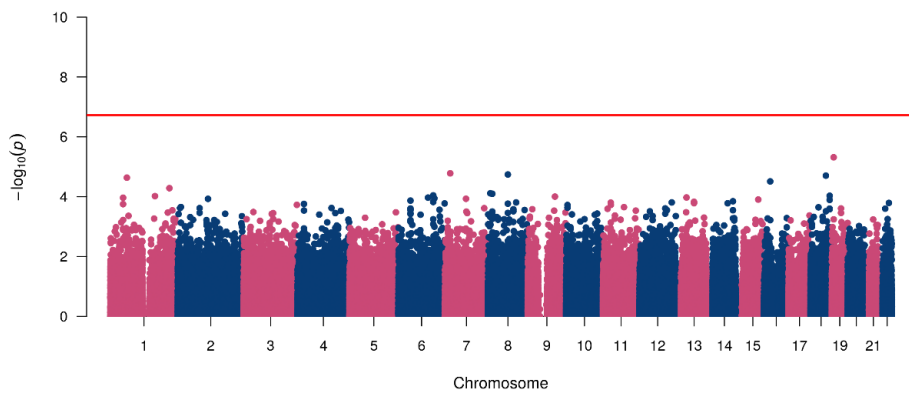| Chr	SNP	BP	Allele	Association p-value	OR	BMIX posterior p-value in Northeast Asian ancestry	BMIX posterior p-value in Southeast Asian ancestry	Gene
				RUN 1				
7	rs7799285	100968363	G	0.0009861	0.6005		0.5479738	*RABL5*
9	rs12683380	9929916	A	0.0005549	1.629	0.8485055		*PTPRD*
21	rs2828104	24764752	C	7.276e-05	1.746	0.5206194	0.7748023	*EEF1A1P1*

Association
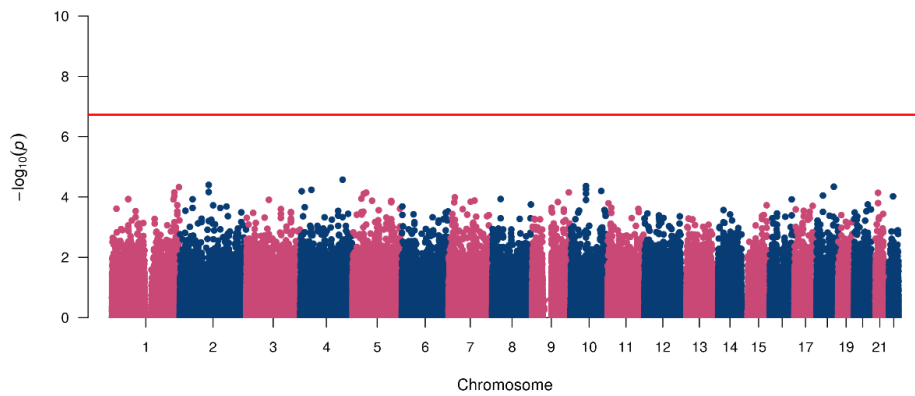
BMIX – NORTHEAST ASIAN
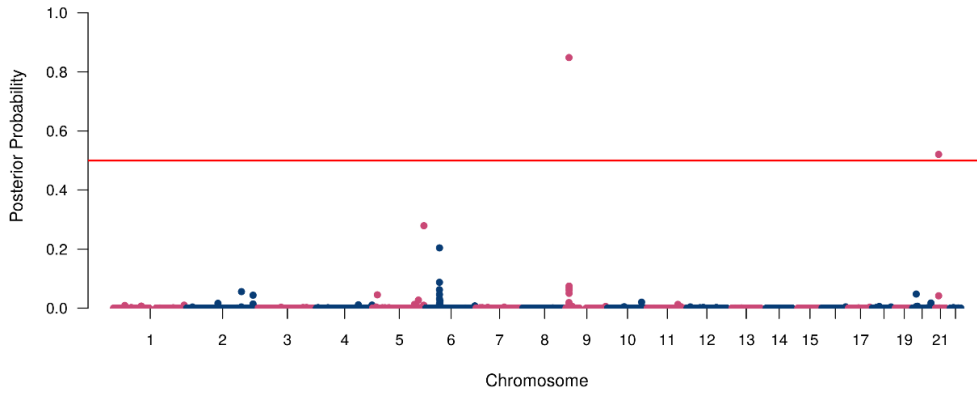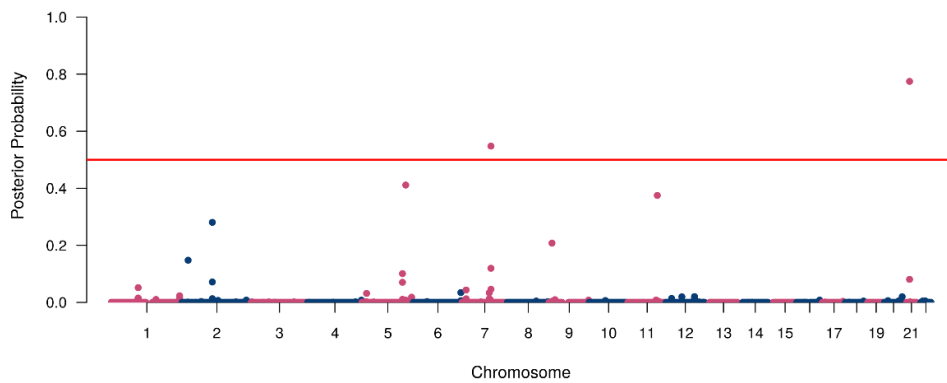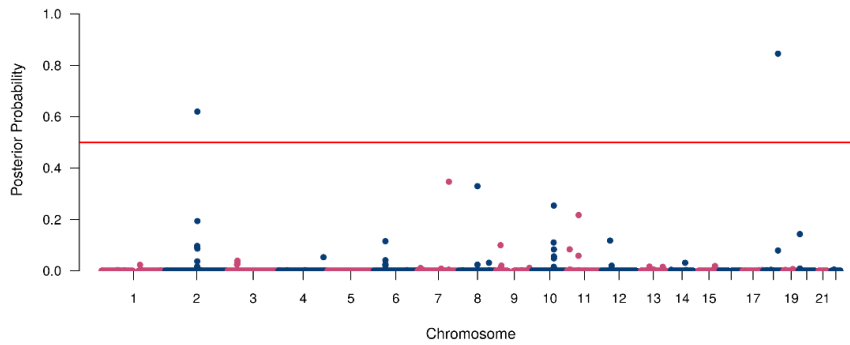

BMIX – SOUTHEAST ASIAN

Chr	SNP	BP	Allele	Association p-value	OR	BMIX posterior p-value in Northeast Asian ancestry	BMIX posterior p-value in Southeast Asian ancestry	Gene
RUN 2								
22	rs5761313	26313745	T	0.0001115	1.919	0.5009741	0.8918858	*MYO18B*
22	rs2301504	26320015	C	0.0002039	1.835		0.5818690	*MYO18B*

Association
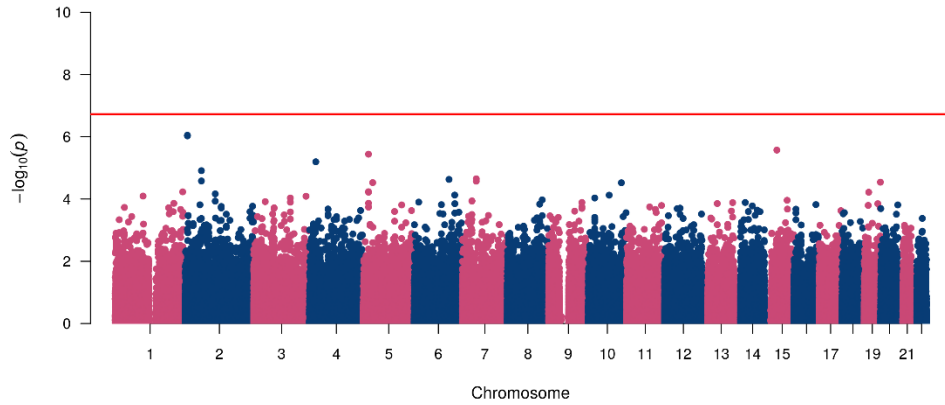

BMIX – NORTHEAST ASIAN
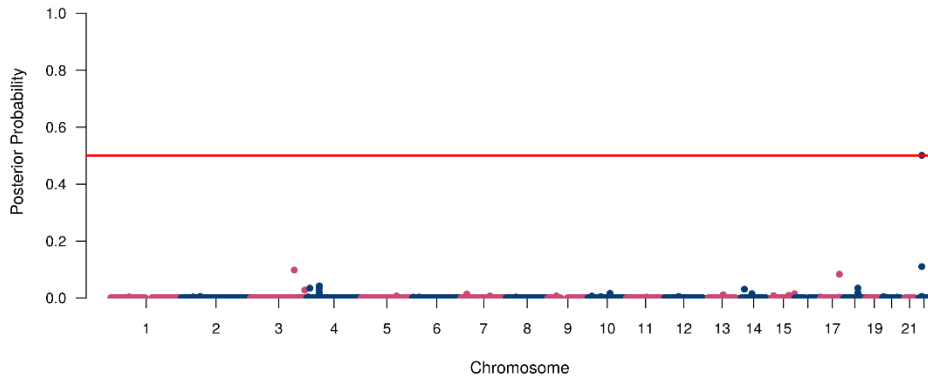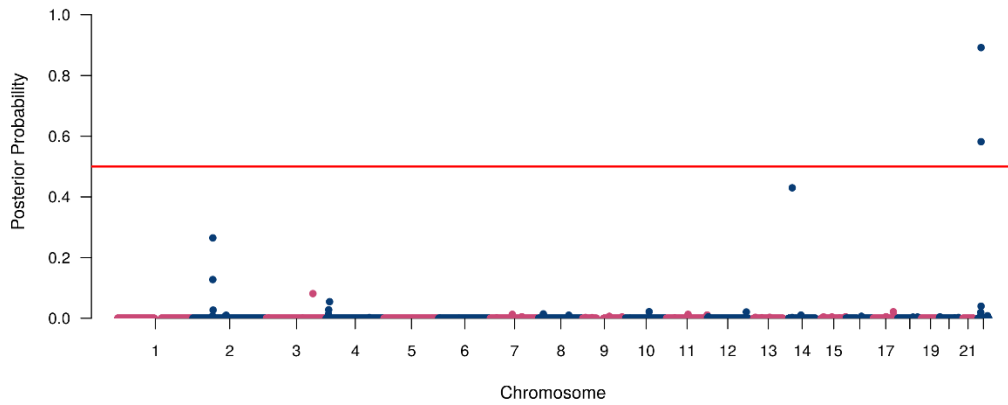
BMIX – SOUTHEAST ASIAN

284

Chr	SNP	BP	Allele	Association p–value	OR	BMIX posterior p–value in Northeast Asian ancestry	BMIX posterior p–value in Southeast Asian ancestry	Gene
RUN 3								
6	rs9496739	144020900	A	3.457e–05	0.551	0.5527984		*PHACTR2*
17	rs12709500	26134974	C	1.326e–05	0.4801		0.7308197	Intergenic
17	rs16949120	48610732	A	8.477e–05	2.491	0.9197636		*EPN3/MYCBPAP*

Association
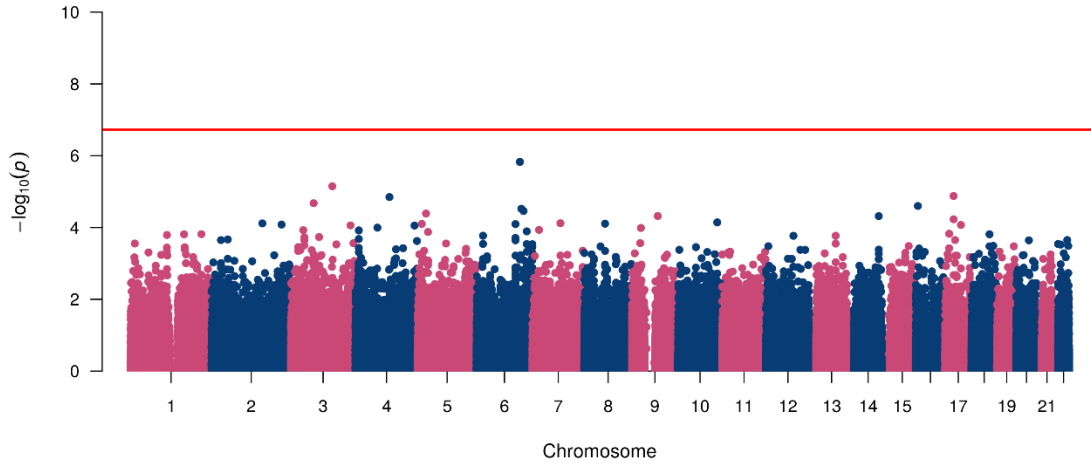
BMIX – NORTHEAST ASIAN
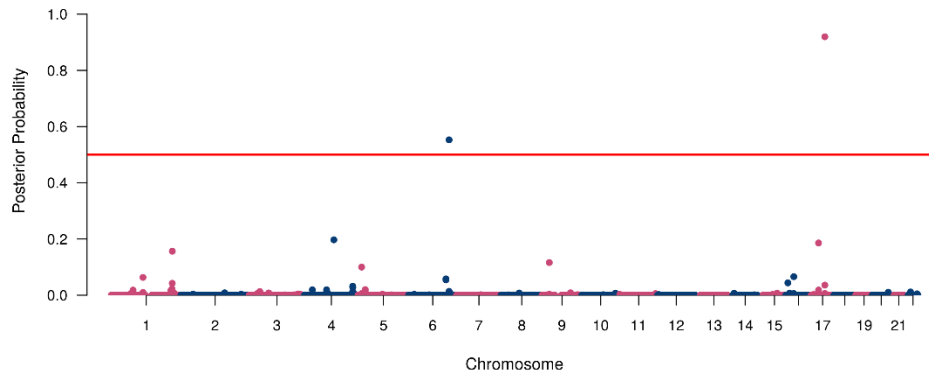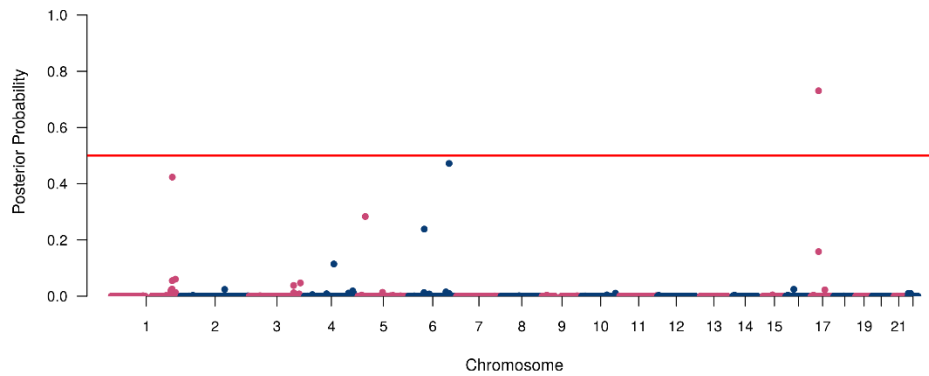
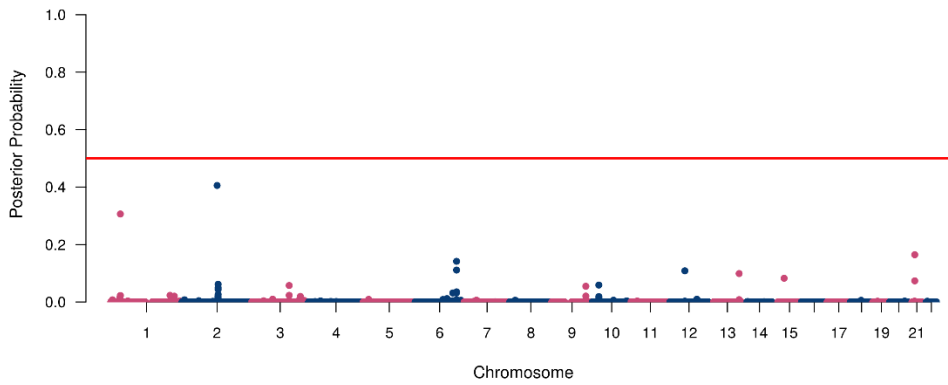

BMIX – SOUTHEAST ASIAN

285

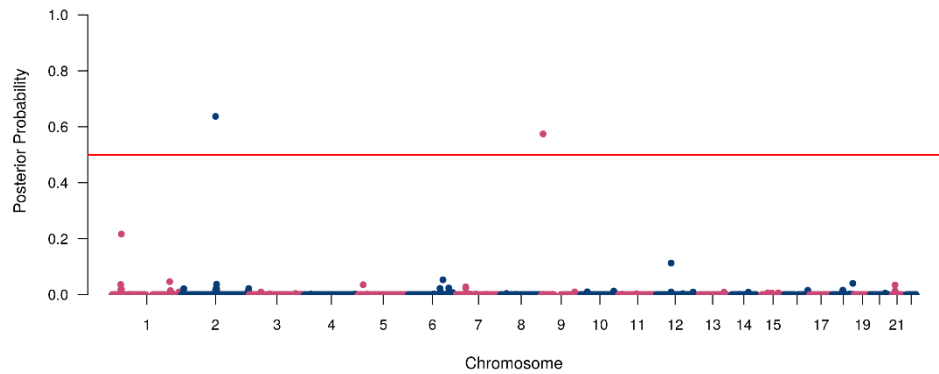| Chr	SNP	BP	Allele	Association p–value	OR	BMIX posterior p–value in Northeast Asian ancestry	BMIX posterior p–value in Southeast Asian ancestry	Gene
				RUN 4				
1	rs2268170	9315847	T	2.19e–05	2.182	0.8404511		*H6PD*
14	rs1456988	98488007	T	0.0005011	1.631		0.6064477	Intergenic
15	rs1553893	26114658	C	0.0003323	2.802	0.6392913		*ATP10A*

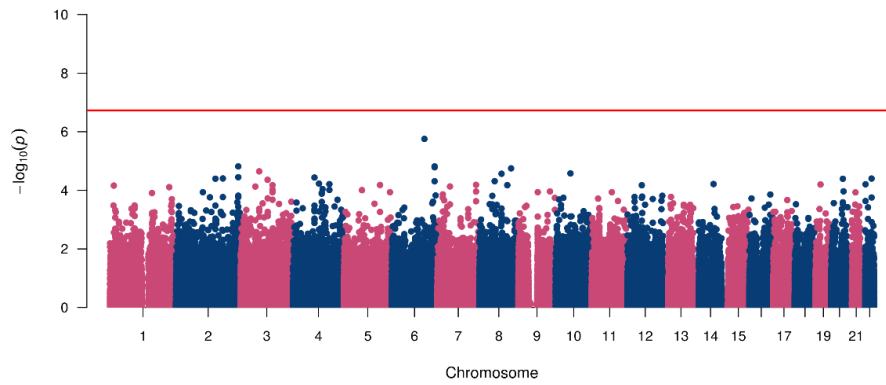## Association
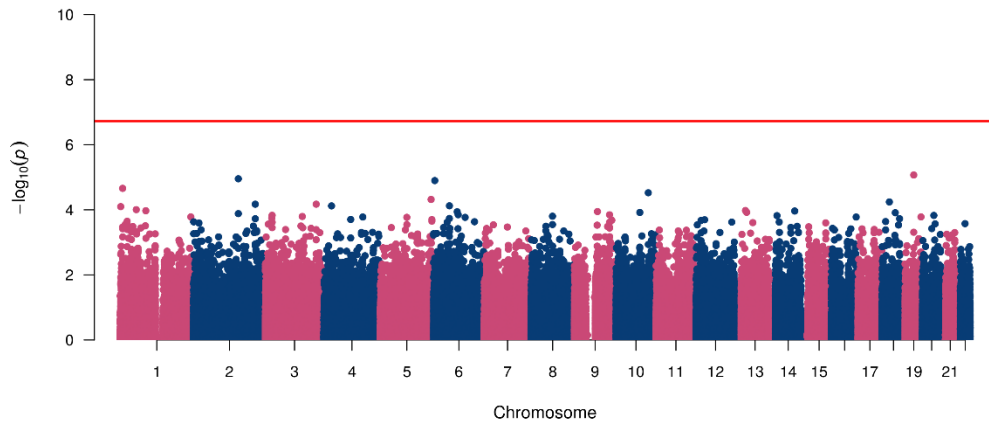
BMIX – NORTHEAST ASIAN
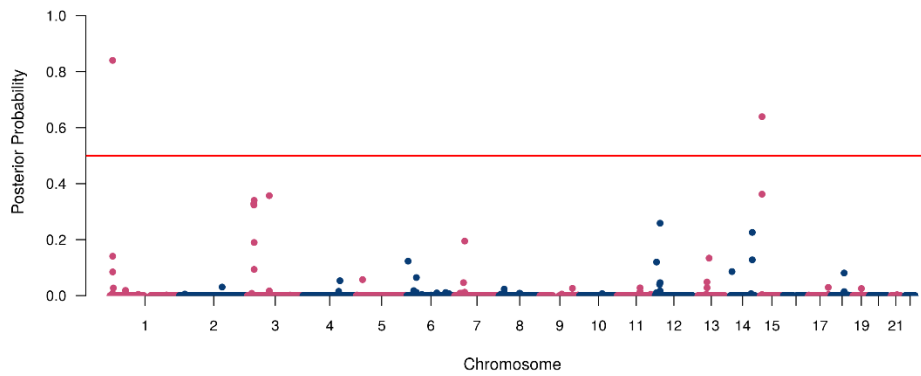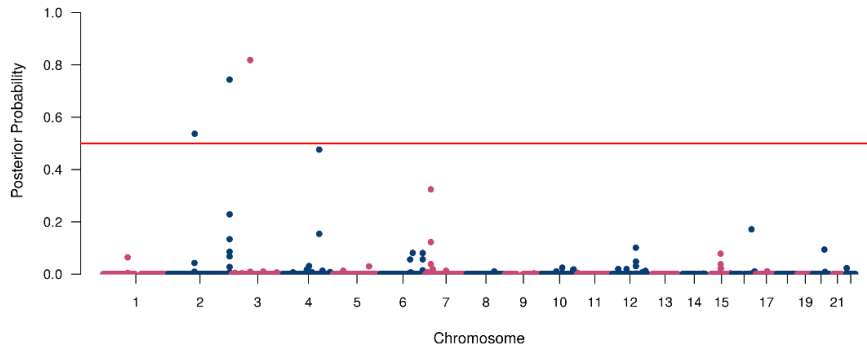

BMIX – SOUTHEAST ASIAN

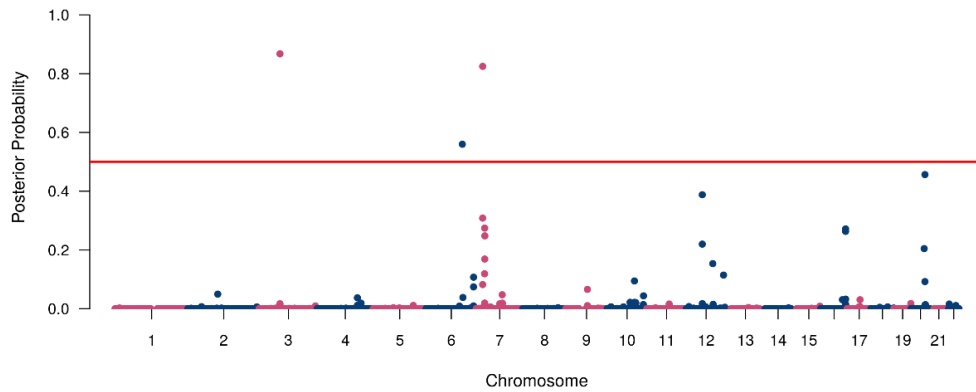| Chr	SNP	BP	Allele	Association p-value	OR	BMIX posterior p-value in Northeast Asian ancestry	BMIX posterior p-value in Southeast Asian ancestry	Gene
				RUN 5				
2	rs12990473	124454519	A	0.002542	0.652	0.6197150		Intergenic
8	rs7002197	73123327	G	1.813e−05	1.978		0.7725738	Intergenic
9	rs12683636	16862691	C	0.0002618	2.008		0.7176504	*BNC2*
11	rs10838216	44038877	C	0.0005432	1.752		0.7115631	Intergenic
12	rs7305703	38138381	C	0.0003085	2.607		0.5054825	Intergenic
18	rs7226876	56532820	T	1.979e−05	2.123	0.8451823		*ZNF532*

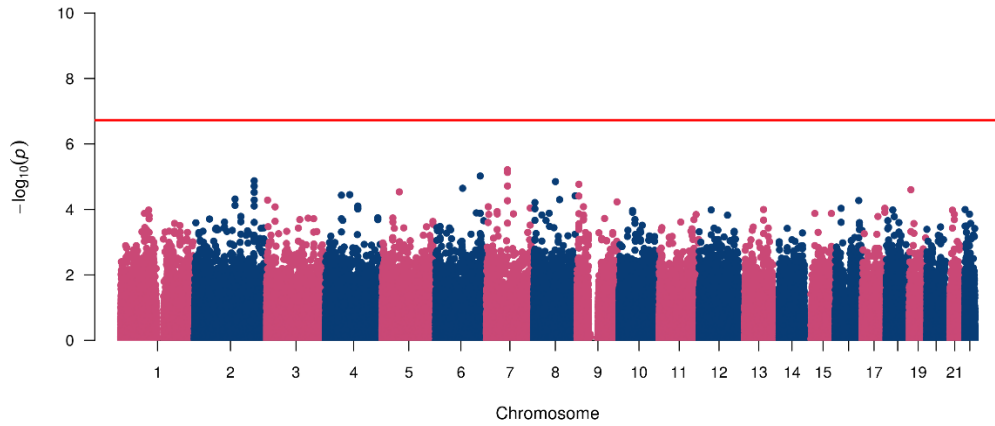## Association

BMIX – NORTHEAST ASIAN
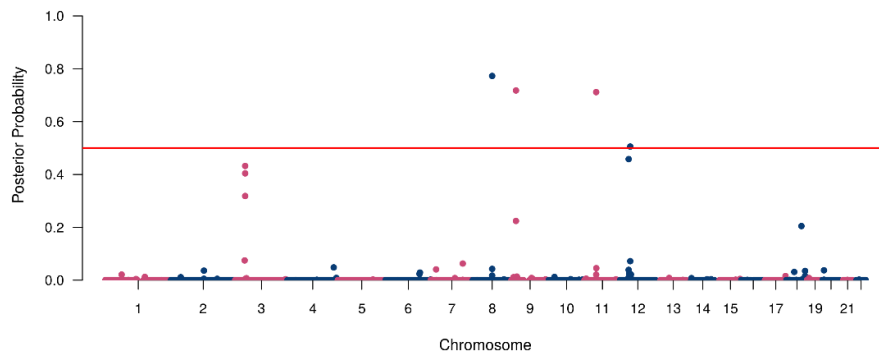
BMIX – SOUTHEAST ASIAN

Chr	SNP	BP	Allele	Association p–value	OR	BMIX posterior p–value in Northeast Asian ancestry	BMIX posterior p–value in Southeast Asian ancestry	Gene
RUN 6								
2	rs7602673	121853648	C	0.0004952	1.633		0.6372767	Intergenic
9	rs2295870	5436461	C	0.0003478	0.5824		0.5744294	*PLGRKT*

Association

BMIX – NORTHEAST ASIAN
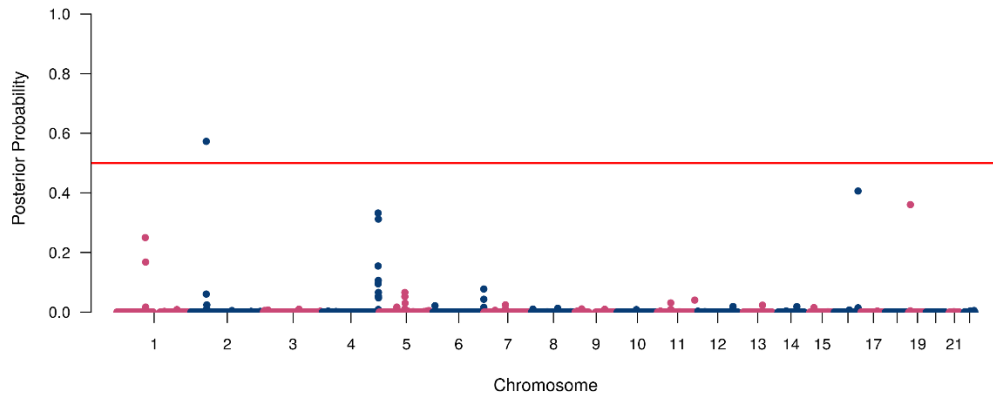

BMIX – SOUTHEAST ASIAN

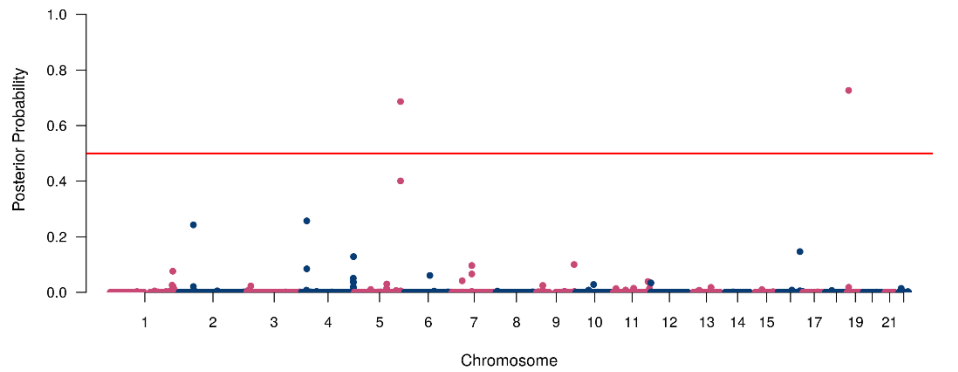| Chr	SNP	BP	Allele	Association p-value	OR	BMIX posterior p-value in Northeast Asian ancestry	BMIX posterior p-value in Southeast Asian ancestry	Gene
RUN 7								
2	rs4851346	101223405	T	0.0001164	1.736	0.5367466		Intergenic
2	rs12988520	234607394	C	1.53e-05	1.870	0.7438419		*UGT1A6/7/8/9/10*
3	rs805478	70427067	T	2.257e-05	2.365	0.8182922	0.8677060	Intergenic
6	rs9385270	123092085	T	1.767e-06	1.983		0.5599361	Intergenic
7	rs7811417	21534152	T	0.0001379	0.580		0.8249133	*SP4*

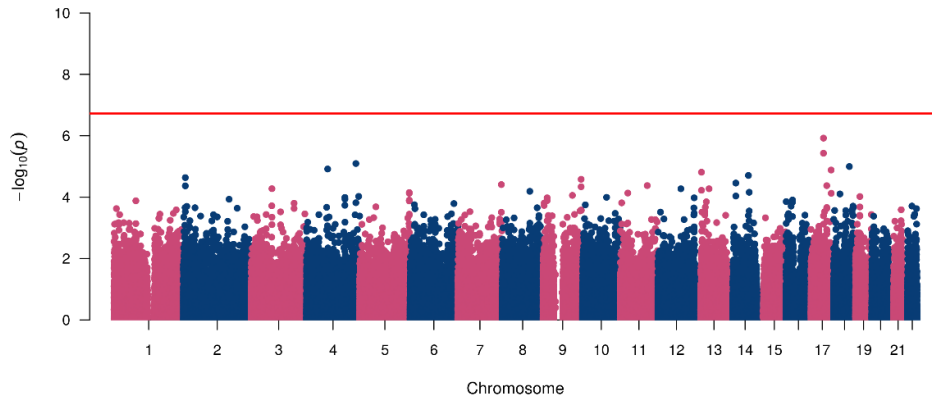## Association

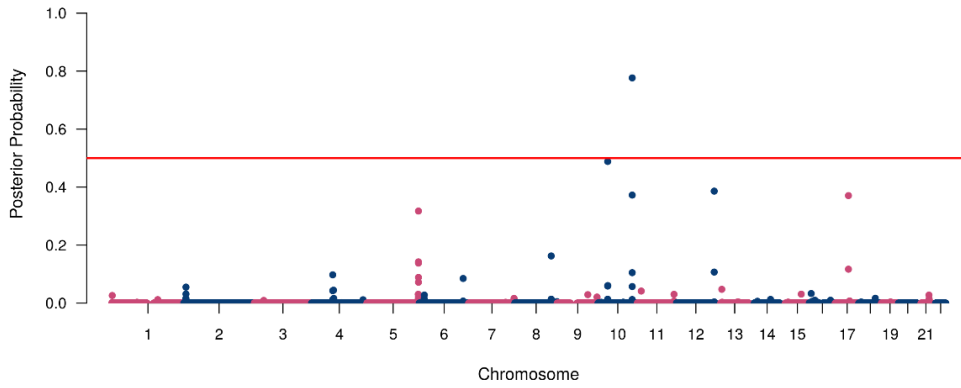## BMIX – NORTHEAST ASIAN

BMIX – SOUTHEAST ASIAN

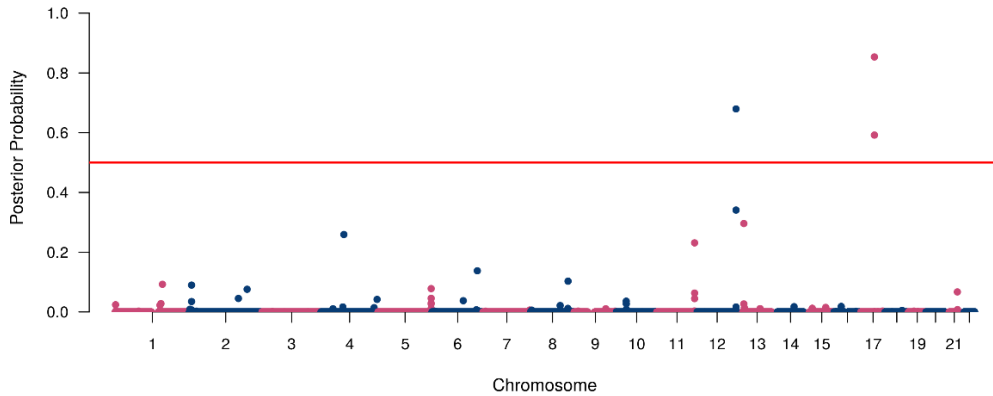| Chr	SNP	BP	Allele	Association p-value	OR	BMIX posterior p-value in Northeast Asian ancestry	BMIX posterior p-value in Southeast Asian ancestry	Gene
RUN 8								
2	rs4246580	50897456	G	0.0001965	1.757	0.5729017		*NRXN1*
5	rs1450624	165956290	G	0.0008588	1.769		0.6864852	Intergenic
19	rs2285963	5591735	A	2.499e−05	2.509		0.7265687	*SAFB2*

## Association

BMIX – NORTHEAST ASIAN

BMIX – SOUTHEAST ASIAN

Chr	SNP	BP	Allele	Association p–value	OR	BMIX posterior p–value in Northeast Asian ancestry	BMIX posterior p–value in Southeast Asian ancestry	Gene
				RUN 9				
10	rs2907567	117618612	T	0.0003373	0.489	0.7767757		*ATRNL1*
12	rs12314724	130274359	C	0.0001051	0.452		0.6794178	*TMEM132D*
17	rs11869840	44996100	A	1.194e–06	1.985		0.8533168	*GOSR2*
17	rs1662576	44996245	T	3.702e–06	1.923		0.5918913	*GOSR2*

Association

BMIX – NORTHEAST ASIAN

BMIX – SOUTHEAST ASIAN


291

| Chr	SNP	BP	Allele	Association p-value	OR	BMIX posterior p-value in Northeast Asian ancestry	BMIX posterior p-value in Southeast Asian ancestry	Gene
				RUN 10				
7	rs9656687	53635159	A	8.489e-05	0.495		0.7990117	Intergenic
7	rs7803594	53654002	T	0.0001859	0.588		0.5609658	Intergenic
7	rs672416	105376750	T	2.829e-06	2.005		0.6579472	*ATXN7L1*
7	rs577004	105383875	C	3.066e-06	2.018		0.5470686	*ATXN7L1*

Association
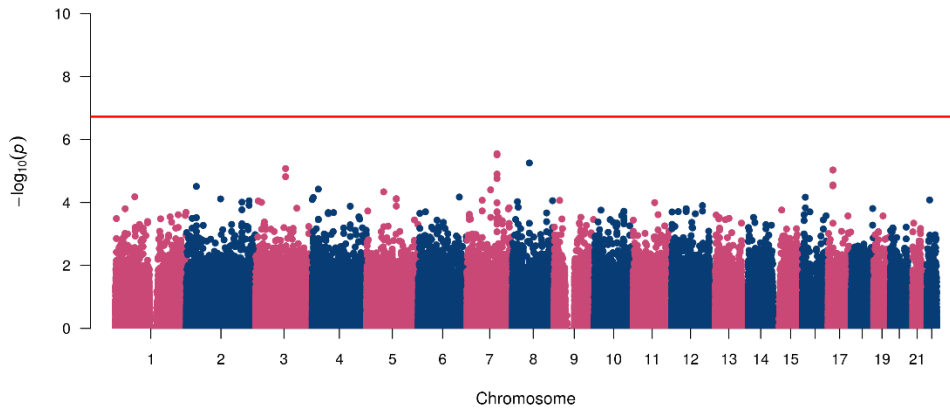
BMIX – NORTHEAST ASIAN
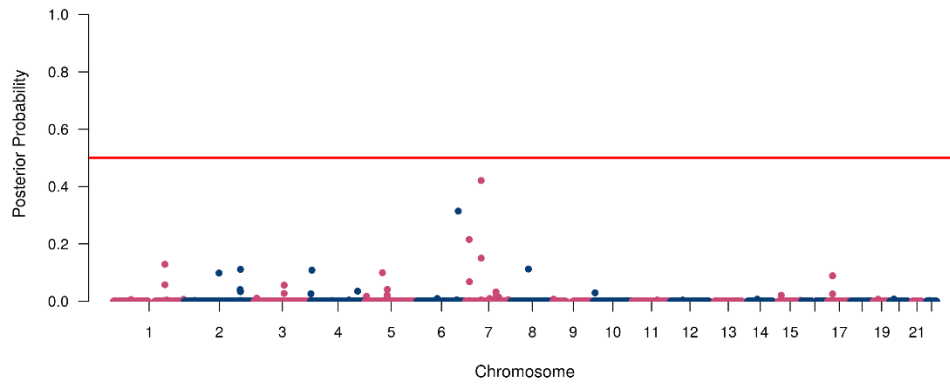
BMIX – SOUTHEAST ASIAN
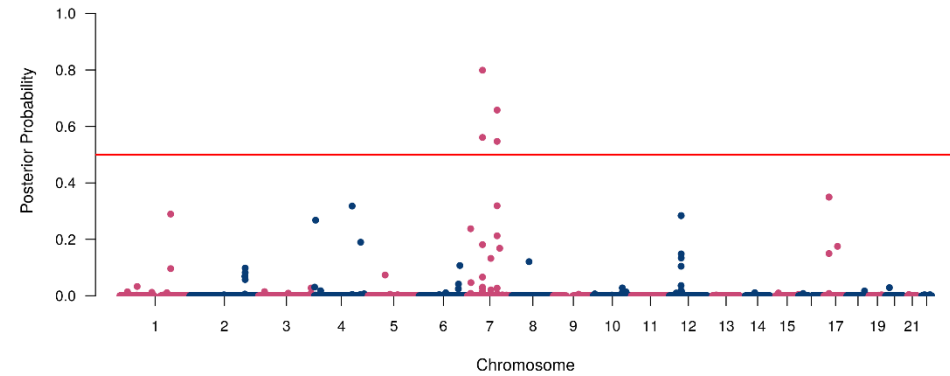
Appendix C – Supplementary information from paper III

Functional evaluation of *GRIP1* and *PPP2R5E* involvement in dengue fever

Marisa Oliveira, Marina Penova, Matthieu Prot, Isabelle Casademont, Yves Jacob, Luisa Pereira, Etienne Simon-Loriere and Anavaj Sakuntabhai

Table C.1 – Primers used for pcDNA3.1 cloning and CRISPR/Cas9 guideRNAs.

GRIP1

	Forward (5'–3')	Reverse (5'–3')
Cloning (restriction enzyme)	ATTGGCTAGCATGATAGCTGTCTCTTTT (NheI)	ACCTACTCGAGCTATAATGTATTAGTGGGT (XhoI)
Crispr	CACCGCTTCATCAGCTCGACGACTG	AAACCAGTCGTCGAGCTGATGAAGC

PPP2R5E

	Forward (5'–3')	Reverse (5'–3')
Assembly – vector (restriction enzyme)	TTCCAACTTAACTCGAGTCTAGAGGGCCC (NheI)	AGGACATTGCAAGCTTAAGTTTAAACGCTAGC (XhoI)
Assembly – insert	ACTTAAGCTTGCAATGTCCTCAGCACCAAC	AGACTCGAGTTAAGTTGGAATTATTCCATCACGTC
Crispr	CACCGAACCTACCCTTGAGGCATCG	AAACCGATGCCTCAAGGGTAGGTTC

Table C.2 – PCR Primers designed for CRISPR/Cas9 cell lines editing confirmation.

GRIP1

	Forward(5'–3')	Reverse (5'–3')
PCR (cell lines)	CCTCCTTGCCGCAGATTAGAT	TGGAAAGAACTTCTGGGGAGGA

PPP2R5E

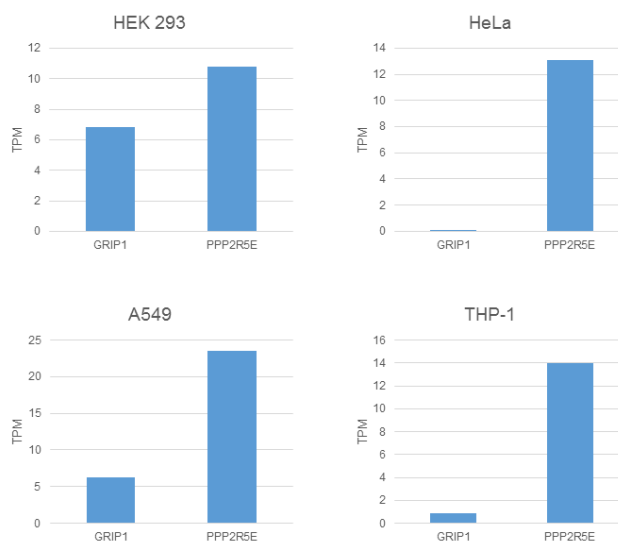|	Forward (5'–3')	Reverse (5'–3')
PCR (cell lines)	GCAGAAGGTGCTTCCCGTTA	GCTCTAGAGGGCAGGATCATA

Figure C.1– mRNA levels (in transcripts per million) for *GRIP1* and *PPP2R5E* in four different cell lines. Data extracted from The Human Protein Atlas (https://www.proteinatlas.org/).
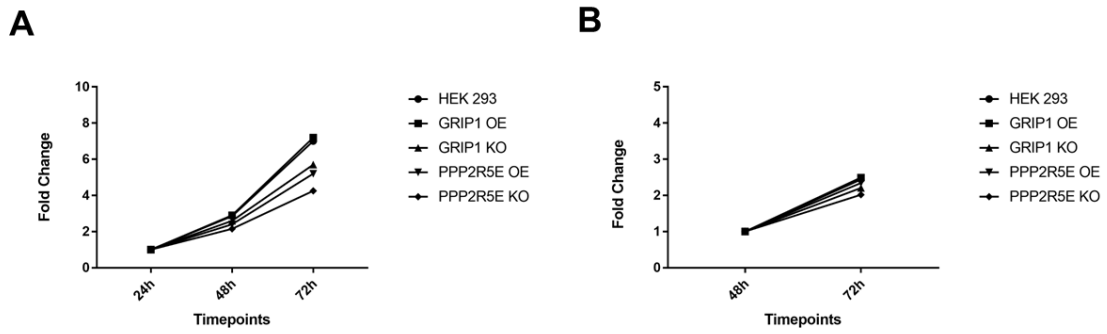
Figure C.2– Proliferation of wild type HEK 293 and transformed cell lines. (A) Fold Change at 48h and 72h after plating (normalize to 24h values); (B) Fold Change at 72h after plating (normalize to 48h values).
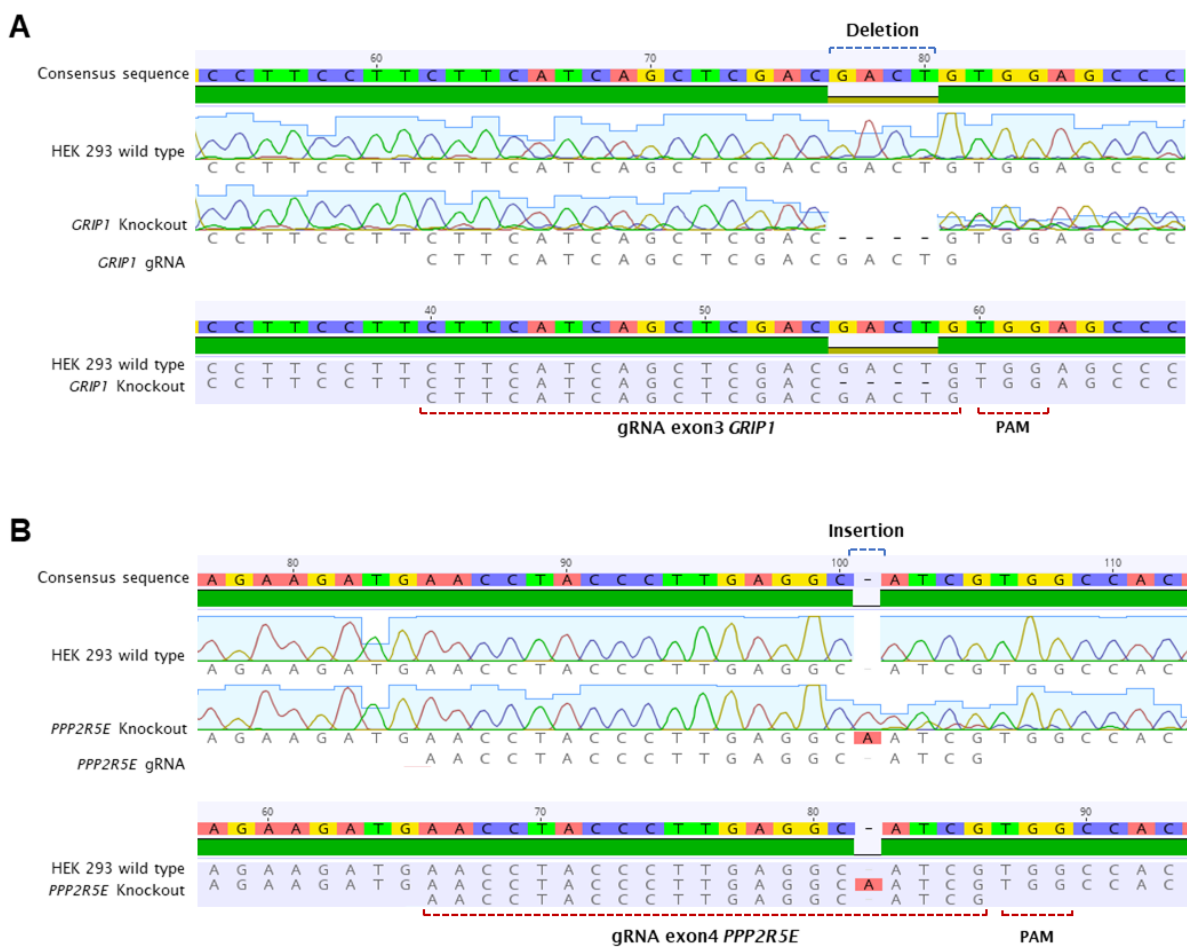

Figure C.3– CRISPR/Cas9 editing confirmation by Sanger sequencing. (A) Detection of a 4 bp deletion in *GRIP1* knockout cell line; (B) Detection of a 1 bp insertion in *PPP2R5E* knockout cell line.