Hannah Love

Investigation of dengue virus envelope gene quasispecies variation in patient samples: implications for virus virulence and disease pathogenesis

Cranfield Health PhD 2011

Supervisors: Prof. David Cullen (Cranfield University)
Dr. Jane Burton (Health Protection Agency)
Dr. Kevin Richards (Health Protection Agency)

This thesis is submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy

September 2011

ABSTRACT

Abstract

Due to the error-prone nature of RNA virus replication, each dengue virus (DV) exists as a quasispecies within the host. To investigate the hypothesis that DV quasispecies populations affect disease severity, serum samples were obtained from dengue patients hospitalised in Ragama, Sri Lanka. From the patient sera, DV envelope glycoprotein (E) genes were amplified by high-fidelity RT-PCR, cloned, and multiple clones per sample sequenced to identify mutations within the quasispecies population. A mean quasispecies diversity of 0.018% was observed, consistent with reported error rates for viral RNA polymerases (0.01%; Smith *et al.*, 1997). However, previous studies reported 8.9 to 21.1-fold greater mean diversities (0.16% to 0.38%; Craig *et al.*, 2003; Lin *et al.*, 2004; Wang *et al.*, 2002a). This discrepancy was shown to result from the lower fidelity of the RT-PCR enzymes used by these groups for viral RNA amplification. Previous studies should therefore be re-examined to account for the high number of mutations introduced by the amplification process.

Nonsynonymous mutation locations were modelled to the crystal structure of DV E, identifying those with the potential to affect virulence due to their proximity to important structural features. No correlation was observed between the extent of quasispecies variation and disease severity. However, genome-defective quasispecies variants, and variants with surface accessible amino acid substitutions, or those proximal to the fusion peptide, proposed receptor binding sites, or other E oligomers, were observed predominantly in patients with severe dengue. Recombinant virus-like particles were produced for nine quasispecies variants, and the effects of the mutations on protein function assessed. Altered heparin binding abilities were demonstrated for four of the nine variants, indicative of differing cell attachment capabilities. Further work is required to assess differences in antibody binding, replication efficiency, virion and oligomer assembly, low pH-induced conformational changes required for fusion, and transmissibility of variants.

ACKNOWLEDGEMENTS

This research project would not have been possible without the support of many people. Firstly I would like to thank my supervisors Prof. David Cullen, Dr. Jane Burton and Dr. Kevin Richards for their guidance and support during this project.

I am grateful to Major Mark Bailey of the Royal Army Medical Corps for allowing me to use the dengue samples and clinical data from his fever study, and for providing medical expertise in the early stages of this project regarding the dengue patient disease severity classifications. Also thanks to his colleagues Dr. Ranjan Premaratna and Prof. Janaka de Silva from the Department of Medicine at the University of Kelaniya in Sri Lanka.

I would also like to express my thanks to my colleagues at HPA CEPR, in particular the Diagnostic Support and C20 Project groups, for giving me the time and invaluable support to complete this project. Special thanks go to Dr. Richard Vipond, Dr. Daniel Bailey and Dr. Brian Dove for their expert advice and troubleshooting over the last four years; Dr. Stuart Dowall for his comments on Chapter 6, Howard Tolley for electron microscopy, Dr. Jane Osborne for help with phylogenetic analysis, Steve Kidd for initial guidance with ELISAs, and also to Mike Stubbington for cloning and sequencing advice.

A huge thank you is due to my friends, for their support and encouragement, but mostly to my fellow PhD students at HPA CEPR, Ruth, Debs, Leanne, Simon, H and Lotte, for sharing the pain and making it a bit more bearable.

Finally the biggest thanks go to my family, for putting up with me and learning when it's best not to ask how it's going; to Jenny, for sending the care package that made me smile; and especially to Simon, for his love, patience and understanding, and for looking after me.

LIST OF CONTENTS

ABSTRACT I

ACKNOWLEDGEMENTS III

LIST OF CONTENTS V

LIST OF FIGURES X

LIST OF TABLES XIV

NOTATION XVII

CHAPTER 1. INTRODUCTION 1

1.1 THESIS OVERVIEW 2
1.2 DENGUE 3
1.2.1 Aetiology of dengue 3
1.2.2 Symptoms of dengue 6
1.2.3 Clinical and laboratory diagnosis of dengue 6
1.3 DENGUE VIRUSES 12
1.3.1 Dengue virus structure and genome organisation 12
1.3.2 Dengue virus replication 13
1.3.3 Dengue virus binding to target host cells 15
1.4 DETERMINANTS OF VIRAL REPLICATION AND DISEASE PATHOGENESIS 17
1.4.1 Antibody-dependent enhancement of dengue virus infection 17
1.4.2 Dengue virus sequence variation and virulence 18
1.4.3 Dengue in Sri Lanka 25
1.4.4 Recombination in dengue viruses 26
1.4.5 Dengue virus as a quasispecies 27
1.5 FLAVIVIRUS ENVELOPE PROTEINS AND THEIR ROLE IN VIRULENCE 30
1.5.1 Flavivirus envelope protein structure 30
1.5.2 Flavivirus envelope gene mutations and altered pathogenesis 33
1.6 HYPOTHESIS, AIMS AND OBJECTIVES 37
1.6.1 Hypothesis 37
1.6.2 Aims and objectives 37

CHAPTER 2. MATERIALS AND METHODS 40

2.1 GENERAL INFORMATION 41
2.1.1 Chemicals and reagents 41
2.1.2 Oligonucleotides 41
2.1.3 Flavivirus RNA used as a positive control template for RT-PCR 41
2.1.4 Dengue patient material 41
2.1.5 Bacterial strains and culture media 42
2.1.6 Insect cell lines and culture medium 42
2.1.7 Insect viruses and virus transfer vectors 43
2.2 IN SILICO TECHNIQUES 44
2.2.1 Acquisition of published Flavivirus sequence data 44
2.2.2 Sequence analysis for primer design 44
2.2.3 RT-PCR, PCR and sequencing PCR primer design 44
2.2.4 Analysis of sequencing data 45
2.2.5 Statistical analysis 45
2.2.6 Protein modelling 45
2.3 NUCLEIC ACID MANIPULATIONS 47
2.3.1 Amplification of nucleic acid sequences using the polymerase chain reaction 47
2.3.2 RNA extraction using commercial spin columns 50
2.3.3 DNA extraction using commercial spin columns 51
2.3.4 Plasmid DNA extraction using commercial filter kits 53
2.3.5 Cloning of dengue virus quasispecies envelope genes, transformations and amplification of positive clones for quasispecies analysis 54

2.3.6 Cloning of dengue virus quasispecies envelope genes, transformations and amplification of positive clones for recombinant baculovirus production *55*
2.3.7 Storage of transformed E. coli *56*
2.3.8 Restriction digestion of plasmid DNA *56*
2.3.9 Quantification of nucleic acids *56*
2.3.10 In vitro transcription of plasmid DNA *57*
2.4 SEQUENCING 58
2.4.1 Sequencing PCR *58*
2.4.2 Sequencing PCR clean-up *58*
2.4.3 Sequencing *58*
2.5 ELECTROPHORETIC TECHNIQUES 59
2.5.1 DNA electrophoresis *59*
2.5.2 RNA electrophoresis *59*
2.5.3 Polyacrylamide gel electrophoresis *59*
2.6 WESTERN BLOTS AND IMMUNODETECTION 61
2.7 BACULOVIRUS MANIPULATIONS 63
2.7.1 Insect cell culture *63*
2.7.2 Production of recombinant baculoviruses *63*
2.7.3 β-galactosidase assay *64*
2.7.4 Amplification and titration of recombinant baculovirus stocks *64*
2.7.5 Recombinant dengue virus-like particle production *66*
2.8 DENGUE VIRUS-LIKE PARTICLE CONCENTRATION, PURIFICATION AND QUANTIFICATION 68
2.8.1 Ultracentrifugation of recombinant dengue virus-like particles *68*
2.8.2 His GraviTrap purification *68*
2.8.3 Protein G purification *68*
2.8.4 Anion exchange column purification *70*
2.8.5 Quantification of total protein by Bradford assay *71*
2.9 ENZYME-LINKED IMMUNOSORBENT ASSAYS (ELISAS) 72
2.9.1 Dengue virus-like particle ELISAs *72*
2.9.2 Heparin and DC-SIGN capture ELISAs *73*
2.10 ELECTRON MICROSCOPY 74

CHAPTER 3. ANALYSIS OF DENGUE VIRUS QUASISPECIES POPULATIONS IN CLINICAL SAMPLES 75

3.1 INTRODUCTION 76
3.2 CHAPTER OBJECTIVES 78
3.3 DESIGN AND TESTING OF PRIMERS FOR AMPLIFICATION OF DENGUE VIRUS ENVELOPE GENES 80
3.4 DENGUE PATIENT SAMPLES USED FOR AMPLIFICATION OF DENGUE VIRUS ENVELOPE GENES AND QUASISPECIES ANALYSIS 85
3.5 HIGH-FIDELITY RT-PCR, CLONING AND SEQUENCING OF DENGUE VIRUS QUASISPECIES ENVELOPE GENES FROM PATIENT SAMPLES 90
3.6 SEQUENCE ANALYSIS OF DENGUE VIRUS QUASISPECIES ENVELOPE GENES 101
3.7 COMPARISON OF MUTATIONS INTRODUCED BY THE RT-PCR PROCESS BETWEEN THIS STUDY AND PUBLISHED DATA 105
3.8 PHYLOGENETIC ANALYSIS OF DENGUE VIRUS QUASISPECIES ENVELOPE GENE AND GLYCOPROTEIN SEQUENCES 109
3.9 LINEAR AND THREE-DIMENSIONAL MAPPING OF DENGUE VIRUS QUASISPECIES ENVELOPE GLYCOPROTEIN AMINO ACID SUBSTITUTIONS 118
3.10 IDENTIFICATION OF DENGUE VIRUS QUASISPECIES ENVELOPE GENES FOR RECOMBINANT PROTEIN PRODUCTION 125
3.11 DISCUSSION 128
3.11.1 Patient samples used for dengue virus quasispecies envelope gene amplification *128*
3.11.2 Amplification of dengue virus envelope genes by high-fidelity RT-PCR *128*
3.11.3 Cloning of dengue virus quasispecies envelope genes *129*
3.11.4 Sequencing of dengue virus quasispecies envelope genes *131*
3.11.5 Analysis of dengue virus envelope gene quasispecies variation *134*

3.11.6 Measurement of mutations introduced via high-fidelity RT-PCR amplification and comparison of methods used in previous studies *136*
3.11.7 Phylogenetic analyses of dengue virus envelope gene quasispecies variation *138*
3.11.8 Linear and three-dimensional mapping of dengue virus quasispecies envelope glycoprotein amino acid substitutions *140*
3.11.9 Identification of dengue virus quasispecies envelope genes for recombinant protein production *144*
3.12 CONCLUSIONS AND SUMMARY 145

CHAPTER 4. PRODUCTION AND PURIFICATION OF RECOMBINANT DENGUE VIRUS ENVELOPE GLYCOPROTEINS 147

4.1 INTRODUCTION 148
4.2 CHAPTER OBJECTIVES 150
4.3 CLONING OF DENGUE VIRUS PRE-MEMBRANE AND QUASISPECIES ENVELOPE GENES INTO THE BACULOVIRUS TRANSFER VECTOR, PBAC-2CP 151
4.4 PRODUCTION OF RECOMBINANT BACULOVIRUSES CONTAINING DENGUE VIRUS PRE-MEMBRANE AND ENVELOPE GENES 156
4.5 OPTIMISATION OF DENGUE VIRUS-LIKE PARTICLE PRODUCTION 159
4.6 CONCENTRATION OF DENGUE VIRUS-LIKE PARTICLES 163
4.7 PURIFICATION OF DENGUE VIRUS-LIKE PARTICLES 166
4.7.1 Immobilised metal affinity chromatography *166*
4.7.2 Immunoprecipitation *166*
4.7.3 Ion exchange chromatography *168*
4.7.4 Removal of baculovirus from virus-like particle samples by immunoprecipitation *171*
4.8 SCALED-UP PRODUCTION AND PURIFICATION OF DENGUE VIRUS-LIKE PARTICLES 173
4.8.1 Detection of dengue virus-like particles using transmission electron microscopy *179*
4.9 DISCUSSION 182
4.9.1 Cloning of dengue virus pre-membrane and envelope genes into the baculovirus transfer vector *183*
4.9.2 Production of recombinant baculoviruses containing dengue virus pre-membrane and envelope genes *183*
4.9.3 Optimisation of dengue virus-like particle production *183*
4.9.4 Concentration of dengue virus-like particles *185*
4.9.5 Purification of dengue virus-like particles *186*
4.9.6 Scaled-up production, purification and detection of dengue virus-like particles *188*
4.9.7 Detection of dengue virus-like particles using transmission electron microscopy *191*
4.10 CONCLUSIONS AND SUMMARY 193

CHAPTER 5. ANALYSIS OF DENGUE VIRUS-LIKE PARTICLES TO DETERMINE THE EFFECTS OF THE QUASISPECIES ENVELOPE GENE MUTATIONS 196

5.1 INTRODUCTION 197
5.2 CHAPTER OBJECTIVES 199
5.3 DEVELOPMENT AND OPTIMISATION OF AN ELISA USING RECOMBINANT DENGUE VIRUS-LIKE PARTICLES AS ANTIGEN 201
5.3.1 Initial ELISA evaluation *201*
5.3.2 Dengue virus-like particle mAb ELISA optimisation *207*
5.3.3 Analysis of the dengue virus-like particle samples using the optimised mAb ELISA *211*
5.4 COMPARISON OF DENGUE VIRUS-LIKE PARTICLE SAMPLES BINDING TO DC-SIGN VIA A CAPTURE ELISA 214
5.5 COMPARISON OF DENGUE VIRUS-LIKE PARTICLE SAMPLES BINDING TO HEPARIN VIA A CAPTURE ELISA 220
5.6 COMPARISON OF DENGUE VIRUS-LIKE PARTICLE SAMPLES OLIGOMERIC REARRANGEMENTS IN RESPONSE TO LOW PH 227
5.7 DENGUE VIRUS-LIKE PARTICLE ELISA USING HUMAN SERUM 229
5.8 DISCUSSION 232
5.8.1 Development and optimisation of an ELISA using recombinant dengue virus-like particles as antigen *232*

5.8.2 Comparison of dengue virus-like particle samples binding to DC-SIGN via a capture ELISA 233
5.8.3 Comparison of dengue virus-like particle samples binding to heparin via a capture ELISA 234
5.8.4 Comparison of dengue virus-like particle samples oligomeric rearrangements in response to low pH 240
5.8.5 Dengue virus-like particle ELISA using human serum 240
5.9 CONCLUSIONS AND SUMMARY 242

CHAPTER 6. FINAL DISCUSSION, CONCLUSIONS AND FUTURE WORK 245

6.1 CONCLUSIONS 255

REFERENCES 256

APPENDICES ..... A

APPENDIX A. DENGUE RT-PCR PRIMER ALIGNMENTS ..... B
APPENDIX B. GEL ELECTROPHORESIS IMAGES OF PLASMID DNA FOR EACH DENGUE VIRUS QUASISPECIES ENVELOPE GENE CLONE. C
APPENDIX C. QUASISPECIES ENVELOPE GENE NUCLEOTIDE SEQUENCE ALIGNMENTS ..... D
APPENDIX D. QUASISPECIES ENVELOPE GLYCOPROTEIN AMINO ACID SEQUENCE ALIGNMENTS ..... E
APPENDIX E. SUMMARY OF DENGUE VIRUS QUASISPECIES ENVELOPE GENE MUTATIONS ..... F
APPENDIX F. ACCUSCRIPT VERSUS EXPAND METHOD COMPARISON: DENGUE VIRUS TYPE 3 CLONE ENVELOPE GENE NUCLEOTIDE SEQUENCE ALIGNMENTS ..... N
APPENDIX G. ACCUSCRIPT VERSUS EXPAND METHOD COMPARISON: DENGUE VIRUS TYPE 3 CLONE ENVELOPE GLYCOPROTEIN AMINO ACID SEQUENCE ALIGNMENTS ..... O
APPENDIX H. EXPAND VERSUS ACCUSCRIPT METHOD COMPARISON: SUMMARY OF DENGUE VIRUS TYPE 3 ENVELOPE GENE MUTATIONS ..... P
APPENDIX I. PERCENTAGE SEQUENCE IDENTITIES CALCULATED FROM DENGUE VIRUS NUCLEOTIDE SEQUENCE ALIGNMENTS FOR PHYLOGENETIC ANALYSIS ..... R
APPENDIX J. ORIGINAL IMAGES ..... U
APPENDIX K. LIST OF PRESENTATIONS V
APPENDIX L. ETHICAL CONSIDERATIONS RELATING TO THE SRI LANKAN PATIENT SAMPLES AND CLINICAL DATA ..... W

LIST OF FIGURES

CHAPTER 1: Introduction

Figure 1.1. Areas at risk of dengue virus transmission in the Western hemisphere (A) and the Eastern hemisphere (B). *4*
Figure 1.2. Flavivirus phylogenetic tree. *5*
Figure 1.3. World Health Organisation dengue case classification guidance. *8*
Figure 1.4. 2009 updated World Health Organisation dengue case classification guidance. *9*
Figure 1.5. Dengue virus genome organisation. *12*
Figure 1.6. Structure of the flavivirus virion. *13*
Figure 1.7. Flavivirus replication. *14*
Figure 1.8. Dengue virus genetic relationships and evolutionary history. *19*
Figure 1.9. Phylogenetic tree showing dengue virus type 1 evolutionary relationships. *22*
Figure 1.10. Phylogenetic tree showing dengue virus type 2 evolutionary relationships. *23*
Figure 1.11. Phylogenetic tree showing dengue virus type 3 evolutionary relationships. *24*
Figure 1.12. Phylogenetic tree showing dengue virus type 4 evolutionary relationships. *25*
Figure 1.13. Dengue haemorrhagic fever in Sri Lanka between 1980 and 1998. ..... *26*
Figure 1.14. Flavivirus envelope protein antigenic model. *31*
Figure 1.15. Tick-borne encephalitis virus envelope protein structure. *32*
Figure 1.16. Clustering of mutations at envelope protein structural domain interfaces. *34*

CHAPTER 2: Materials and Methods

Figure 2.1. Example $TCID^{50}$ calculation *65*

CHAPTER 3: Analysis of dengue virus quasispecies populations in clinical samples

Figure 3.1. Analysis of dengue virus envelope gene RT-PCR primer specificity. *83*
Figure 3.2. Comparison of dengue virus envelope gene reverse transcription and amplification using standard and high-fidelity RT-PCR. *84*
Figure 3.3. Summary of diagnoses for patients enrolled in the Sri Lankan fever study. *86*
Figure 3.4. Dengue virus type 1 patient sample high-fidelity RT-PCR product analysis. *91*
Figure 3.5. Dengue virus type 2 patient sample high-fidelity RT-PCR product analysis. *92*
Figure 3.6. Dengue virus type 3 patient sample high-fidelity RT-PCR product analysis. *93*
Figure 3.7. Dengue virus type 4 patient sample high-fidelity RT-PCR product analysis. *94*
Figure 3.8. M13 PCR validation of plasmid DNA electrophoresis method for verification of dengue virus envelope gene presence in cloned plasmid DNA. *98*
Figure 3.9. Schematic of dengue virus envelope gene sequencing primer read positions. *99*
Figure 3.10. Dengue virus type 2 complete envelope gene phylogenetic tree. *113*
Figure 3.11. Dengue virus type 4 complete envelope gene phylogenetic tree. *114*
Figure 3.12. Dengue virus type 3 complete envelope gene phylogenetic tree. *115*
Figure 3.13. Dengue virus type 3 partial (207 nucleotides at the 5' end) envelope gene phylogenetic tree. *117*
Figure 3.14. Linear schematic representation of the dengue virus envelope glycoprotein, detailing important structural features and the location of Sri Lankan patient sample quasispecies amino acid substitutions. *119*
Figure 3.15. Three-dimensional structure of the dengue virus envelope glycoprotein detailing the locations of the Sri Lankan patient sample quasispecies amino acid substitutions. *121*

CHAPTER 4: Production and purification of recombinant dengue virus envelope glycoproteins

Figure 4.1. pBAC-2cp baculovirus transfer vector circle map and MCS nucleotide sequence *152*
Figure 4.2. Schematic representation of the rBV transfer vector pBAC-2cp containing dengue virus type 1 pre-membrane and dengue patient sample quasispecies envelope genes. *154*

Figure 4.3. Analysis of colonies by PCR for confirmation of the presence of the dengue virus type 1 pre-membrane and patient sample quasispecies envelope gene inserts in pBAC-2cp *154*
Figure 4.4. Sf9 Easy-Titre cells express eGFP when infected with rBV *157*
Figure 4.5. Analysis of dengue rBVs by PCR to confirm the presence of dengue virus type 1 pre-membrane and patient sample quasispecies envelope genes *158*
Figure 4.6. Analysis of optimisation of recombinant dengue virus envelope protein production *160*
Figure 4.7. Analysis of dengue virus envelope gene mRNA transcripts produced in R097A_4 rBV-infected Sf9 cells *162*
Figure 4.8 Immunodetection of dengue virus envelope glycoprotein from clarified medium from rBV.R097A_4-infected cells after concentration by ultracentrifugation *164*
Figure 4.9. Immunodetection of dengue virus envelope glycoproteins from clarified medium from dengue rBV-infected cells after concentration by ultracentrifugation *165*
Figure 4.10. Immunodetection of dengue virus envelope glycoprotein from concentrated clarified medium from rBV.R097A_4-infected cells after purification by immunoprecipitation. *167*
Figure 4.11. Analysis of R097A_4 anion exchange column purification fractions by SDS-PAGE and immunodetection from Western blots. *169*
Figure 4.12. Immunodetection of GP64 in a baculovirus-derived Ebola VLP sample pre and post purification by immunoprecipitation. *172*
Figure 4.13. Analysis of R097A_4 anion exchange column purification fractions obtained using different sample flow-rates. *174*
Figure 4.14. Analysis of R097A_4 and lacZ sample anion exchange column purification fractions using the baculovirus GP64 mAb. *175*
Figure 4.15. Analysis of R097A_4 and lacZ sample anion exchange column purification fractions using the rabbit pAb to dengue virus types 1, 2, 3 and 4. ..... *176*
Figure 4.16. Analysis of baculovirus-derived R097A_4, lacZ and Ebola VLP samples using several different antibodies *177*
Figure 4.17. Glycosylation analysis of R097A_4 and lacZ samples using endoglycosidase H and PNGase F. *178*
Figure 4.18. Analysis of R097A_4 and lacZ samples in their native conformation. *179*
Figure 4.19. Transmission electron micrographs of flavivirus virions and VLPs stained with uranyl acetate. *181*
Figure 4.20. Amino acid sequence alignment of AcMNPV GP64 and the dengue virus envelope glycoprotein. *191*

CHAPTER 5: Analysis of dengue virus-like particles to determine the effects of the quasispecies envelope gene mutations

Figure 5.1. Dengue VLP mAb ELISA, preliminary testing. *204*
Figure 5.2. Dengue VLP pAb ELISA, preliminary testing. *205*
Figure 5.3. Comparison of dengue VLP mAb and pAb ELISAs using both primary and secondary antibodies, and secondary antibodies only. *206*
Figure 5.4. Optimisation of dengue VLP mAb ELISA blocking buffer using dengue VLP samples R097A_13 and R107A_1. ..... *208*
Figure 5.5. Optimisation of dengue VLP mAb ELISA primary and secondary antibody concentrations using dengue VLP samples R097A_13 and R107A_1. ..... *210*
Figure 5.6. Dengue VLP mAb ELISA using optimised antibody concentrations. *212*
Figure 5.7. Schematic diagram showing the dengue VLP mAb ELISA and dengue VLP capture ELISA procedures. *215*
Figure 5.8. Dengue VLP mAb ELISA comparison with DC-SIGN capture ELISA. *216*
Figure 5.9. Optimisation of DC-SIGN coating concentration for the DC-SIGN capture ELISA. *218*
Figure 5.10. Comparison of absorbance signals generated by the DC-SIGN capture ELISA using both primary and secondary antibodies, and secondary antibodies only. *219*
Figure 5.11. Initial dengue VLP mAb ELISA comparison with heparin capture ELISA. *221*
Figure 5.12. Optimisation of heparin coating concentration for the heparin capture ELISA. *222*

Figure 5.13. Dengue VLP mAb ELISA comparison with heparin capture ELISA. 224
Figure 5.14. Expected dengue virus envelope glycoprotein oligomeric state after low pH treatment and chemical cross-linking. 227
Figure 5.15. Dengue VLP ELISA using human serum. 230

APPENDICES

Figure 1. Dengue virus envelope gene clone plasmid DNA gel electrophoresis images..... c
Figure 2. Percentage sequence identities calculated from dengue virus type 2 nucleotide sequence alignments for phylogenetic analysis..... r
Figure 3. Percentage sequence identities calculated from dengue virus type 3 nucleotide sequence alignments for phylogenetic analysis..... s
Figure 4. Percentage sequence identities calculated from dengue virus type 4 nucleotide sequence alignments for phylogenetic analysis..... t

LIST OF TABLES

CHAPTER 2: Materials and Methods

Table 2.1. Standard RT-PCR thermal cycling conditions 47
Table 2.2. AccuScript thermal cycling conditions for PCR amplification of cDNA 48
Table2.3. Expand thermal cycling conditions for PCR amplification of cDNA 49
Table 2.4. Taq PCR core kit thermal cycling conditions 49
Table 2.5. Sequencing PCR thermal cycling conditions 58
Table 2.6. Antibodies and dilutions used in immunodetection from Western blots 62
Table 2.7. Antibodies and dilutions used in ELISAs 72

CHAPTER 3: Analysis of dengue virus quasispecies populations in clinical samples

Table 3.1. Primers tested for RT-PCR amplification of dengue virus envelope genes for quasispecies analysis. 81
Table 3.2. RT-PCR primers used for amplification of dengue virus envelope genes from patient samples for quasispecies analysis. 82
Table 3.3. Classification of disease severity for dengue virus RT-PCR positive patient samples. 88
Table 3.4. Dengue virus envelope genes amplified from patient samples by high-fidelity RT-PCR for cloning. 94
Table 3.5. Presence of dengue virus envelope gene inserts in picked colonies 98
Table 3.6. Sequencing primers designed for dengue virus envelope gene quasispecies analysis. 99
Table 3.7. Summary of dengue virus envelope gene sequence diversity observed in clinical samples. 101
Table 3.8. Clones containing truncated envelope gene sequences. 102
Table 3.9. Summary of dengue virus envelope gene mutation types. 104
Table 3.10. Summary of mutations generated by the Expand and AccuScript RT-PCR methods. 106
Table 3.11. Summary of mutation types generated by the Expand and AccuScript RT-PCR methods. 106
Table 3.12. Comparison of insertion and deletion mutations generated at homopolymeric runs of nucleotides by the Expand and AccuScript RT-PCR methods. 107
Table 3.13. Comparison of mutations generated by the Expand and AccuScript RT-PCR methods. 108
Table 3.14. Published dengue virus envelope gene sequences used for phylogenetic analysis. 110
Table 3.15. Summary of dengue virus envelope glycoprotein amino acid substitution locations identified through linear and three-dimensional modelling. 124
Table 3.16. Summary of quasispecies variants chosen for recombinant protein production. 125
Table 3.17. Predicted prokaryote promoter sites within dengue virus envelope genes from patient samples. 133

CHAPTER 4: Production and purification of recombinant dengue virus envelope glycoproteins

Table 4.1. Primers used for the amplification of the dengue virus type 1 pre-membrane and dengue virus quasispecies envelope genes for generation of rBVs. 153
Table 4.2. rBVs generated for this study 157
Table 4.3. Dengue rBV titres. 158
Table 4.4. Antibodies tested for dengue virus envelope glycoprotein immunodetection 170

CHAPTER 5: Analysis of dengue virus-like particles to determine the effects of the quasispecies envelope gene mutations

Table 5.1. Signal to noise ratios for the dengue VLP samples in the mAb and pAb ELISAs. 207
Table 5.2. Signal to noise ratios for the dengue VLP samples using different blocking buffers in the mAb ELISA. 209
Table 5.3. Signal to noise ratios for the dengue VLP samples using different concentrations of primary and secondary antibody in the mAb ELISA. 211
Table 5.4. Signal to noise ratios for the dengue VLP samples using the optimised mAb ELISA. 213

Table 5.5. Signal to noise ratios for the dengue VLP samples using the dengue VLP mAb ELISA and the DC-SIGN capture ELISA. *217*
Table 5.6. Signal to noise ratios for the dengue VLP sample R107A_1 using different DC-SIGN coating concentrations for the DC-SIGN capture ELISA. *218*
Table 5.7. Signal to noise ratios for the dengue VLP samples using the dengue VLP mAb ELISA and the heparin capture ELISA. *222*
Table 5.8. Signal to noise ratios for the dengue VLP sample R107A_1 using different heparin coating concentrations for the heparin capture ELISA. *223*
Table 5.9. Proportion of available dengue VLPs bound to heparin after normalisation of dengue VLP mAb and heparin capture ELISA data *226*
Table 5.10. Signal to noise ratios for the dengue VLP samples using human serum as primary antibody in the dengue VLP ELISA. *231*

APPENDICES

Table 1. R097A patient sample quasispecies mutations *g*
Table 2. R107A patient sample quasispecies mutations *i*
Table 3. R203A patient sample quasispecies mutations *j*
Table 4. R259A patient sample quasispecies mutations *k*
Table 5. R261A patient sample quasispecies mutations *l*
Table 6. R282A patient sample quasispecies mutations *m*
Table 7. Expand method mutations *p*
Table 8. AccuScript method mutations *q*

NOTATION

<	less than
>	greater than
/	per
µ	micro
°C	degrees celsius
+ve	positive
3D	3-dimensional
Å	ångstrom
A	adenine
ADE	antibody-dependent enhancement
AFRIMS	Armed Forces Research Institute of Medical Sciences
Ala or A	alanine
ALT	alanine transaminase
Arg or R	arginine
Asn or N	asparagine
Asp or D	aspartic acid
AST	aspartate transaminase
ASU	asymmetric unit
BDGP	Berkeley Drosophila Genome Project
BLAST	basic local alignment search tool (NCBI)
bp	base pairs
BSA	bovine serum albumin
C	capsid protein
C	cytosine
CDC	Centers for Disease Control and Prevention
cDNA	complementary DNA
CEPR	Centre for Emergency Preparedness and Response
CNBr	cyanogen bromide
CNS	central nervous system
cys or C	cysteine
D	denatured
dATP	deoxyribose adenosine triphosphate
DC	dendritic cell
DC-SIGN	dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin
DC-SIGNR	DC-SIGN relative
dCTP	deoxyribose cytosine triphosphate
DEN	dengue
DF	dengue fever
dGTP	deoxyribose guanine triphosphate
DHF	dengue haemorrhagic fever
DMP	dimethyl pimelimidate
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
Dr.	doctor
DTT	dithiothreitol
dUTP	deoxyribose uracil triphosphate
DSS	dengue shock syndrome
DV	dengue virus
E	envelope glycoprotein
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetic acid
eGFP	enhanced green fluorescent protein
ELISA	enzyme-linked immunosorbent assay
ER	endoplasmic reticulum
F	forward
FBS	fetal bovine serum
FCS	fetal calf serum
FMDV	foot and mouth disease virus
g	grams
G	guanine
Gln or Q	glutamine
Glu or E	glutamic acid
Gly or G	glycine
HCl	hydrochloride
HCT	haematocrit
Hg	mercury
HI	haemagglutination
His or H	histidine
HIV	human immunodeficiency virus
HPA	Health Protection Agency
HRP	horseradish peroxidase
hrs	hours
HSHS	highly sulphated heparin sulphate
ICAM-3	intracellular adhesion molecule-3
IF	immunofluorescence
Ig	immunoglobulin
Ile or I	isoleucine
IMAC	immobilised metal affinity chromatography
Inc.	incorporated
IPTG	isopropyl-ß-D-thiogalactopyranoside
IRF	immunoreactive fragment
JEV	Japanese encephalitis virus
kb	kilobases
KCl	potassium chloride
kDa	kiloDaltons
L	litre
LB	Luria bertani
Leu or L	leucine
Ltd.	limited
Lys or K	lysine
m	metres
m	milli
M	molar
M	membrane glycoprotein
mAb	monoclonal antibody
M_n	neat media
M_c	concentrated media
MCS	multiple cloning site
Met or M	methionine
MeOH	methanol
MgCl	magnesium chloride
MMLV	Moloney Murine Leukemia Virus
MOI	multiplicity of infection
mRNA	messenger RNA

MVEV	Murray Valley encephalitis virus
MW	molecular weight
n	nano
N	no
N	native
NaCl	sodium chloride
NC	non-coding
NCBI	National Centre for Biotechnology Information
NEB	New England Biolabs
NEG	negative
NIBSC	National Institute for Biological Standards and Control
nm	nanometres
NTC	no template control
NS	non-structural
nt	nucleotide
OET	Oxford Expression Technologies
ORF	open reading frame
p	passage
pAb	polyclonal antibody
PBS	phosphate buffered saline
PD	proportionate distance
pdb	protein data bank
Pfu	*Pyrococcus furiosus*
Pfu	plaque-forming units
Phe or F	phenylalanine
pmol	picomoles
polh	polyhedrin
poly(A)	poly adenylated
POS	positive
pr	pre-membrane peptide
prM	pre-membrane glycoprotein
Pro or P	proline
Prof.	professor
PVDF	polyvinylidene fluoride
R	reverse
RAMC	Royal Army Medical Corps
rBV	recombinant baculovirus
RefSeq	reference sequence (NCBI)
RNA	ribonucleic acid
rpm	revolutions per minute
RT	reverse transcription
RT	reverse transcriptase
PCR	polymerase chain reaction
S	soluble fraction
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser or S	serine
seq	sequencing
Sf	Spodoptera frugiperda
*Sf*9 E-T	*Sf*9 Easy-Titre
SLEV	St. Louis encephalitis virus
STE	Sodium chloride-Tris-EDTA
T	total fraction
T	thymine
T_m	melting temperature
Taq	*Thermus aquaticus*
TBE	tris-borate EDTA
TBEV	tick-borne encephalitis virus
TBS	tris buffered saline
$TCID^{50}$	50% tissue culture infectious dose
TEM	transmission electron microscopy
TGN	trans golgi network
Thr of T	threonine
™	trademark
Tris	tris(hydroxymethyl)-aminomethane
Trp or W	tryptophan
TRX	thioredoxin
Tyr or Y	tyrosine
U	uracil
UK	United Kingdom
USA	United States of America
UTR	untranslated region
UV	ultra violet
V	volts
Val or V	valine
VLP	virus-like particle
vRNA	viral RNA
v/v	volume per volume
WHO	World Health Organisation
WNV	West Nile virus
w/v	weight per volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
YFV	yellow fever virus

CHAPTER 1. INTRODUCTION

1.1 Thesis overview

This thesis describes the investigation of dengue virus (DV) quasispecies populations, specifically examining the envelope glycoprotein (E) gene, within individual patient samples from an outbreak in Sri Lanka in 2006. Chapter 1 provides an introduction to dengue disease and the epidemiology and biology of the aetiological agents, the DVs. The diversity and variability of DVs are examined at the species, genotype and quasispecies level. The effect of this variation on virus virulence and disease pathogenicity is also discussed, with particular attention to the effect of mutations in the E gene. The hypothesis, aims and objectives of the project are stated at the end of Chapter 1. The methods used throughout the course of the investigation are detailed in Chapter 2, and the amplification, cloning and sequencing of the quasispecies E genes from the patient samples is described in Chapter 3. The extent of variation observed in the virus populations within the individual patients, and how the results from this study compare to previous studies is assessed. The effect of the quasispecies mutations on E function is hypothesised based on the proximity of the mutations to important structural features within the E. Chapter 4 describes the construction of recombinant E for several of the quasispecies variants in the form of virus-like particles (VLPs). In Chapter 5, the host cell attachment functions of these recombinant VLPs are assessed, as is their utility as a source of antigen for the serodiagnosis of dengue. A summary and discussion of the results of the investigation, and opportunities for future work are provided in Chapter 6, which details correlations between quasispecies variation, virus virulence and disease severity.

1.2 Dengue

Dengue is the most common mosquito-borne viral disease, with 50 million human cases estimated to occur annually (World Health Organisation, 2009). Dengue is of major global concern as 40% of the world population (2.5 billion people) are at risk of contracting the disease, and the morbidity and mortality associated with dengue constitute an enormous public health burden. There are thought to be around 500,000 cases of severe dengue requiring hospitalisation every year with at least a 2.5% fatality rate (World Health Organisation, 1997b).

1.2.1 Aetiology of dengue

Dengue is caused by four distinct but closely related viruses; dengue virus (DV) types 1, 2, 3 and 4 which are transmitted by female *Aedes* mosquitoes, predominantly *Ae. aegyptii*. The occurrence of dengue has increased 30-fold in the last 50 years due to the failure of vector eradication programmes and the geographical spread of vector and virus populations (World Health Organisation, 2009). Epidemics of severe disease were not recorded until the 1950s and since then have spread geographically from South East Asia to every continent except Europe and Antarctica (World Health Organisation, 1997b). The four DVs are now endemic in more than 100 countries across Africa, the Americas, the Eastern Mediterranean, South East Asia and the Western Pacific (Figure 1.1), with more areas becoming hyperendemic for multiple DV types. Severe disease epidemics that include a higher proportion of severe dengue cases and a greater number of fatalities are becoming more common (World Health Organisation, 2008).

DVs belong to the genus flavivirus within the family *Flaviviridae*. Within the flavivirus genus, the viruses can be classified in terms of their transmission vector as either mosquito-borne, tick-borne or vector unknown (Figure 1.2). Other flaviviruses of global concern include yellow fever virus (YFV), Japanese encephalitis virus (JEV), West Nile virus (WNV) and tick-borne encephalitis virus (TBEV).

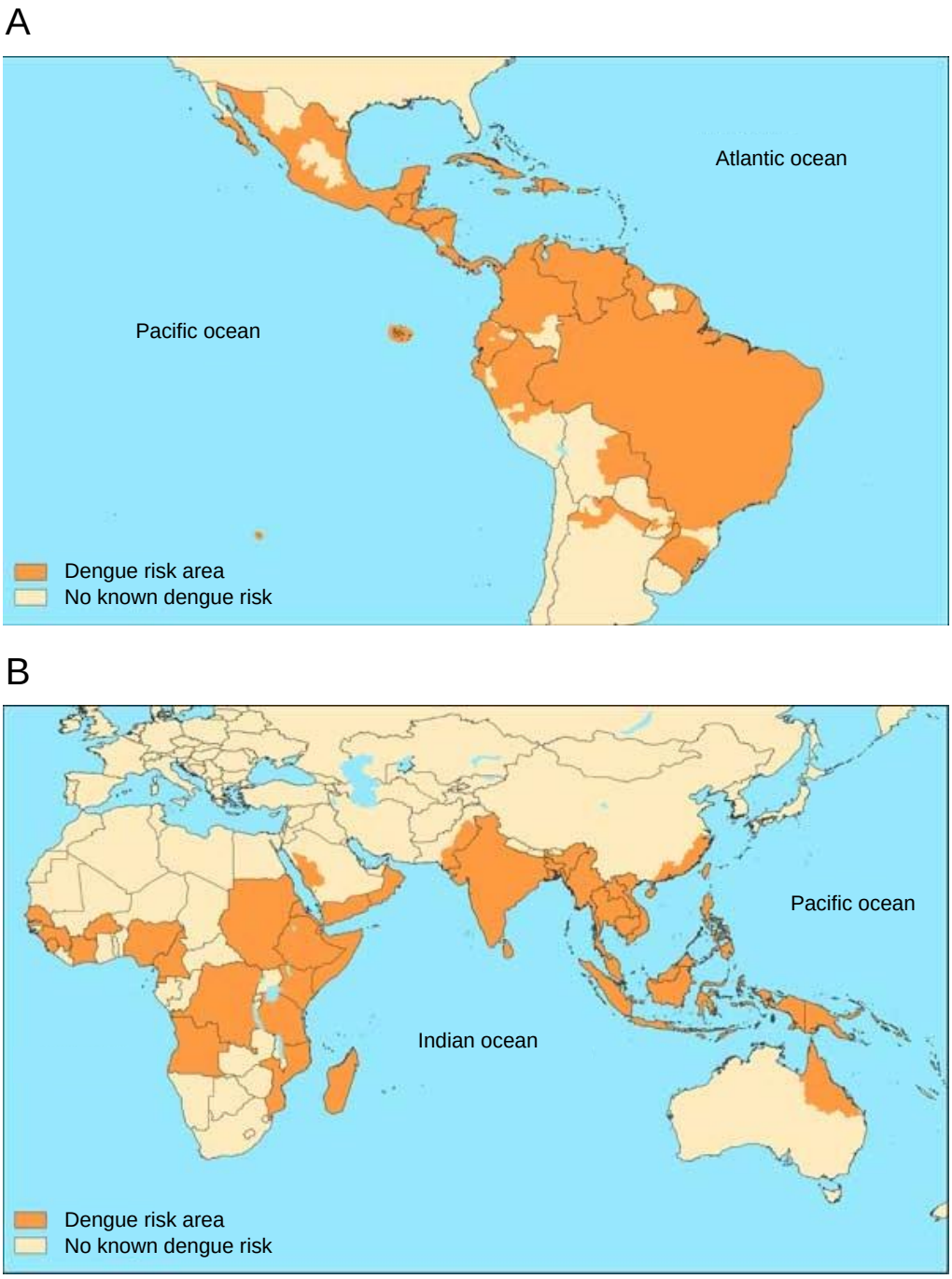


Figure 1.1. Areas at risk of dengue virus transmission in the Western hemisphere (A) and the Eastern hemisphere (B).

The maps show areas at risk of DV transmission based on dengue fever cases reported in 2008 (Centers for Disease Control and Prevention, 2009).

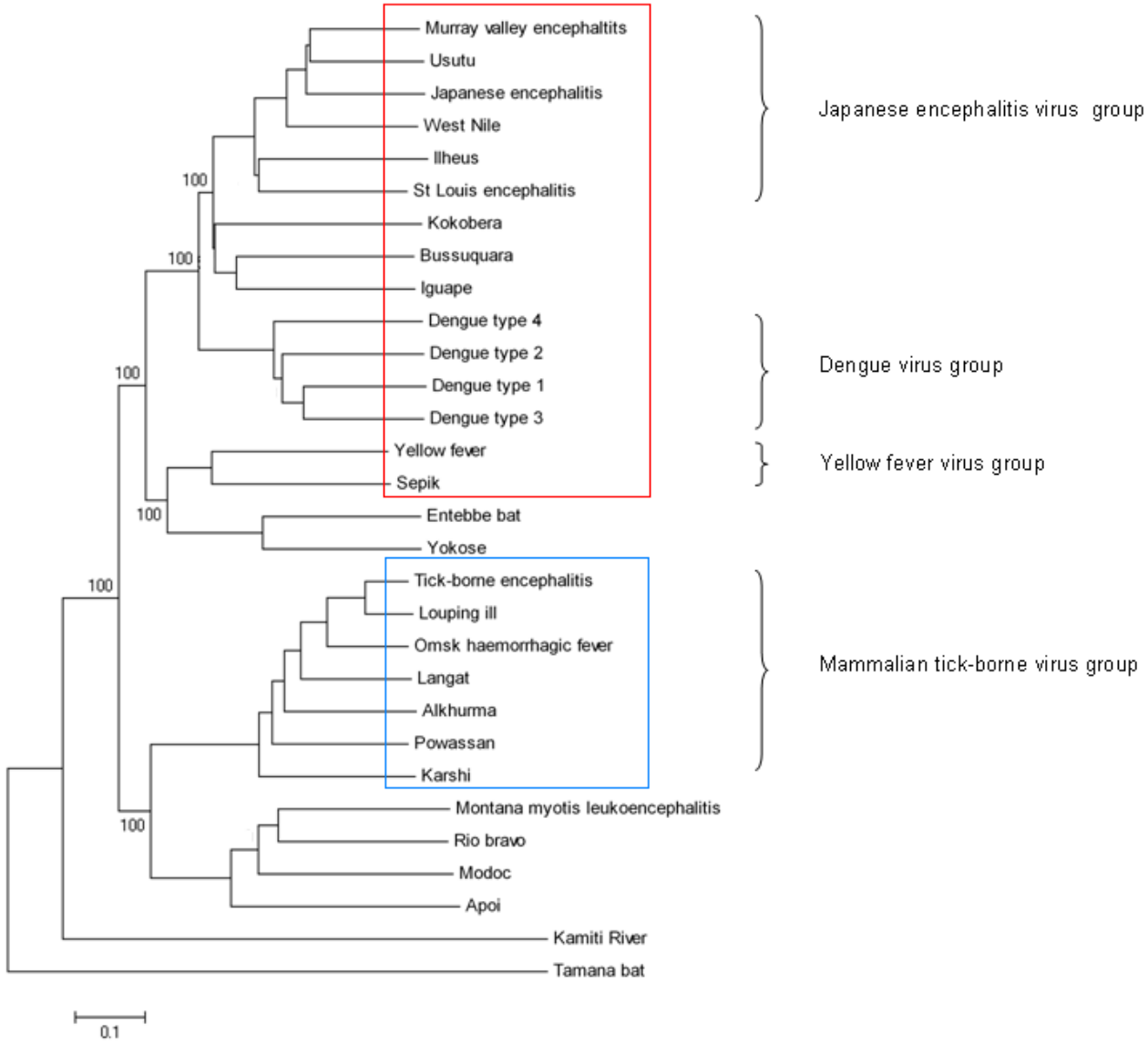


Figure 1.2. Flavivirus phylogenetic tree.

This phylogenetic tree was constructed for this study. The red box encloses mosquito-borne flaviviruses and the blue box encloses mammalian tick-borne flaviviruses. The evolutionary history was inferred from complete genome sequences using the Neighbour-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. Phylogenetic analyses were conducted in MEGA4. Sequences were NCBI RefSeq numbers: NC002031 (Yellow fever virus, YFV), NC001437 (Japanese encephalitis virus, JEV), NC001672 (tick-borne encephalitis virus, TBEV), NC001477 (DV type 1), NC001474 (DV type 2), NC001475 (DV type 3), NC002640 (DV type 4), NC000943 (Murray Valley encephalitis virus, MVEV), NC007580 (St Louis encephalitis virus, SLEV), NC009942 (West Nile virus, WNV), NC006551 (Usutu virus), NC005064 (Kamiti river virus), NC005062 (Omsk haemorrhagic fever virus), NC004355 (Alkhurma virus), NC003687 (Powassan virus), NC003675 (Rio Bravo virus), NC003676 (Apoi virus), NC001809 (Louping ill virus), NC008718 (Entebbe bat virus), NC006947 (Karshi virus), NC004119 (Montana myotis leukoencephalitis virus), NC003670 (Langat virus), NC003635 (Modoc virus), NC009027 (Iguape virus), NC009026 (Bussuquara virus), NC009028 (lheus virus), NC009029 (Kokobera virus), NC008719 (Sepik virus) and NC005039 (Yokose virus).

1.2.2 Symptoms of dengue

In infants and young children, dengue most commonly presents as a non-specific febrile illness with a rash. In older children and adults there can be a mild febrile syndrome or classical disease with pain behind the eyes and a rash. Severe disease occurs in about 1% of cases and symptoms can include high fever, haemorrhagic phenomena, enlarged liver, circulatory failure and hypovolemic shock. Death can occur within 12-24 hours unless appropriate treatment is given (World Health Organisation, 2008). There are no specific antiviral therapies available for dengue and there is no vaccine. Treatment is intensive supportive care such as maintenance of circulating fluid volume and ventilatory support, but this is less widely available in developing countries.

Most flavivirus infections are thought to be under-reported because the wide spectrum of disease severity and a lack of effective case definition make diagnosis difficult. Symptoms common to all flavivirus infections include: fever, headache, chills, muscle aches, loss of appetite, nausea, and vomiting. The most common presentation is of a mild febrile illness and flavivirus infections can sometimes even be asymptomatic. However, severe complications occur in a significant proportion of infected individuals. For infections caused by YFV, JEV, WNV and TBEV these severe complications tend to be neurological rather than the abnormalities of haemostasis and vascular permeability observed with severe dengue.

1.2.3 Clinical and laboratory diagnosis of dengue

During diagnosis, flavivirus infections must be differentiated from other viral infections with similar clinical presentation (for example Chikungunya virus, an alphavirus), bacterial infections (such as *Salmonella typhi*, which causes typhoid fever or *Leptospira* species which cause leptospirosis), or parasitic infections (for example the *Plasmodium* parasites that cause malaria). The diagnosis may then be narrowed to a specific flavivirus infection, although according to the World Health Organisation (WHO) many laboratories only report a diagnosis of ‘acute flavivirus infection’, rather than identifying the precise virus using serological or molecular techniques (World Health Organisation, 1997c).

Cases of symptomatic dengue have historically been classified by severity according to WHO guidelines first published in 1975, which differentiate between cases of dengue fever (DF), dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS; World Health Organisation 1997a; Figure 1.3). However, dengue epidemiology has changed considerably since these guidelines were first published, leading to difficulties with the use of this classification system in a clinical setting (reviewed in Bandyopadhyay, Lum, and Kroeger, 2006; Deen *et al*., 2006). Examples of severe dengue that do not meet the WHO criteria for DHF include: dengue with haemorrhage but without evidence of plasma leakage; dengue with shock but not fulfilling all four DHF criteria; and severe dengue with organ dysfunction but a low degree of plasma leakage (Bandyopadhyay, Lum, and Kroeger, 2006).

Regional WHO publications since 1997 have acknowledged these difficulties by relaxing the criteria in some circumstances to aid clinical assessment of dengue cases. For example, in small hospitals in India, guidelines state that DHF is any probable case of dengue with haemorrhaging and any one (not all) of the DHF indications shown in Figure 1.3 (World Health Organisation, 1999). Another example is in the treatment of children, where the term 'severe dengue' has been proposed to include any case of dengue fever with signs of vascular permeability resulting in plasma leakage (Deen, 2005). In the most recent WHO guidelines (World Health Organisation, 2009), a classification system has been proposed that distinguishes between patients with severe dengue, and those with non-severe dengue with or without warning signs (Figure 1.4). This classification system has been tested in 18 countries by comparison with the existing WHO case classification and published results are expected soon.

Dengue fever

Acute febrile illness with two or more of the following:

- headache
- retro-orbital pain
- myalgia
- arthralgia
- rash
- haemorrhagic manifestations
- leukopenia

and either supportive serology or proximity to laboratory confirmed cases.

Dengue haemorrhagic fever

The following indications must all be present:

- fever or a history of fever
- haemorrhagic tendencies evidenced by one or more of the following:
  - a positive tourniquet test
  - petechiae, ecchymoses or purpura
  - bleeding from the mucosa, gastrointestinal tract, injection sites or other locations
  - haematemesis or melaena
- thrombocytopenia (100,000 cells per mm3 or less)
- evidence of plasma leakage due to increased vascular permeability, manifested by at least one of the following:
  - a rise in the haematocrit equal to or greater than 20% above average for age, sex and population
  - a drop in the haematocrit following volume-replacement treatment equal to or greater than 20% of baseline
  - signs of plasma leakage such as pleural effusion, ascites and hypoproteinaemia

Dengue shock syndrome

All four of the above DHF criteria must be present, plus evidence of circulatory failure manifested by either of the following:

- rapid, weak pulse and narrow pulse pressure (<20mmHg)
- hypotension for age (systolic pressure <80mmHg for those less than 5 years of age, or <90mmHg for those greater than or equal to 5 years of age) and cold, clammy skin and restlessness

Figure 1.3. World Health Organisation dengue case classification guidance.

Adapted from World Health Organisation, 1997a. Cases of dengue are classified as dengue fever (DF), dengue haemorrhagic fever (DHF) or dengue shock syndrome (DSS) according to the clinical observations described in the figure.

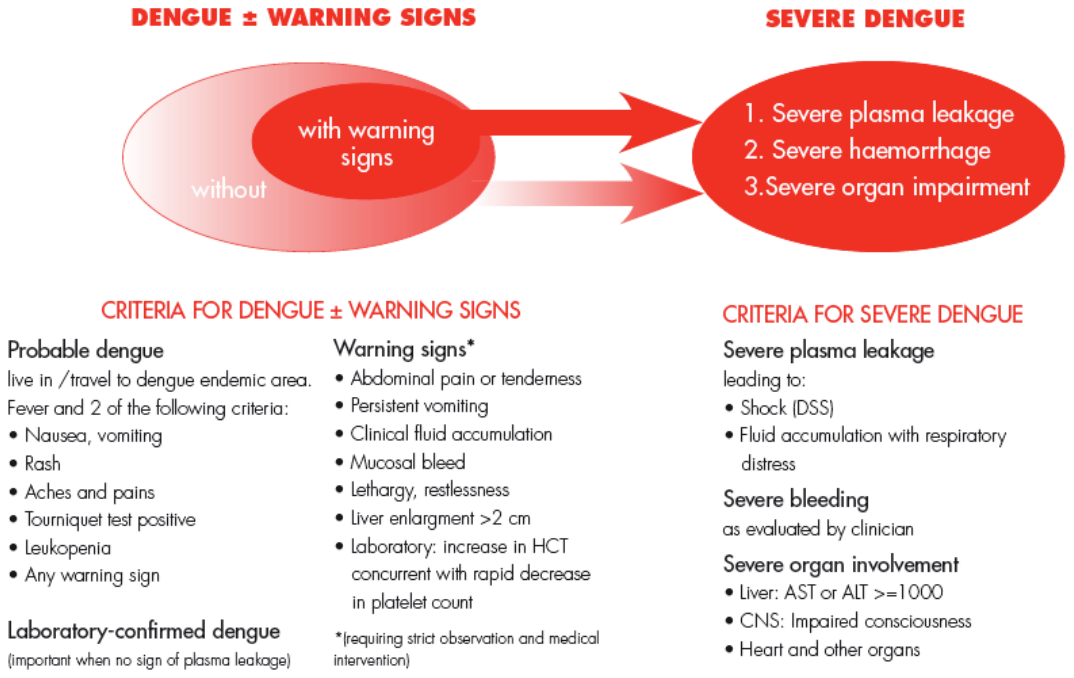


Figure 1.4. 2009 updated World Health Organisation dengue case classification guidance.

Figure is from World Health Organisation, 2009. Cases of dengue are classified as dengue with or without warning signs, and severe dengue according to the clinical observations described in the figure. Abbreviations are: haematocrit (HCT), dengue shock syndrome (DSS), aspartate transaminase (AST), alanine transaminase (ALT), central nervous system (CNS).

Diagnostic laboratory tests most commonly used for dengue include those that detect the DVs (isolation by tissue culture or reverse transcription-polymerase chain reaction, RT-PCR) and those that detect antibodies to the virus (enzyme-linked immunosorbent assay (ELISA), immunofluorescence (IF), haemagglutination inhibition (HI) and neutralisation tests). The gold standard for the laboratory diagnosis of any flavivirus infection is isolation, culture and further characterisation of the virus (for example by antigen detection) from the patient sample which is time consuming, taking over a week to complete. However, RT-PCR allows rapid diagnosis and could therefore improve patient care. Both of these virus detection methodologies are more expensive than serology-based methods and require well-equipped laboratories. They are also only of use within approximately seven days of the onset of symptoms as the virus is present in the sera typically for the duration of fever (World Health Organisation, 1997c). After this time, antibodies to the virus can be identified using serological methods. The antibodies neutralise the virus and therefore it is not

possible to detect or culture the virus once the immune response is significantly underway.

There are several difficulties with serological diagnosis of flavivirus infections. Many different flaviviruses circulate in the same regions and there is a large degree of antigenic cross-reactivity between them, so it can be difficult to identify which flavivirus is causing the current infection. In addition, the presence of antibodies from previous flavivirus infections or vaccinations makes it difficult to diagnose the type of infection (primary or secondary; past, present or recent; Shu and Huang, 2004). Primary antibody responses occur in people who have not previously been infected (or immunised) and are characterised by the presence of immunoglobulin M (IgM) antibodies after several days of fever, or after the fever has passed. Secondary responses occur as a result of a second flavivirus infection (for example sequential infection with different DV types) and it is the IgG response that is dominant (Shu and Huang, 2004). IgG antibodies from secondary dengue infections can persist in the blood for more than 10 months (Gubler, 1996). Serological studies on dengue patients have shown that antibodies to the envelope glycoprotein (E) are more commonly detected than those raised against the capsid (C), pre-membrane (prM), membrane (M), nonstructural (NS) 1, NS3 or NS5 proteins in both primary and secondary infections (Valdes *et al*., 2000). This shows that serology-based diagnostic tests should be directed towards E antigens or antibodies for them to stand the best chance of early detection.

In the UK, travellers falling ill on returning home from flavivirus endemic areas require laboratory diagnosis of their diseases. Currently at the Health Protection Agency (HPA) Centre for Emergency Preparedness and Response (CEPR) approximately 2500 clinical samples are tested for evidence of arboviruses each year, and this includes a flavivirus screen. Samples are received from patients in the early stages of infection (in which antigen is present) and later after the immune response has occurred. Screening for arbovirus infections is done by RT-PCR, IF, and IgM and IgG ELISA. Cross-reactivity between antibodies to the different flaviviruses often complicates interpretation of the

serology results. There is a need for improved reagents for serodiagnosis (flavivirus antibodies and antigens) that are species specific, and do not cross-react.

1.3 Dengue viruses

1.3.1 Dengue virus structure and genome organisation

DVs, like all flaviviruses, are single positive strand ribonucleic acid (RNA)-encoded viruses. DV genomes are around 10,700 nucleotides in length, and act directly as messenger RNA (mRNA) for all viral proteins, which are synthesised in one long open reading frame as a polyprotein which is then proteolytically processed (Figure 1.5). The genome consists of non-coding (NC) regions at the 5' and 3' ends, with the genes encoding the three structural proteins (C, prM and E) followed by those encoding the seven NS proteins (NS1, NS2a, NS2b, NS3, NS4a, NS4b and NS5; Lindenbach and Rice, 2001). The non-structural proteins are involved in replication and processing of the viral polyprotein (Ball, 2001), and inhibition of host interferon responses to infection (Ho *et al*., 2005; Munoz-Jordan *et al*., 2005). Mature virions are spherical and around 50nm in diameter, containing the three structural proteins (C, M and E). M and E are embedded in a host-derived lipid bilayer forming the outer surface of the virion, surrounding the C proteins that encapsulate the viral RNA (Figure 1.6; Kaufmann and Rossmann, 2011).

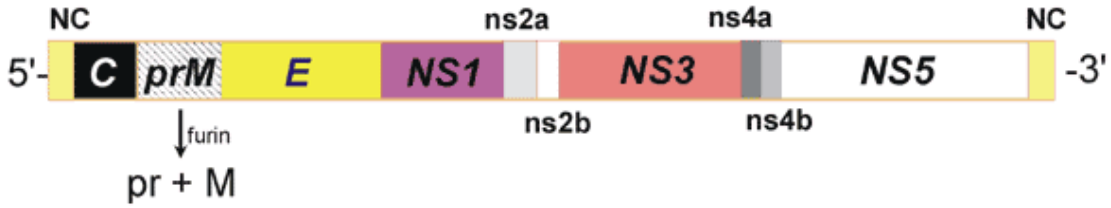


Figure 1.5. Dengue virus genome organisation.

Figure is from Petersen and Roehrig, 2001. The approximately 10.7kb genome has non-coding (NC) regions at the 5' and 3' ends, with the protein-coding genes in the order C, prM, E, NS1, NS2a, NS2b, NS3, NS4a, NS4b and NS5. PrM is the membrane protein precursor which is cleaved by furin to produce the M protein.

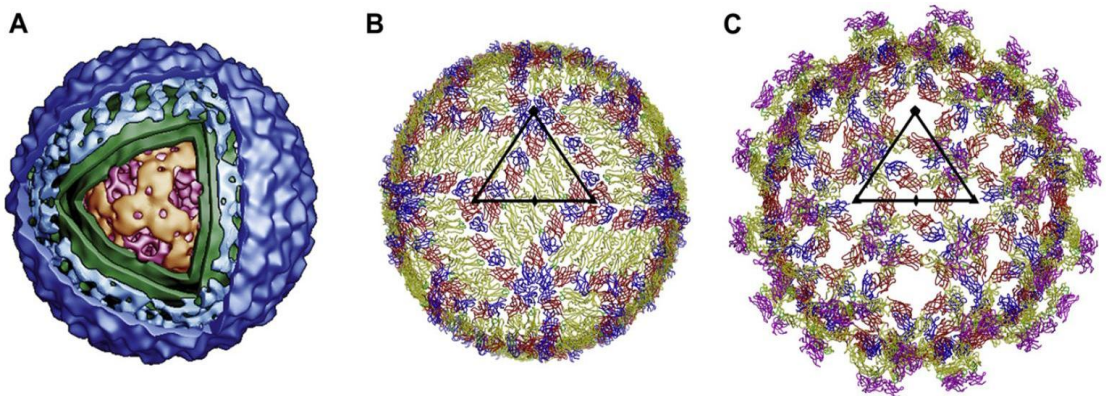


Figure 1.6. Structure of the flavivirus virion.

Figure and legend are from Kaufmann and Rossmann, 2011. (A) Three-dimensional (3D) rendering of a cryoelectron microscopy density map of DV. The cut-out window allows the view inside the virion at the nucleocapsid core (magenta/orange) composed of the C protein and the RNA genome. The core is surrounded by a host-derived lipid envelope (green). The outer icosahedral shell of the virus (blue) is formed by two membrane-anchored glycoproteins, E and M. (B) Pre-fusion arrangement of E on the surface of the mature flavivirion, viewed down an icosahedral twofold axis. Domains *I*, *II*, and *III* of each E monomer are coloured red, yellow, and blue, respectively. The fusion loop is shown in green. A total of 180 E monomers are associated in 30 rafts of three, nearly parallel E homodimers that form a distinct herringbone pattern. One icosahedral asymmetric unit (ASU), the smallest unit from which the particle can be generated by applying icosahedral symmetry, is outlined by a black triangle. The positions of the neighbouring icosahedral five-, three- and two-fold symmetry axes are marked with symbols (pentagons, triangles, and ovals, respectively). Each ASU contains the equivalent of three E monomers. (C) Arrangement of E and pr peptides on the surface of an immature, fusion-incompetent flavivirion, viewed down an icosahedral twofold axis. Domains *I*, *II*, and *III* of each E monomer and the fusion loop are coloured as in (A). The pr peptide is shown in magenta. One ASU is outlined by a black triangle and the positions of the symmetry axes are indicated. E are icosahedrally arranged in 60 trimeric spikes, composed of prM/E heterodimers. The pr peptide of prM covers the fusion peptide in domain II of E (on the tip of each spike) to prevent premature fusion.

1.3.2 Dengue virus replication

The host range of DVs includes humans, monkeys and mosquitoes, and virus replication is able to occur in diverse tissues within all of these organisms. During infection (Figure 1.7), the virus attaches to host cell surface receptors via E (Chen, Maguire, and Marks, 1996) and enters the cell by receptor-mediated endocytosis. The acidic pH within the endosome triggers conformational changes in the viral envelope that allow fusion of viral and cell membranes. The viral RNA is released into the cytoplasm, and once delivered to cellular ribosomes, is translated to form the polyprotein precursor. This undergoes proteolytic processing to release the structural and non-structural proteins. The non-structural proteins catalyse replication via the production of anti-genomic RNA, which is itself replicated to produce more genomic RNA for translation. Virus particles are assembled on the surface of the endoplasmic

reticulum (ER), in the form of immature virions containing prM. These bud off from the ER and are transported through the *trans*-Golgi network (TGN). The prM are then cleaved by the host cell's furin protease, to yield M. This exposes the E fusion peptide, which results in mature infectious virions that are fusion-competent and can be released to infect more host cells (Ball, 2001; Mukhopadhyay, Kuhn, and Rossmann, 2005).

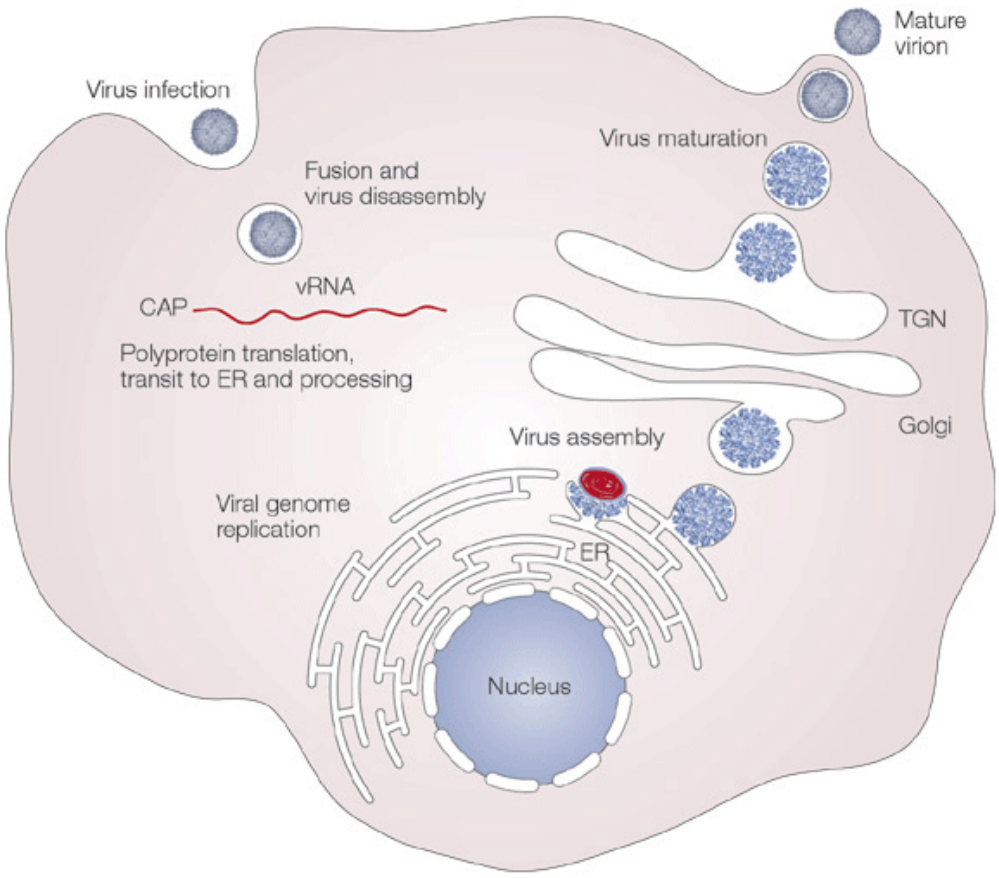


Figure 1.7. Flavivirus replication.

The virus attaches to its target cell and gains access via receptor-mediated endocytosis. Viral and cell membranes fuse, releasing the viral RNA (vRNA) genome into the cell. The genomic RNA is translated into a polyprotein which is then cleaved to release the structural and non-structural proteins. Replication of the viral genome and new virus assembly takes place at the surface of the ER. Immature virions are transported out of the cell via the TGN, and virus maturation occurs. Diagram is from Ball, 2001.

NC regions at the ends of the genome play an important role during replication and translation. Long-range RNA-RNA interactions between the 5' and 3' non-coding regions are thought to circularise the genome, which prevents replication and translation of truncated or damaged mRNAs (Alvarez *et al.*, 2005). Both non-coding

regions are capable of recruiting translation-initiation factors. The 5' end of the viral RNA is capped but the 3' end is not polyadenylated (poly(A)) unlike host cellular mRNAs. However, the 3' non-coding region plays roles in both regulation of translation and mRNA stability similar to that of the cellular mRNA poly(A) tail (Chiu, Kinney, and Dreher, 2005). DVs (and therefore potentially other flaviviruses) are able to initiate viral protein synthesis using interacting 5' and 3' non-coding regions even under conditions where cap-dependent translation is inhibited (Edgil, Polacek, and Harris, 2006). The start codon for translation initiation in dengue and other mosquito-borne flaviviruses is an adenine-uracil-guanine (AUG) sequence at the start of the coding region for the capsid protein. A hairpin loop in the RNA secondary structure near this site is thought to increase site recognition by the ribosome complex by stalling it over the start codon (Clyde and Harris, 2006).

1.3.3 Dengue virus binding to target host cells

Immature dendritic cells (DCs) in the skin are considered the initial site of DV replication in humans after inoculation by the mosquito vector (Ho *et al*., 2001; Palucka, 2000; Wu *et al*., 2000). Migration of infected DCs to the lymph nodes permits antigen presentation leading to an immune response, infection and replication in lymphoid tissues and release of DV into the circulation. Monocytes, macrophages, B and T lymphocytes and endothelial cells have also been shown to support DV infection *in vitro* (Bielefeldt-Ohmann *et al*., 2001; Halstead and O'Rourke, 1977; Lee *et al*., 2006; Miller *et al*., 2008) and different cellular receptors have been implicated in virus-cell attachment. These receptors vary between cell-types, species of origin and DV type. Monocytes and macrophages can be infected by DV in an antibody-dependent manner via binding of non-neutralised virus-antibody complexes to immunoglobulin fcγ receptors on the cell surface (Halstead and O'Rourke, 1977). However, primary infections, and infections of cells that do not have immunoglobulin fcγ receptors on the cell surface must occur via a different mechanism.

Heparan sulphate expressed on the surface of target cells has been shown to bind the DV E (Chen *et al*., 1997). Interactions between heparan sulphate and clusters of basic amino acids within E are thought to concentrate the virus on the cell surface, enabling

subsequent higher-affinity interactions with specific receptors. Lectins such as DC-specific intracellular adhesion molecule-3 (ICAM-3) grabbing nonintegrin (DC-SIGN), DC-SIGN relatives (DC-SIGNR) and the mannose receptor have been proposed as dengue receptors on dendritic cells, endothelial cells and macrophages (Miller *et al*., 2008; Navarro-Sanchez *et al*., 2003; Tassaneetrithep *et al*., 2003). This is consistent with a role for the E glycans in infectivity and virus growth in mammalian cells (Bryant *et al*., 2007; Mondotte *et al*., 2007) as lectins such as DC-SIGN bind high-mannose glycans in the presence of calcium ions (Feinberg *et al*., 2001). Cryoelectron microscopy of the DV/DC-SIGN complex has shown interaction of the lectin with the asparagine (N or Asn)-linked glycan at Asn-67 (Pokidysheva *et al*., 2006).

One mechanism of virus neutralisation by antibodies is prevention of virus attachment to host cells, as the antibody obscures the receptor-binding site on the virion, making it unable to bind to the target host cell (Roehrig, 2003). Viruses are able to escape neutralisation by antibodies if they contain gene mutations that lead to amino acid substitutions that prevent neutralising antibodies from binding.

1.4 Determinants of viral replication and disease pathogenesis

The causes of dengue pathogenesis are poorly understood at present. Work in this area has been hampered by the lack of appropriate animal models for severe disease. Mouse neurovirulence models are commonly used (Gualano *et al*., 1998; Holzmann *et al*., 1990; Kawano *et al*., 1993) even though dengue rarely causes encephalitis in human infections (Solomon *et al*., 2000). Factors in addition to host genetics (reviewed in Coffey *et al*., 2009) that have been proposed as having an effect on pathogenicity include: the antibody-dependent enhancement (ADE) of infection by the presence of antibodies to previous infection with a different DV type, and the virulence of the infecting virus strain (Halstead, 1988; Rosen, 1977). Consistent with both of these hypotheses is that increased disease severity has been correlated with high viraemia early in the course of illness (Libraty *et al*., 2002; Vaughn *et al*., 2000). DVs isolated from patients with severe disease have been observed to replicate more rapidly and to higher titres in DCs than viruses isolated from patients with less severe disease (Takhampunya *et al*., 2009). It has been observed that during the febrile stage, DHF patients display higher dengue viral loads than DF patients, and that during defervescence (when the symptoms of DHF/DSS occur) these high viral loads are maintained in the DHF patients but not the DF patients for up to six days (Wang *et al*., 2003a). This could be indicative of slower virus clearance as a result of ADE or increased efficiency of virus infection or replication by a more virulent strain.

1.4.1 Antibody-dependent enhancement of dengue virus infection

Recovery from infection by one DV type provides lifelong immunity to that virus type and transient immunity lasting just two to three months against the other DV types. There is an increased risk of severe disease after secondary infection with different DV types or in primary infection of infants born to dengue-immune mothers (Halstead, 1988). This is thought to be due to ADE of infection. When there are existing antibodies to a previous infection with a different DV type, these antibodies are unable to neutralise the current infecting virus and there is enhanced uptake of non-neutralised virus into leukocytes. The production and secretion of chemokines and cytokines by the cell in response to the increased viral replication that follows is

thought to lead to the increase in vascular permeability that is the hallmark of severe dengue. During a DV type 2 epidemic in Cuba in 1997, in which there were an estimated 5208 symptomatic (febrile) cases, 92% of the dengue fever cases and over 98% of the 205 severe disease cases had serologically confirmed secondary infections (Guzmán *et al*., 2000). Cuba had no other endemic flaviviruses at this time so the primary infections for the majority of these individuals were attributed to a DV type 1 epidemic 20 years previously. The explanation given for the high occurrence of secondary infections was that the infection rate may have been higher in those that had DV type 1 antibodies than in those that did not.

This ADE hypothesis is the basis for the current emphasis on development of tetravalent dengue vaccines that are active against all four DV types. However, severe cases of dengue also occur as a result of primary infections (Barnes and Rosen, 1974; Gubler *et al*., 1978; Thomas *et al*., 2008), and some ethnic groups appear to be 'immune' to severe disease despite living in areas of hyperendemicity (Halstead *et al*., 2001). A study in Haiti found no recorded cases of DHF amongst the indigenous population (Halstead *et al*., 2001). This was despite the presence of DV type 2 genotype *I* viruses (which are associated with DHF in Southeast Asia and America) and antibodies to multiple DV types in 85% of the children tested, as well as annual transmission rates consistent with countries endemic for DHF. This indicates that dengue pathogenesis is due to a combination of factors rather than just ADE.

1.4.2 Dengue virus sequence variation and virulence

The four DVs are thought to have evolved from a common ancestor of either African or Asian origin around 1000 years ago. A predominantly sylvatic cycle was maintained with non-human primates as hosts until the establishment of endemic transmission to humans as recently as 100 to 320 years ago, coinciding with a large increase in population growth (reviewed in Holmes and Twiddy, 2003); Figure 1.8). All four DV types have the potential to cause severe disease, and each DV type can be divided into genotypes based on the extent of sequence diversity between strains. In some cases the genotypes correspond to particular geographical regions, with strains diverging from a founding strain after its introduction into a new area.

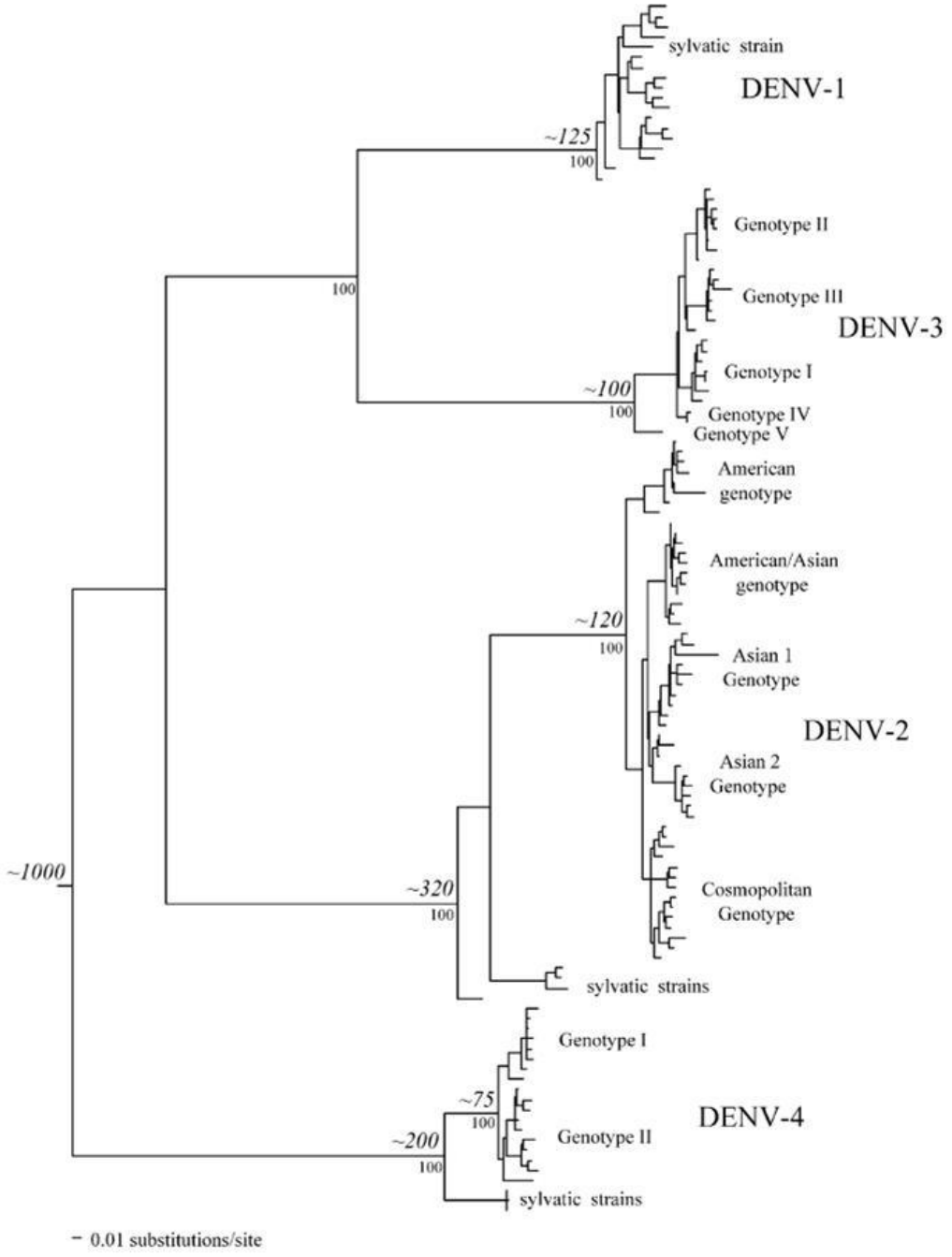


Figure 1.8. Dengue virus genetic relationships and evolutionary history.

Figure is from Holmes and Twiddy, 2003. Phylogenetic analyses were based on E gene comparisons between strains from all four DV types. The approximate divergence times in years are shown above key nodes and the bootstrap values are shown below selected nodes.

Differences in strain virulence within DV types and across genotypes have been attributed to the presence or absence of severe dengue cases during an outbreak (Gubler *et al.*, 1978; Lanciotti *et al.*, 1994; Rico-Hesse *et al.*, 1997b). There is evidence for both DV types 2 and 3 that some genotypes appear to cause only mild disease

whilst others have the potential to cause epidemics of DHF/DSS (Gubler *et al*., 1981; Messer *et al*., 2003a; Rico-Hesse *et al*., 1997b; Vitarana, Jayakuru, and Withane, 1997). Sylvatic DVs are not associated with epidemics in humans. They are transmitted by sylvatic *Aedes* species with non-human primates as hosts. There have been few reports of sylvatic DVs isolated from humans and for the most part they were asymptomatic or the clinical descriptions are not detailed enough to confirm symptomatic dengue (reviewed in Vasilakis, Weaver, and Maramorosch, 2008).

Sequence variation between the five DV type 1 genotypes has been shown to be up to 9%, using comparisons of nucleotide sequence data for a 240 nucleotide region across the E/NS1 junction (Figure 1.9; Rico-Hesse, 1990). When the E gene sequences from 147 strains of DV type 2 were compared, it was found that the five non-sylvatic genotypes differed by an average of 7.3% (Figure 1.10; Twiddy *et al*., 2002). Severe dengue was not epidemic in the Americas until 1981 despite the co-circulation of several DV types and a high incidence of secondary infection within the population. The emergence of epidemic severe dengue coincided with the introduction of the more virulent Southeast Asian genotype of DV type 2, replacing the previous American genotype which was associated with non-severe disease (Rico-Hesse *et al*., 1997a). Subsequent investigations identified sequence differences between the two genotypes in the 5' and 3' NC regions that were predicted to change RNA secondary structures involved in translation initiation and virus replication, and in the E gene which led to an amino acid substitution in the host cell-binding region of E (Leitmeyer *et al*., 1999). In comparison with Southeast Asian genotype DV type 2 strains, American genotype strains exhibit reduced virus output from human monocyte-derived macrophage (Cologna and Rico-Hesse, 2003; Pryor *et al*., 2001) and DC cultures but not mosquito cell cultures (Cologna and Rico-Hesse, 2003). The main determinant of replication efficiency was shown by both groups to be the E amino acid substitution. The American genotype strains have also been shown to replicate and disseminate less efficiently in *Aedes aegypti* mosquitoes when compared to Southeast Asian genotype strains (Armstrong and Rico-Hesse, 2001).

There are four genotypes of DV type 3 (Figure 1.11) with a maximum nucleotide sequence difference between them of around 10% and a protein sequence difference of less than 5% (Lanciotti *et al*., 1994). Genotype *III* clusters into two distinct clades, *IIIa* and *IIIb.* Those strains that comprise genotype *IIIa* are all from outbreaks during or prior to 1989 which were not associated with DHF epidemics. Those that form genotype *IIIb* are all from 1989 or later and have been associated with epidemics of severe dengue (Lanciotti *et al*., 1994; Messer *et al*., 2003a). Interestingly, genotype *IV* strains have never been associated with DHF, although the reason for this is not known (Lanciotti *et al*., 1994). A comparison of 19 DV type 4 E gene sequences from viruses across several decades and geographical regions revealed two genotypes (Figure 1.12) with less than 8% nucleotide sequence diversity and 4% amino acid difference (Lanciotti, Gubler, and Trent, 1997).

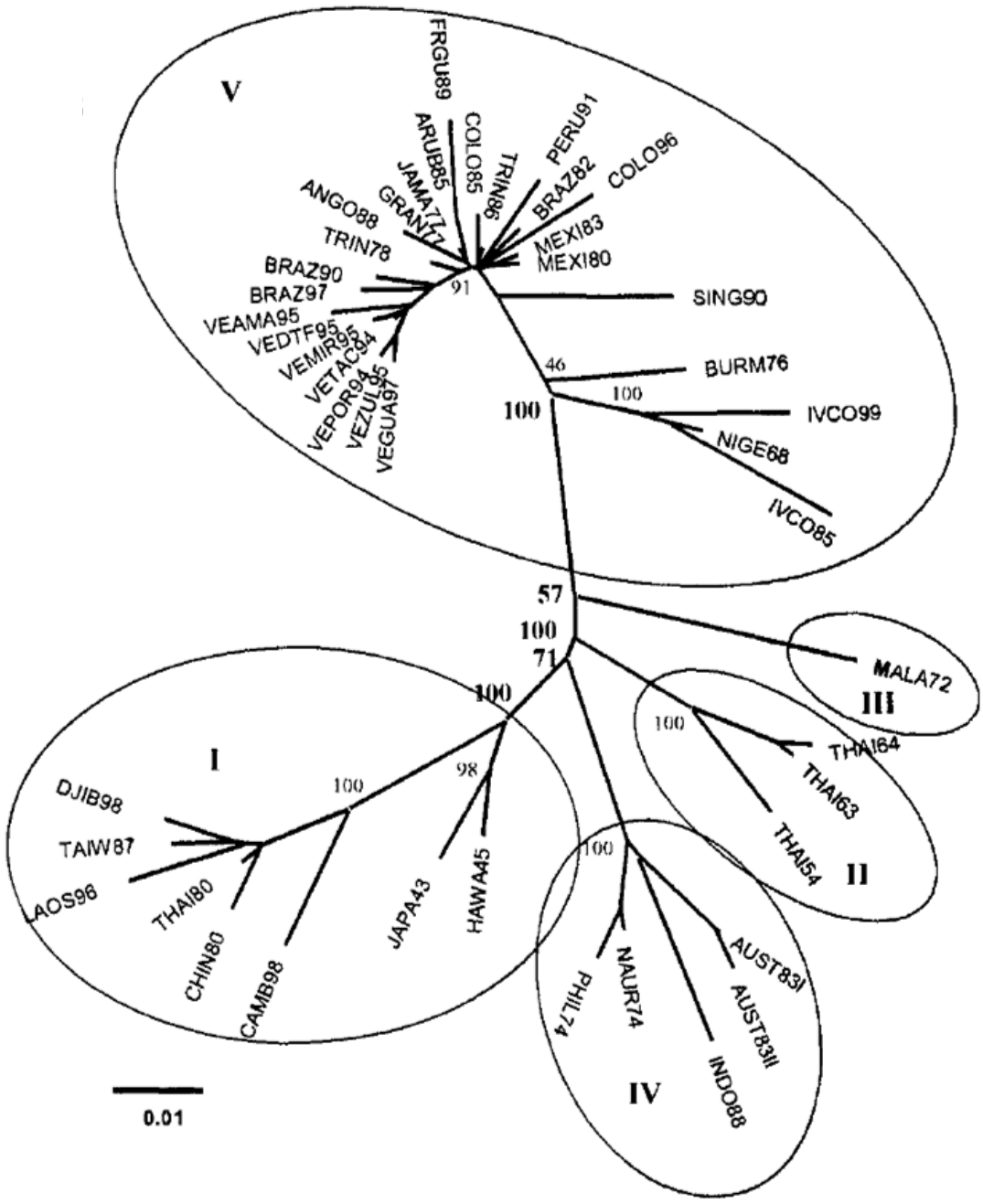


Figure 1.9. Phylogenetic tree showing dengue virus type 1 evolutionary relationships.

Figure is from Goncalvez *et al*., 2002. Genotype *I* contains strains from Japan, Hawaii, Asia and Djibouti. Genotype *II* contains strains from Thailand. Genotype *III* contains a Malaysian sylvatic strain. Genotype *IV* contains strains from Australia, South East Asia and the South Pacific. Genotype *IV* contains strains from Africa, the Americas and South East Asia.

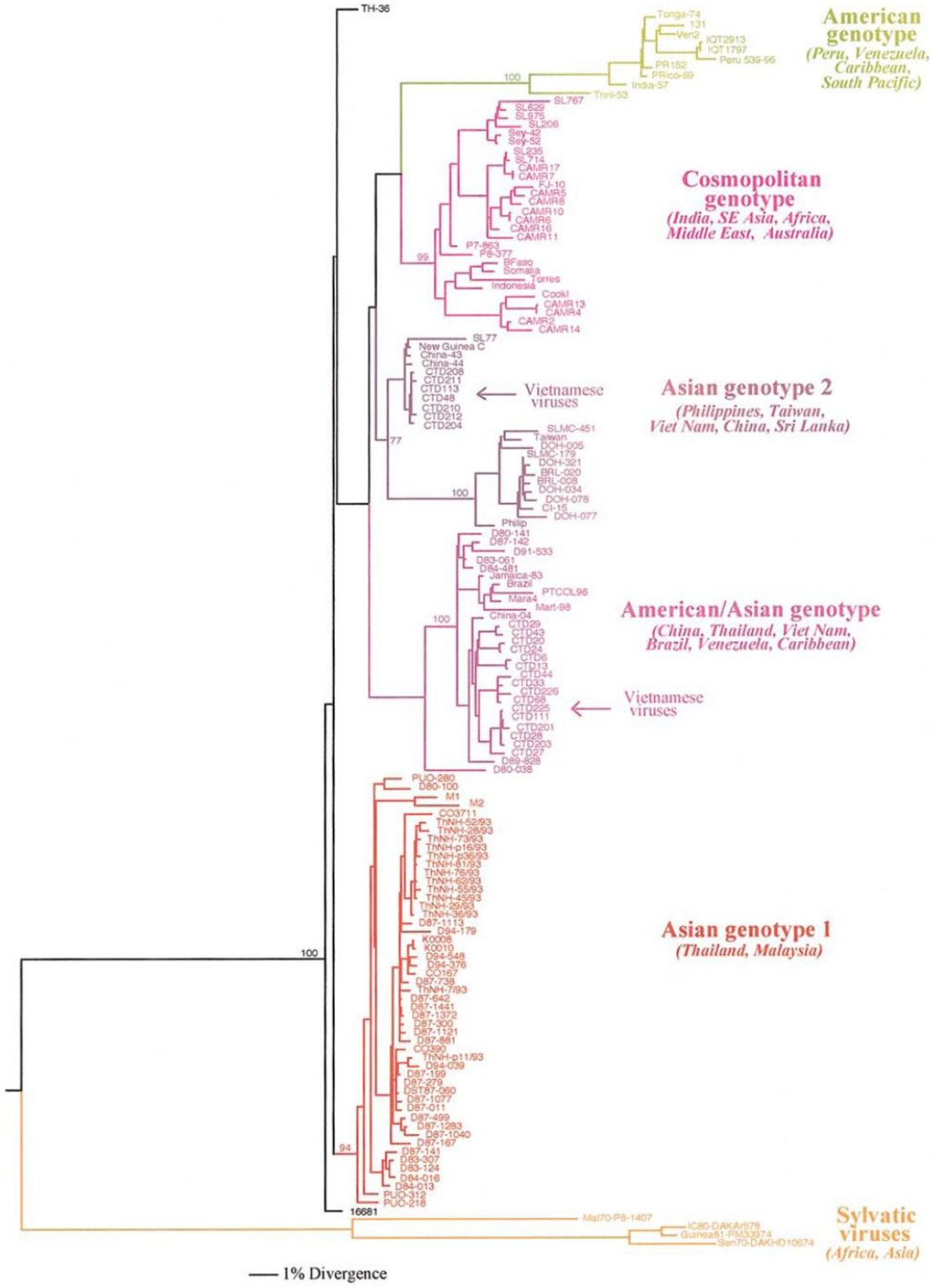


Figure 1.10. Phylogenetic tree showing dengue virus type 2 evolutionary relationships.

Figure is from Twiddy *et al*., 2002. Of the five non-sylvatic DV type 2 genotypes, two contain strains solely from South East Asia, one contains strains from the Americas and older isolates from India and the Pacific Islands, one contains strains from Asia and the Americas and one genotype contains strains with an almost global distribution.

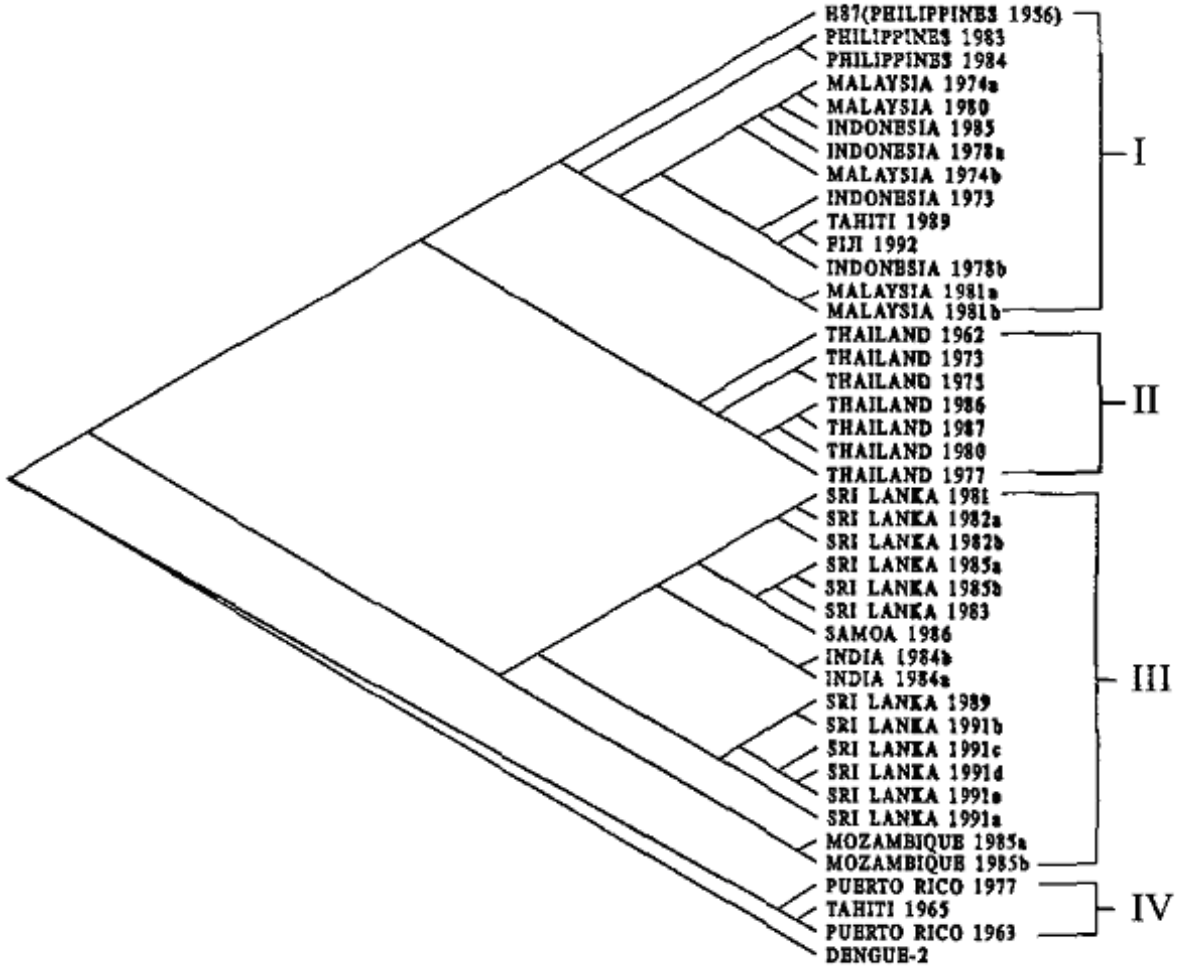


Figure 1.11. Phylogenetic tree showing dengue virus type 3 evolutionary relationships.

Figure is from Lanciotti *et al*., 1994. Genotype *I* strains are mostly from South East Asia. Genotype *II* strains are all from Thailand. Genotype *III* strains are from Sri Lanka, India, Samoa and Mozambique. Genotype *IV* strains are from Puerto Rico and Tahiti.

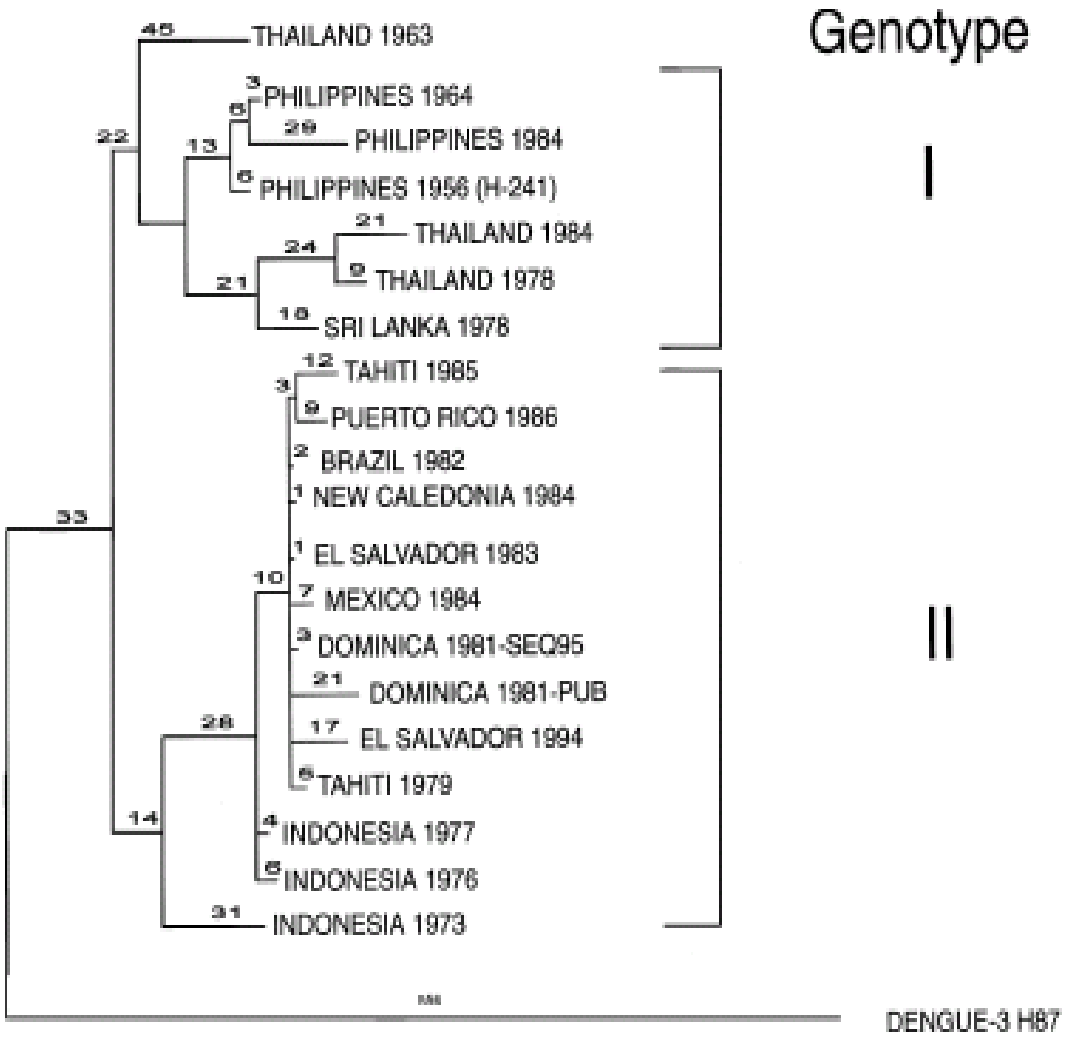


Figure 1.12. Phylogenetic tree showing dengue virus type 4 evolutionary relationships.

Figure is from Lanciotti, Gubler, and Trent, 1997. Genotype *I* comprises strains from Thailand, Sri Lanka and the Philippines. Genotype *II* strains are from Indonesia, Tahiti, the Caribbean islands, and Central and South America.

1.4.3 Dengue in Sri Lanka

Prior to 1989 there was a low incidence of severe dengue in Sri Lanka despite it being a region of hyperendemicity for all four DV types, but since 1989 dengue epidemiology in the region has been characterised by severe disease epidemics (Figure 1.13; Messer *et al*., 2002). No significant differences were found in the relative distribution of DV types, virus transmission rates or the proportion of secondary infections between data collected pre and post 1989 (Messer *et al*., 2002). A genetic shift in DV type 3 genotype *III* strains, favouring genotype *IIIb* strains over the previously dominant

genotype *IIIa* strains, has been proposed as an explanation for the emergence of epidemic severe disease in Sri Lanka since 1989 (Lanciotti *et al.*, 1994). A similar DV type 3 clade replacement has been suggested as the cause of a further increase in severe dengue in Sri Lanka since 2000 (Kanakaratne *et al.*, 2009).

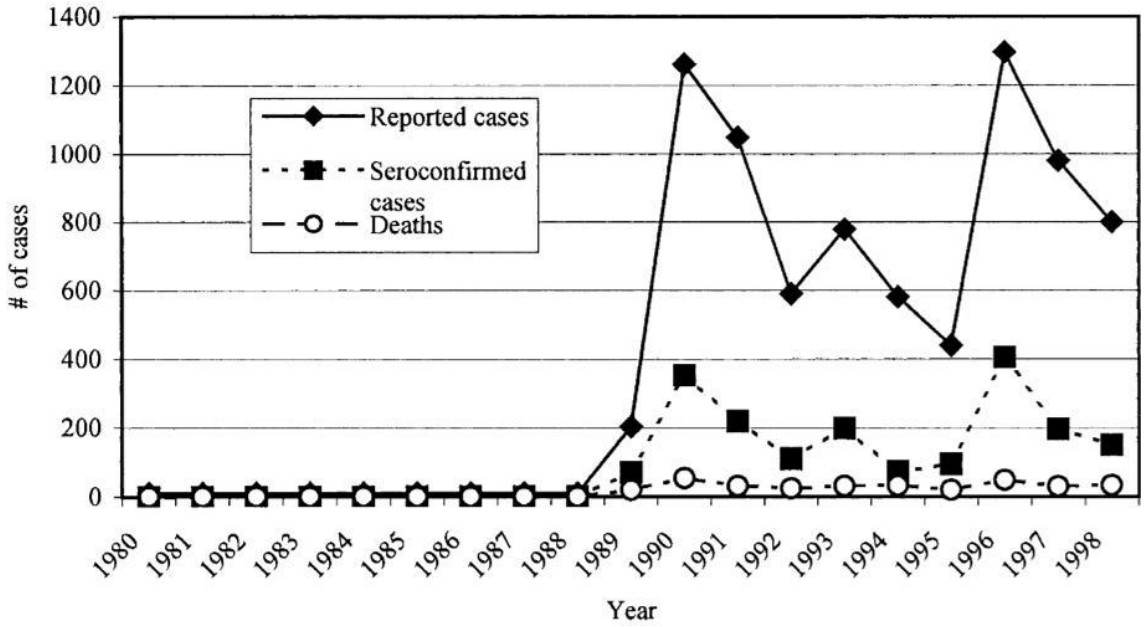


Figure 1.13. Dengue haemorrhagic fever in Sri Lanka between 1980 and 1998.

Figure is from Messer *et al*., 2002. These data are based on cases reported to the Sri Lankan Ministry of Health Epidemiology Unit and samples tested at the Medical Research Institute in Colombo. In this case, the case definition for DHF was at least fever, haemorrhagic symptoms and a platelet count less than 100,000/mm^3. Serological confirmation was based on the HI test.

1.4.4 Recombination in dengue viruses

Phylogenetic analyses of DV genome sequences have demonstrated that DVs undergo recombination between genotypes within the same DV type (AbuBakar, Wong, and Chan, 2002; Holmes, Worobey, and Rambaut, 1999; Tolou *et al*., 2001; Worobey, Rambaut, and Holmes, 1999). This may be as a result of the RNA polymerase changing templates mid-way through replication in a cell that contains more than one DV genotype, as has been suggested for polio virus recombination (Cooper *et al*., 1974). This is able to occur in areas of hyperendemicity where multiple DV types and several genotypes of the same DV type are endemic. Evidence for co-infection with multiple DV types in humans is documented (Bharaj *et al*., 2008; Wang *et al*., 2003b; Wenming

et al., 2005), whereas the evidence for co-infection with multiple genotypes is scarce. This is because diagnostic tests are able to distinguish between DV types whilst not being able to discern the presence of multiple genotypes within a DV type. The presence of multiple DV type 2 genotypes within a single mosquito, including a recombinant virus along with both parental strains has been documented (Craig *et al*., 2003). Similar findings were also reported from a human DV type 1 sample (Aaskov *et al*., 2007).

These findings suggest the potential for emergence of recombinant DVs with altered pathogenicity compared to parental strains. Whilst there is no evidence for recombination between DV types, there may also be implications for vaccine design and safety in relation to potential recombination of attenuated live vaccine strains, either with other strains within a multivalent vaccine or with non-vaccine strains (Worobey, Rambaut, and Holmes, 1999). This could lead to the production of unattenuated recombinant viruses and vaccine-associated disease as has been observed with polio virus (Georgescu *et al*., 1994).

1.4.5 Dengue virus as a quasispecies

Within a single host, a RNA virus exists as a population of many variant forms due to the error-prone nature of RNA virus replication. These populations are termed quasispecies and they contain enough phenotypic variants to be able to respond rapidly to new selection pressures (DeFilippis and Villarreal, 2001; Domingo, 1998). Rates of spontaneous mutation among RNA viruses have been estimated at approximately one mutation in every 10,000 nucleotides ($1x10^{-4}$) during natural replication (Smith *et al*., 1997) which equates to around one nucleotide mutation per DV genome per replication (Drake, 1993). These quasispecies populations must be regulated to maintain genetic and antigenic diversity in a stable manner. The replication rate and mutation rate must remain within acceptable limits to ensure the survival of both virus and host whilst allowing the virus to avoid clearance by the host immune system.

Quasispecies populations have been well studied in viral infections such as human immunodeficiency virus (HIV; Garcia-Arriaza, Domingo, and Briones, 2007), hepatitis E (Grandadam *et al*., 2004), foot and mouth disease virus (FMDV; Domingo *et al*., 2002) and hepatitis C (Martell *et al*., 1992). There are few publications of quasispecies investigations in flaviviruses. DV quasispecies populations were first identified in 2002 in DV type 3 populations from six human plasma samples (from an outbreak in Taiwan in 1998; Wang *et al*., 2002a). The same group then compared DV type 3 populations from naturally infected mosquitoes collected during the 1998 outbreak with the viral populations found in the human samples. Less quasispecies diversity was identified in the mosquito-derived viruses than in the human-derived viruses (Lin *et al*., 2004). A study of DV type 2 populations also examined samples from both humans and mosquitoes (from samples collected in Myanmar between 1998 and 2000), and reported multiple genotypes as well as recombinant viruses within the same host (Craig *et al*., 2003). These studies used RT-PCR, cloning and sequencing of partial (Lin *et al*., 2004; Wang *et al*., 2002a) or complete (Craig *et al*., 2003) E genes to study the extent of sequence diversity within the virus populations of individual hosts.

During HIV-1 infection, the detection of minor quasispecies variants (between 0.1 and 10% of the genome population) is becoming increasingly important in genotyping the virus and predicting the response to different anti-viral treatments throughout the course of the infection (Garcia-Arriaza, Domingo, and Briones, 2007). Little progress has been made in linking elements of the flavivirus quasispecies population to disease pathogenesis, although a higher frequency of genome-defective virus with in-frame stop codons due to insertion or deletion mutations has been reported in patients with DHF compared to patients with DF (Wang *et al*., 2002a).

The nature of the regulatory mechanisms that maintain population stability whilst allowing genetic and antigenic diversity is not well understood but it has been proposed that certain RNA viruses are able to autoregulate genetic mutation via interaction with the viral RNA polymerase itself (Sallie, 2005). The theory is that in viruses capable of persistent infection the actions of the RNA polymerase are

influenced by the viral E gene. Wild-type and variant E genes show different interactions with the polymerase in order to alter fidelity and processivity, and maintain a stable quasispecies population. Accumulation of variant E genes inhibits processivity of the polymerase and improves its fidelity so that the genes that are produced are closer to the original wild-type sequence. These wild-type E genes in turn interact with the RNA polymerase and encourage the production of more variants by increasing processivity and infidelity. This mechanism for mediating stable viral replication and promoting antigenic diversity has been termed 'replicative homeostasis' (Sallie, 2005).

Mosquito-borne flaviviruses are rarely capable of persistent infection in humans, (although it has been documented in neuronal tissues; reviewed in Burke and Monath, 2001) but they do cause persistent non-cytolytic infections in their mosquito vectors (Lindenbach and Rice, 2001). The observation of less DV sequence diversity in mosquitoes compared to human hosts (Lin *et al*., 2004) may be explained by regulatory mechanisms such as replicative homeostasis acting during the persistent infection to maintain evolutionary conservation of the virus. Replicative homeostasis may also play a role in maintaining population stability during acute RNA-virus infections. Arthropod-borne flaviviruses must be able to infect and replicate effectively in both the insect vector and vertebrate host cells, which imposes significant evolutionary constraints on viral genetic variation. This is supported by the finding that high levels of intrahost (quasispecies) short-term variation are not maintained in the long-term at an interhost level (Holmes and Twiddy, 2003).

1.5 Flavivirus envelope proteins and their role in virulence

The majority of the surface of flavivirus virions is composed of E, which are heavily involved in mediating attachment of the virus to the host cell and in membrane fusion. Mutations within the E gene that alter E antibody binding, cell-attachment or membrane fusion therefore have the potential to affect virulence or cell tropism, and consequently disease severity.

1.5.1 Flavivirus envelope protein structure

Much of the early knowledge about the antigenic structure of flavivirus E came from diagnostic serology experiments. These experiments classified the flaviviruses into groups using techniques such as virus neutralisation, heamagglutination inhibition and complement fixation (reviewed in Roehrig, 2003). MAbs were used in conjunction with denaturation, low pH, proteolysis and reduction of disulphide bridges to construct epitope maps (Heinz and Maramorosch, 1986; Mandl *et al*., 1989). Flavivirus gene sequence data began to be reported from 1985 (Deubel, Kinney, and Trent, 1986; Mandl, Heinz, and Kunz, 1988; Mason *et al*., 1987; Rice *et al*., 1985; Zhao *et al*., 1986), which enabled the epitope maps to be linked to the sequence data. The locations of 6 disulphide bridges formed between the 12 cysteine residues in the WNV E were established (Nowak and Wengler, 1987) and were subsequently shown to be conserved in both mosquito and tick-borne flaviviruses (Rey *et al*., 1995). A model of the antigenic structure of flavivirus E was produced based on TBEV (Figure 1.14; Mandl *et al*., 1989). The model proposed that flavivirus E have three antigenic domains; A, B and C containing 16 epitopes. Domain A contains epitopes that are cross-reactive as well as those that are specific for tick-borne flaviviruses. Domain B also contains epitopes that are specific for tick-borne flaviviruses, whereas domain C epitopes are mostly sub-type specific (Mandl *et al*., 1989).

The first flavivirus E crystal structure was determined for TBEV in 1995 (Figure 1.15; Rey *et al*., 1995). This revealed a molecular architecture very different from other viruses (e.g. influenza) known at the time, with E lying flat as homodimers along the outer surface of the virus lipid bilayer, rather than forming spike-like projections. Three structural domains were identified: *I*, *II* and *III*, corresponding to the previously

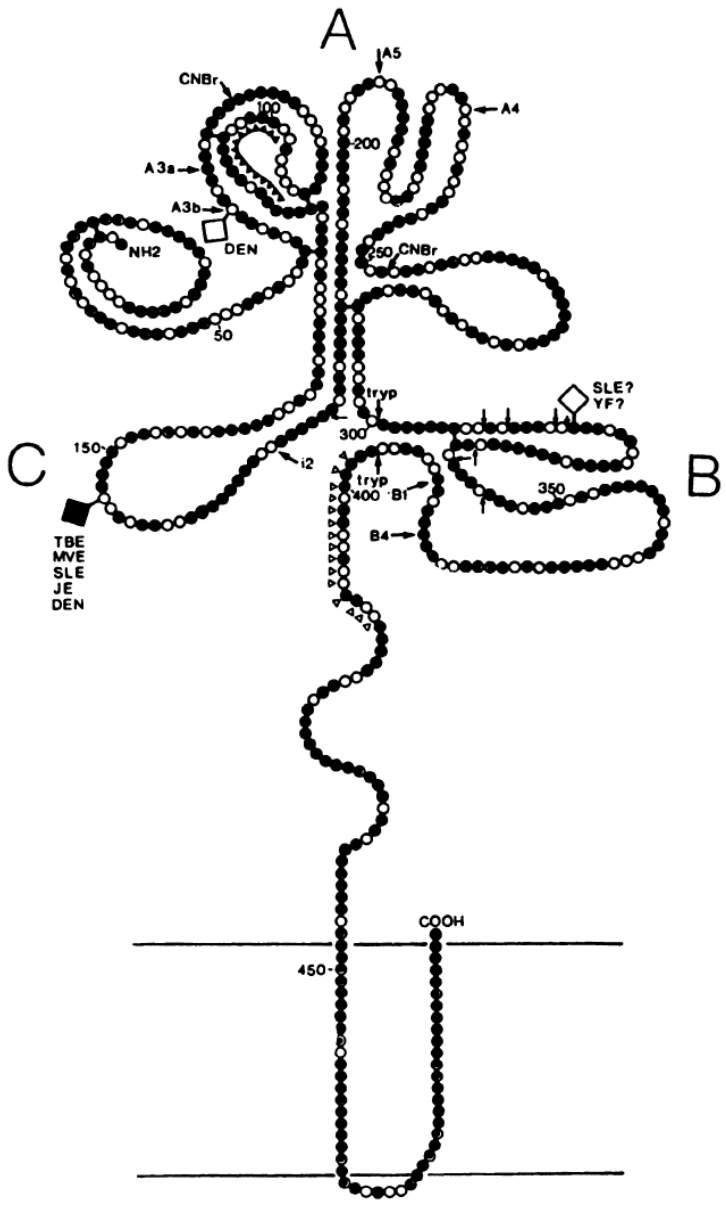


Figure 1.14. Flavivirus envelope protein antigenic model.

Figure and legend are from Mandl *et al.*, 1989. Open circles represent hydrophilic amino acid residues (Arg, Lys, Asn, Asp, Gln, Glu, His), dotted circles show intermediate amino acid residues (Pro, Tyr, Ser, Trp, Thr, Gly), and solid circles show hydrophobic amino acid residues (Ile, Val, Leu, Phe, Cys, Met, Ala). Position numbers are shown every 50 amino acids. Cysteine residues forming disulphide bridges are connected by solid lines. Arrows depict cleavage sites that liberate immune-reactive fragments; IRF1 (trypsin) and IRF3 (CNBr), respectively. Small arrows indicate potential cleavage sites within these fragments that are not utilized. Two solid lines stand for the lipid membrane that is spanned by two transmembrane regions of E. The polypeptide chain is folded to indicate the antigenic domains A, B, and C. Arrows together with the names of neutralizing monoclonal antibodies (mAbs) depict the locations of the mutations identified in the respective antigenic variants of TBEV by sequence analysis. A line of solid triangles indicates the almost perfectly conserved sequence within domain A. A line of open triangles marks the region of a potential T-cell determinant. A solid diamond represents a site of N-glycosylation common to TBEV, MVEV, SLEV, JEV and dengue (DEN) viruses. YFV and SLEV have an additional potential N-glycosylation site within domain B, DVs within domain A.

defined C, A and B antigenic domains respectively (Rey *et al.*, 1995). Domain *I* is an eight-stranded β-barrel with an additional short amino-terminal strand at one edge. It has an N-linked glycosylation site at Asn-153 which is conserved amongst most flaviviruses and is central to the other two domains. Domain *II* is an elongated dimerisation region consisting of twelve β-strands and two α-helices with the fusion peptide at its base. The fusion peptide is within a region (amino acids 98-113) conserved amongst most flaviviruses. Domain *III* is an immunoglobulin-like module with ten β-strands which form a barrel structure closed by an additional sheet. It is thought to contain the flavivirus receptor-binding motif (Rey *et al.*, 1995).

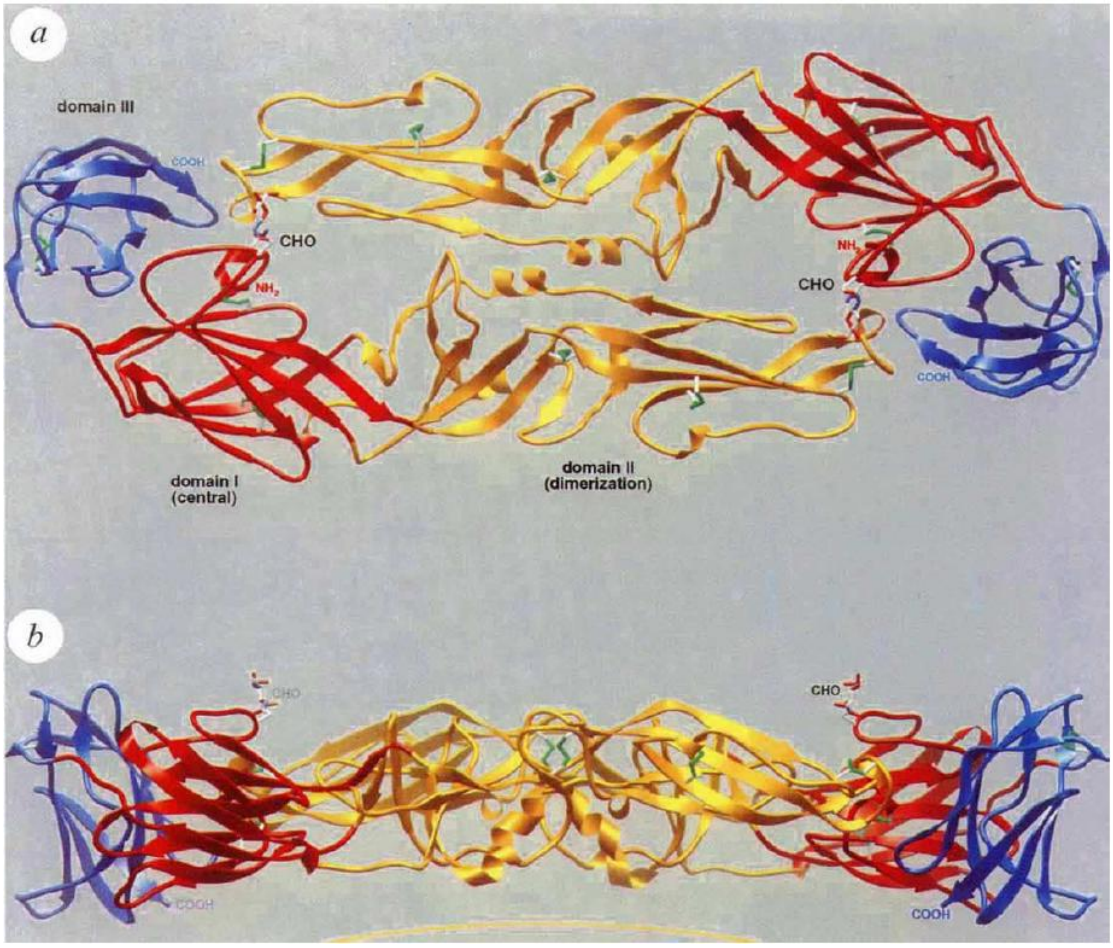


Figure 1.15. Tick-borne encephalitis virus envelope protein structure.

Figure is from Rey *et al.*, 1995. The two E polypeptide chains form a head-to-tail homodimer. β-strands are represented by ribbons with arrowheads; alpha helices by coiled ribbons; connecting loops by thin tubes. The disulphide bridges are shown by ball-and-stick representations of the cystine residues (green). The single glycan on each dimer, labelled CHO is also drawn as ball-and-stick. Domains *I* (red), *II* (yellow) and *III* (blue) are shown. (a) The dimer viewed as if looking down on the outside surface of the virus particle. (b) The dimer viewed from the side as if looking along the surface of the virus particle.

DV E crystal structures were not published until much later (Kuhn *et al*., 2002; Modis *et al*., 2003) but were shown to closely resemble the TBEV E so that determinants of host range, tropism or virulence map to the same locations in the DV and TBEV structures (Modis *et al*., 2003). The major differences between the TBEV and DV E crystal structures are that DVs have an additional N-linked glycosylation site within domain *II* at Asn-67, and an additional four amino acids in domain *III* that form a surface-exposed loop implicated in receptor binding (Modis *et al*., 2003). The DV N-glycosylation sites at Asn-67 and Asn-153 are involved in infection, propagation and secretion of the virus (Mondotte *et al*., 2007). Mutations at these sites could therefore affect infectivity.

1.5.2 Flavivirus envelope gene mutations and altered pathogenesis

For many flaviviruses, single nucleotide mutations within the E gene have been shown to be associated with altered virulence or pathogenicity (Cecilia and Gould, 1991; Guirakhoo *et al*., 2004; Hahn *et al*., 1987; Hasegawa *et al*., 1992; Holzmann *et al*., 1990; Jiang *et al*., 1993; Nitayaphan *et al*., 1990). Mutations in the flavivirus E gene that alter the properties of the virus tend to cluster at three sites on E. These sites are the distal face of domain *III*, the base of domain *II*, and the domain *I*/*III* interface with the domain *II* conserved region at residues 97-113 (Figure 1.16; Rey *et al*., 1995). Mutations causing residue substitutions at the base of domain *II* or at the domain *I*/*III* interface with the domain *II* conserved region at residues 98-113 are thought to affect virulence or antibody neutralisation of virus by impacting upon the conformational rearrangements triggered by fusion (Rey *et al*., 1995). Some of these mutations have been shown to affect fusion by altering the pH required for the fusion-activating conformational change (Mandl *et al*., 1989; Modis *et al*., 2003; Rey *et al*., 1995). Most of these mutations involve substitution of residues with side-chains that project into the ligand-binding pockets between the dimers, and it has been reported that lowering of the pH threshold for conformational change is due to the substitution of longer hydrophobic side chains for shorter ones (Modis *et al*., 2003). During routine passage of a chimeric YFV/DV type 1 vaccine strain, a mutation became incorporated which substituted the lysine residue at E position E204 with an arginine residue (Guirakhoo *et al*., 2004). This resulted in reduced neurovirulence and viraemia in suckling mice and

monkeys respectively. Residue E204 resides within the ligand binding site involved in virus fusion with cell membranes, and the mutation may have attenuated the virus by interfering with this process (Guirakhoo *et al*., 2004). Mutations resulting in substitutions on the distal face of domain *III* can affect virulence or cell tropism and it has been suggested that this is due to interference with cell attachment, as domain *III* is thought to contain the flavivirus receptor-binding motif (Crill and Roehrig, 2001). Neutralisation escape mutations have also been mapped to this part of DV type 2 E (Roehrig, Bolin, and Kelly, 1998), consistent with antibody neutralisation of virus by blocking cell attachment.

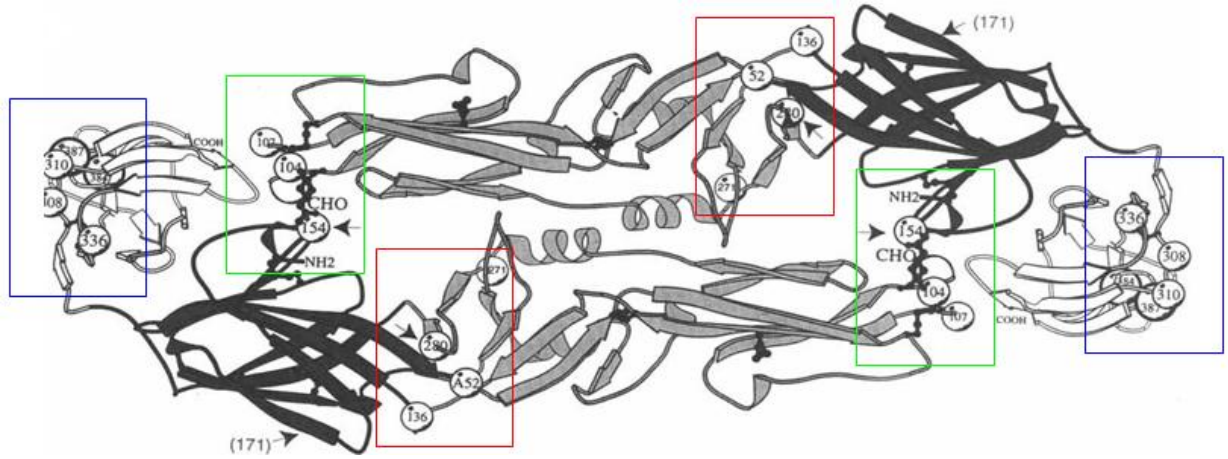


Figure 1.16. Clustering of mutations at envelope protein structural domain interfaces.

Figure adapted from Rey *et al*., 1995. Domain *I* from each molecule of the dimer is in black, domain *II* is in grey, and domain *III* is in white. Blue boxes highlight mutations clustering at the distal face of domain *III*. Red boxes highlight the clustered mutations at the base of domain *II*. Green boxes highlight mutations clustered at the domain *I/III* interface with the domain *II* conserved region at residues 97-113. Mutation positions are labelled according to the positions in TBEV Neudörfl strain. Mutations are from JEV, MVEV, Louping ill, TBEV, DV and TBEV/DV chimera viruses. Arrows mark sites where mutations influence the threshold pH for the fusion-activating conformational change.

Various single amino acid substitutions at residue E390 of DV type 2 E have been implicated in phenotypic and virulence changes in the virus. During plaque purification of a Mexican DV type 2 strain (Sanchez and Ruiz, 1996), differences in plaque size were observed between clones where the E390 aspartic acid (Asp) residue was substituted with either histidine (His) or asparagine (Asn). The Asp to His change was associated with the production of large lytic plaques and the virus was highly virulent in suckling mice, whereas the Asp to Asn change was associated with the production of small lytic

plaques and the virus was less virulent in suckling mice. In American genotype dengue type 2 viruses known only to cause mild disease in humans, the residue at E390 is Asp whereas in Southeast Asian genotype viruses associated with DHF epidemics the residue at E390 is Asn (Leitmeyer *et al*., 1999). Subsequent investigation of these two viruses has shown that infectious clones with Asp at E390 (American genotype) display reduced viral RNA output from infected cells compared to clones with Asn at E390 (Southeast Asian genotype; Cologna and Rico-Hesse, 2003). Substitutions at residue E390 of DV type 2 E are thought to affect the ability of the virus to bind to the host cell surface. This residue occurs within one of two putative glycosaminoglycan binding motifs (E284 to 310 and E386 to 411) containing many basic amino acids (Chen *et al*., 1997), and the change in charge of the substituted residue has the potential to alter attachment (Leitmeyer *et al*., 1999). Amino acid substitutions in this position have also been identified in the YFV 17D strain, which is the attenuated vaccine strain derived from the virulent Asibi strain (Hahn *et al*., 1987), and in attenuated MVEV (Lobigs *et al*., 1990). In JEV, MVEV and YFV 17D, this site forms part of an Arg-Gly-Asp (RGD) sequence implicated in virus-cell attachment. In the case of both YFV 17D and the MVEV strains mentioned above, the E390 mutations were induced by adaptation to culture and resulted in virus attenuation in the original host, which is consistent with the role of the RGD sequence as a receptor-binding motif (Hahn *et al*., 1987; Lobigs *et al*., 1990).

For DVs, E has been shown to contain a binding site (E306-314) for highly sulphated heparin sulphate (HSHS) molecules which are present on the surface of cells targeted by the virus and appear to mediate infectivity (Chen *et al*., 1997; Thullier *et al*., 2001). Neutralisation escape mutations that lead to reduced mouse neurovirulence with Louping ill virus have been located at E positions E308-311 (Jiang *et al*., 1993). Similarly, neutralisation escape mutations in the DV type 2 E have been described at E position E307 (Lin *et al*., 1994). The presence of these neutralisation escape mutations within the HSHS binding site indicates a mechanism for virus neutralisation by antibody binding to the HSHS binding site and thereby preventing attachment of the virus to the host cell.

During adaptation of a human DV type 1 isolate to grow in suckling mouse brain, several amino acid differences were identified between a mouse-passaged variant that caused neurovirulence and the original isolate (Despres *et al.*, 1998). These differences were located within both E and NS3 (the NTPase/helicase). Interestingly, all the substitutions in the neurovirulent strain mapped to the interfaces between the structural domains of E. The virus titres were similar in both strains despite there being a reduced rate of production of viral proteins *in vitro* for the neurovirulent strain. The authors suggested this was due to the E mutations (E196 and E405) positively affecting assembly of glycoprotein oligomers (Duarte dos Santos *et al.*, 2000).

In summary, mutations in the E gene that lead to residue substitutions in E can alter the properties of the virus if they occur:

1. On the surface, in regions involved in antibody recognition, cell attachment or oligomer assembly
2. At the interfaces between structural domains where they can affect the conformational changes required for fusion of virus and host cell membranes.

1.6 Hypothesis, aims and objectives

1.6.1 Hypothesis

DV quasispecies populations affect disease severity in individual patients.

1.6.2 Aims and objectives

The overall aim of this study was to provide in-depth characterisation of DV quasispecies populations in individual patients, in terms of both the extent of DV E gene variation, and the presence of specific variants. This study also aimed to investigate whether there was a relationship between quasispecies variation and disease severity, by analysis of DV E gene sequences. Recombinant E proteins would be produced to further investigate the effects of the quasispecies mutations on protein function. Identification of DV quasispecies E variants that do not bind flavivirus cross-reactive antibodies would lead to improved tools for serodiagnosis of DV infection. This work was subdivided into specific objectives which are described in sections 1.6.2.1 to 1.6.2.5 below.

1.6.2.1 Objective 1: To investigate the overall extent of dengue virus envelope gene quasispecies variation within individual clinical samples and correlations with disease severity.

There have been few studies documenting the extent of variation in DV quasispecies populations in individual hosts (Craig *et al*., 2003; Lin *et al*., 2004; Wang *et al*., 2002a). These groups focused on the DV type 2 or 3 E genes and the methodologies used were not ideal as they permitted the introduction of mutations via the RT-PCR amplification process itself. Both the number of patients (between 3 and 8) and the number of clones analysed per patient was low (between 10 and 21). Only one group (Craig *et al*., 2003) looked at the complete E gene sequence and used high-fidelity PCR for gene amplification. Therefore, this study intended to use high-fidelity reverse transcription and PCR to look at E genes from all four DV types across a greater number of clinical samples from the same outbreak and a greater number of clones per sample. This would determine whether the overall extent of DV E gene variation within a patient sample correlates with disease severity.

1.6.2.2 Objective 2: To investigate the relationship between the presence of specific dengue virus envelope gene variants and disease severity.

Little progress has been made in linking specific variants within the DV quasispecies population to disease severity, although a higher frequency of defective genomes with in-frame stop codons have been reported in patients with DHF compared to those with DF (Wang *et al*., 2002a). This study intended to detail the types of mutation found within the DV E gene variants and investigate correlations with disease severity.

1.6.2.3 Objective 3: To identify the dengue virus genotypes present within the intra and inter-host populations and look for evidence of recombination.

Virus genotype has a direct affect on DV transmission by mosquitoes and pathogenicity in humans (Armstrong and Rico-Hesse, 2001; Cologna and Rico-Hesse, 2003; Leitmeyer *et al*., 1999). The presence of a recombinant DV and both parental DV type 2 strains (from different genotypes) within the same mosquito has been documented (Craig *et al*., 2003). Similar findings were also reported from a human clinical DV type 1 sample (Aaskov *et al*., 2007). Recombinant viruses could have reduced or increased virulence compared to either or both of the parental strains and thus could affect disease transmission and pathogenicity. This study would use phylogenetic analyses to genotype the DV E gene variants from the samples from the Sri Lankan outbreak, and to look for evidence of recombination.

1.6.2.4 Objective 4: To identify dengue virus envelope gene variants with the potential to affect disease severity

The few published studies of quasispecies populations in DV infections have all looked at the E gene sequence but have not progressed to the implications for the E sequence and the folded protein. This study intended to map the locations of the amino acid substitutions resulting from nonsynonymous mutations found within the DV E gene quasispecies variants, using the linear and three-dimensional structural models of the flavivirus E. This was to identify those amino acid substitutions located in regions of the protein involved in virion or oligomer assembly, antibody recognition or virus-cell interactions, with the potential to affect disease severity.

1.6.2.5 Objective 5: To construct recombinant dengue virus envelope glycoproteins (from those identified in the previous objective) that enable comparison of quasispecies consensus and variant proteins through binding studies.

This study intended to compare recombinant DV quasispecies consensus and variant E by ELISA in terms of binding to cell surface molecules implicated in DV-cell binding (HSHS and DC-SIGN), and binding to antibodies (both commercially available and those provided by flavivirus-vaccinated individuals). The antibody and cell surface molecule binding experiments would determine whether specific variants within the quasispecies population would exhibit altered virus-antibody or virus-cell binding affinities compared to the quasispecies consensus E. The antibody binding experiments would aid in comparison of existing commercially available antibodies and could potentially identify DV E variants that did not bind flavivirus cross-reactive antibodies. These proteins could be used as improved tools for serodiagnosis of DV infection as they would overcome the current difficulties with flavivirus cross-reactivity during serological testing. The recombinant quasispecies consensus and variant DV E would also be assessed for differences in their ability to undergo the low pH-induced conformational changes required for fusion of virus and host cell membranes, as this could potentially affect virulence and disease severity.

CHAPTER 2. MATERIALS AND METHODS

2.1 General information

This section provides general information on the chemicals, reagents, oligonucleotides, bacterial strains, insect cell lines and culture medium used throughout this thesis.

2.1.1 Chemicals and reagents

Unless otherwise stated, general chemicals and reagents were obtained from Sigma Chemicals Company Ltd. UK, SAFC Biosciences Ltd. UK, Invitrogen™ UK, Merck Chemicals Ltd. UK, Thermo Fisher Scientific Ltd. UK, Promega Ltd. UK and GE Healthcare Ltd. UK. All solution concentrations were weight per volume (w/v) unless otherwise indicated as volume per volume (v/v) and where required solutions were sterilised by autoclave. Restriction endonucleases and modifying enzymes including the appropriate buffers were obtained from Promega Ltd. UK or New England Biolabs (NEB) Ltd. UK unless otherwise stated. PCR and RT-PCR kits were obtained from Qiagen Ltd. UK, Stratagene Ltd. UK, Applied Biosystems Ltd. UK or Roche Ltd. UK. DNA ladders and loading dyes were obtained from Qiagen Ltd. UK. Antibodies were obtained from Abcam Ltd. UK, eBioscience Ltd. UK or Thermo Fisher Scientific Ltd. UK.

2.1.2 Oligonucleotides

All oligonucleotides were custom synthesised by Sigma-Aldrich Ltd. UK and purified by desalt.

2.1.3 Flavivirus RNA used as a positive control template for RT-PCR

The flavivirus RNA used as positive control template for RT-PCR was obtained from HPA CEPR. Virus strains were: YFV 17D strain, JEV Nakayama strain, TBEV Neudörfl strain, DV type 1 Hawaii A strain, DV type 2 New Guinea C strain, DV type 3 H87 strain and DV type 4 H241 strain.

2.1.4 Dengue patient material

Dengue patient material was kindly donated by Major Mark Bailey of the Royal Army Medical Corps (Royal Centre for Defence Medicine, Birmingham, UK). Serum samples were taken from patients presenting with fever on admission to hospitals in and around Ragama, Sri Lanka. Ethical considerations for the use of these samples and accompanying clinical data are presented in Appendix L.

2.1.5 Bacterial strains and culture media

The modified *Escherichia coli* K12 JM109 strain (Promega) was used for cloning and plasmid propagation. The genotype of the strain is *end*A1, *rec*A1, *gyr*A96, *thi-1, hsd*R17 (r_k^-, m_k^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F´ *tra*D36, *pro*AB, *laq*I^{q}ZΔM15].

Bacterial culture media were obtained from Invitrogen™ UK or bioMérieux UK.

2.1.5.1 Glossary of mutations

*end*A1	This mutation results in improved plasmid yields through inactivation of an endonuclease that would otherwise co-purify with plasmids during purification.
*rec*A1	The *rec*A gene is involved in recombination and DNA repair. This mutation limits recombination of the plasmid with the *E. coli* genome, making the insert more stable.
*gyr*A96	This mutation within the DNA gyrase gene confers resistance to nalidixic acid.
thi-1	This mutation affects thiamine metabolism so there is a requirement for thiamine for growth in minimal media.
*hsd*R17 (r_k^-, m_k^+)	This mutation is within the K strain DNA restriction and methylation system. Cleavage of transformed DNA by endogenous restriction endonucleases is prevented (r_k^-) but methylation by endogenous methylases is permitted (m_k^+).
*rel*A1	This mutation of ppGpp synthetase I enables RNA synthesis in the absence of protein synthesis.
*sup*E44	This mutation suppresses amber (UAG) mutations in termination codons.
Δ *lac*	The β-galactosidase gene is deleted on the chromosome.
*pro*AB	The *pro*AB mutations of proline metabolism confer a requirement for proline for growth in minimal media.
F´	F' strains are sensitive to infection by M13 phage.
*tra*D36	The *tra*D36 transfer factor mutation prevents transfer of the F' episome.
*laq*I^q	*laqI^q* strains over-produce the *lac* repressor protein, inhibiting transcription from the *lac* promoter.
*lac*ZΔM15	The β-galactosidase gene is partially deleted, enabling restoration of β-galactosidase activity via an α-complementation sequence carried by some plasmid vectors and blue/white screening for positive transformants.

2.1.6 Insect cell lines and culture medium

Spodoptera frugiperda (*Sf*) 9 and *Sf*21 insect cell lines were obtained from Invitrogen™ (UK). The *Sf9* Easy-Titre (*Sf*9 E-T) insect cell line was kindly provided by Dr. Ralph Hopkins of the National Cancer Institute at Frederick, Maryland, USA. *Sf*21 and *Sf*9 E-T cells were cultured in *Sf*900 II serum-free and HyClone SFX insect cell culture medium

respectively (Invitrogen™ UK). *Sf*9 cells were cultured in Ex-cell 420 serum-free medium (SFM) with L-glutamine (AFC Biosciences Ltd. UK).

2.1.7 Insect viruses and virus transfer vectors

*Flash*BAC GOLD™ (Oxford Expression Technologies Ltd. UK) and the transfer vector pBAC-2cp (Merck Chemicals Ltd. UK) were used for the construction of recombinant baculoviruses (rBVs).

2.2 In silico techniques

This section details the computer-based methodology used for the acquisition and analysis of sequence data for RT-PCR or PCR primer design, sequencing and protein structure modelling. Also described is the rationale for primer design.

2.2.1 Acquisition of published Flavivirus sequence data

Published sequence data were obtained via the nucleotide or protein sequence databases of the National Centre for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/) for DV types 1, 2, 3 and 4. E gene positions were as defined in the NCBI Reference Sequences (RefSeqs) NC_001477 (type 1), NC_001474 (type 2), NC_001475 (type 3) and NC_002640 (type 4).

2.2.2 Sequence analysis for primer design

Sequence analysis was performed using DNAStar (Lasergene 7 sequence analysis software, DNAStar Inc.). To identify suitable primer sites, multiple sequence alignments were performed on E gene nucleotide or amino acid sequences and the flanking regions using the Clustal W algorithm within the alignment program Megalign.

2.2.3 RT-PCR, PCR and sequencing PCR primer design

Primers were designed to have a length between 18 and 30 nucleotides and a melting temperature (T_m) between 55°C and 80°C. The 3' end of each primer was finished with a guanine-cytosine (GC) clamp, and mismatches and runs of three or more consecutive G or C nucleotides were avoided. Primers were checked using the Sigma-Aldrich DNA Oligo Design Tool (www.sigmaaldrich.com) to estimate their melting temperature and ensure they did not have any significant secondary structure or the ability to dimerise. They were also checked for specificity using the nucleotide database (blastn algorithm) within the NCBI Basic Local Alignment Search Tool (BLAST; http://blast.ncbi.nlm.nih.gov/Blast.cgi).

DV species-specific primers for quasispecies analysis were designed up to 200 nucleotides outside the E gene. This was to allow the whole of the E gene to be sequenced in subsequent experiments. Primers were located in conserved regions of the alignments to enable them to detect the greatest number of strains.

Sequencing primers internal to the DV E genes were designed based on the NCBI RefSeqs (section 2.2.1).

DV species-specific primers for rBV production were designed to amplify only the intended gene target and not the flanking regions. They also incorporated restriction endonuclease recognition sequences to facilitate cloning into the baculovirus transfer vector pBAC-2cp.

2.2.4 Analysis of sequencing data

Sequence chromatograms were viewed and contigs of sequencing reactions constructed using SeqMan (Lasergene 7 sequence analysis software, DNAStar Inc.). Each clone contig consisted of up to eight sequence reads (four primers in each direction) and consensus sequences for each clone were saved as EditSeq files. Multiple sequence alignments were performed for all of the clones from each patient sample. To minimise errors introduced by the sequencing process, any mismatches that were identified from the multiple sequence alignment consensus were compared to the contig chromatogram to ensure only definite nucleotide alterations (confirmed by multiple sequence reads) were counted. EditSeq was used to convert nucleotide sequence into protein sequence and multiple sequence alignments were performed as described in Section 2.2.2.

Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 4 (Tamura *et al.*, 2007), with the neighbour joining method and maximum composite likelihood model. Bootstrapping analysis was conducted using 1000 replicates and where possible the DV type 1 E gene sequence from the NCBI RefSeq NC_001477 was used to root the trees.

2.2.5 Statistical analysis

Statistical tests were applied using Minitab version 15 (Minitab, Coventry, UK).

2.2.6 Protein modelling

DeepView (Swiss-PdbViewer at http://spdbv.vital-it.ch/) was used to model consensus protein sequences onto the appropriate DV E structure and map the locations of the

nonsynonymous mutations found within the quasispecies population. Homology modelling was performed by inputting the sample amino acid sequence and submitting that sequence to a protein data bank (pdb) to find structures with similar sequences. The structure with the best homology match (1UZG for dengue virus type 3 and 1OAN for dengue virus types 2 and 4) was downloaded and the sample amino acid sequence structure was inferred using the downloaded structure as a reference. Transmembrane domains were not included as they are not present in any of the pdb files. To construct the oligomeric protein the sample sequence monomer was opened in two layers so that one of the monomers could be re-positioned as the second monomer of the dimer using the Matrix co-ordinates provided in the pdb text file from the downloaded structure. The two layers were then merged and the dimer coloured according to the protein domains (stated in Rey *et al.*, 1995). Surface accessible residues and residues within 4Å of important structural features were identified.

2.3 Nucleic acid manipulations

2.3.1 Amplification of nucleic acid sequences using the polymerase chain reaction

PCR and RT-PCR amplifications were performed on a GeneAmp 9700 thermocycler (Applied Biosystems). Amplified DNA was visualised using agarose gel electrophoresis (Section 2.5.1).

2.3.1.1 Standard RT-PCR

Standard RT-PCR reactions were performed using the OneStep RT-PCR Kit (Qiagen) according to the manufacturer's one-step instructions for amplification of long RT-PCR products. Reaction volumes of 25µL consisted of final concentrations of 1x reaction buffer (from 5x concentrate of Tris-HCl, KCl, $(NH_4)_2SO_4$, 12.5mM $MgCl_2$, dithiothreitol (DTT); pH 8.7), 1.6mM dNTP mix, 0.6µM each of forward and reverse primer (specific to template RNA), 1µL of enzyme mix (containing Omniscript and Sensiscript reverse transcriptases and HotStar*Taq* DNA Polymerase with 1mM DTT, 0.1mM EDTA, 0.5% Nonidet® P-40, 0.5% Tween® 20, 50% glycerol, stabilizer; pH 9.0) and 5µL of RNA template. Cycling conditions are detailed in Table 2.1.

Table 2.1. Standard RT-PCR thermal cycling conditions

Stage	Temperature (˚C)	Time	Cycles
Reverse transcription	45	30 minutes	
Denaturing	95	15 minutes	
Amplification	94	15 seconds	36 cycles
	60	1 minute	
	68	2.5 minutes	
Extension	68	10 minutes	
Cooling	4	∞	

2.3.1.2 AccuScript high-fidelity RT-PCR

AccuScript high fidelity RT-PCR reactions were performed in duplicate for each sample in 50µL volumes using the AccuScript High-Fidelity RT-PCR System (Stratagene) according to the manufacturer's two-step instructions for synthesising first-strand cDNA and amplifying the cDNA template. First-strand cDNA synthesis reactions (10µL) consisted of a final concentration of 1x AccuScript RT buffer (from 10x concentrate of 0.5M Tris-HCl (pH 8.3), 0.75M KCl, 0.03 M $MgCl_2$), 4µM dNTP mix, 0.625µM reverse

primer (specific to template RNA), 5µL of RNA template, 10mM DTT and 1µL of AccuScript reverse transcriptase (RT). The RT buffer, dNTPs, reverse primer and RNA were incubated at 65˚C for five minutes and then allowed to cool at room temperature for a further five minutes to anneal the primer to the RNA. The DTT and AccuScript RT were then added and the reaction incubated at 42˚C for 30 minutes to allow reverse transcription of the RNA template to cDNA.

Amplification of cDNA was performed in a total reaction volume of 50µL with a final concentration of 1x PCR buffer (formulation not specified by manufacturer), 0.8µM dNTP mix, 0.2µM each of forward and reverse primer (specific to template DNA), 2.5 units of *PfuUltra* high fidelity DNA polymerase and 5µL of cDNA template. Thermal cycling conditions are detailed in Table 2.2.

Table 2.2. AccuScript thermal cycling conditions for PCR amplification of cDNA

Stage	Temperature (˚C)	Time	Cycles
Denaturing	95	1 minute	
Amplification	95	30 seconds	36 cycles
	60	30 seconds	
	68	6 minutes	
Extension	68	10 minutes	
Cooling	4	∞	

2.3.1.3 Expand high-fidelity RT-PCR

Expand high-fidelity RT-PCR reactions were performed using the Expand reverse transcriptase and Expand long template PCR system (Roche) according to the method published in (Craig *et al.*, 2003). First strand synthesis reactions consisted of 20µL total reaction volume with a final concentration of 1x Expand RT buffer (formulation not specified by manufacturer), 4µM dNTP mix, 0.625µM reverse primer (specific to template RNA), 5µL of RNA template, 10mM DTT and 1µL of Expand reverse transcriptase. The reverse primer and RNA were incubated at 65˚C for five minutes and then allowed to cool at room temperature for a further five minutes to anneal the primers to the RNA. The DTT and Expand RT were then added and the reaction incubated at 30˚C for 10 minutes, followed by 42˚C for 30 minutes to allow reverse transcription of the RNA template to cDNA.

Reactions for amplifying the cDNA template consisted of 50µL total reaction volumes with a final concentration of 1x PCR buffer 1 (formulation not specified by manufacturer), 0.8µM dNTP mix, 0.2µM each of forward and reverse primer (specific to template DNA), 3.5 units of Expand high fidelity enzyme mix (containing *Taq* and *Tgo* DNA polymerases) and 5µL of cDNA template. Thermal cycling conditions are detailed in Table 2.3.

Table2.3. Expand thermal cycling conditions for PCR amplification of cDNA

Stage	Temperature (˚C)	Time	Cycles
Denaturing	94	2 minutes	
Amplification	94	10 seconds	36 cycles
	60	30 seconds	
	68	2.5 minutes	
Extension	68	10 minutes	
Cooling	4	∞	

2.3.1.4 Standard PCR

Standard PCR reactions were performed using the *Taq* PCR Core kit (Qiagen) according to the manufacturer's instructions for PCR using *Taq* DNA polymerase. Reaction volumes of 25µL consisted of final concentrations of 1x reaction buffer (from 10x concentrate of Tris-HCl, KCl, $(NH_4)_2SO_4$, 15mM $MgCl_2$; pH 8.7), 0.8mM dNTP mix, 0.3µM each of forward and reverse primer (specific to template RNA), 0.13µL of enzyme mix (20mM Tris-HCl, 100mM KCl, 1mM DTT, 0.1mM EDTA, 0.5% Nonidet P-40, 0.5% Tween 20, 50% glycerol; pH 8.0) and 5µL of RNA template. Cycling conditions were as detailed in Table 2.4.

Table 2.4. *Taq* PCR core kit thermal cycling conditions

Stage	Temperature (˚C)	Time	Cycles
Denaturing	94	3 minutes	
Amplification	94	15 seconds	36 cycles
	60	1 minute	
	68	2.5 minutes	
Extension	68	10 minutes	
Cooling	4	∞	

2.3.2 RNA extraction using commercial spin columns

The RNA extraction methods in this section use lysis under denaturing conditions to inactivate virus and RNases and to ensure isolation of intact RNA. Nucleic acids are then preferentially bound to the silica membrane within a spin column in the presence of a chaotropic salt such as guanidine chloride. Impurities are washed away and RNA is subsequently eluted in a low salt concentration elution buffer.

2.3.2.1 RNA extraction from dengue patient sera

RNA extractions from dengue patient sera were performed using the QIAamp Viral RNA Mini Kit (Qiagen) according to the manufacturer's protocol for purification of viral RNA (Spin protocol). Samples were first lysed for 10 minutes at room temperature using 560µL of AVL buffer containing 5.6µg of carrier RNA. 560µL of 100% ethanol (v/v) was added and the sample was loaded onto the QIAamp Mini spin column and centrifuged (8000 revolutions per minute (rpm) for one minute; Biofuge Pico, Heraeus). Contaminants were then washed away using 500µL of AW1 buffer and centrifugation (8000 rpm for one minute) followed by 500µL of AW2 buffer and centrifugation (13000 rpm for three minutes). RNA was eluted by incubation for one minute in 60µL of RNase-free water and collected by centrifugation (8000 rpm for one minute).

2.3.2.2 RNA extraction from pelleted insect cells

RNA was extracted from pelleted *Sf*9 or *Sf*21 cells using an RNeasy Mini kit (Qiagen) and on-column DNase digestion using an RNase-Free DNase set (Qiagen) according to the manufacturer's microcentrifuge protocol. Samples were lysed with 600µL of RLT buffer containing 1% β-mercaptoethanol (14.3M stock solution) and homogenised by centrifugation through a QIAshredder spin column (2 minutes at 13000 rpm; Biofuge Pico, Heraeus). Ethanol (600µL of 70% v/v) was added and the sample was loaded onto the QIAamp Mini spin column and centrifuged (10000 rpm for 15 seconds). The column was washed with 350µL of RW1 buffer and centrifuged (10000 rpm for 15 seconds) to remove contaminants. DNase I incubation mix (80µL; 3.4 units of DNAse I in RDD buffer) was then added to the spin column and incubated at room temperature for 15 minutes. RW1 buffer (80µL) was added to the spin column and centrifuged

(10000 rpm for 15 seconds), followed by 500µL of RPE buffer and centrifugation (10000 rpm for two minutes). RNA was eluted by incubation for one minute in two 30µL volumes of RNase-free water and collected each time by centrifugation (10000 rpm for one minute).

2.3.2.3 RNA purification and concentration from transcription reactions

RNA was extracted from transcription reactions (section 2.3.10) using an RNeasy MinElute Clean-up kit (Qiagen) according to the manufacturer's protocol for RNA cleanup and concentration. DNase digestion of RNA was performed before RNA clean-up using an RNase-Free DNase Set (Qiagen). DNase I stock solution (6.75 units) and 10µL of RDD buffer were added to each sample and made up to a volume of 100µL with nuclease-free water before incubation at room temperature for 10 minutes. The samples were lysed using 350µL of RLT buffer containing 1% β-mercaptoethanol (14.3M stock solution). Ethanol (500µL of 100% v/v) was added and the sample loaded onto the QIAamp Mini spin column and centrifuged (10000 rpm for 15 seconds; Biofuge Pico, Heraeus). The column was washed with 500µL of RPE buffer and centrifuged (10000 rpm for 15 seconds) to remove contaminants, followed by 500µL of 80% ethanol (v/v) and centrifugation (10000 rpm for two minutes). Residual ethanol was removed by centrifugation (13000 rpm for five minutes with the spin column lids open). RNA was eluted by incubation for one minute in 20µL of RNase-free water and collected by centrifugation (13000 rpm for one minute).

2.3.3 DNA extraction using commercial spin columns

The DNA extraction methods in this section use preferential binding of nucleic acid to the silica membrane within a spin column at pH less than 7.5 in the presence of a chaotropic salt such as guanidine hydrochloride. Impurities are washed away and DNA is subsequently eluted in a low salt concentration pH 7.0 to pH 8.5 elution buffer.

2.3.3.1 Extraction of DNA from agarose gels

Amplified PCR products were visualised using a UV transilluminator (GIBCO BRL, TFX-20M) set at 70% intensity to prevent degradation of DNA. The required band was excised using a sharp sterile scalpel and the DNA extracted using the QIAquick Gel Extraction kit (Qiagen) according to the manufacturer's microcentrifuge protocol. All

centrifugation steps were performed at 13000 rpm for one minute using a Biofuge Pico, Heraeus. To dissolve the agarose, 300µL of QG buffer was added to each extracted band and incubated at 50°C for 10 minutes. The sample was loaded onto the QIAquick column and centrifuged. The DNA from the sample was preferentially bound to the membrane during this centrifugation due to the high salt concentrations in the sample buffer. The column was washed with 500µL of QG buffer and centrifuged to remove contaminants, followed by 750µL of PE buffer and centrifugation. Residual ethanol was removed by centrifugation. DNA was eluted by incubation for one minute in 50µL of RNase-free water and collected by centrifugation.

2.3.3.2 Extraction of DNA from enzymatic reactions

DNA was extracted from restriction digests and PCR reactions using the MinElute PCR purification kit (Qiagen) according to the manufacturer's microcentrifuge protocol. All centrifugation steps were performed at 13000 rpm for one minute using a Biofuge Pico, Heraeus. Five volumes of PB buffer were added to one volume of sample before the sample was loaded onto the QIAquick column and centrifuged. The DNA from the sample was preferentially bound to the membrane during this centrifugation due to high salt concentrations in the sample buffer. The column was washed with 750µL of PE buffer and centrifuged to remove contaminants. Residual ethanol was removed by centrifugation. DNA was eluted by incubation for one minute in 50µL of RNase-free water and collected by centrifugation.

2.3.3.3 Viral DNA extraction from insect cells

Viral DNA was extracted from P2 amplified rBV-infected insect cell culture medium using a HighPure Viral Nucleic Acid Kit (Roche) according to the manufacturer's instructions. Clarified insect cell culture medium (200µL) was incubated for 10 minutes at 72°C with 200µL of carrier RNA-supplemented binding buffer and 50µL of proteinase K solution. Binding buffer (100µL) was then added before applying the mixture to the HighPure filter tube and centrifugation (10000 rpm for one minute; Biofuge Pico, Heraeus). Inhibitor removal buffer (500µL) was applied to the filter tube and centrifuged (10000 rpm for one minute), followed by two 450µL washes with wash buffer and further centrifugation (10000 rpm for one minute). Residual wash buffer

was removed by centrifugation (13000 rpm for 10 seconds). Nucleic acids were eluted in 50µL of elution buffer by centrifugation (10 minutes at 10000 rpm).

2.3.4 Plasmid DNA extraction using commercial filter kits

2.3.4.1 Plasmid DNA extraction for quasispecies analysis

Deep 96-well culture plates containing *E. coli* overnight cultures (section 2.3.5.4) were centrifuged to pellet the bacterial cells (1500 x*g* for seven minutes; Legend RT Plus, Sorvall). The cell culture medium was discarded and plasmid DNA extracted from the cells using the Montage Plasmid Miniprep 96 kit (Millipore) according to manufacturer's instructions. Cell pellets were resuspended in 100µl of Solution 1 and lysed in 100µl of Solution 2 with vigorous mixing. The lysate was neutralised with 100µL of Solution 3 and then drawn through the clearing plate into the plasmid plate by vacuum (8 millimetres of mercury (mmHg); using a Qiagen vacuum pump and manifold). Plasmid DNA was adsorbed to the plasmid plate membrane when the clarified lysate was drawn through the plasmid plate by vacuum (24 mmHg). Plasmid DNA was washed with 200µl of Solution 4 and a vacuum applied (24 mmHg). Plasmid DNA was recovered by adding 85µl of Solution 5 (tris-EDTA) and incubating at room temperature for 30 minutes before removal to a 96-well storage plate.

2.3.4.2 Plasmid DNA extraction for recombinant baculovirus production

Plasmid DNA was extracted from a 200mL culture of *E. coli* transformed with pBAC-2cp (containing DV type 1 prM and patient sample quasispecies E genes; section 2.3.6.2) using a QIAfilter plasmid Maxi kit (Qiagen) according to manufacturer's instructions. Briefly, the bacterial pellets were re-suspended in 10mL P1 buffer before alkaline lysis with 10mL P2 buffer for four minutes at room temperature. P3 buffer (10mL) was added to the lysate before pouring the lysate into the QIAfilter cartridge and incubating at room temperature for 10 minutes. The cell lysate was filtered into the QIAGEN-tips (previously equilibrated with 10mL of buffer QBT; containing an anion-exchange resin to selectively bind supercoiled plasmid DNA at low salt concentrations) and allowed to empty by gravity flow. The plasmid DNA was washed twice with 30mL QC buffer to remove contaminants. DNA was eluted at high salt concentration with 15mL QF buffer into 50mL polypropylene falcon tubes. DNA was precipitated by

adding 10.5mL (0.7 volumes) isopropanol to the eluted DNA and centrifuging (5000 x*g* for 60 minutes at 4°C; Legend RT Plus, Sorvall). Supernatants were carefully decanted and the DNA pellets washed with 5mL of 70% ethanol, before centrifugation (5000 xg for 60 minutes at 4°C) and careful removal of the supernatant. Pellets were air-dried for 5–10 minutes and re-dissolved in 200µL of EB buffer (10mM Tris-HCl, pH 8.5).

2.3.5 Cloning of dengue virus quasispecies envelope genes, transformations and amplification of positive clones for quasispecies analysis

Cloning and transformations for DV quasispecies E gene analysis were performed using the pGEM®-T Easy Vector System *II* (Promega). Confirmation that cloning was successful was achieved by a combination of plasmid DNA extraction (Sections 2.3.4.1 and 2.3.4.2), restriction digest analysis (Section 2.3.8), RT-PCR or PCR (Section 2.3.1), sequencing (Section 2.4) and agarose gel electrophoresis (Section 2.5.1).

2.3.5.1 Poly(A)-tailing of blunt-ended dengue virus envelope gene PCR amplicons

Blunt-ended DNA fragments were generated because of the use of a proofreading DNA polymerase for high-fidelity PCR (sections 2.3.1.2 and 2.3.1.3). These were A-tailed prior to cloning to allow them to ligate into the cloning vector (pGEM®-T Easy). Blunt-ended gel-extracted PCR products were A-tailed using five units of *Taq* DNA polymerase and 1x reaction buffer with $MgCl_2$ (*Taq* PCR Core kit; Qiagen) and dATP (dNTP set, PCR grade; Qiagen) to a final concentration of 0.2mM. A 10µL total reaction volume was made up by the addition of nuclease-free water (Promega). Reactions were incubated at 70°C for 30 minutes.

2.3.5.2 Ligation of poly(A)-tailed dengue virus envelope gene PCR amplicons into the cloning vector pGEM®-T Easy

Ligation reactions consisted of Rapid Ligation Buffer at 1x final concentration, 50ng pGEM®-T Easy Vector, 2µl A-tailed PCR product (section 2.3.5.1), and three units of T4 DNA Ligase. Ligations were incubated at 4°C overnight to maximise the number of transformants.

2.3.5.3 Transformation of E. coli with the cloning vector pGEM®-T Easy containing dengue virus quasispecies envelope gene amplicons

For transformation of *E. coli* (with the cloning vector pGEM®-T Easy containing DV E gene amplicons), 2µl of ligation reaction (section 2.3.5.2) was added to 50µl of JM109 cells on ice and gently mixed before incubation on ice for 20 minutes. The cells were heat-shocked at 42˚C for 45 seconds then placed on ice for two minutes. Room temperature SOC medium (950µl; Invitrogen™) was added and the reaction incubated (37°C at 220 rpm for 1 hour 30 minutes; Minitron shaking incubator, Infors) for 1 hour 30 minutes. Cultures (100µl) were incubated on agar plates (37°C overnight). To allow blue white screening of recombinants, culture plates were made using 200mL L-agar (Biomerieux), 100µg/mL ampicillin (Sigma, 100mg/mL), 80µg/mL X-gal (Promega, 50mg/mL) and 0.5mM IPTG (Novagen, 100mM).

2.3.5.4 Amplification of positive transformants for quasispecies analysis

Transformed (white) colonies (section 2.3.5.3) were picked and inoculated into 96-well deep-well culture plates (Montage Plasmid Miniprep 96 kit; see Section 2.3.3.1) containing 1.5mL L-broth (Biomerieux) with 100µg/mL ampicillin (Sigma). Cultures were then incubated (37°C at 220 rpm for 24 hours; Minitron shaking incubator, Infors). For each patient sample, one 96-well plate consisting of an equal number of colonies from each of the duplicate high-fidelity RT-PCR reactions (section 2.3.1.2) was cultured.

2.3.6 Cloning of dengue virus quasispecies envelope genes, transformations and amplification of positive clones for recombinant baculovirus production

DV quasispecies E genes were amplified using primers containing specific restriction endonuclease sites (section 2.2.3). Purified DV E gene PCR products (Section 2.3.3.1) and pBAC-2cp baculovirus transfer vector (Novagen, Merck) underwent restriction digestion (Section 2.3.8) to generate the appropriate termini for ligation.

2.3.6.1 Ligation of dengue virus quasispecies envelope gene PCR amplicons into the baculovirus transfer vector pBAC-2cp

Ligation reactions were incubated at 4˚C overnight and consisted of a 10µL reaction containing 0.1-1 unit of T4 ligase (Promega), T4 ligase buffer at 1x concentration (from 10x stock of 300mM Tris-HCl; pH 7.8), 100mM MgCl2, 100mM DTT and 10mM ATP;

supplied with the enzyme), 2µL of DV prM gene *Nhe*I/*Eag*I digest product, 4µL of DV E gene *EagI*/*XmaI* digest product and 1µL of pBAC-2cp *NheI*/*XmaI* digest product. Transformations were performed as stated in section 2.3.5.3.

2.3.6.2 Amplification of positive transformants for recombinant baculovirus production

Small-scale culture was performed as stated in Section 2.3.5.4. Larger scale culture was performed once sequencing (section 2.4) had confirmed the presence and authenticity of the DV type 1 prM and quasispecies E gene inserts. A single transformed colony for each sample was used to inoculate a starter culture of 5mL LB medium containing 100µg/mL ampicillin, and incubated (37˚C at 200 rpm for eight hours; Minitron shaking incubator, Infors). The starter culture was diluted 1/500 into 200mL LB medium containing 100µg/mL ampicillin and incubated (37°C for 16 h at 200 rpm) in a 1L conical flask. Bacterial cells were harvested by centrifugation (4100rpm for 15 minutes at 4°C; Legend RT Plus, Sorvall).

2.3.7 Storage of transformed *E. coli*

Glycerol stocks were made for all clones. To 500µL of each transformed culture (section 2.3.5.4 and 2.3.6.2), 500µL of 30% glycerol (Sigma) was added and the bacterial suspension, mixed, and stored at -80°C.

2.3.8 Restriction digestion of plasmid DNA

Digest reactions were incubated (37°C for one to three hours; Minitron shaking incubator, Infors) and consisted of 5-10 units of enzyme, the appropriate buffer as recommended by the enzyme manufacturer at 1x final concentration, 100µg/mL BSA and 0.1-0.3µg of plasmid DNA. Confirmation that DNA was digested appropriately was achieved by DNA agarose gel electrophoresis (Section 2.5.1).

2.3.9 Quantification of nucleic acids

Plasmid DNA was quantified using a ND1000 spectrophotometer (Nanodrop, Thermo Scientific).

2.3.10 In vitro transcription of plasmid DNA

Linearised (section 2.3.8) DV type 3 E gene plasmid DNA (from section 2.3.5) was T7 transcribed *in vitro* using a MEGAScript T7 High Yield Transcription kit (Ambion) according to manufacturer's instructions. The 20µL *in vitro* transcription reaction consisted of 7.5mM each of ATP, CTP, GTP, and UTP; 10x buffer at 1x final concentration and 2µL of T7 enzyme solution with between 0.1 and 0.3µg of plasmid DNA. The reaction was incubated at 37˚C for two hours before the addition of two units of Turbo DNase and incubation for a further 15 minutes. RNA transcript size and purity was assessed using an RNA denaturing gel (section 2.5.2), RT-PCR and PCR (section 2.3.1). The transcription reaction was also performed in the presence or absence of 2µg of thioredoxin (Sigma), which is an accessory protein that improves the processivity and fidelity of T7 polymerase (Kunkel, Patel, and Johnson, 1994).

2.4 Sequencing

2.4.1 Sequencing PCR

Sequencing PCR reactions were performed using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems). pGEM®-T Easy vector M13 forward (5'-GTTTTCCCAGTCACGAC-3') and reverse (5'-CAGGAAACAGCTATGAC-3') primers, and DV sequence specific primers (Table 3.6) were used for sequencing. BigDye PCR reaction mix consisted of 0.5µl of 2.5x Ready Reaction Premix, 3.75µl of 5x sequencing buffer, 3.2pmol of primer, 5µl of plasmid DNA (or gel-extracted high-fidelity RT-PCR product when direct sequencing) and nuclease-free water to a final reaction volume of 20µl. Thermal cycling conditions are detailed in Table 2.5.

Table 2.5. Sequencing PCR thermal cycling conditions

Stage	Temperature (˚C)	Time	Cycles
Denaturing	96	1 minute	
Amplification	96	10 seconds	25 cycles
	50	5 seconds	
	60	4 minutes	
Cooling	4	∞	

2.4.2 Sequencing PCR clean-up

Clean-up of sequencing PCR reactions was performed using the Montage SEQ_{96} Sequencing Reaction Cleanup kit (Millipore). Briefly, BigDye PCR products were diluted with 20µl of injection solution and then added to a SEQ_{96} plate. The SEQ_{96} plate was placed on top of a vacuum manifold and vacuum applied until the wells were empty. Excess liquid was blotted from the bottom of the SEQ_{96} plate before returning it to the vacuum manifold, adding a further 20µl of injection solution to each well and applying the vacuum as before. 20µl of injection solution was used to re-suspend the cleaned-up sequencing reaction, which was then transferred to a sequencing plate (Greiner Bio One).

2.4.3 Sequencing

Sequencing was performed using an ABI PRISM 3700 DNA Analyzer (Applied Biosystems), and results were analysed as stated in Section 2.2.4

2.5 Electrophoretic techniques

DNA and RNA electrophoresis was carried out using GNA100 gel tanks (Pharmacia Biotech). Bands were visualised under UV light using a Gel Doc™ XR Gel Documentation System (Bio-Rad).

2.5.1 DNA electrophoresis

Agarose gels consisted of 1% agarose (Sigma) dissolved in 1x Tris-Borate-EDTA (TBE) running buffer (Sigma) with ethidium bromide solution (10mg/mL) at a final concentration of 9ng/mL (Sigma). DNA ladders used were either 1kb (NEB) or 100bp (GelPilot 100bp Plus Ladder, Qiagen). Samples were mixed with loading dye (GelPilot DNA Loading Dye, Qiagen) before addition to the wells of the gel. Gels were run in 1x TBE running buffer, at 110 volts for 25-60 minutes.

2.5.2 RNA electrophoresis

RNA-denaturing gels consisted of 1.5% agarose-LE (Ambion) gel containing 1x NorthernMax denaturing gel buffer (Ambion). 2µL of transcript RNA was added to 6µL of formaldehyde loading dye (Ambion) and 25ng ethidium bromide (Sigma) and heated at 75˚C for 15 minutes. Samples were run for 90 minutes at 35V in NorthernMax 1x denaturing gel running buffer (Ambion).

2.5.3 Polyacrylamide gel electrophoresis

Polyacrylamide gel electrophoresis (PAGE) was carried out using XCell *SureLock*™ Mini-Cell gel tanks (Invitrogen™).

2.5.3.1 Denaturing and reducing PAGE

Samples (9.75µL) were added to NuPAGE 4x LDS Sample buffer and NuPAGE 10x Reducing Agent (Invitrogen™) to give a 1x final concentration in 15µL and incubated at 70˚C for 10 minutes. Samples were run on NuPAGE 4-12% Bis-Tris gels in 1x NuPAGE MOPS SDS running buffer with 500µL NuPAGE antioxidant (Invitrogen™) in the cathode buffer chamber at 200 volts for 55 minutes. Gels were stained using 20mL SimplyBlue safe stain (Invitrogen™).

2.5.3.2 Native PAGE

Samples were incubated in 1% triton X-100 (Sigma) for 1 hour at room temperature. Samples (9.75µL) were then added to NativePAGE 4x Sample buffer (Invitrogen™) and water (molecular biology grade; Promega) to give a 1x final concentration in 15µL. Samples were run on NativePAGE 3-12% Bis-Tris gels (Invitrogen™) with Light Blue Cathode buffer (1x NativePAGE running buffer with 1x NativePAGE Cathode buffer additive; Invitrogen™) in the inner chamber of the gel tank and anode buffer (1x NativePAGE running buffer) in the outer chamber of the gel tank. Gels were run at 150 volts for two hours. Gels were fixed in 100mL of fixing solution (40% methanol and 10% acetic acid) and microwaved for 45 seconds before incubating (room temperature and 22 rpm for 15 minutes). Gels were de-stained in 100mL of 8% acetic acid and microwaved for 45 seconds before incubating (room temperature and 22 rpm) until the desired background was obtained.

2.6 Western blots and immunodetection

A plate rocker (Mini gyro-rocker SSM3, Stuart) was used during the wash and block steps. SDS-PAGE gels were washed in 1x Tris-Glycine Transfer buffer (Invitrogen™) with 20% methanol for 15 minutes. Native-PAGE gels were washed in 1x Tris-Glycine Transfer buffer (no methanol) for five minutes. 0.2µm polyvinylidene fluoride (PVDF) transfer membranes (Invitrogen™) were immersed in 100% methanol for 15 seconds and then washed in the appropriate transfer buffer. Two sponge pads and one filter paper sandwich (per gel) were immersed in transfer buffer. Transfer apparatus was assembled according to manufacturer's instructions (Mini Trans-Blot Electrophoretic Transfer Cell; Biorad) and wet transfer performed at 100 volts (constant) for two hours with ice packs and constant stirring. After Native PAGE transfer, the PVDF membrane was fixed in 8% acetic acid for 15 minutes then rinsed with deionised water.

Blotted PVDF membranes were blocked in block solution (phosphate buffered saline (PBS) with 0.05% Tween 20 and 5% skimmed milk) for one hour with agitation then stored overnight at 4˚C in block solution. For immunodetection using all antibodies, membranes were incubated with primary antibody diluted in block solution for three hours with gentle agitation. Membranes were washed three times for five minutes in PBS with 0.05% Tween 20 before incubation with HRP-conjugated secondary antibody diluted in block solution for two hours with gentle agitation. Antibodies and concentrations used (based on manufacturer's recommendations) are detailed in Table 2.6. Membranes were again washed three times for five minutes in PBS with 0.05% Tween 20 before detection using ECL Plus™ (Amersham Biosciences) according to manufacturer's instructions. Western blots were imaged using a BioRad ChemiDoc™ XRS Gel Documentation system.

Table 2.6. Antibodies and dilutions used in immunodetection from Western blots

Primary antibody	Dilution	Secondary antibody	Dilution
Mouse monoclonal antibody to polyHistidine [His-1] (Sigma, H1029)	1:3000	Rabbit polyclonal antibody to mouse IgG H&L, HRP conjugated (Abcam, Ab6728)	1:1000
Mouse monoclonal antibody to Flavivirus E [FE1](Abcam, Ab64059; Pierce, Thermo Fisher, MA1-71258)	1:5	Rabbit polyclonal antibody to mouse IgG H&L, HRP conjugated (Abcam, Ab6728)	1:1000
Rabbit polyclonal antibody to DV types 1, 2, 3 and 4 (Abcam, Ab9202)	1:1000	Goat polyclonal antibody to rabbit IgG H&L, HRP conjugated (Abcam, Ab6721)	1:3000
Mouse monoclonal antibody to baculovirus envelope glycoprotein GP64 [AcV5] (eBioscience, 14-6995-82)	1:500	Rabbit polyclonal antibody to mouse IgG H&L, HRP conjugated (Abcam, Ab6728)	1:1000
Mouse monoclonal antibody to DV types 1, 2, 3 and 4 [D1-11(3)] (Abcam, Ab9200)	1:1000	Rabbit polyclonal antibody to mouse IgG H&L, HRP conjugated (Abcam, Ab6728)	1:1000

2.7 Baculovirus manipulations

2.7.1 Insect cell culture

*Sf*9, *Sf*9 E-T or *Sf*21 cells were grown in suspension culture (27˚C at 150 rpm, HT-Infors Labotron orbital shaker) in the appropriate serum-free media as detailed in section 2.1.6. Cell density was determined prior to use with a Neubauer haemocytometer. Cell viability was determined by trypan blue staining using an equal volume of cells to 0.4% (w/v) trypan blue (Sigma). Cells were maintained at a cell density of between $0.5x10^6$ and $6x10^6$ cells/mL and passage number post recovery of cells from frozen stocks did not exceed 30.

For long term storage, $2x10^7$ cells at low passage number were frozen slowly in liquid nitrogen in the appropriate cell culture medium containing 10% dimethyl sulfoxide (DMSO) and 30% fetal bovine serum (FBS). During recovery of cells from frozen stocks cells were rapidly thawed and placed in culture initiation medium containing equal proportions of conditioned and fresh culture medium appropriate to the cell line. Initiation cultures were incubated in suspension (27˚C at 150 rpm, HT-Infors Labotron orbital shaker).

2.7.2 Production of recombinant baculoviruses

Tissue culture dishes (35mm; Corning, Sigma) were seeded with $1x10^6$ *Sf*9 cells. Cells were allowed to settle (for at least 30 minutes at room temperature) and form an even mono-layer. Dishes were observed under an inverted microscope (ID03, Zeiss) to confirm seeding density. To produce rBV, co-transfection reactions containing 5µg of Lipofectin™ (Invitrogen™), 985µL of Ex-cell 420 SFM with L-glutamine, 500ng of transfer vector (section 2.3.6) and 100ng of *Flash*BAC GOLD™ (Oxford Expression Technologies; OET) were set up. Reactions were incubated at room temperature for 40 minutes. Culture medium (2mL) was removed from the seeded 35mm tissue culture dishes and the co-transfection mix gently added. Co-transfections were placed in a sandwich box with a moist tissue to humidify the atmosphere, and incubated at 27°C for 24 hours (MIR-253 incubator, Sanyo). 1mL pre-warmed fresh Ex-cell 420 SFM with L-glutamine was added before re-incubation at 27°C for a further four days. Insect cells were observed under an inverted microscope (ID03, Zeiss) to confirm infection and

passage 0 (P0) recombinant virus stocks (clarified cell culture medium) were harvested by centrifugation (3000 rpm for 15 minutes at 4˚C, Sorvall Megafuge 11R centrifuge) and stored at 4°C.

2.7.3 β-galactosidase assay

After rBV P0 stocks were harvested from the tissue culture dishes (section 2.7.2), 1mL pre-warmed Ex-cell 420 SFM with L-glutamine and 0.3mg X-gal (Sigma) were added to rBV.*lacZ*-infected cells. After re-incubation at 27°C for 24 hours (MIR-253 incubator, Sanyo), the culture was analysed for the presence of a blue precipitate.

2.7.4 Amplification and titration of recombinant baculovirus stocks

2.7.4.1 P1 amplification of P0 recombinant baculoviruses

For each virus, a 250mL Erlenmeyer flask (Corning, Sigma) was seeded with $2x10^6$ *Sf*9 cells/mL in pre-warmed Ex-cell 420 SFM with L-glutamine. P0 recombinant virus stock (1mL; section 2.7.2) was added to give a final culture volume of 50mL. Cultures were incubated in the dark for five days (27˚C at 150 rpm, HT-Infors Labotron orbital shaker, MIR-253 incubator, Sanyo). P1 recombinant virus stock (clarified medium) was harvested by centrifugation (3000 rpm for 15 minutes at 4˚C, Sorvall Megafuge 11R centrifuge).

2.7.4.2 Titration of recombinant baculoviruses

Recombinant baculoviruses were titred by end-point dilution using *Sf*9 E-T cells which produce enhanced green fluorescent protein (eGFP) when infected by baculoviruses (Hopkins and Esposito, 2009). Starting at a dilution of 10^{-3}, the rBVs were serially diluted eight times in HyClone SFX medium (Invitrogen™), in replicates of eight, to a final dilution of 10^{-11}. *Sf*9 E-T cells (7.5×10^4 in 100µL HyClone SFX medium) were added to each well. A negative control was also set up in the same way. Plates were incubated at 27°C without shaking (MIR-253 incubator, Sanyo). At three days post-infection, the plate was scored using a fluorescent microscope (Eclipse TE2000-S Progres C5, Nikon) by monitoring wells containing cells producing eGFP. Wells were scored as positive if they contained foci of cells producing eGFP and negative if they contained either no cells producing eGFP, or only isolated single cells producing eGFP (indicative of replication deficient rBV). All uninfected control wells were negative for

eGFP. Titres were calculated according to the Reed-Muench method (Dulbecco and Ginsberg, 1988) shown in Figure 2.1.

	-3	-4	-5	-6	-7	-8
A	+	+	-	-	-	-
B	+	+	+	+	+	-
C	+	+	+	+	-	-
D	+	+	+	+	-	-
E	+	+	+	+	-	-
F	+	+	+	+	-	-
G	+	+	+	-	-	-
H	+	+	+	+	-	-
Cumulative positive	30	22	14	7	1	0
Cumulative negative	0	0	1	3	10	18
Cumulative % positive	100	100	93	70	9	0

Dilutions were 10 fold (1/10) so log dilution (h) is -1. Log dilution >50% is -6. Log dilution <50% is -7.

Proportionate distance (PD) =

$$h \times \frac{\text{cumulative \% positive at dilution next above 50\% - 50\%}}{\text{cumulative \% positive at dilution next above 50\% - cumulative \% positive at dilution below 50\%}}$$

$$h \times \frac{70-50}{70-9} = -1 \times \frac{20}{61} = -1 \times 0.328 = -0.328$$

$$\text{Log TCID}^{50} = \text{Log dilution} > 50\% + \text{PD} = -6 + -0.328 = -6.328$$

$$\text{TCID}^{50} = 10^{-6.328} \text{ or } 2.13\times10^{6} \qquad (2.13\times10^{6} \text{ is } 1/4.7\times10^{7})$$

$$\text{Pfu} = 0.7 \times 2.13\times10^{6} = 1.49\times10^{6}$$

$$\text{Pfu/ml} = \frac{\text{Pfu}}{0.1\text{ml}} = \frac{1.49\times10^{6}}{0.1} = 1.49\times10^{7}$$

Figure 2.1. Example TCID^{50} calculation

This figure shows the wells scored as positive on the 96-well plate and the resulting calculations to obtain virus titre in Pfu/mL

2.7.4.3 P2 amplification of P1 recombinant baculoviruses

For each virus, a 500mL Erlenmeyer flask was seeded with 2x10^6 *Sf*9 cells/mL in pre-warmed Ex-cell 420 SFM with L-glutamine. P1 recombinant virus stock was added at an MOI of 0.1 to give a final culture volume of 160mL. Cultures were incubated in the dark for five days (27˚C at 150 rpm, HT-Infors Labotron orbital shaker, MIR-253 incubator, Sanyo). P2 recombinant virus stock (clarified medium) was harvested by centrifugation (3000 rpm for 15 minutes at 4˚C, Sorvall Megafuge 11R centrifuge).

Amplified rBVs were titred (section 2.7.4.2) after each amplification and used for dengue virus-like particle (VLP) production if the titre was at least 10^8 Pfu/mL.

2.7.5 Recombinant dengue virus-like particle production

For optimization of recombinant dengue VLP production for each rBV, 250mL Erlenmeyer flasks were seeded with 2x10^6 *Sf*9 cells/mL in pre-warmed Ex-cell 420 SFM with L-glutamine. P2 recombinant virus stock was added at MOIs of 2, 5 and 10 to give a final culture volume of 40mL. Cultures were incubated in the dark for three days (27˚C at 150 rpm, HT-Infors Labotron orbital shaker, MIR-253 incubator, Sanyo). At 48 and 72 hours post infection, cells were monitored for signs of infection using an inverted microscope (ID03, Zeiss) and clarified medium and pelleted cells were harvested from 1mL of culture by centrifugation (3000 rpm for 15 minutes at 4˚C; Hettich Mikro 22R).

Cell pellets from 1mL cultures were re-suspended in Buffer A (50mM tris, 0.5M NaCl, 8mM imidazole) with 1x protease inhibitor (Calbiochem, Merck) in a total volume of 333µL. Samples were then lysed by the addition of 33.3µL Insect PopCulture (Novagen, Merck) and 10 units/mL of benzonase (Novagen, Merck) for 15 minutes at room temperature with occasional vigorous vortexing. A 150µL sample of total supernatant (T) was removed from each sample before centrifugation (13,000 rpm for 20 minutes at 4˚C, Hettich Mikro 22R) to pellet insoluble material. A 150µL sample of the soluble fraction (S) was removed from each sample without disturbing the pellets, which were discarded.

Clarified insect cell culture medium from 1mL cultures (M_{neat} or M_n) were concentrated from 500µL to 50µL (M_{conc} or M_c) using Ultrafree-0.5 10 kDa centrifugal filter devices (Millipore).

For purification of recombinant dengue VLPs 250mL or 1L Erlenmeyer flasks were seeded with $2x10^6$ *Sf*9 cells/mL in pre-warmed Ex-cell 420 SFM with L-glutamine. P2 rBV stock was added at a MOI of 5 to give a final culture volume of 40mL or 200mL for larger scale culture. Cultures were incubated in the dark for two days (27˚C at 150 rpm, HT-Infors Labotron orbital shaker, MIR-253 incubator, Sanyo). Cells were monitored for signs of infection using an inverted microscope (ID03, Zeiss). Clarified medium was harvested from culture by centrifugation (3000 rpm for 15 minutes at 4˚C, Sorvall Megafuge 11R centrifuge).

Recombinant dengue VLP production was analysed using SDS-PAGE and immunodetection from Western blots as described previously (sections 2.5.3 and 2.6).

2.8 Dengue virus-like particle concentration, purification and quantification

2.8.1 Ultracentrifugation of recombinant dengue virus-like particles

Recombinant dengue VLPs (section 2.7.5) were concentrated through a 20% sucrose cushion using ultracentrifugation. Sucrose (6.3mL of 20% sucrose in PBS, Sigma) was added to 25x89mm Ultra-Clear open-top centrifuge tubes (Beckman). Clarified medium (approximately 32mL) was gently layered on top of the sucrose cushion to the top of the tube. Samples were centrifuged using a Beckman L7-55 Ultracentrifuge and SW32 Ti rotor at 32,000 rpm for two hours. Pellets were re-suspended in 2mL PBS.

2.8.2 His GraviTrap purification

His GraviTrap affinity columns (GE Healthcare) were equilibrated with 10mL of binding buffer (for native purification 20mM sodium phosphate, 500mM NaCl and 20mM imidazole, pH 7.4; for denaturing purification 20mM Tris-HCl, 8M urea, 500mM NaCl, 5mM imidazole and 1mM β-mercaptoethanol, pH 8.0). Recombinant dengue VLPs (concentrated by ultracentrifugation, section 2.8.1) were diluted 1:8 using the appropriate binding buffer and added to the columns. Columns were washed with 10mL of the appropriate binding buffer. Elution buffer (3mL; for native purification 20mM sodium phosphate, 500mM NaCl and 500mM imidazole, pH 7.4; for denaturing purification 20mM Tris-HCl, 8M urea, 500mM NaCl, 500mM imidazole and 1mM β-mercaptoethanol, pH 8.0) was applied to the columns and the eluate collected. For the denaturing purification the elution step was performed twice.

2.8.3 Protein G purification

2.8.3.1 Crosslink protocol using denaturing or native buffers

Recombinant dengue VLPs (concentrated by ultracentrifugation, section 2.8.1) were purified via immunoprecipitation using a mouse monoclonal antibody to flavivirus envelope glycoprotein (Pierce, Thermo Fisher; Table 2.6), Protein G mag sepharose and Protein A/G buffer kit (GE Healthcare). The bead storage solution was removed from 25µL of magnetic beady slurry (containing 5µL of magnetic beads) and the beads equilibrated with 500µL of binding buffer (50mM tris, 150mM NaCl, pH 7.5 (TBS), before removal of the liquid. Antibody solution (diluted 1:5 in 200µL binding buffer)

was added to the beads and incubated with slow end-over-end mixing for one hour before removal of the liquid. Beads were washed with 500µL of binding buffer and the liquid removed. Crosslink solution A (500µL of 200mM triethanolamine pH 8.9) was added to the beads, and the liquid removed. Crosslink solution A with dimethyl pimelimidate dihydrochloride (DMP; 500µL of 200mM triethanolamine pH 8.9 with 50mM DMP) was added to the beads and incubated with slow end-over-end mixing for one hour before removal of the liquid. Beads were washed with 500µL of crosslink solution A and the liquid removed. Beads were blocked with 500µL of crosslink solution B (1 M ethanolamine, pH 8.9) and incubated with slow end-over-end mixing for 15 minutes before removal of the liquid. Non-bound antibody was removed using 500µL of elution buffer (0.1M glycine-HCl with 2M urea, pH 2.9 for the denaturing protocol; 0.1M glycine-HCl, pH 2.9 for the native protocol) before removal of the liquid and two washes with 500µL of binding buffer. Sample was diluted in binding buffer, added to the beads and incubated with slow end-over-end mixing for one hour before removal and collection of the non-bound fraction. The beads were washed three times with 500µL of wash buffer (TBS with 2M urea, pH 7.5 for the denaturing protocol; TBS with 1% octylglucoside, pH 7.5 for the native protocol). The liquid was removed and collected after each wash. During the third wash the magnetic bead solution was transferred to a fresh tube to prevent potential elution of proteins bound non-specifically to the plastic of the tube. Proteins were eluted twice for five minutes in 50µL of the appropriate elution buffer for the native or denaturing protocol. Eluted proteins were pH neutralized using 5µL Tris pH 8.5.

2.8.3.2 Classic protocol

Baculovirus virions in the recombinant dengue VLP samples (concentrated by ultracentrifugation, section 2.8.1) were purified via immunoprecipitation using mouse monoclonal baculovirus envelope glycoprotein GP64 mAb (eBioscience, Table 2.6), Protein G mag sepharose and Protein A/G buffer kit (GE Healthcare). The bead storage solution was removed from 25µL of magnetic beady slurry (containing 5µL of magnetic beads) and the beads equilibrated with 500µL of binding buffer (50mM tris, 150mM NaCl, pH 7.5; TBS), before removal of the liquid. Antibody solution (diluted 1:5 in

200µL binding buffer) was added to the beads and incubated with slow end-over-end mixing for 15 minutes before removal of the liquid. Beads were washed with 500µL of binding buffer and the liquid removed. Sample was diluted in binding buffer, added to the beads and incubated with slow end-over-end mixing for one hour before removal and collection of the non-bound fraction (containing the dengue VLPs). The beads were washed three times with 500µL of wash buffer (TBS with 1% octylglucoside, pH 7.5). The liquid was removed and collected after each wash. During the third wash the magnetic bead solution was transferred to a fresh tube to prevent potential elution of proteins bound non-specifically to the plastic of the tube. Baculovirus virions were eluted twice for five minutes in 50µL of elution buffer (0.1M glycine-HCl, pH 2.9). Eluted proteins were pH neutralized using 5µL Tris pH 8.5.

2.8.4 Anion exchange column purification

Recombinant DV proteins (concentrated by ultracentrifugation, section 2.8.1) were purified via ion exchange chromatography using HiTrap CaptoQ 1mL high capacity strong anion exchange columns (GE Healthcare). Initial purifications were performed manually using syringes and later purifications were performed using an ÄKTA*FPLC* automated system. The sample (1.5-2mL) was first dialysed into 20mM Tris-HCl at pH 8.5 using a 0.5-3.0mL 10 kDa slide-a-lyzer dialysis cassette (Pierce, Thermo Scientific). CaptoQ columns were washed at a flow rate between 0.1 and 1mL/minute with five column volumes each of distilled water, start buffer (20mM Tris-HCl at pH 8.5), 100% elution buffer (20mM Tris-HCl with 1M NaCl at pH 8.5) and then start buffer again before addition of sample. The sample was bound to the column and washed with five column volumes of start buffer before gradient elution using a stepped (manual purification) or linear (automated purification) ionic strength gradient of NaCl (5 column volumes each of elution buffer at 10%, 20%, 30%, 40%, and 50% (v/v diluted in start buffer). Any remaining ionically bound material was eluted using five column volumes of 100% elution buffer. The column was re-equilibrated using five column volumes of start buffer and the flow-through from each stage was collected and analysed by SDS-PAGE and immunodetection from Western blots (sections 2.5.3 and

2.6). Columns were prepared for storage at room temperature using five column volumes each of distilled water followed by 20% ethanol (v/v).

2.8.5 Quantification of total protein by Bradford assay

Total protein content of recombinant DV protein samples (sections 2.7.5, 2.8.1, 2.8.2 and 2.8.4) was quantified using the Coomassie Plus (Bradford) assay kit (Pierce, Thermo Fisher) according to manufacturer's instructions. Briefly, 10µL of sample or BSA protein standard was added to 300µL of Coomassie Plus reagent, and incubated at room temperature for 10 minutes. Sample optical densities were read at 595nm using a ND1000 spectrophotometer (Nanodrop, Thermo Scientific).

2.9 Enzyme-linked immunosorbent assays (ELISAs)

All ELISAs were performed using MaxiSorp® flat-bottom 96 well plates (Nunc) and read using a Multiskan EX microplate photometer (Thermo Scientific). Cut-off points or thresholds, above which the samples were deemed positive, were calculated using the average background absorbance plus three standard deviations, as this accounts for 99.7% of normally distributed data (Swinscow 1996).

2.9.1 Dengue virus-like particle ELISAs

ELISA plate wells were coated with antigen at 1µg/mL in 0.05M carbonate:bicarbonate buffer, using 100µL per well, and stored at 4°C overnight. Wells were washed three times using PBS with 0.05% Tween 20 and blocked for 1 hour 30 minutes using blocking buffer (300µL per well; PBS with 0.05% Tween 20 and 5% skimmed milk unless otherwise stated). Wells were incubated at 37°C with the appropriate primary antibody (100µL per well; Table 2.7) for two hours before washing three times using PBS with 0.05% Tween 20. Wells were incubated at 37°C with the appropriate secondary antibody (100µL per well; Table 2.7) for two hours before washing three times using PBS with 0.05% Tween 20. Equal volumes of ABTS™ HRP substrate solutions A and B (KPL) were mixed together and added to the wells (100µL per well), and incubated at 37°C in the dark for one hour. ABTS™ stop solution was added (100µL per well), and the plate read immediately.

Table 2.7. Antibodies and dilutions used in ELISAs

Primary antibody	Dilution	Secondary antibody	Dilution
Rabbit polyclonal antibody to DV types 1, 2, 3 and 4 (Abcam, Ab9202)	1:1000	Goat polyclonal antibody to rabbit IgG H&L, HRP conjugated (Abcam, Ab6721)	1:3000
Mouse monoclonal antibody to DV types 1, 2, 3 and 4 [D1-11(3)] (Abcam, Ab9200)	1:1000 to 1:250	Rabbit polyclonal antibody to mouse IgG H&L, HRP conjugated (Abcam, Ab6728)	1:2000 to 1:500
Human polyclonal antibody to DV types 1, 2, 3 and 4. WHO International Standard Reference Reagent (NIBSC, 02/186)	1:8	Goat polyclonal antibody to human IgG, HRP conjugated (Abcam, Ab81202)	1:1000
Human serum negative for DV antibodies (NIBSC, 02/184)	1:8	Goat polyclonal antibody to human IgG, HRP conjugated (Abcam, Ab81202)	1:1000

2.9.2 Heparin and DC-SIGN capture ELISAs

ELISA plate wells were coated with recombinant DC-SIGN (recombinant human CD209, Fc chimera; Creative Biomart) at 1-100µg/mL or heparin sodium salt (Sigma) at 1-100mg/mL in 0.05M carbonate:bicarbonate buffer, using 100µL per well, and stored at 4°C overnight. Wells were washed three times using PBS with 0.05% Tween 20 and blocked for 1 hour 30 minutes using blocking buffer (300µL per well; PBS with 0.05% Tween 20 and 5% skimmed milk unless otherwise stated). Wells were incubated at 37°C with the sample antigens at 1µg/mL in blocking buffer, for two hours before washing three times using PBS with 0.05% Tween 20. Wells were incubated at 37°C with primary antibody (100µL per well; mouse monoclonal antibody to DV types 1, 2, 3 and 4 at 1:1000 dilution) for two hours before washing three times using PBS with 0.05% Tween 20. Wells were incubated at 37°C with secondary antibody (100µL per well; rabbit polyclonal antibody to mouse IgG H&L, HRP conjugated at 1:750 dilution) for two hours before washing three times using PBS with 0.05% Tween 20. Equal volumes of ABTS™ HRP substrate solutions A and B (KPL) were mixed together and added to the wells (100µL per well), and incubated at 37°C in the dark for one hour. ABTS™ stop solution was added (100µL per well), and the plate read immediately.

2.10 Electron microscopy

Transmission electron microscopy (TEM) analysis was conducted on the dengue VLP samples by Howard Tolley (HPA CEPR). Briefly, a shallow meniscus (between 2 and 3μL) of sample suspension was applied to a filmed TEM grid (400 mesh, old pattern, formvar/carbon coated copper grid) and left to adhere for approximately 30 seconds. Surplus sample suspension was removed by touching the edge of the grid with moist filter paper. Distilled water (between 2 and 3μL) was applied to wash the grid and was removed after approximately 10 seconds with moist filter paper. Negative stain (between 2 and 3μL of 2% uranyl acetate) was applied to the grid and removed after 1 to 10 seconds using moist filter paper. Grids were air dried prior to TEM examination using a CM100 TEM (Philips) operating at 80kV using a lanthanum hexaboride tip.

CHAPTER 3. ANALYSIS OF DENGUE VIRUS QUASISPECIES POPULATIONS IN CLINICAL SAMPLES

3.1 Introduction

A wide spectrum of disease from subclinical to life-threatening is observed in dengue virus (DV) infected individuals. The proposed causes of dengue pathogenicity is a subject of much debate and involves complex interactions between host genetics and immunity, as well as viral genetic factors.

In Sri Lanka, the first severe dengue epidemics did not occur until after 1989 (Messer *et al*., 2002). This was despite Sri Lanka being an area of hyperendemicity for all four DV types for many years, and exhibiting large numbers of secondary infections among the population. This coincided with clade replacements in the circulating DV type 3 genotype 3 strains (Lanciotti *et al*., 1994). DV sequence variations at the genotype level and the emergence of strains into new areas have also been associated with changing patterns in the severity of disease in other parts of the world (Rico-Hesse *et al*., 1997a).

Recombination of DVs within DV types has been documented in both humans and mosquitoes (Aaskov *et al*., 2007; Craig *et al*., 2003). This predicts the possible emergence of recombinant strains that differ in pathogenic potential to the parent strains. In addition, the low-fidelity of viral RNA polymerases leads to the incorporation of mutations into progeny RNA, and the production of a quasispecies population of virus variants within each infected host (DeFilippis and Villarreal, 2001; Domingo, 1998). This intrahost viral sequence diversity suggests that some viral variants will be better adapted for survival in the host than others, and that this will be reflected in disease pathogenesis. Quasispecies investigations with poliovirus have demonstrated a link between the extent of sequence diversity and pathogenicity; too many or too few virus variants within the population results in reduced viral fitness and attenuation (Crotty, Cameron, and Andino, 2001; Vignuzzi *et al*., 2006). Important selective constraints acting on DVs in particular include the requirement for viral replication in both humans and mosquitoes, and the antibody dependent enhancement of infection by cross-reactive antibodies raised against a previous infection with a different DV type.

Quasispecies populations have been demonstrated for DV types 2 and 3 in clinical samples from patients in Taiwan and Myanmar and also in mosquitoes (Craig *et al.*, 2003; Lin *et al.*, 2004; Wang *et al.*, 2002a). As a result of so few studies, little is known about how the quasispecies population dynamics vary between different individuals, host populations and infecting virus strains. Previous investigations have focused on the DV envelope glycoprotein (E) gene as it encodes the E, which comprises the majority of the outer surface of the virion. E is involved in antibody recognition, attachment to host cells and in membrane fusion. Variations in E constitute potential mechanisms for alterations in virus virulence and pathogenicity.

Studies of DV quasispecies population dynamics are essential for in-depth understanding of the biology, epidemiology and evolution of DV. This has implications for disease pathogenesis, vaccine design and safety, and the development of novel therapeutic and diagnostic tools. Improved understanding of dengue epidemiology and virus evolution in different regions will enable more effective control programmes to be implemented.

3.2 Chapter objectives

The primary objective of this chapter was to investigate DV quasispecies populations (by E gene sequence analysis) both within individual patient samples and between samples from the same outbreak in Sri Lanka, in 2006. Specifically, the correlation between disease severity and either the extent of quasispecies variation or the presence of specific variants was to be investigated. Quasispecies variants with the potential to affect virus virulence or disease pathogenesis would be identified.

Previous studies of DV quasispecies populations used RT-PCR, ligation independent cloning, and sequencing of partial or complete viral E genes; examining between 10 and 21 clones per sample and reporting mutation rates of between 0.16% and 0.38% (Craig *et al*., 2003; Lin *et al*., 2004; Wang *et al*., 2002a). For this study complete E genes would be amplified directly from patient serum without *in-vitro* passage of the virus to avoid the introduction of mutations (Chen, Wu, and Chiou, 2003) using DV type-specific primers and high-fidelity RT-PCR. The use of high-fidelity RT-PCR was deemed essential to minimise the number of mutations introduced by the amplification process itself. Amplification of complete rather than partial E genes would enable potential identification of recombinant viruses via phylogenetic comparison of different regions of the E genes. The amplified E genes would be cloned and multiple clones per sample sequenced. A greater number of samples would be analysed in this study than in previous studies, to provide more data for statistical analysis. Sequence data would be analysed to define E gene variants and to investigate potential relationships between quasispecies variation and disease severity.

Using an estimated viral RNA polymerase error rate of $1x10^{-4}$ (or 0.01%; Smith *et al*., 1997) and a flavivirus E gene length of 1485 nucleotides (correct for DV types 1, 2 and 4; DV type 3 E gene length is 1479 nucleotides), one mutant clone was expected to be found for every 6.7 clones analysed. From each sample, 96 clones would be studied, with the expectation that 15% of clones would contain mutations. Mutations in up to 70% of clones have been reported previously (Craig *et al*., 2003).

Phylogenetic analyses would be performed on the consensus and variant E gene clone sequences from each patient sample to investigate the variability of the DVs circulating during the Sri Lankan outbreak in 2006. The DV genotypes present would be identified, and the variant E gene clone sequences assessed for evidence of recombination. Protein sequence changes caused by the E gene mutations would be mapped to the DV E sequences and three-dimensional structures. This was to identify mutations in locations with the potential to affect virulence or pathogenicity, through altering virus attachment to the host cell, fusion of viral and host cell membranes or antibody recognition and attachment to the virus. These variant DV E genes would be used to construct recombinant DV E (Chapter 4) for use in studies to determine the effect of the mutations (Chapter 5).

3.3 Design and testing of primers for amplification of dengue virus envelope genes

To enable the design of primers for DV E gene amplification, published sequences were obtained from the NCBI database and aligned for all four DV types (sections 2.2.1 and 2.2.2). To ensure detection of the greatest number of strains within each virus type, and to allow the entire E gene to be sequenced in subsequent experiments without introducing mutations into the E gene sequence via the primers, primers were designed up to 200 nucleotides outside of the E gene in conserved regions of the sequence alignments (section 2.2.3). Primer sequences were uploaded to the Sigma-Aldrich DNA Oligo Design Tool to estimate their melting temperature, secondary structure and ability to dimerise. Primer specificity was predicted using BLAST (NCBI), and only primers with 100% correct matches to the appropriate DV type were tested further.

Combinations of primers (Table 3.1) were tested for each DV type using RNA from the DV strains detailed in section 2.1.3. RNA was transcribed to cDNA, and the cDNA amplified using standard RT-PCR (section 2.3.1.1); amplicons were analysed by DNA agarose gel electrophoresis (section 2.5.1). Each primer set was optimised for primer concentration and annealing temperature before being tested for specificity by RT-PCR using RNA from YFV, JEV, TBEV and DV types 1, 2, 3 and 4 as template material (section 2.1.3 details source and strain information). As summarised in Table 3.1 (DNA gel electrophoresis images not shown), the first DV type 1 (Dengue 1 forward (F) 1 and reverse (R) 1) and 4 (Dengue 4 F1 and R1) primer sets tested produced amplicons of the expected size (1808 and 1799 nucleotides respectively). However, the initial DV type 2 primer set (Dengue 2 F1 and R1) produced an amplicon of the expected size for both DV type 2 and type 3 RNA. Due to the non-specific nature of this primer set, attempts were made to improve specificity by substituting the forward primer (Dengue 2 F1) for an alternative forward primer (Dengue 2 F2). The new primer set was specificity tested by RT-PCR as before, and only produced amplicons of the expected size (1730 nucleotides) using DV type 2 RNA. The initial DV type 3 primer set (Dengue 3 F1 and R1) failed to produce an amplicon of the expected size, although several

smaller amplicons were observed. These primers were replaced and the new primer set (Dengue 3 F2 and R2) produced an amplicon of the expected size although multiple non-specific amplicons were also present. Attempts to remove the additional non-specific amplicons by annealing temperature optimisation were unsuccessful, so two new reverse primers (Dengue 3 R3 and R4) were designed and tested. Both combinations of primers (Dengue 3 F2 and R3 or R4) were able to produce amplicons of the expected size (1801 and 1764 nucleotides respectively). Specificity tests showed the primer set using the R3 primer produced multiple amplicons for all flavivirus RNA tested, with the exception of DV type 1; the primer set using the R4 primer was specific for DV type 3 RNA.

Table 3.1. Primers tested for RT-PCR amplification of dengue virus envelope genes for quasispecies analysis.

Primer	Sequence (5' to 3')	Reason for replacement
Dengue 1 F1	GGT CTA GAA ACA AGA ACC GAA ACG TGG	Not replaced
Dengue 1 R1	GAC AGT CTT TTT GGG GAG TCA G	Not replaced
Dengue 2 F1	GGA CTG GAG ACA CGA ACT GAA AC	Cross-reactivity observed with DV type 3 RNA
Dengue 2 F2	TGA GAC ATC CAG GCT TTA CCA TAA TG	Not replaced
Dengue 2 R1	AGC TAG TTT TGA AGG GGA TTC TGG	Not replaced
Dengue 3 F1	CGC ACC CAA ACC TGG ATG TC	Failed to produce an amplicon of the expected size
Dengue 3 F2	TGG ATG TCG GCT GAA GGA GCT TGG AG	Not replaced
Dengue 3 R1	CCG AAT ATT ATT TGG TGC ACC ATT TTC CC	Failed to produce an amplicon of the expected size
Dengue 3 R2	CCT ATT CCA ATT TTC ATT ATC CAA G	Multiple bands present at all annealing temperatures and primer concentrations
Dengue 3 R3	TCC AGG TGT GGA CCT CAT TAG	Cross reactivity observed with all flavivirus RNA tested except DV type 1
Dengue 3 R4	GGA GTC TGC TTG AAA TTT GTA TTG CTC	Not replaced
Dengue 4 F1	TCA GTA GCT TTA ACA CCA CAT TCA GG	Not replaced
Dengue 4 R1	TCT GTC CAA GTG TGC ACG TTG TC	Not replaced

Finalised primer sets are summarised in Table 3.2. These primers were chosen to amplify DV E genes from clinical samples for quasispecies analysis on the basis of their specificity. The alignment reports for these primer sets are presented in Appendix A. Figure 3.1 shows the DNA agarose gel electrophoresis images from the specificity test RT-PCRs. No cross reactivity between flaviviruses was observed as each primer set

produced amplicons consistent in size with the DV E gene and flanking regions (between 1.5 and 2.0kb) for only its intended DV type. An annealing temperature of 60°C was adopted for all primer sets to enable parallel amplification of the different DV types. At this annealing temperature the DV type 1, 2 and 3 primer sets each produced two amplicons (Figure 3.1), but the complete E gene amplicon could be identified between 1.7 and 1.8kb (depending on the primers used). For DV types 1 and 2, these additional amplicons were the result of non-specific binding due to the presence of excess template RNA in the RT-PCR reaction. They are not visible at lower template concentrations in the high-fidelity RT-PCR positive control reactions (Figure 3.4 and Figure 3.5). For DV type 3, reducing the concentration of template RNA eliminated some, but not all of the non-specific amplicons.

Table 3.2. RT-PCR primers used for amplification of dengue virus envelope genes from patient samples for quasispecies analysis.

Primer	Sequence (5' to 3')	Amplicon size (bp)
Dengue 1 F	GGT CTA GAA ACA AGA ACC GAA ACG TGG	1808
Dengue 1 R	GAC AGT CTT TTT GGG GAG TCA G	
Dengue 2 F	TGA GAC ATC CAG GCT TTA CCA TAA TG	1730
Dengue 2 R	AGC TAG TTT TGA AGG GGA TTC TGG	
Dengue 3 F	TGG ATG TCG GCT GAA GGA GCT TGG AG	1764
Dengue 3 R	GGA GTC TGC TTG AAA TTT GTA TTG CTC	
Dengue 4 F	TCA GTA GCT TTA ACA CCA CAT TCA GG	1799
Dengue 4 R	TCT GTC CAA GTG TGC ACG TTG TC	

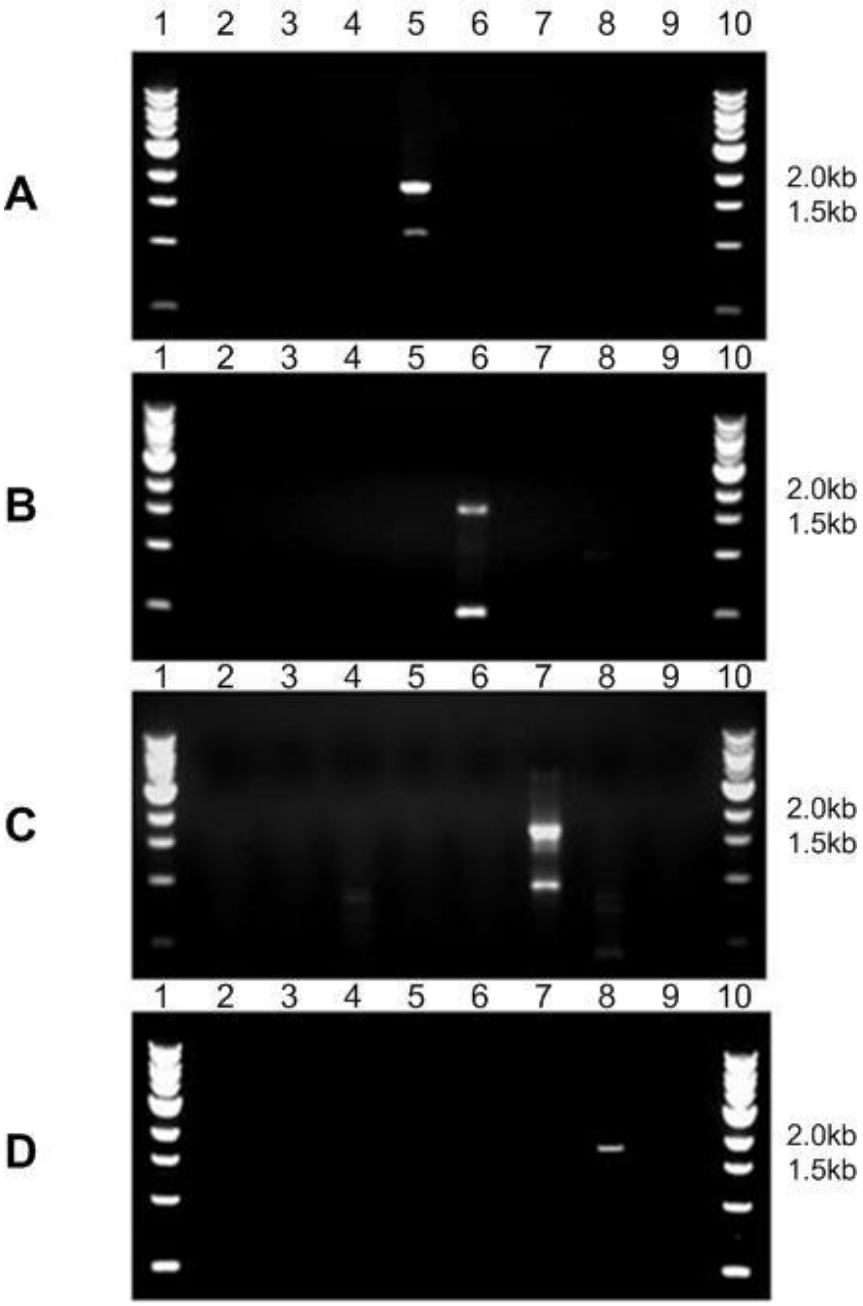


Figure 3.1. Analysis of dengue virus envelope gene RT-PCR primer specificity.

RT-PCRs were performed using the primers from Table 3.2, and RNA from several flaviviruses to assess primer specificity. Panels A to D show the DV types 1, 2, 3 and 4 RT-PCR primer sets respectively. Lanes 1 and 10 are 1kb DNA ladder, with the 1.5 and 2.0kb fragments highlighted. Lanes 2 to 8 are the RT-PCR amplicons generated using YFV, JEV, TBEV, DV type 1, DV type 2, DV type 3, and DV type 4 RNA respectively as RT-PCR template. Lane 9 is a negative template control.

For quasispecies analysis it was important that the number of mutations introduced into the E genes as part of the RT-PCR process itself was as low as possible. The standard RT-PCR kit (section 2.3.1.1) used for the testing and optimisation of the DV E

gene primers contained *Thermus aquaticus* (*Taq*) polymerase which lacks 3'-5' proofreading activity (Tindall and Kunkel, 1988). The AccuScript high-fidelity RT-PCR kit (Stratagene; section 2.3.1.2) was used for this study because it includes a high-fidelity reverse transcriptase as well as a high-fidelity polymerase (*PfuUltra*). The reverse transcriptase combines Moloney murine leukaemia virus (MMLV) reverse transcriptase with the *E. coli* DNA polymerase III ε subunit to give an error rate of 1 mutation per 62,200 nucleotides (Arezi and Hogrefe, 2007). *PfuUltra* comprises *Pyrococcus furiosus* (*Pfu*) DNA polymerase and a *Pfu* dUTPase to improve yield (compared to *Pfu* DNA polymerase alone; Hogrefe *et al.*, 2002) and gives an error rate of 1 mutation per 2,325,581 nucleotides (Arezi and Hogrefe, 2007). Initial tests using the high-fidelity RT-PCR showed it to be less sensitive than the standard RT-PCR (Figure 3.2 panels B and A respectively), as determined by visual assessment of comparative band intensity, but the DV E gene amplicon was clearly visible between the 1.5kb and 2.0kb ladder fragments. This decreased sensitivity would be expected due to the lower processivity of the high-fidelity enzymes.

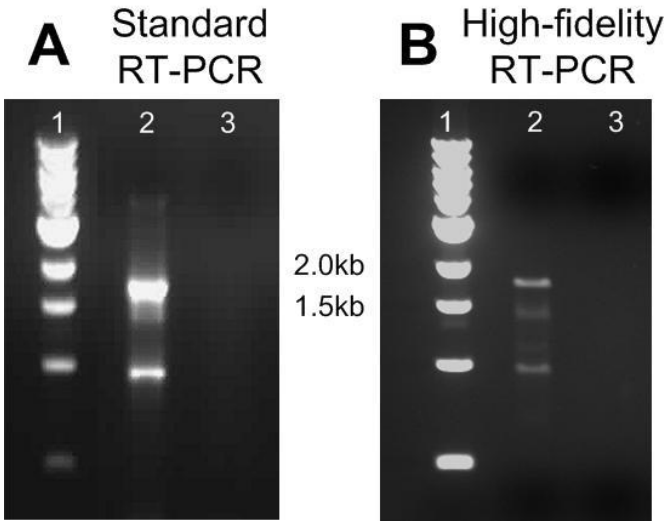


Figure 3.2. Comparison of dengue virus envelope gene reverse transcription and amplification using standard and high-fidelity RT-PCR.

A 1kb DNA ladder is shown in lane 1, with the 1.5 and 2.0kb fragments highlighted. Positive RT-PCR reactions (lane 2) used DV type 3 RNA as template material. Negative reactions (lane 3) used molecular grade water instead of template RNA. The same DV type 3 RNA and number of PCR amplification cycles (36) was used for both methods.

3.4 Dengue patient samples used for amplification of dengue virus envelope genes and quasispecies analysis

RNA used as template material for the high-fidelity RT-PCRs was extracted (section 2.3.2.1) from serum samples obtained from dengue patients admitted to hospitals in and around Ragama, Sri Lanka, during a fever study conducted in 2006 by Major Mark Bailey (Royal Army Medical Corps, RAMC) and his colleagues Dr. Ranjan Premaratna and Prof. Janaka de Silva (Department of Medicine, University of Kelaniya, Sri Lanka). Samples were obtained on admission from 617 patients presenting with fever, or a history of fever, and sent to the Armed Forces Research Institute of Medical Sciences (AFRIMS) in Thailand and Health Protection Agency (HPA) Centre for Emergency Preparedness and Response (CEPR) for laboratory diagnosis. A summary of the spectrum of disease diagnosed in these patients is shown in Figure 3.3.

Of these patients, 143 were diagnosed with dengue by RT-PCRs or ELISAs performed previously by HPA CEPR and AFRIMS. Of these 143, 31 were also diagnosed with chikungunya (virus RNA detected by RT-PCR), one also diagnosed with leptospirosis (*Leptospira* bacteria detected by culture and/or antibodies by IgM ELISA), one also diagnosed with an unidentified bacterial infection (detected by blood culture), and three also diagnosed with Q fever (*Coxiella burnetii* bacteria DNA detected by PCR and/or antibodies by ELISA). For the purposes of this study, only samples from those patients whose records showed a positive result for DV RNA by RT-PCR were used, which amounted to 95 samples. Patients with only positive dengue serology results were not included because their samples were less likely to still contain DV, as an immune response had already developed (clearing the virus and making E gene amplification less likely). All four DV types were circulating in Sri Lanka at the time the fever study samples were collected. Of the 95 samples positive for DV RNA, 67% were DV type 3 infections, 28% were DV type 2 infections and DV type 1 and 4 infections represented 2% and 3% of the patients respectively. There were no occurrences of dual infection with multiple DV types.

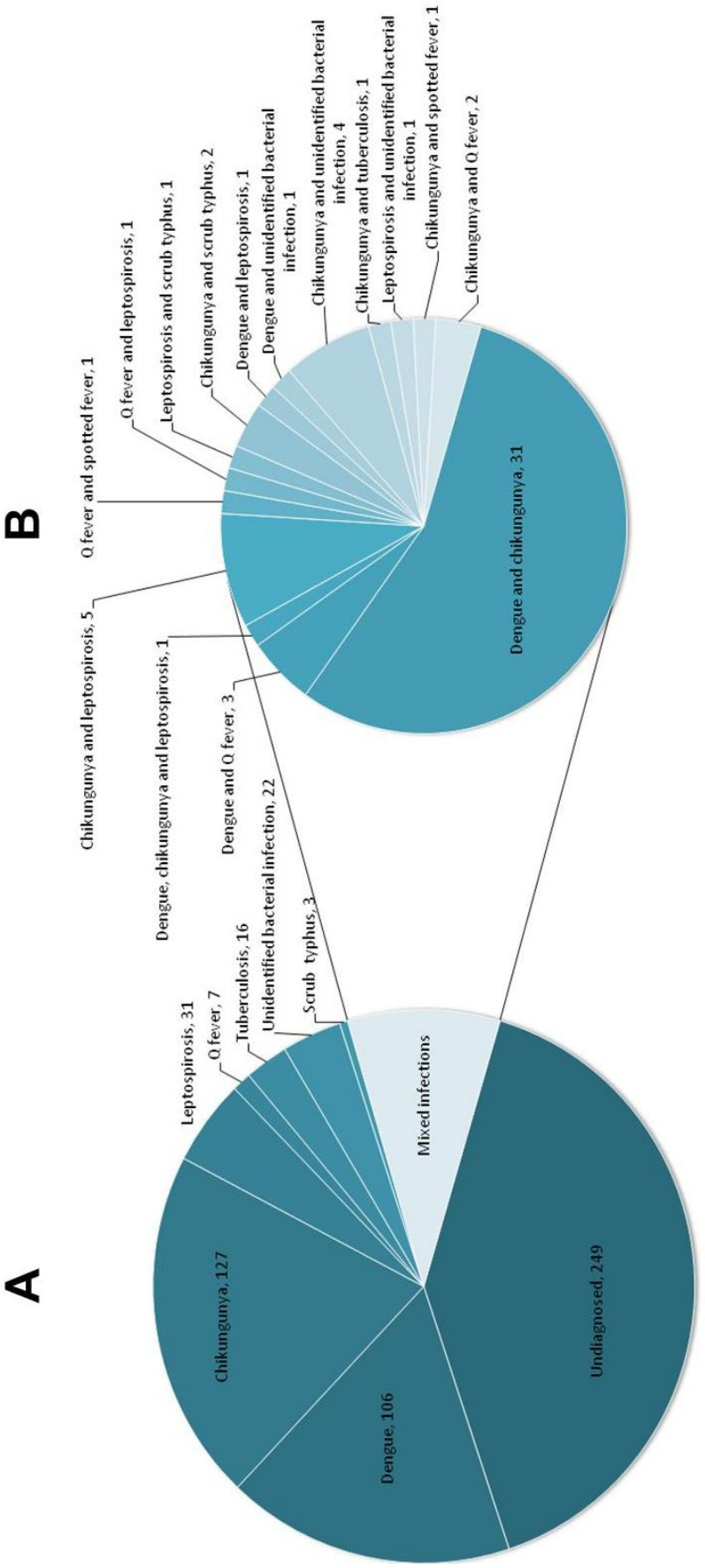


Figure 3.3. Summary of diagnoses for patients enrolled in the Sri Lankan fever study.

Diagnoses and number of patients are shown in pie chart A with the mixed infections represented in detail in pie chart B (diagram generated for this study; Major Mark Bailey, unpublished data). Taken together, dengue and dengue mixed infections account for 143 (23%) of the patients.

Disease severity for each of the fever study patients was classified (with guidance from Major Mark Bailey, RAMC) as mild or severe based on the more relaxed WHO guidelines used in the treatment of children or in small hospitals in India, where the term 'severe dengue' includes any case of dengue displaying fever and one or more of theDHF or DSS indicators detailed in the 1997 WHO guidelines (section 1.2.3; World Health Organisation, 1997a). Table 3.3 summarises the clinical data for each sample, including the infecting DV type and patient's serological status, the presence of haemorrhage, low platelet count or hypotension and the resulting disease severity classification. Patient samples were designated as severe cases where the clinical history included haemorrhaging, thrombocytopenia (<100,000 platelets/μL) or hypotension. Mild cases were those in which these symptoms were absent. Of the 95 DV RT-PCR positive clinical samples, 57 (60%) were classified as severe cases and 38 (40%) as mild cases.

Primary antibody responses occur in people who have not previously been infected (or immunised) and are characterised by the presence of IgM antibodies after several days of fever, or after the fever has passed. Secondary responses occur as a result of a second flavivirus infection (e.g. sequential infection with different DV types) and it is the IgG response that is dominant (Shu and Huang, 2004). IgG antibodies from secondary DV infections can persist in the blood for over 10 months (Gubler, 1996). Negative serology is indicative of an early-stage infection prior to immune response development. None of these patients met the 1997 WHO criteria for distinguishing between dengue fever and dengue haemorrhagic fever or dengue shock syndrome (section 1.2.3; World Health Organisation, 1997a). This can be attributed to incomplete testing or recording of observations such as pleural effusion, ascites or hypoproteinaemia and repeat hematocrit results by the hospital staff in Sri Lanka. This information may have enabled four of the patients (R006, R014, R024 and R380) to meet the 1997 WHO criteria for DHF and DSS.

Table 3.3. Classification of disease severity for dengue virus RT-PCR positive patient samples.

Sample	DV type	Haemorrhage	Platelets <100,000/µL	Hypotension	Severity classification
R004	3	No	No	Yes	Severe
R006	3	Yes	Yes	Yes	Severe
R007	3	No	Yes	Yes	Severe
R008	3	Yes	Yes	No	Severe
R011	2	No	No	No	Mild
R012	3	No	Yes	No	Severe
R014	3	Yes	Yes	Yes	Severe
R016	3	No	Yes	No	Severe
R022	2	No	No	No	Mild
R024	3	Yes	Yes	Yes	Severe
R036	3	No	No	No	Mild
R038	3	No	Yes	No	Severe
R041	3	No	No	No	Mild
R054	3	Yes	Yes	No	Severe
R060	3	No	Yes	Yes	Severe
R062	2	No	No	No	Mild
R077	3	Yes	Yes	No	Severe
R080	3	Yes	Yes	No	Severe
R095	3	No	No	No	Mild
R097	3	No	Yes	No	Severe
R100	3	No	Yes	No	Severe
R103	2	No	Yes	No	Severe
R105	2	No	No	No	Mild
R107	2	Yes	No	Yes	Severe
R108	3	No	No	No	Mild
R109	3	No	No	Yes	Severe
R114	3	No	Yes	No	Severe
R120	3	No	Yes	No	Severe
R123	3	No	No	Yes	Severe
R126	2	No	No	No	Mild
R128	3	No	No	No	Mild
R146	2	No	Yes	Yes	Severe
R149	3	Yes	Yes	No	Severe
R153	4	No	Yes	No	Severe
R156	3	No	Yes	No	Severe
R157	3	No	Yes	No	Severe
R167	2	No	Yes	No	Severe
R170	2	No	Yes	Yes	Severe
R172	3	No	Yes	No	Severe
R178	2	No	Yes	No	Severe
R179	3	No	Yes	No	Severe
R191	2	No	No	No	Mild
R193	4	No	No	Yes	Severe
R197	3	No	Yes	No	Severe
R198	2/3	No	Yes	No	Severe
R203	3	No	Yes	No	Severe
R211	2	No	Yes	No	Severe
R216	3	No	No	No	Mild
R217	3	No	No	No	Mild
R221	3	No	No	No	Mild
R230	3	No	No	No	Mild

Sample	DV type	Haemorrhage	Platelets <100,000/µL	Hypotension	Severity classification
R232	3	No	No	No	Mild
R254	3	No	Yes	No	Severe
R255	3	No	No	No	Mild
R259	2	Yes	Yes	No	Severe
R261	3	No	No	No	Mild
R265	1	No	No	No	Mild
R267	2	No	Yes	No	Severe
R282	4	No	No	No	Mild
R286	2	Yes	Yes	No	Severe
R290	3	No	No	No	Mild
R291	2	Yes	Yes	No	Severe
R299	3	Yes	No	No	Severe
R316	2	No	No	No	Mild
R325	3	Yes	No	No	Severe
R330	3	No	Yes	No	Severe
R334	3	No	No	No	Mild
R337	3	No	Yes	No	Severe
R342	3	No	No	No	Mild
R352	3	No	No	No	Mild
R355	2	No	Yes	No	Severe
R358	2	No	Yes	Yes	Severe
R363	3	No	No	No	Mild
R373	3	No	No	No	Mild
R380	3	Yes	Yes	Yes	Severe
R384	3	No	No	No	Mild
R388	3	No	No	Yes	Severe
R390	3	No	No	No	Mild
R394	2	No	No	No	Mild
R410	2	No	Yes	No	Severe
R421	2	No	No	No	Mild
R424	3	Yes	Yes	No	Severe
R433	2	No	No	No	Mild
R443	3	No	No	No	Mild
R447	3	No	No	No	Mild
R452	3	No	Yes	No	Severe
R478	3	Yes	No	No	Severe
R512	3	No	Yes	No	Severe
R513	1	No	No	No	Mild
R515	3	No	Yes	No	Severe
R516	2	No	No	No	Mild
R517	3	Yes	Yes	No	Severe
R592	3	No	No	No	Mild
R605	2	No	No	No	Mild
R617	3	No	Yes	No	Severe

Table 3.3 continued from page 88.

3.5 High-fidelity RT-PCR, cloning and sequencing of dengue virus quasispecies envelope genes from patient samples

High-fidelity RT-PCR (section 2.3.1.2) was performed in duplicate on the 95 DV RNA positive samples using the appropriate DV RNA as a positive control (section 2.1.3); molecular biology grade water was used in place of template RNA as a negative control. Previous testing of the clinical samples by RT-PCR (conducted by HPA CEPR and AFRIMS) identified the DV type for each sample (Table 3.3), which enabled the appropriate type-specific primers to be used for E gene amplification. E gene amplification from sample R198A was attempted using both DV type 2 and DV type 3 primer sets, as previous HPA and AFRIMS RT-PCRs had given conflicting results (data not shown). Analysis of RT-PCR amplicons by DNA agarose gel electrophoresis yielded 21 amplified DV E genes of between 1.5 and 2.0 kb (Figures 3.4 to 3.7). All positive and negative control reactions yielded the expected results. Two samples (R265A and R513A) were previously identified by diagnostic RT-PCR as containing DV type 1 RNA. DV E genes from these samples were unable to be amplified by high-fidelity RT-PCR using the DV type 1-specific primers (Figure 3.4). Twenty-eight samples were previously identified by diagnostic RT-PCR as containing DV type 2 RNA. Of these, DV E genes from only 6 (R105A, R107A, R126A, R178A, R259A and R267A) were amplified by high-fidelity RT-PCR using the DV type 2-specific primers (Figure 3.5). R105A, R126A, R178A and R267A produced very faint positive results which are not easily visible in Figure 3.5. Sixty-five samples were previously identified by diagnostic RT-PCR as containing DV type 3 RNA. Of these, DV E genes from only 14 (R004A, R080A, R097A, R109A, R114A, R123A, R197A, R203A, R232A, R254A, R255A, R261A, R299A and R334A) were amplified by high-fidelity RT-PCR using the DV type 3-specific primers (Figure 3.6). R004A, R080A, R109A, R123A, and R254A produced very faint positive results which are not easily visible in this figure. Three samples were previously identified by diagnostic RT-PCR as containing DV type 4 RNA. The DV E gene from only one of these (R282A) was amplified by high-fidelity RT-PCR using the DV type 4-specific primers (Figure 3.7). Additional amplicons were present in some of the PCR products that were not the correct size for the DV E gene. These are possibly the result of non-specific amplification where excess RNA template was used. DV E gene amplicons

from 12 of the 21 samples were considered strong positives as they appeared as high intensity bands on the gels (summarised in Table 3.4). DNA amplicons were visualised briefly using low intensity UV light to limit any damage to the DNA, and the appropriate bands excised and the DNA purified for cloning (section 2.3.3.1).

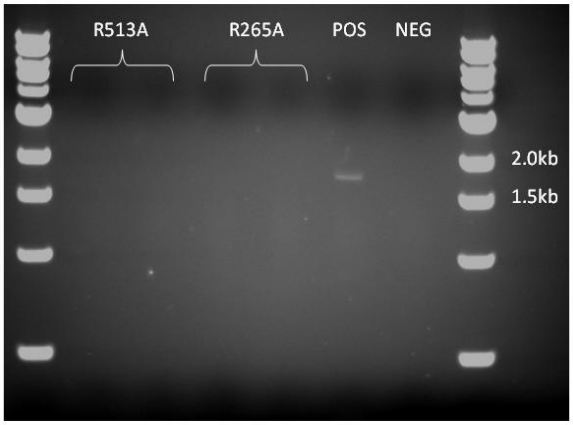


Figure 3.4. Dengue virus type 1 patient sample high-fidelity RT-PCR product analysis.

Positive control RNA (POS) was DV type 1 Hawaii A strain. Negative (NEG) template control reactions showed no amplification.

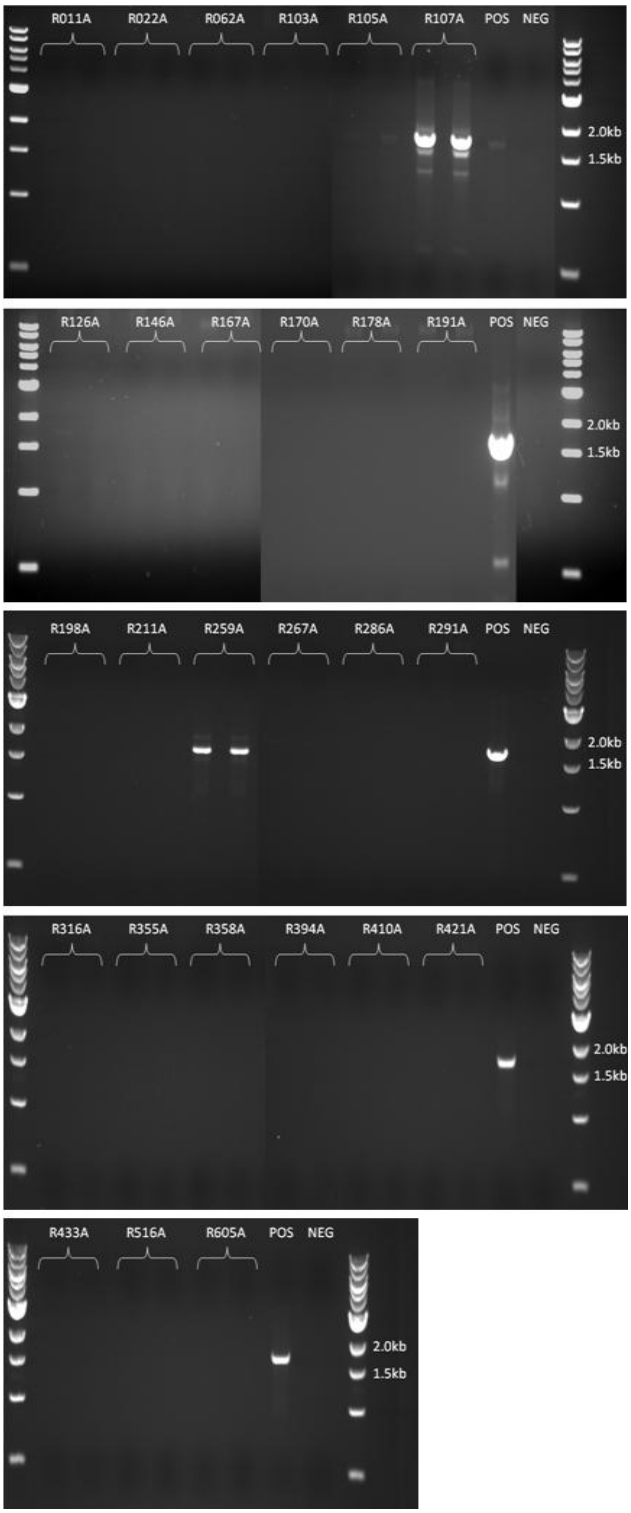


Figure 3.5. Dengue virus type 2 patient sample high-fidelity RT-PCR product analysis.

Positive control RNA (POS) was DV type 2 New Guinea C strain. Negative (NEG) template control reactions showed no amplification. The original (unprocessed) gel images used to construct this figure are included electronically in Appendix J.

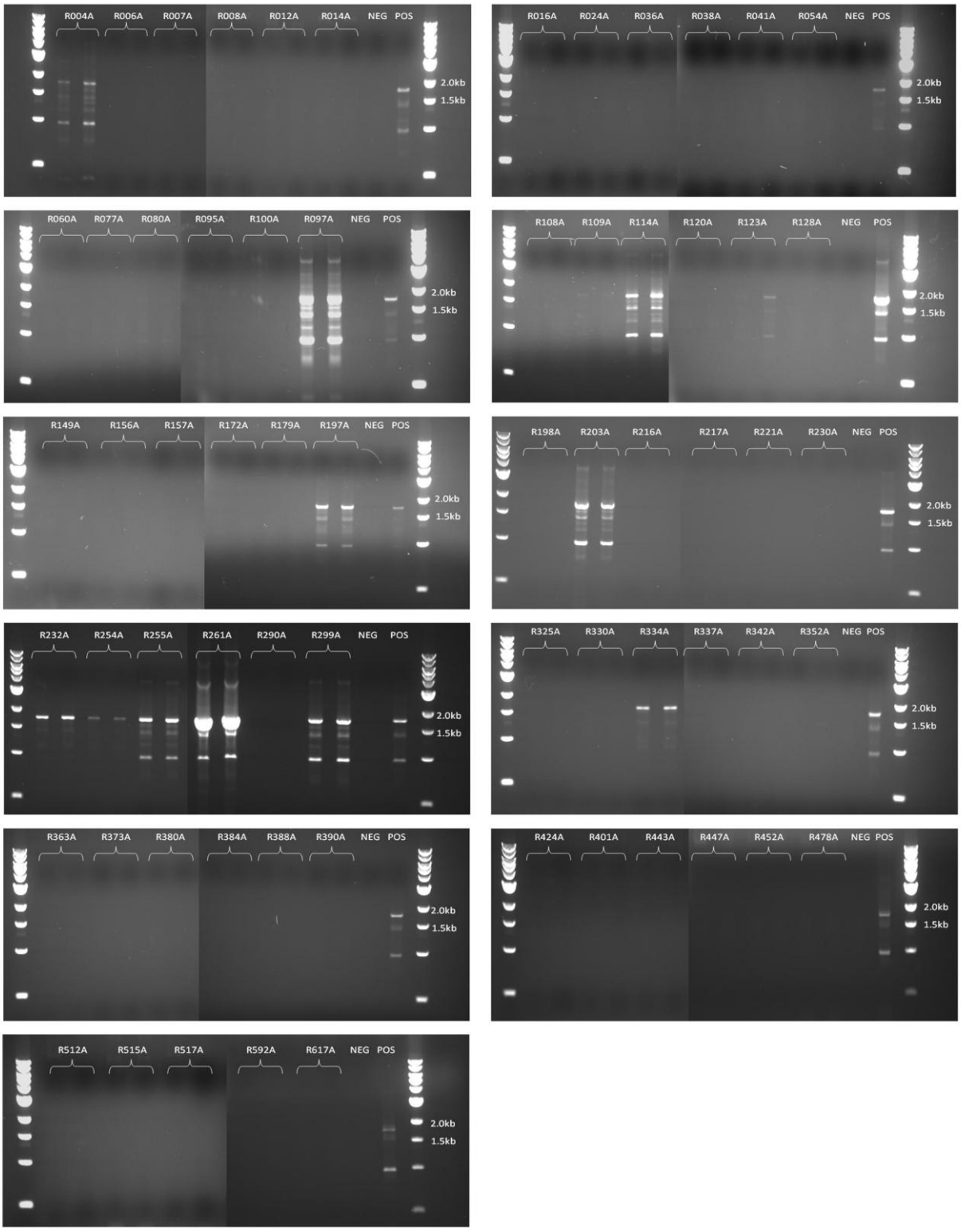


Figure 3.6. Dengue virus type 3 patient sample high-fidelity RT-PCR product analysis.

Positive control RNA (POS) was DV type 3 H87 strain. Negative (NEG) template control reactions showed no amplification. The original (unprocessed) gel images used to construct this figure are included electronically in Appendix J.

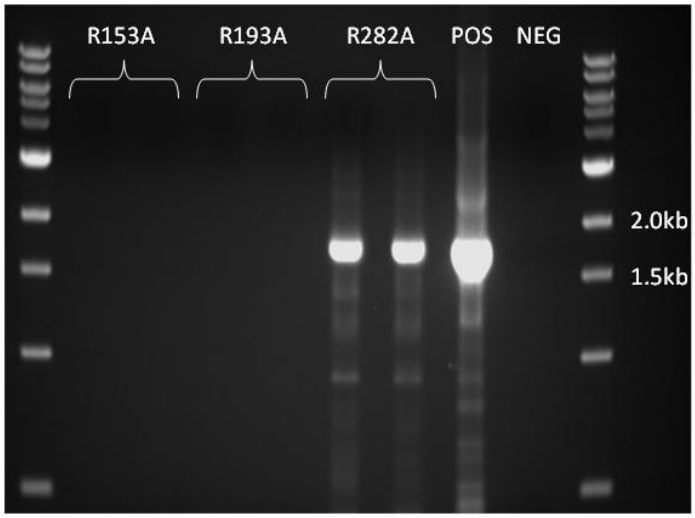


Figure 3.7. Dengue virus type 4 patient sample high-fidelity RT-PCR product analysis.

Positive control RNA (POS) was DV type 4 H241 strain. Negative (NEG) template control reactions showed no amplification.

Table 3.4. Dengue virus envelope genes amplified from patient samples by high-fidelity RT-PCR for cloning.

Sample number	Dengue virus type	Severity classification
R004A	3	Severe
R080A	3	Severe
R097A	3*	Severe
R105A	2	Mild
R107A	2*	Severe
R109A	3	Severe
R114A	3*	Severe
R123A	3	Severe
R126A	2	Mild
R178A	2	Severe
R197A	3*	Severe
R203A	3*	Severe
R232A	3*	Mild
R254A	3	Severe
R255A	3*	Mild
R259A	2*	Severe
R261A	3*	Mild
R267A	2	Severe
R282A	4*	Mild
R299A	3*	Severe
R334A	3*	Mild

From the 95 patient samples known to contain DV, only 21 DV E genes were amplified. Strong positives are marked with *, and disease severity classifications are shown.

The DV E gene amplicons produced by the high-fidelity RT-PCRs were blunt-ended so were A-tailed (section 2.3.5.1) prior to cloning. This enabled ligation into the cloning vector via the 3'-terminal thymidine overhangs at the insertion site. The A-tailed E

gene amplicons were ligated (section 2.3.5.2) into linearised pGEM®-T Easy vector (Promega) and transformed into JM109 competent cells (section 2.3.5.3). The JM109 strain was chosen for its high transformation efficiency (greater than 10^8 colony forming units per μg).

JM109 cells transformed with pGEM®-T Easy containing the E gene inserts were cultured overnight on selective agar plates (section 2.3.5.3). From the twenty-one high-fidelity RT-PCR-amplified E genes (Table 3.4), nine (R004A, R080A, R105A, R109A, R123A, R126A, R178A, R254A and R267A) produced less than ten positive transformants using blue/white screening of recombinants. This could be attributed to the comparatively low yield of amplicon from the high-fidelity RT-PCRs performed on these samples (Table 3.4). This resulted in a lower concentration of DNA being used for cloning, and consequently fewer transformants. These samples did not produce enough transformants to allow assessment of quasispecies variation.

Work was continued using the twelve cloned DV E gene samples that produced enough positive transformants to enable assessment of the quasispecies variation. These twelve samples consisted of: one DV type 4 (R282A), two DV type 2 (R107A and R259A) and nine DV type 3 (R097A, R114A, R197A, R203A, R232A, R255A, R261A, R299A, and R334A) E genes (Table 3.4; denoted *). The DV type 2 and 4 E gene amplicons were cloned successfully at the first attempt and a 96-well plate of transformants was obtained for each sample. Cloning of the DV type 3 E gene amplicons proved difficult, with some cloning reactions yielding low numbers of positive transformants despite the DNA concentration appearing sufficient for cloning, and control reactions confirming that the problems were not due to faulty ligation, transformation or culture components. Attempts were made to optimise cloning efficiency and increase the number of positive transformants; method optimisation included varying the vector:insert ratios, methods of transformation (chemical and electrical), and host cell types (ABLE C and K cells which reduce the cloning vector copy number to improve cloning of toxic genes; Agilent Technologies). Incubation at a reduced incubation

temperature (28°C) was also attempted, but with little success. E gene samples R114A and R197A, yielded only 64 colonies instead of the intended 96.

Colonies were cultured for 24 hours in selective medium (section 2.3.5.4) before plasmid DNA was extracted (section 2.3.4.1). The resulting plasmid DNA was used for quasispecies analysis. In order to process the large number of clones generated for each sample, the presence of the DV E genes was verified by DNA agarose gel electrophoresis (section 2.5.1) of plasmid DNA from all clones, by monitoring differences in plasmid DNA migration. This method of screening was validated by PCR. Thirty DV E gene clones from three different patient samples were amplified by PCR (section 2.3.1.4) from plasmid DNA using M13 primers (detailed in section 2.4.1). Undigested plasmid DNA and PCR products were compared by DNA gel electrophoresis (Figure 3.8, panels A and B respectively; results shown are from eight of the thirty clones tested). The PCR products (Figure 3.8, panel B) showed amplicons of between 1.5 and 2.0kb only for clones containing the DV E gene. For each clone's plasmid DNA (Figure 3.8, panel A), two bands were observed corresponding to nicked circular plasmid DNA (>10kb) and supercoiled plasmid DNA either with (4.8kb) or without (3kb) the E gene insert. The plasmid DNA and PCR product DNA gel electrophoresis results correlated for all clones tested with no exceptions, confirming that DNA gel electrophoresis of the plasmid DNA was an acceptable method for screening clones for the presence of the DV E gene.

The presence of the DV E gene in the rest of the clones was then verified by DNA agarose gel electrophoresis of plasmid DNA (Appendix B; summarised in Table 3.5). In total, 1088 white colonies from 12 patient samples were screened for DV E genes by DNA gel electrophoresis of plasmid DNA (Appendix B). DV E genes were confirmed in 677 (62%) of these. For each patient sample, the percentage of white colonies found to contain the DV E gene ranged from 2% to 99%. For the DV type 2 and 4 clones, almost all of the white colonies picked (>95%) contained the DV E gene. For the DV type 3 clones, however, the number of white colonies containing the insert varied with each

patient sample from 2% to 99%. These white colonies were ampicillin-resistant, so contained vector and an insert at some point during their growth.

To provide sufficient DV E genes from each sample for quasispecies analysis, it was decided to proceed with the samples that had been successfully cloned with more than 50% of white colonies containing inserts of the expected size (Table 3.5). Samples identified for sequencing of DV E gene quasispecies variants were R097A, R107A, R114A, R197A, R203A, R259A, R261A and R282A. Sequencing was performed on clone plasmid DNA and directly from RT-PCR products (section 2.4). The first sequencing reactions were performed using M13 primers, which flank the multiple cloning site (MCS) in the cloning vector. To obtain sequence data covering the entire E gene (1.5kb) internal sequencing primers were designed based on the NCBI RefSeqs (section 2.2.1) and used (Table 3.6). The RT-PCR primers listed in Table 3.2 were also used for sequencing directly from the RT-PCR amplicons. This was performed to measure the fidelity of the cloning procedure, and to investigate whether the differences in DV type 3 sample cloning efficiency correlated with sequence differences between the infecting strains.

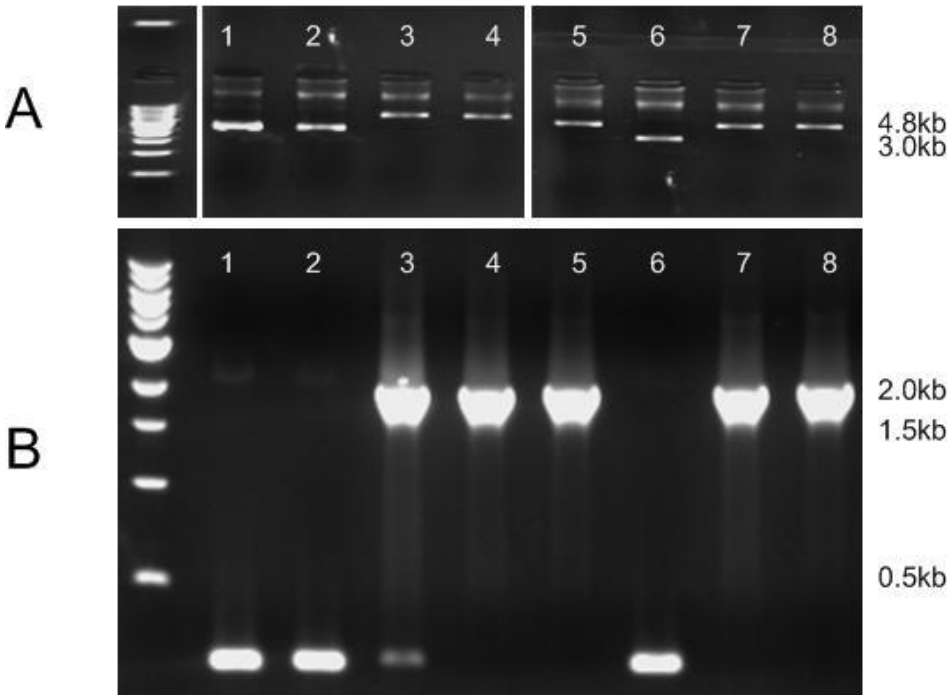


Figure 3.8. M13 PCR validation of plasmid DNA electrophoresis method for verification of dengue virus envelope gene presence in cloned plasmid DNA.

Panel A shows the DNA gel electrophoresis of undigested plasmid DNA for eight of the thirty clones analysed. Two bands are observed for each clone, corresponding to nicked circular plasmid DNA (>10 kb) and supercoiled plasmid DNA either with (4.8 kb) or without (3 kb) the E gene insert. Panel B shows the plasmid DNA M13 PCR product analysis by DNA agarose gel electrophoresis for the same 8 clones shown in panel A. Bands of 1.5 kb or 0.5 kb are shown where the DV E gene insert is present or absent respectively. The original (unprocessed) gel images used to construct panel A of this figure are included electronically in Appendix J.

Table 3.5. Presence of dengue virus envelope gene inserts in picked colonies

Sample number	Dengue virus type	No. of inserts/No. of white colonies	% of white colonies containing inserts
R107A	2	94/96	98%
R259A	2	91/96	95%
R097A	3	95/96	99%
R114A	3	52/64	81%
R197A	3	47/64	73%
R203A	3	50/96	52%
R232A	3	23/96	24%
R255A	3	47/96	49%
R261A	3	68/96	71%
R299A	3	14/96	15%
R334A	3	2/96	2%
R282A	4	94/96	98%

Table 3.6. Sequencing primers designed for dengue virus envelope gene quasispecies analysis.

Primer	Sequence (5' to 3')
Dengue 2 seq F1	TTG GAA TAC ACC ATT GTG ATA AC
Dengue 2 seq R1	GGT TGA CTG TAA TCA GGC GAC C
Dengue 2 seq F2	AGG AAT GTC ATA CTC TAT GTG CAC
Dengue 2 seq R2	AGC ACT CCA TCG TGA CAG TG
Dengue 3 seq F1	AGT GGT GCA ACA TGA GAA CCT C
Dengue 3 seq R1	GGA ATC TTG CAG GGT GCA TC
Dengue 3 seq F2	TGA GCT ATG CAA TGT GCT TGA ATA C
Dengue 3 seq R2	GCT TCA GCG GTT GAT GCC TG
Dengue 4 seq F1	ATT GAG AAC CTT GAA TAC ACA GTG G
Dengue 4 seq R1	CGA TGG TGA CCT GGG AGT TAT C
Dengue 4 seq F2	AAT GTC ATA CAC GAT GTG CTC AGG
Dengue 4 seq R2	CCT TGT TCA CAT CTC TTA TCT CTA TGG

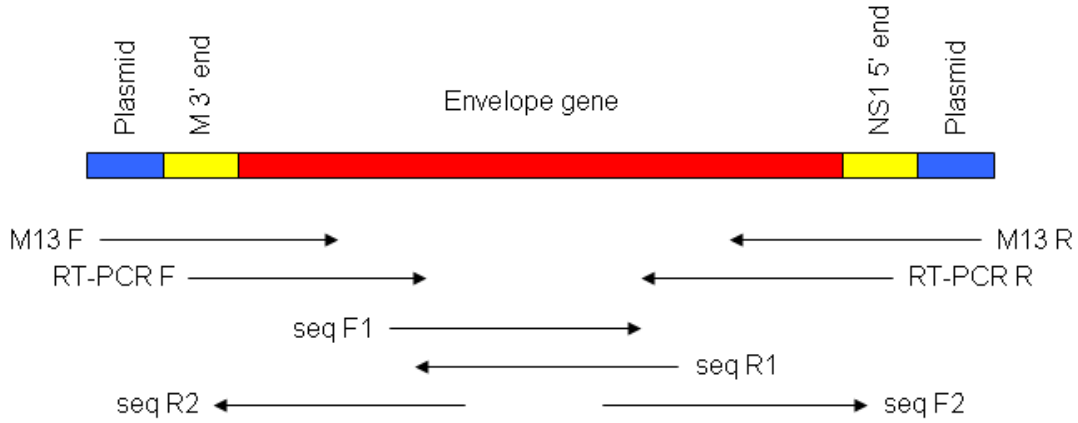


Figure 3.9. Schematic of dengue virus envelope gene sequencing primer read positions.

M13 (section 2.4.1), RT-PCR (Table 3.2) and sequencing (Table 3.6) primers were used to obtain sequence data covering the entire E gene. E gene sequence was flanked by DV prM and NS1 gene sequence, which contained the binding sites for the RT-PCR primers. M13 primer binding sites were specific to the plasmid and sequencing (seq) primers were internal to the E gene sequences.

Sequence data obtained by sequencing (of both the dengue patient sample clones, and directly from the RT-PCR products) was verified as DV E gene sequence using BLAST (NCBI). The following samples produced DV E gene sequence data for most of the insert-containing clones: R097A, R107A, R203A, R259A, R261A, and R282A. This equates to one DV type 4, two DV type 2 and three DV type 3 samples. These data were assembled into separate contigs (section 2.2.4) for each clone. When sequencing the clones from samples R114A and R197A, the majority of the sample reads obtained were of insufficient length, which meant they were either too short to incorporate into

contigs, or the contigs assembled did not cover the entire E gene. These two samples were thus not included in subsequent sequence analyses as the E gene sequence data were incomplete. Direct sequencing of the RT-PCR products and contig assembly were successful for samples R004A, R097A, R107A, R114A, R197A, R203A, R232A, R254A, R255A, R259A, R261A, R299A, R282A and R334A. Sequencing directly from the RT-PCR product was unsuccessful for samples R080A, R105A, R109A, R123A, R126A, R178A and R267A. This could be attributed to the low DNA concentration of these samples.

3.6 Sequence analysis of dengue virus quasispecies envelope genes

To compare the DV quasispecies E gene sequences from each patient sample, the contigs from each clone (section 3.5) were aligned as both nucleotide and amino acid sequences (section 2.2.4; see Appendices C and D for sequence alignments). Tables in Appendix E detail the types and positions of mutations in each sample. This information is summarised in Table 3.7. In total, 299 E gene clones were analysed across the six patient samples, a total of 436,909 nucleotides (nt) were sequenced and 64 nucleotide changes were observed. Overall mean diversity (number of mutations per nucleotide) was 0.013%, whereas mean diversity values for individual samples ranged from 0.0034% to 0.053% (average 0.018%). For all 14 of the patient samples where direct sequencing of the RT-PCR products was successful, the sequences obtained were shown to match the consensus sequence of the clones for that sample (data not shown). This confirmed that no major changes to any of the E gene sequences were introduced during the cloning process.

Table 3.7. Summary of dengue virus envelope gene sequence diversity observed in clinical samples.

Sample	Infection type	No. of clones analysed	No. of clones containing mismatches	No. of nt substitutions	No. of nt deletions or insertions	Total No. of nt sequenced	Mean diversity (mutations/ nt)
R097A	Dengue 3 severe	82	16 (20%)	15	3	119,864	0.015%
R107A	Dengue 2 severe	40	2 (5%)	1	1	59,400	0.0034%
R203A	Dengue 3 severe	25	6 (24%)	10	9	35,913	0.053%
R259A	Dengue 2 severe	45	6 (13%)	3	4	66,822	0.01%
R261A	Dengue 3 mild	47	13 (28%)	12	1	65,977	0.02%
R282A	Dengue 4 mild	60	5 (8%)	5	0	88,933	0.0056%

Differences in mean diversity between the severe and the mild dengue cases were compared using the Mann-Whitney two-sample rank test (section 2.2.5). This showed no statistically significant difference between the samples from each group (p=1.0).

Of the 299 cloned DV E genes analysed, nine contained incomplete E gene sequences (Table 3.8 and Appendices C and D). Sequence data from these clones showed pGEM®-T Easy vector sequence flanking the incomplete E gene sequence confirming

that the problem was not with the sequencing reaction. These truncated E gene clones were considered artefacts of the cloning process rather than genome defective quasispecies variants. All nine clones containing incomplete E gene sequences were from patients with DV type 3 infections. Two common points of E gene truncation were observed at E gene positions E772-3 and E1128-32. For these truncated clones, the portion of the E gene sequence up to the point of truncation was included in the clone alignments.

Table 3.8. Clones containing truncated envelope gene sequences.

Sample	Clone	Point of truncation within the envelope gene
R097A	28	E1128
	93	E773
R203A	42	E773
	77	E1132
R261A	72	E772
	75	E772
	79	E772
	80	E772
	84	E772

Of the 299 clones analysed, 48 contained at least one mutation, which amounts to 16% of the clones differing from their respective consensus sequences. Substitution, insertion and deletion mutations were all observed, with substitutions accounting for 72% of all nucleotide changes. Deletion mutations were deletions of one or two nucleotides, and accounted for 16% of all nucleotide changes. Substitution and insertion mutations were predominantly single nucleotide mutations, with the exception being one clone (R203A clone 54), which carried a mutation where a sequence of eight nucleotides were replaced with a different sequence of 14 nucleotides (substitution of eight nucleotides and insertion of six nucleotides), resulting in a protein sequence change of three amino acids and an insertion of another two amino acids. Three clones contained mutations at two separate sites (R203A clone 54, E605 and E1032-40; R097A clones 37 and 41, E291 and E736, and E163 and E229 respectively), and two clones contained deletions of two consecutive nucleotides (R203A clone 21 E914-5; R259A clone 54 E370-1). For patient sample R261A, there were two sites where multiple clones carried the same mutation (E602

and E1146; 6 and 3 clones respectively). These two groups of clones represent distinct DV lineages compared to the consensus, and their relative abundance within all of the clones analysed amounts to 13% and 6.5% respectively of the quasispecies population. In the larger group of clones the mutation causes an amino acid substitution in the protein sequence, whereas the mutation in the smaller group of clones does not affect the amino acid sequence.

A transition mutation is one where a purine (A or G) is substituted with another purine, or a pyrimidine (C, T or U) with another pyrimidine. In contrast, a transversion mutation is one in which a purine is substituted with a pyrimidine or vice versa. Nonsynonymous mutations lead to amino acid sequence changes, whereas synonymous mutations do not. Of the nucleotide substitutions, 29% were transversion mutations, and 61% of all the mutations (46 substitutions, 8 insertions and 10 deletions) were nonsynonymous (Table 3.9). Conservative amino acid changes occur when an amino acid is substituted for one with similar properties, but drastic protein changes occur when amino acids with different properties are substituted (for example, a polar amino acid such as threonine for a nonpolar amino acid such as leucine). Of these nonsynonymous mutations, 67% were drastic rather than conservative, and 50% of these drastic mutations caused frameshifts by insertion or deletion of nucleotides (Table 3.9).

Table 3.9. Summary of dengue virus envelope gene mutation types.

Sample	Infection type	Transversion mutations / total substitution mutations	Nonsynonymous mutations / total mutations	Drastic mutations / Nonsynonymous mutations	Frameshift mutations / Drastic mutations
R097A	Dengue 3 severe	5/15	9/18	9/9	3/9
R107A	Dengue 2 severe	1/1	2/2	1/2	1/1
R203A	Dengue 3 severe	0/2*	4/5	3/4	2/3
R259A	Dengue 2 severe	1/3	4/6	3/4	3/3
R261A	Dengue 3 mild	2/12	8/13	2/8	1/2
R282A	Dengue 4 mild	2/5	3/5	2/3	0/2
Total		11/38 (29%)	30/49 (61%)	20/30 (67%)	10/20 (50%)

* The multiple nucleotide substitution and insertion mutation observed in patient sample R203A clone 54 was not included in the transversion/substitution mutations calculation as it was not clear which nucleotides were substituted and which were inserted. This mutation was counted as one mutation rather than per nucleotide in all of the subsequent calculations, as were two nucleotide deletion mutations.

3.7 Comparison of mutations introduced by the RT-PCR process between this study and published data

Previous studies reported error frequencies for mutations introduced by the RT-PCR process of between 0.023 and 0.055% (Lin *et al*., 2004; Wang *et al*., 2002a). These values were determined by RT-PCR, cloning and sequencing of DV E gene T7 transcript RNA. To assess how much sequence variation was introduced using the high-fidelity RT-PCR amplification methodology used for this study, DV type 3 E gene plasmid DNA from a single colony of known sequence was transcribed *in vitro* using T7 polymerase (section 2.3.10). The resulting transcript was subjected to high fidelity RT-PCR, cloning and sequencing using the same methods applied to the Sri Lankan patient samples in this study (AccuScript method; section 2.3.1.2). To enable comparison of mutation frequencies between these results and those obtained previously by other groups, the same DV E gene RNA control transcript was also subjected to high fidelity RT-PCR, cloning and sequencing using published methods (Craig *et al*., 2003); Expand method; section 2.3.1.3). Sequence contigs were assembled and the clones from each sample aligned as both nucleotide and amino acid sequences (section 2.2.4; see Appendices F and G for alignments). Tables in Appendix H detail the types and positions of mutations in each sample. These results are summarised in Table 3.10. Analysis of the RNA transcript data obtained using the Expand and AccuScript RT-PCR methods showed that of the clones analysed, 55% and 33% respectively contained at least one mutation (Table 3.7). These values are 3.4-fold (Expand method) and 2.1-fold (AccuScript method) higher than the 16% of clones containing mutations observed using the AccuScript method and dengue patient samples (Table 3.7). Alignment consensus sequences from both methods using the RNA transcript were shown to be identical (data not shown). The AccuScript method generated fewer mutations than the Expand method, but both methods that used the RNA transcript as RT-PCR template material produced a greater number of clones containing mutations, and a higher mean diversity than was found using the AccuScript method with the dengue patient samples as RT-PCR template material.

Substitution, insertion and deletion mutations were observed for the methods using the RNA transcript, with substitutions accounting for 68% (Expand method) and 46% (AccuScript method) of all nucleotide changes. Deletion mutations were deletions of one or two nucleotides. Substitution and insertion mutations generated using the Expand method were all single nucleotide mutations. Substitution and insertion mutations generated using the AccuScript method were predominantly single nucleotide mutations, with the exceptions being clones 4, 21, 27 and 38 (Appendix H, Table 8).

Table 3.10. Summary of mutations generated by the Expand and AccuScript RT-PCR methods.

Method	No. of clones analysed	No. of clones containing mismatches	No. of nt substitutions	No. of nt deletions or insertions	Total No. of nt sequenced	Mean diversity (mutations/nt)
Expand RNA transcript	55	30 (55%)	41	19	56,639	0.11%
AccuScript RNA transcript	55	18 (33%)	18	21	60,453	0.065%
AccuScript Patient samples	299	48 (16%)	38	18	436,909	0.013%

Table 3.11. Summary of mutation types generated by the Expand and AccuScript RT-PCR methods.

Method	Transversion mutations / total substitution mutations	Nonsynonymous mutations / total mutations	Drastic mutations / Nonsynonymous mutations	Frameshift mutations / Drastic mutations
Expand RNA transcript	5/41 (12%)	48/59 (81%)	37/48 (77%)	17/37 (46%)
AccuScript RNA transcript	6/8 (75%)*	26/26 (100%)	19/26 (73%)	16/19 (84%)
AccuScript Patient samples	11/38 (29%)	30/49 (61%)	20/30 (67%)	10/20 (50%)

* The multiple nucleotide substitution and insertion mutation observed in clone 21 (AccuScript method) was not included in the transversion/substitution mutations calculation as it was not clear which nucleotides were substituted and which were inserted. This mutation was counted as one mutation rather than per nucleotide in all of the subsequent calculations, as were multiple nucleotide insertion or deletion mutations.

Of the substitution mutations, 12% (Expand method) and 75% (AccuScript method) were transversion rather than transition mutations. For both methods a high percentage of the mutations were nonsynonymous (Expand 81%; AccuScript 100%) and drastic (Expand 77%; AccuScript 73%), which is in accordance with the findings from the patient samples (see Tables 3.9 and 3.11). The percentages of insertion or deletion mutations at homopolymeric runs of nucleotides were very similar between

the methods using the RNA transcript (Table 3.12). For the AccuScript method using patient samples as RT-PCR template material, compared to the methods using the RNA transcript, the percentage of mutations at homopolymeric runs of nucleotides was up to 3-fold lower for insertion mutations and up to 1.66-fold higher for deletion mutations. The proportion of these mutations (at homopolymeric runs of nucleotides) of the total insertion and deletion mutations was also lower for the patient samples than for the methods using RNA transcript.

Table 3.12. Comparison of insertion and deletion mutations generated at homopolymeric runs of nucleotides by the Expand and AccuScript RT-PCR methods.

Method	Insertion mutations at homopolymeric runs / total insertion mutations	Deletion mutations at homopolymeric runs / total deletion mutations	Insertion and deletion mutations at homopolymeric runs / total insertion and deletion mutations
Expand RNA transcript	6/8 (75%)	3/10 (30%)	9/18 (50%)
AccuScript RNA transcript	12/18 (67%)	1/3 (33%)	13/21 (62%)
AccuScript Patient samples	2/8 (25%)	5/10 (50%)	7/18 (39%)

For the patient samples, overall rates of insertion and deletion mutations were 0.0041% compared to rates of 0.032% (Expand) and 0.035% (AccuScript) for the methods using RNA transcript as RT-PCR template material (Table 3.13). Substitution mutation rates were 0.074%, 0.03% and 0.0087% for the Expand RNA transcript method, AccuScript RNA transcript method and patient samples respectively (Table 3.13). Therefore, although rates of insertion and deletion mutations, and rates of substitution mutations were consistently higher for the methods using RNA transcript than when the patient samples were used, it is the incorporation of substitution mutations rather than insertion and deletion mutations that makes the difference between the two RT-PCR methods using the RNA transcript. The Expand method had a substitution mutation rate 2.47-fold greater than the AccuScript method.

The difference between the mutation rates observed using RNA transcript versus patient samples indicated that the presence of such mutations might be attributed to the RNA transcript used rather than the high-fidelity RT-PCR, cloning and sequencing. Thioredoxin is an accessory protein that improves the processivity of T7 polymerase,

reducing substitution mutations by 3-fold, deletion mutations of one and two nucleotides by 27- and 9-fold respectively and insertion mutations at homopolymeric runs of nucleotides by 46-fold (Kunkel, Patel, and Johnson, 1994).

Table 3.13. Comparison of mutations generated by the Expand and AccuScript RT-PCR methods.

Method	Insertion/deletion mutation rates	Substitution mutation rates
Expand RNA transcript	0.032%	0.074%
AccuScript RNA transcript	0.035%	0.030%
AccuScript Patient samples	0.0041%	0.0087%

Substitution mutation rates were calculated by dividing the number of substitution mutations by the total number of nucleotides sequenced.

To test whether the mutations in the RNA transcript could be attributed to the infidelity of the T7 polymerase during transcript generation, the *in vitro* transcription reaction was repeated using new reagents in the presence and absence of thioredoxin (TRX; section 2.3.10). The resulting DV type 3 E gene transcripts (E^{TRX+} and E^{TRX-}) were subjected to high fidelity RT-PCR, cloning and sequencing using the same methods applied to the Sri Lankan patient samples in this study (AccuScript method; section 2.3.1.2). Sequence contigs were assembled and the clones aligned as both nucleotide and amino acid sequences (section 2.2.4). In total, 23 clones were analysed; 12 and 11 clones generated using E^{TRX+} and E^{TRX-} respectively. There were no mutations in any of the clones generated; all of the clones matched the consensus sequence (data not shown). This consensus sequence also matched those obtained during the AccuScript versus Expand method comparison. Although there was no difference in the numbers of mutations between E^{TRX+} and E^{TRX-}, no mutations were found in clones generated from either transcript. It is hypothesised that the source of the high number of mutations observed during the AccuScript versus Expand method comparison was the RNA transcript as previously suggested, and that using new reagents for the *in vitro* transcription reaction resulted in the generation of a homogenous RNA transcript population. An error frequency for mutations introduced by the AccuScript RT-PCR of less than 0.003% was calculated, based on 33,644 nucleotides sequenced and no mutations were found (1/33,645 = 0.003%).

3.8 Phylogenetic analysis of dengue virus quasispecies envelope gene and glycoprotein sequences

To investigate sequence variation between the DVs from the dengue patient samples, multiple sequence alignments were constructed and phylogenetic analyses performed (section 2.2.4). Alignments were constructed using published DV sequences of known genotype and sequences from Sri Lanka (Table 3.14) to identify the genotypes present within the DV types circulating in Sri Lanka at the time of the outbreak. These sequences were obtained from NCBI (section 2.2.1). Consensus E gene sequences for the dengue samples from this study were included in their respective multiple sequence alignments (separately by DV type) alongside E gene sequences for clones containing mutations. This was to identify whether any of the mutations were present in the clones as a result of recombination with a co-infecting genotype. Evidence of recombination would have been observed if any of the mutant clones displayed a vastly different percentage sequence identity, or a shift in phylogeny compared to their respective consensus sequences.

The percentage sequence identity between each of the strains was calculated from the nucleotide sequence alignments (Appendix I). The two DV type 2 patient sample (R107A and R259A) consensus sequences share 99.5% sequence identity and are identical in protein sequence. The R107A and R259A clones containing mutations share 99.9% sequence identity with their respective consensus sequences. Sequence identities between all of the DV type 2 samples used to construct the alignment were between 80.7 and 99.9%, with the lowest sequence identities observed with the sylvatic strains P8-1407 and Sen70-DAKHD10674. Sequence identities between the DV type 2 sequences and the DV type 1 sequence included in the alignment were around 66% as would be expected because they are different viruses.

The DV type 4 patient sample (R282A) clones containing mutations share 99.9% sequence identity with the sample consensus sequence. Sequence identities between all of the DV type 4 samples used to construct the alignment were between 91.7 and 99.9%, with the lowest sequence identities observed with the genotype *II* strain El Salvador 1994. Sequence identities between the DV type 4 sequences and the DV type

Table 3.14. Published dengue virus envelope gene sequences used for phylogenetic analysis.

NCBI accession no.	Strain	DV type	Genotype	Reference
X54319	Tonga 1974	2	American	(Twiddy *et al.*, 2002)
AF264054	Puerto Rico 1969	2	American	"
L10053	Trinidad 1953	2	American	"
L10047	Seychelles 1977 (Sey-42)	2	Cosmopolitan	"
AF10374	India 1994 (CAMR 10)	2	Cosmopolitan	"
AF410376	Far East 1998 (CAMR 13)	2	Cosmopolitan	"
M29095	New Guinea C	2	Asian 2	"
AF410364	Vietnam 1998 (CTD211)	2	Asian 2	"
AF295697	Philippines (DOH-005)	2	Asian 2	"
AF295698	Philippines (DOH-034)	2	Asian 2	"
AF195040	Thailand (D91-533)	2	American/Asian	"
AF163096	PTCOL96	2	American/Asian	"
AF410350	Vietnam 1997 (CTD20)	2	American/Asian	"
AF410367	Vietnam 1998 (CTD226)	2	American/Asian	"
AF022436	ThNH-52/93	2	Asian 1	"
AF022434	ThNH-7/93	2	Asian 1	"
AF022437	ThNH-p11/93	2	Asian 1	"
AF264053	Thailand 1980 (PUO-312)	2	Asian 1	"
AF231717	P8-1407	2	Sylvatic	"
AF231720	Sen70-DAKHD10674	2	Sylvatic	"
AF410372	Sri Lanka 1994 (CAMR 7)	2	Cosmopolitan	"
AB194883	Sri Lanka 2004	2	Unknown	
NC_001475	Sri Lanka 2000	3	Unknown	
AF547234	Sri Lanka 1993	3	IIIb	(Messer *et al.*, 2003a)
AY099336	Sri Lanka 2000a	3	Unknown	
DQ518679	Sri Lanka 1999	3	Unknown	
L11438	Sri Lanka 1991	3	IIIb	(Lanciotti *et al.*, 1994)
AF547230	Sri Lanka 1989	3	IIIb	(Messer *et al.*, 2003a)
AF547242	Sri Lanka 1997	3	IIIb	"
AF547232	Sri Lanka 1989b	3	IIIb	"
AF547243	Sri Lanka 1998	3	IIIb	"
AF547235	Sri Lanka 1994	3	IIIb	"
L11429	Malaysia 1974 (1300)	3	I	(Lanciotti *et al.*, 1994)
L11437	Sri Lanka 1989c	3	IIIa	"
AF547229	Sri Lanka 1984	3	IIIa	(Messer *et al.*, 2003a)
AF547225	Sri Lanka 1983	3	IIIa	"
AF547226	Sri Lanka 1983a	3	IIIa	"
AF547228	Sri Lanka 1983c	3	IIIa	"
AF547241	Sri Lanka 1985	3	IIIa	"
AF547231	Sri Lanka 1989a	3	IIIa	"
L11436	Sri Lanka 1985a	3	IIIa	(Lanciotti *et al.*, 1994)
AF547227	Sri Lanka 1983b	3	IIIa	(Messer *et al.*, 2003a)
L11431	Sri Lanka 1981	3	IIIa	(Lanciotti *et al.*, 1994)
L11434	Puerto Rico 1977 (1340)	3	IV	"
L11442	Thailand 1987 (MK315)	3	II	"
U18437	Sri Lanka 1978a	4	I	(Lanciotti, Gubler, and Trent, 1997)
U18442	Thailand 1984	4	I	
AY550909	Sri Lanka 1978	4	Unknown	
U18441	Thailand 1978	4	I	(Lanciotti, Gubler, and Trent, 1997)
U18433	Philippines 1956	4	I	
U18434	Philippines 1964	4	I	"
U18435	Philippines 1984	4	I	"
U18426	El Salvador 1983	4	II	"
U18432	New Caledonia 1984	4	II	"
U18431	Mexico 1984	4	II	"
U18425	Brazil 1982	4	II	"
U18436	Puerto Rico 1986	4	II	"
U18439	Tahiti 1985	4	II	"
U18438	Tahiti 1979	4	II	"
U18427	El Salvador 1994	4	II	"
U18429	Indonesia 1976	4	II	"
U18430	Indonesia 1977	4	II	"
U18428	Indonesia 1973	4	II	"
U18440	Thailand 1963	4	I	"

All sequences were obtained from NCBI (section 2.2.1). Genotype information was obtained from the publications listed.

1 sequence included in the alignment were around 63% as would be expected because they are different viruses.

The three DV type 3 patient sample (R097A, R203A and R261A) consensus sequences share a sequence identity of 99.1 to 99.6%. R097A and R203A each differ from the R261A sequence by one amino acid. The R097A and R203A samples share greater sequence identity with each other than to the R261A sample. The clones containing mutations share 99.9% sequence identity with their respective consensus sequences. R261A clones 4, 29, 34, 37 and 46 share 100% sequence identity, as do clones 49, 52 and 60. Sequence identities between all of the DV type 3 samples used to construct the alignment were between 92.6 and 100%, with the lowest sequence identities observed with the genotype *IV* strain Puerto Rico 1977.

Sequence analysis of the RT-PCR amplicons from the DV type 3 patient samples that were not cloned (R004A, R114A, R197A, R232A, R254A, R255A, R299A and R334A) showed that most of these samples share 99% or more sequence identity with the samples that were cloned. The exceptions to this were samples R197A and R254A, which had average sequence identities of 94 and 47% respectively. These were lower because the DV E gene sequences for these samples were incomplete (Appendix C). Compensating for the missing portions of sequence (5% and 53% for R197A and R254A respectively) increases the sequence identities to over 99% for both samples. Samples R299A and R114A share 99.9% sequence identity with the cloned sample R097A; with the R097A and R299A samples having an identical protein sequence and sample R114A differing by one amino acid. All three of these samples were from severe dengue cases (Table 3.3). Sample R004A shares 99.9% sequence identity with the cloned sample R203A and the protein sequences were identical. Both of these samples were from severe dengue cases. Samples R197A, R232A, R254A, R255A and R334A share the highest sequence identity with the cloned sample R261A (94.3, 99.5, 46.9, 99.4, and 99.5% respectively), which was from a mild dengue case. Samples R232A, R255A and R334A were also from mild dengue cases, but samples R197A and R254A were from severe dengue cases.

Phylogenetic trees were constructed based on the E gene sequence alignments (Figures 3.10 to 3.12; alignments not shown). Bootstrapping statistical analyses of 1000 replicates was used to assess the integrity of the phylogenetic trees. Percentage values are shown at the major nodes of each tree and values of 70% and above are considered acceptable, because 70% of the 1000 generated trees contained that clade.

Published sequence data from Sri Lankan DV type 2 strains has shown only strains from the Cosmopolitan genotype or Asian genotype 2 (Twiddy *et al*., 2002). Both of the DV type 2 E gene samples, and all of the mutant clones from this study cluster together within the Cosmopolitan genotype (Figure 3.10). The genotype of the Sri Lanka 2004 strain was previously undetermined and the results of this study show that it also belongs to the Cosmopolitan genotype. This genotype contains other strains from Sri Lanka as well as strains from India, Southeast Asia, Africa, the Middle East and Australia. Sri Lankan DV type 4 sequences published to date have shown only strains from genotype *I* (Lanciotti, Gubler, and Trent, 1997). The DV type 4 E gene sample from this study, and all of the mutant clones cluster together within genotype *I* as expected (Figure 3.11). The genotype of the Sri Lanka 1978 strain was previously undetermined and the results of this study show that it also belongs to genotype *I*, alongside other Sri Lankan strains, and strains from Thailand and the Philippines..

Previously published E gene sequences from Sri Lankan DV type 3 strains have shown no strains from genotypes *I*, *II* or *IV*, only genotype *III* (Messer *et al*., 2003a). These genotype *III* strains cluster into two distinct groups: *IIIa* and *IIIb.* Those strains that comprise genotype *IIIa* are all from outbreaks during or prior to 1989, which were not associated with epidemics of severe dengue. Those that form genotype *IIIb* are all from 1989 or later, and have been associated with epidemics of severe dengue (Lanciotti *et al*., 1994; Messer *et al*., 2003a). It has been suggested that this genetic shift may have increased the virulence of the Sri Lankan strains, and be responsible for the increase in cases of severe disease seen on the island after 1988 (Lanciotti *et al*., 1994). A similar DV type 3 genotype *IIIb* clade replacement has been suggested as the cause of a further increase in severe dengue in Sri Lanka since 1999 (Kanakaratne *et*

al., 2009). The DV type 3 E gene samples from this study were identified as genotype *IIIb* strains, and cluster together with their respective mutant clones within the clade containing the more recent samples (since 1999), as expected (Figure 3.12). The genotypes of the Sri Lanka 1999 and 2000 strains were previously undetermined and the results of this study show that they also belong to genotype *IIIb*.

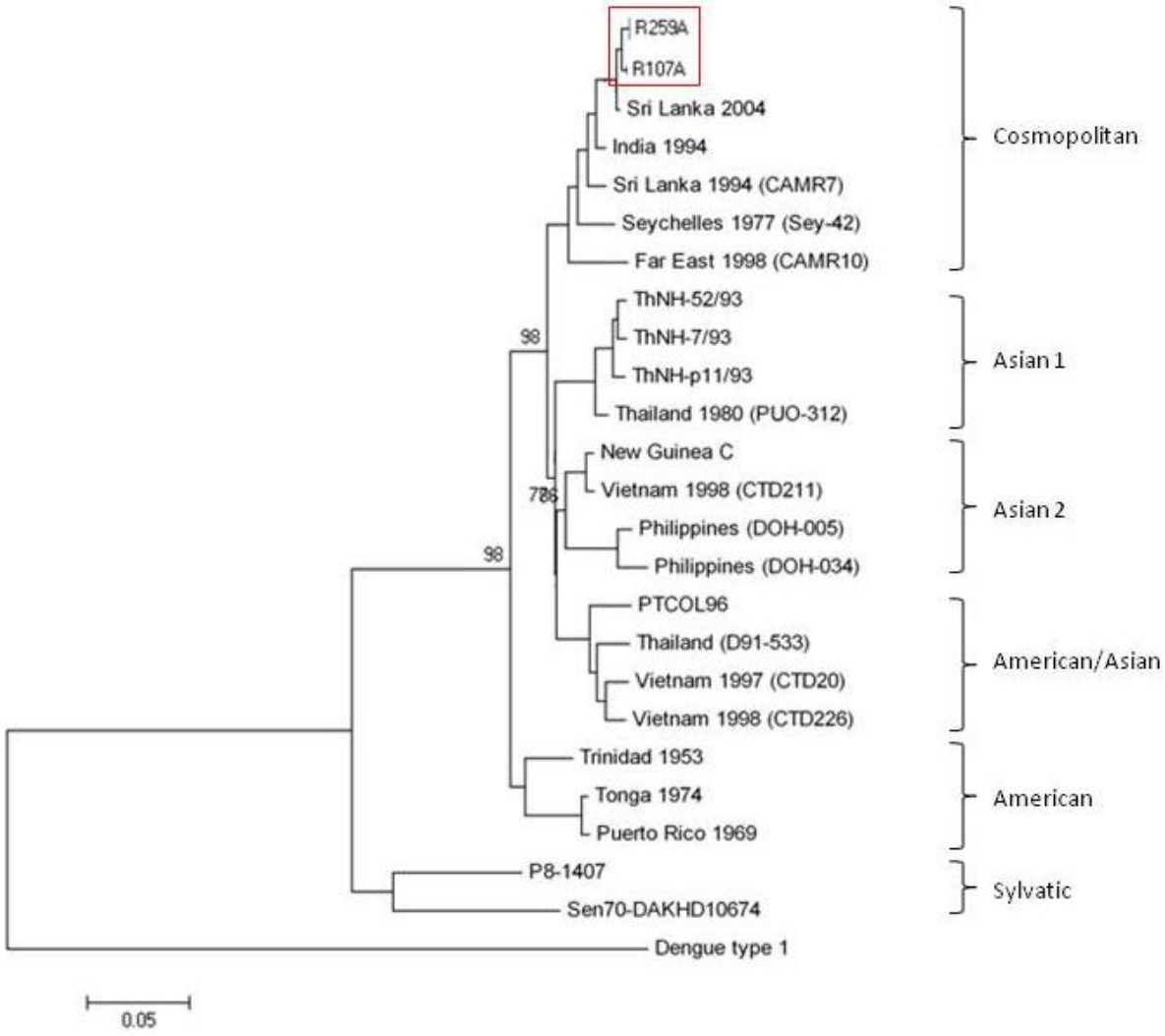


Figure 3.10. Dengue virus type 2 complete envelope gene phylogenetic tree.

Samples from this study are highlighted with a red box. The scale represents the evolutionary distance in units of the number of base substitutions per site.

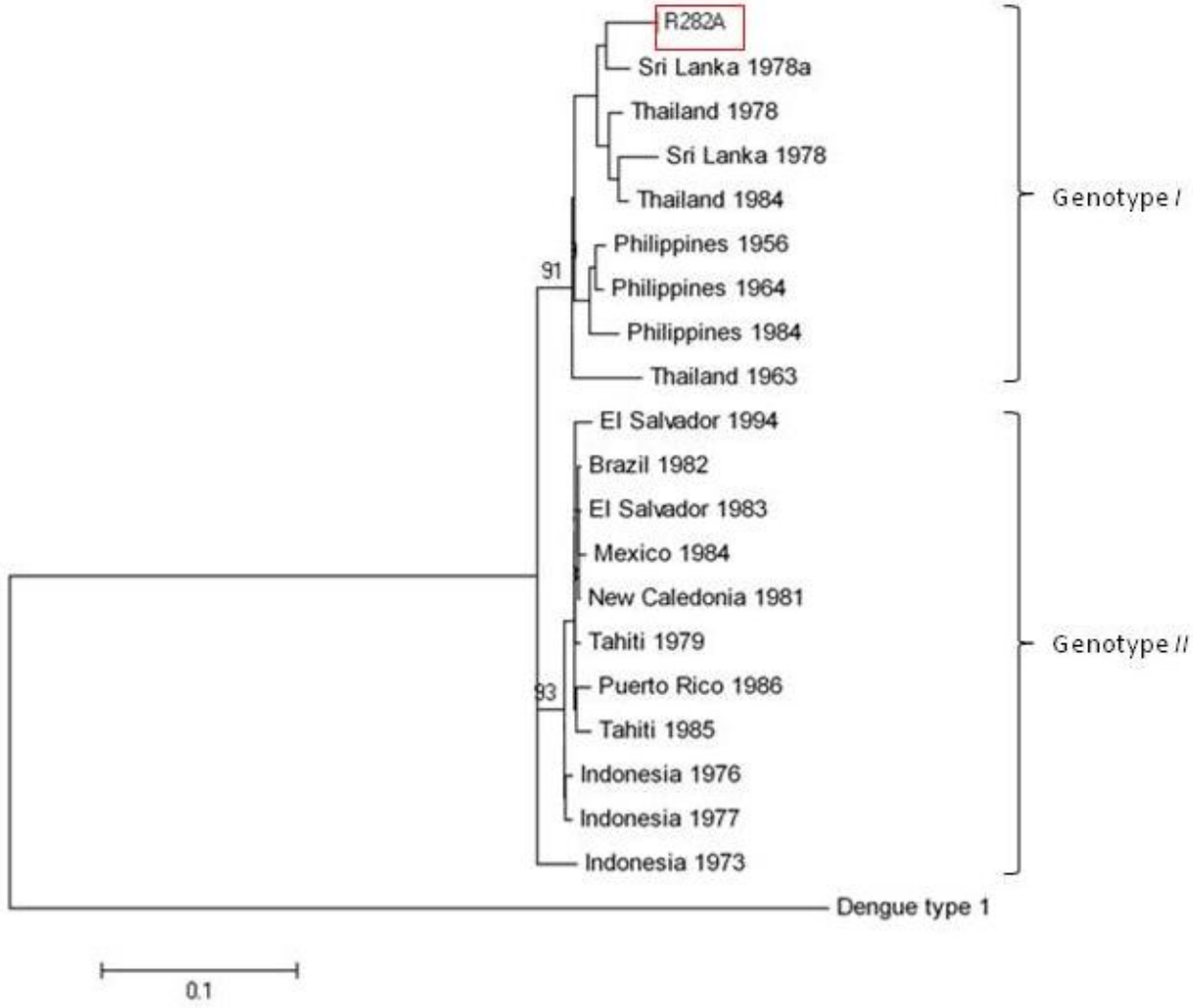


Figure 3.11. Dengue virus type 4 complete envelope gene phylogenetic tree.

The sample from this study is highlighted with a red box. The scale represents the evolutionary distance in units of the number of base substitutions per site.

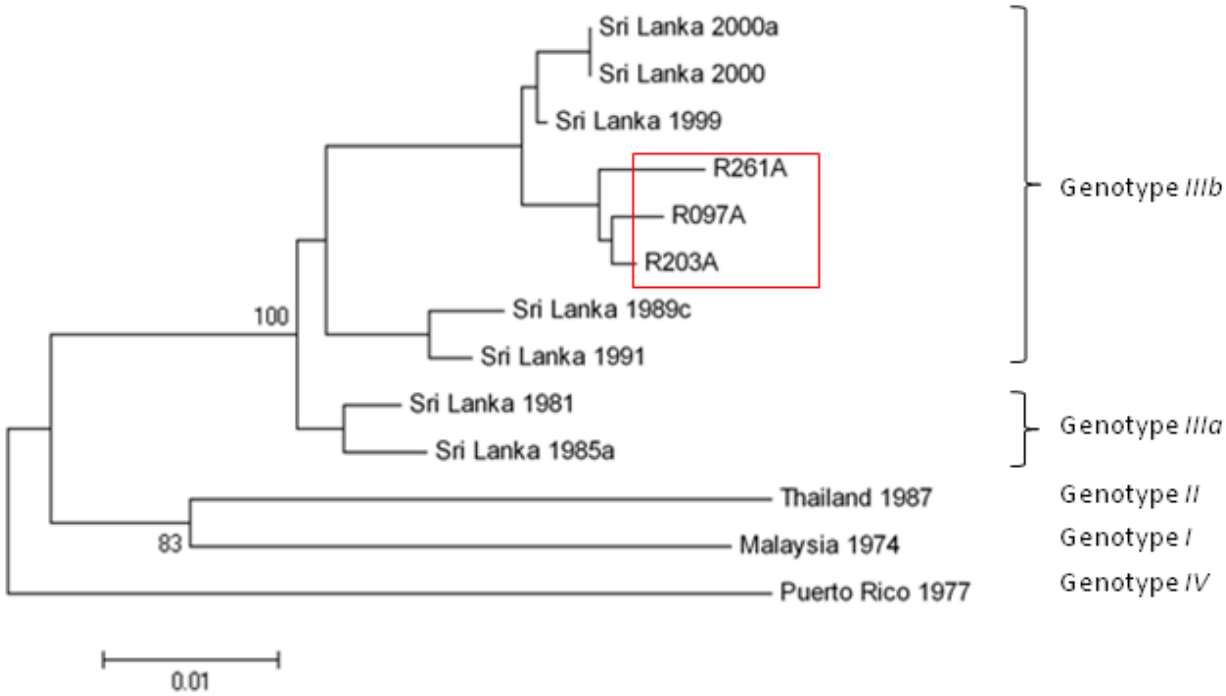


Figure 3.12. Dengue virus type 3 complete envelope gene phylogenetic tree.

Samples from this study are highlighted with a red box. The scale represents the evolutionary distance in units of the number of base substitutions per site.

For the DV type 3 samples, a second tree (Figure 3.13) was constructed to incorporate an increased number of Sri Lankan DV sequences (from Messer *et al*., 2003a) using a 207 nucleotide region from the 5' end of the E gene. In this tree the genotype *IIIb* strains cluster in two temporally distributed clades as with the complete E gene tree (Figure 3.12). There is no indication from these E gene data that any of the mutant DV clones result from recombination events between co-infecting genotypes, as the samples and mutant clones from this study all cluster as expected within the clade containing the more recent samples. If there had been recombinant E genes present, a difference in phylogeny would have been observed between the consensus and mutant sequences across the complete and partial E gene trees. However, compared to the complete E gene tree, in this region of the sequence the R261A sample is less closely related to the R097A and R203A strains from this study as well as the Sri Lanka 2000 and 2000a strains. Also, in the DV type 3 complete E gene phylogram (Figure 3.12), the genotype *I* strain Malaysia 1974 appeared most closely related to the genotype *II* strain Thailand 1987, but in the DV type 3 partial E gene phylogram (Figure

3.13) the genotype *I* strain appears more closely related to the genotype *IIIb* strains. The Sri Lanka 1989c strain also shifts phylogeny between the two phylogenetic trees. In the complete E gene tree (Figure 3.12) it clusters with Sri Lanka 1991 in genotype *IIIb*, whereas in the partial E gene tree (Figure 3.13) it appears more closely related to Sri Lanka 1984 in genotype *IIIa*. A shift in phylogeny when different regions of a sequence are used for the analysis is indicative of recombination. In the partial E gene tree (Figure 3.13), the presence of a DV strain from 1993 within the clade containing the more recent samples (since 1999) is a strong indication that the clade is likely derived from strains indigenous to Sri Lanka rather than being imported (Kanakaratne *et al*., 2009).

The bootstrap values for all of the complete E gene phylogenetic trees were above 70% at the major nodes, which provides statistical confirmation that the trees are accurate. The bootstrap values for the DV type 3 partial E tree (Figure 3.13) are below 70%, but this is not uncommon for trees based on shorter sequence lengths and would be expected given that DVs are known to undergo recombination within DV types. This does emphasise the importance of using larger regions of sequence for phylogenetic analyses as trees based on shorter sequence lengths are difficult to validate statistically. Ideally, these analyses would be based on complete genome sequences, but that was not feasible for the present study; using the complete E gene sequences yielded an appropriate level of statistical confidence.

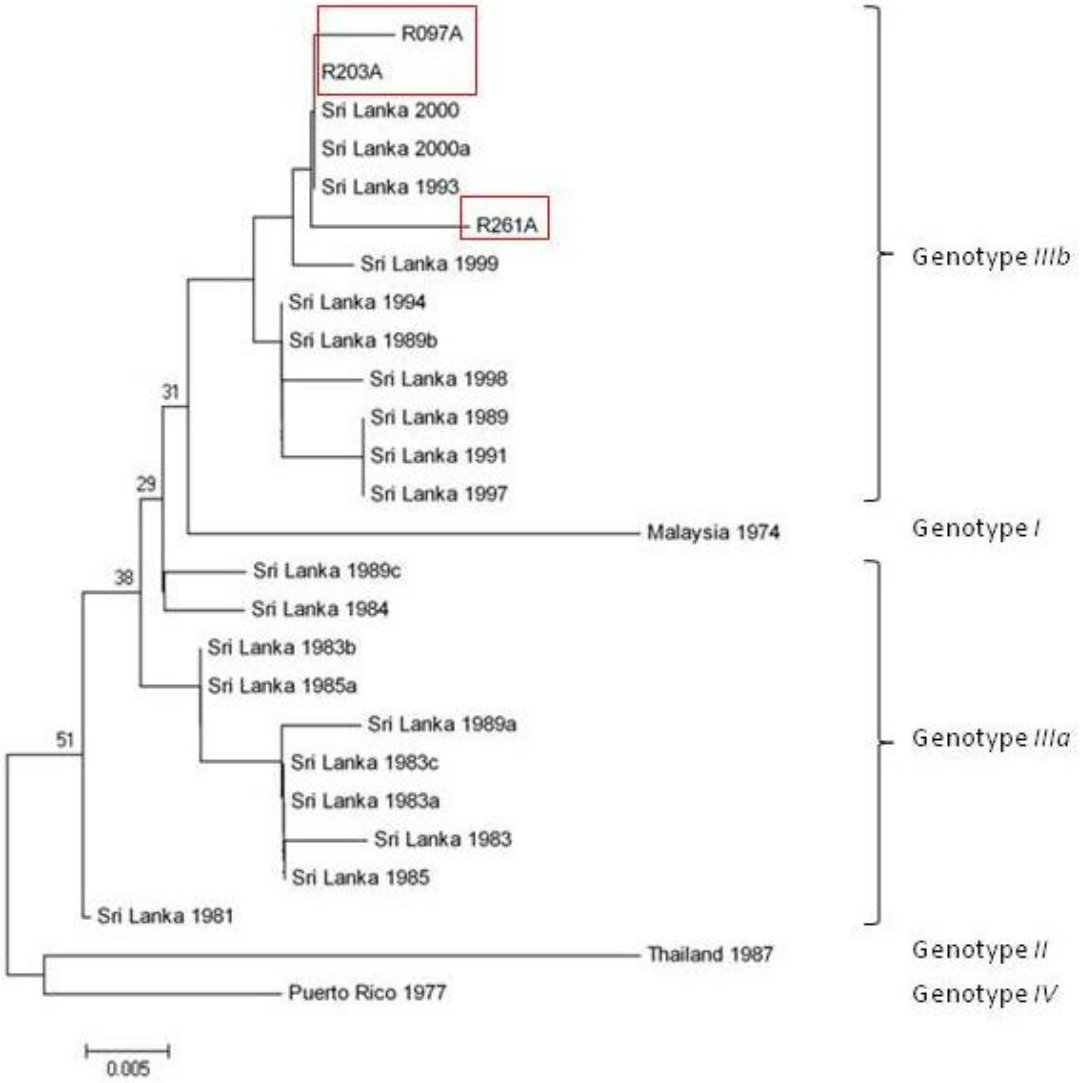


Figure 3.13. Dengue virus type 3 partial (207 nucleotides at the 5' end) envelope gene phylogenetic tree.

Samples from this study are highlighted with red boxes. The scale represents the evolutionary distance in units of the number of base substitutions per site.

3.9 Linear and three-dimensional mapping of dengue virus quasispecies envelope glycoprotein amino acid substitutions

All of the amino acid substitutions resulting from nonsynonymous E gene mutations (identified by sequence analysis in section 3.6; detailed in Appendices C to E) were mapped to the DV E linear and three-dimensional forms to identify those with the potential to affect pathogenicity. This is, to the best of my knowledge, the first report where mapping of DV quasispecies mutations to the structure of E has been conducted. A linear map of DV E was constructed using the flavivirus E antigenic structure based on TBEV (Figure 3.14, panel A; Mandl *et al*., 1989) and the crystal structures of TBEV and DV (Figure 3.14, panels B and C respectively; Modis *et al*., 2003; Rey *et al*., 1995). The consensus protein sequence for each patient sample was applied to the crystal structure of the appropriate DV E using DeepView Swiss Pdb-Viewer (section 2.2.6 Arnold *et al*., 2006) to construct a three-dimensional rendering of the protein, and the locations of the amino acid substitutions were highlighted (Figure 3.15). Substituted amino acids located on the surface of E, and therefore potentially involved in interactions with other oligomers, antibodies or host cell receptors were identified. Amino acid substitutions that were within 4Å of the other E protein chain within the dimer were also highlighted, as these residues could be involved in oligomer assembly. Similarly, substituted residues within 4Å of the putative highly-sulphated heparan sulphate (HSHS) binding site thought to be involved in virus-cell interactions (Chen *et al*., 1997; Thullier *et al*., 2001), or within 4Å of the fusion peptide, were also noted (Table 3.15).

From a linear perspective, the quasispecies mutations cause amino acid substitutions that occur throughout DV E, from residue E29 through to E455 (Figure 3.15, panel D). Six quasispecies variant clones were identified with amino acid substitutions located in structural domain *I*, fifteen clones with amino acid substitutions in domain *II*, seven clones with amino acid substitutions in domain *III* and two clones with amino acid substitutions in the transmembrane domain.

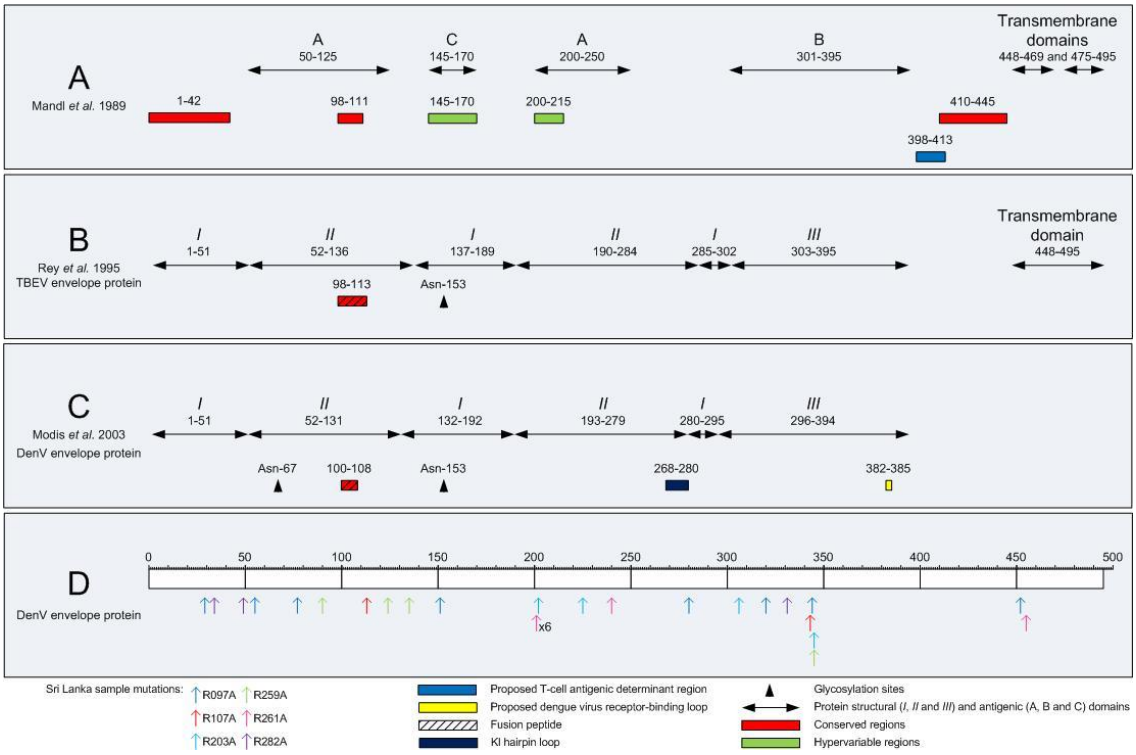
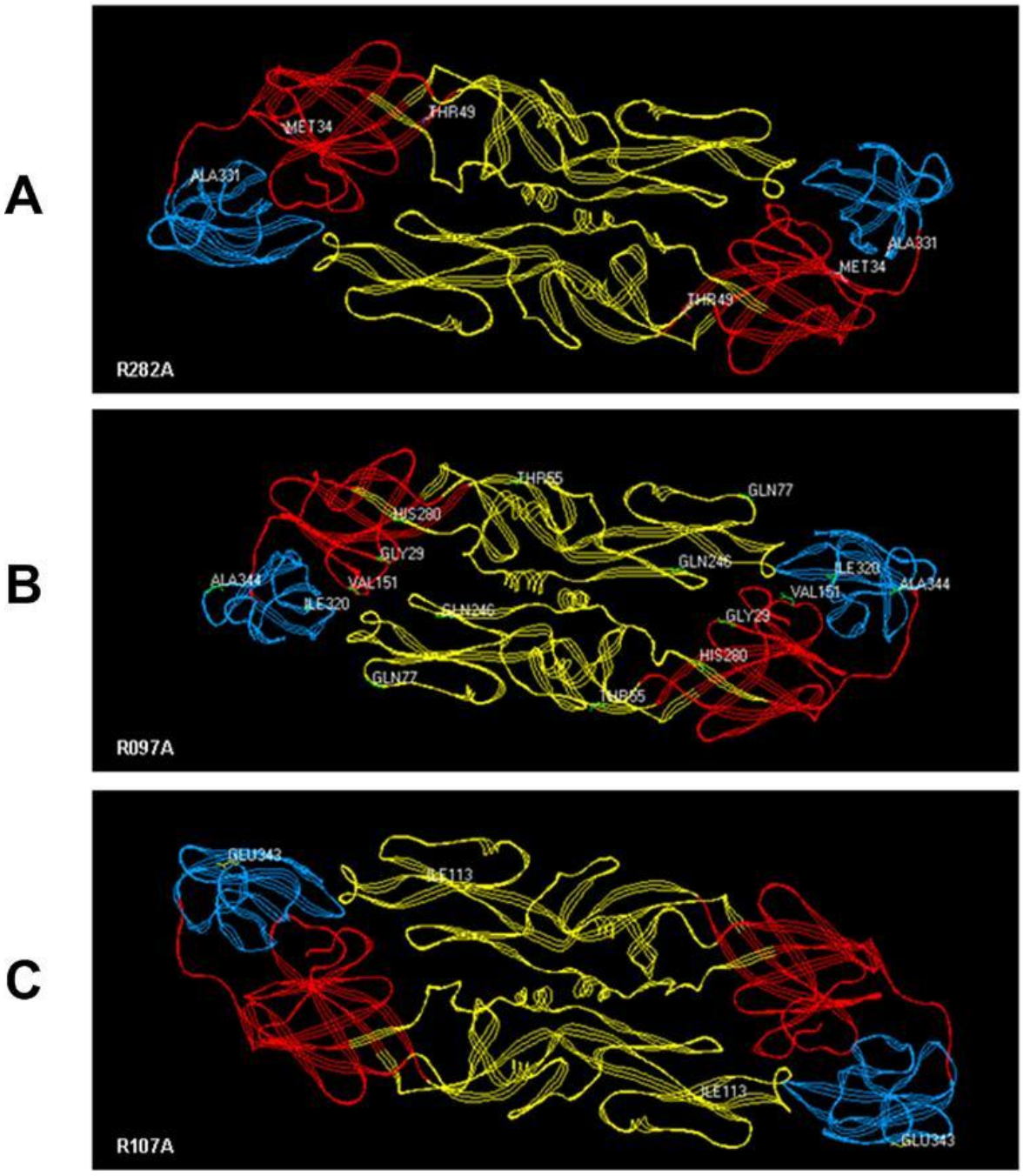


Figure 3.14. Linear schematic representation of the dengue virus envelope glycoprotein, detailing important structural features and the location of Sri Lankan patient sample quasispecies amino acid substitutions.

(A) Antigenic structure of flavivirus E using TBEV as a model (Mandl *et al*., 1989). Three antigenic domains (A, B and C) and two transmembrane domains are shown. Regions conserved amongst most flaviviruses are shown in red. Hypervariable regions are shown in green. The location of a proposed T-cell antigenic determinant is shown in blue. (B) TBEV E structural domains based on crystal structure determination (Rey *et al*., 1995). Three structural domains (*I*, *II* and *III*) are shown, corresponding to the previously identified antigenic domains A, C and B respectively (see panel A). A glycosylation site common to most flaviviruses is marked with a triangle. A region highly conserved in flaviviruses and implicated in membrane fusion is shown in red with black stripes. (C) DV (DenV) type 2 E structural domains based on crystal structure determination (Modis *et al*., 2003). Three structural domains (*I, II* and *III*) are shown, corresponding largely with those identified for TBEV (panel B). Two glycosylation sites are marked with triangles. One is common to most flaviviruses (Asn-153) and the other is specific to DVs (Asn-67). The highly conserved fusion peptide is shown in red with black stripes. A proposed DV receptor-binding loop not present in tick-borne flaviviruses is shown in yellow. The *kl* hairpin loop at the domain *I*/*II* interface (shown in dark blue) has been implicated in oligomer rearrangements during fusion. (D) DV (DenV) E showing the locations of the Sri Lankan patient sample amino acid substitutions (Appendix E).


A
THR49
MET34
ALA331
ALA331
MET34
THR49
R282A
B
THR55
GLN77
HIS280
GLY29
GLN246
ILE320
ALA344
VAL151
ALA344
ILE320
VAL151
GLN246
GLY23
HIS280
GLN77
THR55
R097A
C
GLU343
ILE113
ILE113
GLU343
R107A

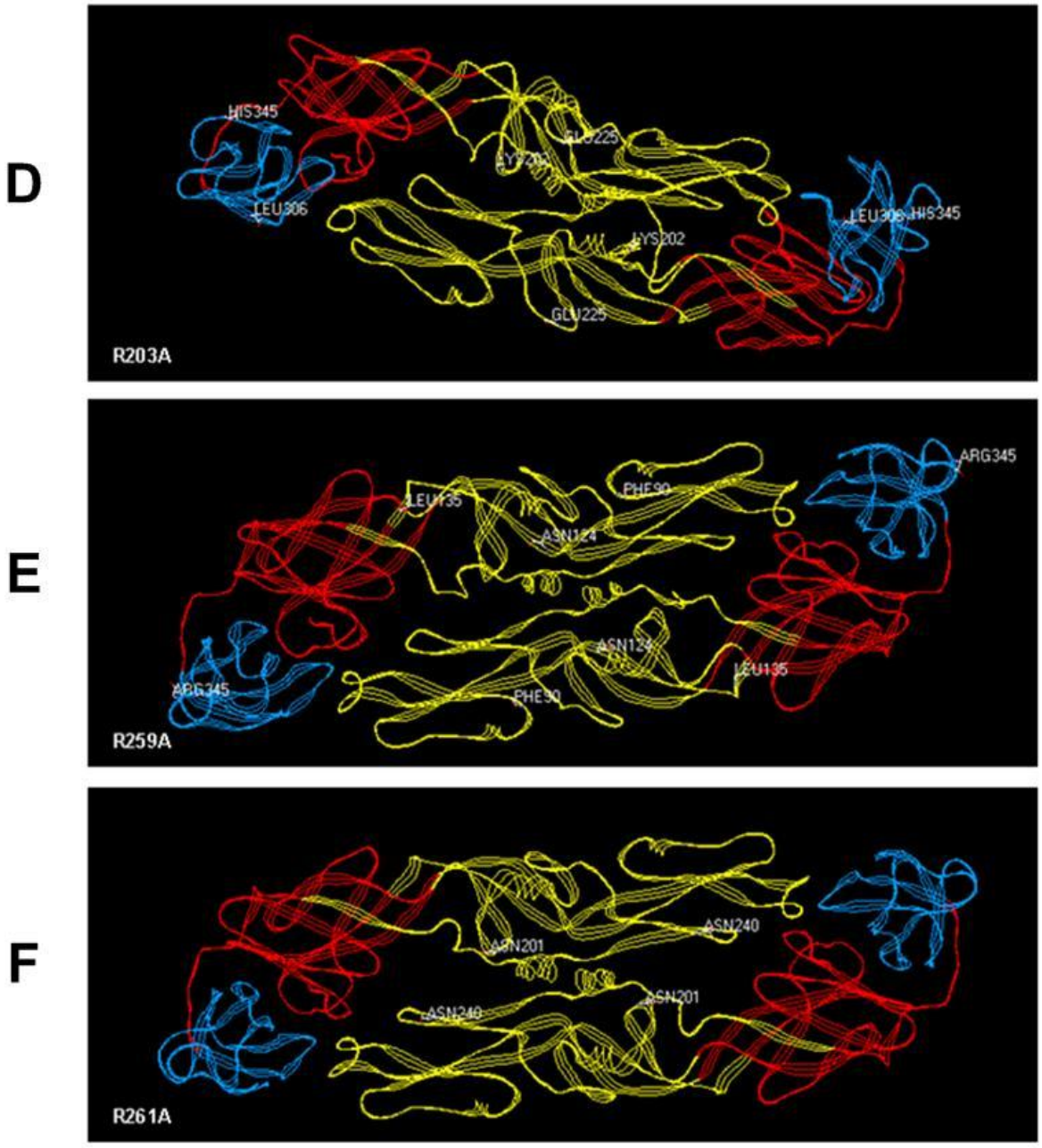


Figure 3.15. Three-dimensional structure of the dengue virus envelope glycoprotein detailing the locations of the Sri Lankan patient sample quasispecies amino acid substitutions.

Panel A shows the three amino acid substitutions for the R282A DV type 4 patient sample (Appendix E, Table 6). Panel B shows the nine amino acid substitutions for the R097A DV type 3 patient sample (Appendix E, Table 1). Panel C shows the two amino acid substitutions for the R107A DV type 2 patient sample (Appendix E, Table 2). Panel D shows the four amino acid substitutions for the R203A DV type 3 patient sample (Appendix E, Table 3). Panel E shows the four amino acid substitutions for the R259A DV type 2 patient sample (Appendix E, Table 4). Panel F shows the two amino acid substitutions for the R261A DV type 3 patient sample (Appendix E, Table 5). Images were constructed using DeepView Swiss PDB-viewer (section 2.2.6).

As expected, none of the amino acid substitutions were observed in positions occupied by the 12 conserved cysteine residues required for protein structure (Nowak and Wengler, 1987), or within the fusion peptide at residues 100-108 (Modis *et al*., 2003). None of the amino acid substitutions were located at the N-glycosylation sites at residues Asn-67 or Asn-153 (Mondotte *et al*., 2007). However, one of the amino acid substitutions (R107A clone 22, position E113) occurs within the highly conserved region surrounding the fusion peptide (residues 98-113; Rey *et al*., 1995), and two of the amino acid substitutions (from patient samples R097A and R282A; positions E29 and E34 respectively) occur within the conserved region at residues 1-42 (Mandl *et al*., 1989). Two amino acid substitutions (R097A E452 and R261A E455) map to the transmembrane domain at E position E450 to E472, which functions as the membrane anchor for E (Allison *et al*., 1999). This region is required for the incorporation of E into virus particles and stabilisation of interactions between prM and E in immature virus particles (Allison *et al*., 1999). Three of the R097A amino acid substitutions (E29, E55 and E77) and one of the R107A amino acid substitutions (E113) occur at computer-predicted potential post-translational modification sites for protein kinase C phosphorylation, casein kinase II phosphorylation, or N-myristoylation; although the significance of these modifications is unknown (Amin *et al*., 2010). Six E gene clones from patient sample R261A carry the same conservative amino acid substitution at residue E201 and patient sample R203A has one clone with a conservative amino acid substitution at residue E202. These occur within a region of domain *II* (E200-215) that has been suggested to be hypervariable amongst flaviviruses (Mandl *et al*., 1989). Three-dimensional modelling identified the E201 amino acid substitution as surface accessible and the E202 amino acid substitution as within 4Å of the other E protein chain that makes up the dimer (Table 3.15). Therefore, these amino acid substitutions may be of importance in oligomer assembly or oligomeric rearrangements during virus fusion with cell membranes. One clone from patient sample R097A contains an amino acid substitution (E151) in another hypervariable region of domain *I* (E145-170; Mandl *et al*., 1989). This amino acid substitution was also identified as being within 4Å of the other E protein chain that makes up the dimer.

Three-dimensional mapping of the amino acid substitutions identified seven of these (R097A, E77 and E280; R203A, E225 and E345; R261A, E201; R107A, E343; R259A, E345) to be situated on the surface of E (surface accessible, defined as 30% or more of the residue surface is accessible to solvent; Table 3.15). These amino acid substitutions were from five different patient samples, across DV types 2 and 3, and may be involved in interactions with the DV prM or M protein, the other E protein chain that makes up the dimer, other E oligomers, antibodies or host cell receptors.

Mutations leading to amino acid substitutions in the flavivirus E that alter the properties of the virus (e.g. virulence) tend to cluster at three sites: the distal face of domain *III*, the base of domain *II*, and the domain *I*/*III* interface with the domain *II* conserved region at residues 98-113 (section 1.5.1; Rey *et al*., 1995). Three-dimensional mapping showed that several of the amino acid substitutions identified in the dengue patient sample quasispecies variants also map to these sites (Table 3.15).

Mutations leading to amino acid substitutions at the base of domain *II*, or at the domain *I*/*III* interface with the domain *II* conserved region at residues 98-113 are thought to affect virulence by impacting upon the conformational rearrangements that lead to fusion of virus and host cell membranes (Rey *et al*., 1995). Three amino acid substitutions from two patient samples (R097A E151 and E246; R107A E343) were identified by three-dimensional mapping as within 4Å of the fusion domain.

One of the clones isolated from patient sample R097A (clone 43) has a mutation causing an amino acid substitution at DV E position E280, which is part of the *kl* hairpin loop at the domain *I*/*II* interface, which is implicated in conformational rearrangements during fusion (Modis *et al*., 2003). Amino acid substitutions at DV E structural domain interfaces have also been proposed to affect oligomer assembly (Duarte dos Santos *et al*., 2000). Amino acid substitutions on the distal face of domain *III* can affect virulence or cell tropism, and it has been suggested that this is due to interference with cell attachment as domain *III* is thought to contain the flavivirus receptor-binding motif (Crill and Roehrig, 2001). One of the clones from patient sample R203A (clone 21) has a mutation leading to an amino acid substitution at DV E

position E306, which is suggested to form part of a binding site (E306-314) for HSHS molecules, which are present on the surface of cells targeted by the virus and appear to mediate infectivity (Chen *et al.*, 1997; Thullier *et al.*, 2001). The amino acid substitution at E306 therefore has the potential to directly affect virus attachment to the host cell. A clone from patient sample R097A (clone 13) has a mutation leading to an amino acid substitution at E position E320, which was determined by three-dimensional mapping to be within 4Å of the HSHS binding site (Table 3.15). It is possible that disruption of the binding site by this amino acid substitution could either enhance or attenuate virulence by altering virus infectivity.

Table 3.15. Summary of dengue virus envelope glycoprotein amino acid substitution locations identified through linear and three-dimensional modelling.

Sample	Mutated residue	Mutation location	Surface accessible	Within 4Å of another protein chain	Within 4Å of the HSHS binding site	Within 4Å of the fusion domain
R097A	E29	Domain *I* conserved region at E1-42	No	No	No	No
	E55	Base of domain *II* at the interface with domain *I*	No	No	No	No
	E77$_f$	Domain *II*	Yes	No	No	No
	E151$_f$	Domain *I* at the interface of domain *III* and the domain *II* conserved region at E98-113 on the other monomer	No	Yes	No	Yes
	E246$_f$	Domain *II*	No	No	No	Yes
	E280	Base of domain *II* at the interface with domain *I*; within the *kl* hairpin	Yes	No	No	No
	E320	Domain *III*	No	No	Yes	No
	E344	Distal face of domain *III*	No	No	No	No
	E452	Transmembrane domain	No	No	No	No
R203A	E202	Base of domain *II* at the interface with domain *I*, within the domain *II* proposed hypervariable region at E200-215	No	Yes	No	No
	E225$_f$	Domain *II*	Yes	No	No	No
	E306$_f$	Distal face of domain *III*	No	No	Yes	No
	E345	Distal face of domain *III*	Yes	No	No	No
R261A	E201	Base of domain *II* at the interface with domain *I*, within the domain *II* proposed hypervariable region at E200-215	Yes	No	No	No
	E240	Domain *II*	No	No	No	No
	E455$_f$	Transmembrane domain	No	No	No	No
R107A	E113	Domain *II* within the conserved region at 98-113	No	No	No	No
	E343$_f$	Distal face of domain *III*	Yes	No	No	Yes
R259A	E90	Domain *II*	No	No	No	No
	E124$_f$	Domain *II*	No	No	No	No
	E135$_f$	Base of domain *II* at the interface with domain *I*	No	No	No	No
	E345$_f$	Distal face of domain *III*	Yes	No	No	No
R282A	E34	Domain *I* conserved region at E1-42	No	No	No	No
	E49	Base of domain *II* at the interface with domain *I*	No	No	No	No
	E331	Distal face of domain *III*	No	No	No	No

E gene mutations leading to a shift in the reading frame are denoted with $_f$.

3.10 Identification of dengue virus quasispecies envelope genes for recombinant protein production

Quasispecies variants were chosen from the DV E gene clones for recombinant protein production based on the data obtained from the mapping of the mutations (Table 3.15). The main criteria for selecting clones for recombinant E production were that the clones contained nonsynonymous non-frameshift mutations leading to amino acid substitutions with the potential to alter the behaviour of the protein (Table 3.16). Clones representative of the DV quasispecies E gene consensus sequence for each sample were also selected for recombinant E production. These would be used in future studies of E function, as a control for comparison to the quasispecies variant E.

Table 3.16. Summary of quasispecies variants chosen for recombinant protein production.

Sample	Clone	Mutated E protein residue(s)
R097A	4	None. Represents sample consensus.
	13	E320; drastic mutation; domain *III*; within 4Å of the HSHS site
	43	E280; drastic mutation; surface accessible at the base of domain *II* at the interface with domain *I*; within the *kl* hairpin
R203A	2	None. Represents sample consensus.
	54	E202; conservative mutation; domain *II*; within 4Å of the other protein chain E345; drastic mutation; surface accessible on the distal face of domain *III*
R261A	1	None. Represents sample consensus.
	4	E201: conservative mutation; domain *II*; surface accessible
R107A	1	None. Represents sample consensus.
	22	E113; conservative mutation; domain *II*; within the conserved region surrounding the fusion peptide (E98-113)

Clone 4 from DV type 3 sample R097A (R097A_4) represents the consensus E gene sequence for the R097A sample set. Clone 13 (R097A_13) contains a drastic amino acid substitution at protein position E320, which was mapped to within 4Å of the residues of the putative HSHS binding site thought to be involved in virus-cell interactions. This amino acid change involves the substitution of a nonpolar amino acid (isoleucine) with a large hydrophobic side chain, with a polar amino acid (threonine) (Appendix E; Table 1). Isoleucine will usually orient itself to the interior of the folded protein and is important for the correct folding of the protein, whereas threonine tends to be located on the outer surface of the protein. It is possible that this amino acid substitution could disrupt the HSHS binding site and impact on virus infectivity. Clone 43 (R097A_43) contains a drastic amino acid substitution at position

E280, which is situated on the surface of the protein and therefore may be involved in antibody or cell binding. This substituted residue also forms part of the *kl* hairpin loop at the interface between E structural domains *I* and *III*, which is implicated in conformational rearrangements during fusion (Modis *et al*., 2003). Mutations leading to amino acid substitutions within the *kl* hairpin loop are known to cause virus attenuation (Lee, Weir, and Dalgarno, 1997; Monath *et al*., 2002), and amino acid substitutions at the domain *I*/*III* interface with the domain *II* fusion peptide on the other monomer are known to affect virulence, with some changing the pH required to trigger the conformational changes that lead to membrane fusion (Rey *et al*., 1995).

Clone 2 from DV type 3 sample R203A (R203A_2) represents the consensus E gene sequence for the R203A sample set. Clone 54 (R203A_54) contains two amino acid substitutions at E202 and E345 (Appendix E; Table 3). The E202 conservatively substituted residue is within 4Å of the other protein chain that makes up the E dimer and involves the substitution of a basic amino acid (lysine) with a functionally similar basic amino acid (arginine). The difference in side chains may be of importance in oligomer assembly or fusion, as this residue has been proposed to map to a putative hinge region between domains *I* and *II*, which may be important during fusion (Lee, Weir, and Dalgarno, 1997). The substitution of lysine with arginine at E position E202 has been shown to occur during routine serial passage of flaviviruses in mammalian cells, and therefore may represent an adaptation to growth in mammalian cells (Lee, Weir, and Dalgarno, 1997). The E345 drastic amino acid substitution involves the substitution of three amino acids and the insertion of two amino acids (histidine-asparagine-glycine becomes isoleucine-glutamine-isoleucine-serine-serine; Appendix E; Table 3). Incorporating the two extra residues into the protein chain will impact on protein structure. E345 is situated on the surface of the protein at the distal face of domain *III*. This part of the domain is thought to contain the flavivirus receptor-binding motif (Crill and Roehrig, 2001) and amino acid subtitutions in this part of the domain are known to affect virulence and cell tropism, presumably through interference with cell attachment.

Clone 1 from DV type 3 sample R261A (R261A_1) represents the consensus E gene sequence for the R261A sample set. Clone 4 (R261A_4) contains a conservative amino acid substitution at E position E201 which is situated on the surface of the protein. This residue change involves the substitution of a polar amino acid with a long hydrophilic side chain (asparagine), with a polar amino acid (serine; Appendix E; Table 5).

Clone 1 from DV type 2 sample R107A_1 represents the consensus E gene sequence for the R107A sample set. Clone 22 (R107A_22) contains a conservative amino acid substitution at E position E113, which is located in domain *II* within the conserved region (E98-113) surrounding the fusion peptide. This residue change involves the substitution of a nonpolar amino acid with a large hydrophobic side chain (isoleucine), with a nonpolar amino acid with a side chain containing a sulphur atom (methionine; Appendix E; Table 2).

3.11 Discussion

DV E gene quasispecies variation was determined by RT-PCR, cloning and sequencing of the DV E gene population in individual clinical samples as described in previous studies (Craig *et al*., 2003; Lin *et al*., 2004; Wang *et al*., 2002a).

3.11.1 Patient samples used for dengue virus quasispecies envelope gene amplification

Quasispecies populations were studied in clinical samples from dengue patients admitted to hospitals in and around Ragama in Sri Lanka, in 2006. The detailed clinical histories linked to these samples enabled disease severity to be classified as mild or severe based on WHO guidelines. These patient samples were limited to clinically-apparent DV infections from hospitalised patients. Samples from non-hospitalised dengue patients or asymptomatic patients are rarely studied and reported in the literature because it requires surveillance of a prospective cohort in the form of regular testing for DV.

3.11.2 Amplification of dengue virus envelope genes by high-fidelity RT-PCR

A relatively low number (21/95) of DV E genes were amplified from the patient samples by high-fidelity RT-PCR despite them having been confirmed previously by HPA CEPR and AFRIMS as DV RT-PCR positive. For each DV type the RT-PCR primers were designed using an alignment of sequences from the NCBI database which included examples of each genotype (see Appendix A). It was only possible to test the primers against reference strain RNA (section 2.1.3), as RNA from other strains was not available. Every effort was made *in silico* to ensure that the primer sequences matched as many strains as possible, but it is possible that the samples that were not amplified contained strains that were not detected by the primers used in this experiment due to nucleotide mismatches between the primer and template sequences. The samples had also been freeze-thawed for diagnostic use prior to this study, which may have degraded the viral RNA sufficiently to make low concentrations of virus undetectable by the RT-PCR used in this experiment. High-fidelity RT-PCR is not used for diagnostic purposes because the stringency of the high-fidelity enzymes makes it less sensitive and more time-consuming than diagnostic methods. The

diagnostic RT-PCR used by both HPA CEPR and AFRIMS was nested, with the first round of amplification used to amplify a 511bp amplicon from any DV type and the second round of amplification used to amplify type-specific amplicons of varying sizes (119-482bp; Lanciotti *et al*., 1992). The relatively large (up to 1.8kb) DV E gene target in this study requires more intact RNA for amplification, compared to the smaller amplicons in the AFRIMS and HPA CEPR diagnostic nested RT-PCR. In support of this, HPA CEPR detected fewer DV RNA positive samples during diagnostic testing than had been detected previously by AFRIMS using the same method. Real-time RT-PCRs (amplifying a 70bp amplicon) were also performed at HPA CEPR on the same samples, and DV RNA was detected in more samples using this method than using the nested RT-PCR. These observations are consistent with sample degradation after repeated freeze-thaw cycles.

3.11.3 Cloning of dengue virus quasispecies envelope genes

To eliminate bias due to the preferential amplification of certain templates in a single PCR, DNA used for cloning was obtained from two separate PCR reactions per sample. Nine of the twenty-one DV E genes amplified by high-fidelity RT-PCR produced low numbers (less than 10) of positive transformants during cloning. Assessment of quasispecies variation required 96 positive transformants per sample, which was not achievable from these nine high-fidelity RT-PCR-amplified E genes. This was attributed to the comparatively low yield of amplicon from the high-fidelity RT-PCRs performed on these samples (Table 3.4), resulting in a lower concentration of DNA being used for cloning. Quantification of high-fidelity RT-PCR-amplified E genes and standardisation of the amount of DNA used for cloning were not feasible due to the large numbers of samples being processed. This problem of low amplicon yield may have been overcome with the design of primers for a second round of amplification (a nested PCR), but this would potentially have introduced more mutations into the E gene amplicons via the PCR method itself, so the results would not be comparable to those produced using the non-nested amplification method. It was decided to continue using the 12 samples that were cloned and produced sufficient positive transformants to enable assessment of the DV quasispecies E gene variation.

Compared to the DV type 2 and 4 samples, the type 3 sample E gene amplicons proved more difficult to clone in terms of obtaining enough positive transformants to enable the assessment of quasispecies variation. Some cloning reactions yielded low numbers of positive transformants despite the DNA concentration appearing sufficient for cloning, and control reactions confirming that the problems were not due to faulty ligation, transformation or culture components. Attempts were made to optimise cloning efficiency and increase the number of positive transformants. Cloning was attempted repeatedly using the same sample amplicons with varying results each time. Attempts made to optimise cloning efficiency and increase the number of positive transformants were unsuccessful, so between 64 and 96 white colonies were able to be picked for each sample. Each colony was cultured for 24 hours in selective medium before plasmid DNA extraction and verification of the presence of the E gene inserts. Over 95% of the colonies picked from the DV types 2 and 4 samples contained the E gene insert. However, for the DV type 3 clones, the number of white colonies containing the insert ranged from 2% to 99% (Table 3.5). Flavivirus cDNA has been reported to be unstable in *E. coli* (Ward and Davidson, 2008) and differences between cloning of DV types 1 to 4 structural genes have been observed, with cloning of the type 3 genes being more problematic (Chen *et al*., 1995), which is in accordance with the findings of this study. Plasmid instability can manifest with rearrangement or ejection of part or the entire insert, insertion of foreign sequences into the insert, or segregational instability (the production of plasmid-free cells when plasmid-containing cells replicate). This could explain the absence of the DV E gene insert in some of the white colonies. It is thought that either the flavivirus cDNA itself is toxic to *E. coli*, or that spurious transcription occurring from cryptic bacterial promoters internal to the flavivirus structural gene sequences leads to the synthesis of viral proteins toxic to *E. coli* and destabilises the plasmid (Ward and Davidson, 2008; Yamshchikov, Mishin, and Cominelli, 2001). Several putative prokaryotic promoter sequences have been demonstrated within the E gene of a DV type 2 strain (Pu *et al*., 2011).

3.11.4 Sequencing of dengue virus quasispecies envelope genes

To provide enough DV E genes from each sample for quasispecies analysis, it was decided to proceed with sequencing the samples that had been successfully cloned with more than 50% of white colonies containing inserts (Table 3.5). Problems were experienced when sequencing the clones from samples R114A and R197A, as many of the sample reads obtained were of insufficient length, which meant they were either too short to incorporate into contigs, or the contigs assembled did not cover the entire E gene. This was due to an equipment failure, and sequencing could not be repeated as the fault was unable to be rectified. These two samples were not included in subsequent sequence analyses as the E gene sequence data were incomplete.

The reason why insufficient positive transformants were obtained from some of the DV type 3 patients samples remains unclear. Direct sequencing of the RT-PCR amplicons from the DV type 3 patient samples that did not produce sufficient numbers of positive transformants for quasispecies analysis revealed that these samples shared a high level of sequence identity (greater than 99%) with the samples that were successfully cloned. The failure to produce enough positive transformants from these samples therefore cannot be attributed to major sequence differences between the samples.

Of the 299 dengue patient sample E gene clones successfully sequenced and analysed, nine clones contained truncated E gene sequences flanked by vector sequence. These E gene clones were considered artefacts of the cloning process rather than genome defective quasispecies variants. All nine clones containing incomplete E gene sequences were from patients with DV type 3 infections and two common points of E gene truncation were identified at E gene positions E772-3 and E1128-1132. The full-length E gene templates used for cloning were extracted from appropriately sized RT-PCR amplicons, so it is unlikely that these truncated sequences originate from the cDNA insert. It is likely that the portions of E gene sequence past the truncation point were originally present in the transformed *E. coli*, but were not tolerated by subsequent generations and were rejected due to plasmid instability. A search for prokaryote promoter sites in the E gene sequences obtained from the dengue patient

samples using the Berkeley Drosophila Genome Project (BDGP) Neural Network Promoter Predictor (http://fruitfly.org/seq_tools/promoter.html) revealed several putative promoter sequences in the patient sample sequences (Table 3.17). This suggests that spurious transcription occurring from these prokaryote promoter sites may induce the synthesis of part of E downstream of the promoter site. If these proteins are not tolerated by the *E. coli*, for example if they are inherently toxic, the plasmid can be destabilised and part of it ejected, resulting in the E gene truncations observed.

Table 3.17. Predicted prokaryote promoter sites within dengue virus envelope genes from patient samples.

Dengue patient sample	Start	End	Promoter sequence
R097A R203A	312	357	TTGTGGTTTGTTTGGCAAAGGAAGCTTGGTAACATGTGCG<u>A</u>AATTTCAAT
R004A* R114A*	513	558	CATCTTGCCTGAATATGGAACCCTTGGGCTAGAATGCTCA<u>C</u>CACGGACAG
R232A* R255A*	1131	1176	AATTGGAATTGGAGACAACGCCTTGAAAATCAACTGGTAC<u>A</u>AGAAGGGGA
R299A* and	1362	1407	AATTGGAATAGGTGTCCTCTTGACTTGGATAGGGTTGAAT<u>T</u>CAAAGAACA
R334A*	1420	1465	TCATTTTCATGCATTGCGATAGGAATCATTACACTCTATC<u>T</u>GGGAGCTGT
R107A and R259A	352	397	ATGTTCACATGCAAAAAAAACATGGAAGGGAAAATTGTGC<u>A</u>ACCAGAAAA
	413	458	CCATTGTGGTAACACCTCATTCAGGGGAAGAGAATGCAGT<u>T</u>GGAAATGAC
	575	620	ACTTCAATGAGATGGTGTTGCTGCAAATGGAAAATAAGGC<u>T</u>TGGCTGGTG
	917	962	TTAAAGTTGTAAAGGAAATAGCAGAAACACAACATGGAAC<u>A</u>ATAGTTGTT
	1201	1246	ATGTTCGAGACAACAATGAGAGGAGCGAAGAGAATGGCCA<u>T</u>TTTAGGTGA
R261A	312	357	TTGTGGTTTGTTTGGCAAAGGAAGCTTGGTAACATGTGCG<u>A</u>AATTTCAAT
	513	558	CATCTTGCCTGAATATGGAACCCTTGGGCTAGAATGCTCA<u>C</u>CACGGACAG
	1131	1176	AATTGGAATTGGAGATAACGCCTTGAAAATCAACTGGTAC<u>A</u>AGAAGGGGA
	1307	1352	AAATATTTGGAAGTGCCTACACAGCCCTGTTTAGTGGAGT<u>C</u>TCTTGGGTG
	1362	1407	AATTGGAATAGGTGTCCTCTTGACTTGGATAGGATTGAAT<u>T</u>CAAAGAACA
R282A	841	886	GGACATTTGAAGTGCAAAGTTCGCATGGAGAAATTGAGGA<u>T</u>CAAGGGAAT
	1336	1381	ACTATGTTTGGAGGAGTCTCATGGATGGTTAGAATCCTAA<u>T</u>CGGGTTCTT
	1359	1404	GATGGTTAGAATCCTAATCGGGTTCTTAGTATTGTGGATT<u>G</u>GCACGAGTT
R197A*	312	357	TTGTGGTTTGTTTGGCAAAGGAAGCTTGGTAACATGTGCG<u>A</u>AATTTCAAT
	513	558	CATCTTGCCTGAATATGGAACCCTTGGGCTAGAATGCTCA<u>C</u>CACGGACAG
	1034	1109	ATAGTAATTGGAATTGGAGATAACGCCTTGAAAATCAACT<u>G</u>GTACAA

Prokaryote promoter sites were predicted from the DV E gene consensus sequence data from the patient samples using the Berkeley Drosophila Genome Project (BDGP) Neural Network Promoter Predictor (http://fruitfly.org/seq_tools/promoter.html). No prokaryote promoter sites were predicted for sample R254A. The transcription start sites are underlined in the promoter sequences. The start and end positions of the promoter sequence are shown; numbers refer to the DV E gene sequence for the respective patient samples.

* denotes patient samples that were amplified by high-fidelity RT-PCR but were unable to be sufficiently cloned and sequenced to enable quasispecies analysis

3.11.5 Analysis of dengue virus envelope gene quasispecies variation

To compare the DV quasispecies E gene sequences, the clones from each patient sample were aligned as both nucleotide and amino acid sequences, and types and positions of mutations were detailed. Analysis of the clone sequences revealed less variation within the virus population of individual patients than was expected based on published data. Nucleotide sequences differing from the consensus have been reported for more than 70% of the clones analysed (Craig *et al*., 2003), whilst this study found just 16% differed from the consensus. Average values for the mean diversity of viral quasispecies variants of 0.38% (Lin *et al*., 2004; Wang *et al*., 2002a) and 0.16% (Craig *et al*., 2003) were previously reported compared to the current findings of 0.018% (21.1- and 8.9-fold lower respectively; Table 3.10). However, the results of this study, both in terms of the number of clones found to contain mutations and the average mean diversity values, are consistent with reported error rates for viral RNA polymerases during natural replication. Mutation rates of 1/10,000 nucleotides (Smith *et al*., 1997) give a mean diversity value of 0.01%, and for a 1500 nucleotide E gene, one mutation will be present for every 6.7 clones analysed, corresponding to 15% of clones differing from the consensus.

Further investigations are required to elucidate how viral quasispecies populations vary between the infecting DVs or in the infected human populations studied. Previous studies have looked at DV quasispecies populations in patients from Thailand (Lin *et al*., 2004; Wang *et al*., 2002a) and Myanmar, (Craig *et al*., 2003) who may have an entirely different history of exposure to DV types and genotypes than the Sri Lankan patient samples utilised for this study.

Similar methodologies were used for both studies, although this study used high-fidelity reverse transcriptase and polymerase enzymes in the RT-PCR. The difference between the amounts of variation observed between this and previous studies could be due to the fidelity of the enzymes used for RT-PCR. Previous studies (Lin *et al*., 2004; Wang *et al*., 2002a) used nested PCR after the RT step (using a standard-fidelity RT enzyme), with two rounds of amplification for 30 cycles using *Taq* polymerase. This produced an average mean diversity value that is 21.1 times greater than the results

obtained from this study. The combination of low-fidelity enzymes and increased number of PCR amplification cycles could explain the increased variation observed compared to the present study. Another group (Craig *et al*., 2003) used a blend of *Taq* and high-fidelity *Tgo* polymerase after the RT step (using a standard-fidelity RT enzyme), over 36 cycles of PCR amplification, and reported an average mean diversity value that is 8.9 times greater than this study. A high fidelity RT-PCR kit (AccuScript) was used in this study to try and minimise the number of errors introduced to the DV E gene during amplification by RT-PCR. This combination of high-fidelity reverse transcriptase and polymerase enzymes has been shown to produce up to an 8-fold reduction in mutation rate when compared to a *Taq* and high-fidelity DNA polymerase blend (Arezi and Hogrefe, 2007). This is consistent with the 8.9-fold difference in average mean diversity between the two studies.

In accordance with all the previous studies (Craig *et al*., 2003; Lin *et al*., 2004; Wang *et al*., 2002a) this study found that the extent of sequence diversity within the viral quasispecies population varies between patients. It also concluded that there was no statistical difference ($p=1.0$) between the extent of sequence diversity in the mild compared to the severe dengue cases, which concurs with previous findings (Wang *et al*., 2002a).

Previous reports stated that 5.8% of the clones analysed contained genome-defective DV E gene sequence with insertion or deletion mutations, or substitution mutations resulting in in-frame stop codons, and these were found at a higher frequency in samples from patients with more severe cases of the disease (Wang *et al*., 2002a). These results were based on data obtained from a 430 nucleotide region of the E gene, which codes for all of structural domain *III* and the hinge junction to domain *II*, whereas this study examined the entire E gene. This study observed that 3.4% of the total number of clones analysed contained genome-defective virus gene sequence with insertion or deletion mutations (10/299 clones), and all but one of these were found in patients with severe disease. In this study none of the substitution mutations in the patient samples resulted in in-frame stop codons. Long-term and widespread

transmission of a genome defective DV type 1 strain containing a stop codon within E has been reported (Aaskov *et al*., 2006). The likely mechanism of transmission of the defective lineage was proposed to be complementation of the defective virus by coinfection of cells with functional virus quasispecies variants. The prevalence of the defective lineage increased in quasispecies populations from different patients in successive years in the region sampled and was observed in dengue patients thousands of miles away from the initial site. The increased prevalence of the defective lineage was associated with a major reduction in the occurrence of DV type 1 infections, indicating that the persistence of such deleterious mutations in a population has adverse effects on virus fitness (Aaskov *et al*., 2006).

3.11.6 Measurement of mutations introduced via high-fidelity RT-PCR amplification and comparison of methods used in previous studies

Previous studies estimated error frequencies for mutations introduced by the RT-PCR process of 0.023-0.055% (Lin *et al*., 2004; Wang *et al*., 2002a). These were based on previously published error rates for reverse transcriptase and *Taq* polymerase enzymes, and 60 cycles of PCR amplification. In contrast, the patient samples in this study were both reverse transcribed and amplified using high-fidelity enzymes, and 36 cycles of PCR amplification were conducted, so error frequencies for mutations introduced by the RT-PCR process would be expected to be lower. To determine the error frequencies for mutations introduced by this RT-PCR method, DV type 3 E gene plasmid DNA (from a single colony of known sequence) was transcribed *in vitro* using T7 polymerase. The resulting transcripts were subjected to high-fidelity RT-PCR, cloning and sequencing as described previously. High-fidelity RT-PCRs were conducted using both the AccuScript method applied to the Sri Lankan patient samples in this study and the high-fidelity Expand method used in previous publications (Craig *et al*., 2003). The standard fidelity method used by previous groups was not tested as there was not enough information in either of the publications (Lin *et al*., 2004; Wang *et al*., 2002a) to identify the RT-PCR reagents used.

The Expand method generated more mutations than the AccuScript method, with 1.7-fold higher mean diversity and number of clones containing mutations. Comparison

of insertion and deletion mutation rates between the AccuScript and Expand methods using the RNA transcript revealed them to be very similar (AccuScript 0.032%; Expand 0.035%). In contrast, Expand method substitution mutation rates were 2.47-fold greater than those generated using the AccuScript method. These data support the hypothesis that differences in the amounts of variation observed between this and previous studies are due to the fidelity of the enzymes used for RT-PCR. In light of these results, published data on DV quasipecies variation need to be re-examined to account more accurately for mutations introduced by the amplification process itself.

Both methods using the RNA transcript showed a greater number of clones containing mutations (AccuScript 33%; Expand 55%) and a higher overall mean diversity (AccuScript 0.065%; Expand 0.11%) than was observed with the Sri Lankan patient samples (16% of clones analysed contained mutations; overall mean diversity 0.013%). The difference between these mutation rates observed using RNA transcript versus patient samples indicated that the presence of so many mutations might be attributed to the RNA transcript used rather than the high-fidelity RT-PCR, cloning and sequencing. To test this hypothesis, the *in vitro* transcription reaction was repeated using new reagents in the presence and absence of thioredoxin (section 3.11.6). The resulting transcripts were subjected to high-fidelity RT-PCR, cloning and sequencing using the AccuScript method applied to the Sri Lankan patient samples in this study. In total, 23 clones were analysed: 12 and 11 clones from RNA transcript produced in the presence and absence of thioredoxin respectively. All of these clones matched the consensus sequence as no mutations were identified. This RNA transcript consensus sequence also matched those obtained during the AccuScript versus Expand method comparison. Although no difference was observed between the RNA transcripts generated in the presence and absence of thioredoxin, the fact that no mutations were found in any of the clones this time indicated that the RNA transcript used for the AccuScript and Expand method comparisons was not a homogeneous population of E gene transcripts to begin with. This explains the increased mutation rates observed initially using RNA transcript versus patient samples. The comparison between the mutation rates generated by the AccuScript and Expand methods is still valid, despite

the heterogeneous RNA transcript population, as they were conducted at the same time using the same RNA transcript as template. Presumably, using new reagents for the *in vitro* transcription reactions performed in the presence and absence of thioredoxin enabled the generation of a homogeneous transcript population. A maximum error rate of 0.003% was determined for mutations introduced by the AccuScript RT-PCR. A more accurate determination of the error rate would require sequencing of additional clones until a mutation was found. All of the patient samples have mean diversity values greater than 0.003%, which exceeds the calculated maximum error rate for mutations introduced by the AccuScript RT-PCR.

3.11.7 Phylogenetic analyses of dengue virus envelope gene quasispecies variation

Phylogenetic analyses were conducted to investigate the variability of the DVs circulating during the Sri Lankan outbreak in 2006. All four DV types were circulating at the time of the outbreak, but the genotypic composition of strains had not been identified. Alignments and phylogenetic trees were constructed for each DV type using E gene sequences of known genotype (Lanciotti, Gubler, and Trent, 1997; Lanciotti *et al*., 1994; Messer *et al*., 2003b; Twiddy *et al*., 2002) alongside the quasispecies consensus and mutant clone sequences identified from the patient samples in this study. Nucleotide sequence identities between DV strains from the Sri Lankan patient samples were shown to be over 99% within DV types. The two DV type 2 patient sample consensus sequences were identical in protein sequence, whilst the three DV type 3 patient sample consensus sequences each differed by one amino acid.

It has been reported that intra-virus type recombination occurs in DVs and can produce viruses with different properties to the parental strains (AbuBakar, Wong, and Chan, 2002; Holmes, Worobey, and Rambaut, 1999; Tolou *et al*., 2001; Uzcategui *et al*., 2001; Worobey, Rambaut, and Holmes, 1999). The presence of multiple DV genotypes and recombinant viruses has been reported in individual mosquito and human hosts (Aaskov *et al*., 2007; Craig *et al*., 2003). Evidence of recombination would have been observed if any of the mutant clones displayed a shift in phylogeny compared to their respective consensus sequences. All of the patient samples for each DV type studied

were identified as the same genotype, and all of the mutant clone sequences were shown to cluster together with their respective consensus sequences within genotypes containing strains of a similar geographical origin to the samples. There was no evidence of the presence of multiple genotypes in the circulating Sri Lankan strains, and there was no indication from these E gene data that any of the mutant dengue clones result from recombination events between co-infecting genotypes. Within DV types, all of the samples from this study emerge from a common node on the tree, suggesting divergence from a common ancestor. Therefore, Sri Lankan DVs appear to be restricted to single genotypes within a DV type, despite the abundance of different genotypes that are present on the mainland; recent introduction of strains from outside Sri Lanka is not apparrent from the phylogenetic analyses conducted for this study. This limitation in the number of geneotypes per DV type present in Sri Lanka may explain why no evidence of recombination was found. Given the similarity between the patient sample E gene sequences, there was no evidence to suggest that the severity of disease experienced by the patient was related to the infecting DV genotype.

The first DV type 3 phylogram (Figure 3.12) was constructed using complete E gene sequences. A second phylogram (Figure 3.13) was constructed using a 207 nucleotide region at the 5' end of the gene to incorporate partial Sri Lankan DV sequences available from NCBI (Messer *et al*., 2003a). No NCBI sequences were found for Sri Lankan strains that fell into genotypes *I*, *II* or *IV*, which concurs with previous findings (Messer *et al*., 2003a). As shown previously (Lanciotti *et al*., 1994; Messer *et al*., 2003a), genotype *III* clusters into two distinct groups: *IIIa* and *IIIb.* Those strains that comprise genotype *IIIa* are all from outbreaks during or prior to 1989, which were not associated with DHF epidemics. Those that form genotype *IIIb* are all from 1989 or later and have been associated with epidemics of severe dengue (Lanciotti *et al*., 1994; Messer *et al*., 2003a). It has been suggested that this genetic shift may have increased the virulence of the Sri Lankan strains, and be responsible for the increase in cases of severe disease seen on the island after 1988 (Lanciotti *et al*., 1994). A similar DV type

3 genotype *IIIb* clade replacement has been suggested as the cause of a further increase in severe dengue in Sri Lanka since 1999 (Kanakaratne *et al*., 2009).

In the first DV type 3 phylogram (Figure 3.12), the genotype *I* strain Malaysia 1974 appeared most closely related to the genotype *II* strain Thailand 1987. However, in the second DV type 3 phylogram (Figure 3.13) the genotype *I* strain appears between the genotype *IIIa* and *b* clades. The Sri Lanka 1989c strain also shifts phylogeny between the two phylogenetic trees. In the complete E gene tree (Figure 3.12) it clusters with Sri Lanka 1991 in genotype *IIIb*, whereas in the partial E gene tree (Figure 3.13) it appears more closely related to Sri Lanka 1984 in genotype *IIIa.* A shift in phylogeny when different regions of a sequence are used for the analysis is indicative of recombination. A recombination event may have occurred between a genotype *IIIa* strain and a genotype *I* strain to produce genotype *IIIb* strains with increased virulence compared to *IIIa*. The Sri Lanka 1989c strain may represent an early example of this recombination as it shifts phylogeny between the genotype *IIIa* and *b* strains when different regions of the E gene are used for the analysis. There is no evidence of genotype *I* strains in Sri Lanka so this recombinant must have been imported. The bootstrap values are far from conclusive for the partial E gene phylogenetic tree, but further evidence to support this hypothesis could be obtained by comparison of the complete genome sequences from strains pre- and post-1989 in order to try and identify differences that could have resulted in this increase in virulence. In addition, computational analyses could be performed in order to try and identify recombination break points within the post-1989 sequences. This work is not possible at the present time due to a lack of available virus strains from before 1989.

3.11.8 Linear and three-dimensional mapping of dengue virus quasispecies envelope glycoprotein amino acid substitutions

The amino acid residue substitutions resulting from the nonsynonymous mutations in the patient samples were shown to be distributed throughout the DV E by mapping the substitutions to the linear protein sequence. Six clones were identified with amino acid substitutions that map to structural domain *I*; fifteen clones with substitutions that map to domain *II*; seven clones with substitutions that map to domain *III*, and two

clones with substitutions that map to the transmembrane domain. As expected, none of the residue substitutions were in positions occupied by the 12 conserved cysteine residues required for protein structure (Nowak and Wengler, 1987), or within the fusion peptide at residues E100-108 (Modis *et al*., 2003). Mapping of the amino acid substitutions also confirmed that they did not occur at or proximal to the N-glycosylation sites at amino acids Asn-67 or Asn-153 (Mondotte *et al*., 2007).

Amino acid substitutions with the potential to affect virulence or cell tropism were identified on the basis of their proximity to important structural features within the three-dimensional structure of DV E. Virulence or cell tropism can be affected through modulation of oligomer or virion assembly, cell attachment, fusion of virus and cell membranes, or antibody recognition.

Two of the substituted residues (E306 and E320) were shown to be located on domain *III*, at or within 4Å of the binding site for HSHS molecules on the target cell surface which mediate infectivity (Chen *et al*., 1997; Thullier *et al*., 2001). Neutralisation escape mutations in multiple flaviviruses map to this HSHS binding site (Jiang *et al*., 1993; Lin *et al*., 1994), so whilst these substituted residues are not surface accessible they could potentially alter the protein structure enough to disrupt the HSHS-binding site and either enhance or attenuate virulence by altering virus infectivity. Four of the substituted residues (E49, E55, E135 and E280) were located at the base of domain *II* at the interface with domain *I*, with one of these residues (E280) shown to be surface accessible and within the *kl* hairpin loop implicated in conformational rearrangements during fusion (Modis *et al*., 2003). Attenuated viruses with single amino acid substitutions in the *kl* hairpin loop have been obtained by flavivirus passage in cell culture (Lee, Weir, and Dalgarno, 1997; Monath *et al*., 2002), and neutralising antibody epitopes have also been mapped to this region (Beasley and Aaskov, 2001). A further three substituted residues (E151, E246, and E343) were located within 4Å of the fusion domain, with one of these located to domain *I* at the interface of domain *III* and the domain *II* conserved region on the other monomer at residues 98-113, which houses the fusion peptide (Modis *et al*., 2003; Rey *et al*., 1995). One additional substituted

residue (E113) was found to be within this conserved region at residues 98-113. Mutations leading to amino acid substitutions at the base of domain *II* or at the domain *I*/*III* interface with the domain *II* conserved region at residues 98-113 are thought to affect virulence by impacting upon the conformational rearrangements that lead to fusion of virus and host cell membranes (Rey *et al*., 1995). Some of these mutations have been shown to affect fusion by altering the pH required to induce the conformational changes (Mandl *et al*., 1989; Modis *et al*., 2003; Rey *et al*., 1995). Most of these mutations involve residues with side-chains that project into the ligand-binding pockets between the dimers and it has been reported that lowering of the pH threshold for fusion is due to the substitution of longer hydrophobic side chains for shorter ones (Modis *et al*., 2003).

Seven of the amino acid substitutions (at positions E77, E201, E225 E280, E343 and two at E345) from five different patient samples were situated on the surface of the protein and therefore may be involved in antibody or cell binding. Amino acid substitutions in flavivirus E that affect the binding of neutralising mAbs have been shown to map to the surface of the dimer across all three structural domains (Rey *et al*., 1995). Of these surface accessible residues, four (E331, E343 and two at E345) were on the distal face of domain *III*, which is reported to contain the flavivirus receptor-binding motif (Crill and Roehrig, 2001). The region between E residues E335 and E351 has been reported to contain an epitope for the binding of neutralising antibodies for DV type 2 (Roehrig, Bolin, and Kelly, 1998), so the amino acid substitutions at E343 and E345 found in this study could affect virus neutralisation by antibodies. Substitutions at the distal face of domain *III* have been shown to affect virulence or cell tropism via interference with cell attachment (Rey *et al*., 1995) and could also affect virus neutralisation by antibodies.

In the DV type 3 sample from a patient with mild disease (R261A), three distinct lineages of DV were observed within the quasispecies population. The consensus sequence was represented by 72% of clones, 13% of clones carried the same nonsynonymous mutation causing a conservative amino acid substitution at E position

E201, and 6.5% of clones carried the same synonymous mutation at E gene position E1146. There were also 4 clones (8.5%) shown to contain unique mutations. The relative abundance of the two lineages that differ from the consensus sequence within the quasispecies population suggests that either they were derived from virus with the consensus sequence, but replicate more efficiently than other variants within the population, or they were present in the DV quasispecies population in the mosquito along with the consensus sequence and were introduced together when the patient was bitten. Mutations leading to amino acid substitutions at E position E202 commonly occur during routine serial passage in mammalian cells and have been proposed to map to a putative hinge region between domains *I* and *II*, which may be important during fusion (Lee, Weir, and Dalgarno, 1997). Taken together with the results of this study this suggests that the amino acid substitutions observed at E201 may represent a beneficial adaptation of the virus to growth in mammalian cells, via a mechanism related to membrane fusion. Quasispecies investigations with poliovirus have demonstrated a link between the extent of sequence diversity and pathogenicity; too many or too few virus variants within the population results in reduced viral fitness and attenuation (Crotty, Cameron, and Andino, 2001; Vignuzzi *et al*., 2006). Whilst the mean diversity of the R261A sample is comparable to those of the other DV type 3 samples, the presence of so many clones containing the same mutations means that this sample actually exhibits less quasispecies variation than samples where each of the variant clones contains different mutations. This restricted quasispecies population could explain why patient R261A exhibited mild disease.

Amino acid substitutions located at regions involved in cell attachment, membrane fusion or antibody recognition were scarce in the samples from patients exhibiting mild disease. In the DV type 4 sample from a patient with mild disease (R282A), none of the mutations were surface accessible, or mapped to predicted virulence determinants. With the DV type 3 sample from a patient with mild disease (R261A), the only mutation that was located in a region of potential importance for virulence was surface accessible, but resulted in a conservative rather than drastic amino acid

substitution in E. This is consistent with a role for individual quasispecies variants in the pathogenesis of disease, but is by no means conclusive.

Because the outer surfaces of mature flavivirus virions are covered with a dense network of E dimers, any conformational changes in E are likely to induce concerted reorganisation across the surface of the virion (Crill and Chang, 2004; Kuhn *et al*., 2002; Modis *et al*., 2003). Three-dimensional mapping was limited in that the proximity of substituted residues to important sites was only established for pre-fusion E. More amino acid substitutions with the potential to affect virulence or cell tropism may have been identified by three-dimensional mapping using post-fusion crystal structures.

3.11.9 Identification of dengue virus quasispecies envelope genes for recombinant protein production

Nine of the patient sample DV E gene clones were chosen for recombinant E production (Chapter 4) in preparation for binding studies to determine the effects of the mutations (Chapter 5). These clones were from four different patient samples, three with severe and one with a mild case of disease, across DV types 2 and 3. For each patient sample, a clone was identified that represents the consensus sequence for the quasispecies population for that patient. These consensus clones would also be used to produce recombinant E as a comparison alongside clone variants with nonsynonymous, non-frameshift mutations in regions identified as determinants of DV pathogenicity.

3.12 Conclusions and summary

The primary objective of the work presented in this chapter was to investigate DV quasispecies populations, both within individual patient samples and between samples from the same outbreak in Sri Lanka, in 2006. Previous studies of DV quasispecies populations used RT-PCR, ligation independent cloning, and sequencing of partial or complete DV E genes obtained directly from patient serum or mosquitoes; examining between 10 and 21 clones per sample and reporting average mean diversities of between 0.16% and 0.38% (Craig *et al*., 2003; Lin *et al*., 2004; Wang *et al*., 2002a). The work presented in this chapter differed from previously published work in that it was intended that more dengue patient samples, and a greater number of complete E gene clones per sample would be analysed than in previous studies, to provide more data for statistical analysis. The use of high-fidelity reverse transcriptase and PCR enzymes was deemed essential to minimise the number of mutations introduced by the amplification process itself. Potential relationships between quasispecies variation and disease severity were also investigated.

From the 95 dengue patient samples where DV E gene amplification, cloning and quasispecies analysis was attempted, only six produced enough clones to permit quasispecies analysis. Therefore the aim of this study to examine more patient samples than previous studies was not fulfilled. This was due primarily to difficulties with RNA integrity, and plasmid instability during cloning. However, from these six patient samples, 299 clones were analysed (range 25 to 82 clones per sample), so the intention to study a greater number of complete E gene clones per sample than previous studies was successful. Across the six samples, an average mean diversity of 0.018% (range 0.0034% to 0.053%) was obtained, and no statistically significant difference was observed between the extent of quasispecies variation and disease severity. This is less variation than was expected based on previously reported average mean diversities of between 0.16% and 0.38% (Craig *et al*., 2003; Lin *et al*., 2004; Wang *et al*., 2002a). However, the results presented in this chapter are more consistent with reported error rates for viral RNA polymerases during natural replication (0.01%, Smith *et al*., 1997). The differences in the amounts of sequence variation between this study

and previous studies, were shown to be directly related to the fidelity of the enzymes used by each group for RT-PCR amplification of the viral RNA. This emphasises the importance of using high-fidelity enzymes for quasispecies analysis. In light of these results, previously reported levels of variation in DV quasispecies populations need to be re-examined to account more accurately for mutations introduced by the amplification process itself.

Phylogenetic analyses were conducted to investigate the variability of the DVs circulating during the Sri Lankan outbreak in 2006. Within each DV type studied, all of the DV strains from the Sri Lankan patient samples were identified as the same genotype, appropriate to their geographic origin, and with nucleotide sequence identities of greater than 99%. There was no evidence of the presence of multiple genotypes in the circulating Sri Lankan strains, and there was no indication from these E gene data that any of the mutant dengue clones resulted from recombination of co-infecting genotypes. Therefore, no relationship was observed between the infecting DV genotype and disease severity.

Nine of the ten clones observed to contain genome-defective DV E gene sequences (with insertion or deletion mutations; 10/299 clones) were found in patients with severe disease. Linear and three-dimensional mapping of the amino acid substitutions within the DV quasispecies populations identified residues in locations with the potential to affect virulence or pathogenicity. These were more prevalent in the samples from patients with severe disease. These observations are consistent with a role for individual quasispecies variants in the pathogenesis of disease.

Nine clones from four patient samples were chosen for recombinant E production on the basis that they contain a mutation with the potential to alter the behaviour of the protein, or represent the consensus sequence for that sample for comparison. Studying the effect of these mutations on protein function will provide further insight into potential links between quasispecies variation and disease severity.

CHAPTER 4. PRODUCTION AND PURIFICATION OF RECOMBINANT DENGUE VIRUS ENVELOPE GLYCOPROTEINS

4.1 Introduction

Recombinant flavivirus envelope glycoproteins (E) have been transiently produced using various expression systems. Bacterial, yeast, insect and mammalian cells have been employed as hosts for recombinant protein production from flavivirus E genes delivered by plasmid or viral vectors (Allison *et al*., 1995; AnandaRao *et al*., 2006; Bielefeldt-Ohmann *et al*., 1997; Bray *et al*., 1989; Chang *et al*., 2003; Delenda *et al*., 1994; Liu *et al*., 2010). Stable expression of flavivirus E genes using transformed cell lines has also been demonstrated, but production of mature, secreted protein is complicated in mammalian cells by the inherent toxicity of flavivirus E (Konishi and Fujii, 2002). Similar problems have been reported with flavivirus protein production using genome-length infectious clones (Pu *et al*., 2011; Yamshchikov, Mishin, and Cominelli, 2001) which require the same level of biocontainment as live virus (Containment Level 3).

It has been demonstrated that the production of secreted recombinant flavivirus E with the correct structural conformation requires either co-expression of the pre-membrane glycoprotein (prM), or truncation of the E gene to remove the C-terminal membrane anchor (Allison *et al*., 1995). This allows the protein complex to be trafficked through the acidic environment of the endoplasmic reticulum and secreted. Co-expression of flavivirus E with prM has been shown to produce a higher yield of secreted protein than C-terminal truncation of E (Allison *et al*., 1995), and results in the production of spherical virus-like particles (VLPs; Mason *et al*., 1991; Schalich *et al*., 1996). These VLPs are approximately 30nm in diameter, and are comparable to virions in terms of their antigenic and oligomeric structures, and their ability to undergo structural rearrangements at acidic pH (Schalich *et al*., 1996). Recombinant VLPs or E secreted into cell culture medium are easier to purify and represent a more homogenous population than intracellularly, where proteins are at different stages of processing and maturation. Secreted VLPs have an additional advantage over secreted E alone, as they are more resistant to proteolytic degradation (Sugrue *et al*., 1997). During natural flavivirus infections, in addition to infectious virions, non-infectious viral particles are produced. These contain flavivirus membrane protein (M) and E but lack

the nucleocapsid (Russell, Brandt, and Dalrymple, 1980), and are therefore essentially the same as recombinant VLPs.

The use of eukaryotic expression systems enables the appropriate post-translational modifications (such as disulphide bond formation, N-linked glycosylation and furin cleavage) for the production of secreted flavivirus E. The level of post-translational modification achieved using insect or mammalian cells as hosts is more complex than can be achieved using yeast cells. However, there are significant differences in the post-translational processing of glycoproteins by insects and higher eukaryotes, in particular with regards to N-linked glycosylation (reviewed in Jarvis, 2003). In light of the fact that flaviviruses are able to infect and replicate naturally in both human and mosquito cells, using an insect or mammalian expression system provides the best chance of producing functional flavivirus E with the correct post-translational modifications.

The baculovirus expression system produces increased protein yields and is less expensive compared to protein production using a mammalian system. Flavivirus E produced using a baculovirus expression system have been shown to be antigenically indistinguishable from those produced by the flavivirus itself (Shiu *et al.*, 1991), and are immunogenic in mice (Delenda *et al.* 1994; Despres *et al.* 1991; Yang *et al.* 2005). For these reasons, it was decided to use the baculovirus expression system for the work described in this chapter.

4.2 Chapter objectives

The primary objectives of this chapter were to produce and purify recombinant dengue virus (DV) E, using the DV quasispecies E gene clones identified in Chapter 3.

The recombinant DV quasispecies E would be produced as secreted VLPs to enable accurate antigenic presentation of E. The VLPs produced would be used in cell surface receptor binding studies to determine the effect of the quasispecies mutations on protein function, and as a source of antigen in the serodiagnosis of DV infection (Chapter 5). The dengue VLPs would also provide a model for studying the effects of the quasispecies E gene mutations on the low pH-induced oligomeric rearrangements required for fusion of virus and host cell membranes. To the best of my knowledge this is the first report of recombinant flavivirus E or VLP production from quasispecies variants.

DV quasispecies E gene clones identified in the previous chapter (section 3.10; Table 3.16) would be amplified by high-fidelity PCR, before cloning with the DV type 1 prM gene into the baculovirus transfer vector, pBAC-2cp. These modified transfer vectors would be used to construct recombinant baculoviruses (rBVs) containing the DV prM and E genes, which would enable secreted recombinant dengue VLP production upon infection of insect cells. The recombinant dengue VLPs would be purified using His-tags inserted at the C-terminal of E.

4.3 Cloning of dengue virus pre-membrane and quasispecies envelope genes into the baculovirus transfer vector, pBAC-2cp

To produce recombinant proteins using the baculovirus expression system, the genes encoding the proteins of interest needed to be amplified and cloned into a transfer vector suitable for the production of rBVs. pBAC-2cp is a 5411bp commercially-available baculovirus transfer vector (Novagen; Figure 4.1). This transfer vector contains the baculovirus polyhedrin (polh) promoter region and an ATG start codon upstream of the MCS, which is flanked by N-terminal His- and S-tag, and C-terminal His-tag coding sequences. Cloning of the DV prM and quasispecies E genes into the MCS, and subsequent co-transfection of these plasmids with baculovirus DNA into insect cells, was expected to result in the production of rBVs which could then be used to produce secreted recombinant dengue VLPs. The N-terminal His- and S-tags were predicted to be detached when prM was cleaved to produce M. E were expected to retain the C-terminal His-tags to aid detection and purification.

To ensure that any differences in function between the DV quasispecies VLPs could be attributed to variations in E, the source of the prM gene was kept consistent for all of the DV quasispecies E genes. DV type 1 (Hawaii strain) prM gene was chosen for co-expression with all of the DV quasispecies E genes. PrM and E gene primers (Table 4.1) were designed (sections 2.2.2 and 2.2.3) to amplify only the gene of interest and not the flanking regions. Unique endonuclease restriction sites were incorporated six bases from the ends of the primers to enable directional ligation-dependent cloning of the two genes into pBAC-2cp (Figure 4.2), mimicking the order they are presented in the flavivirus genome. The DV type 1 prM gene forward primer contains a *Nhe*I restriction enzyme sequence for ligation into the pBAC-2cp MCS. The reverse primer incorporates an *Eag*I restriction enzyme sequence to enable ligation with the *Eag*I restriction enzyme sequences within the E gene forward primers, ensuring the E gene is inserted downstream of the prM gene. E gene reverse primers contain *Xma*I restriction enzyme sequences to ligate into the pBAC-2cp MCS downstream of the *Nhe*I site. Restriction sites were chosen on the basis that they produced overhanging rather than blunt ends, and were unique in pBAC-2cp, but absent in the DV prM and E

gene sequences. Due to the presence of the *Eag*I restriction site, the codons for two amino acids (arginine and proline) were inserted between the prM and E genes.

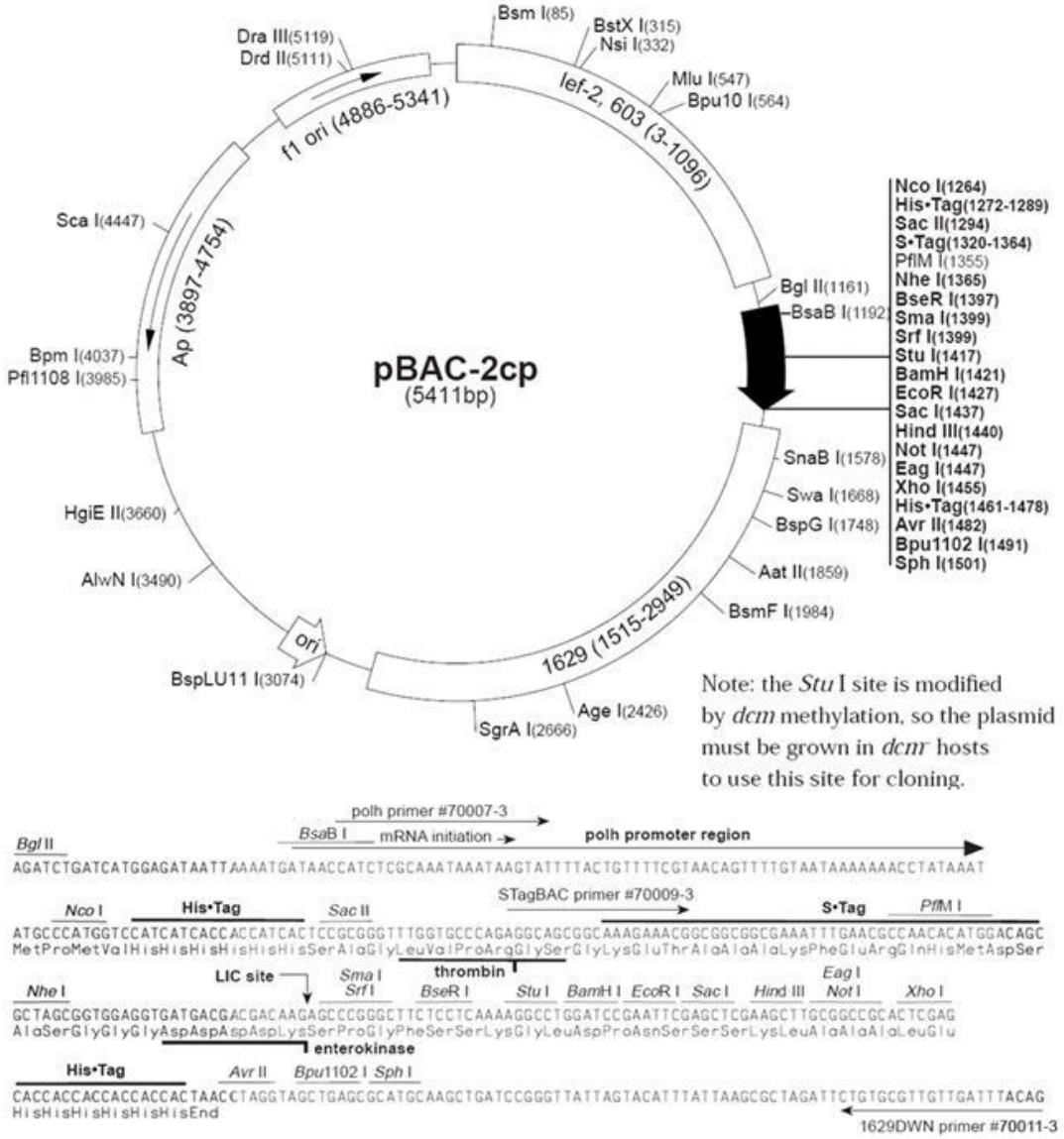


Figure 4.1. pBAC-2cp baculovirus transfer vector circle map and MCS nucleotide sequence

Unique restriction sites and sequence landmarks are indicated on the circle map with nucleotide sequence positions in brackets). The baculovirus polyhedrin (polh) promoter region (1177-1259) contains the polh transcription start (1210-1211) and wild type polh 5'UTR -1 position (1259). His-tag (1271-1289 and 1461-1478) and S-Tag (1320-1364) coding sequences are positioned flanking the MCS (*Sma*I to *Xho*I, 1397-1460). Diagram is from product insert (Novagen).

Using the appropriate primers (Table 4.1), the DV type 1 prM gene was amplified from DV type 1 Hawaii strain RNA (section 2.1.3) by high-fidelity RT-PCR (section 2.3.1.2), and the DV quasispecies E genes were amplified from plasmid DNA (section 3.5) using

only the PCR step of the high-fidelity RT-PCR. DNA from PCR amplicons of the correct size was gel extracted (section 2.3.3.1) before sequential restriction digestion with the appropriate restriction enzymes (*Nhe*I and *Eag*I for the prM gene; *Eag*I and *Xma*I for the E genes), DNA purification, and ligation into pBAC-2cp, which had undergone a double digest with *Nhe*I and *Xma*I (sections 2.3.8, 2.3.3.2 and 2.3.6.1).

After transformation of each pBAC-2cp ligation reaction into JM109 cells and overnight culture on selective agar plates (section 2.3.5.3), standard PCR (section 2.3.1.4) was performed on the resulting colonies using the pBAC-2cp polh forward primer (Figure 4.1) and the appropriate DV quasispecies E gene reverse primer for each patient sample (Table 4.1). DNA agarose gel electrophoresis (section 2.5.1) of the PCR amplicons (Figure 4.3) showed an amplicon size of 2kb, confirming that the colonies contained the pBAC-2cp transfer vector with the prM (0.5kb) and E (1.5kb) genes in the correct order and orientation.

Table 4.1. Primers used for the amplification of the dengue virus type 1 pre-membrane and dengue virus quasispecies envelope genes for generation of rBVs.

Primer	Sequence (5' to 3')	Product size (bp)
DV type 1 prM *Nhe*I F	CAC AGC GCT AGC TTC CAT CTG ACC AC	528
DV type 1 M *Eag*I R	CAC GCA CGG CCG GGC CAT GGA TG	
R097A E *Eag*I F	CGG GAA CGG CCG ATG AGA TGT GTG	1500
R097A E *Xma*I R	TTC ACT CCC GGG ATA GCT TGT ACC ACA GC	
R203A/R261A E *Eag*I F	CCA TCC CGG CCG ATG AGA TGT GTG GGA G	1503
R203A/R261A E *Xma*I R	ACA CCC CCC GGG AGC TTG TAC CAC AGC	
R107A/R259A E *Eag*I F	CCT TCA CGG CCG ATG CGT TGT ATT GG	1509
R107A/R259A E *Xma*I R	ACA ACC CCC GGG AGC CTG CAC CAT AAC	
R282A E *Eag*I F	CCA TCC CGG CCG ATG CGA TGC GTA G	1509
R282A E *Xma*I R	ACA ACC CCC GGG TGC ATG AAC TGT G	

E gene primer sequences were derived from the quasispecies E gene sequencing data (section 3.6, Appendix C). DV type 1 prM gene primer sequences were derived from the DV type 1 RefSeq (NCBI; section 2.2.1). Primers are labelled according to the gene template source and whether they target prM or E genes. Restriction sites within the primer sequence are highlighted and PCR amplicon product size is shown in nucleotide base pairs (bp). Primers were also used for confirmatory PCRs and sequencing.

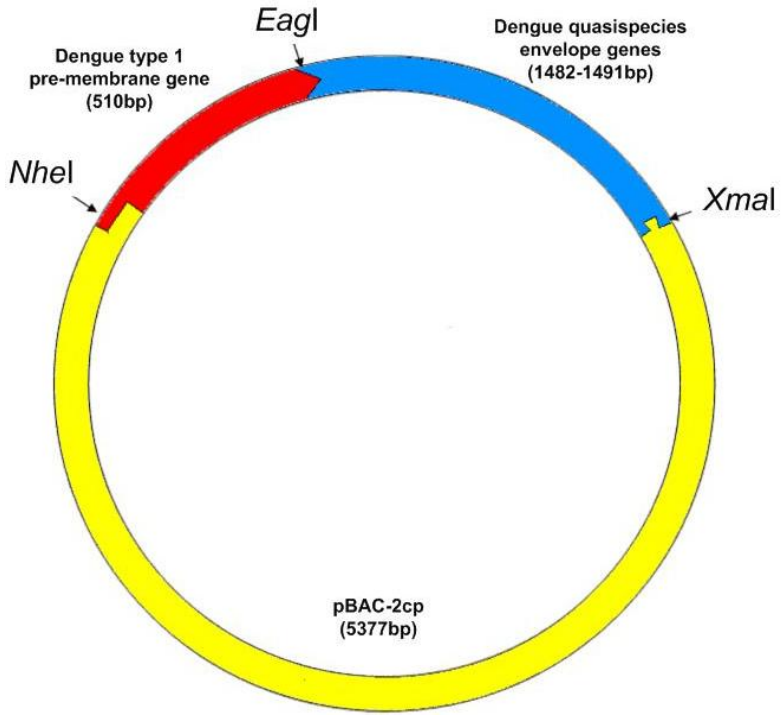


Figure 4.2. Schematic representation of the rBV transfer vector pBAC-2cp containing dengue virus type 1 pre-membrane and dengue patient sample quasispecies envelope genes.

Prior to ligation, pBAC-2cp (yellow) was digested with *Nhe*I and *Xma*I, the DV type 1 prM gene (red) with *Nhe*I and *Eag*I, and the DV quasispecies E genes (blue) with *Eag*I and *Xma*I. Digest product sizes are shown in nucleotide base pairs (bp) in brackets.

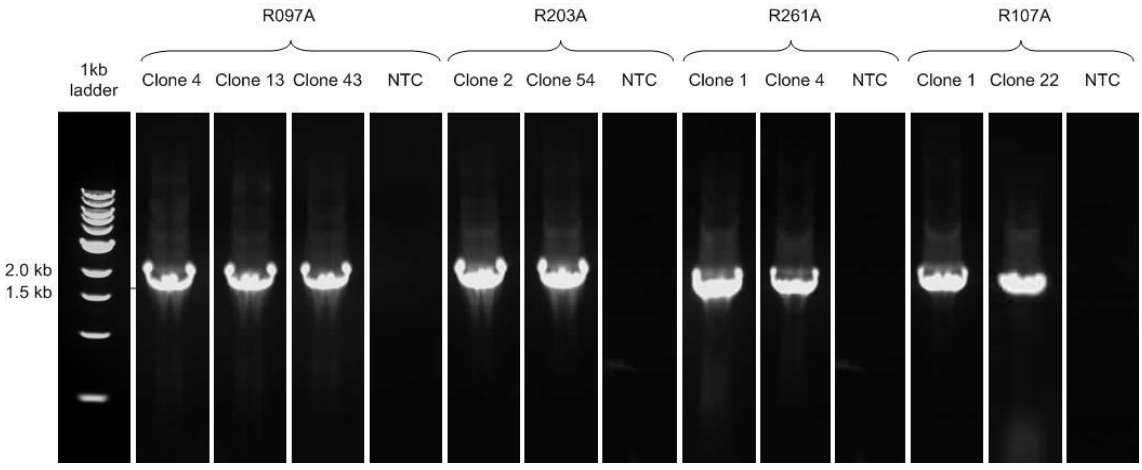


Figure 4.3. Analysis of colonies by PCR for confirmation of the presence of the dengue virus type 1 pre-membrane and patient sample quasispecies envelope gene inserts in pBAC-2cp

PCR was performed using the polh forward primer (Figure 4.1) and the appropriate DV E gene reverse primer for each dengue patient sample (Table 4.1). Lanes are labelled according to the patient sample clone used to source the E gene (section 3.10). A 1kb DNA ladder is shown for comparison, with the 1.5 and 2.0kb fragments highlighted. PCR amplicons of 2kb are consistent with the size of the prM (0.5kb) and E (1.5kb) genes. No amplicons were detected for any of the no template control (NTC) reactions.

Colonies with the correct insert underwent further culture and plasmid DNA extraction (sections 2.3.6.2 and 2.3.4.1 respectively). To confirm the DV type 1 prM gene sequence, and the presence or absence of the respective quasispecies E gene mutations, the plasmid DNA was sequenced (section 2.4) using the Polh and 1629DWN primers (Figure 4.1), and primers internal to the E gene sequences (Table 3.6). The sequence of the prM-E reading frame was verified to confirm that the genes were in-frame with the pBAC-2cp start codon and His- and S-tags (data not shown). Culture of the sequence verified clones was then scaled-up (section 2.3.6.2) and plasmid DNA purified (section 2.3.4.2) for subsequent rBV production.

4.4 Production of recombinant baculoviruses containing dengue virus pre-membrane and envelope genes

To maximise the yield of secreted dengue VLPs, *Flash*BAC GOLD™ (Oxford Expression Technologies, OET) was chosen as the source of parental baculovirus DNA for transfection. *Flash*BAC GOLD™ is modified *Ac*MNPV DNA lacking the polyhedrin gene, part of open reading frame (ORF) *1629* (which renders it non-infectious), and the non-essential auxiliary genes *chiA* and *v-cath*, which would otherwise compete with the recombinant protein for cellular resources during gene expression. Removal of the auxiliary genes improves the efficiency of the secretory pathway leading to a greater yield of secreted or membrane-targeted recombinant proteins (Hitchman *et al*., 2010).

The transfer vectors containing the DV prM and E genes were co-transfected with *Flash*BAC GOLD™ baculovirus DNA into *Sf*9 cells to generate rBV (section 2.7.2). A baculovirus transfer vector containing the *lacZ* gene was used as a positive control for the transfection reactions, as a β-galactosidase assay (section 2.7.3) could then be performed to determine the transfection reagents were working correctly. The *lacZ* rBV-infected culture produces β-galactosidase (116 kDa for each of the four subunits) intracellularly under the control of the baculovirus polyhedrin promoter. When 5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside (X-gal) is added to the culture it is cleaved by β-galactosidase into galactose and 5-bromo-4-chloro-3-hydroxyindole, which is then oxidized to 5,5'-dibromo-4,4'-dichloro-indigo, an insoluble blue product. Recombinant baculoviruses were designated rBV.*xxxxx* according to either the patient sample clone yielding the quasispecies E gene, or the inserted gene in the case of the *LacZ rBV* (Table 4.2).

Each dengue rBV underwent two successive amplifications (passages 1 and 2; P1 and P2) and was titred (sections 2.7.4.1 and 2.7.4.3). Titration of rBVs was performed by end-point dilution using *Sf*9 *Easy-Titre* (*E-T*) cells (kindly provided by Dr. Ralph Hopkins, National Cancer Institute at Frederick, Maryland, USA; Hopkins and Esposito, 2009). This cell line consists of *Sf*9 cells stably transfected with the enhanced green fluorescent protein (*eGFP*) gene under the control of the baculovirus polyhedrin promoter. Fluorescence can be detected from rBV-infected *Sf*9 *E-T* cells due to the

activation of the polyhedrin promoter/*eGFP* complex by baculovirus early gene products expressed during the infection. To verify that the cells express *eGFP* when infected with rBV, the *Sf*9 *E-T* cells were infected with rBV.*lacZ*. After a three day incubation at 27°C in the dark, cells were visualised using a fluorescent microscope and green cell foci were observed (Figure 4.4). Confirmation of β-galactosidase production (and therefore rBV infection) was obtained by β-galactosidase assay (section 2.7.3). End-point dilution titrations were carried out according to the method detailed in section 2.7.4.2. The results are shown in Table 4.3.

Table 4.2. rBVs generated for this study

Inserted genes	Recombinant baculovirus
prM and R097A_4 E	rBV.R097A_4
prM and R097A_13 E	rBV.R097A_13
prM and R097A_43 E	rBV.R097A_43
prM and R203A_2 E	rBV.R203A_2
prM and R203A_54 E	rBV.R203A_54
prM and R261A_1 E	rBV.R261A_1
prM and R261A_4 E	rBV.R261A_4
prM and R107A_1 E	rBV.R107A_1
prM and R107A_22 E	rBV.R107A_22
lacZ	rBV.lacZ

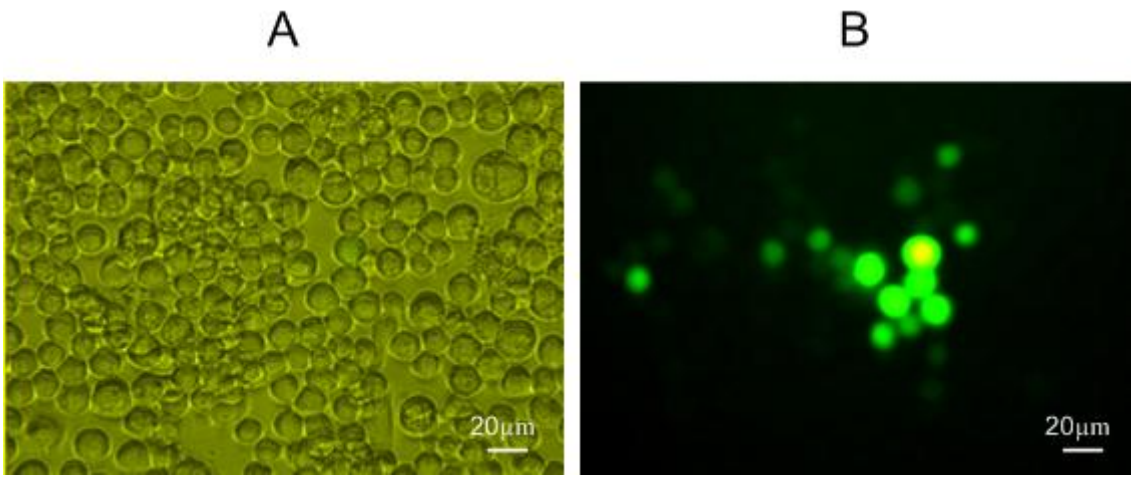


Figure 4.4. *Sf*9 *Easy-Titre* cells express *eGFP* when infected with rBV

*Sf*9 *Easy-Titre* cells infected by rBV.*lacZ* are shown under white light (panel A) and blue light (wavelength 488nm; panel B) using a fluorescent microscope. Green (wavelength 509nm) cell foci (panel B) are indicative of replication-competent rBV-infection.

Table 4.3. Dengue rBV titres

Recombinant baculovirus	$TCID^{50}$ (10^{exp})	$TCID^{50}$	pfu/mL
rBV.R097A_4	8.8	$6.81x10^{8}$	$4.77x10^{9}$
rBV.R097A_13	8.33	$2.13x10^{8}$	$1.49x10^{9}$
rBV.R097A_43	7.6	$3.99x10^{7}$	$2.79x10^{8}$
rBV.R203A_2	10.1	$1.26x10^{10}$	$8.81x10^{10}$
rBV.R203A_54	8.71	$5.12x10^{8}$	$3.59x10^{9}$
rBV.R261A_1	8.61	$4.11x10^{8}$	$2.88x10^{9}$
rBV.R261A_4	7.34	$2.17x10^{7}$	$1.52x10^{8}$
rBV.R107A_1	7.5	$3.16x10^{7}$	$2.21x10^{8}$
rBV.R107A_22	10.12	$1.32x10^{10}$	$9.23x10^{10}$

The presence of the DV prM and E genes in the rBVs was confirmed by PCR of DNA extracted from culture medium from insect cells infected with P2 amplified rBVs (Figure 4.5). After DNA extraction (section 2.3.3.3), standard PCR (section 2.3.1.4) was performed using the polh forward primer (Figure 4.1) and the appropriate DV E gene reverse primer for each patient sample (Table 4.1). DNA agarose gel electrophoresis (section 2.5.1) of the PCR amplicons (Figure 4.5) showed an amplicon size of 2kb, confirming that the rBVs contained the prM (0.5kb) and appropriate E (1.5kb) genes in the correct order and orientation.

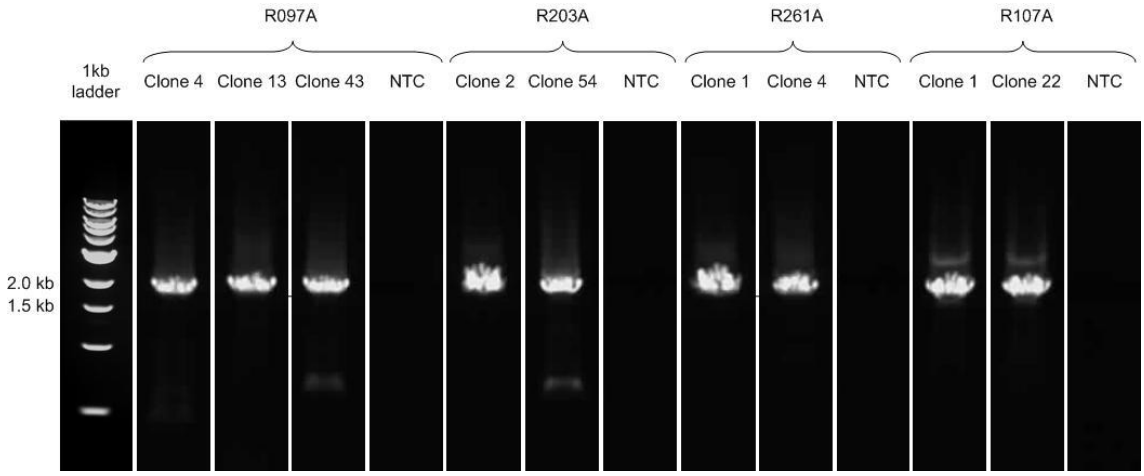


Figure 4.5. Analysis of dengue rBVs by PCR to confirm the presence of dengue virus type 1 pre-membrane and patient sample quasispecies envelope genes

PCR was performed on P2 rBV-infected insect cell culture medium using the polh forward primer (Figure 4.1) and the appropriate DV E gene reverse primer for each dengue patient sample (Table 4.1). Lanes are labelled according to the patient sample clone used to source the E gene (section 3.10). A 1kb DNA ladder is shown for comparison, with the 1.5 and 2.0kb fragments highlighted. PCR amplicons of 2kb are consistent with the size of the DV prM (0.5kb) and E (1.5kb) genes. No PCR products were detected for any of the no template control (NTC) reactions. The original (unprocessed) gel images used to construct this figure are included electronically in Appendix J.

4.5 Optimisation of dengue virus-like particle production

Secreted DV E (C-terminally truncated) have been successfully produced using the baculovirus expression system with rBVs infecting *Sf*21 cells at a multiplicity of infection (MOI) greater than five, and proteins harvested from the culture supernatant at 48 hours post infection (Bielefeldt-Ohmann *et al*., 1997). For this study, optimisation experiments were conducted to determine the optimal cell line, multiplicity of infection (MOI) and post infection interval before harvesting the dengue proteins as secreted VLPs (section 2.7.5). Infections were set up using rBV.R097A_4 at three MOIs (2, 5 and 10) in both *Sf*9 and *Sf*21 cells. Samples were taken at 48 and 72 hours post infection and both the lysed cell pellet and clarified cell culture medium were analysed via SDS-PAGE for the presence of the DV 54 kDa E (section 2.5.3.1; Figure 4.6). Proteins of approximately 64 kDa and 54 kDa were visible in the infected clarified cell culture medium samples, but not visible in the mock-infected control samples. The 64 kDa protein was present at higher concentration than the 54 kDa protein (as judged by visual analysis of stained SDS-PAGE gels), and represented a secreted baculovirus or insect cell protein, as comparison of these results to similar in-house experiments using rBVs containing genes for non-DV proteins (Ebola virus glycoprotein (rBV.EBOV-GP) and 40 kDa viral protein (rBV.EBOV-VP40), and Crimean-Congo haemorrhagic fever virus nucleoprotein (rBV.CCHFV-NP) revealed that the 64 kDa protein was consistently present (data not shown). The 54 kDa protein was consistent with the size of DV E, and was not present in samples from cultures infected with rBV.EBOV-GP, rBV.EBOV-VP40 or rBV.CCHFV-NP. For the 54 kDa protein, there was no observable difference in protein production between cell lines. However, from visual analysis of protein band intensity, the 54 kDa protein production appeared greatest using a MOI of 5, and harvesting protein at 48 hours post infection.

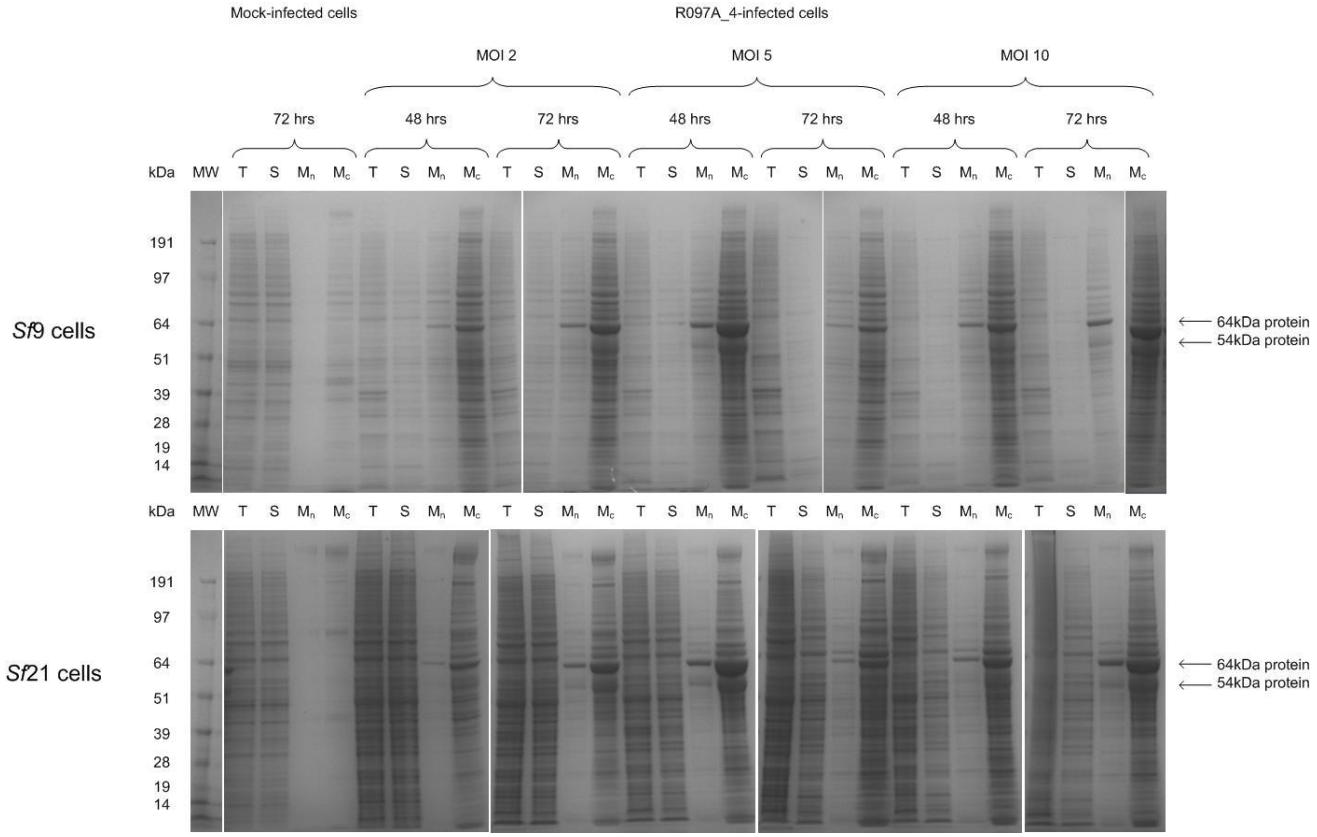


Figure 4.6. Analysis of optimisation of recombinant dengue virus envelope protein production

Infections with rBV.R097A_4 were conducted in *Sf*9 and *Sf*21 cell lines, at MOIs of 2, 5 and 10, and proteins harvested at 48 and 72 hours post infection. Total cell lysate (T), the soluble fraction of the total cell lysate (S) and clarified insect cell culture medium, neat (M_n) or concentrated using a centrifugal filter device (M_c), were analysed by SDS-PAGE. The original (unprocessed) gel images used to construct this figure are included electronically in Appendix J.

To confirm the presence of His-tagged DV E, denatured proteins from the SDS-PAGE gels were transferred onto a PVDF membrane by Western blot and immunodetection attempted using a monoclonal antibody (mAb) to the His-tag (section 2.6). Unexpectedly, His-tagged protein was not detected in any of the samples (data not shown). Only the His-tagged control protein was detected, confirming that the Western blot and immunodetection antibodies were working. This led to two hypotheses. The first hypothesis was that the C-terminal His-tag on DV E was either not present or not accessible to the His-tag mAb, preventing the DV E from being detected using this antibody. A different antibody was sourced to directly detect the DV E (flavivirus E mAb, Abcam) so as not to rely on the presence of the His-tag. The second hypothesis was that R097A_4 E was either not being produced or only being

produced at very low levels, below the limit of detection using the His-tag mAb. From previously published work it became apparent that concentration of the rBV-infected insect cell culture medium was required to detect E (Bielefeldt-Ohmann *et al.*, 1997; Jaaskelainen *et al.*, 2003). It was decided that samples would be concentrated via ultracentrifugation through sucrose before attempting detection of DV E again, as this method had been used successfully previously to concentrate flavivirus VLPs from cell culture supernatant (Jaaskelainen *et al.*, 2003; Schalich *et al.*, 1996).

As previously described (section 4.4, page 158, Figure 4.5), PCR performed on the rBV DNA had confirmed that the rBV genomes contained the DV prM and E genes. To determine whether these genes were being transcribed in preparation for translation and protein production, total RNA (including mRNA) was isolated (section 2.3.2.2) from rBV.R097A_4-infected *Sf*9 cells harvested for the protein production optimisation experiment (Figure 4.6). These samples were subjected to RT-PCR (section 2.3.1.1) using the R097A_4 E gene primers (Table 4.1). PCR (section 2.3.1.4) was also performed using the same primers to ensure there was no DNA contamination of the RNA extract (so no baculovirus DNA containing the DV genes was present). Analysis of the RT-PCR and PCR products by DNA gel electrophoresis (section 2.5.1) showed DV E mRNA (1.5kb) at all MOIs and time-points from infected, but not mock-infected cells (Figure 4.7). No DNA contamination of the RNA extract was detected in the PCR products (data not shown).

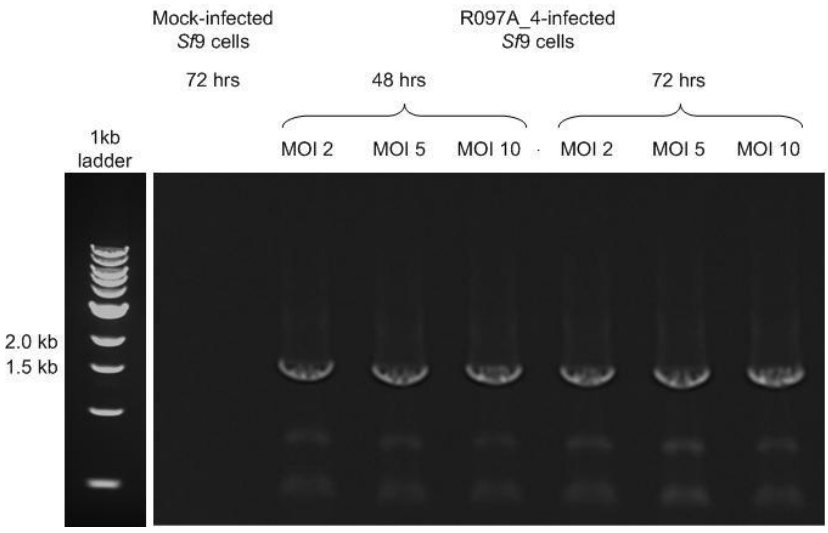


Figure 4.7. Analysis of dengue virus envelope gene mRNA transcripts produced in R097A_4 rBV-infected *Sf*9 cells

RT-PCR (section 2.3.1.1) was performed on total RNA extracted from rBV.R097A_4-infected *Sf*9 cells (section 2.3.2.2) using R097A_4 E gene primers (Table 4.1). Amplicons of 1.5 kb are consistent with the DV E gene. The original (unprocessed) gel images used to construct this figure are included electronically in Appendix J.

4.6 Concentration of dengue virus-like particles

A review of published literature revealed that previous studies concentrated rBV-infected insect cell culture medium prior to detection of recombinant flavivirus E (Bielefeldt-Ohmann *et al*., 1997; Jaaskelainen *et al*., 2003). Secreted flavivirus VLPs have been concentrated successfully from cell culture medium by ultracentrifugation through sucrose (Jaaskelainen *et al*., 2003; Schalich *et al*., 1996). For this study, clarified medium from rBV.R097A_4-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion (section 2.8.1). Negative control samples were included from similarly concentrated clarified medium from mock-infected *Sf*9 cells, and *Sf*9 cells infected with rBV.CCHFV-NP, rBV.EBOV-GP or rBV.EBOV-VP40. This was to confirm that the antibodies used for immunodetection were specific for flavivirus E, and did not detect other viral proteins or components of the insect cell culture medium. Samples were taken pre and post concentration and were analysed by SDS-PAGE and Western blot (sections 2.5.3.1 and 2.6). Immunodetection was conducted using antibodies to the His-tag or flavivirus E, using a recombinant His-tagged WNV E (Abcam) as a positive control. Immunodetection using the His-tag mAb only showed the presence of the His-tagged WNV E positive control. As before, R097A_4 E was not detected in the pre-concentration sample. It was also undetected in the concentrated sample (data not shown). However, immunodetection using the flavivirus E mAb was successful, as a protein of between 50 and 60 kDa was detected in the concentrated R097A_4 sample (Figure 4.8). The His-tagged WNV E positive control (50 kDa) was also detected using this mAb. This flavivirus E mAb was shown to be specific for flavivirus E, as it did not detect any proteins in samples from mock-infected, rBV.CCHFV-NP, rBV.EBOV-GP or rBV.EBOV-VP40-infected *Sf*9 cells. This confirmed that the rBV.R097A_4-infected *Sf*9 cells were producing DV E, and it was being secreted into the culture medium. These results also indicated that the DV E C-terminal His-tag was either not present or accessible to the His-tag mAb, even under the denaturing conditions used for SDS-PAGE and Western blot, suggesting that purification of the dengue VLPs using the E His-tags might not be possible.

At this point dengue VLPs were produced using the rest of the rBVs from section 4.4 (Table 4.2). Infections were performed using *Sf*9 cells infected at a MOI of 5, with proteins harvested from the culture medium 48 hours post infection, in accordance with the SDS-PAGE results in Figure 4.6. Proteins were concentrated by ultracentrifugation through a 20% sucrose cushion, and detected using the flavivirus E mAb (Figure 4.9). Immunodetection using the His-tag mAb was unsuccessful for all of the dengue VLP samples tested (data not shown). Quasispecies E from samples R097A_4, R097A_13 and R261A_1 were detected using the original batch of flavivirus E mAb, which was obtained from Abcam. However, Abcam removed this mAb from their catalogue so it had to be sourced from elsewhere. Subsequent suppliers provided the same mAb, but titre and stability were reduced compared to the original mAb sourced from Abcam. The mAb concentration used for immunodetection could not be increased to improve the sensitivity of detection as the dilution of mAb used for immunodetection was already 1:5. DV E from samples R097A_43, R203A_2, R203A_54, R261A_4, R107A_1 and R107A_22 were detected, although only as low intensity chemiluminescence, which was attributed to reliability issues with the replacement flavivirus E mAb.

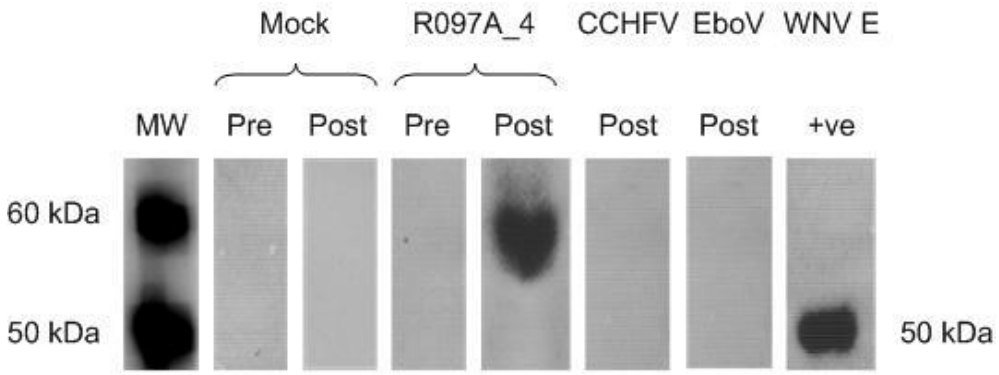


Figure 4.8 Immunodetection of dengue virus envelope glycoprotein from clarified medium from rBV.R097A_4-infected cells after concentration by ultracentrifugation

Immunodetection was performed using a flavivirus E mAb. Clarified medium from mock or rBV-infected (rBV.R097A_4, rBV.CCHFV-NP, or rBV.EBOV-GP and rBV.EBOV-VP40) *Sf*9 cells was analysed pre and post concentration by ultracentrifugation through a 20% sucrose cushion. The 50 and 60 kDa fragments of a molecular weight marker (MW) are shown for comparison, as is a 50kDa WNV E positive (+ve) control. The original (unprocessed) blot images used to construct this figure are included electronically in Appendix J.

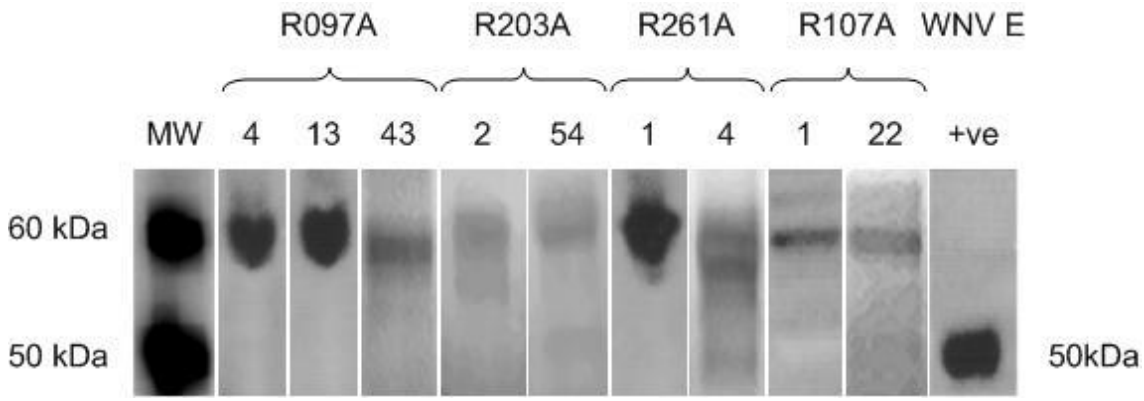


Figure 4.9. Immunodetection of dengue virus envelope glycoproteins from clarified medium from dengue rBV-infected cells after concentration by ultracentrifugation

Immunodetection was performed using a flavivirus E mAb on proteins secreted from *Sf*9 cells infected with rBV.R097A_4, rBV.R097A_13, rBV.R097A_43, rBV.R203A_2, rBV.R203A_54, rBV.R261A_1, rBV.R261A_4, rBV.R107A_1 and rBV.R107A_22. Clarified medium from dengue rBV-infected *Sf*9 cells was analysed post concentration by ultracentrifugation through a 20% sucrose cushion**.** The 50 and 60 kDa fragments of a molecular weight marker (MW) are shown for comparison, as is a 50kDa WNV E positive (+ve) control. The original (unprocessed) blot images used to construct this figure are included electronically in Appendix J.

4.7 Purification of dengue virus-like particles

The dengue VLPs were originally intended to be purified using immobilised metal affinity chromatography (IMAC) to bind the C-terminal His-tags on the DV E. However, as the DV E were unable to be detected using the His-tag mAb (sections 4.5 and 4.6), it was likely that purification using this method would be unsuccessful. To address this issue, it was decided to compare several protein purification methods to determine which method yielded the best results in terms of quantity and purity. The methods chosen included IMAC, immunoprecipitation (via Protein G and flavivirus E mAb) and ion exchange chromatography (using CaptoQ resin). Each method was evaluated using clarified medium from rBV.R097A_4-infected *Sf*9 cells which had been concentrated by ultracentrifugation through a 20% sucrose cushion (section 4.6). Similarly concentrated clarified medium from mock-infected *Sf*9 cells was included as a negative control. The R097A_4 sample was diluted to the limit of immunodetection using the flavivirus E mAb, and then purification was attempted using each of the methods.

4.7.1 Immobilised metal affinity chromatography

Purification by IMAC is dependent on the His-tags attached to the DV E binding to nickel ions in a sepharose gravity-flow column. Purifications were performed under native and denaturing conditions (section 2.8.2) but neither method yielded purified DV E (data not shown). This result was not unexpected given the inability to detect DV E from Western blots using the His-tag mAb (sections 4.5 and 4.6).

4.7.2 Immunoprecipitation

The immunoprecipitation method chosen relies on the binding of DV E to the flavivirus E mAb, which is bound to sepharose and protein G-coated magnetic beads. Protein G was chosen over Protein A as a ligand because the flavivirus E mAb is raised in mouse and the Protein G ligand has a higher affinity for mouse IgG than Protein A. Purifications were performed under native and denaturing conditions (section 2.8.3.1) using a protocol that cross-linked the flavivirus E mAb to the Protein G ligand in an attempt to elute only the purified DV E or dengue VLPs (depending on whether the denaturing or native protocols were used respectively), and not the flavivirus E mAb.

Using the flavivirus E mAb for immunodetection, a protein of between 50 and 60 kDa, which is consistent in size with DV E (54 kDa) was detected in the R097A_4 sample pre-purification (Figure 4.10). A protein of lower molecular weight (50 kDa) was observed in the post-purification native and denaturing elution fractions. This 50 kDa protein was also present in the post-purification native and denaturing elution fractions from the mock-infected negative control sample. The 50 kDa protein was not consistent with the size of DV E. To determine whether the flavivirus E mAb was being eluted and detected, it was analysed by SDS-PAGE, and Western blot, and was immunodetected as a 50 kDa protein (Figure 4.10; FV E mAb). This indicated that despite cross-linking the flavivirus E mAb to the Protein G ligand, the antibody was in fact being eluted (using both the native and denaturing protocols), and was subsequently detected by the anti-mouse IgG secondary pAb (pAb) used for immunodetection.

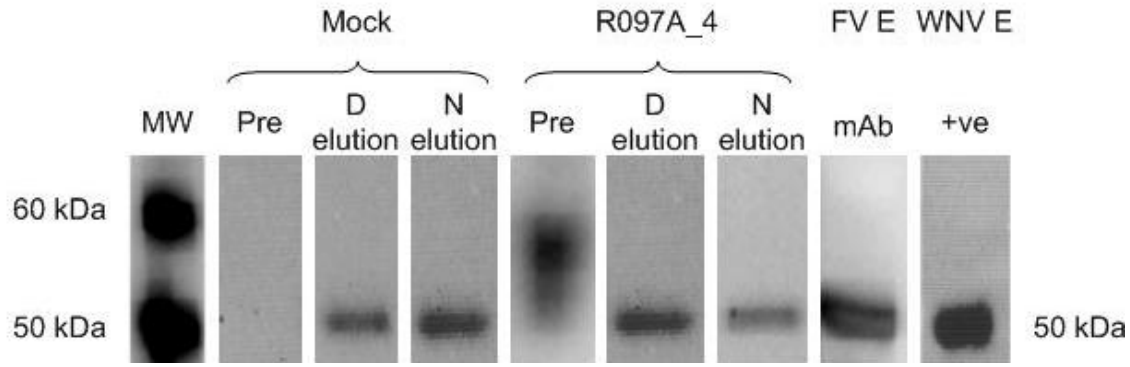


Figure 4.10. Immunodetection of dengue virus envelope glycoprotein from concentrated clarified medium from rBV.R097A_4-infected cells after purification by immunoprecipitation.

Immunodetection was performed using the flavivirus E mAb. Clarified medium from mock or rBV.R097A_4-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion**.** Pre-purification samples (Pre) were diluted 1/20 (to the limit of R097A_4 E immunodetection using the flavivirus (FV) E mAb) and then immunoprecipitated (section 2.8.3.1). The 50 and 60 kDa fragments of a molecular weight marker (MW) are shown for comparison, as is a 50kDa WNV E positive (+ve) control. The original (unprocessed) blot images used to construct this figure are included electronically in Appendix J.

Further investigation of the purification fractions revealed that the DV E in the R097A_4 sample was not binding to the flavivirus E mAb during the sample binding step, as it was observed in the non-bound fraction sampled after incubation of the sample with the Protein G-bound antibody (data not shown). Increasing the incubation times during cross-linking and sample binding appeared to solve the

problem of flavivirus E mAb elution, but did not improve sample binding (data not shown). With hindsight, it is probable that variability in the quality of the flavivirus E mAb was responsible for these observations (section 4.6, discussed further in sections 4.7.3 and 4.9.5).

4.7.3 Ion exchange chromatography

The principle of purification by ion exchange chromatography is that the charged protein to be purified binds to the ion exchange media, contaminants are washed away, and the purified protein is eluted using an ionic strength gradient. The isoelectric points and pH-dependent net charges of DV E were predicted based on the amino acid sequences using the EMBL WWW Gateway to Isoelectric Point Service (http://www3.embl.de/cgi/pi-wrapper.pl). This determined that DV E have a net negative charge at pH 7.0 and above. As a result, the CaptoQ strong anion exchange media was deemed to be the most suitable media for purification of the dengue VLPs.

Initial purification attempts were performed manually using a syringe (section 2.8.4), and the purification fractions analysed by SDS-PAGE and Western blot. Both flavivirus E mAb and baculovirus envelope glycoprotein (GP64) mAb were used for immunodetection to determine the success of the purification (Figure 4.11). From the SDS-PAGE results it was clear that the anion exchange column purification removed many of the impurities that were in the pre-purification sample. The greatest amount of purified protein was present in the fraction eluted using 20% sodium chloride (20% NaCl; Figure 4.11, panel A, lane 5) with decreasing amounts of protein also present in the 30 to 50% NaCl elution fractions. Using the flavivirus E mAb, a protein of 50 to 60 kDa was detected in the 20 to 50% NaCl elution fractions, and pre-purification sample (Figure 4.11, panel B). Consistent with the SDS-PAGE image, the greatest amount of purified protein was present in the 20% NaCl elution fraction.

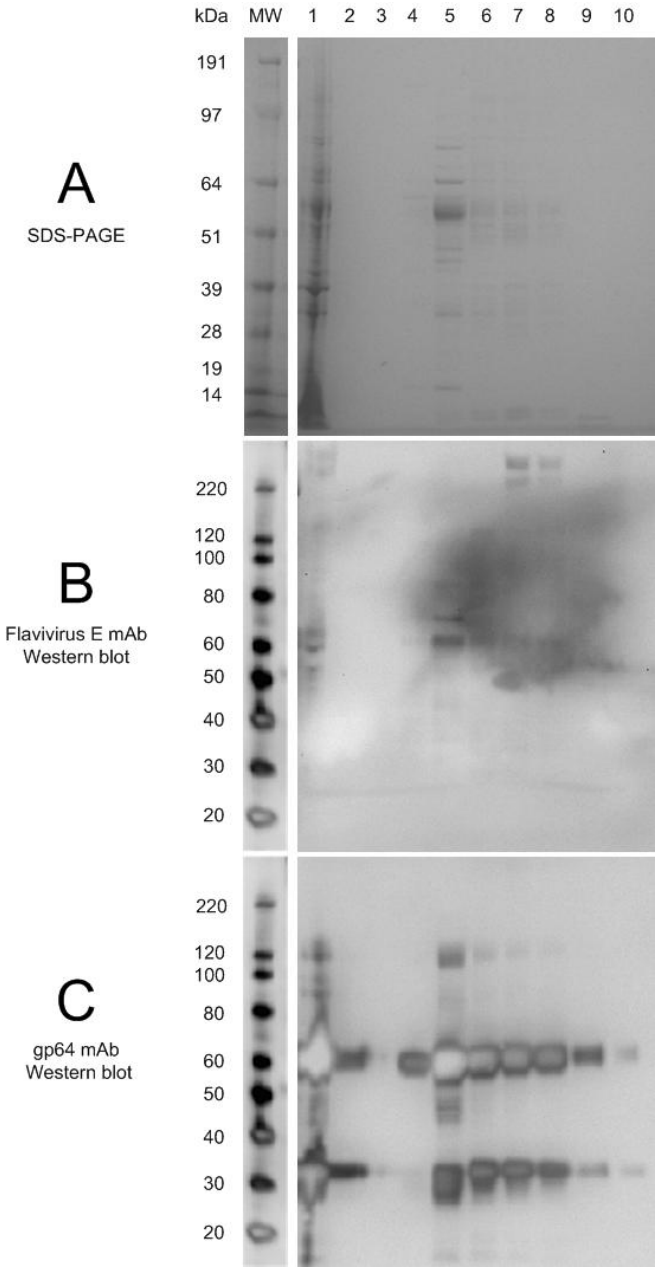


Figure 4.11. Analysis of R097A_4 anion exchange column purification fractions by SDS-PAGE and immunodetection from Western blots.

Immunodetection was performed using mAbs to flavivirus E or baculovirus GP64. Clarified medium from rBV.R097A_4-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion**.** Pre-purification sample (lane 1) was diluted 1/20 (to the limit of R097A_4 immunodetection using the flavivirus E mAb) and then purified by anion exchange column chromatography (section 2.8.4). Lane 1 shows R097A_4 pre-purification sample; Lane 2 shows flow-through (non-bound) from the sample loading step; Lane 3 shows flow-through from the sample wash step; lanes 4 to 9 are the fractions eluted using 10, 20, 30, 40, 50 and 100% NaCl respectively; lane 10 is flow-through from the column re-equilibration step. The original (unprocessed) gel and blot images used to construct this figure are included electronically in Appendix J.

Using the GP64 mAb, proteins consistent in size with GP64 (64 kDa) were detected in all fractions except the flow-through from the sample wash step (Figure 4.11, panel C, lane 3). Interestingly, GP64 was even detected in lanes where it was not visible on the SDS-PAGE gel, indicating that the GP64 mAb is extremely sensitive. As with the SDS-PAGE and immunodetection results obtained using the flavivirus E mAb, the greatest amount of purified protein was present in the 20% elution fraction (Figure 4.11, lane 5). There were also decreasing amounts of protein present in the 30-50%, 10% and 100% elution fractions and the flow-through fraction from the sample load step respectively. The presence of GP64 in the flow-through from the sample load step indicates that not all of the baculovirus present in the sample was binding to the column.

Despite using a new batch of flavivirus E mAb for immunodetection (Figure 4.11, panel B), the level of detection was poor compared to that achieved with the same pre-purification sample previously (see Figure 4.8). At this point it was decided to source a new antibody for the detection of the DV quasispecies E because this one had become so unreliable. Several primary antibodies were tested (Table 4.4), and a pAb raised in rabbit against DV types 1, 2, 3 and 4 was chosen based on its superior detection sensitivity compared to the other antibodies tested (data not shown).

Anion exchange column purification was chosen as the method that would be used to purify the dengue VLPs. This was because despite being shown to co-purify GP64 and dengue VLPs, a reduction in other background proteins had been demonstrated by SDS-PAGE analysis (Figure 4.11). It was intended that the dengue VLPs and baculovirus would be separated post anion exchange column purification, by immunoprecipitation.

Table 4.4. Antibodies tested for dengue virus envelope glycoprotein immunodetection

Antibody	Source	Immunogen
Mouse monoclonal to flavivirus E [FE1][1]	Abcam (Ab64059)	Flavivirus E
Mouse monoclonal to flavivirus E [FE1][2]	Pierce (MA1-71258)	Flavivirus E
Dengue antibody positive human control serum	Focus Diagnostics dengue IgG DxSelect ELISA kit (EL1500G)	DV
Mouse monoclonal to DV (pan) E [dengue 1-11(3)]*	Acris Antibodies, GmbH (AM01108PU)	DV types 1, 2, 3 and 4
Mouse monoclonal to DV E [dengue 1-11(3)]*	Antibodies-online (ABIN180831)	DV types 1, 2, 3 and 4
Rabbit polyclonal to DV types 1, 2, 3 and 4[3]	Abcam (Ab9200)	DV types 1, 2, 3 and 4
Mouse monoclonal to DV E	Abcam (Ab41349)	Purified DV type 2 E
Rabbit polyclonal to DV types 1, 2, 3 and 4	Abcam (Ab26837)	DV types 1, 2, 3 and 4
Mouse monoclonal to DV E [dengue 1-11(3)]*	Abcam (Ab9202)	DV types 1, 2, 3 and 4

After the original flavivirus E mAb[1] used for this work was discontinued by the supplier, a new flavivirus E mAb[2] was sourced to replace it. This antibody exhibited reduced detection sensitivity and stability compared to the original antibody. Commercially-available antibodies were tested until a replacement[3] was identified.

*detects a DV E specific band of 61 kDa under reducing conditions from Western blots. A weak secondary band of 80 kDa may also be apparent.

4.7.4 Removal of baculovirus from virus-like particle samples by immunoprecipitation

To assess the feasibility of removing the contaminating baculovirus from the dengue VLP samples, a baculovirus-derived Ebola VLP preparation was subjected to immunoprecipitation using the baculovirus GP64 mAb bound to Protein G and sepharose-coated magnetic beads. The Ebola VLP preparation was used in place of the dengue VLPs to avoid using valuable dengue VLP material whilst testing the immunoprecipitation method. Purifications were performed using a classic, non-cross-linking protocol (section 2.8.3.2) under native buffer conditions. Immunodetection was performed using both Ebola virus (data not shown) and baculovirus GP64 mAbs (Figure 4.12). The Ebola VLPs were detected in the non-bound and wash fractions from the sample binding step, but not the elution fractions (data not shown), confirming that they did not bind to the GP64 mAb. Baculovirus GP64 was detected in the elution fractions, indicating that the purification was at least partly successful. However, comparatively more GP64 was present in the non-bound and wash fractions

from the sample binding step, suggesting that only a small proportion of the baculovirus contamination was removed from the VLP preparation using this method. Therefore this method was unsuitable for removing the contaminating baculovirus from the dengue VLP samples. However, it was still hoped that the dengue VLPs and baculovirus could be separated post anion exchange column purification, by immunoprecipitation using a DV antibody.

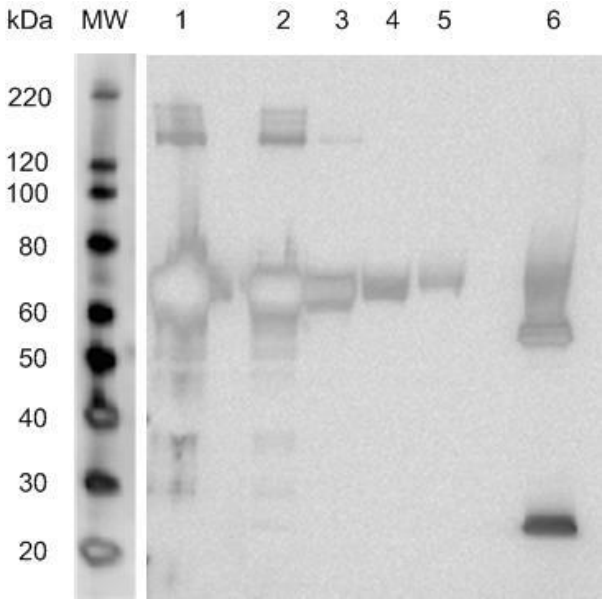


Figure 4.12. Immunodetection of GP64 in a baculovirus-derived Ebola VLP sample pre and post purification by immunoprecipitation.

Immunodetection was performed using a mAb to baculovirus GP64. Clarified medium from rBV.EBOV-GP and rBV.EBOV-VP40-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion, and purified by immunoprecipitation using the GP64 mAb (section 2.8.3.2). Proteins were eluted using a stepped ionic strength gradient of 10 to 100% NaCl. A molecular weight marker (MW) is shown for comparison. Lane 1 shows pre-purification Ebola VLPs; Lane 2 shows flow-through (non-bound) from the sample loading step; lanes 3, 4 and 5 show flow-through from sample wash steps 1, 2 and 3 respectively; lane 6 shows the elution fraction. The original (unprocessed) blot image used to construct this figure is included electronically in Appendix J.

4.8 Scaled-up production and purification of dengue virus-like particles

Production of all of the dengue VLPs (section 4.6, Figure 4.9) was scaled up from 50mL to 200mL cultures using the same conditions described previously (section 2.7.5). The clarified medium from rBV-infected insect cell culture was concentrated by ultracentrifugation through a 20% sucrose cushion prior to anion exchange column purification using an ÄKTA*FPLC* automated system (section 2.8.4). The automated liquid chromatography system was advantageous for two reasons. Firstly, using this system it was possible to reduce the sample flow-rate in an attempt to improve sample binding to the column. Secondly, it was thought that using a linear ionic elution gradient (rather than a stepped gradient as had been used previously when purifying manually) might enable further separation of the dengue VLPs and baculovirus into different elution fractions.

To determine whether reducing the sample flow-rate improved sample binding to the column, the concentrated R097A_4 sample was purified using sample flow rates of 1mL per minute and 0.1mL per minute during the sample load and wash steps. The purification fractions were analysed by SDS-PAGE and Western blot. Using the rabbit pAb to DV types 1, 2, 3 and 4 for immunodetection, a large amount of protein was shown to be present in the flow-through from the sample load and wash steps when the sample flow-rate was 1mL per minute (Figure 4.13; Panel A). Decreasing the sample flow-rate to 0.1mL per minute (Figure 4.13; Panel B) resulted in a reduction of protein in the flow-through from these steps, indicating that sample binding to the column had increased. It also resulted in improved purification quality as judged by visual inspection of protein staining after SDS-PAGE.

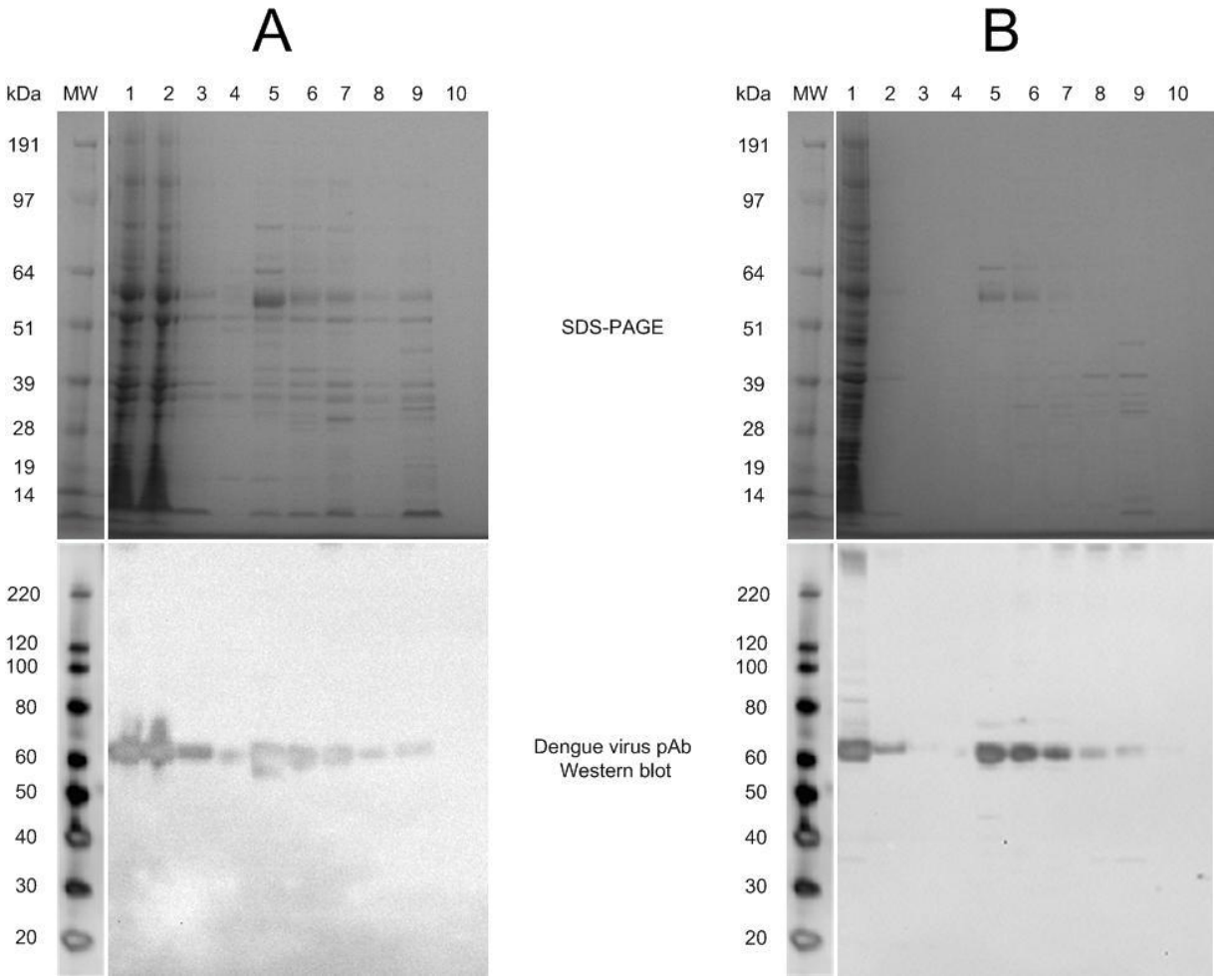


Figure 4.13. Analysis of R097A_4 anion exchange column purification fractions obtained using different sample flow-rates.

Immunodetection was performed using a rabbit pAb to DV types 1, 2, 3 and 4**.** Clarified medium from rBV.R097A_4-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion before anion exchange column purification. Proteins were eluted using a linear ionic strength gradient of 10 to 100% NaCl. Two molecular weight markers (MW) are shown for comparison. The purification in Panel A was performed using a sample flow-rate of 1.0mL per minute during the sample loading step; the purification in Panel B used a sample flow-rate of 0.1mL per minute. Lane 1 shows R097A_4 pre-purification; lanes 2 and 3 show flow-through from the sample loading and wash steps respectively; lanes 4 to 9 show the fractions eluted using 0-10, 10-20, 20-30, 30-40, 40-50 and 50-100% NaCl respectively; lane 10 shows flow-through from the column re-equilibration step. The original (unprocessed) gel and blot images used to construct this figure are included electronically in Appendix J.

To serve as a control sample representative of the background of baculovirus and insect cell-generated proteins present in the dengue VLP samples, the clarified medium from rBV.*lacZ*-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion and purified by anion exchange column purification in the same way as the dengue VLPs. This will be referred to from this point on as the

lacZ sample. Purification fractions from the lacZ and R097A_4 samples showed similar profiles after analysis by immunodetection using the GP64 mAb. This was to be expected, as they both contain baculovirus (Figure 4.14). However, they also unexpectedly demonstrated a similar profile after immunodetection using the rabbit pAb to DV types 1, 2, 3 and 4 (Figure 4.15.), indicating that this antibody was not specifically detecting the DV proteins. Clarified medium from mock-infected insect cells, concentrated by ultracentrifugation through a 20% sucrose cushion was not detected by this antibody (data not shown), suggesting that the source of the cross-reactivity was a baculovirus protein.

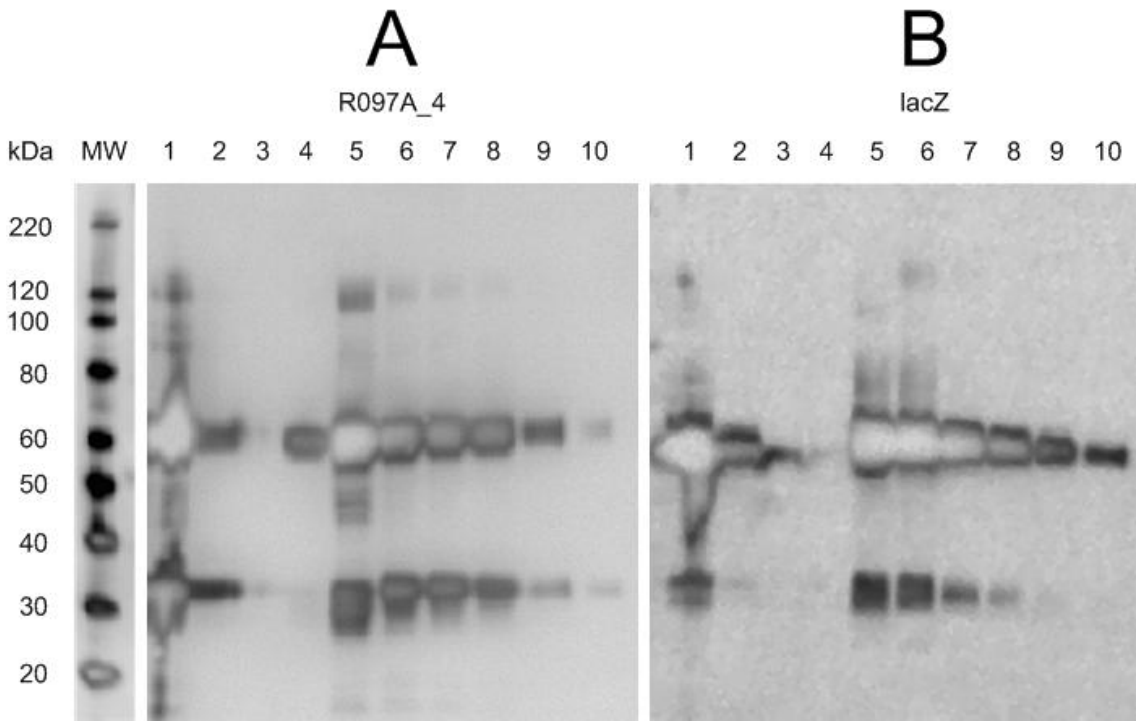


Figure 4.14. Analysis of R097A_4 and lacZ sample anion exchange column purification fractions using the baculovirus GP64 mAb.

Immunodetection was performed using a baculovirus GP64 mAb**.** Clarified medium from rBV.R097A_4 and rBV.*lacZ*-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion before anion exchange column purification (section 2.8.4). The sample flow-rate was 0.1mL per minute, and proteins were eluted using a linear ionic strength gradient of 10 to 100% NaCl. A molecular weight marker (MW) is shown for comparison. Panel A shows the purification fractions from the R097A_4 sample. Panel B shows the purification fractions from the lacZ sample. Lane 1 shows the pre-purification 20% sucrose pellet concentrated sample; lanes 2 and 3 show flow-through from the sample loading and wash steps respectively; lanes 4 to 9 show the fractions eluted using 0-10, 10-20, 20-30, 30-40, 40-50 and 50-100% NaCl respectively; lane 10 shows flow-through from the column re-equilibration step. The original (unprocessed) blot images used to construct this figure are included electronically in Appendix J.

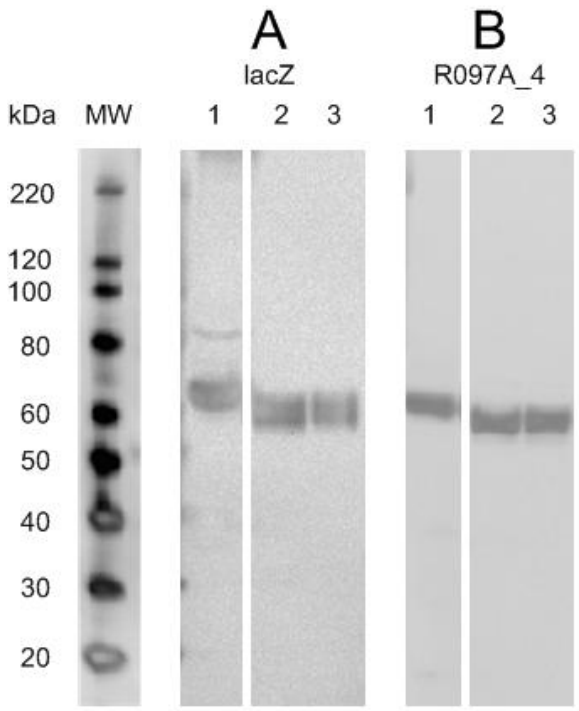


Figure 4.15. Analysis of R097A_4 and lacZ sample anion exchange column purification fractions using the rabbit pAb to dengue virus types 1, 2, 3 and 4.

Immunodetection was performed using a rabbit pAb to DV types 1, 2, 3 and 4**.** Clarified medium from rBV.R097A_4 and rBV.*lacZ*-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion before anion exchange column purification. The sample flow-rate was 0.1mL per minute, and proteins were eluted using a linear ionic strength gradient of 10 to 100% NaCl. A molecular weight marker (MW) is shown for comparison. Panel A shows the purification fractions from the lacZ sample. Panel B shows the purification fractions from the R097A_4 sample. Lane 1 shows the pre-purification 20% sucrose pellet concentrated sample; Lane 2 shows the 20% NaCl elution fraction; Lane 3 shows the 30% NaCl elution fraction. The original (unprocessed) blot images used to construct this figure are included electronically in Appendix J.

To determine whether the DV E could be reliably detected against the background of non-DV proteins, immunodetection was performed using three different commercially-available antibodies (mouse mAbs to the flavivirus and DV E, and the rabbit pAb to DV types 1, 2, 3 and 4) and baculovirus-derived samples (concentrated clarified medium from rBV.R097A_4, rBV.*lacZ*, and rBV.EBOV-GP and rBV.EBOV-VP40-infected *Sf*9 cells). All three antibodies detected a 60 kDa protein in all of the samples tested (Figure 4.16). As this protein was present in all three samples it was concluded not to be a DV protein. Additionally, using the flavivirus and DV E mAbs, a protein consistent in size with the predicted molecular weight of the DV E was detected in the R097A_4 sample but not the two non-dengue baculovirus-derived samples (lacZ and Ebola VLPs). This

protein was not detected using the rabbit pAb to DV types 1, 2, 3 and 4. From these results it was concluded that none of the antibodies tested were specific enough to unambiguously detect DV E from the denaturing Western blots used up to this point. It was hypothesised that the 60 kDa protein reacting non-specifically during immunodetection using the DV antibodies could be baculovirus GP64. Purified GP64 (kindly provided by Prof. Ian Jones, Reading University) was analysed using the mouse mAbs to the flavivirus and DV E, and the rabbit pAb to DV types 1, 2, 3 and 4. Baculovirus GP64 was identified as the source of antibody cross-reactivity (data not shown).

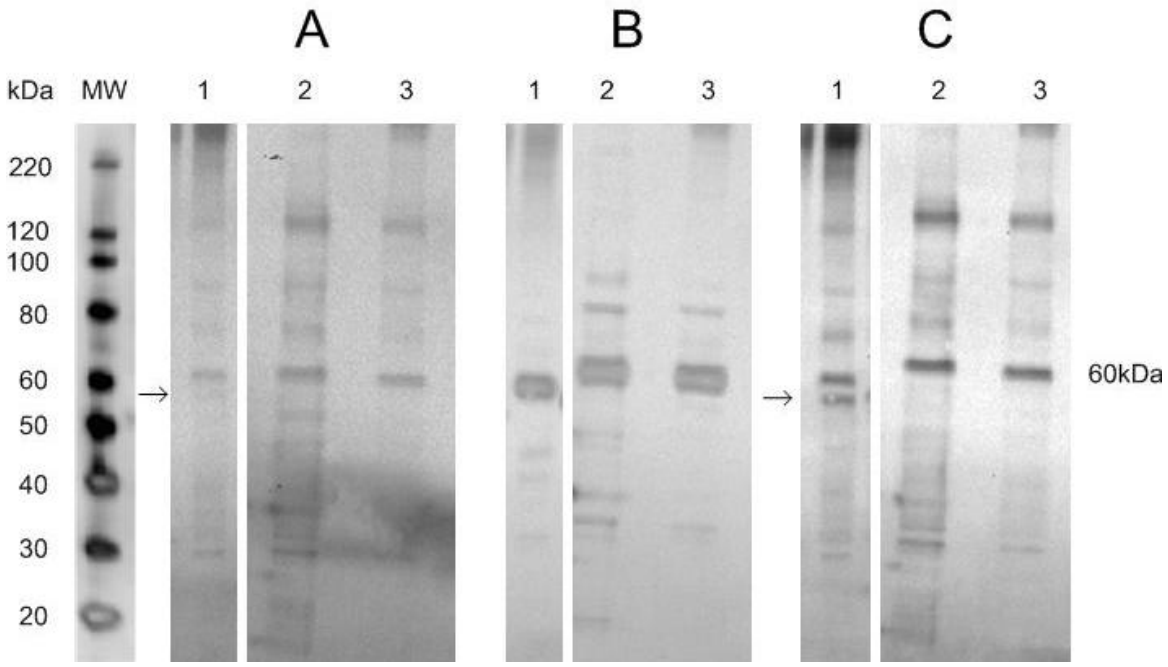


Figure 4.16. Analysis of baculovirus-derived R097A_4, lacZ and Ebola VLP samples using several different antibodies

Immunodetection was performed using mouse mAbs to the flavivirus and DV E (Panel A and Panel C respectively), and a rabbit pAb to DV types 1, 2, 3 and 4 (Panel B). For each sample, clarified medium from rBV-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion before analysis. A molecular weight marker (MW) is shown for comparison. Lane 1 shows R097A_4; Lane 2 shows Ebola VLPs; Lane 3 shows lacZ. The original (unprocessed) blot images used to construct this figure are included electronically in Appendix J.

To try to confirm the presence of DV E in the dengue VLP samples, other methods were attempted. These included glycoprotein deglycosylation analysis to attempt to distinguish between baculovirus GP64 and DV E. Endoglycosidase H cleaves only high

mannose N-linked glycans, whereas peptide:N-glycosidase F (PNGase F) is also able to cleave more complex glycans. Concentrated clarified medium from rBV.R097A_4 and rBV.*lacZ*-infected *Sf*9 cells, was denatured and digested using either endoglycosidase H or PNGase F, and analysed using the rabbit pAb to DV types 1, 2, 3 and 4 (Figure 4.17). Proteins of around 64 kDa were detected in both the R097A_4 and lacZ undigested samples. These proteins demonstrated increased mobility upon digestion with endoglycosidase H, and a further increase in mobility upon digestuib with PNGaseF. These findings are consistent with the size of baculovirus GP64 (64 kDa), which contains four N-linked glycans, at least one of which is endoglycosidase H resistant (Jarvis and Garcia, 1994). Several lower molecular weight proteins between 45 and 55 kDa were detected in the R097A_4 sample but were absent in the lacZ sample, one of which may represent DV E. The DV E contains two N-linked glycans, one of which is endoglycosidase H sensitive and one is resistant (Hacker, White, and de Silva, 2009), consistent with the electrophoretic shifts observed in the digested sample using the different glycosidases.

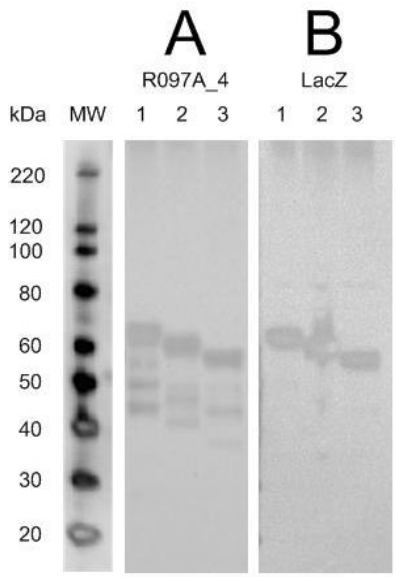


Figure 4.17. Glycosylation analysis of R097A_4 and lacZ samples using endoglycosidase H and PNGase F.

Immunodetection was performed using the rabbit pAb to DV types 1, 2, 3 and 4. For each sample, clarified medium from rBV-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion before analysis. A molecular weight marker (MW) is shown for comparison. Panel A is the R097A_4 sample. Panel B is the lacZ sample. Lane 1 shows undigested (glycosylated) sample; Lane 2 shows endoglycosidase H-digested sample; Lane 3 shows PNGase F-digested sample. The original (unprocessed) gel images used to construct this figure are included electronically in Appendix J.

To determine whether antibodies detecting conformational epitopes would exhibit less cross-reactivity with baculovirus GP64 than antibodies detecting linear epitopes, the R097A_4 and lacZ samples were analysed by native PAGE and Western blot (sections 2.5.3.2 and 2.6). Several DV antibodies were tested, including pAbs, and mAbs known to bind to conformation-dependent epitopes. Using the rabbit pAb to DV types 1, 2, 3 and 4 used previously, a protein was detected in the R097A_4 sample but not the lacZ sample, indicating that it is DV E (Figure 4.18). This suggests that the baculovirus GP64 cross-reactivity observed previously with this antibody may be dependent on epitopes that are only accessible under denaturing conditions.

The samples used for the native PAGE and Western blot analysis were first incubated with 1% triton X-100 for one hour to solubilise the lipid membranes. Previous attempts to perform native PAGE and Western blot analysis on the R097A_4 sample without first incubating with 1% triton X-100 resulted in immunodetection of a high molecular weight protein (data not shown). Taken together, these results support the presence of DV E in a lipid-enveloped form, consistent with VLPs.

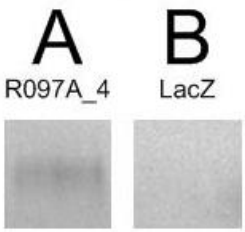


Figure 4.18. Analysis of R097A_4 and lacZ samples in their native conformation.

Immunodetection was performed using the rabbit pAb to DV types 1, 2, 3 and 4. For each sample, clarified medium from rBV-infected *Sf*9 cells was concentrated by ultracentrifugation through a 20% sucrose cushion before analysis. Samples were incubated in 1% triton X-100 for 1 hour to solubilise the lipid membranes, before native PAGE and immunodetection from a Western blot. Panel A shows the R097A_4 sample; Panel B shows the lacZ sample. The original (unprocessed) blot images used to construct this figure are included electronically in Appendix J.

4.8.1 Detection of dengue virus-like particles using transmission electron microscopy

To visualise the dengue VLPs, concentrated clarified medium from R097A_4 rBV-infected *Sf*9 cells was sent to Howard Tolley (HPA CEPR) for analysis by transmission

electron microscopy (section 2.10; Figure 4.19, panels D and E). Previous studies demonstrated that flavivirus virions are observed as spherical structures of approximately 50nm in diameter, whilst flavivirus VLPs range in diameter between 30 and 55nm (Schalich 1996; Lin & Wu 2005). Spherical structures consistent in size with published transmission electron micrographs of flavivirus VLPs (Figure 4.19, panels B and C) were observed in the dengue VLP sample (Figure 4.19, panels D and E). Whilst similar structures were also observed in a baculovirus-derived Ebola VLP sample (Figure 4.19, panel F) included as a negative control, these were fewer in number and were smaller (less than 25nm) than those in the dengue VLP sample.

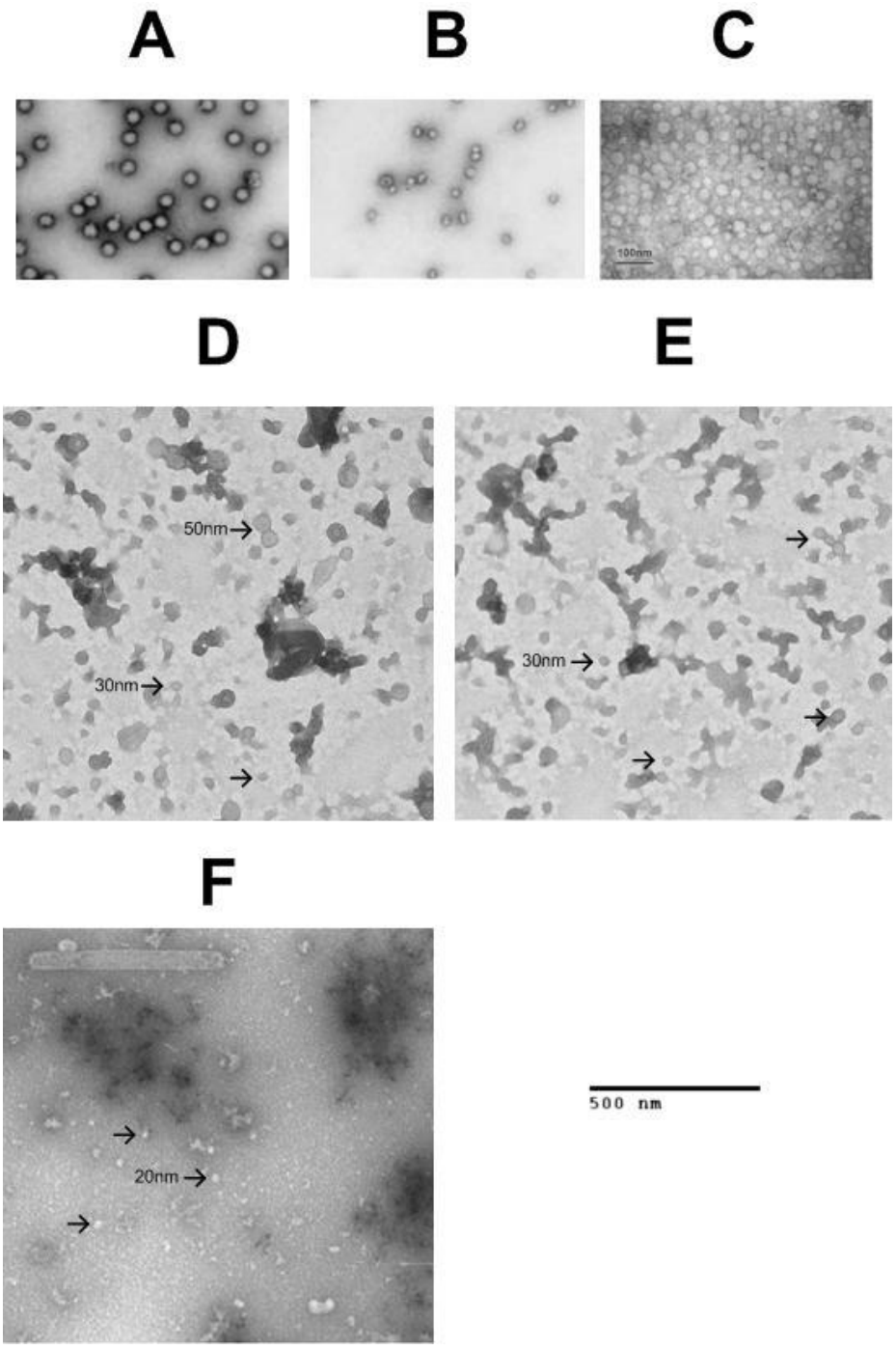


Figure 4.19. Transmission electron micrographs of flavivirus virions and VLPs stained with uranyl acetate.

All images are at the same magnification, and a 500nm scale bar is shown. Panels A and B are purified TBEV virions and VLPs respectively (Schalich *et al*., 1996). Panel C is purified JEV VLPs (Lin and Wu, 2005). Panels D and E are unpurified dengue VLPs from this study (clarified medium from rBV.R097A_4-infected *Sf*9 cells, concentrated by ultracentrifugation through a 20% sucrose cushion). Examples of dengue VLPs are marked with arrows, and the diameters shown. Panel F is unpurified Ebola virus VLPs, included as a negative control (kindly provided by Dr. Kevin Richards, HPA CEPR).

4.9 Discussion

The objective for this chapter was to produce recombinant DV E from those quasispecies E genes containing mutations with the potential to affect virulence or pathogenicity, as identified in the previous chapter. It was decided that the DV E would be produced in the form of secreted VLPs, consisting of DV type 1 membrane and quasispecies E. This was to replicate as closely as possible the natural antigenic presentation of E within the dengue virion, as flavivirus VLPs have been shown to exhibit the same antigenic and oligomeric structure as virions (Schalich *et al*., 1996). Studies using recombinant vaccinia viruses encoding combinations of JEV prM, E and NS1 genes, demonstrated that the co-expression of prM and E genes induced the highest levels of neutralising antibody and protection in mice, and these antibody levels correlated with the ability of the recombinant viruses to induce the synthesis of extracellular VLPs *in vitro* (Konishi *et al*., 1991; Konishi *et al*., 1992). DV E in the form of VLPs is more resistant to proteolytic degradation than secreted E alone (Sugrue *et al*., 1997). VLPs secreted into the insect cell culture medium represent a more homogenous VLP population than intracellular VLPs, which are at different stages of processing and maturation. The production of DV quasispecies E as VLPs also provides a model for studying the effects of the quasispecies E gene mutations on E conformational changes and membrane fusion in response to low pH.

A eukaryotic expression system was required to enable the appropriate post-translational modifications for the production of authentic secreted DV E. The use of a yeast expression system was discounted because its level of post-translational modification is less than can be achieved using baculovirus or mammalian systems. Yeast cells have also been shown to accumulate dengue VLPs intracellularly as secretion is prevented (Sugrue *et al*., 1997). Despite the differences in post-translational processing between the baculovirus and mammalian expression systems, and since DV is able to infect and replicate naturally in both mosquito (insect) and human (mammalian) cells, the baculovirus expression system was chosen over a mammalian expression system based on its comparatively higher protein yields, ease of use and lesser expense.

4.9.1 Cloning of dengue virus pre-membrane and envelope genes into the baculovirus transfer vector

The pBAC-2cp baculovirus transfer vector was used for the production of rBVs as it contains an ATG start codon and incorporated a His-tag sequence onto the C-terminal of the DV E genes inserted into the MCS, which was intended to be used for VLP purification. Rather than using the corresponding prM genes from the patient samples, the same DV type 1 prM gene was ligated into the baculovirus transfer vector with each DV quasispecies E gene. This was to ensure any differences between the dengue VLPs were not attributed to differences in prM. Ligation-dependent cloning of DV prM and E genes with restriction enzyme-digested 5' and 3' ends enabled directional cloning of the genes into the MCS so that the genes were inserted in the correct order (prM followed by E as in the DV genome), and in-frame with the pBAC-2cp start codon and C-terminal His-tag coding sequences. This was confirmed for each of the clones by sequencing.

4.9.2 Production of recombinant baculoviruses containing dengue virus pre-membrane and envelope genes

Transfer vectors containing DV prM and quasispecies E genes were co-transfected with baculovirus DNA into insect cells to generate nine rBVs (rBV.R097A_4, rBV.R097A_13, rBV.R097A_43, rBV.R203A_2, rBV.R203A_54, rBV.R261A_1, rBV.R261A_4, rBV.R107A_1 and rBV.R107A_22), containing quasispecies consensus and variant E genes from four of the dengue patient samples. To serve as a control for the transfection reactions and rBV titrations, rBV.*lacZ* was also generated. In subsequent experiments, culture medium from rBV.*lacZ*-infected cells was used as a control to represent the background of non-DV proteins present in the dengue VLP preparations. After successive rBV amplifications and titrations, the presence of the DV prM and E genes in the rBVs was confirmed by PCR.

4.9.3 Optimisation of dengue virus-like particle production

Optimisation experiments were conducted to determine the optimal cell line, multiplicity of infection (MOI) and post infection interval before harvesting the dengue VLPs. *Sf*9 and *Sf*21 cells were infected with rBV.R097A_4 at MOIs of 2, 5 and 10, and

samples of lysed cell pellet and clarified culture medium were analysed at 48 and 72 hours post infection. SDS-PAGE analysis revealed proteins of approximately 64 kDa and 54 kDa in the clarified medium from rBV-infected cells, but these were not visible in mock-infected samples. Comparison of these results to similar in-house experiments using rBVs containing genes for non-DV proteins (rBV.EBOV-GP, rBV.EBOV-VP40 and rBV.CCHFV-NP) revealed that the 64 kDa protein was present in all of these samples, suggesting that this was a baculovirus protein rather than a DV protein. The 54 kDa protein was not present in samples from the non-dengue rBVs, and was consistent with the size of DV E.

To confirm the presence of the His-tagged DV E, immunodetection was performed using a mAb to the His-tag. His-tagged protein was not detected in any of the samples. Only the His-tagged WNV E positive control protein was visible, confirming that the method and reagents used were working effectively. To overcome this problem, a different mAb was sourced to directly detect the DV quasispecies E in case the His-tag was either not present or accessible. The clarified medium from rBV-infected cells was also further concentrated by ultracentrifugation through a sucrose cushion before re-analysing, in case production of only low quantities of dengue VLPs had made the His-tags undetectable. Ultracentrifugation through a sucrose cushion had been used previously to concentrate secreted flavivirus VLPs (Jaaskelainen *et al*., 2003; Schalich *et al*., 1996).

As the recombinant dengue VLPs were not able to be detected using this His-tag mAb, the presence of mRNA for DV E was confirmed by RT-PCR in the rBV.R097A_4-infected insect cells at all MOIs and both post infection harvest intervals. The presence of mRNA for the DV E was not confirmation of protein production but was a positive indication that the gene was transcribed, ready for protein synthesis.

Optimal conditions for the production of dengue VLPs were determined based on the SDS-PAGE results, as immunodetection of DV E using the His-tag mAb had not been successful. Subsequent production of dengue VLPs using rBV.R097A_4, rBV.R097A_13, rBV.R097A_43, rBV.R203A_2, rBV.R203A_54, rBV.R261A_1, rBV.R261A_4,

rBV.R107A_1 and rBV.R107A_22 was performed in *Sf*9 cells infected at a MOI of 5, and culture medium harvested at 48 hours post infection. Although no difference in protein production was observed between the cell-lines used in the optimisation experiments, *Sf*9 cells were chosen due to their increased replication rate and ability to reach higher cell densities compared to the *Sf*21 cells they are derived from (Smith and Summers, 1987).

4.9.4 Concentration of dengue virus-like particles

Clarified medium from rBV.R097A_4-infected insect cell culture was concentrated by ultracentrifugation through a 20% sucrose cushion before immunodetection using either His-tag or flavivirus E mAbs. DV E was not detected using the His-tag mAb, but was detected using the flavivirus E mAb. At this stage of the work, the flavivirus E mAb was shown to be specific for DV E, as it did not detect any proteins in concentrated clarified medium from rBV.CCHFV-NP or rBV.EBOV-GP and rBV.EBOV-VP40-infected cells.

These results indicated that the C-terminal His-tags on the R097A_4 DV E were either not present or not accessible to the His-tag mAb, even under the denaturing conditions used for SDS-PAGE. This suggested that purification of the dengue VLPs using the DV E His-tags as originally intended might not be possible. It has been proposed that the C-terminal of flavivirus E may contain a signal for retention of this protein in intracellular compartments (Allison *et al*., 1995). Cleavage of this signal may be required for VLP assembly and/or secretion, which would explain the inability to detect the C-terminal His-tag in the secreted VLPs under denaturing conditions. In support of this, C-terminal His-tags have been used previously to purify JEV E and prM produced separately from cell lysates, but failed in purification of these proteins from insect cell culture medium (Yang *et al*., 2005).

At this point recombinant VLPs were concentrated and detected for the rest of the dengue VLPs (R097A_13, R097A_43, R203A_2, R203A_54, R261A_1, R261A_4, R097A_1 and R107A_22; Figure 4.9).

4.9.5 Purification of dengue virus-like particles

Three methods of purification were compared using the R097A_4 dengue VLP sample and native conditions to preserve the structure of the dengue VLPs. These were: immobilised metal affinity chromatography (IMAC), immunoprecipitation (Protein G-bound flavivirus E mAb) and anion exchange chromatography (CaptoQ resin). The IMAC purification method was unsuccessful. This was expected because this purification method relies on the presence of the His-tag on the DV E, which had not been detected using the His-tag mAb.

The immunoprecipitation purification method was also unsuccessful despite optimisation of antibody cross-linking to protein G and sample incubation durations. This was shown to be the result of poor sample binding to the flavivirus E mAb. This was the first indication of problems with the replacement flavivirus E mAb. This antibody had been originally sourced from Abcam, and was shown to be sensitive enough to detect DV E once samples had been concentrated by ultracentrifugation through a 20% sucrose cushion (Figure 4.8). This antibody was also shown to be specific, with no cross-reactivity observed with samples from infections with rBVs containing genes for non-DV proteins (rBV.CCHFV-NP or rBV.EBOV-GP and rBV.EBOV-VP40). When this antibody was discontinued from the Abcam catalogue, a replacement monoclonal antibody from the same hybridoma clone was subsequently sourced from Thermo Fisher (Pierce), but the antibody titre was much lower and antibody stability varied between lots compared to when sourced from Abcam. It was intended that the immunoprecipitation method would be revisited once a new DV E antibody was sourced.

Anion exchange column purification of the dengue VLPs was attempted, and performed manually initially using syringes and a stepped ionic elution gradient. Analysis of the SDS-PAGE results showed a reduction in the presence of background protein in the elution fractions compared to the pre-purification sample (Figure 4.11). Proteins of between 51 and 64 kDa were clearly visible and detected using the flavivirus E mAb. However, immunodetection of the same samples using the

baculovirus GP64 mAb demonstrated co-purification of baculovirus in the same elution fractions as the dengue VLPs.

The level of detection achieved using the replacement flavivirus E mAb was poor compared to that achieved previously using the original antibody and the same R097A_4 pre-purification sample (Figure 4.8 and Figure 4.11 respectively). Comparison of the SDS-PAGE results with those from previous experiments using the R097A_4 sample (data not shown), indicated that the loss of sensitivity during detection was not due to sample degradation, but was likely due to the flavivirus E mAb. As a result of the declining quality of this antibody it was decided that a new antibody was required for the detection of DV E. After testing several potential replacement antibodies, a pAb raised in rabbit against DV types 1, 2, 3 and 4 was chosen based on its superior detection sensitivity compared to the other antibodies tested (Table 4.4).

Attempts to remove the baculovirus virions from baculovirus-derived VLP preparations by immunoprecipitation using a baculovirus GP64 mAb (bound to Protein G-coated magnetic beads) were unsuccessful, despite the high sensitivity of the antibody. Immunodetection from Western blots showed that although some of the baculovirus virions were immunoprecipitated, the vast majority did not bind to the antibody and remained with the VLPs in the un-bound fraction during sample loading (Figure 4.12). Some of the GP64 is present in the dengue VLP preparation in the form of baculovirus virions which might be able to be separated from the dengue VLPs by further purification, for example by size-exclusion, immunoprecipitation (once a suitable antibody is sourced) or further sucrose density gradient ultracentrifugation. However, some of the GP64 could be present within the envelope of the dengue VLPs due to insertion into the lipid bilayer during VLP formation, so would be unable to be separated from the DV proteins without dissolving the dengue VLPs.

4.9.6 Scaled-up production, purification and detection of dengue virus-like particles

To produce enough dengue VLPs for analysis, production of the dengue VLPs was scaled-up by increasing the volume of rBV-infected insect cell culture from 50mL to 200mL. To concentrate and purify the dengue VLPs, ultracentrifugation through a 20% sucrose cushion, and anion exchange column purification were used as described previously. Unlike previous manual purifications, anion exchange column purification was this time performed using an ÄKTA*FPLC* automated system. This was to enable further optimisation of the sample loading step, by reducing the sample flow-rate, thereby increasing sample binding to the column. Also, it was hoped that by using a low sample flow-rate, and a linear rather than stepped ionic elution gradient, further separation of the dengue VLPs and baculovirus into different elution fractions might be possible. As a control representative of the non-dengue proteins present in the baculovirus-derived samples, clarified medium from rBV.*lacZ*-infected cultures was concentrated and purified in the same way as described for the dengue VLP preparations. Anion exchange column purified lacZ samples were analysed alongside corresponding dengue samples by immunodetection using the baculovirus GP64 mAb, and both were shown to have similar profiles. This confirmed that the lacZ sample was an appropriate background control for the baculovirus proteins present in the dengue VLP samples. Unexpectedly, comparison of the same lacZ and dengue samples using the rabbit pAb to DV types 1, 2, 3 and 4, also demonstrated similar immunodetection profiles for both samples. This indicated that the rabbit pAb to DV types 1, 2, 3 and 4 was cross-reacting with a non-DV protein of around 64 kDa. This left the anion exchange column purification results from the scaled-up cultures (section 4.8) in question, as it was not clear what protein was being detected.

Subsequent testing of the rabbit pAb to DV types 1, 2, 3 and 4, the flavivirus E mAb used previously, and another commercially available mAb to DV E, showed all three antibodies cross-reacted with a non-DV protein consistent in size with the known molecular weight of baculovirus GP64 (64 kDa). However, the two mAbs tested were both able to detect an additional protein with a lower molecular weight in the

R097A_4 sample, but not the lacZ sample. This protein was consistent in size with the known molecular weight of DV E (54 kDa). Baculovirus GP64 was confirmed as the source of the antibody cross-reactivity by immunodetection of purified GP64 using the mouse mAbs to the flavivirus and DV E, and the rabbit pAb to DV types 1, 2, 3 and 4. To the best of my knowledge, cross-reactivity between flavivirus antibodies and GP64 has not been previously reported, and further purification of VLP preparations may resolve the issue. Antibody cross-reactivity with GP64 may have been expected had the antibodies been generated using baculovirus-derived antigen, but the antigens used to produce the antibodies used in this experiment were derived from live flaviviruses grown in mosquito cell lines. It is feasible that there could be antibody cross-reactivity to insect cell-derived protein because of the use of mosquito cell lines to culture the flaviviruses used for antibody production.

To assess whether the observed antibody cross-reactivity was due to sequence homology between DV E and baculovirus GP64, their respective amino acid sequences were aligned, which revealed a reasonable degree of sequence similarity between them (Figure 4.20). This is not unexpected as they are both major viral envelope glycoproteins responsible for receptor-binding and membrane fusion, and both bind to insect cells. A review of the literature has shown that a crystallographically-determined structural model for pre-fusion GP64 has not been eluded to date, so a structural comparison was not able to be made with the DV E pre-fusion structure. Structural domain *III* of flavivirus E is known to contain an immunoglobulin C-like fold common to cell adhesion proteins that has also been identified on the outer surface of an insect virus, although this virus was non-enveloped (Rey *et al*., 1995).

Analysis of protein glycosylation was performed using glycosidases to attempt to distinguish between baculovirus GP64 and DV E in the lacZ and R097A_4 samples. The glycosylation profile produced by the 64 kDa protein in both samples was consistent with that of GP64, which contains four N-linked glycans, at least one of which is endoglycosidase H resistant (Jarvis and Garcia, 1994). Several proteins with lower molecular weights of between 45 and 55 kDa were detected in the R097A_4 sample,

but were absent in the lacZ sample, one of which could represent DV E. The DV E contains two N-linked glycans, one of which is endoglycosidase H sensitive and one is resistant (Hacker, White, and de Silva, 2009), consistent with the electrophoretic shifts observed in the digested sample using the different glycosidases.

The lacZ and R097A_4 samples were also analysed by native PAGE and immunodetection from Western blots, using the rabbit pAb to DV types 1, 2, 3 and 4. A protein was detected in the R097A_4 sample but not the *lacZ* sample, indicating that it is DV E (Figure 4.18). This suggested that the GP64 cross-reactivity observed previously with this antibody is dependent on epitopes that are only accessible under denaturing conditions, so there may be a better chance of conclusively detecting the dengue VLPs using methods that enable the proteins to remain in a native conformation, for example by ELISA or immunofluorescence.

Figure 4.20. Amino acid sequence alignment of AcMNPV GP64 and the dengue virus envelope glycoprotein.

Amino acid sequences were obtained from NCBI for baculovirus GP64 (AcMNPV; GenBank accession L22858) and DV type 3 E (GenBank accession NC_001475), and aligned using ClustalW (section 2.2.4). Amino acids are colour-coded according to physiochemical properties, with red representing nonpolar, green representing polar, pink representing basic, and blue representing acidic. Aligned matching amino acids are marked *****, and amino acids with similar properties are marked **.** or **:** depending on the degree of similarity (**:** indicating more similarity than **.**).

4.9.7 Detection of dengue virus-like particles using transmission electron microscopy

Transmission electron micrographs of the R097A_4 dengue VLP sample confirmed that spherical structures consistent in size with published transmission electron micrographs of flavivirus VLPs were present (Figure 4.19, panels D and E). Whilst similar structures were also observed in a baculovirus-derived Ebola VLP sample (Figure 4.19, panel F) included as a negative control, these were fewer in number and smaller in diameter than those observed in the dengue VLP sample. Confirmation that

the spherical structures seen in the dengue VLP sample were actually dengue VLPs could be obtained using immunogold labelling. Provided a primary antibody could be sourced that was specific for the dengue VLPs, and did not cross-react with baculovirus, a gold-conjugated secondary antibody could then be added and the sample re-analysed by transmission electron microscopy. Accumulation of gold particles on the spherical structures would confirm that the particles observed are in fact the dengue VLPs.

4.10 Conclusions and summary

The intention for this part of the study was to generate rBVs containing the DV type 1 prM gene and the quasispecies E genes identified in Chapter 3 (section 3.10; R097A_4, R097A_13, R097A_43, R203A_2, R203A_54, R261A_1, R261A_4, R107A_1 and R107A_22). These would be used to produce dengue VLPs for each of the E gene variants, which could be purified in preparation for use in studies to determine the effect of the mutations on protein function, and assessment of their value as diagnostic tools.

The DV quasispecies E genes identified in the previous chapter were each successfully cloned alongside the DV type 1 prM gene into the baculovirus transfer vector, pBAC-2cp. Co-transfection of pBAC-2cp containing the inserted genes with baculovirus DNA into insect cells resulted in the production of rBVs, which were subsequently amplified and titred. Optimisation experiments were conducted to determine the optimal cell line, multiplicity of infection (MOI) and post infection interval before harvesting the dengue VLPs. However, problems were encountered in detecting the dengue VLPs using the C-terminal His-tags on the DV E, which meant that optimised expression conditions could only be determined based on the SDS-PAGE results.

Clarified medium from dengue rBV-infected cells (containing secreted dengue VLPs) was concentrated by ultracentrifugation and re-analysed by immunodetection. Again, the DV E were not detected using the His-tag mAb, but proteins of 50 to 60 kDa, consistent with the known molecular weight of DV E (54 kDa) were detected for all of the dengue VLP samples using a mouse mAb demonstrated to be specific for flavivirus E.

Purification of the dengue VLPs as originally intended using IMAC to bind the C-terminal His-tags on the DV E was thought to be unlikely, given the failure to detect the DV E using the His-tag mAb. As a result three methods of purification were attempted, namely IMAC, immunoprecipitation and anion exchange column chromatography. Purification by IMAC proved to be unsuccessful as predicted. Purification by immunoprecipitation was hampered by problems with the quality of

the flavivirus E mAb used. Preliminary anion exchange column purification results were promising, but the flavivirus E mAb was no longer sensitive and reliable enough for use as determined by repeat analysis of the same sample over time (data not shown), and co-purification of baculovirus was demonstrated. A rabbit pAb to DV types 1, 2, 3 and 4 was sourced to analyse the anion exchange column purification fractions, but on closer examination, it was found to cross react with a non-DV protein in all of the baculovirus-derived samples tested. The source of the cross-reactivity was determined to be baculovirus GP64. Similar cross-reactivity was subsequently demonstrated using two other commercially available mouse mAbs (flavivirus and DV E antibodies), although an additional protein of lower molecular weight was also detected. The size of this additional protein was consistent with the known molecular weight of DV E (54 kDa), and it was not evident in non-DV rBV-derived samples. There are not many commercially available antibodies for DV proteins, and of those tested for this study, poor sensitivity and stability were commonly encountered. These results identify a need for better quality DV-specific antibodies to be made available to the wider scientific community. Other purification methods would need to be attempted to obtain dengue VLP preparations that do not contain baculovirus, which should solve the problem of antibody cross-reactivity. Additionally, the use of virus-free expression systems for protein production (using plasmid rather than viral gene-delivery vectors) would remove the potential for virus-contamination of the dengue VLP preparations, thus simplifying purification.

DV E was detected in clarified insect cell culture medium from rBV-infected cultures concentrated by ultracentrifugation through a 20% sucrose cushion. Detection of DV E in subsequent purified samples is ambiguous due to antibody cross-reactivity. DV E was also detected in the R097A_4, but not lacZ samples analysed by native PAGE and Western blots. These data also provided supportive evidence that the DV E are present in a lipid enveloped form consistent with VLPs. This was further supported by transmission electron micrographs, which showed spherical structures consistent in size with published transmission electron micrographs of flavivirus VLPs. Future detection of dengue VLPs may need to be undertaken by ELISA, as this may circumvent

the antibody cross-reactivity problem as it allows the proteins to remain in their native conformation.

CHAPTER 5. ANALYSIS OF DENGUE VIRUS-LIKE PARTICLES TO DETERMINE THE EFFECTS OF THE QUASISPECIES ENVELOPE GENE MUTATIONS

5.1 Introduction

Dengue virus (DV) infectivity is directly related to the ability of the DV envelope glycoprotein (E) to evade neutralisation by antibodies, attach to target cells within the host, undergo low pH-induced conformational change, and mediate fusion of virus and host cell membranes. Mechanisms of DV neutralisation by host antibodies include preventing virus-cell attachment and inhibiting membrane fusion (Roehrig, 2003; Roehrig, Bolin, and Kelly, 1998). Single nucleotide mutations within the E gene that cause amino acid substitutions in E have been shown to be associated with altered DV virulence and disease pathogenicity (Cecilia and Gould, 1991; Guirakhoo *et al*., 2004; Hahn *et al*., 1987; Hasegawa *et al*., 1992; Holzmann *et al*., 1990; Jiang *et al*., 1993; Nitayaphan *et al*., 1990). The amino acid substitutions caused by these mutations cluster on the outer surface of the virion and at the interfaces between E structural domains (Rey *et al*., 1995). They are thought to act either through altering accessibility to neutralising antibodies, or through direct modulation of the processes of attachment to host cells or membrane fusion (Rey *et al*., 1995).

Although a definitive antibody-independent cellular receptor for DV has yet to be elucidated, dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN) and other lectins have been proposed as DV receptors on DCs, endothelial cells and macrophages (Miller *et al*., 2008; Navarro-Sanchez *et al*., 2003; Tassaneetrithep *et al*., 2003). Cryoelectron microscopy of the DV E-DC-SIGN complex has shown interaction of the lectin with the N-linked glycan at Asn-67 (Pokidysheva *et al*., 2006). The DV E glycans have been shown to play an important role in infectivity and virus growth in mammalian cells (Bryant *et al*., 2007; Mondotte *et al*., 2007). Highly sulphated heparan sulphate (HSHS), a glycosaminoglycan on the surface of target cells has also been shown to bind DV E (Chen *et al*., 1997), and is thought to concentrate the virus on the cell surface, enabling subsequent higher-affinity interactions with specific receptors (Martinez-Barragan and del Angel, 2001). Two HSHS-binding sites have been proposed for DV E at positions E284 to 310, and E386 to 411 (Chen *et al*., 1997). The first of these is on the distal face of E domain *III*, where

gene mutations leading to amino acid substitutions in several flaviviruses have been shown to affect virus virulence (Rey *et al*., 1995).

DV E is a class II fusion protein that exists as a dimer in the pre-fusion state. Cleavage of the pre-membrane glycoprotein (prM) by furin is required to prime E for fusion. In response to the acidic environment of the endosome, the E dimer undergoes an irreversible conformational change to a trimeric form that exposes the previously inaccessible fusion peptide, and enables fusion of virus and host cell membranes (Modis *et al*., 2004). E gene mutations that cause amino acid substitutions at the interface of E structural domains *I* and *II* have been shown to affect fusion by altering the pH required for fusion-induced conformational change (Modis *et al*., 2003).

In addition to their use in studying the effect of E gene mutations on protein function, the dengue virus-like particles (VLPs) produced in Chapter 4 may be a useful source of antigen for the serodiagnosis of DV infection. The similarity between flavivirus Es leads to significant antibody cross-reactivity between different flaviviruses. In addition, flaviviruses often co-circulate in the same regions, further complicating serodiagnosis if antibodies are present from a previous infection or vaccination against a different flavivirus to the current infection. The use of recombinant proteins as antigen permits the introduction of gene mutations that lead to amino acid substitutions in the protein that may alleviate the antibody cross-reactivity observed using wild-type proteins. It also avoids the need to culture and purify large quantities of DV, traditionally sourced from tissue culture or suckling mouse brain, which requires high-containment facilities (Containment Level 3). A further advantage of recombinant dengue VLPs is that they closely resemble the native conformation of dengue virions, so antigen presentation is likely to be superior compared to using recombinant E alone. Enzyme-linked immunosorbent assays (ELISAs) using flavivirus VLP antigens (St. Louis encephalitis virus, SLEV; West Nile virus, WNV; DV types 1 to 4 and Japanese encephalitis virus, JEV) have demonstrated similar or improved sensitivity and specificity compared to those that use suckling mouse brain-derived antigens (Chang *et al*., 2003; Holmes *et al*., 2005; Hunt, Cropp, and Chang, 2001; Purdy, Noga, and Chang, 2004).

5.2 Chapter objectives

The primary objective for this chapter was to use the recombinant dengue VLPs produced in Chapter 4 to investigate the effects of the DV quasispecies E gene mutations on protein function. The specific functions examined would be dengue VLP attachment to host cell surface molecules (DC-SIGN and HSHS), and the low pH-induced DV E oligomeric rearrangements required for virus and host cell membrane fusion.

In Chapter 4, using immunodetection from Western blots, the denatured DV E were unable to be clearly distinguished from the baculovirus proteins present in the dengue VLP preparations, due to antibody cross-reactivity (discussed in section 4.9.6). However, using the same detection antibodies, immunodetection of native dengue VLPs from Western blots was successful in detecting only the DV proteins. This indicated that the antibody cross-reactivity observed was restricted to denatured but not native dengue VLPs. An ELISA would be developed and optimised to enable detection of the recombinant dengue VLPs in their native conformation. The antibody cross-reactivity observed previously would be overcome during ELISA development and optimisation. To enable differences in cell attachment capabilities between the quasispecies variant and consensus dengue VLPs to be identified, this ELISA was then to be performed with an additional capture step using recombinant DC-SIGN, or heparin (as a HSHS analogue).

Dengue VLPs would also be subjected to low pH, and the oligomeric conformation assessed. This was to compare the ability of the quasispecies consensus and variant VLPs to undergo low pH-induced E oligomeric rearrangements.

One of the aims of this work from the beginning was to compare the quasispecies consensus and variant VLPs in terms of cross-reactivity with antibodies to other non-dengue flaviviruses. Initial ELISAs would be performed using the dengue VLPs as antigen, and human serum containing antibodies to DV. This was to assess the affinity of the dengue VLPs for human DV antibodies and lead to subsequent examinations of sensitivity, specificity and cross-reactivity. To our knowledge, such investigations using

quasispecies variants have not been reported. However, since beginning this work, flavivirus VLPs constructed using SLEV, WNV and JEV prM and E genes with mutations in the fusion peptide, have been shown to exhibit dramatically reduced flavivirus antibody cross-reactivity, compared to wild-type VLPs using IgM capture ELISAs (Chiou *et al*., 2008; Roberson, Crill, and Chang, 2007).

5.3 Development and optimisation of an ELISA using recombinant dengue virus-like particles as antigen

Previous attempts to detect the denatured dengue VLPs by immunodetection from Western blots had proved difficult, as the antibodies used for immunodetection were unexpectedly shown to cross-react with baculovirus proteins present in the VLP preparations (discussed in section 4.9.6). By developing and optimising an indirect ELISA, using the dengue VLPs in their native conformation as antigen, it was thought that the cross-reactivity observed previously could be overcome. Concentrated clarified medium from rBV.*lacZ*-infected *Sf*9 cells was used as a background control antigen (this protein sample is referred to as lacZ throughout the chapter), as it was previously shown to be representative of the non-DV proteins present in the dengue VLP preparations (section 4.8). The dengue VLP samples had also been concentrated from clarified insect cell culture medium by ultracentrifugation through a sucrose cushion. The total protein content of the dengue VLP and lacZ samples was quantified by Bradford assay (section 2.8.5), and ELISA plates were consistently coated with samples at a concentration of 1µg/mL (0.1µg/well), unless otherwise stated.

5.3.1 Initial ELISA evaluation

Three primary antibodies were used in the previous chapter. These were: a mouse monoclonal antibody (mAb) to flavivirus E, a rabbit polyclonal antibody (pAb) to DV types 1, 2, 3 and 4, and a mouse mAb to DV E. The mouse mAb to flavivirus E was not tested for use in the ELISA because it had proved to be unreliable previously (discussed in section 4.9.5). The mouse mAb to DV E and rabbit pAb to DV types 1, 2, 3 and 4 were compared in preliminary ELISAs (mAb and pAb ELISA; Figure 5.1 and Figure 5.2 respectively) to determine which provided a better distinction between the dengue VLPs and baculovirus proteins present in the VLP preparations (section 2.9.1). The secondary antibodies used were horse radish peroxidise (HRP)-conjugated rabbit anti-mouse IgG and goat anti-rabbit IgG, respectively. All samples were tested in triplicate and the bar charts display the average absorbance for each sample, with error bars representing one standard deviation. A positive threshold value was calculated based on the average lacZ sample absorbance plus three standard deviations. Absorbance readings above this threshold were considered positive and those below were negative

for dengue VLP detection. For the mAb ELISA (Figure 5.1), this threshold was 0.228, and five of the dengue VLP samples (R203A_2, R261A_1, R261A_4, R107A_1 and R107A_22) displayed average absorbance values greater than this, so were considered positive for dengue VLPs. For the pAb ELISA (Figure 5.2), the threshold was 0.728, and six of the dengue VLP samples (R097A_4, R203A_2, R203A_54, R261A_1, R107A_1 and R107A_22) displayed average absorbance values greater than this, so were considered positive for dengue VLPs.

In addition to the lacZ sample background control, purified baculovirus GP64 and concentrated clarified medium from mock-infected cells were used as negative controls. Both ELISAs showed similar absorbance readings for the lacZ and purified GP64 samples, providing further confirmation that GP64 is the source of the antibody cross-reactivity observed. The mock-infected sample served to establish a background level of absorbance from the culture medium of uninfected insect cells. For the mAb ELISA (Figure 5.1), this signal was below that of any of the other samples, as would be expected. However, for the pAb ELISA (Figure 5.2), this signal was greater than all but one of the samples tested.

In both ELISAs, uncoated wells lacking sample antigen were included as negative controls and no binding was observed (data not shown). For each sample antigen, control wells lacking the primary antibodies were also included to measure any direct binding of the secondary antibodies to the samples (Figure 5.3). For the mAb ELISA, the absorbance signals from lacZ sample and purified GP64-coated wells in which both primary and secondary antibodies were used were similar to dengue VLP, lacZ and purified GP64 sample-coated wells containing secondary antibody only. This indicates that the antibody cross-reactivity observed for this mAb ELISA is baculovirus GP64-specific, and due to the secondary rather than the primary antibody as was first thought. For the pAb ELISA, the absorbance signals from wells containing secondary antibody only were similarly low for dengue VLP, lacZ and purified GP64 sample-coated wells, indicating that the signal observed is background fluorescence and the secondary antibody does not bind significantly to them. However, high absorbance

signals were observed for the pAb ELISA where both primary and secondary antibodies were used in wells coated with the dengue VLPs or any of the background or negative control samples (mock, lacZ and purified GP64). This confirms that the primary antibody for the pAb ELISA is binding non-specifically to baculovirus GP64 and a component of the insect cell culture medium. The similarity between the dengue VLP sample signals and the lacZ sample signals generated using only the mAb ELISA secondary antibody (non-specifically detecting background proteins rather than DV E) indicates that all of these samples contain a similar amount of background protein.

Signal to noise ratios were calculated for each dengue VLP sample from the preliminary mAb and pAb ELISA data, by dividing the average absorbance signal from the dengue VLP sample replicates by the average signal from the lacZ background sample replicates (Table 5.1). Using the mAb ELISA, all but one of the dengue VLP samples (R097A_13) had a signal to noise ratio greater than 1.0. Using the pAb ELISA, all but two of the dengue VLP samples (R097A_43 and R261A_4) had a signal to noise ratio greater than 1.0. In general, the mAb ELISA produced better signal to noise ratios than the pAb ELISA. This difference was statistically significant using a paired t-test ($p=0.026$).

The mAb ELISA was chosen to take forward for further optimisation. This was because the positive threshold value for this ELISA was 3-fold lower than the pAb ELISA (0.228 compared to 0.728), and the signal to noise ratios for each sample were better for the mAb ELISA. The pAb ELISA also displayed a high level of reactivity to the mock-infected sample, which was not observed with the mAb ELISA. Also, the antibody cross-reactivity in the mAb ELISA was shown to result from the secondary detection antibody rather than the primary DV-specific antibody (which was not the case with the pAb ELISA), indicating that further optimisation of the ELISA might improve the signal to noise ratio.

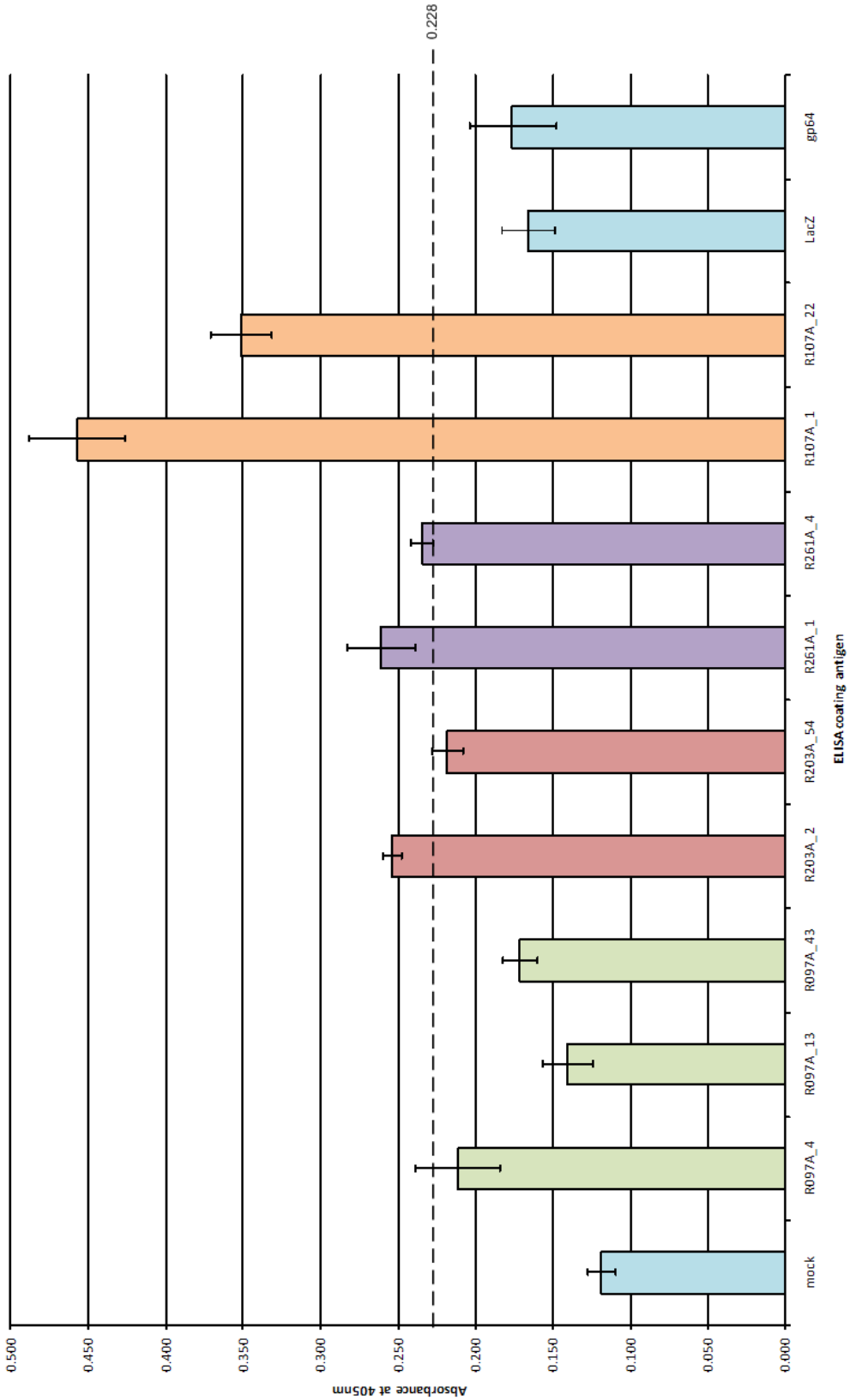


Figure 5.1. Dengue VLP mAb ELISA, preliminary testing.

The primary antibody was a mouse mAb to DV E. The secondary antibody was HRP-conjugated rabbit anti-mouse IgG. The positive threshold value (0.228) was calculated based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three replicates for each sample.

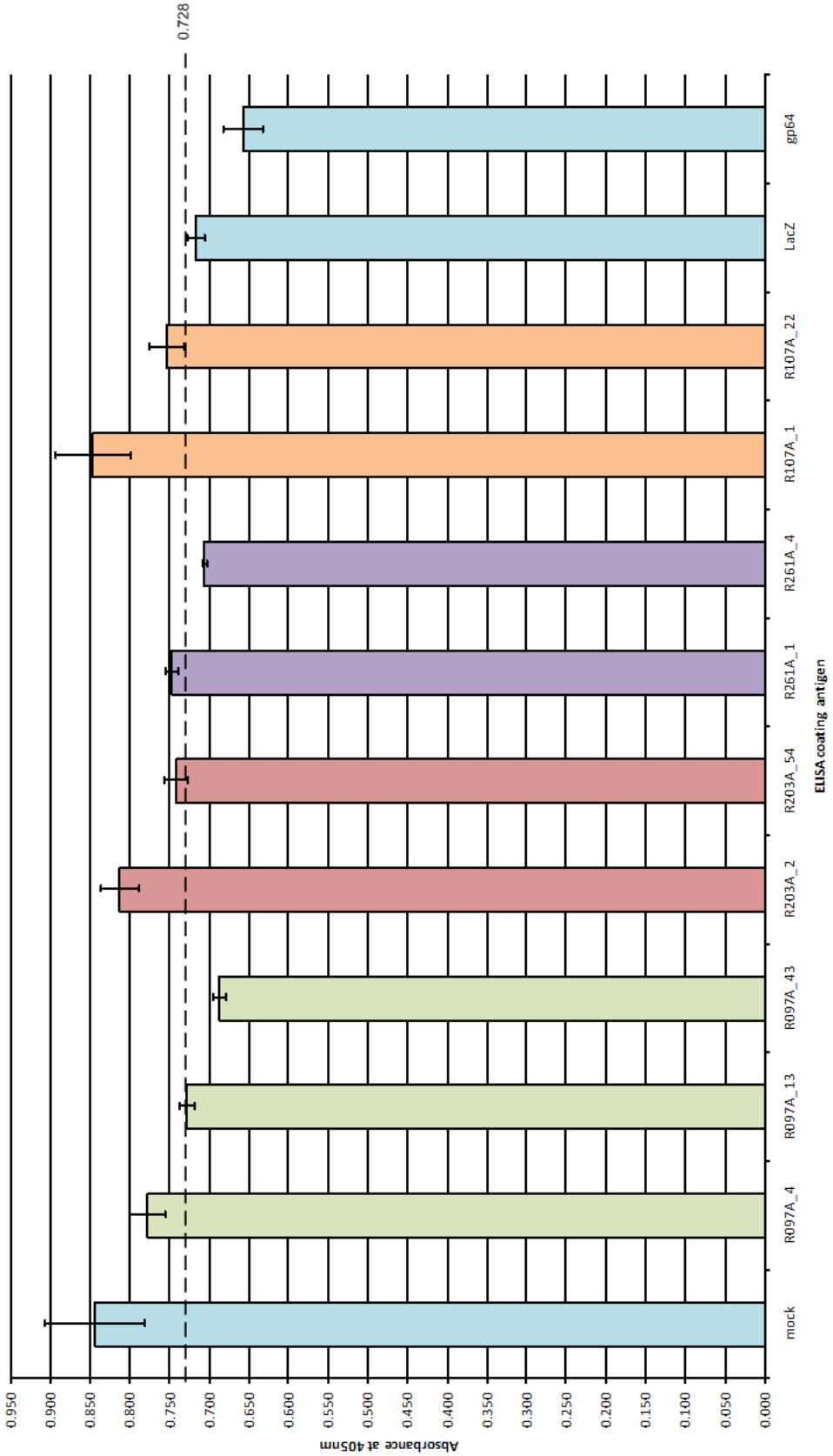


Figure 5.2. Dengue VLP pAb ELISA, preliminary testing.

The primary antibody was a rabbit pAb to DV types 1, 2, 3 and 4. The secondary antibody was HRP-conjugated goat anti-rabbit IgG. The positive threshold value (0.728) was calculated based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three replicates for each sample.

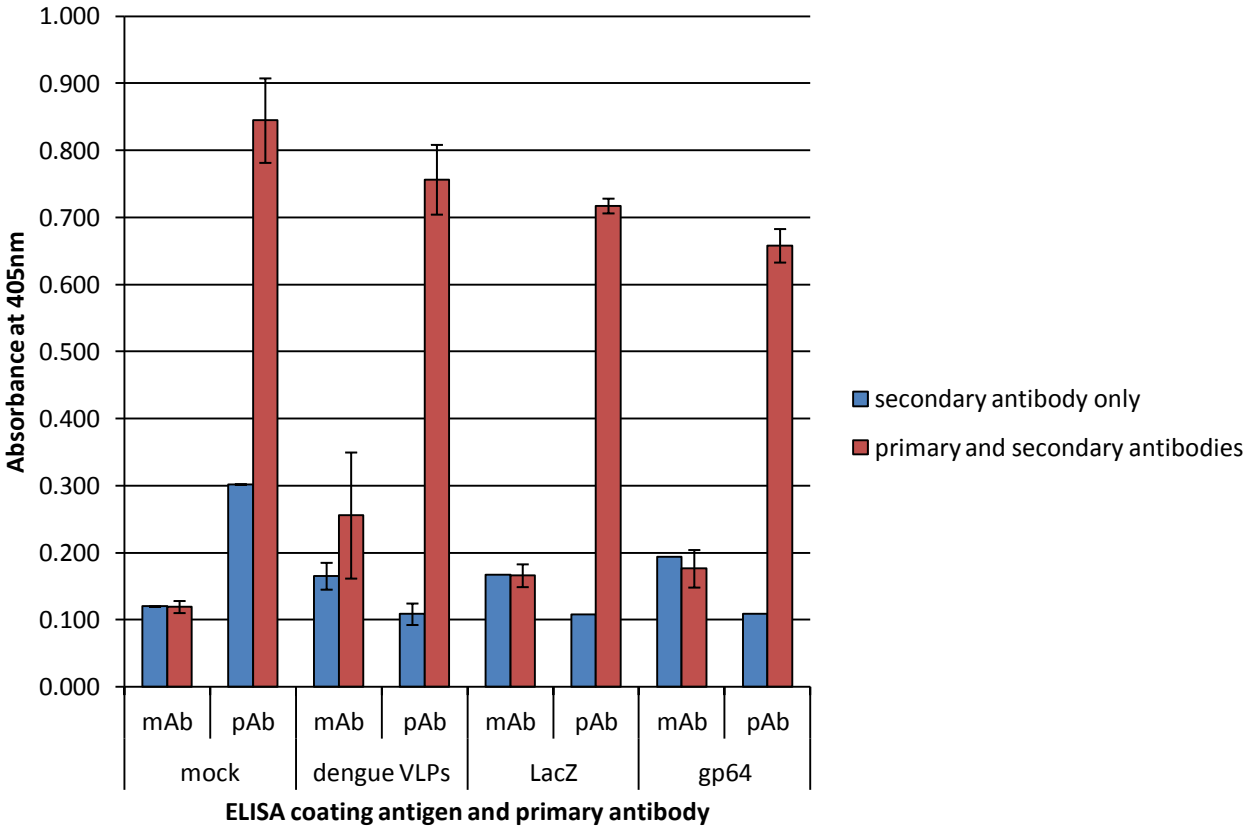


Figure 5.3. Comparison of dengue VLP mAb and pAb ELISAs using both primary and secondary antibodies, and secondary antibodies only.

The primary and secondary antibodies were as detailed in Figure 5.1 and Figure 5.2 for the mAb and pAb ELISAs respectively. Error bars represent the standard deviation of the replicates for each sample. For the mock, lacZ and purified GP64 negative and background control samples three replicates were performed, whereas the dengue VLP error bars correspond to the standard deviation for all of the dengue VLP samples in triplicate. Where error bars are not shown (wells coated with negative and background control sample antigens and probed using the secondary antibody only), no replicate samples were tested.

Table 5.1. Signal to noise ratios for the dengue VLP samples in the mAb and pAb ELISAs.

Dengue VLP sample	mAb ELISA S:N	pAb ELISA S:N
R097A_4	1.28	1.09
R097A_13	0.85	1.02
R097A_43	1.03	0.96
R203A_2	1.53	1.14
R203A_54	1.32	1.04
R261A_1	1.57	1.04
R261A_4	1.42	0.99
R107A_1	2.76	1.18
R107A_22	2.12	1.05

5.3.2 Dengue virus-like particle mAb ELISA optimisation

Optimisation experiments were conducted using the R097A_13 and R107A_1 dengue VLP samples because they displayed the lowest and highest signal to noise ratios respectively in the initial experiment (Table 5.1). The lacZ sample was included as a background control as before. Four different blocking buffers were tested to optimise the blocking step of the mAb ELISA (Figure 5.4). These were 10% fetal calf serum (FCS) in phosphate buffered saline with 0.05% tween 20 (PBST), protein-free (Thermo Fisher), StartingBlock T20 (Thermo Fisher) and 5% skimmed milk in PBST. Positive threshold values were 0.080 using both 10% FCS in PBST and 5% skimmed milk in PBST, 0.099 using StartingBlock T20, and 0.193 using the protein-free blocking buffer. The R097A_13 sample average absorbance signals were not greater than the positive threshold values for any of the blocking buffers tested, but the R107A_1 sample average absorbance signals were greater than the positive threshold values for all of the blocking buffers tested. Signal to noise ratios were calculated for the two dengue VLP samples (Table 5.2), with the best ratio observed for the R107A_1 sample using the StartingBlock blocking buffer (1.97) and for the R097A_13 sample using the protein-free blocking buffer (1.27).

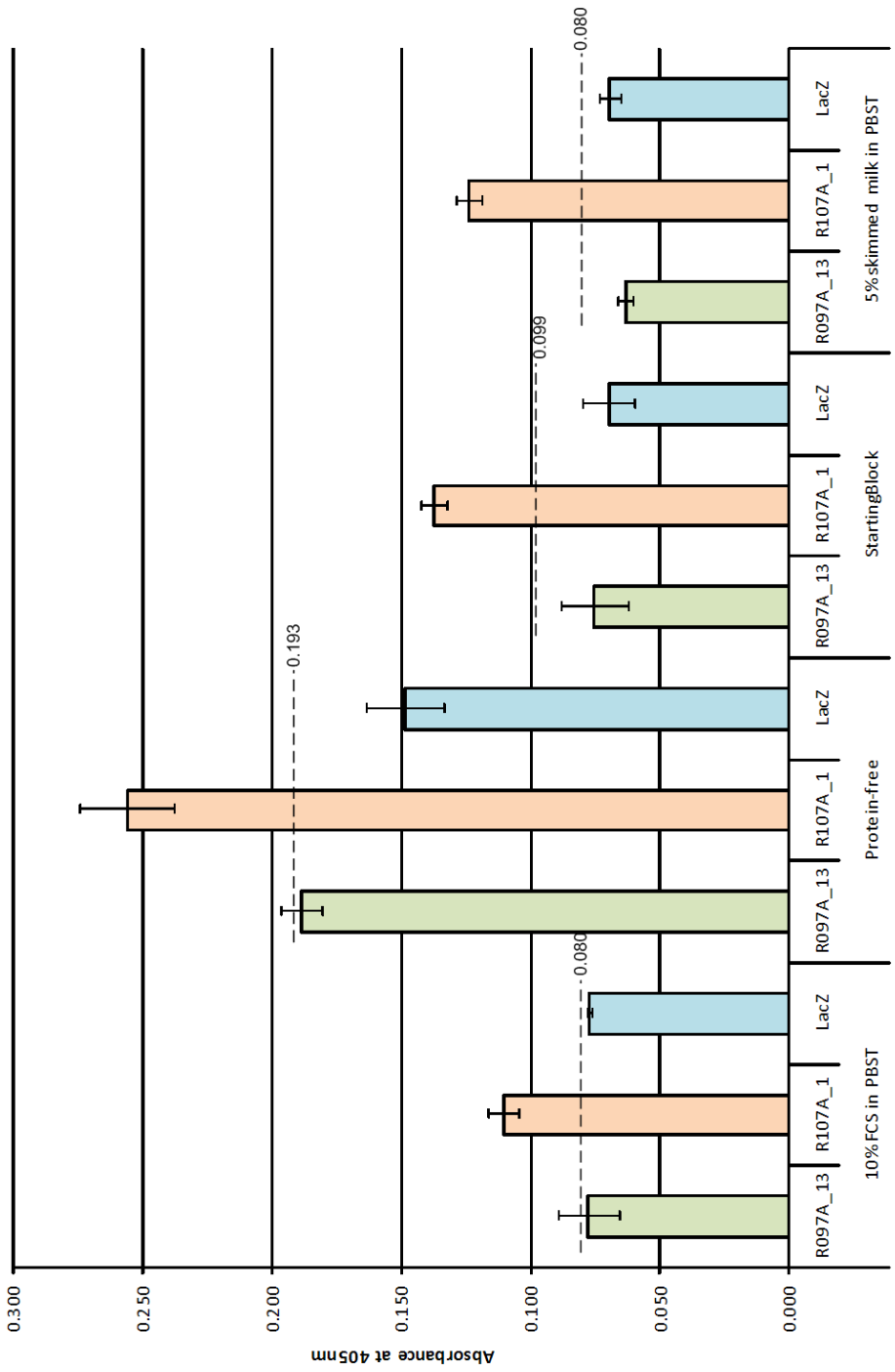


Figure 5.4. Optimisation of dengue VLP mAb ELISA blocking buffer using dengue VLP samples R097A_13 and R107A_1.

The primary and secondary antibodies were as detailed in Figure 5.1. The positive threshold value (broken line) was calculated for each blocking buffer based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three replicates for each sample.

Table 5.2. Signal to noise ratios for the dengue VLP samples using different blocking buffers in the mAb ELISA.

Dengue VLP sample	10% FCS in PBST S:N	Protein-free S:N	StartingBlock S:N	5% skimmed milk in PBST S:N
R097A_13	1.00	1.27	1.08	0.91
R107A_1	1.43	1.73	1.97	1.78

Future mAb ELISAs were conducted using 5% skimmed milk in PBST as the blocking buffer. This was chosen over the other blocking buffers primarily due to the low positive threshold value (0.080), which it shared with the 10% FCS in PBST blocking buffer. The signal to noise ratio for the R107A_1 dengue VLP sample was greater using 5% skimmed milk in PBST than for 10% FCS in PBST.

The concentration of primary and secondary antibodies was optimised in the same experiment by checkerboard titration (Figure 5.5). This was to attempt to improve signal to noise ratios by reducing the background detection of non-DV proteins, and consequently reducing the positive threshold value. The lowest positive threshold value (0.189) was observed at primary and secondary antibody concentrations of 1/500 and 1/2000, respectively. Both the R097A_13 and R107A_1 average sample absorbance signals were greater than the positive threshold values at all but four of the antibody concentration combinations tested (the exceptions being primary and secondary antibody concentration combinations respectively of 1/1000 and 1/2000, 1/1000 and 1/1000, 1/750 and 1/2000, and 1/750 and 1/1000). Signal to noise ratios were calculated for each dengue VLP sample at all of the antibody concentrations used (Table 5.3). The greatest signal to noise ratios were achieved for sample R107A_1 (3.28) using primary and secondary antibody concentrations of 1/1000 and 1/750 respectively, and for sample R097A_13 (2.12) using primary and secondary antibody concentrations of 1/750 and 1/500 respectively.

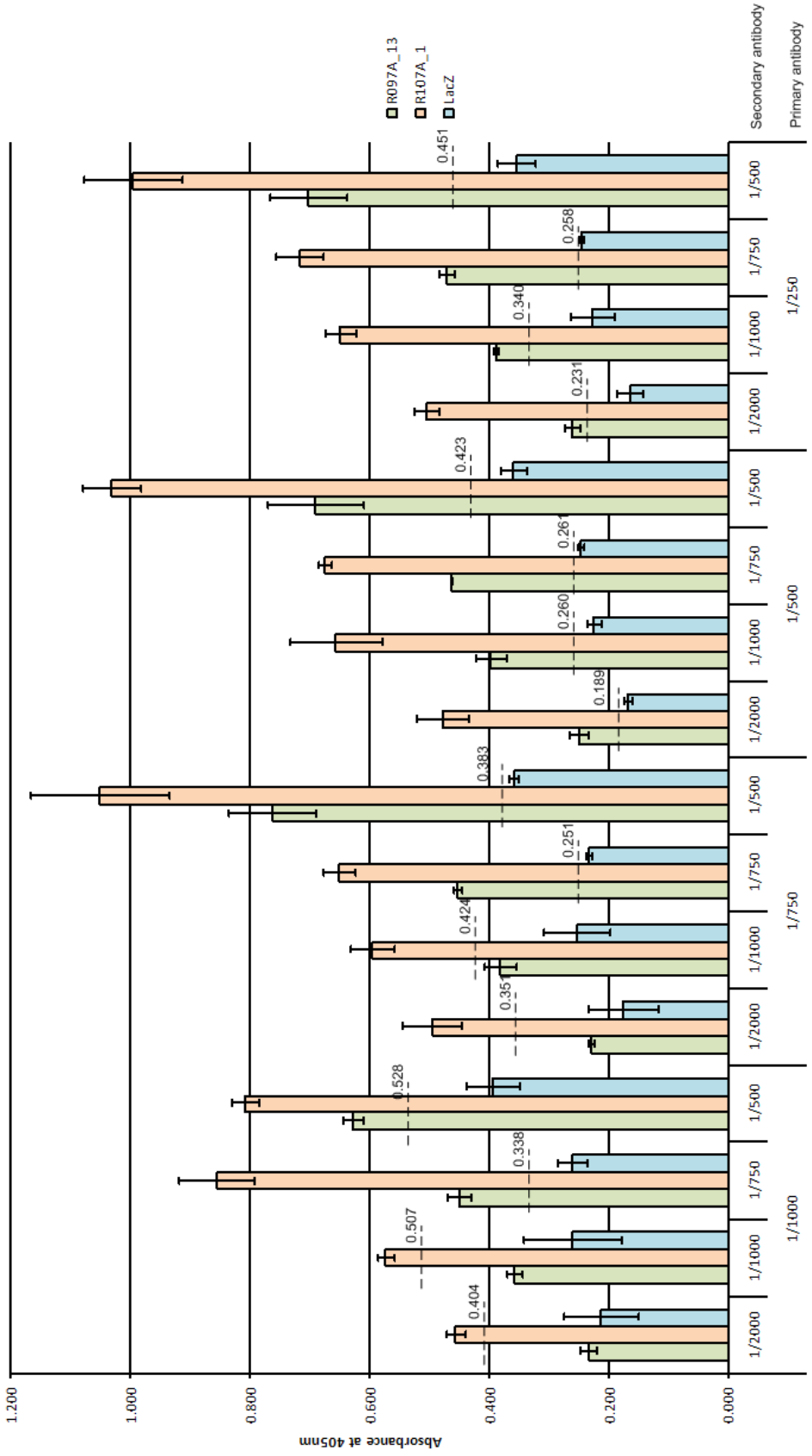


Figure 5.5. Optimisation of dengue VLP mAb ELISA primary and secondary antibody concentrations using dengue VLP samples R097A_13 and R107A_1.

The primary and secondary antibodies were as detailed in Figure 5.1, using the dilutions shown in the figure. Positive threshold values (broken lines) were calculated for each combination of antibody concentrations based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three replicates for each sample.

Table 5.3. Signal to noise ratios for the dengue VLP samples using different concentrations of primary and secondary antibody in the mAb ELISA.

Dengue VLP sample	Primary antibody concentration	Secondary antibody concentration			
		1/2000	1/1000	1/750	1/500
R097A_13	1/1000	1.10	1.37	1.72	1.59
	1/750	1.30	1.50	1.93	2.12
	1/500	1.48	1.76	1.87	1.92
	1/250	1.58	1.70	1.92	1.98
R107A_1	1/1000	2.13	2.19	3.28	2.05
	1/750	2.80	2.34	2.77	2.93
	1/500	2.83	2.91	2.74	2.87
	1/250	3.06	2.84	2.92	2.80

It was decided to use a primary antibody concentration of 1/1000 and a secondary antibody concentration of 1/750 for future dengue VLP mAb ELISAs. Although the positive threshold value at these antibody concentrations (0.338) was not as low as with other antibody concentrations, this combination of antibody concentrations had produced the greatest signal to noise ratio (3.28) for sample R107A_1. This combination of antibody concentrations also enabled conservative use of the limited stocks of primary antibody that were available.

5.3.3 Analysis of the dengue virus-like particle samples using the optimised mAb ELISA

Following optimisation of the blocking buffer and antibody concentrations, the dengue VLP mAb ELISA was performed using all of the dengue VLP samples (Figure 5.6). All nine of the dengue VLP samples were detected and distinguished as positive compared to the lacZ background control sample. Signal to noise ratios were calculated (Table 5.4), and ranged from 1.54 to 3.60.

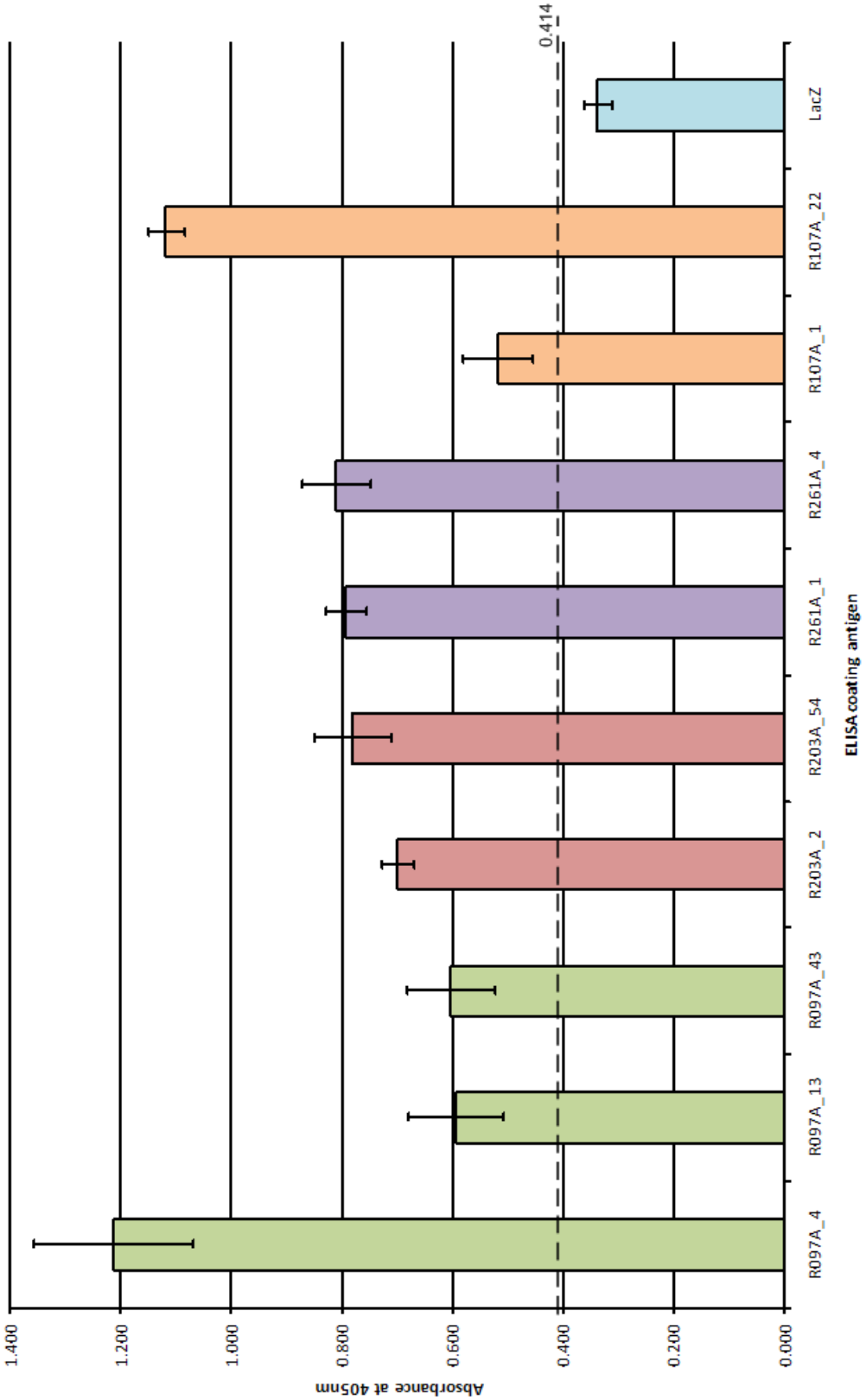


Figure 5.6. Dengue VLP mAb ELISA using optimised antibody concentrations.

The primary and secondary antibodies were as detailed in Figure 5.1, at 1/1000 and 1/750 dilutions respectively. The positive threshold value (0.414) was calculated based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three replicates for each sample.

Table 5.4. Signal to noise ratios for the dengue VLP samples using the optimised mAb ELISA.

Dengue VLP sample	Signal to noise ratio
R097A_4	3.60
R097A_13	1.76
R097A_43	1.79
R203A_2	2.08
R203A_54	2.31
R261A_1	2.35
R261A_4	2.40
R107A_1	1.54
R107A_22	3.31

Comparing these results (Figure 5.6) to those obtained during the optimisation of antibody concentrations (Figure 5.4), the signal generated by the R107A_1 sample has decreased (from 0.569 to 0.181 after subtraction of the background signal), along with the signal to noise ratio (from 3.28 to 1.54). In contrast, the signal from the R097A_13 sample (0.190 and 0.257 after subtraction of the background signal) and signal to noise ratios (1.72 and 1.76) have increased slightly. This is likely due to the freeze-thaw status of the samples as a fresh aliquot of sample R097A_13 was used for each experiment, whereas the R107A_1 sample had been freeze-thawed between experiments. For future ELISAs, fresh sample aliquots were used for each experiment wherever possible, but most of the optimisation experiments were performed using freeze-thawed sample aliquots.

5.4 Comparison of dengue virus-like particle samples binding to DC-SIGN via a capture ELISA

To compare differences in DC-SIGN-binding capabilities between the variant and consensus dengue VLPs, an additional capture step was added to the dengue VLP mAb ELISA optimised in section 5.3 by coating the ELISA plate with DC-SIGN instead of the dengue VLP samples (Figure 5.7; section 2.9.2). Plates were then blocked using 5% skimmed milk with PBST before addition of antigen (dengue VLP samples or lacZ background control sample). Primary and secondary antibodies (mouse mAb to DV E and HRP-conjugated rabbit pAb to mouse IgG respectively) were used at 1/1000 and 1/750 dilutions respectively, as these were previously determined as optimal.

Because the dengue VLP sample antigens were unpurified, it was not clear what proportion of the total protein concentration measured by Bradford assay consisted of dengue VLPs and what proportion consisted of non-DV proteins. However, the amount of non-DV proteins present that were able to bind non-specifically to the secondary antibody and cause background absorbance signals had been previously shown to be consistent across the dengue VLP and lacZ background control samples (Figure 5.3). Initial ELISAs were performed on all of the dengue VLP samples using both the dengue VLP mAb ELISA and the DC-SIGN capture ELISA. For the DC-SIGN capture ELISA, wells were initially coated with DC-SIGN at 1µg/mL (0.1µg/well). The dengue VLP mAb ELISA was performed to provide a standard measurement of VLP detection by the antibodies. This enabled comparisons to be made between the relative signals generated by the consensus and variant proteins for each patient sample. The DC-SIGN capture ELISA results were compared to the dengue VLP mAb ELISA results to see if the relative signal intensities changed due to differences in the dengue VLPs binding to DC-SIGN (Figure 5.8). Positive threshold values and signal to noise ratios (Table 5.5) were calculated for each ELISA as before. All of the dengue VLP sample average absorbance signals from the dengue VLP mAb ELISA were greater than the calculated positive threshold value (0.180), and signal to noise ratios ranged from 1.87 to 4.19. However, for the DC-SIGN capture ELISA, none of the dengue VLP sample signals were greater than the calculated positive threshold value (0.212) and poor

signal to noise ratios (ranging from 0.92 to 1.18) were observed. This indicated that the dengue VLP samples were not binding to the DC-SIGN-coated ELISA plates.

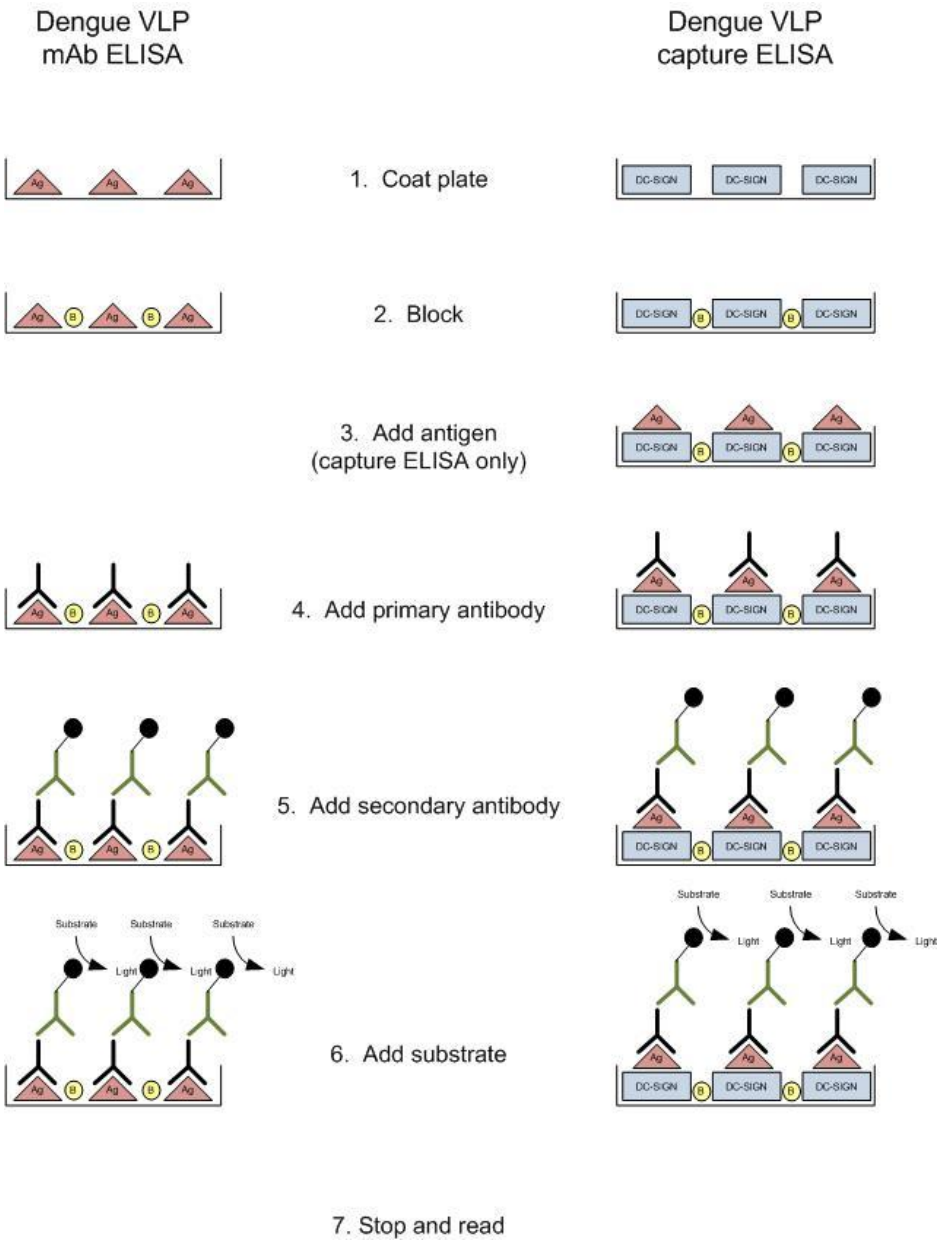


Figure 5.7. Schematic diagram showing the dengue VLP mAb ELISA and dengue VLP capture ELISA procedures.

An additional capture step was added to the dengue VLP mAb ELISA to compare differences in DC-SIGN-binding capabilities between the variant and consensus dengue VLPs (section 2.9.2). Wells were washed with PBST between each of steps 1 to 6. For the dengue VLP mAb ELISA, plates were coated overnight with dengue VLP sample antigens (Ag; step 1), before blocking with 5% skimmed milk in PBST (B; step 2), addition of primary and secondary antibodies (steps 4 and 5), substrate addition, stopping the reaction and reading the absorbance at 405nm (steps 6 and 7). For the capture ELISA, plates were coated overnight with DC-SIGN (step 1), before blocking with 5% skimmed milk in PBST (B; step 2), and addition of the dengue VLP sample antigens (Ag; step 3). Steps 4 to 7 were as described for the dengue VLP mAb ELISA.

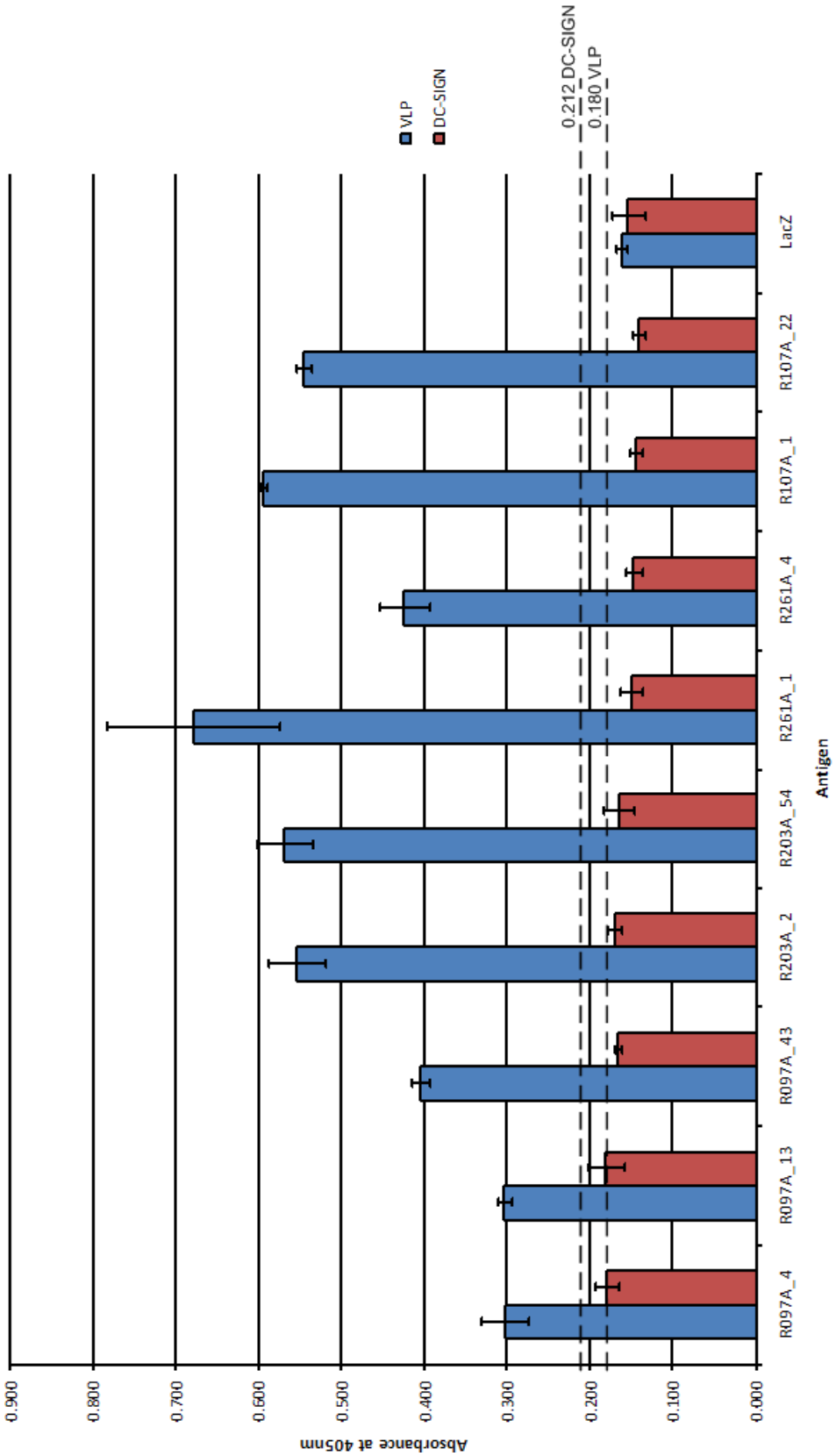


Figure 5.8. Dengue VLP mAb ELISA comparison with DC-SIGN capture ELISA.

ELISA plate wells were coated using DC-SIGN or dengue VLP sample antigens at 1µg/mL (0.1µg/well). For the capture ELISA, dengue VLP sample antigens were added at 1µg/mL (0.1µg/well). The primary and secondary antibodies were as detailed in Figure 5.1, at 1/1000 and 1/750 dilutions respectively. The positive threshold values were calculated for the dengue VLP ELISA and the DC-SIGN capture ELISA (0.180 and 0.212 respectively) based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three replicates for each sample.

Table 5.5. Signal to noise ratios for the dengue VLP samples using the dengue VLP mAb ELISA and the DC-SIGN capture ELISA.

Dengue VLP sample	Dengue VLP mAb ELISA S:N	DC-SIGN capture ELISA S:N
R097A_4	1.87	1.17
R097A_13	1.87	1.18
R097A_43	2.50	1.08
R203A_2	3.42	1.11
R203A_54	3.51	1.08
R261A_1	4.19	0.98
R261A_4	2.62	0.96
R107A_1	3.67	0.94
R107A_22	3.37	0.92

To attempt to improve dengue VLP binding to DC-SIGN, the DC-SIGN capture ELISA was repeated using the R107A_1 dengue VLP and lacZ background control samples at DC-SIGN coating concentrations of 1, 10, 50 and 100µg/mL (Figure 5.9). Positive threshold values and signal to noise ratios (Table 5.6) were calculated for each DC-SIGN coating concentration as before. As with the previous DC-SIGN capture ELISA, the dengue VLP sample average absorbance signals were not greater than the calculated positive threshold value at any of the DC-SIGN coating concentrations tested. Average signals from both the R107A_1 dengue VLP and lacZ background control samples increased with increasing DC-SIGN coating concentration, and signal to noise ratios decreased. This showed that the dengue VLP samples were not binding to DC-SIGN at any of the DC-SIGN concentrations tested.

For the DC-SIGN capture ELISA, uncoated wells lacking DC-SIGN were included as negative controls and no binding was observed (data not shown). Control wells coated with DC-SIGN but lacking sample antigen were used to measure any direct binding of the antibodies to DC-SIGN (Figure 5.10). The absorbance signals from these wells showed very little difference to those containing either dengue VLP or lacZ background control sample antigens. In addition, similar absorbance signals were generated from wells in which both primary and secondary antibodies were used, compared to wells in which only secondary antibody was used. This indicated that the signal was due to the secondary antibody binding directly to DC-SIGN, and as a result it was decided to move on to testing the dengue VLP samples using a heparin capture ELISA.

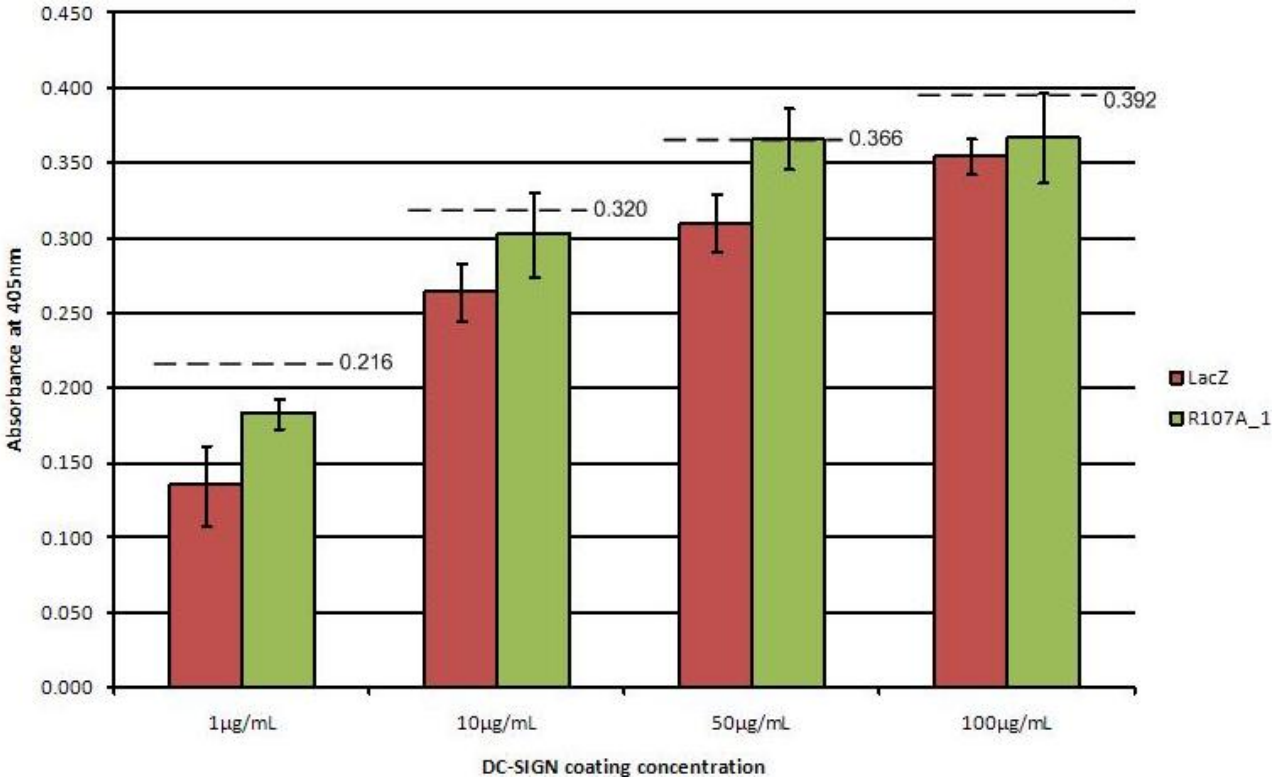


Figure 5.9. Optimisation of DC-SIGN coating concentration for the DC-SIGN capture ELISA.

ELISA plate wells were coated using DC-SIGN at 1, 10, 50 and 100µg/mL (0.1, 1, 5 and 10µg/well), and lacZ background control or dengue VLP (R107A_1) sample antigens added at 1µg/mL (0.1µg/well). The primary and secondary antibodies were as detailed in Figure 5.1, at 1/1000 and 1/750 dilutions respectively. Positive threshold values (broken lines) were calculated for each DC-SIGN-coating concentration based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three sample replicates.

Table 5.6. Signal to noise ratios for the dengue VLP sample R107A_1 using different DC-SIGN coating concentrations for the DC-SIGN capture ELISA.

DC-SIGN coating concentration	Signal to noise ratio
1µg/mL	1.36
10µg/mL	1.15
50µg/mL	1.18
100µg/mL	1.04

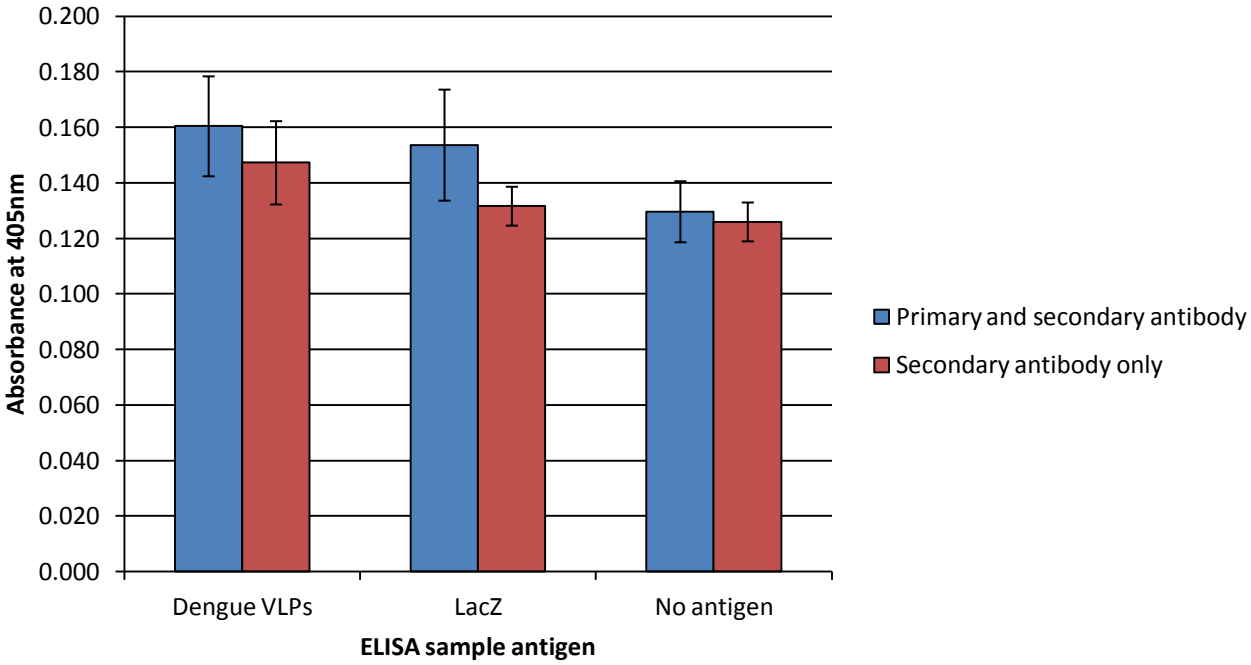


Figure 5.10. Comparison of absorbance signals generated by the DC-SIGN capture ELISA using both primary and secondary antibodies, and secondary antibodies only.

ELISA plate wells were coated using DC-SIGN, and sample antigens added at 1µg/mL (0.1µg/well). The primary and secondary antibodies were as detailed in Figure 5.1, at 1/1000 and 1/750 dilutions respectively. Error bars represent the standard deviation of the three replicates for the lacZ and No sample antigen wells, whereas the dengue VLP error bars correspond to the standard deviation for all of the dengue VLP samples in triplicate.

5.5 Comparison of dengue virus-like particle samples binding to heparin via a capture ELISA

To compare differences in HSHS-binding capabilities between the variant and consensus dengue VLPs, a heparin capture ELISA was performed in the same way as the DC-SIGN capture ELISA (Figure 5.7), except that ELISA plates were coated with heparin rather than DC-SIGN. Initial ELISAs were performed on all of the dengue VLP samples using both the dengue VLP mAb ELISA and the heparin capture ELISA. For the heparin capture ELISA, wells were initially coated with heparin at 1mg/mL (0.1mg per well).

As before, the dengue VLP mAb ELISA was performed as a reference, and the heparin capture ELISA results were compared to the dengue VLP mAb ELISA results (Figure 5.11). This was to see if the relative signal intensities changed due to differences in the dengue VLPs binding to heparin. Positive threshold values and signal to noise ratios (Table 5.7) were calculated for each ELISA as before. All of the dengue VLP sample average absorbance signals from the dengue VLP mAb ELISA were greater than the calculated positive threshold value (0.180), and signal to noise ratios ranged from 1.87 to 4.19. For the heparin capture ELISA, only one of the dengue VLP sample signals (R107A_22) was greater than the calculated positive threshold value (0.109) and signal to noise ratios ranged from 1.15 to 1.41.

To attempt to improve dengue VLP binding to heparin, the heparin capture ELISA was repeated using the R107A_1 dengue VLP and lacZ background control samples at heparin coating concentrations of 1, 10, 50 and 100mg/mL (Figure 5.12). Positive threshold values and signal to noise ratios (Table 5.7) were calculated for each heparin coating concentration as before. The dengue VLP sample average absorbance signals at the 1mg/mL and 10mg/mL heparin coating concentrations were not greater than the calculated positive threshold value for these coating concentrations (0.121 and 0.150, respectively). However, the dengue VLP sample signals at the 50mg/mL and 100mg/mL heparin coating concentrations were greater than the calculated positive threshold value for these coating concentrations (0.144 and 0.151, respectively).

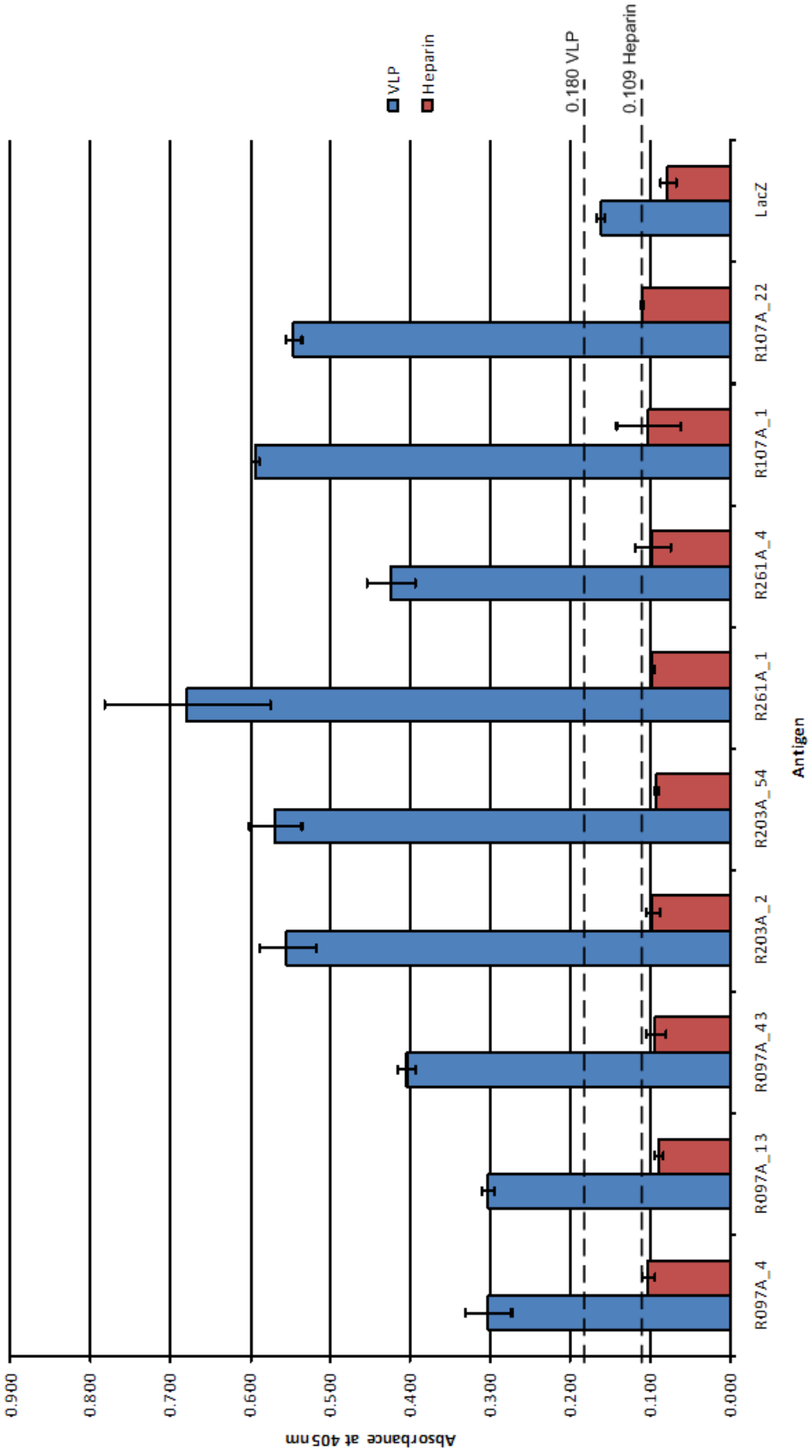


Figure 5.11. Initial dengue VLP mAb ELISA comparison with heparin capture ELISA.

ELISA plate wells were coated using heparin at 1mg/mL (0.1mg/well), or dengue VLP sample antigens at 1µg/mL (0.1µg/well). For the capture ELISA, sample antigens were added at 1µg/mL (0.1µg/well). The primary and secondary antibodies were as detailed in Figure 5.1, at 1/1000 and 1/750 dilutions respectively. The positive threshold values were calculated for the dengue VLP ELISA and the heparin capture ELISA (0.180 and 0.109 respectively) based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three replicates for each sample.

Table 5.7. Signal to noise ratios for the dengue VLP samples using the dengue VLP mAb ELISA and the heparin capture ELISA.

Dengue VLP sample	Dengue VLP mAb ELISA S:N	Heparin capture ELISA S:N
R097A_4	1.87	1.32
R097A_13	1.87	1.15
R097A_43	2.50	1.19
R203A_2	3.42	1.24
R203A_54	3.51	1.18
R261A_1	4.19	1.25
R261A_4	2.62	1.24
R107A_1	3.67	1.32
R107A_22	3.37	1.41

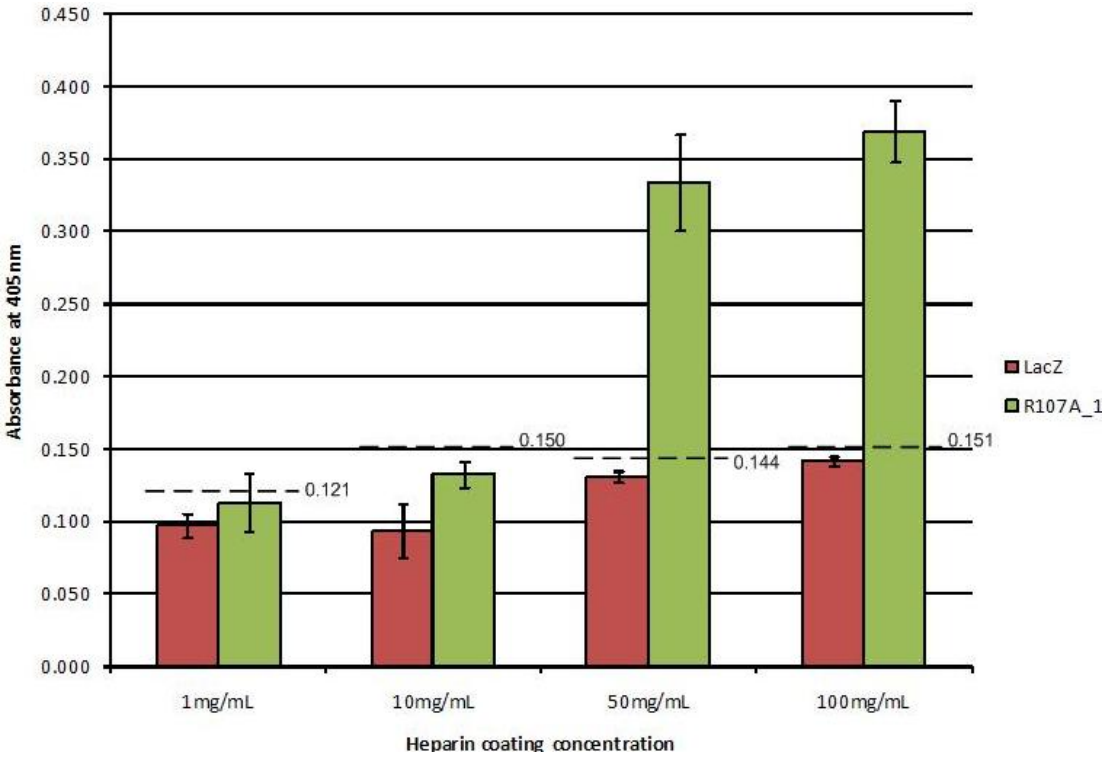


Figure 5.12. Optimisation of heparin coating concentration for the heparin capture ELISA.

ELISA plate wells were coated using heparin at 1, 10, 50 and 100mg/mL (0.1, 1, 5 and 10mg/well), and lacZ or dengue VLP (R107A_1) sample antigens added at 1µg/mL (0.1µg/well). The primary and secondary antibodies were as detailed in Figure 5.1, at 1/1000 and 1/750 dilutions respectively. Positive threshold values (broken lines) were calculated for each heparin-coating concentration based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three sample replicates.

In general, average absorbance signals from both the R107A_1 dengue VLP and lacZ background control samples increased with increasing heparin coating concentration; the exception being the lacZ sample signal from the 10mg/mL heparin coating concentration, which also exhibited greater standard deviation of sample replicates at

this heparin coating concentration compared to the other concentrations. Signal to noise ratios (Table 5.8) also increased with increasing heparin coating concentration, ranging from 1.16 to 2.60. This demonstrated that at the higher heparin coating concentrations (50 and 100 mg/mL), the dengue VLP sample was binding to the heparin with a greater binding affinity than the lacZ background control sample.

Table 5.8. Signal to noise ratios for the dengue VLP sample R107A_1 using different heparin coating concentrations for the heparin capture ELISA.

Heparin coating concentration	Signal to noise ratio
1mg/mL	1.16
10mg/mL	1.41
50mg/mL	2.54
100mg/mL	2.60

The heparin capture ELISA was then performed in parallel with the dengue VLP mAb ELISA using new aliquots (not previously freeze-thawed) of all of the dengue VLP samples and the lacZ background control sample (Figure 5.13). For the heparin capture ELISA, a heparin coating concentration of 100mg/mL was used because at this concentration the lacZ sample background was acceptably low (Figure 5.12; positive threshold value 0.151) and the signal to noise ratio using the R107A_1 dengue VLP sample was greatest (Table 5.8; signal to noise ratio 2.60). Positive threshold values were calculated as before for both ELISAs. All of the dengue VLP sample signals were greater than the calculated positive threshold (0.345) for the dengue VLP mAb ELISA, and all but one of the dengue VLP sample signals (the exception being sample R097A_13) were also greater than the calculated positive threshold value (0.237) for the heparin capture ELISA.

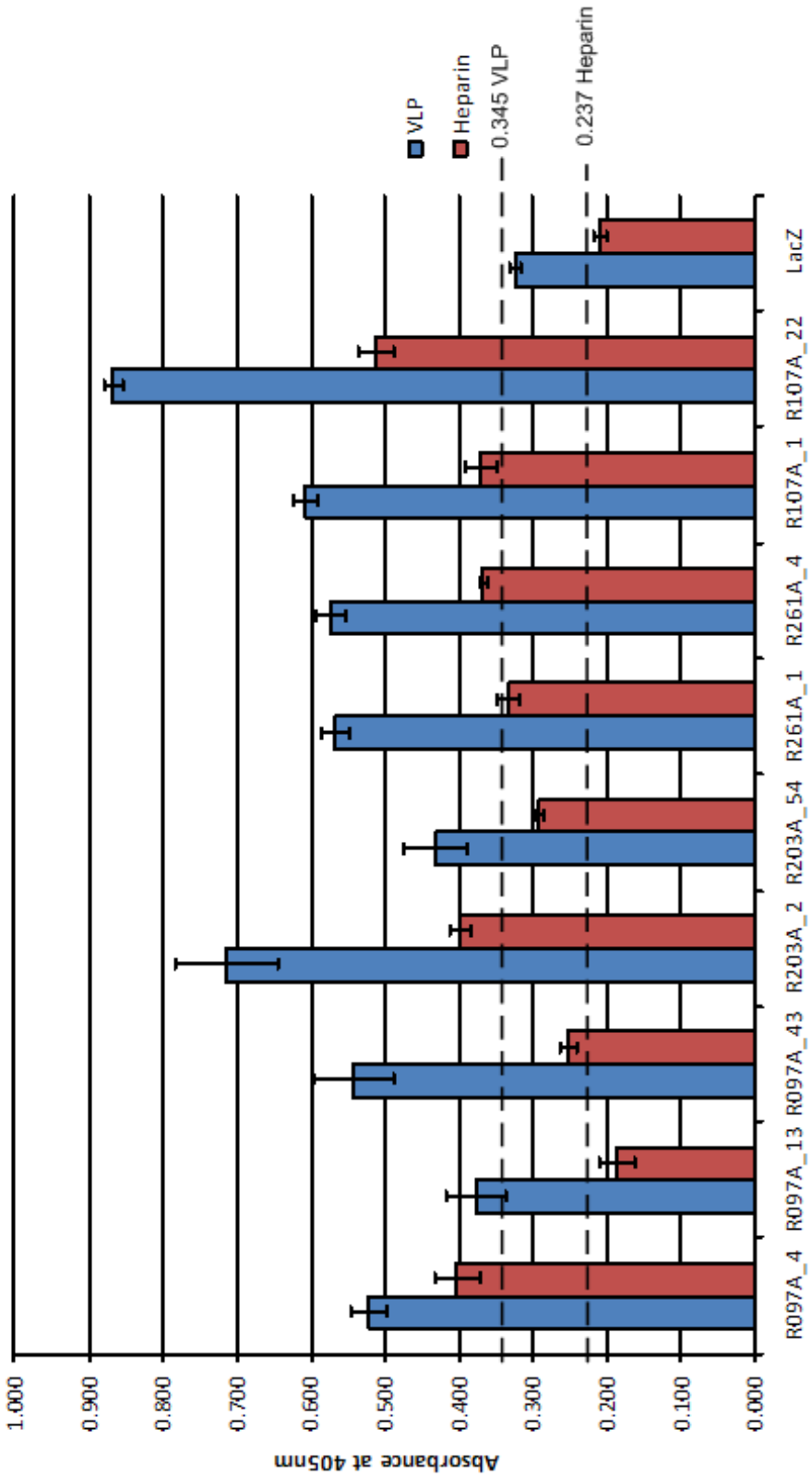


Figure 5.13. Dengue VLP mAb ELISA comparison with heparin capture ELISA.

ELISA plate wells were coated using heparin at 100mg/mL (10mg/well), or dengue VLP sample antigens at 1µg/mL (0.1µg/well). For the capture ELISA, sample antigens were added at 1µg/mL (0.1µg/well). The primary and secondary antibodies were as detailed in Figure 5.1, at 1/1000 and 1/750 dilutions respectively. The positive threshold values were calculated for the dengue VLP mAb ELISA and the heparin capture ELISA (0.345 and 0.237 respectively) based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three replicates for each sample.

The amount of non-DV proteins present that were able to bind non-specifically to the secondary antibody and cause background absorbance signals had been previously shown to be consistent across the dengue VLP and lacZ background control samples (Figure 5.3). All of the dengue VLP mAb ELISA results (obtained using both primary and secondary antibodies) show differences in the absorbance values between the dengue VLP samples, indicating that the amount of dengue VLPs present varies between samples. To simplify analysis of the results from the dengue VLP mAb and heparin capture ELISAs (Figure 5.13), the data was first normalised to account for the background signal generated by the secondary antibody binding non-specifically to non-DV proteins (by subtracting the lacZ negative control signal from the dengue VLP signal; Table 5.9). The normalised average dengue VLP absorbance values from the dengue VLP mAb ELISA were used as a measure of the amount of dengue VLPs available to bind to the heparin in the capture ELISA. The normalised average dengue VLP absorbance values from the heparin capture ELISA were used as a measure of the amount of dengue VLPs actually bound to the heparin in the capture ELISA. The proportion of available dengue VLPs bound to heparin was then calculated (Table 5.9). For each dengue VLP sample set, the consensus protein representative sample (R097A_4, R203A_2, R261A_1 and R107A_1) was used as a basis for comparison with the variant protein samples.

For the R097A sample set, the consensus protein sample R097A_4 showed 97% of the available dengue VLPs bound to heparin, whereas for the variant protein samples R097A_13 and R097A_43, 0% and 20% respectively of the available dengue VLPs bound to the heparin. For the R203A sample set, the consensus protein sample R203A_2 demonstrated 49% of the available dengue VLPs bound to heparin, whereas for the variant protein sample R203A_54, 78% of the available dengue VLPs bound to the heparin. For the R261A sample set, the consensus protein sample R261A_1 showed 51% of the available dengue VLPs bound to heparin whereas for the variant protein sample R261A_4, 64% of the available dengue VLPs bound to the heparin. For the R107A sample set, the consensus protein sample R107A_1 demonstrated 57% of the available dengue VLPs bound to heparin whereas for the variant protein sample

R107A_22, 56% of the available dengue VLPs bound to the heparin. When this experiment was repeated similar results were obtained (data not shown).

Table 5.9. Proportion of available dengue VLPs bound to heparin after normalisation of dengue VLP mAb and heparin capture ELISA data

Dengue VLP sample	Available dengue VLPs	Heparin-bound dengue VLPs	Proportion of available dengue VLPs bound to heparin
R097A_4[c]	0.199	0.194	97%
R097A_13	0.053	-0.023	0%
R097A_43	0.219	0.043	20%
R203A_2[c]	0.391	0.191	49%
R203A_54	0.108	0.084	78%
R261A_1[c]	0.245	0.125	51%
R261A_4	0.250	0.159	64%
R107A_1[c]	0.284	0.162	57%
R107A_22	0.544	0.304	56%

5.6 Comparison of dengue virus-like particle samples oligomeric rearrangements in response to low pH

It was intended as part of this chapter to determine whether the low pH-induced DV E oligomeric rearrangements (required for virus and host cell membrane fusion) were prevented in any of the dengue VLP quasispecies variant samples with residue substitutions in areas of E known to be important during conformational change. Initial experiments were intended to be performed as described previously (Allison *et al.*, 1999) by subjecting the dengue VLPs to low pH (pH 6.0), before neutralising and cross-linking oligomers together using dimethyl pimelimidate. The lipid bilayer would then be solubilised using 1% triton X-100 for 1 hour before SDS-PAGE, Western blot, and immunodetection of proteins. It was expected from previously published experiments using TBEV VLPs, that low pH-treated VLPs would exhibit DV E in trimeric form, whereas without low pH-treatment, E would be predominantly in dimeric form, and no trimers would be evident (Figure 5.14; Allison *et al.*, 1999).

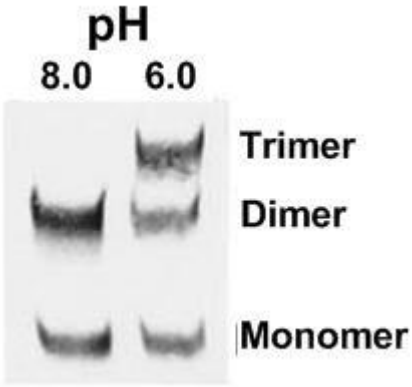


Figure 5.14. Expected dengue virus envelope glycoprotein oligomeric state after low pH treatment and chemical cross-linking.

Figure modified from Allison *et al.*, 1999. Flavivirus (TBE) VLPs pretreated at pH 6.0 before cross-linking and detergent solubilisation show E in three distinct oligomeric states: trimer, dimer and monomer. Without low pH pre-treatment (pH 8.0), only dimer and monomer form are visible. At pH 6.0, the dimer and monomer forms, and at pH 8.0, the monomer form are visible because cross-linking is incomplete.

However, baculovirus was known to be present in the dengue VLP samples, as purification attempts had failed to separate the baculovirus from the dengue VLPs (Chapter 4). Baculovirus GP64 is a class III fusion protein, which forms a trimeric structure in both the pre- and post-fusion states (Kadlec *et al.*, 2008). Baculovirus

GP64 and DV E are similar in size (64 kDa and 54 kDa respectively), and the antibodies used in Chapter 4 for immunodetection of DV E from Western blots were shown to cross-react with baculovirus GP64 (discussed in section 4.9.6). It was therefore decided that it was unfeasible at this time to attempt analysis of low pH-induced DV E oligomeric rearrangements using this method, as clear distinction from baculovirus GP64 was not possible. Once purification of the dengue VLPs is successful it is intended that this experiment will be re-visited.

5.7 Dengue virus-like particle ELISA using human serum

To assess the suitability of the dengue VLPs as a source of diagnostic antigen, an ELISA was performed using human serum containing DV antibodies (Figure 5.15; section 2.9.1). DV antibody negative human serum was used as a negative control. ELISA plates were coated with dengue VLP or lacZ samples at 1µg/mL (0.1µg/well), as with the dengue VLP mAb ELISAs performed previously (sections 5.3 to 5.5). DV antibody positive and negative human sera were obtained from the National Institute for Biological Standards and Control (NIBSC) for use as the primary antibody. The secondary antibody used was HRP-conjugated goat polyclonal anti-human IgG. A positive threshold value (1.019) was calculated as with previous ELISAs based on the average lacZ sample absorbance plus three standard deviations. This positive threshold value was far greater than any observed previously with the dengue VLP mAb ELISA as the average signal from the lacZ sample was so high, indicating non-specific binding to a non-DV protein was occurring. Only four of the dengue VLP samples (R097A_4, R097A_13, R097A_43 and R203A_2) enabled detection of the DV antibodies with average absorbance signals that were greater than this positive threshold value. Whilst the DV antibody negative sera produced average signals for all of the dengue VLPs (range 0.521 to 0.892) that were less than the positive threshold value, the negative serum had not been expected to bind to the dengue VLP samples at all as it did not contain DV antibodies. Signal to noise ratios (Table 5.10) were calculated as before, and were shown to be similar in range for both DV antibody positive (range 0.89 to 1.53) and negative (range 0.88 to 1.50) human serum.

For each sample antigen, control wells lacking the DV antibody positive or negative human serum were also included to measure any direct binding of the secondary antibody to the sample antigens. The absorbance values from these wells were the same as for non-coated wells (data not shown), confirming that the secondary antibody was not binding directly to the dengue VLP or lacZ sample antigens. This indicated that the absorbance signals observed from the lacZ sample-coated wells were due to non-specific binding of the human serum to non-DV proteins in the samples.

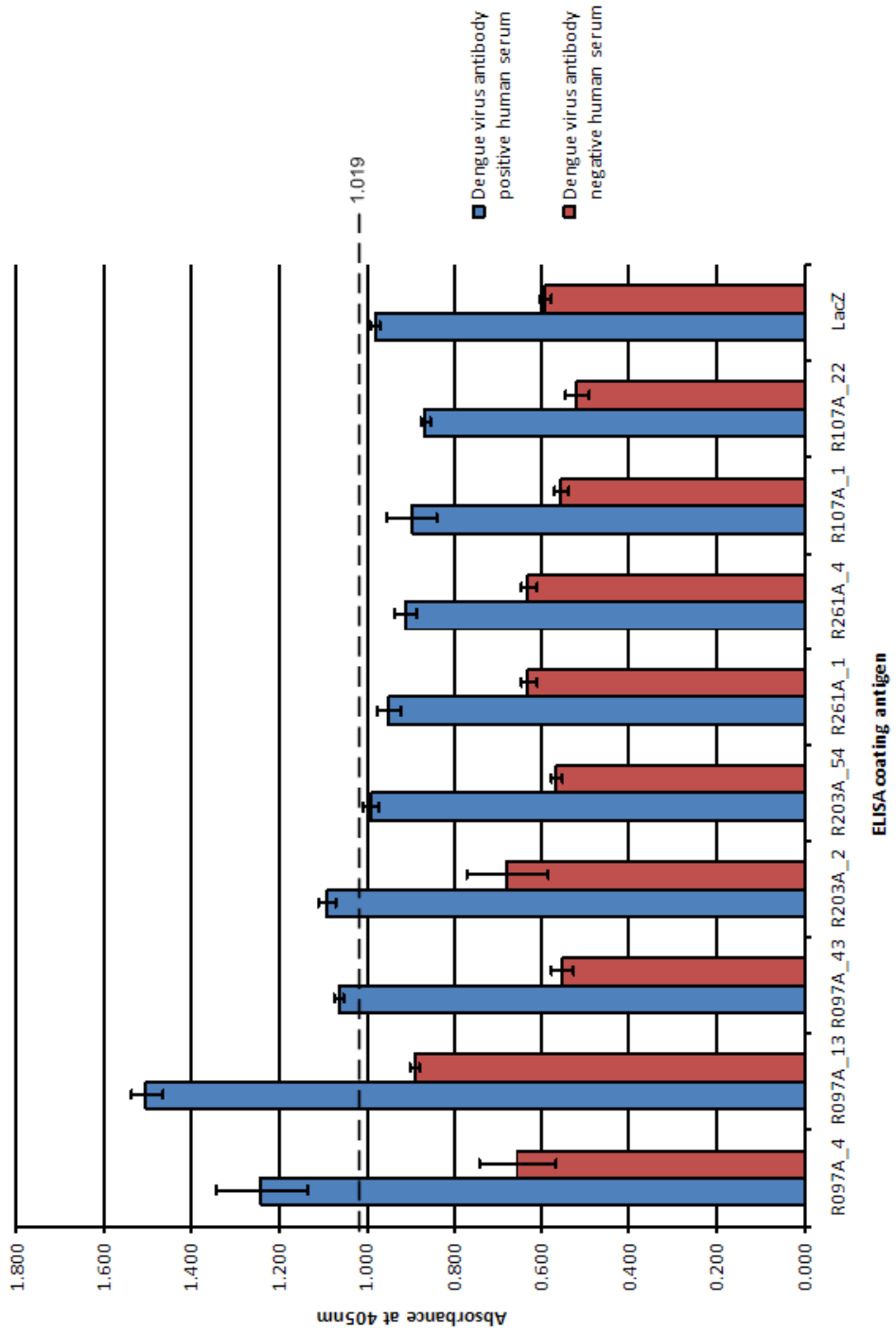


Figure 5.15. Dengue VLP ELISA using human serum.

The primary antibodies were DV antibody positive and negative human serum, and the secondary antibody was HRP-conjugated goat pAb to human IgG. The positive threshold value (1.019) was calculated based on the average lacZ sample absorbance plus three standard deviations. Error bars represent the standard deviation of the three replicates for each sample.

Table 5.10. Signal to noise ratios for the dengue VLP samples using human serum as primary antibody in the dengue VLP ELISA.

Dengue VLP sample	DV antibody positive human serum S:N	DV antibody negative human serum S:N
R097A_4	1.27	1.11
R097A_13	1.53	1.50
R097A_43	1.09	0.93
R203A_2	1.12	1.15
R203A_54	1.01	0.95
R261A_1	0.97	1.06
R261A_4	0.93	1.06
R107A_1	0.92	0.94
R107A_22	0.89	0.88

Optimisation of the human serum concentration may have reduced the observed cross-reactivity and improved signal to noise ratios between the dengue VLP and lacZ samples. However, it was decided not to proceed until the dengue VLPs samples had been further purified to reduce the potential for cross reactivity.

5.8 Discussion

The first objective for this chapter was to develop and optimise a dengue VLP ELISA using the recombinant dengue VLPs produced in Chapter 4 as antigen. This was to enable detection of the dengue VLPs over the non-DV background proteins present in the dengue VLP samples. The optimised ELISA was then intended to be used as the basis for capture ELISAs using recombinant DC-SIGN and heparin, as DC-SIGN and HSHS (a heparin analogue) are implicated in DV attachment to host cells. These capture ELISAs would allow comparison of the host cell attachment capabilities of the DV quasispecies consensus and variant VLPs within each patient sample set. To compare the ability of the quasispecies consensus and variant VLPs to undergo low pH-induced E oligomeric rearrangements, dengue VLPs were intended to be subjected to low pH and the oligomeric conformation assessed. The dengue VLPs were also to be tested by ELISA using human serum containing DV antibodies, to assess their suitability as source of antigen for diagnostic ELISAs.

5.8.1 Development and optimisation of an ELISA using recombinant dengue virus-like particles as antigen

Initially two ELISAs were tested to compare dengue VLP and non-DV protein background detection using either a mouse mAb to DV E (mAb ELISA; Figure 5.1) or a rabbit pAb to DV types 1, 2, 3 and 4 (pAb ELISA; Figure 5.2) as the primary antibody. The lacZ sample (concentrated clarified medium from rBV.*lacZ*-infected cells) was used as a background control, representative of the non-DV proteins present in the dengue VLP samples. A statistical positive threshold value was calculated for each ELISA based on the average lacZ sample absorbance (from three replicates) plus three standard deviations, as this encompasses 99.7% of normally distributed data (Swinscow, 1996). The mAb ELISA proved superior to the pAb ELISA both in terms of a lower positive threshold value and greater signal to noise ratios (Table 5.1) for all but one (R097A_13) of the dengue VLP samples tested. However, both ELISAs still displayed significant antibody cross-reactivity with the lacZ background control sample.

Baculovirus GP64 was confirmed as the source of the antibody cross-reactivity observed using purified GP64 as antigen (Figures 5.1 to 5.3). For the pAb ELISA, results indicated that it was the primary antibody (rabbit pAb to DV types 1, 2, 3 and 4) that

was binding non-specifically to baculovirus GP64 (Figure 5.3). For the mAb ELISA, the secondary antibody (HRP-conjugated rabbit pAb to mouse IgG) was identified as the source of the non-specific binding (Figure 5.3). Reactivity to baculovirus GP64 was not observed when the pAb ELISA secondary antibody (HRP-conjugated goat pAb to mouse IgG) was used alone. Taken together, these results suggest that a component of rabbit serum is able to bind to baculovirus GP64, and that sourcing antibodies raised in a species other than rabbit might alleviate the cross-reactivity. No polyclonal DV antibodies raised in mouse or goat were able to be sourced commercially, so the pAb ELISA was not tested further. Non-specific antibody reactivity to non-DV proteins in the samples was shown to be consistent across all of the dengue VLP samples and the lacZ background control sample, based on the absorbance values generated by the mAb ELISA using only the secondary antibody (Figure 5.3). This confirmed that the lacZ sample was an appropriate background control for these ELISAs. Even without substituting the secondary antibody for the mAb ELISA, after optimisation of blocking conditions, and primary and secondary antibody concentrations, all nine of the dengue VLP samples were able to be detected over the non-DV protein background signal.

5.8.2 Comparison of dengue virus-like particle samples binding to DC-SIGN via a capture ELISA

DC-SIGN and other similar lectins have been proposed as DV receptors both at the initial site of infection (DCs; Tassaneetrithep *et al*., 2003) and also at secondary sites (endothelial cells and macrophages; Miller *et al*., 2008; Navarro-Sanchez *et al*., 2003). The ability of the dengue VLP samples to bind to recombinant DC-SIGN was assessed using the dengue VLP mAb ELISA with an additional DC-SIGN capture step. The dengue VLP mAb ELISA was performed in parallel to the capture ELISA using the same samples, so the relative absorbance signals could be directly compared between the two ELISAs. Lectins such as DC-SIGN bind high-mannose glycans in the presence of calcium ions (Feinberg *et al*., 2001). To the best of my knowledge, binding of baculovirus GP64 to DC-SIGN has not been shown, but was a possibility due to GP64s glycosylated status.

The initial DC-SIGN capture ELISA (Figure 5.8) was performed using all of the dengue VLP samples, and ELISA plates coated with DC-SIGN at 1 μg/mL. The dengue VLP

samples displayed a similarly low affinity for DC-SIGN as the lacZ sample, so the ELISA was repeated (Figure 5.9) using just the R107A_1 dengue VLP and lacZ samples, and ELISA plates coated with DC-SIGN at increasing concentrations (1 μg/mL to 100 μg/mL). The dengue VLP sample average absorbance signal was not greater than the calculated positive threshold value at any of the DC-SIGN coating concentrations tested. Although the average signals increased, the signal to noise ratios (Table 5.6) decreased with increasing DC-SIGN coating concentration. The absorbance signals generated were shown to result from the secondary antibody (HRP-conjugated rabbit pAb to mouse IgG) binding directly to DC-SIGN. The recombinant DC-SIGN used for the capture ELISA was a chimeric fusion protein, with the extracellular domain of human DC-SIGN fused to the Fc region of human IgG1. It may be that the secondary antibody was binding directly to the human IgG1 Fc region of the chimeric protein rather than the DC-SIGN region.

Both the dengue VLP and lacZ samples showed no affinity for DC-SIGN at any of the concentrations tested. This was unexpected for the dengue VLP samples and may indicate that the recombinant DV E glycans were not sufficiently processed by the insect cells, affecting binding of the dengue VLPs to DC-SIGN. It may be that further increasing the DC-SIGN coating concentration would enable the dengue VLPs to bind to the DC-SIGN. However, as an alternative recombinant DC-SIGN protein (not fused to the human IgG1 Fc region) would be required to overcome the problem of the secondary antibody binding to the recombinant DC-SIGN, it was decided to move on to testing the dengue VLP samples using a heparin capture ELISA.

5.8.3 Comparison of dengue virus-like particle samples binding to heparin via a capture ELISA

HSHS expressed on the surface of target cells has been shown to bind DV E (Chen *et al*., 1997), and is thought to concentrate the virus on the cell surface, enabling subsequent higher-affinity interactions with specific receptors. Heparin is also highly sulphated, with close structural homology to heparan sulphate. As such, heparin is an acceptable surrogate for HSHS in binding studies, and has been previously used in investigations of DV E binding to host cells (Chen *et al*., 1997; Marks *et al*., 2001;

Thullier *et al*., 2001). The ability of the dengue VLP samples to bind to heparin was assessed using the dengue VLP mAb ELISA with an additional heparin capture step. The dengue VLP mAb ELISA was performed in parallel to the capture ELISA using the same samples, and the lacZ sample was used as a background control.

The initial heparin capture ELISA (Figure 5.11) was performed using ELISA plates coated with heparin at 1 mg/mL. As the dengue VLP samples displayed a similarly low affinity for heparin as the lacZ sample, the ELISA was repeated (Figure 5.12) using just the R107A_1 dengue VLP and lacZ samples, with ELISA plates coated with increasing concentrations of heparin (1 mg/mL to 100 mg/mL. The dengue VLP sample average absorbance signal was greater than the calculated positive threshold at the highest heparin coating concentrations (50 and 100 mg/mL). The signal to noise ratios (Table 5.8) also increased with increasing heparin coating concentration, so that at the 50 and 100mg/mL heparin coating concentrations, the signal from the dengue VLP sample was clearly distinguishable from the background signal measured using the lacZ sample.

The heparin capture ELISA was then performed on all of the dengue VLP samples using ELISA plates coated with heparin at 100mg/mL, which had been shown to produce an optimal signal to noise ratio, and a low positive threshold value (Figure 5.12). The dengue VLP mAb ELISA was performed in parallel to the capture ELISA using the same samples so the absorbance signals could be directly compared between the two ELISAs (Figure 5.13). A measure of the dengue VLPs for each sample that were available to bind to the heparin was obtained using the dengue VLP mAb ELISA, and the proportion of these actually bound to heparin was calculated from the heparin capture ELISA results (Table 5.9). The R097A sample set showed the greatest discrepancy between the amount of available dengue VLPs and those able to bind to the heparin. The contrast between the quasispecies consensus and variant VLP samples was also more extreme for the R097A sample set than was seen with any of the other sample sets. For the consensus protein sample R097A_4, 97% of the available dengue VLPs bound to heparin, whilst for the variant protein samples R097A_13 and R097A_43, 0% and 20% respectively of the available dengue VLPs bound to heparin. The R203A and

R261A sample sets both exhibited an increase in the proportion of available dengue VLPs bound to heparin in the variant (R203A_54, 78%; R261A_4, 64%) compared to consensus protein samples (R203A_2, 49%; R261A_1, 51%). The R107A sample set showed almost no difference in heparin binding between consensus (R107A_1, 57%) and variant (R107A_22, 56%) protein samples.

The R097A_13 quasispecies variant VLP sample contains a single residue substitution in E compared to the R097A_4 quasispecies consensus VLP sample. This residue substitution is at position E320, and involves the substitution of a nonpolar amino acid with a large aliphatic hydrophobic side chain (isoleucine) with a polar amino acid (threonine). DV E position E320 was demonstrated by three-dimensional modelling (section 3.9) to be within 4Å of a proposed HSHS binding site implicated in DV binding to host cells (Chen *et al*., 1997; Thullier *et al*., 2001). Neutralisation escape mutations in multiple flaviviruses map to this HSHS binding site (Jiang *et al*., 1993; Lin *et al*., 1994). Isoleucine is often important for the correct folding of protein, so it may be that the substitution of isoleucine for threonine is sufficient to disrupt the protein structure around the HSHS binding site so that binding to heparin is prevented. This would be consistent with the heparin capture ELISA results, which showed no heparin binding for this dengue VLP sample.

The R097A_43 quasispecies variant VLP sample also contains a single residue substitution in E compared to the R097A_4 quasispecies consensus VLP sample. This residue substitution is at position E280 and involves the substitution of a basic amino acid with an imidazole ring (histidine) for a nonpolar amino acid with a large aliphatic hydrophobic side chain (leucine). DV E position E280 was demonstrated by three-dimensional modelling (section 3.9) to be surface accessible, and located within the *kl* hairpin loop at the domain *I*/*II* interface, which is implicated in conformational rearrangements during fusion (Modis *et al*., 2003). Removal of the histidine residue at this location was predicted to impact on the protein's response to acidic environments, not cell attachment. However, although E position E280 is distant from the HSHS binding site, it does reside at the interface between structural domains *I* and *II.*

Flavivirus E amino acid substitutions that affect virulence have been shown to cluster at the interfaces between E structural domains (Rey *et al*., 1995); Figure 1.16). In this case, the substituted amino acid leucine will tend to orient to the interior of the folded protein due to its hydrophobic nature, which may impact upon protein folding. Changes in structure at protein domain interfaces could have magnified consequences for other domains within the protein as the interfaces between domains often act as hinge regions during conformational change. This residue substitution at such an important location may alter the relative orientations of the other protein domains such that the HSHS binding site is less accessible (relative to the quasispecies consensus VLP sample R097A_4). This would be consistent with the heparin capture ELISA results for this sample, which showed only 20% of the available dengue VLPs bound to heparin.

Decreased ability to bind HSHS might be expected to confer a selective disadvantage for the DV quasispecies variant, as the cell attachment process would be less efficient. This would lead to decreased infectivity compared to variants demonstrating a higher affinity for HSHS. However, quasispecies variants with lower HSHS affinity would still be able to enter cells in an antibody-dependent manner by forming a complex with non-neutralising antibodies from a previous DV infection. Such quasispecies variants would then enter leukocytes via fc receptor-mediated endocytosis, but remain in a non-neutralised state and be able to replicate and disseminate more virus. The production and secretion of chemokines and cytokines by the cell in response to the subsequent increase in viral replication is thought to lead to the increase in vascular permeability that is the hallmark of severe dengue (Halstead, 1988). The R097A sample was obtained from a patient with severe dengue, which was later confirmed to be a secondary infection. Therefore, this patient had pre-existing antibodies to a previous DV infection, which would have enabled uptake of quasispecies variants with low heparin binding affinity into the host cells.

The R203A_54 and R261A_4 quasispecies variant VLPs each contain amino acid substitutions within E compared to their respective quasispecies consensus VLPs

(R203A_2 and R261A_1). For R203A_54 there are amino acid substitutions at two locations. One of these is at position E202, and involves a conservative amino acid change where a basic amino acid (lysine) is replaced with a functionally similar basic amino acid (arginine). Three-dimensional modelling (Table 3.15) demonstrated DV E position E202 to be within 4Å of the other protein chain that makes up the dimer. Mutations leading to amino acid substitutions at E position E202 commonly occur during routine serial passage in mammalian cells (Lee, Weir and Dalgarno 1997). Therefore, as suggested in Chapter 3, this mutation may represent a beneficial adaptation of the virus to growth in mammalian cells. The other amino acid change in R203A_54 is at position E345, and involves a drastic amino acid substitution where three amino acids (histidine-asparagine-glycine) are replaced with five amino acids (isoleucine-glutamine-isoleucine-serine-serine). Incorporating the two extra residues into the protein chain was predicted to impact on protein structure (section 3.10). Three-dimensional modelling (section 3.9) demonstrated DV E position E345 to be situated on the surface of the protein at the distal face of domain *III*. This part of the domain is thought to contain the flavivirus receptor-binding motif (Crill and Roehrig, 2001) and amino acid subtitutions in this part of the domain are known to affect virulence and cell tropism, presumably through interference with cell attachment. This is consistent with the heparin capture ELISA results for this sample, which showed increased ability to bind heparin compared to the quasispecies consensus dengue VLP for this sample set (R203A_2).

The R261A_4 amino acid substitution is at position E201 and involves a conservative amino acid change where a polar amino acid with a long hydrophilic side chain (asparagine) is replaced with a polar amino acid (serine). Three-dimensional modelling (Table 3.15) demonstrated DV E position E201 to be surface accessible, and located within the same hinge region at the domain *I*/*II* interface as the R097A_43 drastic amino acid substitution and the R203A_54 conservative substitutions. This conservative protein change at this location may alter the relative orientations of the other protein domains, such that the HSHS binding site is more accessible (relative to the quasispecies consensus VLP sample R261A_1). This would be consistent with the

heparin capture ELISA results, which showed increased ability to bind heparin for the R261A_4 dengue VLP sample compared to the quasispecies consensus VLP for this sample set (R261A_1).

Increased ability to bind HSHS could be predicted to confer a selective advantage for the DV quasispecies variant, as the cell attachment process would be more efficient. This could lead to increased infectivity compared to variants demonstrating a lower affinity for HSHS, which would be reflected in the quasispecies population, as variants with high affinity for HSHS would be expected to replicate faster and comprise a higher proportion of the population. In support of this, the residue substitution at E position E201 was identified in six (13%) of the clones from the DV quasispecies population from patient R261A. In Chapter 3 it was proposed that the relative abundance of this lineage within the quasispecies population for this sample could be due to increased replication efficiency. However, routine passage of a DV type 4 vaccine candidate in fetal rhesus lung cells yielded a virus variant with increased HSHS binding affinity, which predominated over other quasipecies variants within three passages, but exhibited reduced infectivity and immunogenicity in rhesus monkeys (Anez *et al*., 2009). This virus variant became less dominant within the quasispecies population when the virus population was transferred from mammalian cell culture into a mosquito cell line (C6/36); selection of the parental wild-type virus was favoured in the mosquito cells. Rapid clearance from the bloodstream and reduced ability to cause disseminated infection were proposed as the explanation for virus attenuation despite increased HSHS binding affinity (Anez *et al*., 2009). HSHS is a component of extracellular matrices so it is conceivable that increased heparin binding ability could result in the virus becoming trapped extracellularly and therefore being cleared more rapidly from the host.

The R203A sample was obtained from a patient with severe dengue, which was later confirmed to be a secondary infection. The R261A sample was obtained from a patient with mild dengue, but follow-up samples after discharge from hospital were not taken, so serological confirmation of the type of infection (primary or secondary) could not be

obtained. If patient R261 had a primary infection, and therefore no pre-existing DV antibodies, this could explain the difference in disease severity between patient R261 and patient R203 (who had a secondary infection), despite the increased heparin binding ability of the quasispecies variants.

5.8.4 Comparison of dengue virus-like particle samples oligomeric rearrangements in response to low pH

It had been intended to compare the quasispecies consensus and variant dengue VLPs in terms of their ability to undergo conformational rearrangements in response to low pH. These conformational rearrangements are required during DV infection to enable fusion of virus and cell membranes and subsequent replication of the virus genome by the host cell. However, given the similarities between DV E and baculovirus GP64, it was decided that it would be too difficult to distinguish between them using this method. This experiment was therefore put on hold until the dengue VLPs and baculovirus had been separated by further purification.

5.8.5 Dengue virus-like particle ELISA using human serum

To assess the suitability of the dengue VLPs as source of antigen for diagnostic ELISAs, a dengue VLP ELISA was performed using human serum containing antibodies to all four DV types (Figure 5.15). High levels of binding were observed between the human serum and the lacZ background control sample, indicating that the human serum was binding non-specifically to non-DV proteins within the sample. The same problem had been observed previously in the dengue VLP pAb ELISA, which used rabbit polyclonal serum containing DV antibodies as the primary antibody, and to a lesser extent in the dengue VLP mAb ELISA, which used rabbit pAb to mouse IgG as the secondary detection antibody. In both of these cases, baculovirus GP64 was shown to be the source of the cross-reactivity, using purified GP64, so it is likely that is also the case with the human serum.

It was therefore demonstrated that these dengue VLP samples were unsuitable for use as antigen in diagnostic ELISAs due to the cross-reactivity observed. Although this cross-reactivity may have been reduced by optimisation of the human serum

concentration, it was decided not to proceed with this until the dengue VLP samples had been further purified to remove the contaminating baculovirus.

5.9 Conclusions and summary

As set out in the objectives for this chapter, an ELISA was developed and optimised that enabled detection of all of the recombinant dengue VLPs produced in Chapter 4. With this ELISA it was demonstrated that similar levels of background (non-DV) protein were present in the dengue VLP and lacZ background control samples, confirming the suitability of the lacZ sample as a background control using this ELISA method. Baculovirus GP64 was indicated as the primary cause of the antibody cross-reactivity using purified GP64.

To assess the cell-attachment capabilities of the dengue VLPs, and perform comparisons between the quasispecies consensus and variant VLPs from each sample set, DC-SIGN and heparin capture ELISAs were performed. The DC-SIGN capture ELISA was unsuccessful as no dengue VLP binding to DC-SIGN was detected over that observed from the lacZ background control. The absorbance signals generated were shown to result from the secondary antibody binding directly to the recombinant DC-SIGN. It was decided not to proceed with the DC-SIGN capture ELISA until an alternative source of either recombinant DC-SIGN or secondary antibody that overcame these problems could be acquired.

The heparin capture ELISA was developed and optimised, and showed differences in heparin binding ability of four of the variant VLP samples compared to their respective consensus VLPs. Reductions in heparin binding ability were observed for both of the variant VLP samples (R097A_13 and R097A_43) from the R097A sample set. The R097A_13 sample contained a drastic residue substitution within 4Å of the proposed HSHS binding site, so alterations in heparin binding had been predicted previously (Table 3.15) for this sample. The R097A_43 sample contained a drastic residue substitution at the interface of DV E structural domains *I* and *II*, which had been predicted to impact upon low pH-induced conformational change and fusion rather than cell attachment. Improvements in heparin binding ability were observed for the variant VLP samples (R203A_54 and R261A_4) from both the R203A and R261A sample sets. The R203A_54 VLP sample contained amino acid substitutions in two locations. One of these was a drastic residue substitution on the distal face of domain *III*, which

had been predicted to potentially alter cell attachment. The other amino acid substitution in the R203A_54 sample was a conservative residue substitutions at E201. A similar conservative amino acid change was observed in the R261A_4 VLP sample at E202, within the interface between structural domains *I* and *II*, as described for the R097A_43 VLP sample. It is proposed that, as the interfaces between DV structural domains function as hinge regions during conformational changes, residue substitutions in these positions may alter the relative orientation of other domains within the protein. This could be either detrimental (as in the case of the R097A_43 drastic residue substitution) or beneficial (as in the case of the R203A_54 and R261A_4 conservative residue substitutions) to heparin binding affinity, by altering the accessibility of the HSHS binding site. Further support could be lent to these conclusions by computational modelling of the effects of these residue substitutions on the structure of DV E. These results are a preliminary indication of differences in infectivity possessed by DV variants within the quasispecies population. Linking these observations to disease severity is complex, as quasispecies variants with low heparin binding affinity can still enter cells in an antibody-dependent manner and subsequently replicate, provided pre-existing antibodies to a previous DV infection are present. Secondary DV infections were confirmed in patients R097 and R203, who both experienced severe disease. However, for patient R261, who experienced mild disease, but also exhibited an abundance of quasispecies variants with increased heparin binding affinity, serological confirmation of the type of infection (primary or secondary) could not be obtained.

To assess the suitability of the dengue VLPs as source of antigen for diagnostic ELISAs, assays were performed using human serum containing antibodies to DV as the primary antibody. These ELISAs were unsuccessful as it was found that human serum (both positive and negative for DV antibodies) bound to the lacZ background control sample similarly to the dengue VLP samples. Similar cross-reactivity was observed previously using rabbit serum in the initial dengue VLP mAb and pAb ELISAs, and baculovirus GP64 identified as the likely cause using purified GP64. Further analysis was deemed

unfeasible until the dengue VLPs had been effectively purified to remove the contaminating baculovirus.

It was also intended as part of this chapter to investigate the ability of the dengue VLPs to undergo low pH-induced oligomeric rearrangements required for membrane fusion. However, due to the similarities between DV E and baculovirus GP64, this experiment was postponed until the dengue VLPs could be further purified.

CHAPTER 6. FINAL DISCUSSION, CONCLUSIONS AND FUTURE WORK

The work presented in this thesis was undertaken to improve the available knowledge surrounding dengue virus (DV) quasispecies populations and to investigate links with disease severity. The hypothesis was that DV quasispecies populations affect disease severity in individual patients. Work previous to this thesis demonstrated that DVs exist as quasispecies populations within individual mosquito and human hosts (Chao *et al*., 2005; Craig *et al*., 2003; Lin *et al*., 2004; Wang *et al*., 2002a; Wang *et al*., 2002b). Initial studies looked at DVs from patients or mosquitoes in Taiwan (Lin *et al*., 2004; Wang *et al*., 2002a) and Myanmar (Craig *et al*., 2003), and focused on the DV envelope glycoprotein (E) gene, as it encodes E, which comprises the majority of the outer surface of the virion. Subsequent work has shown that the DV E gene region exhibits greater heterogeneity than any other region of the genome (Chao *et al*., 2005; Wang *et al*., 2002b). As a result of so few studies, little is known about how DV quasispecies population dynamics vary between different individuals, host populations and infecting virus strains. Expanding the knowledge in this area is essential for in-depth understanding of the biology, epidemiology and evolution of DV. This has implications for disease pathogenesis, vaccine design and safety, the implementation of effective control and treatment programs, and the development of novel therapeutic and diagnostic tools.

The primary aim of the work presented in this thesis was to investigate DV quasispecies populations, both within individual patient samples and between samples from the same outbreak in Sri Lanka, in 2006. Relationships between quasispecies variation and disease severity were also examined. Across the six dengue patient samples successfully analysed, an average mean diversity of 0.018% (range 0.0034% to 0.053%) was obtained for the intrahost quasispecies populations, which is consistent with reported error rates for viral RNA polymerases during natural replication (0.01%; Smith *et al*., 1997). No statistically significant correlations were observed between the extent of quasispecies variation and disease severity. Compared to previously reported mean diversities of between 0.22% and 0.47% (Craig *et al*., 2003; Lin *et al*., 2004; Wang *et al*., 2002a), the results from this study show considerably less quasispecies variation in all of the patient samples analysed. This difference in the

extent of quasispecies sequence variation was shown to be directly related to the fidelity of the enzymes used by each group for RT-PCR amplification of the DV RNA. These results have demonstrated clearly that future work of this kind must be conducted using high-fidelity enzymes for both reverse transcription and amplification of cDNA by PCR. In light of the results from this study, previously reported DV quasispecies mean diversities need to be re-examined to account more accurately for mutations introduced by the amplification process itself.

Consistent with previous reports (Wang *et al*., 2002a), genome-defective quasispecies variants with insertion or deletion mutations were observed predominantly in samples from patients with severe cases of dengue. These observations are consistent with a role for specific quasispecies variants in the pathogenesis of disease; however, further work would be required to determine their contribution to disease pathogenesis. It has been suggested previously that the transmission of defective quasispecies variants may be selectively advantageous for co-infecting non-defective variants by ensuring the presence of viruses differentially adapted to human or mosquito cells, modulating host immune responses, or allowing the production of extra C and prM (Aaskov *et al*., 2006). Longer-term studies of DV quasispecies populations from humans and mosquitoes in Sri Lanka would be required to determine whether the genome-defective quasispecies variants identified in this study were stably transmitted, and were present after the 2006 dengue season.

Given the technological advances that have occurred since beginning this work, in particular with regards to next-generation sequencing, it is likely that studies like this will be conducted very differently in future. Massively parallel ultra-deep pyrosequencing has been used for the detailed analysis of human immunodeficiency virus (HIV) quasispecies populations from individual patients (Rozera *et al*., 2009). PCR products were clonally amplified on capture beads in water-in-oil emulsion micro-reactors, abolishing the need for cloning into *E. coli*. The quantity of data generated (13,456 reads per sample) was several orders of magnitude higher than any previously obtained using conventional approaches based on cloning of PCR products or limiting

dilution PCR. Similarly, next-generation sequencing has been used to examine quasispecies populations of foot and mouth disease virus (Wright *et al.*, 2011) and the 2009 pandemic influenza A virus (Kuroda *et al.*, 2010). The current limitations to this approach include the requirement for large quantities of DNA (µg) as starting material, short sequence read lengths, high intrinsic error frequency, high cost, and difficulties with computational analysis of such large data sets. These are likely to be improved as development of the technology progresses.

To investigate relationships between the infecting DV genotype and disease severity, phylogenetic analyses of the intrahost quasispecies populations were conducted for the six dengue patient samples using the quasispecies E gene sequences. Within each DV type, strains (quasispecies consensus sequences) from different patients were identified as having greater than 99% sequence homology, and were shown to be the same genotype, appropriate to the geographic origin of the samples. No relationship was observed between the infecting DV genotype and disease severity, because within DV types, all of the quasispecies consensus and variant E gene sequences were identified as the same genotype, regardless of the severity of disease experienced by the patient. Consistent with this observation, there was no evidence of intra-DV type (inter-genotype) recombination in the quasispecies populations from any of the patient samples. However, sequencing the rest of the genome for the DV quasispecies populations from these patient samples could yield further differences between them that could account for the differences in disease severity, and would enable identification of recombination events outside of the E gene. Mutations in the 5' and 3' non-coding regions and the non-structural-5 (NS5) gene of DVs have been shown to be associated with changes in viral replication efficiency (Leitmeyer *et al*., 1999; Tajima *et al*., 2007; Zhao *et al*., 2010). Increased viral replication efficiency has been linked to high viral loads in blood, and severe disease presentation (Wang *et al*., 2006).

The locations of amino acid substitutions encoded by the DV quasispecies E gene mutations were mapped to the three-dimensional structure of DV E. This was to identify those mutations with the potential to affect virulence or cell tropism on the

basis of their proximity to important structural features. Amino acid substitutions that were surface accessible, or proximal to the fusion peptide, proposed HSHS binding site or the other protein chain within the E dimer, were highlighted. Amino acid substitutions were less frequently observed at these locations in the samples from patients with mild disease than in those from patients with severe disease, which is also consistent with a role for specific quasispecies variants in the pathogenesis of disease. Using this method, nine E gene clones were identified for recombinant protein production to further investigate the effects of the mutations on protein function. These nine variants were from four different patient samples, and were chosen because they either represented the quasispecies consensus E gene sequence for that patient sample, or they contained a nonsynonymous non-frameshift mutation with the potential to affect virulence or cell tropism.

Recombinant baculoviruses (rBVs) were produced containing the DV type 1 pre-membrane glycoprotein (prM) gene, and each of the quasispecies variant E genes. Upon infection of insect cells with the rBVs, secreted dengue virus-like particles (VLPs) were produced, comprising DV type 1 membrane (M) and quasispecies E. The source of the prM gene was kept consistent so that any differences in protein function between the VLPs could be attributed to variations in E. Attempts to purify the dengue VLPs were unsuccessful, and there were problems with immunodetection of the dengue VLPs from Western blots due to antibody cross-reactivity with baculovirus GP64. Production of dengue VLPs in a plasmid vector based expression system would provide a cleaner VLP preparation in that there would not be baculovirus, or other viral vectors present. This would overcome the problems observed with antibody cross-reactivity with baculovirus GP64, thus enabling efficient purification by immunoprecipitation and also improving immunodetection.

The use of recombinant proteins as antigen avoids the need to culture and purify large quantities of DV, traditionally sourced from tissue culture or suckling mouse brain, which requires high-containment facilities (Containment Level 3). It also permits the introduction of gene mutations that lead to amino acid substitutions in the protein

that may alleviate the flavivirus antibody cross-reactivity observed using wild-type proteins. To assess the suitability of the dengue VLPs as a source of antigen for diagnostic ELISAs, assays were performed using human serum containing antibodies to DV. Unexpectedly the human serum was shown to bind non-specifically to a non-DV component of the clarified culture medium from rBV-infected insect cells. This cross-reactivity prevented assessment of the dengue VLPs for sensitivity and specificity in diagnostic ELISAs using human serum. Similar cross-reactivity was observed previously using rabbit serum, and baculovirus GP64 was identified as the cause. Further purification of the dengue VLPs to remove the contaminating baculovirus would enable these assays to be repeated, and for the dengue VLPs to be tested against serum collected from individuals vaccinated against non-DV flaviviruses. Quasispecies variants demonstrating reduced flavivirus cross-reactivity could then be used to develop improved assays for the diagnosis of DV infection, as was intended during this study. To the best of my knowledge, quasispecies variants have not been previously assessed in this way. However, during the course of this work it was reported that flavivirus VLPs constructed using SLEV, WNV and JEV prM and E genes with mutations in the fusion peptide exhibit dramatically reduced flavivirus antibody cross-reactivity compared to wild-type VLPs using IgM capture ELISAs (Chiou *et al*., 2008; Roberson, Crill, and Chang, 2007). The work presented in this study also identified a need for more DV-specific antibodies to be made available to the wider scientific community, as many of the commercially available antibodies tested for this work were found to be not sensitive, specific or reliable enough for use. Purified quasispecies variant dengue VLPs demonstrating reduced flavivirus cross-reactivity could also be used as immunogen to produce more effective DV-specific antibodies.

This study demonstrated alterations in heparin binding ability in four of the quasispecies variant VLP samples (R097A_13, R097A_43, R203A_54 and R261A_4) compared to their respective consensus VLPs (R097A_4, R203A_2 and R261A_1). Both samples displaying reductions in heparin binding ability were from the same patient (R097A), who exhibited severe disease. The two samples that showed increased heparin binding ability were from different patients, one of which had mild disease

(R261A), and one had severe disease (R203A). The R097A_13 and R203A_54 dengue VLP samples contained nonsynonymous E gene mutations in close proximity to the putative highly-sulphated heparan sulphate (HSHS) binding region of E. The R097A_43 and R261A_4 dengue VLP samples contained nonsynonymous E gene mutations at the interface of E structural domains *I* and *II,* which is thought to act as a hinge region during low pH-induced conformational changes that lead to fusion of virus and host cell membranes. It is hypothesised that residue substitutions in these positions may alter the relative orientation of other domains within the protein, thus altering the accessibility of the HSHS binding site. Given the status of HSHS as a putative receptor for DV, it is tempting to predict that increased heparin binding ability would lead to increased viral replication and correlate with severe disease. However, quasispecies variants with low heparin binding affinity can still enter cells in an antibody-dependent manner and subsequently replicate, provided pre-existing antibodies to a previous DV infection are present. Also, HSHS is a constituent of extracellular matrices, as well as being present on the surface of cells targeted by DVs. Therefore, increased heparin binding affinity could result in the virus becoming trapped extracellularly. This may aid the formation of immune complexes containing DV, but whether this would result in a more rapid clearance of the virus from the system, or whether the increased heparin binding affinity would slow down the process of virus clearance remains to be determined. During defervescence, slower rates of clearance of DV and immune complexes have been associated with subsequent immune activation and progression to severe disease (Wang *et al*., 2006). Further work is required to determine whether the differences in heparin binding ability observed during this study result in altered cell attachment capabilities and subsequent changes in the efficiency of viral replication. Potential future work includes using surface plasmon resonance (SPR) to monitor the interactions between the quasispecies dengue VLPs and heparin in real-time and obtain more data regarding the binding affinity and kinetics. Cells expressing HSHS on their surface could also be used in place of heparin, to provide a better measure of the cell attachment capabilities of the quasispecies variant dengue VLPs. This method could also be used to investigate potential anti-viral compounds that act

by inhibiting virus attachment to host cells, as has been recently shown for noroviruses, using VLPs (Rademacher *et al*., 2011).

For this study, a limited number of the identified quasispecies variants were produced as recombinant DV E in the form of dengue VLPs, to study the effects of the E gene mutations on E function. An alternative methodology would have been to use the quasispecies E gene variants to construct genome-length infectious clones. Positive strand RNA virus (e.g. flavivirus) genomes are infectious, as RNA acts directly as mRNA for protein synthesis. Therefore, when flavivirus cDNA is transfected into susceptible eukaryotic cells, it undergoes nuclear transcription to RNA, which is then replicated and viral protein synthesis is initiated. Flavivirus genome-length infectious clones require the same level of bio-containment (Containment Level 3) as wild-type flaviviruses, and are difficult to produce, as flavivirus cDNA is unstable in *E. coli* (Ward and Davidson, 2008). To circumvent the problem of instability, infectious clones have been produced *in vitro* by ligation or long-template high-fidelity PCR (Campbell and Pletnev, 2000; Sumiyoshi, Hoke, and Trent, 1992), or in *E. coli* with the addition of short introns to prevent spurious expression (Yamshchikov, Mishin, and Cominelli, 2001). The use of genome-length infectious clones incorporating the quasispecies variant E genes would enable the cell-attachment capabilities of E to be tested using live virus. It had been intended as part of this study to subject the quasispecies dengue VLPs to low pH and assess their ability to form E trimers; however, due to the problems experienced with dengue VLP purification and immunodetection, this was not possible. This methodology would have yielded limited results as it would only have identified dengue VLP quasispecies variants where trimerisation was either permitted or prohibited due to the mutations present. Using genome-length infectious clones, E oligomer and virion assembly, and the ability of E to undergo low pH-induced oligomeric rearrangements required for membrane fusion could be studied. The use of genome-length infectious clones and cell-based fusion assays would permit more in-depth analysis of the pH required, and time taken for each of the quasispecies variant E to undergo the low pH-induced oligomeric rearrangements

required for membrane fusion. The subsequent fusion, replication and virus dissemination processes could also be examined by monitoring the cells.

Quasispecies variants that escape neutralisation by pre-existing heterotypic antibodies will be able to replicate, and therefore when inoculated into a new host, they will comprise a greater proportion of the quasispecies population. This could potentially lead to increased disease severity and case fatality rates as an epidemic progresses (Guzmán *et al*., 2000). Genome-length infectious clones incorporating the quasispecies variant E genes could be used in neutralisation assays to assess neutralisation of the specific variants within the quasispecies population using serum from the Sri Lankan patients. This would identify quasispecies variants with neutralisation escape mutations, and enable assessment of potential links between the presence of such variants and disease severity in the patient. Differences in replication efficiency between quasispecies consensus and variant viruses could also be determined using these genome-length infectious clones and primary human and mosquito cells. To determine transmissibility of different variants within the quasispecies population, genome-length infectious clones representing different quasispecies variants could be co-infected into live mosquitoes and used to determine the susceptibility of the mosquito population to infection with the different quasispecies variants. When this method was used to compare the replication efficiency and dissemination of South East Asian and American genotype DV type 2 strains, the more virulent South East Asian genotype strains were shown to replicate to higher titres in the midgut and disseminate to the salivary glands (and therefore be available to transmit to humans) up to seven days faster than the less virulent American genotype strains, which do not cause epidemics of severe disease (Anderson and Rico-Hesse, 2006). This increase in vectorial capacity for the more virulent strains was proposed as a mechanism by which the less virulent American genotype strains have been displaced by the South East Asian genotype strains causing severe disease (Anderson and Rico-Hesse, 2006). It would be interesting to investigate whether the same was true for the quasispecies variants compared to their respective consensus viruses. Restricted quasispecies populations have been reported for naturally infected

mosquitoes compared to those found in patient samples during the same outbreak (Lin *et al*., 2004). This suggests that the quasispecies variants observed in the patient samples will experience selection pressure when transmitted to the mosquito vector, as some will be better adapted to survival in the mosquito cells than others.

6.1 Conclusions

In conclusion, the hypothesis that DV quasispecies populations affect disease severity in individual patients was neither proven nor disproven. No correlation was observed between the overall extent of quasispecies variation and disease severity. However, genome-defective quasispecies variants, and variants with mutations leading to amino acid substitutions that were surface accessible, or proximal to the fusion peptide, proposed HSHS binding site, or the other protein chain within the E dimer, were observed predominantly in samples from patients with severe cases of dengue. For four of these quasispecies variants, the mutations in the E gene were shown to cause alterations in the heparin binding ability of the recombinant E. Further work will be required to confirm that these differences represent varying cell attachment capabilities of the quasispecies variants compared to their respective consensus VLPs. It is important to remember that for this study, recombinant dengue VLPs were only able to be produced for a very small subset of the DV quasispecies populations, so it is difficult to link these observations with disease severity, as the net effect of the entire population needs to be taken into account. The production of dengue VLPs, or genome-length infectious clones for more of the quasispecies variants identified in this study would subsequently provide more data for analysis.

It is extremely difficult to correlate *in vitro* observations of DV behaviour to disease pathogenesis and epidemiology. This is predominantly because dengue pathogenicity results from the complex interactions of both viral and host factors, and there is a lack of suitable animal models for dengue, as DVs only cause symptomatic disease in humans. The majority of the available information regarding DVs has been obtained from patients with clinically apparent illness, and there are very few samples available from asymptomatic human infections. For these reasons, the natural spectrum of DV variation, and the relative virulence and transmissibility of strains, let alone quasispecies variants are difficult to determine with any confidence. The mechanisms of dengue pathogenesis are therefore likely to remain controversial until appropriate *in vivo* models representing the entire spectrum of disease have been developed.

REFERENCES

Aaskov, J., Buzacott, K., Field, E., Lowry, K., Berlioz-Arthaud, A., and Holmes, E. C. (2007). Multiple recombinant dengue type 1 viruses in an isolate from a dengue patient. *Journal of General Virology* **88**(12)**,** 3334-3340.

Aaskov, J., Buzacott, K., Thu, H. M., Lowry, K., and Holmes, E. C. (2006). Long-Term Transmission of Defective RNA Viruses in Humans and *Aedes* Mosquitoes *Science* **311,** 236-238.

AbuBakar, S., Wong, P. F., and Chan, Y. F. (2002). Emergence of dengue virus type 4 genotype IIA in Malaysia. *Journal of General Virology* **83**(Pt 10)**,** 2437 - 2442.

Allison, S. L., Stadler, K., Mandl, C. W., Kunz, C., and Heinz, F. X. (1995). Synthesis and secretion of recombinant tick-borne encephalitis virus protein E in soluble and particulate form. *Journal of Virology* **69**(9)**,** 5816-5820.

Allison, S. L., Stiasny, K., Stadler, K., Mandl, C. W., and Heinz, F. X. (1999). Mapping of Functional Elements in the Stem-Anchor Region of Tick-Borne Encephalitis Virus Envelope Protein E. *Journal of Virology* **73**(7)**,** 5605-5612.

Alvarez, D. E., Lodeiro, M. F., Luduena, S. J., Pietrasanta, L. I., and Gamarnik, A. V. (2005). Long-Range RNA-RNA Interactions Circularize the Dengue Virus Genome. *Journal of Virology* **79**(11)**,** 6631-6643.

Amin, M. R., Mahbub, A. H. M., Sikder, A. R., and Karim, M. M. (2010). Prediction of the post-translational modification sites on dengue virus E protein and deciphering their role in pathogenesis. *International Journal of Bioinformatics Research and Applications* **6**(5)**,** 508-521.

AnandaRao, R., Swaminathan, S., Fernando, S., Jana, A. M., and Khanna, N. (2006). Recombinant Multiepitope Protein for Early Detection of Dengue Infections. *Clinical and Vaccine Immunology.* **13**(1)**,** 59-67.

Anderson, J. R., and Rico-Hesse, R. (2006). *Aedes aegypti* vectorial capacity is determined by the infecting genotype of dengue virus. *The American Journal of Tropical Medicine and Hygiene* **75**(5)**,** 886-892.

Anez, G., Men, R., Eckels, K. H., and Lai, C.-J. (2009). Passage of Dengue Virus Type 4 Vaccine Candidates in Fetal Rhesus Lung Cells Selects Heparin-Sensitive Variants That Result in Loss of Infectivity and Immunogenicity in Rhesus Macaques. *Journal of Virology* **83**(20)**,** 10384-10394.

Arezi, B., and Hogrefe, H. H. (2007). *Escherichia coli* DNA polymerase III ε subunit increases Moloney murine leukemia virus reverse transcriptase fidelity and accuracy of RT-PCR procedures. *Analytical Biochemistry* **360**(1)**,** 84-91.

Armstrong, P. M., and Rico-Hesse, R. (2001). Differential Susceptibility of *Aedes aegypti* to Infection by the American and Southeast Asian Genotypes of Dengue Type 2 Virus. *Vector-Borne and Zoonotic Diseases* **1**(2)**,** 159-168.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**(2)**,** 195-201.

Ball, L. A. (2001). Replication Strategies of RNA Viruses. 4th ed. *In* "*Fields Virology*" (D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus, Eds.), Vol. 1, pp. 105-118. 2 vols. Lippincott Williams & Wilkins, Philadelphia.

Bandyopadhyay, S., Lum, L. C., and Kroeger, A. (2006). Classifying dengue: a review of the difficulties in using the WHO case classification for dengue haemorrhagic fever. *Tropical Medicine and International Health* **11**(8)**,** 1238-1255.

Barnes, W. J. S., and Rosen, L. (1974). Fatal Hemorrhagic Disease and Shock Associated with Primary Dengue Infection on a Pacific Island. *The American Journal of Tropical Medicine and Hygiene* **23**(3)**,** 495-506.

References

Beasley, D. W. C., and Aaskov, J. G. (2001). Epitopes on the Dengue 1 Virus Envelope Protein Recognized by Neutralizing IgM Monoclonal Antibodies. *Virology* **279**(2)**,** 447-458.

Bharaj, P., Chahar, H. S., Pandey, A., Diddi, K., Lalit, D., Randeep, G., Kabra, S. K., and Broor, S. (2008). Concurrent infections by all four dengue virus serotypes during an outbreak of dengue in 2006 in Delhi, India. *Virology Journal* **5,** 1.

Bielefeldt-Ohmann, H., Beasley, D. W., Fitzpatrick, D. R., and Aaskov, J. G. (1997). Analysis of a recombinant dengue-2 virus-dengue-3 virus hybrid envelope protein expressed in a secretory baculovirus system. *Journal of General Virology* **78**(11)**,** 2723-2733.

Bielefeldt-Ohmann, H., Meyer, M., Fitzpatrick, D. R., and Mackenzie, J. S. (2001). Dengue virus binding to human leukocyte cell lines: receptor usage differs between cell types and virus strains. *Virus Research* **73**(1)**,** 81-89.

Bray, M., Zhao, B. T., Markoff, L., Eckels, K. H., Chanock, R. M., and Lai, C. J. (1989). Mice immunized with recombinant vaccinia virus expressing dengue 4 virus structural proteins with or without nonstructural protein NS1 are protected against fatal dengue virus encephalitis. *Journal of Virology* **63**(6)**,** 2853-2856.

Bryant, J. E., Calvert, A. E., Mesesan, K., Crabtree, M. B., Volpe, K. E., Silengo, S., Kinney, R. M., Huang, C. Y. H., Miller, B. R., and Roehrig, J. T. (2007). Glycosylation of the dengue 2 virus E protein at N67 is critical for virus growth in vitro but not for growth in intrathoracically inoculated *Aedes aegypti* mosquitoes. *Virology* **366**(2)**,** 415-423.

Burke, D. S., and Monath, T. P. (2001). Flaviviruses. 4th ed. *In* "*Fields Virology*" (D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus, Eds.), Vol. 1. 2 vols. Lippincott Williams & Wilkins, Philadelphia, Pa, Philadelphia.

Campbell, M. S., and Pletnev, A. G. (2000). Infectious cDNA Clones of Langat Tick-Borne Flavivirus That Differ from Their Parent in Peripheral Neurovirulence. *Virology* **269**(1)**,** 225-237.

Cecilia, D., and Gould, E. A. (1991). Nucleotide changes responsible for loss of neuroinvasiveness in Japanese encephalitis virus neutralization-resistant mutants. *Virology* **181**(1)**,** 70-77.

Centers for Disease Control and Prevention (2009). "CDC Health Information for International Travel 2010 (the Yellow Book), Elsevier, Atlanta.

Chang, G.-J. J., Hunt, A. R., Holmes, D. A., Springfield, T., Chiueh, T.-S., Roehrig, J. T., and Gubler, D. J. (2003). Enhancing biosynthesis and secretion of premembrane and envelope proteins by the chimeric plasmid of dengue virus type 2 and japanese encephalitis virus. *Virology* **306**(1)**,** 170-180.

Chao, D.-Y., King, C.-C., Wang, W.-K., Chen, W.-J., Wu, H.-L., and Chang, G.-J. J. (2005). Strategically examining the full-genome of dengue virus type 3 in clinical isolates reveals its mutation spectra. *Virology Journal* **2,** 72.

Chen, W., Kawano, H., Men, R., Clark, D., and Lai, C. J. (1995). Construction of intertypic chimeric dengue viruses exhibiting type 3 antigenicity and neurovirulence for mice. *Journal of Virology* **69**(8)**,** 5186-5190.

Chen, W. J., Wu, H. R., and Chiou, S. S. (2003). E/NS1 modifications of dengue 2 virus after serial passages in mammalian and/or mosquito cells. *Intervirology* **46**(5)**,** 289-295.

Chen, Y., Maguire, T., Hileman, R. E., Fromm, J. R., Esko, J. D., Linhardt, R. J., and Marks, R. M. (1997). Dengue virus infectivity depends on envelope protein binding to target cell heparan sulfate. *Nature Medicine* **3,** 866-871.

Chen, Y., Maguire, T., and Marks, R. M. (1996). Demonstration of binding of dengue virus envelope protein to target cells. *Journal of Virology* **70**(12)**,** 8765-8772.

Chiou, S.-S., Crill, W. D., Chen, L.-K., and Chang, G.-J. J. (2008). Enzyme-Linked Immunosorbent Assays Using Novel Japanese Encephalitis Virus Antigen Improve the Accuracy of Clinical Diagnosis of Flavivirus Infections. *Clinical and Vaccine Immunology.* **15**(5)**,** 825-835.

Chiu, W.-W., Kinney, R. M., and Dreher, T. W. (2005). Control of Translation by the 5'- and 3'-Terminal Regions of the Dengue Virus Genome. *Journal of Virology* **79**(13)**,** 8303-8315.

Clyde, K., and Harris, E. (2006). RNA Secondary Structure in the Coding Region of Dengue Virus Type 2 Directs Translation Start Codon Selection and Is Required for Viral Replication. *Journal of Virology* **80**(5)**,** 2170-2182.

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

Cologna, R., and Rico-Hesse, R. (2003). American Genotype Structures Decrease Dengue Virus Output from Human Monocytes and Dendritic Cells. *Journal of Virology* **77**(7)**,** 3929-3938.

Cooper, P. D., Steiner-Pryor, A., Scotti, P. D., and Delong, D. (1974). On the Nature of Poliovirus Genetic Recombinants. *Journal of General Virology* **23**(1)**,** 41-49.

Craig, S., Thu, H. M., Lowry, K., Wang, X.-f., Holmes, E. C., and Aaskov, J. (2003). Diverse Dengue Type 2 Virus Populations Contain Recombinant and Both Parental Viruses in a Single Mosquito Host. *Journal of Virology* **77**(7)**,** 4463-4467.

Crill, W. D., and Chang, G.-J. J. (2004). Localization and Characterization of Flavivirus Envelope Glycoprotein Cross-Reactive Epitopes. *Journal of Virology* **78**(24)**,** 13975-13986.

Crill, W. D., and Roehrig, J. T. (2001). Monoclonal Antibodies That Bind to Domain III of Dengue Virus E Glycoprotein Are the Most Efficient Blockers of Virus Adsorption to Vero Cells. *Journal of Virology* **75**(16)**,** 7769-7773.

Crotty, S., Cameron, C. E., and Andino, R. (2001). RNA virus error catastrophe: Direct molecular test by using ribavirin. *Proceedings of the National Academy of Sciences* **98**(12)**,** 6895-6900.

Deen, J. L. (2005). Dengue, dengue haemorrhagic fever and dengue shock syndrome in the context of the integrated management of childhood illness. World Health Organisation, Geneva.

Deen, J. L., Harris, E., Wills, B., Balmaseda, A., Hammond, S. N., Rocha, C., Dung, N. M., Hung, N. T., Hien, T. T., and Farrar, J. J. (2006). The WHO dengue case classification and case definitions: time for a reassessment. *The Lancet* **368**(9530)**,** 170-173.

DeFilippis, V. R., and Villarreal, L. P. (2001). Virus Evolution. 4th ed. *In* "*Fields Virology*" (D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus, Eds.), Vol. 1. 2 vols. Lippincott Williams & Wilkins, Philadelphia, Pa, Philadelphia.

Delenda, C., Staropoli, I., Frenkiel, M.-P., Cabanie, L., and Deubel, V. (1994). Analysis of C-terminally truncated dengue 2 and dengue 3 virus envelope glycoproteins: processing in insect cells and immunogenic properties in mice. *Journal of General Virology* **75**(7)**,** 1569-1578.

Despres, P., Frenkiel, M.-P., Ceccaldi, P.-E., Duarte Dos Santos, C., and Deubel, V. (1998). Apoptosis in the Mouse Central Nervous System in Response to Infection with Mouse-Neurovirulent Dengue Viruses. *Journal of Virology* **72**(1)**,** 823-829.

Deubel, V., Kinney, R. M., and Trent, D. W. (1986). Nucleotide sequence and deduced amino acid sequence of the structural proteins of dengue type 2 virus, Jamaica genotype. *Virology* **155**(2)**,** 365-377.

Domingo, E. (1998). Quasispecies and the implications for virus persistence and escape. *Clinical and Diagnostic Virology* **10**(2-3)**,** 97-101.

Domingo, E., Baranowski, E., Escarmís, C., and Sobrino, F. (2002). Foot-and-mouth disease virus. *Comparative Immunology, Microbiology and Infectious Diseases* **25**(5-6)**,** 297-308.

Drake, J. W. (1993). Rates of Spontaneous Mutation Among RNA Viruses. *Proceedings of the National Academy of Sciences* **90**(9)**,** 4171-4175.

Duarte dos Santos, C. N., Frenkiel, M.-P., Courageot, M.-P., Fernando S. Rocha, C., Vazeille-Falcoz, M.-C., Wien, M. W., Rey, F. A., Deubel, V., and Desprès, P. (2000). Determinants in the Envelope E Protein and Viral RNA Helicase NS3 That Influence the Induction of Apoptosis in Response to Infection with Dengue Type 1 Virus. *Virology* **274**(2)**,** 292-308.

Dulbecco, R., and Ginsberg, H. S. (1988). The nature of viruses. 2nd ed. *In* "Virology", pp. 22-25. J.B. Lippincott company, Philadelphia.

Edgil, D., Polacek, C., and Harris, E. (2006). Dengue Virus Utilizes a Novel Strategy for Translation Initiation When Cap-Dependent Translation Is Inhibited. *Journal of Virology* **80**(6)**,** 2976-2986.

Feinberg, H., Mitchell, D., Drickamer, K., and Weis, W. (2001). Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science* **294,** 2163-2166.

Garcia-Arriaza, J., Domingo, E., and Briones, C. (2007). Characterization of minority subpopulations in the mutant spectrum of HIV-1 quasispecies by successive specific amplifications. *Virus Research* **129,** 123-134.

Georgescu, M. M., Delpeyroux, F., Tardy-Panit, M., Balanant, J., Combiescu, M., Combiescu, A. A., Guillot, S., and Crainic, R. (1994). High diversity of poliovirus strains isolated from the central nervous system from patients with vaccine-associated paralytic poliomyelitis. *Journal of Virology* **68**(12)**,** 8089-8101.

Goncalvez, A. P., Escalante, A. A., Pujol, F. H., Ludert, J. E., Tovar, D., Salas, R. A., and Liprandi, F. (2002). Diversity and evolution of the envelope gene of dengue virus type 1. *Virology* **303**(1)**,** 110 - 119.

Grandadam, M., Tebbal, S., Caron, M., Siriwardana, M., Larouze, B., Koeck, J. L., Buisson, Y., Enouf, V., and Nicand, E. (2004). Evidence for hepatitis E virus quasispecies. *Journal of General Virology* **85,** 3189-3194.

Gualano, R. C., Pryor, M. J., Cauchi, M. R., Wright, P. J., and Davidson, A. D. (1998). Identification of a major determinant of mouse neurovirulence of dengue virus type 2 using stably cloned genomic-length cDNA. *Journal of General Virology* **79**(3)**,** 437-446.

Gubler, D. J. (1996). Serologic diagnosis of dengue/dengue haemorrhagic fever. *Dengue Bulletins* **20,** 20-23.

Gubler, D. J., Reed, D., Rosen, L., and Hitchcock, J. C., Jr. (1978). Epidemiologic, Clinical, and Virologic Observations on Dengue in the Kingdom of Tonga. *The American Journal of Tropical Medicine and Hygiene* **27**(3)**,** 581-589.

Gubler, D. J., Suharyono, W., Lubis, I., Eram, S., and Gunarso, S. (1981). Epidemic Dengue 3 in Central Java, Associated with Low Viremia in Man. *The American Journal of Tropical Medicine and Hygiene* **30**(5)**,** 1094-1099.

Guirakhoo, F., Zhang, Z., Myers, G., Johnson, B. W., Pugachev, K., Nichols, R., Brown, N., Levenbook, I., Draper, K., Cyrek, S., Lang, J., Fournier, C., Barrere, B., Delagrave, S., and Monath, T. P. (2004). A Single Amino Acid Substitution in the Envelope Protein of Chimeric Yellow Fever-Dengue 1 Vaccine Virus Reduces Neurovirulence for Suckling Mice and Viremia/Viscerotropism for Monkeys. *Journal of Virology* **78**(18)**,** 9998-10008.

Guzmán, M. G., Kouri, G., Valdes, L., Bravo, J., Alvarez, M., Vazques, S., Delgado, I., and Halstead, S. B. (2000). Epidemiologic Studies on Dengue in Santiago de Cuba, 1997. *American Journal of Epidemiology* **152**(9)**,** 793-799.

Hacker, K., White, L., and de Silva, A. M. (2009). N-Linked glycans on dengue viruses grown in mammalian and insect cells. *Journal of General Virology* **90**(9)**,** 2097-2106.

Hahn, C. S., J, M. D., J, H. S., and C, M. R. (1987). Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. *Proceedings of the National Academy of Sciences of the United States of America* **84**(7)**,** 2019-2023.

Halstead, S. B. (1988). Pathogenesis of dengue: challenges to molecular biology. *Science* **239**(4839)**,** 476-481.

Halstead, S. B., and O'Rourke, E. J. (1977). Dengue viruses and mononuclear phagocytes. I. Infection enhancement by non-neutralizing antibody. *The Journal of Experimental Medicine* **146**(1)**,** 201-217.

Halstead, S. B., Streit, T. G., Lafontant, J. G., Putvatana, R., Russell, K., Sun, W., Kanesa-Thasan, N., Hayes, C. G., and Watts, D. M. (2001). Haiti: absence of dengue hemorrhagic fever despite hyperendemic dengue virus transmission. *The American Journal of Tropical Medicine and Hygiene* **65**(3)**,** 180-183.

Hasegawa, H., Yoshida, M., Shiosaka, T., Fujita, S., and Kobayashi, Y. (1992). Mutations in the envelope protein of japanese encephalitis virus affect entry into cultured cells and virulence in mice. *Virology* **191**(1)**,** 158-165.

Heinz, F. X., and Maramorosch, F. A. M. a. A. J. S. (1986). Epitope Mapping of Flavivirus Glycoproteins. *In* "Advances in Virus Research", Vol. Volume 31, pp. 103-168. Academic Press.

Hitchman, R. B., Possee, R. D., Siaterli, E., Richards, K. S., Clayton, A. J., Bird, L. E., Owens, R. J., Carpentier, D. C., King, F. L., Danquah, J. O., Spink, K. G., and King, L. A. (2010). Improved expression of secreted and membrane-targeted proteins in insect cells. *Biotechnology and Biochemistry* **56**(3)**,** 85-93.

Ho, L.-J., Hung, L.-F., Weng, C.-Y., Wu, W.-L., Chou, P., Lin, Y.-L., Chang, D.-M., Tai, T.-Y., and Lai, J.-H. (2005). Dengue Virus Type 2 Antagonizes IFN-Î± but Not IFN-Î³ Antiviral Effect via Down-Regulating Tyk2-STAT Signaling in the Human Dendritic Cell. *The Journal of Immunology* **174**(12)**,** 8163-8172.

Ho, L.-J., Wang, J.-J., Shaio, M.-F., Kao, C.-L., Chang, D.-M., Han, S.-W., and Lai, J.-H. (2001). Infection of Human Dendritic Cells by Dengue Virus Causes Cell Maturation and Cytokine Production. *The Journal of Immunology* **166**(3)**,** 1499-1506.

Hogrefe, H. H., Hansen, C. J., Scott, B. R., and Nielson, K. B. (2002). Archaeal dUTPase enhances PCR amplifications with archaeal DNA polymerases by preventing dUTP incorporation. *Proceedings of the National Academy of Sciences of the United States of America* **99**(2)**,** 596-601.

Holmes, D. A., Purdy, D. E., Chao, D.-Y., Noga, A. J., and Chang, G.-J. J. (2005). Comparative Analysis of Immunoglobulin M (IgM) Capture Enzyme-Linked Immunosorbent Assay Using Virus-Like Particles or Virus-Infected Mouse Brain Antigens To Detect IgM Antibody in Sera from Patients with Evident Flaviviral Infections. *Journal of Clinical Microbiology* **43**(7)**,** 3227-3236.

Holmes, E. C., and Twiddy, S. S. (2003). The origin, emergence and evolutionary genetics of dengue virus. *Infection, Genetics and Evolution* **3,** 19-28.

Holmes, E. C., Worobey, M., and Rambaut, A. (1999). Phylogenetic evidence for recombination in dengue virus. *Molecular Biology and Evolution* **16**(3)**,** 405-409.

Holzmann, H., Heinz, F. X., Mandl, C. W., Guirakhoo, F., and Kunz, C. (1990). A single amino acid substitution in envelope protein E of tick-borne encephalitis virus leads to attenuation in the mouse model. *Journal of Virology* **64**(10)**,** 5156-5159.

Hopkins, R. F., and Esposito, D. (2009). A rapid method for titrating baculovirus stocks using the Sf-9 Easy Titer cell line. *BioTechniques* **47**(3)**,** 785-788.

Hunt, A. R., Cropp, C. B., and Chang, G.-J. J. (2001). A recombinant particulate antigen of Japanese encephalitis virus produced in stably-transformed cells is an effective noninfectious antigen and subunit immunogen. *Journal of Virological Methods* **97**(1-2)**,** 133-149.

Jaaskelainen, A., Han, X., Niedrig, M., Vaheri, A., and Vapalahti, O. (2003). Diagnosis of Tick-Borne Encephalitis by a μ-Capture Immunoglobulin M-Enzyme Immunoassay Based on Secreted Recombinant Antigen Produced in Insect Cells. *Journal of Clinical Microbiology* **41**(9)**,** 4336-4342.

Jarvis, D. L. (2003). Developing baculovirus-insect cell expression systems for humanized recombinant glycoprotein production. *Virology* **310**(1)**,** 1-7.

Jarvis, D. L., and Garcia, A. (1994). Biosynthesis and Processing of the Autographa californica Nuclear Polyhedrosis Virus GP64 Protein. *Virology* **205**(1)**,** 300-313.

Jiang, W. R., Lowe, A., Higgs, S., Reid, H., and Gould, E. A. (1993). Single Amino Acid Codon Changes Detected in Louping Ill Virus Antibody-Resistant Mutants with Reduced Neurovirulence. *Journal of General Virology* **74**(5)**,** 931-935.

Kadlec, J., Loureiro, S., Abrescia, N. G. A., Stuart, D. I., and Jones, I. M. (2008). The postfusion structure of baculovirus GP64 supports a unified view of viral fusion machines. *Nature Structural and Molecular Biology* **15**(10)**,** 1024-1030.

Kanakaratne, N., Wahala, M. P. B. W., Messer, W. B., Tissera, H. A., Shahani, A., Abeysinghe, N., De Silva, A. M., and Gunasekera, M. (2009). Severe Dengue Epidemics in Sri Lanka, 2003–2006. *Emerging Infectious Diseases* **15**(2)**,** 192-199.

Kaufmann, B., and Rossmann, M. G. (2011). Molecular mechanisms involved in the early steps of flavivirus cell entry. *Microbes and Infection* **13**(1)**,** 1-9.

Kawano, H., Rostapshov, V., Rosen, L., and Lai, C. J. (1993). Genetic determinants of dengue type 4 virus neurovirulence for mice. *Journal of Virology* **67**(11)**,** 6567-6575.

Konishi, E., and Fujii, A. (2002). Dengue type 2 virus subviral extracellular particles produced by a stably transfected mammalian cell line and their evaluation for a subunit vaccine. *Vaccine* **20**(7-8)**,** 1058-1067.

Konishi, E., Pincus, S., Fonseca, B. A. L., Shope, R. E., Paoletti, E., and Mason, P. W. (1991). Comparison of protective immunity elicited by recombinant vaccinia viruses that synthesize E or NS1 of Japanese encephalitis virus. *Virology* **185**(1)**,** 401-410.

Konishi, E., Pincus, S., Paoletti, E., Shope, R. E., Burrage, T., and Mason, P. W. (1992). Mice immunized with a subviral particle containing the japanese encephalitis virus prM/M and E proteins are protected from lethal JEV infection. *Virology* **188**(2)**,** 714-720.

Kuhn, R. J., Zhang, W., Rossmann, M. G., Pletnev, S. V., Corver, J., Lenches, E., Jones, C. T., Mukhopadhyay, S., Chipman, P. R., Strauss, E. G., Baker, T. S., and Strauss, J. H. (2002). Structure of Dengue Virus: Implications for Flavivirus Organization, Maturation, and Fusion. *Cell* **108**(5)**,** 717-725.

Kunkel, T. A., Patel, S. S., and Johnson, K. A. (1994). Error-prone replication of repeated DNA sequences by T7 DNA polymerase in the absence of its processivity subunit. *Proceedings of the National Academy of Sciences of the United States of America* **91**(15)**,** 6830–6834.

Kuroda, M., Katano, H., Nakajima, N., Tobiume, M., Ainai, A., Sekizuka, T., Hasegawa, H., Tashiro, M., Sasaki, Y., Arakawa, Y., Hata, S., Watanabe, M., and Sata, T. (2010). Characterization of quasispecies of pandemic 2009 influenza A virus (A/H1N1/2009) by *de novo* sequencing using a next-generation DNA sequencer. *PLoS One* **5**(4)**,** e10256.

Lanciotti, R. S., Calisher, C. H., Gubler, D. J., Chang, G. J., and Vorndam, A. V. (1992). Rapid detection and typing of dengue viruses from clinical samples by using reverse transcriptase-polymerase chain reaction. *Journal of Clinical Microbiology* **30**(3)**,** 545-551.

Lanciotti, R. S., Gubler, D. J., and Trent, D. W. (1997). Molecular evolution and phylogeny of dengue-4 viruses. *Journal of General Virology* **78**(9)**,** 2279-2284.

Lanciotti, R. S., Lewis, J. G., Gubler, D. J., and Trent, D. W. (1994). Molecular evolution and epidemiology of dengue-3 viruses. *Journal of General Virology* **75**(1)**,** 65-75.

Lee, E., Pavy, M., Young, N., Freeman, C., and Lobigs, M. (2006). Antiviral effect of the heparan sulfate mimetic, PI-88, against dengue and encephalitic flaviviruses. *Antiviral Research* **69**(1)**,** 31-38.

Lee, E., Weir, R. C., and Dalgarno, L. (1997). Changes in the Dengue Virus Major Envelope Protein on Passaging and Their Localization on the Three-Dimensional Structure of the Protein. *Virology* **232**(2)**,** 281-290.

Leitmeyer, K. C., Vaughn, D. W., Watts, D. M., Salas, R., Villalobos de Chacon, I., Ramos, C., and Rico-Hesse, R. (1999). Dengue virus structural differences that correlate with pathogenesis. *Journal of Virology* **73**(6)**,** 4738-4747.

Libraty, D. H., Young, P. R., Pickering, D., Endy, T. P., Kalayanarooj, S., Green, S., Vaughn, D. W., Nisalak, A., Ennis, F. A., and Rothman, A. L. (2002). High Circulating Levels of the Dengue Virus Nonstructural Protein NS1 Early in Dengue Illness Correlate with the Development of Dengue Hemorrhagic Fever. *The Journal of Infectious Diseases* **186**(8)**,** 1165-1168.

Lin, B., Parrish, C. R., Murray, J. M., and Wright, P. J. (1994). Localization of a Neutralizing Epitope on the Envelope Protein of Dengue Virus Type 2. *Virology* **202**(2)**,** 885-890.

Lin, S.-R., Hsieh, S.-C., Yueh, Y.-Y., Lin, T.-H., Chao, D.-Y., Chen, W.-J., King, C.-C., and Wang, W.-K. (2004). Study of Sequence Variation of Dengue Type 3 Virus in Naturally Infected Mosquitoes and Human Hosts: Implications for Transmission and Evolution. *Journal of Virology* **78**(22)**,** 12717-12721.

Lin, Y.-J., and Wu, S.-C. (2005). Histidine at Residue 99 and the Transmembrane Region of the Precursor Membrane prM Protein Are Important for the prM-E Heterodimeric Complex Formation of Japanese Encephalitis Virus. *Journal of Virology* **79**(13)**,** 8535-8544.

Lindenbach, B. D., and Rice, C. M. (2001). Flaviviridae: the viruses and their replication. 4th ed. *In* "*Fields Virology*" (D. M. Knipe, and P. M. Howley, Eds.), Vol. 1, pp. 991-1041. 2 vols. Lippincott Williams & Wilkins, Philadelphia, Pa.

Liu, W., Jiang, H., Zhou, J., Yang, X., Tang, Y., Fang, D., and Jiang, L. (2010). Recombinant dengue virus-like particles from *Pichia pastoris*: efficient production and immunological properties. *Virus Genes* **40**(1)**,** 53-59.

Lobigs, M., Usha, R., Nestorowicz, A., Marshall, I. D., Weir, R. C., and Dalgarno, L. (1990). Host cell selection of Murray Valley encephalitis virus variants altered at an RGD sequence in the envelope protein and in mouse virulence. *Virology* **176**(2)**,** 587-595.

Mandl, C. W., Guirakhoo, F., Holzmann, H., Heinz, F. X., and Kunz, C. (1989). Antigenic structure of the flavivirus envelope protein E at the molecular level, using tick-borne encephalitis virus as a model. *Journal of Virology* **63**(2)**,** 564-571.

Mandl, C. W., Heinz, F. X., and Kunz, C. (1988). Sequence of the structural proteins of tick-borne encephalitis virus (Western subtype) and comparative analysis with other flaviviruses. *Virology* **166**(1)**,** 197-205.

Marks, R. M., Lu, H., Sundaresan, R., Toida, T., Suzuki, A., Imanari, T., Hernaįiz, M. a. J., and Linhardt, R. J. (2001). Probing the Interaction of Dengue Virus Envelope Protein with Heparin: Assessment of Glycosaminoglycan-Derived Inhibitors. *Journal of Medicinal Chemistry* **44**(13)**,** 2178-2187.

Martell, M., Esteban, J. I., Quer, J., Genesca, J., Weiner, A., Esteban, R., Guardia, J., and Gomez, J. (1992). Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *Journal of Virology* **66**(5)**,** 3225-3229.

Martinez-Barragan, J. d. J., and del Angel, R. M. (2001). Identification of a Putative Coreceptor on Vero Cells That Participates in Dengue 4 Virus Infection. *Journal of Virology* **75**(17)**,** 7818-7827.

Mason, P. W., McAda, P. C., Mason, T. L., and Fournier, M. J. (1987). Sequence of the dengue-1 virus genome in the region encoding the three structural proteins and the major nonstructural protein NS1. *Virology* **161**(1)**,** 262-267.

Mason, P. W., Pincus, S., Fournier, M. J., Mason, T. L., Shope, R. E., and Paoletti, E. (1991). Japanese encephalitis virus-vaccinia recombinants produce particulate forms of the structural membrane proteins and induce high levels of protection against lethal JEV infection. *Virology* **180**(1)**,** 294-305.

Messer, W., Gubler, D. J., Harris, E., Sivananthan, K., and de Silva, A. (2003a). Emergence and Global Spread of a Dengue Serotype 3, Subtype III Virus. *Emerging Infectious Diseases* **9**(7)**,** 800-809.

Messer, W. B., Gubler, D. J., Harris, E., Sivananthan, K., and de Silva, A. M. (2003b). Emergence and global spread of a dengue serotype 3, subtype III virus. *Emerging Infectious Diseases* **9**(7)**,** 800 - 809.

Messer, W. B., Vitarana, U. T., Sivananthan, K., Elvtigala, J., Preethimala, L. D., Ramesh, R., Withana, N., Gubler, D. J., and De Silva, A. M. (2002). Epidemiology of dengue in Sri Lanka before and after the emergence of epidemic dengue hemorrhagic fever. *The American Journal of Tropical Medicine and Hygiene* **66**(6)**,** 765-773.

Miller, J. L., deWet, B. J. M., Martinez-Pomares, L., Radcliffe, C. M., Dwek, R. A., Rudd, P. M., and Gordon, S. (2008). The Mannose Receptor Mediates Dengue Virus Infection of Macrophages. *PLoS Pathogens* **4**(2)**,** e17.

Modis, Y., Ogata, S., Clements, D., and Harrison, S. C. (2003). A ligand-binding pocket in the dengue virus envelope glycoprotein. *Proceedings of the National Academy of Sciences* **100**(12)**,** 6986-6991.

Modis, Y., Ogata, S., Clements, D., and Harrison, S. C. (2004). Structure of the dengue virus envelope protein after membrane fusion. *Nature* **427**(6972)**,** 313-319.

Monath, T. P., Arroyo, J., Levenbook, I., Zhang, Z.-X., Catalan, J., Draper, K., and Guirakhoo, F. (2002). Single Mutation in the Flavivirus Envelope Protein Hinge Region Increases Neurovirulence for Mice and Monkeys but Decreases Viscerotropism for Monkeys: Relevance to Development and Safety Testing of Live, Attenuated Vaccines. *Journal of Virology* **76**(4)**,** 1932-1943.

Mondotte, J. A., Lozach, P.-Y., Amara, A., and Gamarnik, A. V. (2007). Essential Role of Dengue Virus Envelope Protein N Glycosylation at Asparagine-67 during Viral Propagation. *Journal of Virology* **81**(13)**,** 7136-7148.

Mukhopadhyay, S., Kuhn, R. J., and Rossmann, M. G. (2005). A structural perspective of the flavivirus life cycle. *Nature Reviews Microbiology* **3,** 13-22.

Munoz-Jordan, J. L., Laurent-Rolle, M., Ashour, J., Martinez-Sobrido, L., Ashok, M., Lipkin, W. I., and Garcia-Sastre, A. (2005). Inhibition of Alpha/Beta Interferon Signaling by the NS4B Protein of Flaviviruses. *Journal of Virology* **79**(13)**,** 8004-8013.

Navarro-Sanchez, E., Altmeyer, R., Amara, A., Schwartz, O., Fieschi, F., Virelizier, J.-L., Arenzana-Seisdedos, F., and Desprès, 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**(7)**,** 723-728.

Nitayaphan, S., Grant, J. A., Chang, G.-J. J., and Trent, D. W. (1990). Nucleotide sequence of the virulent SA-14 strain of Japanese encephalitis virus and its attenuated vaccine derivative, SA-14-14-2. *Virology* **177**(2)**,** 541-552.

Nowak, T., and Wengler, G. (1987). Analysis of Disulphides Present in the Membrane Proteins of the West Nile Flavivirus. *Virology* **156,** 127-137.

Palucka, A. K. (2000). Dengue virus and dendritic cells. *Nature Medicine* **6**(7)**,** 816-820.

Petersen LR, and Roehrig, JT. (2001). West Nile virus: a reemerging global pathogen. *Emerging Infectious Diseases* **7**(4)**,** 611-614.

Pokidysheva, E., Zhang, Y., Battisti, A. J., Bator-Kelly, C. M., Chipman, P. R., Xiao, C., Gregorio, G. G., Hendrickson, W. A., Kuhn, R. J., and Rossmann, M. G. (2006). Cryo-EM Reconstruction of Dengue Virus in Complex with the Carbohydrate Recognition Domain of DC-SIGN. *Cell* **124**(3)**,** 485-493.

Pryor, M. J., Carr, J. M., Hocking, H., Davidson, A. D., Li, P., and Wright, P. J. (2001). Replication of dengue virus type 2 in human monocyte-derived macrophages: comparisons of isolates and recombinant viruses with substitutions at amino acid 390 in the envelope glycoprotein. *The American Journal of Tropical Medicine and Hygiene* **65**(5)**,** 427-34.

Pu, S.-Y., Wu, R.-H., Yang, C.-C., Jao, T.-M., Tsai, M.-H., Wang, J.-C., Lin, H.-M., Chao, Y.-S., and Yueh, A. (2011). Successful Propagation of Flavivirus Infectious cDNAs by a Novel Method to Reduce the Cryptic Bacterial Promoter Activity of Virus Genomes. *Journal of Virology***,** JVI.01986-10.

Purdy, D. E., Noga, A. J., and Chang, G.-J. J. (2004). Noninfectious Recombinant Antigen for Detection of St. Louis Encephalitis Virus-Specific Antibodies in Serum by Enzyme-Linked Immunosorbent Assay. *Journal of Clinical Microbiology* **42**(10)**,** 4709-4717.

Rademacher, C., Guiard, J., Kitov, P. I., Fiege, B., Dalton, K. P., Parra, F., Bundle, D. R., and Peters, T. (2011). Targeting Norovirus Infection—Multivalent Entry Inhibitor Design Based on NMR Experiments. *Chemistry – A European Journal* **17**(27)**,** 7442-7453.

Rey, F. A., Heinz, F. X., Mandl, C. W., Kunz, C., and Harrison, S. C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2A resolution. *Nature* **375,** 291-298.

Rice, C. M., Lenches, E. M., Eddy, S. R., Shin, S. J., Sheets, R., and Strauss, J. H. (1985). Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution *Science* **229,** 726-733.

Rico-Hesse, R. (1990). Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* **174**(2)**,** 479-493.

Rico-Hesse, R., Harrison, L. M., Salas, R. A., Tovar, D., Nisalak, A., Ramos, C., Boshell, J., de Mesa, M. T., Nogueira, R. M., and da Rosa, A. T. (1997a). Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. *Virology* **230**(2)**,** 244 - 251.

Rico-Hesse, R., Harrison, L. M., Salas, R. A., Tovar, D., Nisalak, A., Ramos, C., Boshell, J., de Mesa, M. T. R., Nogueira, R. M. R., and Rosa, A. T. d. (1997b). Origins of Dengue Type 2 Viruses Associated with Increased Pathogenicity in the Americas. *Virology* **230**(2)**,** 244-251.

Roberson, J. A., Crill, W. D., and Chang, G.-J. J. (2007). Differentiation of West Nile and St. Louis Encephalitis Virus Infections by Use of Noninfectious Virus-Like Particles with Reduced Cross-Reactivity. *Journal of Clinical Microbiology* **45**(10)**,** 3167-3174.

Roehrig, J. (2003). "Antigenic Structure of Flavivirus Proteins." *Advances in Virus Research*, 59.

Roehrig, J. T., Bolin, R. A., and Kelly, R. G. (1998). Monoclonal Antibody Mapping of the Envelope Glycoprotein of the Dengue 2 Virus, Jamaica. *Virology* **246**(2)**,** 317-328.

Rosen, L. (1977). The Emperor's New Clothes Revisited, or Reflections on the Pathogenesis of Dengue Hemorrhagic Fever. *The American Journal of Tropical Medicine and Hygiene* **26**(3)**,** 337-343.

Rozera, G., Abbate, I., Bruselles, A., Vlassi, C., D'Offizi, G., Narciso, P., Chillemi, G., Prosperi, M., Ippolito, G., and Capobianchi, M. (2009). Massively parallel pyrosequencing highlights minority variants in the HIV-1 env quasispecies deriving from lymphomonocyte sub-populations. *Retrovirology* **6**(1)**,** 15.

Russell, P. K., Brandt, W. E., and Dalrymple, J. M. (1980). Chemical and antigenic structure of flaviviruses. *In* "*The togaviruses*" (R. W. Schlesinger, Ed.), pp. 503–529. Academic Press, New York.

References

Sallie, R. (2005). Replicative Homeostasis: A fundamental mechanism mediating selective viral replication and escape mutation. *Virology Journal* **2,** 10.

Sanchez, I. J., and Ruiz, B. H. (1996). A single nucleotide change in the E protein gene of dengue virus 2 Mexican strain affects neurovirulence in mice. *Journal of General Virology* **77**(10)**,** 2541-2545.

Schalich, J., Allison, S. L., Stiasny, K., Mandl, C. W., Kunz, C., and Heinz, F. X. (1996). Recombinant subviral particles from tick-borne encephalitis virus are fusogenic and provide a model system for studying flavivirus envelope glycoprotein functions. *Journal of Virology* **70**(7)**,** 4549-4557.

Shiu, S. Y. W., Morikawa, S., Buckley, A., Higgs, S., Karunakarannair, V., Blachere, C., and Gould, E. A. (1991). 17D yellow fever vaccine virus envelope protein expressed by recombinant baculovirus is antigenically indistinguishable from authentic viral protein. *Journal of General Virology* **72**(6)**,** 1451-1454.

Shu, P.-Y., and Huang, J.-H. (2004). Current advances in dengue diagnosis. *Clinical Diagnostic Laboratory Immunology* **11**(4)**,** 642-650.

Smith, D. B., McAllister, J., Casino, C., and Simmonds, P. (1997). Virus 'quasispecies': making a mountain out of a molehill? *Journal of General Virology* **78**(7)**,** 1511-1519.

Smith, G. E., and Summers, M. D. (1987). A manual of methods for baculovirus vectors and insect cell culture procedures, Vol. 1555. Texas Agricultural Experiment Station Bulletin.

Solomon, T., Dung, N. M., Vaughn, D. W., Kneen, R., Thao, L. T. T., Raengsakulrach, B., Loan, H. T., Day, N. P. J., Farrar, J., Myint, K. S. A., Warrell, M. J., James, W. S., Nisalak, A., and White, N. J. (2000). Neurological manifestations of dengue infection. *The Lancet* **355**(9209)**,** 1053-1059.

Sugrue, R. J., Fu, J., Howe, J., and Chan, Y. C. (1997). Expression of the dengue virus structural proteins in *Pichia pastoris* leads to the generation of virus-like particles. *Journal of General Virology* **78**(8)**,** 1861-1866.

Sumiyoshi, H., Hoke, C. H., and Trent, D. W. (1992). Infectious Japanese encephalitis virus RNA can be synthesized from in vitro-ligated cDNA templates. *Journal of Virology* **66**(9)**,** 5425-5431.

Swinscow, T. D. V. (1996). "Statistics at square one." 9th ed. (M. J. Campbell, Ed.) BMJ Publishing Group, London.

Tajima, S., Nukui, Y., Takasaki, T., and Kurane, I. (2007). Characterization of the variable region in the 3' non-translated region of dengue type 1 virus. *Journal of General Virology* **88**(8)**,** 2214-2222.

Takhampunya, R., Palmer, D. R., McClain, S., Barvir, D. A., Lynch, J., Jarman, R. G., Thomas, S., Gibbons, R. V., Burgess, T. H., Sun, P., Kamau, E., Putnak, R., and Zhang, C. (2009). Phenotypic analysis of dengue virus isolates associated with dengue fever and dengue hemorrhagic fever for cellular attachment, replication and interferon signaling ability. *Virus Research* **145**(1)**,** 31-38.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. . *Molecular Biology and Evolution* **24,** 1596-1599.

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., and Marovich, M. A. (2003). DC-SIGN (CD209) Mediates Dengue Virus Infection of Human Dendritic Cells. *Journal of Experimental Medicine* **197**(7)**,** 823-829.

Thomas, L., Verlaeten, O., Cabie, A., Kaidomar, S., Moravie, V., Martial, J., Najioullah, F., Plumelle, Y., Fonteau, C., Dussart, P., and Cesaire, R. (2008). Influence of the Dengue Serotype, Previous Dengue Infection, and Plasma Viral Load on Clinical Presentation and Outcome During a Dengue-2 and Dengue-4 Co-Epidemic. *The American Journal of Tropical Medicine and Hygiene* **78**(6)**,** 990-998.

Thullier, P., Demangel, C., Bedouelle, H., Megret, F., Jouan, A., Deubel, V., Mazie, J.-C., and Lafaye, P. (2001). Mapping of a dengue virus neutralizing epitope critical for the infectivity of all serotypes: insight into the neutralization mechanism. *Journal of General Virology* **82**(8)**,** 1885-1892.

Tindall, K. R., and Kunkel, T. A. (1988). Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. *Biochemistry* **27**(16)**,** 6008-6013.

Tolou, H. J. G., Couissinier-Paris, P., Durand, J. P., Mercier, V., de Pina, J. J., de Micco, P., Billoir, F., Charrel, R. N., and de Lamballerie, X. (2001). Evidence for recombination in natural populations of dengue virus type 1 based on the analysis of complete genome sequences. *Journal of General Virology* **82**(6)**,** 1283-1290.

Twiddy, S. S., Farrar, J. J., Vinh Chau, N., Wills, B., Gould, E. A., Gritsun, T., Lloyd, G., and Holmes, E. C. (2002). Phylogenetic Relationships and Differential Selection Pressures among Genotypes of Dengue-2 Virus. *Virology* **298**(1)**,** 63-72.

Uzcategui, N. Y., Camacho, D., Comach, G., Cuello de Uzcategui, R., Holmes, E. C., and Gould, E. A. (2001). Molecular epidemiology of dengue type 2 virus in Venezuela: evidence for in situ virus evolution and recombination. *Journal of General Virology* **82**(12)**,** 2945-2953.

Valdes, K., Alvarez, M., Pupo, M., Vazquez, S., Rodriguez, R., and Guzman, M. G. (2000). Human Dengue Antibodies against Structural and Nonstructural Proteins. *Clinical Diagnostic Laboratory Immunology* **7**(5)**,** 856-857.

Vasilakis, N., Weaver, S. C., and Maramorosch, A. J. S. a. F. A. M. (2008). The History and Evolution of Human Dengue Emergence. *Advances in Virus Research* **Volume 72,** 1-76.

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., and Nisalak, A. (2000). Dengue Viremia Titer, Antibody Response Pattern, and Virus Serotype Correlate with Disease Severity. *The Journal of Infectious Diseases* **181**(1)**,** 2-9.

Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E., and Andino, R. (2006). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population *Nature* **439,** 344-348.

Vitarana, T., Jayakuru, W., and Withane, N. (1997). *Historical account of dengue haemorrhagic fever in Sri Lanka*. World Health Organisation, Geneva

Wang, W.-K., Chao, D.-Y., Kao, C.-L., Wu, H.-C., Liu, Y.-C., Li, C.-M., Lin, S.-C., Ho, S.-T., Huang, J.-H., and King, C.-C. (2003a). High Levels of Plasma Dengue Viral Load during Defervescence in Patients with Dengue Hemorrhagic Fever: Implications for Pathogenesis. *Virology* **305**(2)**,** 330-338.

Wang, W.-K., Chao, D.-Y., Lin, S.-R., King, C.-C., and Chang, S.-C. (2003b). Concurrent infections by two dengue virus serotypes among dengue patients in Taiwan *Journal of Microbiology, Immunology and Infection* **36**(2)**,** 89-95.

Wang, W.-K., Chen, H.-L., Yang, C.-F., Hsieh, S.-C., Juan, C.-C., Chang, S.-M., Yu, C.-C., Lin, L.-H., Huang, J.-H., and King, C.-C. (2006). Slower Rates of Clearance of Viral Load and Virus-Containing Immune Complexes in Patients with Dengue Hemorrhagic Fever. *Clinical Infectious Diseases* **43**(8)**,** 1023-1030.

Wang, W.-K., Lin, S.-R., Lee, C.-M., King, C.-C., and Chang, S.-C. (2002a). Dengue type 3 virus in plasma is a population of closely related genomes: quasispecies. *Journal of Virology* **76**(9)**,** 4662-4665.

Wang, W.-K., Sung, T.-L., Lee, C.-N., Lin, T.-Y., and King, C.-C. (2002b). Sequence Diversity of the Capsid Gene and the Nonstructural Gene NS2B of Dengue-3 Virus in Vivo. *Virology* **303**(1)**,** 181-191.

Ward, R., and Davidson, A. D. (2008). Reverse genetics and the study of dengue virus. *Future Medicine* **3**(3)**,** 279-290.

Wenming, P., man, Y., Baochang, F., Yongqiang, D., Tao, J., Hongyuan, D., and Ede, Q. (2005). Simultaneous infection with dengue 2 and 3 viruses in a Chinese patient return from Sri Lanka. *Journal of Clinical Virology* **32**(3)**,** 194-198.

World Health Organisation (1997a). Clinical Diagnosis. 2nd ed. *In* "*Dengue haemorrhagic fever: diagnosis, treatment, prevention and control*", pp. 12-23. World Health Organisation, Geneva.

World Health Organisation (1997b). General Considerations. 2nd ed. *In* "*Dengue haemorrhagic fever: diagnosis, treatment, prevention and control*", pp. 1-11. World Health Organisation, Geneva.

World Health Organisation (1997c). Laboratory Diagnosis. 2nd ed. *In* "*Dengue haemorrhagic fever: diagnosis, treatment, prevention and control*", pp. 34-47. World Health Organisation, Geneva.

World Health Organisation (1999). Recognition of Dengue Fever/Dengue haemorrhagic Fever. *In* "*Guidelines for treatment of dengue fever/dengue haemorrhagic fever in small hospitals*", pp. 2-5. WHO Regional Office for Southeast Asia, New Delhi.

World Health Organisation (2008). Dengue and dengue haemorrhagic fever. *In* "*WHO Factsheet*", Vol. 117.

World Health Organisation (2009). "*Dengue: guidelines for diagnosis, treatment, prevention and control.*" 3rd ed. World Health Organisation, Geneva.

Worobey, M., Rambaut, A., and Holmes, E. C. (1999). Widespread intra-serotype recombination in natural populations of dengue virus. *Proceedings of the National Academy of Sciences* **96**(13)**,** 7352-7357.

Wright, C. F., Morelli, M. J., Thébaud, G., Knowles, N. J., Herzyk, P., Paton, D. T., and King, D. P. (2011). Beyond the consensus: dissecting within-host viral population diversity of foot-and-mouth disease virus by using next-generation genome sequencing. *Journal of Virology* **85**(5)**,** 2266-2275.

Wu, S.-J. L., Grouard-Vogel, G., Sun, W., Mascola, J. R., Brachtel, E., Putvatana, R., Louder, M. K., Filgueira, L., Marovich, M. A., Wong, H. K., Blauvelt, A., Murphy, G. S., Robb, M. L., Innes, B. L., Birx, D. L., Hayes, C. G., and Frankel, S. S. (2000). Human skin Langerhans cells are targets of dengue virus infection. *Nature Medicine* **6**(7)**,** 816-820.

Yamshchikov, V., Mishin, V., and Cominelli, F. (2001). A New Strategy in Design of (+)RNA Virus Infectious Clones Enabling Their Stable Propagation in *E. coli*. *Virology* **281**(2)**,** 272-280.

Yang, D. K., Kweon, C. H., Kim, B. H., Lim, S. I., Kwon, J. H., Kim, S. H., Song, J. Y., and Han, H. R. (2005). Immunogenicity of baculovirus expressed recombinant proteins of Japanese encephalitis virus in mice. *Journal of veterinary science.* **6**(2)**,** 125-133.

Zhao, B., Mackow, E., Buckler-White, A., Markoff, L., Chanock, R. M., Lai, C.-J., and Making, Y. (1986). Cloning full-length dengue type 4 viral DNA sequences: Analysis of genes coding for structural proteins. *Virology* **155**(1)**,** 77-88.

Zhao, R., Chinnawirotpisan, P., Klungthong, C., Zhang, C., and Putnak, R. (2010). Evidence for inter-and intra-genotypic variations in dengue serotype 4 viruses representing predominant and non-predominant genotypes co-circulating in Thailand from 1977 to 2001. *Virus Genes* **41**(1)**,** 5-13.

APPENDICES

Appendix A. Dengue RT-PCR primer alignments

These alignments are included as an electronic version only.

Nucleotide sequence alignments for the regions flanking the dengue virus envelope genes are included, with the NCBI accession number for each sequence shown at the start of each row. These alignments were used to design forward and reverse primers for envelope gene amplification. For each alignment, nucleotides matching the consensus sequence are shown as dots, with mismatches shown in full. The final primer sequences used for dengue virus envelope gene amplification from patient samples are highlighted in the consensus (Majority) sequence.

Appendix B. Gel electrophoresis images of plasmid DNA for each dengue virus quasispecies envelope gene clone.

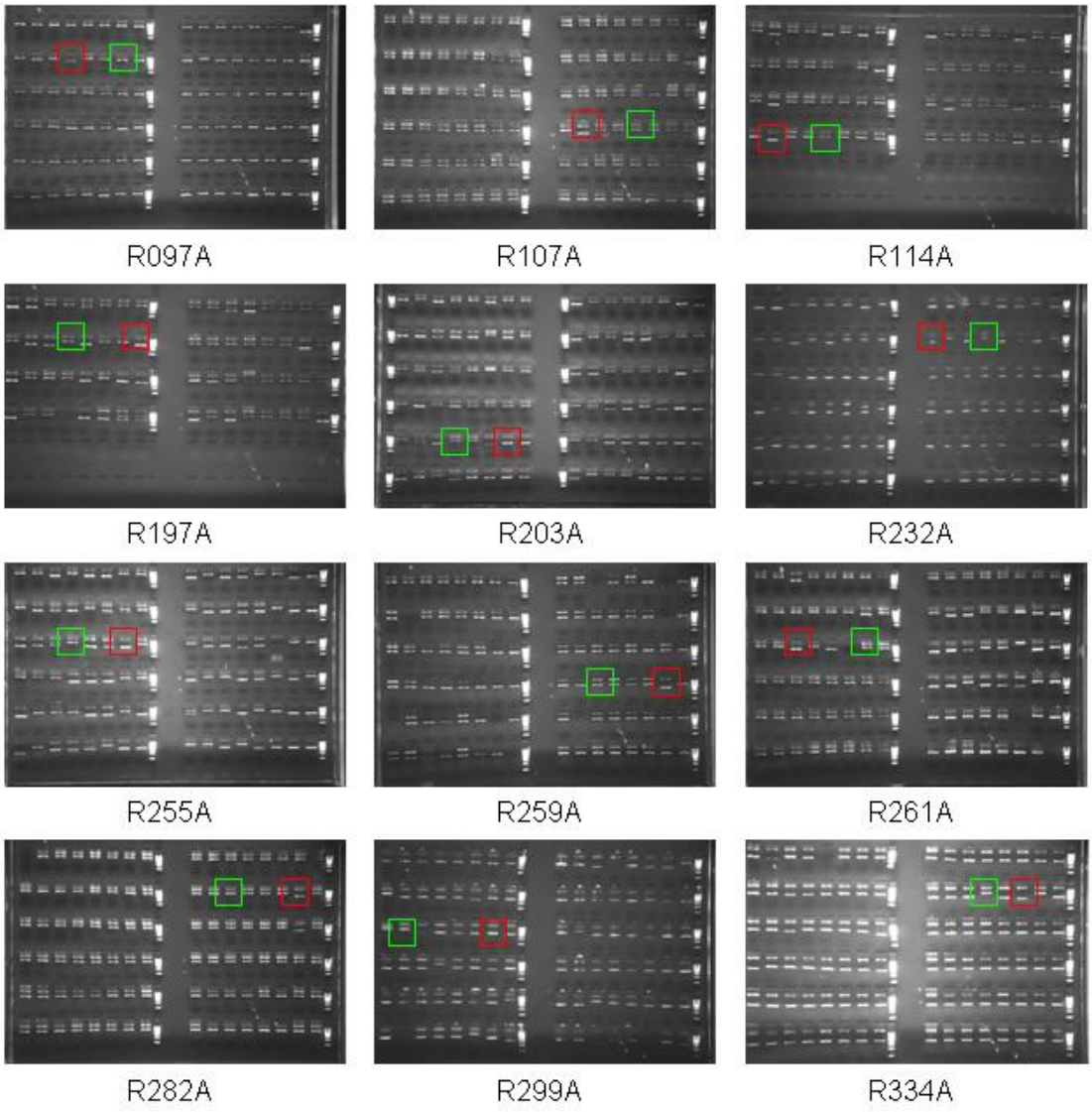


Figure 1. Dengue virus envelope gene clone plasmid DNA gel electrophoresis images.

For each patient sample (R097A, R107A, R114A, R197A, R203A, R232A, R255A, R259A, R261A, R282A, R299A, R334A), each lane consists of eight clones with a 1kb ladder at one end. Ladder band sizes from the bottom upwards are 0.5kb, 1.0kb, 1.5kb, 2.0kb, 3.0kb (brightest band), 4.0kb, 5.0kb, 6.0kb, 8.0kb and 10.0kb. For each clone two bands were observed corresponding to nicked circular plasmid DNA (>10 kb) and supercoiled plasmid DNA either with (4.8 kb; example shown in a green box) or without (3 kb; example shown in a red box) the envelope gene insert.

Appendix C. Quasispecies envelope gene nucleotide sequence alignments

These alignments are included as an electronic version only.

Nucleotide sequence alignments for the cloned dengue virus quasispecies envelope genes are included below for each patient sample (R097A, R107A, R203A, R259A, R261A, R282A). The clone number for each sequence is shown at the start of each row. These alignments were used to identify mutations within the quasispecies populations. For each alignment, nucleotides matching the consensus sequence are shown as dots, with mismatches shown in full. The quasispecies consensus (Majority) sequence is shown along the top of each alignment. Mutations are summarized in Appendix E, tables 1 to 6.

Appendix D. Quasispecies envelope glycoprotein amino acid sequence alignments

These alignments are included as an electronic version only.

Amino acid sequence alignments for the cloned dengue virus quasispecies envelope genes are included below for each patient sample (R097A, R107A, R203A, R259A, R261A, R282A). The clone number for each sequence is shown at the start of each row. These alignments were used to identify amino acid sequence changes resulting from the envelope gene mutations within the dengue virus quasispecies populations that were identified in the alignments in Appendix 3. For each alignment, amino acids matching the consensus sequence are shown as dots, with mismatches shown in full; stop codons are represented with a slash (/). The quasispecies consensus (Majority) sequence is shown along the top of each alignment. Protein sequence changes are summarized in Appendix E, tables 1 to 6.

Appendix E. Summary of dengue virus quasispecies envelope gene mutations

The dengue virus quasispecies envelope gene mutations and amino acid sequence changes identified from the sequence alignments shown in Appendices C and D are summarised for each dengue patient sample (R097A, R107A, R203A, R259A, R261A, R282A). A transition mutation is one where a purine (A or G) is substituted with another purine, or a pyrimidine (C, T or U) with another pyrimidine. In contrast, a transversion mutation is one in which a purine is substituted with a pyrimidine or vice versa. Nonsynonymous mutations lead to protein sequence changes whereas synonymous mutations do not. Conservative protein changes occur when an amino acid is substituted for one with similar properties, but drastic protein changes occur when amino acids with different properties are substituted (for example, a polar amino acid such as glycine for a nonpolar amino acid such as leucine).

Table 1. R097A patient sample quasispecies mutations

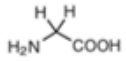
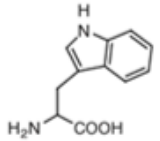
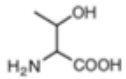
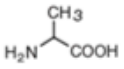
Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic	Original amino acid	Substituted amino acid	Consequence
2	E85	G→T	Transversion	Nonsynonymous	E29	G→W	Drastic	 Glycine (Gly, G)	 Tryptophan (Trp, W)	Substitution of a small polar amino acid for a large nonpolar (hydrophobic) amino acid with rigid aromatic rings on the side chain. Glycine adds flexibility to the protein chain where tryptophan will tend to orient towards the interior of the folded protein.
74	E63	C→T	Transition	Synonymous	n/a	n/a	n/a			
85	E63	C→T	Transition	Synonymous	n/a	n/a	n/a			
41	E163	A→G	Transition	Nonsynonymous	E55	T→A	Drastic	 Threonine (Thr, T)	 Alanine (Ala, A)	Substitution of a polar (hydrophilic) amino acid for a small non-polar amino acid which due to its size can be located on either the interior or exterior of the folded protein.
	E229	C deletion	n/a	Nonsynonymous	E77	frameshift	Drastic			
37	E291 E736	A→G A	Transition n/a	Synonymous Nonsynonymous	n/a E246	n/a frameshift	n/a Drastic			
88	E291	A→G	Transition	Synonymous	n/a	n/a	n/a			
50	E452	G	n/a	Nonsynonymous	E151	frameshift	Drastic			
24	E480	T→A	Transversion	Synonymous	n/a	n/a	n/a			

Table continued on page h.

Table continued from page g

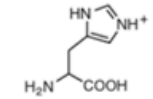
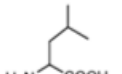
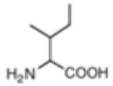
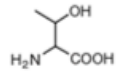
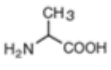
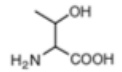
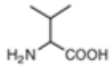
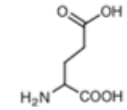
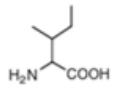
Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic	Original amino acid	Substituted amino acid	Consequence
43	E839	A→T	Transversion	Nonsynonymous	E280	H→L	Drastic	 Histidine (His, H)	 Leucine (Leu, L)	Substitution of a basic amino acid with an imidazole ring for a nonpolar amino acid with a large aliphatic hydrophobic side chain. Leucine will orient to the interior of the folded protein which may impact upon protein folding. Removal of the histidine residue may impact on the protein's response to acidic environments (e.g. fusion).
44	E843	A→G	Transition	Synonymous	n/a	n/a	n/a			
13	E959	T→C	Transition	Nonsynonymous	E320	I→T	Drastic	 Isoleucine (Ile, I)	 Threonine (Thr, T)	Substitution of a nonpolar amino acid with a large aliphatic hydrophobic side chain with a polar amino acid. Isoleucine is important for the correct folding of proteins.
3	E1020	A→T	Transversion	Synonymous	n/a	n/a	n/a			
59	E1030	G→A	Transition	Nonsynonymous	E344	A→T	Drastic	 Alanine (Ala, A)	 Threonine (Thr, T)	Substitution of a small non-polar amino acid (which due to its size can be located on either the interior or exterior of the folded protein) with a polar amino acid
69	E1095	T→C	Transition	Synonymous	n/a	n/a	n/a			
28	E1119	A→G	Transition	Synonymous	n/a	n/a	n/a			
71	E1355	T→A	Transversion	Nonsynonymous	E452	V→E	Drastic	 Valine (Val, V)	 Glutamic Acid (Glu, E)	Substitution of a nonpolar amino acid with a large aliphatic hydrophobic side chain with an acidic amino acid with a long hydrophilic side chain with a strong negative charge.

Table 2. R107A patient sample quasispecies mutations

Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic	Original amino acid	Substituted amino acid	Consequence
22	E339	T→G	Transversion	Nonsynonymous	E113	I→M	Conservative	 Isoleucine (Ile, I)	 Methionine (Met, M)	Substitution of a nonpolar amino acid with a large aliphatic hydrophobic side chain with another nonpolar amino acid containing a sulphur atom.
39	E1027	G	n/a	Nonsynonymous	E343	frameshift	Drastic			

Appendix E

Table 3. R203A patient sample quasispecies mutations

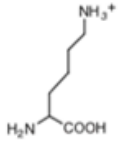
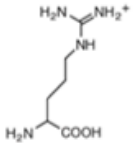
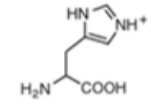
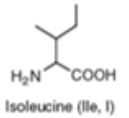
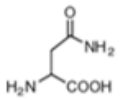
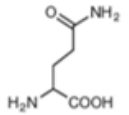
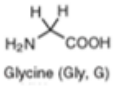
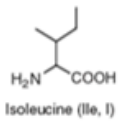
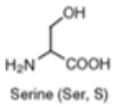
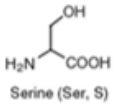
Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic	Original amino acid	Substituted amino acid	Consequence
54	E605	A→G	Transition	Nonsynonymous	E202	K→R	Conservative	 Lysine (Lys, K)	 Arginine (Arg, R)	Substitution of a basic amino acid for a functionally similar basic amino acid.
	E1032-1040	TCACAA TG→AA TACAAA TTTCAA	n/a	Nonsynonymous	E345	HNG→IQIS S	Drastic	 Histidine (His, H)	 Isoleucine (Ile, I)	Substitution of three amino acids (a basic amino acid with an imidazole ring, a polar amino acid with a long hydrophilic side chain and a small polar amino acid) with five amino acids (two nonpolar amino acid with large aliphatic hydrophobic side chains, a large polar amino acid with a long hydrophilic side chain and two polar amino acids). Incorporating two extra amino acids into the protein chain will impact upon protein structure.
								 Asparagine (Asn, N)	 Glutamine (Gln, Q)	
								 Glycine (Gly, G)	 Isoleucine (Ile, I)	
									 Serine (Ser, S)	
									 Serine (Ser, S)	
58	E673	G	n/a	Nonsynonymous	E225	frameshift	Drastic			
21	E914-915	TG	n/a	Nonsynonymous	E306	frameshift	Drastic			
77	E1119	A→G	Transition	Synonymous	n/a	n/a	n/a			

Table 4. R259A patient sample quasispecies mutations

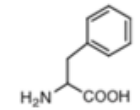
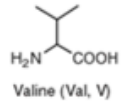
Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic	Original amino acid	Substituted amino acid	Consequence
69	E198	C→T	Transition	Synonymous	n/a	n/a	n/a			
25	E268	T→G	Transversion	Nonsynonymous	E90	F→V	Conservative	 Phenylalanine (Phe, F)	 Valine (Val, V)	Substitution of a large nonpolar amino acid with a rigid aromatic side chain with a rigid nonpolar amino acid with a large aliphatic hydrophobic side chain.
54	E370-371	AA	n/a	Nonsynonymous	E124	frameshift	Drastic			
52	E403	C	n/a	Nonsynonymous	E135	frameshift	Drastic			
83	E619	C→T	Transition	Synonymous	n/a	n/a	n/a			
4	E1033	A	n/a	Nonsynonymous	E345	frameshift	Drastic			

Table 5. R261A patient sample quasispecies mutations

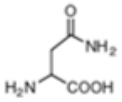
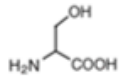
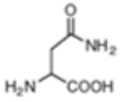
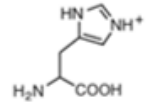
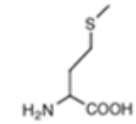
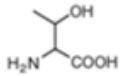
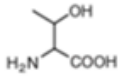
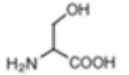
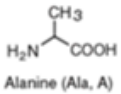
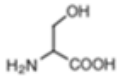
Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic	Original amino acid	Substituted amino acid	Consequence
8	E426	A→T	Transversion	Synonymous	n/a	n/a	n/a			
4	E602	A→G	Transition	Nonsynonymous	E201	N→S	Conservative	 Asparagine (Asn, N)	 Serine (Ser, S)	Substitution of a polar amino acid with a long hydrophilic side chain with a polar amino acid.
29	E602	A→G	Transition	Nonsynonymous	E201	N→S	Conservative			
34	E602	A→G	Transition	Nonsynonymous	E201	N→S	Conservative			
37	E602	A→G	Transition	Nonsynonymous	E201	N→S	Conservative			
46	E602	A→G	Transition	Nonsynonymous	E201	N→S	Conservative			
84	E602	A→G	Transition	Nonsynonymous	E201	N→S	Conservative			
44	E718	A→C	Transversion	Nonsynonymous	E240	N→H	Drastic	 Asparagine (Asn, N)	 Histidine (His, H)	Substitution of a polar amino acid with a long hydrophilic side chain with a basic amino acid with an imidazole ring. Substitution of the histidine residue may impact on the protein's response to acidic environments (e.g. fusion).
57	E1140	T→C	Transition	Synonymous	n/a	n/a	n/a			
49	E1146	T→C	Transition	Synonymous	n/a	n/a	n/a			
52	E1146	T→C	Transition	Synonymous	n/a	n/a	n/a			
60	E1146	T→C	Transition	Synonymous	n/a	n/a	n/a			
62	E1363	A	n/a	Nonsynonymous	E455	frameshift	Drastic			

Table 6. R282A patient sample quasispecies mutations

Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic	Original amino acid	Substituted amino acid	Consequence
59	E101	T→C	Transition	Nonsynonymous	E34	M→T	Drastic	 Methionine (Met, M)	 Threonine (Thr, T)	Substitution of a nonpolar amino acid containing a sulphur atom with a polar amino acid.
57	E144	A→T	Transversion	Nonsynonymous	E49	T→S	Conservative	 Threonine (Thr, T)	 Serine (Ser, S)	Substitution of a polar amino acid with a functionally similar polar amino acid.
95	E381	C→T	Transition	Synonymous	n/a	n/a	n/a			
29	E468	C→T	Transition	Synonymous	n/a	n/a	n/a			
73	E991	G→T	Transversion	Nonsynonymous	E331	A→S	Drastic	 Alanine (Ala, A)	 Serine (Ser, S)	Substitution of a small nonpolar amino acid with a polar amino acid.

Appendix F. AccuScript versus Expand method comparison: dengue virus type 3 clone envelope gene nucleotide sequence alignments

These alignments are included as an electronic version only.

Nucleotide sequence alignments for the cloned dengue virus type 3 envelope genes are included below for each RT-PCR amplification method tested (AccuScript or Expand). The clone number for each sequence is shown at the start of each row. These alignments were used to identify differences in the numbers of mutations generated using the two methodologies. For each alignment, nucleotides matching the consensus sequence are shown as dots, with mismatches shown in full. The consensus (Majority) sequence is shown along the top of each alignment. Mutations are summarized in Appendix 9, tables 7 and 8.

Appendix G. AccuScript versus Expand method comparison: dengue virus type 3 clone envelope glycoprotein amino acid sequence alignments

These alignments are included as an electronic version only.

Amino acid sequence alignments for the cloned dengue virus type 3 envelope genes are included below for each RT-PCR amplification method tested (AccuScript or Expand). The clone number for each sequence is shown at the start of each row. These alignments were used to identify amino acid sequence changes resulting from the envelope gene mutations identified in the alignments shown in Appendix 7. For each alignment, amino acids matching the consensus sequence are shown as dots, with mismatches shown in full; stop codons are represented with a slash (/). The consensus (Majority) sequence is shown along the top of each alignment. Protein sequence changes are summarized in Appendix 9, tables 1 and 2.

Appendix H. Expand versus AccuScript method comparison: summary of dengue virus type 3 envelope gene mutations

The dengue virus type 3 envelope gene mutations and amino acid sequence changes identified from the sequence alignments shown in Appendices F and G are summarised for each RT-PCR amplification method (Expand or AccuScript). A transition mutation is one where a purine (A or G) is substituted with another purine, or a pyrimidine (C, T or U) with another pyrimidine. In contrast, a transversion mutation is one in which a purine is substituted with a pyrimidine or vice versa. Nonsynonymous mutations lead to protein sequence changes whereas synonymous mutations do not. Conservative protein changes occur when an amino acid is substituted for one with similar properties, but drastic protein changes occur when amino acids with different properties are substituted (for example, a polar amino acid such as glycine for a nonpolar amino acid such as leucine).

Table 7. Expand method mutations

Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic
5	E7	T→C	Transition	Nonsynonymous	E3	C→R	Conservative
	E231	A→G	Transition	Synonymous	E77	None	
	E531	A→G	Transition	Synonymous	E177	None	
	E1365	T→C	Transition	Synonymous	E455	None	
91	E27	A→G	Transition	Synonymous	E9	None	
32	E37	G→A	Transition	Nonsynonymous	E13	E→K	Drastic
	E245	T deletion	N/A	Nonsynonymous	E82	Frameshift	Drastic
	E1037	A→G	Transition	Synonymous	E346	N/A	
92	E131	A→G	Transition	Nonsynonymous	E44	E→G	Drastic
	E248	C deletion	N/A	Nonsynonymous	E83	Frameshift	Drastic
	E964	G→T	Transversion	Nonsynonymous	E322	N/A	
	E1069	A→G	Transition	Nonsynonymous	E357	N/A	
52	E139	A→G	Transition	Nonsynonymous	E47	K→E	Drastic
3	E222	T→C	Transition	Synonymous	E74	None	
	E734	A deletion	N/A	Nonsynonymous	E245	Frameshift	Drastic
	E1051	A→G	Transition	Nonsynonymous	E351	H→R	Conservative
31	E304-305	GG deletion	N/A	Nonsynonymous	E102	Frameshift	Drastic
	E357	T→C	Transition	Nonsynonymous	E119	N/A	
	E374	T→A	Transversion	Nonsynonymous	E125	N/A	
	E735	A insertion	N/A	Nonsynonymous	E246	N/A	
49	E277	A→G	Transition	Nonsynonymous	E93	K→E	Drastic
25	E319	T→C	Transition	Synonymous	E107	None	
24	E324	T insertion	N/A	Nonsynonymous	E109	Frameshift	Drastic
	E438	A→G	Transition	Nonsynonymous	E146	N/A	
	E1410	A insertion	N/A	Nonsynonymous	E470	N/A	
30	E355	T→C	Transition	Nonsynonymous	E119	F→L	Conservative
	E735	A insertion	N/A	Nonsynonymous	E246	Frameshift	Drastic
59	E329	A→G	Transition	Nonsynonymous	E110	K→R	Conservative
	E1114	T insertion	N/A	Nonsynonymous	E372	Frameshift	Drastic
93	E319	T→C	Transition	Synonymous	E107	None	
1	E426	A→G	Transition	Synonymous	E142	None	Drastic
	E651	A→G	Transition	Synonymous	E217	None	
16	E407	A→T	Transversion	Nonsynonymous	E136	K→I	Drastic
83	E407	A→G	Transition	Nonsynonymous	E136	K→R	Conservative
	E530	G deletion	N/A	Nonsynonymous	E177	Frameshift	Drastic
19	E489	A→G	Transition	Synonymous	E163	None	
65	E500	C→T	Transition	Nonsynonymous	E167	S→L	Drastic

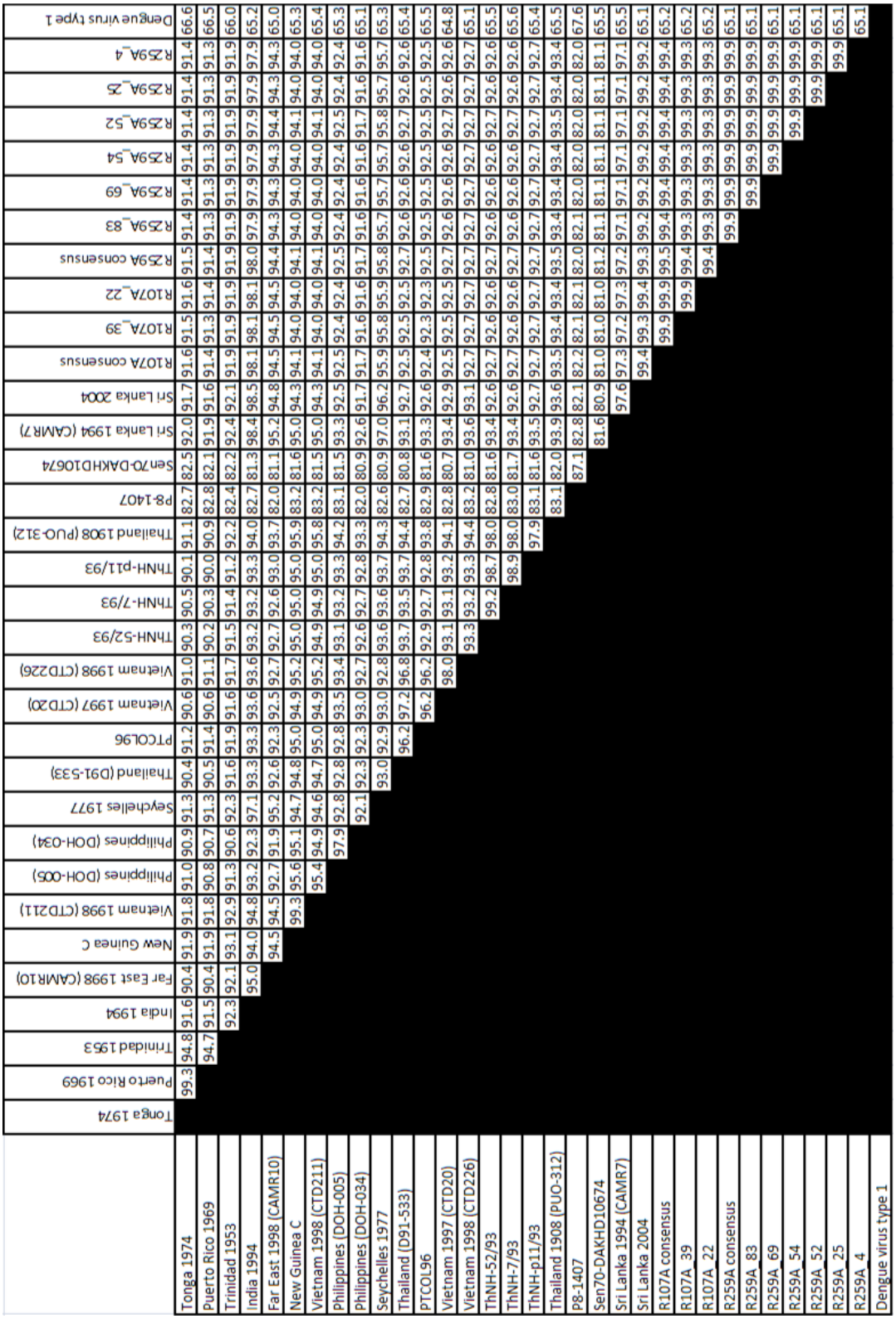
Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic
48	E550	T→C	Transition	Nonsynonymous	E184	S→P	Drastic
	E774	G→A	Transition	Nonsynonymous	E258	M→I	Conservative
	E1410	A insertion	N/A	Nonsynonymous	E470	Frameshift	Drastic
12	E734	A deletion	N/A	Nonsynonymous	E245	Frameshift	Drastic
	E1242	A deletion	N/A	Nonsynonymous	E414	Frameshift	Drastic
20	E695	A→G	Transition	Nonsynonymous	E232	K→R	Conservative
	E1009	A→G	Transition	Nonsynonymous	E337	T→A	Drastic
50	E667	A→T	Transversion	Nonsynonymous	E223	T→S	Conservative
	E1009	A→G	Transition	Nonsynonymous	E337	T→A	Drastic
71	E715	A→T	Transversion	Nonsynonymous	E239	K→Stop	Drastic
	E750	C→T	Transition	Nonsynonymous	E521	N/A	
	E936	A→G	Transition	Nonsynonymous	E313	N/A	
72	E697	G→A	Transition	Nonsynonymous	E233	E→K	Drastic
	E720	A insertion	N/A	Nonsynonymous	E240	Frameshift	Drastic
94	E735	A insertion	N/A	Nonsynonymous	E246	Frameshift	Drastic
	E1295	A→G	Transition	Nonsynonymous	E432	N/A	
14	E766	G→A	Transition	Nonsynonymous	E256	G→R	Drastic
33	E812	T deletion	N/A	Nonsynonymous	E272	Frameshift	Drastic
41	E774	G→A	Transition	Nonsynonymous	E258	M→I	Conservative
11	E964	G deletion	N/A	Nonsynonymous	E322	Frameshift	Drastic
	E998	C insertion	N/A	Nonsynonymous	E334	Frameshift	Drastic
42	E1062	A→G	Transition	Synonymous	E354	None	

Table 7 continued from page p

Table 8. AccuScript method mutations

Clone No.	Mutation position (nt)	Nt change	Transition / Transversion	Synonymous / Nonsynonymous	Mutation position (protein)	Protein change	Conservative / drastic
15	E95	C→G	Transversion	Nonsynonymous	E32	T→S	Conservative
84	E79	C→A	Transversion	Nonsynonymous	E27	H→N	Drastic
	E1410	A insertion	N/A	Nonsynonymous	E470	Frameshift	Drastic
71	E539	T insertion	N/A	Nonsynonymous	E180	Frameshift	Drastic
	E1427	T insertion	N/A	Nonsynonymous	E476	Frameshift	Drastic
89	E573	T insertion	N/A	Nonsynonymous	E191	Frameshift	Drastic
4	E677	AA insertion	N/A	Nonsynonymous	E226	Frameshift	Drastic
91	E673	G→A	Transition	Nonsynonymous	E225	E→K	Drastic
35	E767	G deletion	N/A	Nonsynonymous	E256	Frameshift	Drastic
73	E847	A insertion	N/A	Nonsynonymous	E282	Frameshift	Drastic
31	E1005	T insertion	N/A	Nonsynonymous	E335	Frameshift	Drastic
37	E1005	T insertion	N/A	Nonsynonymous	E335	Frameshift	Drastic
21	E1038-1039	TG→ATTTCAA	N/A	Nonsynonymous	E346	Frameshift	Drastic
52	E1166	G→A	Transition	Nonsynonymous	E389	W→Stop	Drastic
27	E1349	C→G	Transversion	Nonsynonymous	E450	S→C	Conservative
	E1352	G→T	Transversion	Nonsynonymous	E451	W→L	Conservative
	E1355	T deletion	N/A	Nonsynonymous	E452	Frameshift	Drastic
	E1359-62	GAAA→TTGT	Both	Nonsynonymous	E453-4	M→F	Conservative
38	E1349	C→G	Transversion	Nonsynonymous	E450	S→C	Conservative
	E1352	G→T	Transversion	Nonsynonymous	E451	W→L	Conservative
	E1355	T deletion	N/A	Nonsynonymous	E452	Frameshift	Drastic
	E1359-62	GAAA→TTGT	Both	Nonsynonymous	E453-4	M→F	Conservative
50	E1410	A insertion	N/A	Nonsynonymous	E470	Frameshift	Drastic
51	E1410	A insertion	N/A	Nonsynonymous	E470	Frameshift	Drastic
53	E1410	A insertion	N/A	Nonsynonymous	E470	Frameshift	Drastic
65	E1410	A insertion	N/A	Nonsynonymous	E470	Frameshift	Drastic

Appendix I. Percentage sequence identities calculated from dengue virus nucleotide sequence alignments for phylogenetic analysis



	Tonga 1974	Puerto Rico 1969	Trinidad 1953	India 1994	Far East 1998 (CAMR10)	New Guinea C	Vietnam 1998 (CTD211)	Philippines (DOH-005)	Philippines (DOH-034)	Seychelles 1977	Thailand (D91-533)	PTCOL96	Vietnam 1997 (CTD20)	Vietnam 1998 (CTD226)	ThNH-52/93	ThNH-7/93	ThNH-p11/93	Thailand 1908 (PUO-312)	P8-1407	Sen70-DAKHD10674	Sri Lanka 1994 (CAMR7)	Sri Lanka 2004	R107A consensus	R107A_39	R107A_22	R259A consensus	R259A_83	R259A_69	R259A_54	R259A_52	R259A_25	R259A_4	Dengue virus type 1
Tonga 1974																																	
Puerto Rico 1969	99.3																																
Trinidad 1953	94.8	94.7																															
India 1994	91.6	91.5	92.3																														
Far East 1998 (CAMR10)	90.4	90.4	92.1	95.0																													
New Guinea C	91.9	91.9	93.1	94.0	94.5																												
Vietnam 1998 (CTD211)	91.8	91.8	92.9	94.8	94.5	99.3																											
Philippines (DOH-005)	91.0	90.8	91.3	93.2	92.7	95.6	95.4																										
Philippines (DOH-034)	90.9	90.7	90.6	92.3	91.9	95.1	94.9	97.9																									
Seychelles 1977	91.3	91.3	92.3	97.1	95.2	94.7	94.6	92.8	92.1																								
Thailand (D91-533)	90.4	90.5	91.6	93.3	92.6	94.8	94.7	92.8	92.3	93.0																							
PTCOL96	91.2	91.4	91.9	93.3	92.3	95.0	94.9	92.8	92.3	92.9	96.2																						
Vietnam 1997 (CTD20)	90.6	90.6	91.6	93.6	92.5	94.9	95.2	93.5	93.0	93.0	97.2	96.2																					
Vietnam 1998 (CTD226)	91.0	91.1	91.7	93.6	92.7	95.2	94.9	93.4	92.7	92.8	96.8	96.2	98.0																				
ThNH-52/93	90.3	90.2	91.5	93.2	92.7	95.0	94.9	93.1	92.6	93.6	93.7	92.9	93.1	93.3																			
ThNH-7/93	90.5	90.3	91.4	93.2	92.6	95.0	94.9	93.2	92.7	93.6	93.5	92.7	93.1	93.2	99.2																		
ThNH-p11/93	90.1	90.0	91.2	93.3	93.0	95.0	95.0	93.3	92.8	93.7	93.7	92.8	93.2	93.3	98.7	98.9																	
Thailand 1908 (PUO-312)	91.1	90.9	92.2	94.0	93.7	95.9	95.8	94.2	93.3	94.3	94.4	93.8	94.1	94.4	98.0	98.0	97.9																
P8-1407	82.7	82.8	82.4	82.7	82.0	83.2	83.2	83.1	82.0	82.6	80.8	82.9	82.8	83.2	82.8	83.0	83.1	83.1															
Sen70-DAKHD10674	82.5	82.1	82.2	81.3	81.1	81.6	81.5	81.5	80.9	80.9	80.8	81.6	81.6	81.0	81.6	81.7	81.6	82.0	87.1														
Sri Lanka 1994 (CAMR7)	92.0	91.9	92.4	98.4	95.2	95.0	95.0	93.3	92.6	97.0	93.1	93.3	92.9	93.6	93.4	93.6	93.5	93.9	82.8	81.6													
Sri Lanka 2004	91.7	91.6	92.1	98.5	94.8	94.3	94.1	92.5	91.7	96.2	92.7	92.4	92.5	93.1	92.6	92.6	92.7	93.6	82.1	80.9	97.6												
R107A consensus	91.6	91.4	91.9	98.1	94.5	94.1	94.0	92.4	91.7	95.9	92.5	92.2	92.5	92.7	92.7	92.7	92.7	93.5	82.2	81.0	97.3	99.4											
R107A_39	91.5	91.3	91.9	98.1	94.5	94.0	94.0	92.4	91.6	95.8	92.5	92.3	92.5	92.7	92.6	92.6	92.7	93.4	82.1	81.0	97.2	99.3	99.9										
R107A_22	91.6	91.4	91.9	98.1	94.5	94.0	94.0	92.4	91.6	95.9	92.5	92.4	92.5	92.7	92.6	92.6	92.7	93.4	82.1	81.0	97.3	99.4	99.9	99.9									
R259A consensus	91.5	91.4	91.9	98.0	94.4	94.1	94.0	92.5	91.7	95.8	92.7	92.3	92.5	92.7	92.7	92.7	92.7	93.5	82.0	81.2	97.2	99.5	99.4	99.4	99.4								
R259A_83	91.4	91.3	91.9	97.9	94.3	94.0	94.0	92.4	91.6	95.7	92.6	92.5	92.6	92.7	92.6	92.6	92.7	93.4	82.0	81.1	97.1	99.4	99.4	99.3	99.3	99.9							
R259A_69	91.4	91.3	91.9	97.9	94.3	94.0	94.0	92.4	91.6	95.7	92.6	92.5	92.6	92.7	92.6	92.6	92.7	93.4	82.0	81.1	97.1	99.4	99.4	99.3	99.3	99.9	99.9						
R259A_54	91.4	91.3	91.9	97.9	94.3	94.0	94.0	92.4	91.6	95.7	92.6	92.5	92.6	92.7	92.6	92.6	92.7	93.4	82.0	81.1	97.1	99.2	99.4	99.3	99.3	99.9	99.9	99.9					
R259A_52	91.4	91.3	91.9	97.9	94.4	94.1	94.0	92.5	91.7	95.8	92.7	92.5	92.5	92.7	92.7	92.7	92.7	93.5	82.0	81.1	97.1	99.2	99.4	99.3	99.3	99.9	99.9	99.9	99.9				
R259A_25	91.4	91.3	91.9	97.9	94.3	94.0	94.0	92.4	91.6	95.7	92.6	92.5	92.5	92.6	92.6	92.6	92.7	93.4	82.0	81.1	97.1	99.2	99.4	99.3	99.3	99.9	99.9	99.9	99.9	99.9			
R259A_4	91.4	91.3	91.9	97.9	94.3	94.0	94.0	92.4	91.6	95.7	92.6	92.5	92.5	92.6	92.6	92.6	92.7	93.4	82.0	81.1	97.1	99.2	99.4	99.3	99.3	99.9	99.9	99.9	99.9	99.9	99.9		
Dengue virus type 1	66.6	66.5	66.0	65.2	65.0	65.3	65.4	65.3	65.1	65.3	65.4	65.5	64.8	65.1	65.5	65.6	65.4	65.5	67.6	65.5	65.5	65.1	65.2	65.2	65.2	65.1	65.1	65.1	65.1	65.1	65.1	65.1	

Figure 2. Percentage sequence identities calculated from dengue virus type 2 nucleotide sequence alignments for phylogenetic analysis

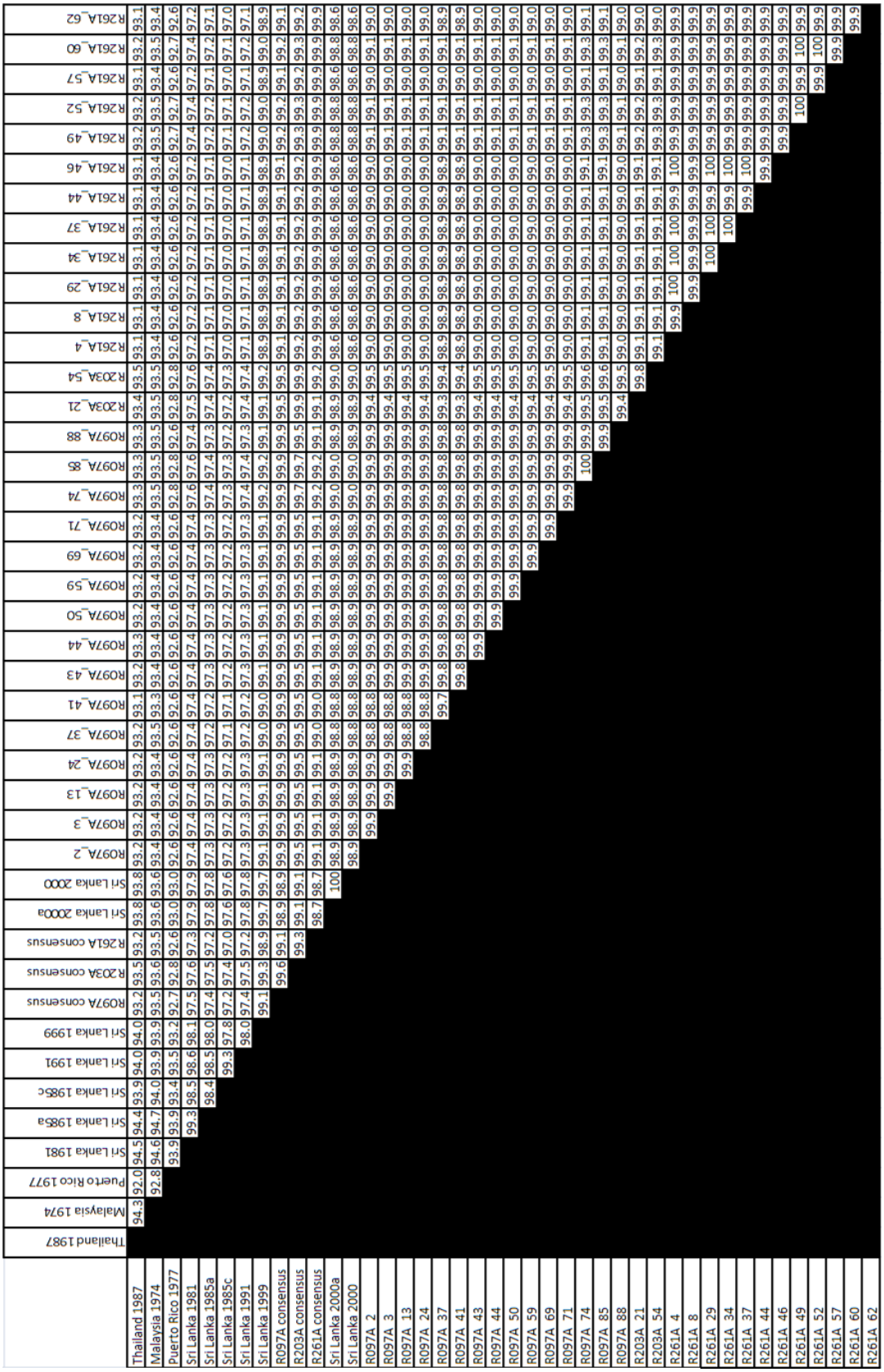
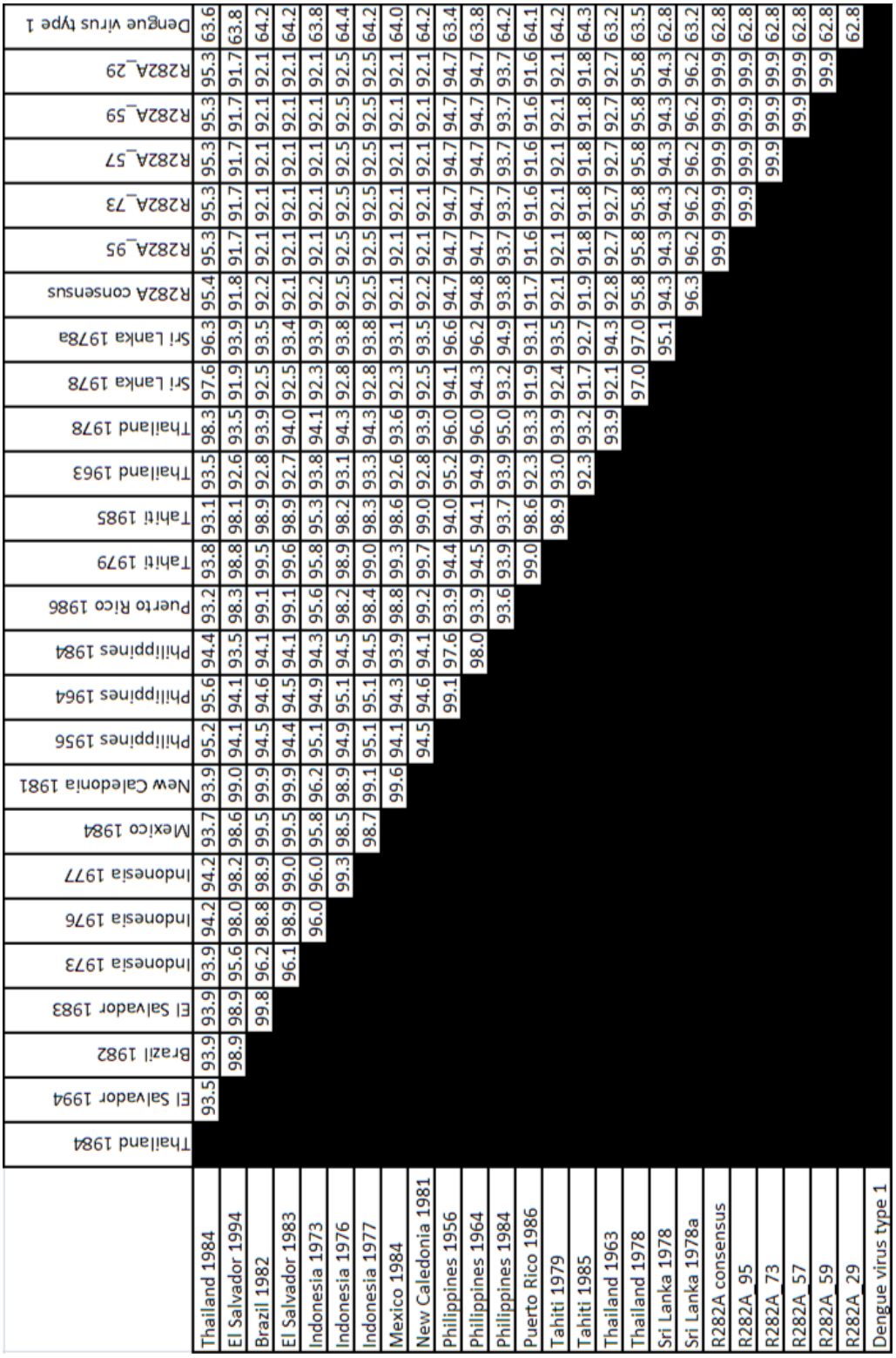


Figure 3. Percentage sequence identities calculated from dengue virus type 3 nucleotide sequence alignments for phylogenetic analysis

Appendix I



	Thailand 1984	El Salvador 1994	Brazil 1982	El Salvador 1983	Indonesia 1973	Indonesia 1976	Indonesia 1977	Mexico 1984	New Caledonia 1981	Philippines 1956	Philippines 1964	Philippines 1984	Puerto Rico 1986	Tahiti 1979	Tahiti 1985	Thailand 1963	Thailand 1978	Sri Lanka 1978	Sri Lanka 1978a	R282A consensus	R282A_95	R282A_73	R282A_57	R282A_59	R282A_29	Dengue virus type 1
Thailand 1984		93.5	93.9	93.9	93.9	94.2	94.2	93.7	93.9	95.2	95.6	94.4	93.2	93.8	93.1	93.5	98.3	97.6	96.3	95.4	95.3	95.3	95.3	95.3	95.3	63.6
El Salvador 1994	93.5		98.9	98.9	95.6	98.0	98.2	98.6	99.0	94.1	94.1	93.5	98.3	98.8	98.1	92.6	93.5	91.9	93.9	91.8	91.7	91.7	91.7	91.7	91.7	63.8
Brazil 1982	93.9	98.9		99.8	96.2	98.8	98.9	99.5	99.9	94.5	94.6	94.1	99.1	99.5	98.9	92.8	93.9	92.5	93.5	92.2	92.1	92.1	92.1	92.1	92.1	64.2
El Salvador 1983	93.9	98.9	99.8		96.1	98.9	99.0	99.5	99.9	94.4	94.5	94.1	99.1	99.6	98.9	92.7	94.0	92.5	93.4	92.1	92.1	92.1	92.1	92.1	92.1	64.2
Indonesia 1973	93.9	95.6	96.2	96.1		96.0	96.0	95.8	96.2	95.1	94.9	94.3	95.6	95.8	95.3	93.8	94.1	92.3	93.9	92.2	92.1	92.1	92.1	92.1	92.1	63.8
Indonesia 1976	94.2	98.0	98.8	98.9	96.0		99.3	98.5	98.9	94.9	95.1	94.5	98.2	98.9	98.2	93.1	94.3	92.8	93.8	92.5	92.5	92.5	92.5	92.5	92.5	64.4
Indonesia 1977	94.2	98.2	98.9	99.0	96.0	99.3		98.7	99.1	95.1	95.1	94.5	98.4	99.0	98.3	93.3	94.3	92.8	93.8	92.5	92.5	92.5	92.5	92.5	92.5	64.2
Mexico 1984	93.7	98.6	99.5	99.5	95.8	98.5	98.7		99.6	94.1	94.3	93.9	98.8	99.3	98.6	92.6	93.6	92.3	93.1	92.1	92.1	92.1	92.1	92.1	92.1	64.0
New Caledonia 1981	93.9	99.0	99.9	99.9	96.2	98.9	99.1	99.6		94.5	94.6	94.1	99.2	99.7	99.0	92.8	93.9	92.5	93.5	92.2	92.1	92.1	92.1	92.1	92.1	64.2
Philippines 1956	95.2	94.1	94.5	94.4	95.1	94.9	95.1	94.1	94.5		99.1	97.6	93.9	94.4	94.0	95.2	96.0	94.1	96.6	94.7	94.7	94.7	94.7	94.7	94.7	63.4
Philippines 1964	95.6	94.1	94.6	94.5	94.9	95.1	95.1	94.3	94.6	99.1		98.0	93.9	94.5	94.1	94.9	96.0	94.3	96.2	94.8	94.7	94.7	94.7	94.7	94.7	63.8
Philippines 1984	94.4	93.5	94.1	94.1	94.3	94.5	94.5	93.9	94.1	97.6	98.0		93.6	93.9	93.7	93.9	95.0	93.2	94.9	93.8	93.7	93.7	93.7	93.7	93.7	64.2
Puerto Rico 1986	93.2	98.3	99.1	99.1	95.6	98.2	98.4	98.8	99.2	93.9	93.9	93.6		99.0	98.6	92.3	93.3	91.9	93.1	91.7	91.6	91.6	91.6	91.6	91.6	64.1
Tahiti 1979	93.8	98.8	99.5	99.6	95.8	98.9	99.0	99.3	99.7	94.4	94.5	93.9	99.0		98.9	93.0	93.9	92.4	93.5	92.1	92.1	92.1	92.1	92.1	92.1	64.2
Tahiti 1985	93.1	98.1	98.9	98.9	95.3	98.2	98.3	98.6	99.0	94.1	94.1	93.7	98.6	98.9		92.3	93.2	91.7	92.7	91.9	91.8	91.8	91.8	91.8	91.8	64.3
Thailand 1963	93.5	92.6	92.8	92.7	93.8	93.1	93.3	92.6	92.8	95.2	94.9	94.5	92.3	93.0	92.3		93.9	92.1	94.3	92.8	92.7	92.7	92.7	92.7	92.7	63.2
Thailand 1978	98.3	93.5	93.9	94.0	94.1	94.3	94.3	93.6	93.9	96.0	96.0	95.0	93.3	93.9	93.2	93.9		97.0	97.0	95.8	95.8	95.8	95.8	95.8	95.8	63.5
Sri Lanka 1978	97.6	91.9	92.5	92.5	92.3	92.8	92.8	92.3	92.5	94.1	94.3	93.2	91.9	92.4	91.7	92.1	97.0		95.1	94.3	94.3	94.3	94.3	94.3	94.3	62.8
Sri Lanka 1978a	96.3	93.9	93.8	93.8	93.1	93.8	93.8	92.3	93.5	96.6	96.2	94.7	93.1	93.5	92.7	94.3	97.0	95.1		96.3	96.2	96.2	96.2	96.2	96.2	63.2
R282A consensus	95.4	91.8	92.5	92.5	92.1	92.5	92.5	92.1	92.2	94.7	94.8	93.8	91.7	91.9	91.9	92.8	95.8	94.3	96.3		99.9	99.9	99.9	99.9	99.9	62.8
R282A_95	95.3	91.7	92.1	92.1	92.1	92.5	92.5	92.1	92.1	94.7	94.7	93.7	91.6	91.6	91.8	92.7	95.8	94.3	96.2	99.9		99.9	99.9	99.9	99.9	62.8
R282A_73	95.3	91.7	92.1	92.1	92.1	92.5	92.5	92.1	92.1	94.7	94.7	93.7	91.6	91.6	91.8	92.7	95.8	94.3	96.2	99.9	99.9		99.9	99.9	99.9	62.8
R282A_57	95.3	91.7	92.1	92.1	92.1	92.5	92.5	92.1	92.1	94.7	94.7	93.7	91.6	91.6	91.8	92.7	95.8	94.3	96.2	99.9	99.9	99.9		99.9	99.9	62.8
R282A_59	95.3	91.7	92.1	92.1	92.1	92.5	92.5	92.1	92.1	94.7	94.7	93.7	91.6	91.6	91.8	92.7	95.8	94.3	96.2	99.9	99.9	99.9	99.9		99.9	62.8
R282A_29	95.3	91.7	92.1	92.1	92.1	92.5	92.5	92.1	92.1	94.7	94.7	93.7	91.6	91.6	91.8	92.7	95.8	94.3	96.2	99.9	99.9	99.9	99.9	99.9		62.8
Dengue virus type 1	63.6	63.8	64.2	64.2	63.8	64.4	64.2	64.0	64.2	63.4	63.8	64.2	64.1	64.2	64.3	63.2	63.5	62.8	63.2	62.8	62.8	62.8	62.8	62.8	62.8	

Figure 4. Percentage sequence identities calculated from dengue virus type 2 nucleotide sequence alignments for phylogenetic analysis

Appendix J. Original images

These images are included as an electronic version only.

The original (unprocessed) gel and blot images used to construct figures 3.5, 3.6, 3.8a, 4.3, and 4.5 to 4.18 are included in the file labelled Appendix J on the disc included with this thesis.

Appendix K. List of presentations

Flaviviruses.

Oral presentation to the Novel and Dangerous Pathogens group at HPA CEPR on November 15th 2007.

Investigation of envelope gene variation in clinical samples from patients with dengue virus infections.

Oral presentation to the Novel and Dangerous Pathogens group at HPA CEPR on February 26th 2009.

Investigation of dengue virus envelope gene variation in clinical samples from Sri Lanka.

Hannah Love, Dr Jane Burton, Dr Daniel Bailey, Prof. David Cullen. Conference poster presentation for FEMS (Federation of European Microbiological Societies) 3rd congress of European microbiologists. Gothenburg, Sweden; June 28th – July 2nd 2009.

Appendix L. Ethical considerations relating to the Sri Lankan patient samples and clinical data

Copies of letters confirming that ethical approval was obtained from both the University of Kelaniya in Sri Lanka and the Liverpool School of Tropical Medicine in the UK, and the corresponding summaries of the Sri Lanka Fever Study protocol are provided as an electronic version only, in the file labelled Appendix L on the disc included with this thesis.

Briefly, ethical approval was obtained for extra blood samples to be taken from consenting adults with fever, or fever history who were admitted to hospitals in Ragama, Sri Lanka. The aim of the study was to identify the specific infections that cause febrile illnesses in Sri Lanka. This would assist in making clinical diagnoses of febrile illnesses, provide useful microbiological surveillance data and indicate which infections require local laboratory confirmation in the future. It would also enable various serological tests to be calibrated for use in local populations. Approval was granted for the anonymised samples and clinical data to be sent to collaborating institutions, including HPA CEPR in the UK, for testing for arboviruses including dengue virus, and for these anonymised electronic records and biomedical samples to be kept permanently by the laboratories involved for possible non-commercial use in the future.

For the work presented in this thesis, therefore, further ethical approval was not required for the following reasons:

1. The samples and clinical data were anonymised.

2. The work was restricted to amplification of the dengue viruses already known to be present in the samples, and the clinical data pertaining to the dengue virus infection.

The work would lead to diagnostic service improvement at HPA CEPR.