**YMTHE, Volume *30***

## Supplemental Information

# Immune profile and responses of a novel dengue DNA vaccine encoding an EDIII-NS1 consensus design based on Indo-African sequences

**Arun Sankaradoss, Suraj Jagtap, Junaid Nazir, Shefta E. Moula, Ayan Modak, Joshuah Fialho, Meenakshi Iyer, Jayanthi S. Shastri, Mary Dias, Ravisekhar Gadepalli, Alisha Aggarwal, Manoj Vedpathak, Sachee Agrawal, Awadhesh Pandit, Amul Nisheetha, Anuj Kumar, Mahasweta Bordoloi, Mohamed Shafi, Bhagyashree Shelar, Swathi S. Balachandra, Tina Damodar, Moses Muia Masika, Patrick Mwaura, Omu Anzala, Kar Muthumani, Ramanathan Sowdhamini, Guruprasad R. Medigeshi, Rahul Roy, Chitra Pattabiraman, Sudhir Krishna, and Easwaran Sreekumar**

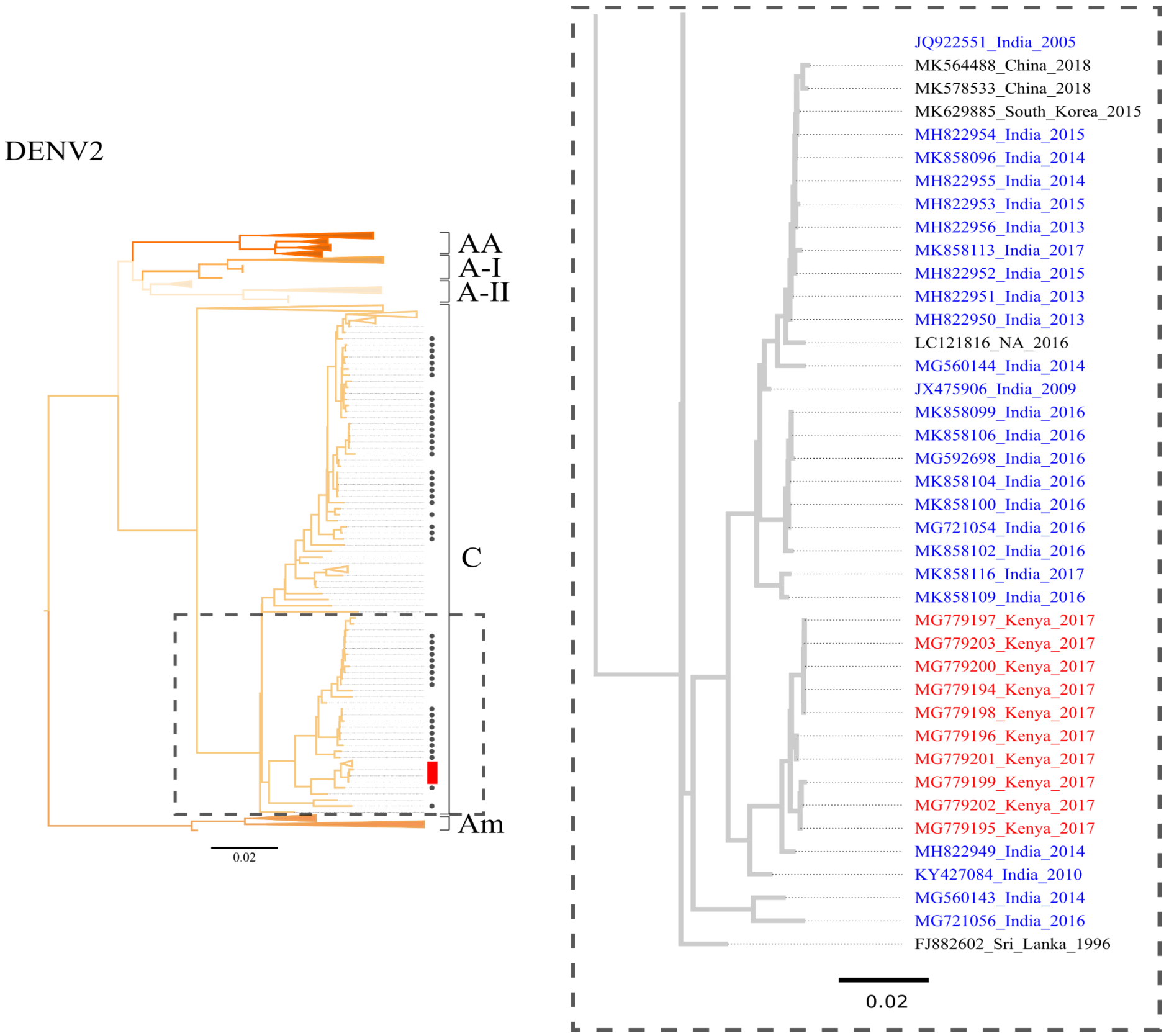



**Figure S1. Genetic relationship of Indo-Africa DENV2 strains**

Phylogenetic tree of DENV-2 are highlighted to show where Indian and African strains cluster together. Indian strains are represented by blue, while African strains are represented by red**.**

| DENV serotype and genotype information | Non-structural protein 1-Amino acid position | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 5 | 6 | 9 | 21 | 22 | 50 | 54 | 60 | 62 | 71 | 84 | 90 | 93 | 94 | 95 | 98 |
| DENV1 (n=40) Genotypes III and I | V | I | K | V | T | W | V | A | R | I | M | A | A | A/N | R | T |
| DENV2 (n=48) Cosmopolitan | V | V | K | I | A | Y | M | A | G | I | I/A | T | I | K | G | Q |
| DENV3 (n=22) Genotype III | V | I | K | V | T | W | V | T | R | G | I | V | S/I | I | G | Q |
| DENV4 (n=9) Genotype I | A | V | N | I | I | H | V | T | R | I | H | T | V | K | G | S/F |

| 103 | 105 | 109 | 117 | 126 | 127 | 128 | 131 | 139 | 144 | 145 | 146 | 152 | 162 | 171 | 174 | 175 | 178 | 192 | 194 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| T | R | M | N | A | E | T | S | D | S | D | D | I | V | L | S | H | M | K | V |
| S | R | T | T | I | E | P/H | Q | E | P | S | T | S | I | L | R | Q | F | G | I |
| T | T | M | T | A | E | I | S | N | P | S | A | V | V | L | V | Y | M | R | |
| A | A | N | T | P | E | T | S | D | P | N | E | F | M | L | G | S | V | K | |

| 202 | 213 | 214 | 217 | 220 | 222 | 227 | 233 | 246 | 247 | 251 | 261 | 264 | 265 | 276 | 278 | 281 | 290 | 296 | 307 | 324 | 347 | 351 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | V | R | F | V | T | R | S | I | Y | I | | I/S | A | D | N | E | H | A | I | R | R | S |
| M | E | K | F | I | S | K | S | N | F | A | Y | T | A | D | D | E | N | P | L | K | N | T |
| | E | K | F | V | T | K | S | S | L | I | | T | A | D | N | E | N | P | L | M | K | S |
| | E | R | L | V | T | K | S | A | Y | F | | T | M | E | G | P | D | P | L | L | K | S |

| Variant frequency | 0-10% | 10-20% | 20-30% | >50% |
|---|---|---|---|---|

**Figure S2: DENV NS1 amino acid variation in our study clinical isolates for all four serotypes.** Frequency based representation of all the variable sites in NS1 for each serotype within our study sequences. Amino acid variants residue identified relative to (NC_001477(DENV1), NC_001474 (DENV2), NC_001475 (DENV3), NC_002640 (DENV4). Colours assigned based on the % frequency of mutations in the given clinical isolates.

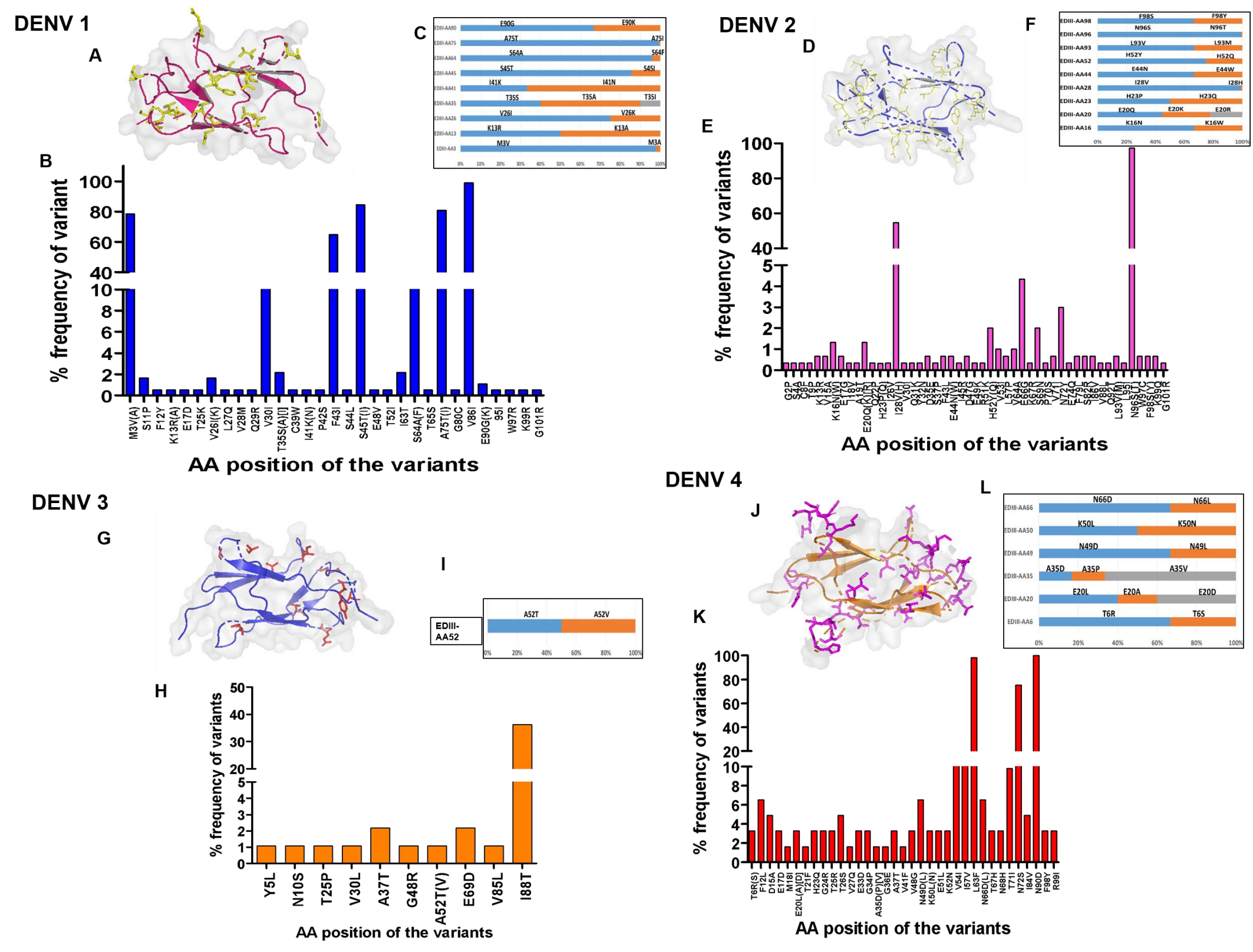
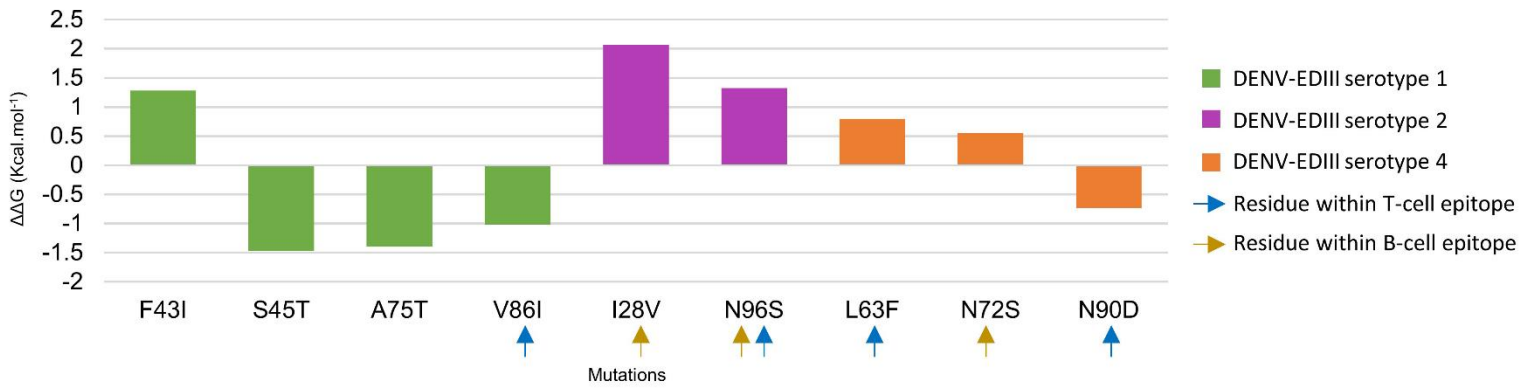



**Figure S3: DENV amino acid variation among Indian strains of all four serotypes isolated from 2000-2020**

(A, D, G and J) The genetic diversity within EDIII of DENV1, DENV2, DENV3 and DENV4 Indian strain variants spotted on EDIII PDB structure (ribbon). Stick and ball representation on PDB structure indicates variants at the particular position. PDB: 4gt0.1, 4ut6.2.8,4Gsx.1.A and 5BIC.1.A, used to annotate the genetic variants of DENV. (B, E, H and K ) Frequencies of variants within Indian DENV1, DENV2, DENV3, DENV4 strains, respectively. Amino acid variants residue identified relative to (NC_001477(DENV1), NC_001474 (DENV2), NC_001475 (DENV3), NC_002640 (DENV4). (C, F, I and L) Relative frequencies of multiple amino acid substitution at a given position in DENV1, DENV2, DENV3,and DENV4, respectively, represented in different colors.

A




B

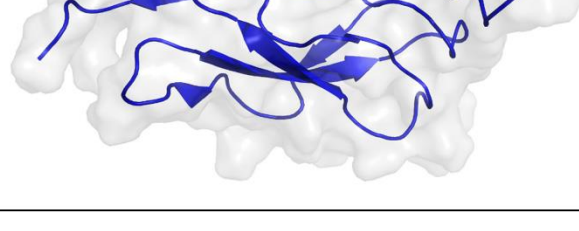
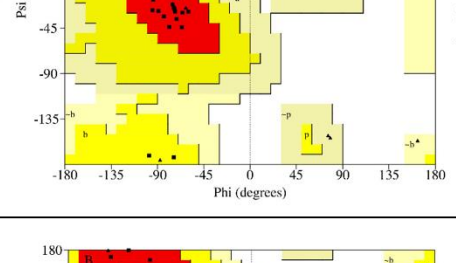
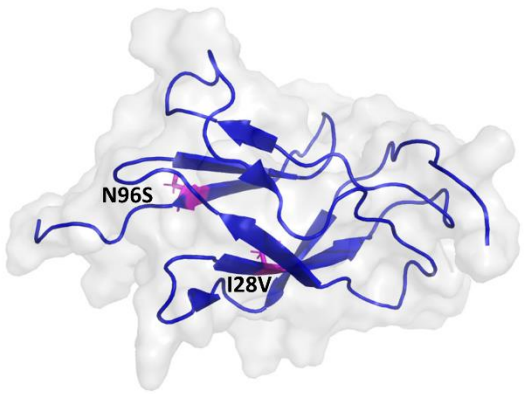
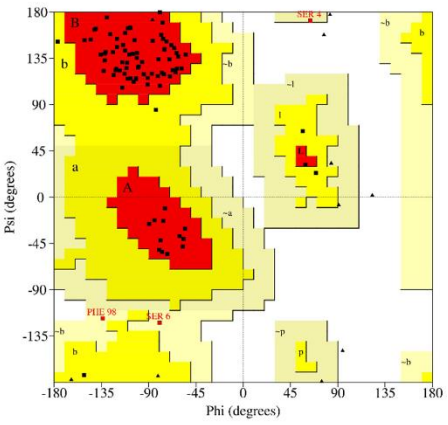
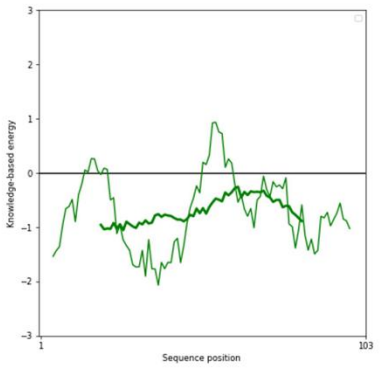
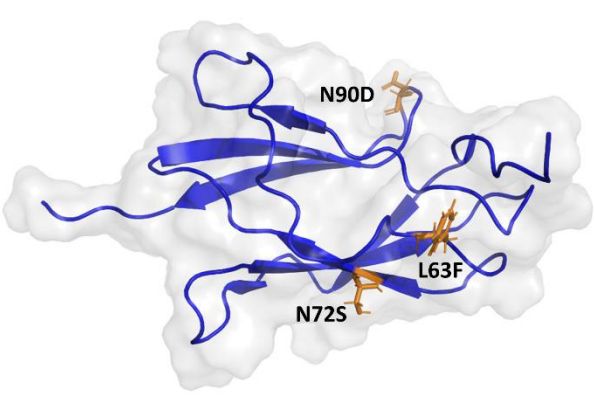
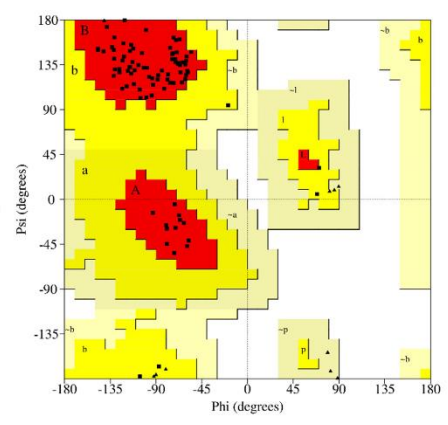
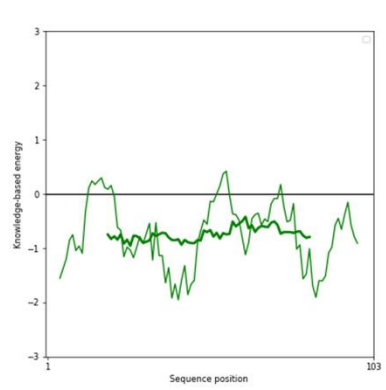
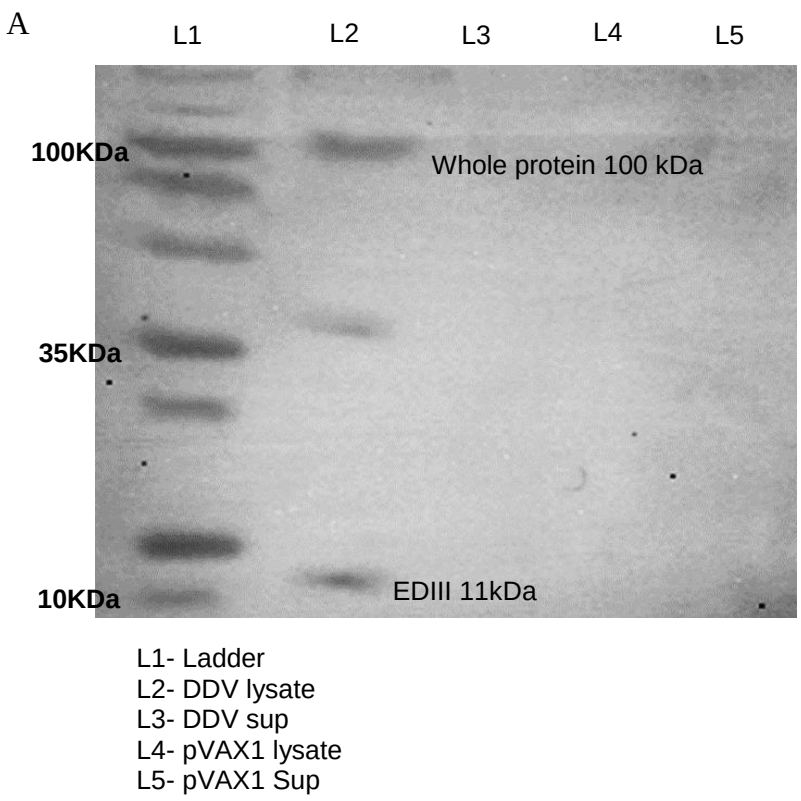
| | Modelled structure (EDIII) | Ramachandran plot | ProSA energy plot |
|---|---|---|---|
| DENV1 |   |   |   |
| DENV2 |   |   |   |
| DENV3 |  |   |   |
| DENV4 |   |   |   |

**Figure S4. Construct structure modelling and stability analysis: (A) The graph depicts the effect of mutations on structure stability**. The ΔΔG free energy was calculated for each mutation using FoldX. Mutations that occurred within B-Cell and T-Cell epitope regions were also marked. (B) The models for DENV-EDIII consensus serotypes were generated using Modeller program. The mutations are highlighted in the structure. Ramachandran map shows that 100% of the residues are in the allowed region for the modelled structures, the predicted z-score based on ProSA analysis is negative and within the ranges for the solved crystal structures.



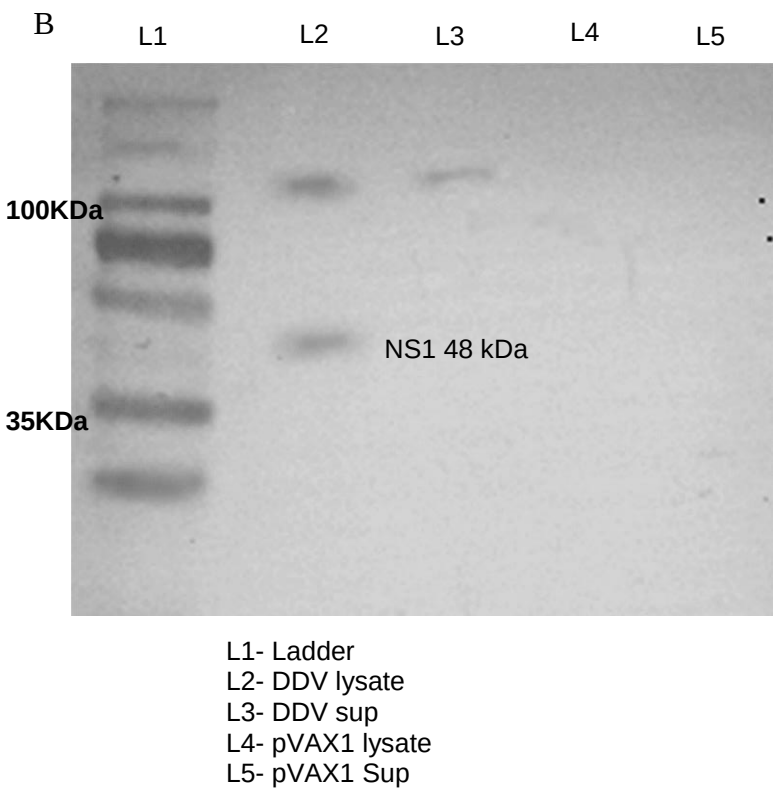
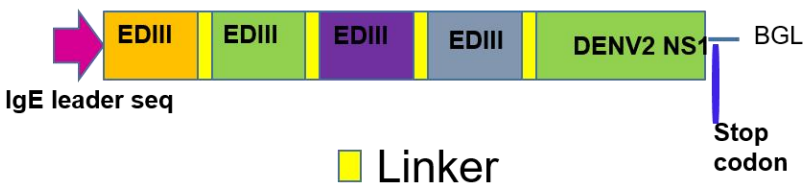
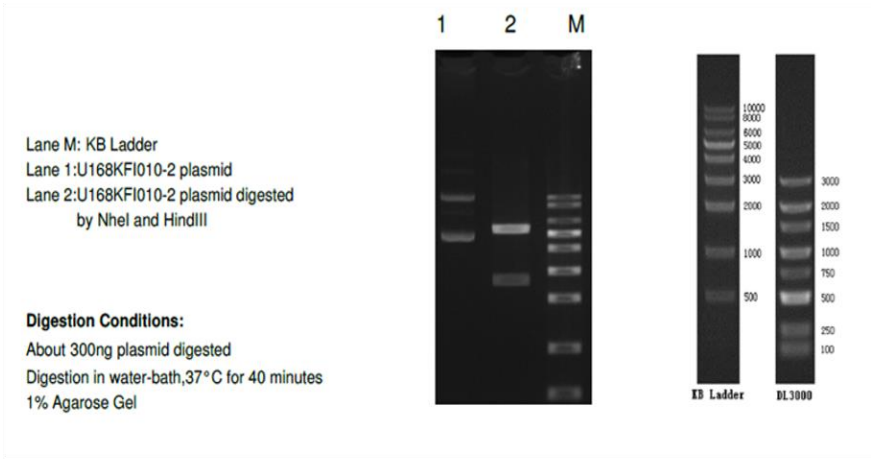




**Figure S5: *In vitro* antigen expression and localization using monoclonal EDIII and NS1 antibodies.** Analysis of in vitro expression of ED III (A) and NS1 (B) protein after transfection of 293T cells with DDV or plasmid control by Western blot. 293T cells supernatant and lysates resolved on a gel and probed with EDIII and NS1 monoclonal antibodies. Blots were stripped then probed with β-actin loading control.

A




B




C

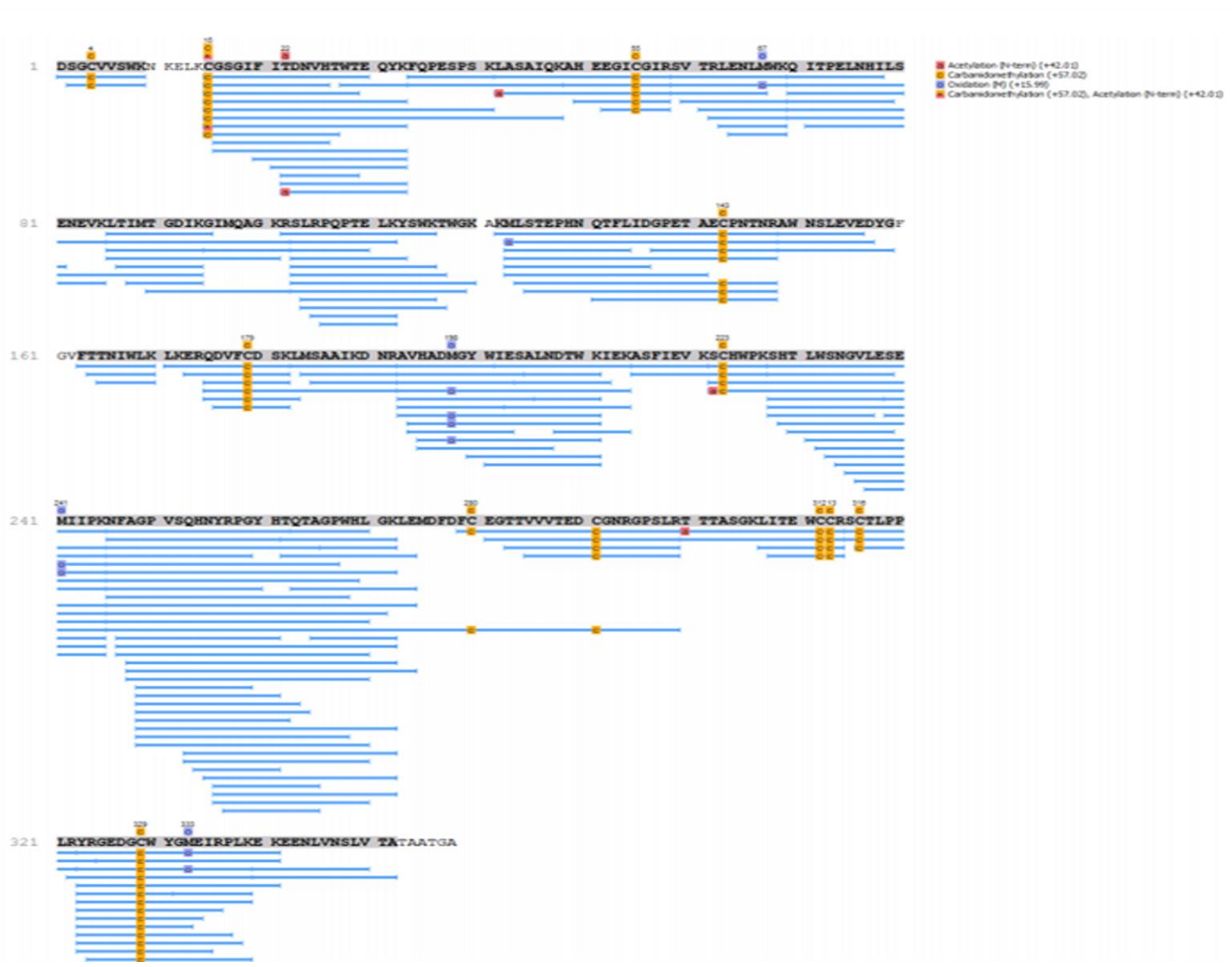



**Figure S6. Recombinant protein synthesis.** A. Schematic representation of DENV DNA vaccine recombinant protein synthesis design. (B) Enzyme Digestion: About 300ng plasmid digested. Digestion in water-bath,37°C for 40 minutes and analysed in 1% Agarose Gel (C) Mass spec profile of His-purified recombinant protein

**Table S1. Clinical characteristics and biochemical correlates of sequenced samples**

| Variables / Characteristics* | All patients (N=319)[a] | Dengue Fever (N=114) | Dengue with warning signs N=128 | Severe dengue N=28 | p value | Statistical test |
|---|---|---|---|---|---|---|
| **Patient characteristics** | | | | | | |
| Age | 23 (0.6, 83) | 24 (4,66) | 25 (0.66,83) | 21 (0.75, 69) | 0.09 | Kruskal Wallis test |
| **Sex** | | | | | | |
| Female; n (%) | 140 (43.8) | 43 (37.7) | 58 (45.3) | 9 (32.1) | | |
| Male; n (%) | 173 (54.23) | 57 (50) | 57 (44.5) | 14 (50) | 0.42 | Chi-square test |
| **Onset symptoms** | | | | | | |
| Average days of fever; mean ±SD | 3.35±1.96 | 2.89±1.38 | 3.41±1.5 | 3.42±1.34 | 0.35 | ANOVA test |
| Abdominal Pain; n (%) | 54 (16.92) | 13 (20.63) | 15 (13.27) | 10 (35.71) | **0.021** | **Chi-square test** |
| Rash; n (%) | 26 (8.15) | 10 (17.24) | 12 (19.04) | 4 (21.05) | 0.92 | Chi-square test |
| Retro Orbital Pain; n (%) | 54 (16.92) | 22 (35.55) | 16 (34.04) | 8 (47.05) | 0.431 | Chi-square test |
| Hepatomegaly; n (%) | 30 (9.4) | 11 (24.13) | 8 (20.68) | 3 (22.22) | 0.77 | Chi-square test |
| **Serological classification** | | | | | | |
| Primary infection; n(%) | 191 (59.8) | 72 (63.1) | 70 (54.68) | 13 (46.42) | | |
| Secondary infection; n(%) | 62 (19.43) | 17 (14.91) | 29 (22.65) | 11 (39.28) | | |
| Unknown; n(%)[b] | 66 (20.68) | 25(21.92) | 29(22.65) | 4(14.28) | **0.024** | **Chi-square test** |
| **Biochemical parameters** | | | | | | |
| Hemoglobin (g/dL) mean (SD), | 12.79±2.09 | 12.85±2.21 | 12.89±1.91 | 11.87±2.35 | 0.06 | ANOVA test |
| TC /cumm | 4300 (1260,18800) | 4600 (2000,15400) N=101 | 4100 (1260,18800) N=121 | 4415 (2000, 10010) N=28 | 0.2657 | Kruskal Wallis |
| Hematocrit; %. mean (SD) | 37.70 ± 5.51 | 35.97 ±4.90 N=55 | 38.0 ± 4.62 N=55 | 42.48 ± 7.38 N=16 | **<0.0001** | ANOVA test |
| Platelet count at admission/ cumm | | 136 (41,302) N=105 | 97 (11,346) N=115 | 39 (14,225) N=28 | **4.398E-11 (<0.0001)** | Kruskall Wallis |
| Platelet count (least)/cumm | | 82 (38,185) N=13 | 31 (6, 182) N=41 | 32 (15, 45) N=7 | **0.000814** | Kruskall Wallis |
| Aspartate aminotransferase; IU/L | 52.5 (10, 1480) N=206 | 38 (10, 673) N=73 | 66 (13, 562) N=105 | 150 (30, 1480) N=24 | **0.0000000003102 (<0.0001)** | Kruskall Wallis |
| Alanine aminotransferase; IU/L | 40 (12,802) N=206 | 32 (13, 743) N=73 | 41 (12, 454) N=105 | 87 (13, 802) N=24 | **0.0001922** | Kruskall Wallis |

*****Values are Median unless stated otherwise. SD- standard deviation

[a] Total number of samples sequenced. Since there was no clinical and biochemical information, WHO disease severity classification could not be done for 49 cases.

[b] Unknown baseline serostatus across stratifications: Inadequate samples to perform serum dengue-specific IgM and IgG Elisa

**Table S2a. EDIII diversity in DENV 1-4 genotype strains**

| DENV1 | | | DENV2 | | | DENV3 | | | DENV4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Genotype | Median and range | IQR | Genotype | Median and range | IQR | Genotype | Median and range | IQR | Genotype | Median and range | IQR |
| I | 2.91 (0.97-4.85) | 2.91-3.88 | American (16) | 1.94 (0.00-2.91) | 1.94-2.18 | I (62) | 3.88 (2.91-5.83) | 3.88-3.88 | I (45) | 2.91 (1.94-4.85) | 2.91-3.88 |
| III | 4.85 (3.88-5.83) | 4.85-4.85 | Asian I | 0.97 (0.97-1.94) | 0.97-0.97 | II | 1.94 (1.94-2.91) | 1.94-1.94 | IIA (34) | 1.94 (1.94-4.85) | 1.94-3.88 |
| IV (47) | 1.94 (0.97-4.85) | 1.94-2.91 | Asian II (29) | 0.97 (0.00-2.91) | 0.00-1.94 | III | 0.97 (0.00-2.91) | 0.97-0.97 | IIB | 1.94 (0.00-3.88) | 0.97-1.94 |
| V | 2.91 (1.94-3.88) | 2.91-2.91 | Asian American | 1.94 (0.97-4.85) | 0.97-1.94 | | | | | | |
| | | | Cosmopolitan | 0.97 (0.97-2.91) | 0.97-1.94 | | | | | | |

**Table S2b. NS1 diversity in DENV 1-4 genotype strains**

| DENV1 | | | DENV2 | | | DENV3 | | | DENV4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Genotype | Median (Range) | IQR | Genotype | Median (Range) | IQR | Genotype | Median (Range) | IQR | Genotype | Median (Range) | IQR |
| I | 1.42 (0.85-3.41) | 1.42-1.7 | American (16) | 4.55 (4.26-8.81) | 4.47-4.9 | I (62) | 2.27 (1.7-3.13) | 1.99-2.56 | I (45) | 4.26 (3.13-5.97) | 3.69-4.83 |
| III | 3.69 (2.56-4.55) | 3.69-3.69 | Asian I | 1.7 (1.14-3.13) | 1.7-1.7 | II | 2.84 (1.99-5.11) | 2.56-3.13 | IIA (34) | 2.27 (1.99-2.84) | 1.99-2.56 |
| IV (47) | 1.42 (0.57-2.56) | 1.14-1.99 | Asian II (29) | 1.99 (1.7-3.98) | 1.99-3.69 | III | 1.42 (0.85-2.56) | 1.14-1.7 | IIB | 1.7 (0.28-2.27) | 0.85-1.7 |
| V | 2.84 (2.27-4.55) | 2.56-3.13 | Asian American | 2.27 (1.99-3.13) | 2.27-2.56 | | | | | | |
| | | | Cosmopolitan | 3.13 (2.27-6.82) | 2.56-3.41 | | | | | | |

**Table S3. Structural significance and immune interactions of mutations detected in the EDIII of our study clinical isolates**

| Mutations | Predicted structural significance and immune interactions | Amino acid class |
|---|---|---|
| **DENV1** | | |
| M3V | Resides within a known T cell epitope (IEDB identifier : 867378); Fold-X stability calculations suggested this mutation to be mildly destabilizing | Hydrophobic |
| V30I | Resides within a discontinuous B cell epitope (IEDB identifiers; 504136) | Hydrophobic |
| F43I | Fold-X stability calculations suggested this mutation to be mildly destabilizing | Hydrophobic |
| S45T | Fold-X stability calculations suggested this mutation to be mildly stabilizing | Polar uncharged |
| I63T | Resides within a known T cell epitope (IEDB identifiers: 195646), Fold-X stability calculations suggested this mutation to be mildly destabilizing | Hydrophobic to polar uncharged |
| V64A | Resides within a known T cell epitope (IEDB identifiers; 195646), Fold-X stability calculations suggested this mutation to be mildly destabilizing | Hydrophobic |
| A75T | Fold-X stability calculations suggested this hydrophobic amino acid substitution to be mildly stabilizing | Hydrophobic to polar uncharged |
| V86I | Resides within a known T cell epitope (IEDB identifiers: 869085), Fold-X stability calculations suggested this mutation to be mildly stabilizing | Hydrophobic |
| E90G | Resides within known discontinuous lateral ridge B cell epitope (IEDB identifiers: 591576, 858293, 858294, 914619, 628716, 173906) | negatively charged to aliphatic |
| **DENV2** | | |
| K13R | Resides within known discontinuous and linear B cell epitopes (IEDB identifiers; 504134, 538883) | Positively charged |
| V15A | Resides within known discontinuous and linear B cell epitopes (IEDB identifiers; 504134, 538883, 538991), Fold-X stability calculations suggested this mutation to be mildly destabilizing | Hydrophobic |
| I28V | Resides within linear B cell epitopes (IEDB identifiers; 22716, 22717, 23961, 28415, 29315, 538991, 69106), Fold-X stability calculations suggested this mutation to be destabilizing | Hydrophobic |
| H52Y | Resides within linear B cell epitopes (IEDB identifiers; 144940, 538856), Resides within known T cell epitopes (IEDB identifiers; 144940, 183781, 184755, 184337), Fold-X stability calculations suggested this ring amino acid mutation to be mildly stabilizing | Positively charged to polar uncharged |
| E66G | Resides within a linear B cell epitope (IEDB identifiers; 538997), Resides within known T cell epitopes (IEDB identifiers; 869123, 196140, 195638) | Negatively charged to hydrophobic |
| V71I | Resides within linear B cell epitopes (IEDB identifiers; 538994, 538997), Resides within a known T cell epitope (IEDB identifiers; 869123) | Hydrophobic |
| N96S | Resides within linear B cell epitopes (IEDB identifiers; 36947, 51391, 51392, 144885), Resides within a known T cell epitope (IEDB identifiers; 180433), Fold-X stability calculations suggested this this mutation to be mildly destabilizing | Polar uncharged |
| **DENV3** | | |
| E69D | Negatively charged amino acid substitution, Resides within a known T cell epitope (IEDB identifiers; 195416), Resides within a linear B cell epitope (IEDB identifiers; 919522) | Negatively charged |
| I88T | Resides within a lateral ridge epitope. Fold-X stability calculations suggested this this mutation to be mildly destabilizing | Hydrophobic to polar uncharged |
| **DENV4** | | |
| A37T | Fold-X stability calculations suggested this mutation to be destabilizing | Hydrophobic to polar uncharged |
| V54I | Resides within a known T cell epitope (IEDB identifiers;167808) | Hydrophobic |
| I57V | Resides within a known T cell epitope (IEDB identifiers;167808), Fold-X stability calculations suggested this mutation to be mildly destabilizing | Hydrophobic |
| L63F | Resides within a known T cell epitope (IEDB identifiers;167808), Fold-X stability calculations suggested this mutation to be mildly destabilizing | Hydrophobic |
| N72S | Resides within a discontinuous B cell epitope(IEDB identifiers; 504135), Fold-X stability calculations suggested this mutation to be mildly destabilizing | Polar uncharged |
| I84V | Resides within a known T cell epitope (IEDB identifiers;195227), Fold-X stability calculations suggested this mutation to be mildly destabilizing | Hydrophobic |
| N90D | Resides within a known T cell epitope (IEDB identifiers;195227), Fold-X stability calculations suggested this mutation to be mildly stabilizing | Polar uncharged to negatively charged |
| F98Y | Aromatic amino acid substitution | Hydrophobic |

In FoldX stability calculations, substitutions with ddG >−2 and <= −0.5 are labelled as mildly stabilizing, while those with ddG >= 0.5 and <2 are labelled as mildly destabilizing. Substitutions with ddG <= −2 or >= 2 were labelled as stabilizing and destabilizing respectively.

**Table S4. Comparison of multiple EDIIIs within each serotype with the EDIII consensus sequences used to design DENV DNA vaccine**

| DENV DNA vaccine EDIIIs | Specific DENV strain (n=1000) | Identity (%) |
|---|---|---|
| **EDIII-1** | Global DENV 1 strains | 95.15 (range: 66.99-100) |
| **EDIII-2** | Global DENV 2 strains | 97.09 (range: 93.20-100) |
| **EDIII-3** | Global DENV 3 strains | 97.09 (range: 59.22-100) |
| **EDIII-4** | Global DENV 4 strains | 98.06 (range: 86.41-100) |

DENV EDIII sequences used in designing DDV were aligned with the global dengue sequences for each serotype (n=1000) in the VIPR database to obtain the range of percent identity within members of the given serotype.

**Table S5a-d. B-cell structural epitopes predicted for the constructs across serotypes.**

DiscoTope tool implemented in the IEDB website was used for the analysis.

**Table S5a.** Predictions for the Serotype 1 construct sequence

| Residue ID | Residue Name | Contact Number | Propensity Score | Discotope Score |
|---|---|---|---|---|
| 317 | HIS | 14 | -0.622 | -7.622 |
| 342 | GLU | 16 | 0.822 | -7.178 |
| 343 | LYS | 10 | 1.189 | -3.811 |
| 344 | GLY | 11 | 1.67 | -3.83 |
| 345 | VAL | 13 | 1.881 | -4.619 |
| 346 | THR | 13 | 1.482 | -5.018 |
| 347 | GLN | 14 | 2.237 | -4.763 |
| 348 | ASN | 11 | 1.448 | -4.052 |
| 349 | GLY | 14 | 1.247 | -5.753 |
| 362 | GLU | 12 | 1.659 | -4.341 |
| 363 | LYS | 13 | 0.749 | -5.751 |
| 374 | GLY | 13 | -0.101 | -6.601 |
| 394 | LYS | 12 | 0.257 | -5.743 |
| 395 | GLY | 10 | 0.482 | -4.518 |
| 396 | SER | 11 | -0.217 | -5.717 |
| 397 | SER | 11 | -0.313 | -5.813 |

**Table S5b. Predictions for the Serotype 2 construct sequence**

| Chain ID | Residue ID | Residue Name | Contact Number | Propensity Score | Discotope Score |
|---|---|---|---|---|---|
| A | 311 | GLU | 11 | -1.429 | -6.929 |
| A | 318 | GLY | 16 | 0.341 | -7.659 |
| A | 327 | GLU | 12 | -0.658 | -6.658 |
| A | 329 | ASP | 13 | -1.091 | -7.591 |
| A | 342 | LEU | 13 | -0.803 | -7.303 |
| A | 343 | GLU | 9 | -0.548 | -5.048 |
| A | 345 | ARG | 9 | -0.09 | -4.59 |
| A | 346 | HIS | 12 | -0.739 | -6.739 |
| A | 360 | GLU | 15 | 1.117 | -6.383 |
| A | 362 | ASP | 12 | 1.48 | -4.52 |
| A | 363 | SER | 16 | 1.569 | -6.431 |
| A | 373 | PHE | 14 | 0.264 | -6.736 |
| A | 374 | GLY | 11 | 0.583 | -4.917 |
| A | 375 | ASP | 15 | 0.462 | -7.038 |
| A | 383 | GLU | 11 | -2.124 | -7.624 |
| A | 384 | PRO | 10 | -1.814 | -6.814 |

**Table S5c.** Predictions for the Serotype 3 construct sequence

| Chain ID | Residue ID | Residue Name | Contact Number | Propensity Score | Discotope Score |
|---|---|---|---|---|---|
| A | 317 | HIS | 14 | -0.622 | -7.622 |
| A | 342 | GLU | 16 | 0.822 | -7.178 |
| A | 343 | LYS | 10 | 1.189 | -3.811 |
| A | 344 | GLY | 11 | 1.67 | -3.83 |
| A | 345 | VAL | 13 | 1.881 | -4.619 |
| A | 346 | THR | 14 | 0.973 | -6.027 |
| A | 347 | GLN | 14 | 2.237 | -4.763 |
| A | 348 | ASN | 11 | 1.448 | -4.052 |
| A | 349 | GLY | 14 | 1.247 | -5.753 |
| A | 362 | GLU | 12 | 1.659 | -4.341 |
| A | 363 | LYS | 13 | 0.749 | -5.751 |
| A | 374 | GLY | 13 | -0.101 | -6.601 |

**Table S5d.** Predictions for the Serotype 4 construct sequence

| Residue ID | Residue Name | Contact Number | Propensity Score | Discotope Score |
|---|---|---|---|---|
| 313 | THR | 12 | 0.558 | -5.442 |
| 314 | GLN | 8 | 0.512 | -3.488 |
| 315 | HIS | 8 | 0.238 | -3.762 |

| | | | | |
|---|---|---|---|---|
| 316 | GLY | 11 | 0.538 | -4.962 |
| 340 | VAL | 16 | 1.087 | -6.913 |
| 341 | ASN | 10 | 0.871 | -4.129 |
| 342 | LYS | 13 | 0.693 | -5.807 |
| 343 | GLU | 13 | 1.073 | -5.427 |
| 344 | LYS | 15 | 0.508 | -6.992 |
| 345 | VAL | 17 | -0.678 | -9.178 |
| 346 | VAL | 11 | -0.114 | -5.614 |
| 347 | GLY | 10 | -0.547 | -5.547 |
| 352 | SER | 9 | -1.881 | -6.381 |
| 371 | PHE | 14 | 1.85 | -5.15 |
| 372 | GLY | 13 | 1.766 | -4.734 |
| 373 | ASP | 17 | 1.623 | -6.877 |
| 382 | ASN | 8 | -3.355 | -7.355 |

**Table S6a-d. Predicted T-cell epitopes (MHCI) and the corresponding HLA alleles**

We predicted the MHCI epitopes from the four serotype EDIII construct sequence using the NetMHCpan (v4.0) implemented in IEDB.

**Table S6a.** Predictions for the Serotype 1 construct sequence

| **MHCI epitope** | **HLA allele** |
|---|---|
| AETQHGTVL | HLA-B*40:01, HLA-B*44:02, HLA-B*44:03 |
| AETQHGTVLV | HLA-B*40:01 |
| AGEKALKLSW | HLA-B*44:02, HLA-B*44:03, HLA-B*57:01 |
| ALKLSWFKK | HLA-A*03:01, HLA-A*30:01 |
| CTGSFKLEK | HLA-A*11:01 |
| EPPFGESYI | HLA-B*51:01 |
| ETEPPFGESY | HLA-A*01:01, HLA-A*26:01, HLA-B*44:02, HLA-B*44:03 |
| ETQHGTVLV | HLA-A*68:02 |
| EVAETQHGT | HLA-A*68:02 |
| EVAETQHGTV | HLA-A*68:02 |
| GEKALKLSW | HLA-B*44:02, HLA-B*44:03 |
| GEKALKLSWF | HLA-B*44:02, HLA-B*44:03 |
| GTVLVQVKY | HLA-A*01:01, HLA-A*26:01, HLA-A*30:02, HLA-A*32:01, HLA-B*15:01, HLA-B*58:01 |
| HGTVLVQVKY | HLA-A*26:01 |
| ITANPIVTDK | HLA-A*03:01, HLA-A*11:01, HLA-A*30:01, HLA-A*68:01 |
| KALKLSWFK | HLA-A*03:01, HLA-A*11:01, HLA-A*30:01, HLA-A*31:01 |
| RLITANPIV | HLA-A*02:01, HLA-A*02:03, HLA-A*02:06 |
| SYVMCTGSF | HLA-A*23:01, HLA-A*24:02 |
| TANPIVTDK | HLA-A*11:01, HLA-A*68:01 |
| TEPPFGESY | HLA-B*44:02, HLA-B*44:03 |
| TQHGTVLVQV | HLA-A*02:06 |
| TQNGRLITA | HLA-A*02:03, HLA-A*02:06 |
| VAETQHGTVL | HLA-B*40:01 |
| YVMCTGSF | HLA-A*11:01 |

**Table S6b.** Predictions for the Serotype 2 construct sequence

| **MHCI epitope** | **HLA allele** |
|---|---|
| AEPPFGDSY | HLA-B*44:02, HLA-B*44:03 |
| AETQHGTIV | HLA-B*40:01, HLA-B*44:02, HLA-B*44:03 |
| AETQHGTIVV | HLA-B*40:01 |
| DLEKRHVLGR | HLA-A*33:01 |
| EAEPPFGDSY | HLA-A*01:01, HLA-A*26:01, HLA-B*35:0, HLA-B*44:02, HLA-B*44:03 |
| EIAETQHGT | HLA-A*68:02 |
| EIAETQHGTI | HLA-A*68:02 |
| EIMDLEKRHV | HLA-A*68:02 |
| EPGQLKLSW | HLA-B*53:01 |
| EPGQLKLSWF | HLA-B*53:01 |
| ETQHGTIVV | HLA-A*68:02 |
| ETQHGTIVVR | HLA-A*33:01, HLA-A*68:01 |
| GQLKLSWFK | HLA-A*03:01, HLA-A*11:01 |
| GTIVVRVQY | HLA-A*11:01, HLA-A*26:0, HLA-A*30:0, HLA-A*32:0, HLA-B*15:0, HLA-B*57:0, HLA-B*58:01 |
| HGTIVVRVQY | HLA-A*26:01 |

| | |
|---|---|
| HVLGRLITV | HLA-A*02:01, HLA-A*02:03, HLA-A*02:06, HLA-A*32:01, HLA-A*68:01, HLA-B*08:01 |
| IIGVEPGQLK | HLA-A*03:01, HLA-A*11:01 |
| IMDLEKRHVL | HLA-B*08:01 |
| ITVNPIVTEK | HLA-A*03:01, HLA-A*11:01, HLA-A*30:01, HLA-A*68:01 |
| MDLEKRHVL | HLA-B*08:01 |
| MSYSMCTGK | HLA-A*03:01, HLA-A*11:01 |
| QLKLSWFKK | HLA-A*03:01, HLA-A*30:01 |
| RHVLGRLITV | HLA-A*02:06 |
| RLITVNPIV | HLA-A*02:01, HLA-A*02:03, HLA-A*02:06 |
| SPCKIPFEI | HLA-B*07:02, HLA-B*51:01, HLA-B*53:01 |
| SPCKIPFEIM | HLA-B*07:02 |
| SYSMCTGKF | HLA-A*23:01, HLA-A*24:02 |
| TEKDSPVNI | HLA-B*40:01, HLA-B*44:02, HLA-B*44:03 |
| TQHGTIVVR | HLA-A*31:01 |
| TVNPIVTEK | HLA-A*03:01, HLA-A*11:01, HLA-A*30:01, HLA-A*31:01, HLA-A*33:01, HLA-A*68:01 |
| VEPGQLKLSW | HLA-B*44:02, HLA-B*44:03, HLA-B*53:01 |
| YEGDGSPCKI | HLA-B*40:01 |

**Table S6c.** Predictions for the Serotype 3 construct sequence

| **MHCI epitope** | **HLA allele** |
|---|---|
| AEPPFGESNI | HLA-B*40:01 |
| ALKINWYKK | HLA-A*03:01, HLA-A*30:01 |
| AMCTNTFVLK | HLA-A*03:01 |
| CTNTFVLKK | HLA-A*03:01, HLA-A*11:01 |
| ETQHGTILI | HLA-A*68:02 |
| ETQHGTILIK | HLA-A*68:01 |
| EVSETQHGT | HLA-A*68:02 |
| EVSETQHGTI | HLA-A*68:02 |
| GDNALKINW | HLA-B*44:02, HLA-B*44:03 |
| GEDAPCKIPF | HLA-B*40:01, HLA-B*44:02, HLA-B*44:03 |
| GESNIVIGI | HLA-B*40:01, HLA-B*44:02, HLA-B*44:03 |
| GTILIKVEY | HLA-A*01:01, HLA-A*11:01, HLA-A*26:01, HLA-A*30:02, HLA-A*32:01, HLA-B*15:01, HLA-B*57:01, HLA-B*58:01 |
| GTILIKVEYK | HLA-A*11:01 |
| HGTILIKVEY | HLA-A*26:01 |
| IGDNALKINW | HLA-B*57:01, HLA-B*58:01 |
| ITANPVVTK | HLA-A*03:01, HLA-A*11:01, HLA-A*30:01, HLA-A*68:01 |
| ITANPVVTKK | HLA-A*03:01, HLA-A*11:01, HLA-A*30:01, HLA-A*68:01 |
| KEEPVNIEA | HLA-B*40:01 |
| LITANPVVTK | HLA-A*03:01, HLA-A*11:01, HLA-A*68:01 |
| MCTNTFVLKK | HLA-A*03:01, HLA-A*11:01 |
| MSYAMCTNTF | HLA-A*23:01, HLA-A*24:02, HLA-B*57:01, HLA-B*58:01 |
| NALKINWYK | HLA-A*11:01, HLA-A*33:01, HLA-A*68:01 |
| NTFVLKKEV | HLA-A*68:02 |
| RLITANPVV | HLA-A*02:01, HLA-A*02:03, HLA-A*02:06 |
| SETQHGTIL | HLA-B*40:01, HLA-B*44:02, HLA-B*44:03 |
| SETQHGTILI | HLA-B*40:01, HLA-B*44:02 |
| SYAMCTNTF | HLA-A*23:01, HLA-A*24:02 |
| TANPVVTKK | HLA-A*03:01, HLA-A*11:01, HLA-A*30:01, HLA-A*68:01 |
| TILIKVEYK | HLA-A*03:01, HLA-A*11:01 |
| TQHGTILIK | HLA-A*03:01, HLA-A*11:01 |
| VLKKEVSET | HLA-A*02:03 |
| VSETQHGTIL | HLA-B*40:01 |

**Table S6d.** Predictions for the Serotype 4 construct sequence

| **MHCI epitope** | **HLA allele** |
|---|---|
| AENTNSVTSI | HLA-B*40:01, HLA-B*44:02, HLA-B*44:03 |
| AETQHGTTV | HLA-B*40:01, HLA-B*44:02, HLA-B*44:03 |
| AETQHGTTVV | HLA-B*40:01 |
| ALTLHWFRK | HLA-A*03:01 |
| APCKVPIEI | HLA-B*07:02, HLA-B*51:01, HLA-B*53:01 |
| DSALTLHWF | HLA-A*26:01 |
| DSALTLHWFR | HLA-A*33:01, HLA-A*68:01 |
| DVNKEKVVGR | HLA-A*33:01, HLA-A*68:01 |
| EIRDVNKEK | HLA-A*68:01 |

| | |
|---|---|
| ELEPPFGDSY | HLA-A*01:01, HLA-A*26:01, HLA-B*44:02, HLA-B*44:03 |
| ETQHGTTVV | HLA-A*68:02 |
| ETQHGTTVVK | HLA-A*68:01 |
| GDSALTLHW | HLA-B*44:02, HLA-B*44:03 |
| GTTVVKVKY | HLA-A*01:01, HLA-A*30:02 |
| KEKVVGRII | HLA-B*44:02 |
| LEPPFGDSY | HLA-B*44:02, HLA-B*44:03 |
| NTNSVTSIEL | HLA-A*68:02 |
| RIISSTPFA | HLA-A*02:06, HLA-A*30:01 |
| SALTLHWFR | HLA-A*31:01, HLA-A*33:01, HLA-A*68:01 |
| SYTMCSGKF | HLA-A*23:01, HLA-A*24:02 |
| TPFAENTNSV | HLA-B*51:01 |
| TQHGTTVVK | HLA-A*11:01, HLA-A*30:01 |
| TSIELEPPF | HLA-B*35:01, HLA-B*53:01, HLA-B*58:01 |
| VGDSALTLHW | HLA-B*57:01, HLA-B*58:01 |
| VNKEKVVGR | HLA-A*31:01, HLA-A*33:01 |

**Table S7a-d.** Predicted T-cell epitopes (MHCII) and the corresponding HLA alleles

We predicted the MHCII epitopes from the four serotype EDIII construct sequence using combination of NetMHCIIpan 4.0, NN-align 2.3 and SMMalign (MHCII) implemented in IEDB.

**Supplementary Table 7a: Predictions for the Serotype 1 construct sequence**

| **MHCII epitope** | **HLA allele** |
|---|---|
| ESYIVIGAGEKALKL | HLA-DRB5*01:01 |
| FGESYIVIGAGEKAL | HLA-DRB5*01:01 |
| GESYIVIGAGEKALK | HLA-DRB5*01:01 |
| GVSYVMCTGSFKLEK | HLA-DRB5*01:01 |
| KGVSYVMCTGSFKLE | HLA-DRB5*01:01 |
| PFGESYIVIGAGEKA | HLA-DRB5*01:01 |
| PPFGESYIVIGAGEK | HLA-DRB5*01:01 |
| SYIVIGAGEKALKLS | HLA-DRB5*01:01 |
| SYVMCTGSFKLEKEV | HLA-DRB5*01:01 |
| VSYVMCTGSFKLEKE | HLA-DRB5*01:01 |
| YIVIGAGEKALKLSW | HLA-DRB5*01:01 |
| YVMCTGSFKLEKEVA | HLA-DRB5*01:01 |
| AETQHGTVLVQVKYE | HLA-DQA1*01:02/DQB1*06:02 |
| AGEKALKLSWFKKGS | HLA-DPA1*02:01/DPB1*05:01 |
| ANPIVTDKEKPVNIE | HLA-DRB1*13:02, HLA-DRB3*01:01 |
| DKEKPVNIETEPPFG | HLA-DQA1*03:01/DQB1*03:02 |
| EKPVNIETEPPFGES | HLA-DQA1*03:01/DQB1*03:02 |
| EPPFGESYIVIGAGE | HLA-DPA1*01:03/DPB1*02:01 |
| ESYIVIGAGEKALKL | HLA-DQA1*05:01/DQB1*03:01 |
| ETEPPFGESYIVIGA | HLA-DPA1*01:03/DPB1*02:01 |
| ETQHGTVLVQVKYEG | HLA-DQA1*01:02/DQB1*06:02 |
| EVAETQHGTVLVQVK | HLA-DQA1*01:02/DQB1*06:02 |
| FGESYIVIGAGEKAL | HLA-DQA1*05:01/DQB1*03:01 |
| GAGEKALKLSWFKKG | HLA-DPA1*02:01/DPB1*05:01 |
| GEKALKLSWFKKGSS | HLA-DPA1*02:01/DPB1*05:01 |
| GESYIVIGAGEKALK | HLA-DQA1*05:01/DQB1*03:01 |
| GRLITANPIVTDKEK | HLA-DQA1*04:01/DQB1*04:02, HLA-DRB1*13:02, HLA-DRB3*02:02 |
| GVSYVMCTGSFKLEK | HLA-DRB1*07:01 |
| GVTQNGRLITANPIV | HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*13:02, HLA-DRB3*02:02, HLA-DRB4*01:01 |
| IETEPPFGESYIVIG | HLA-DPA1*01:03/DPB1*02:01 |
| IGAGEKALKLSWFKK | HLA-DPA1*02:01/DPB1*05:01 |
| IVIGAGEKALKLSWF | HLA-DQA1*05:01/DQB1*03:01 |
| IVTDKEKPVNIETEP | HLA-DQA1*03:01/DQB1*03:02 |
| KEKPVNIETEPPFGE | HLA-DQA1*03:01/DQB1*03:02 |
| KEVAETQHGTVLVQV | HLA-DQA1*01:02/DQB1*06:02 |
| KGVSYVMCTGSFKLE | HLA-DRB1*07:01 |
| KGVTQNGRLITANPI | HLA-DRB1*13:02 |
| LITANPIVTDKEKPV | HLA-DRB1*13:02 |
| NGRLITANPIVTDKE | HLA-DPA1*02:01/DPB1*14:01, HLA-DQA1*04:01/DQB1*04:02, HLA-DRB1*13:02, HLA-DRB3*02:02 |
| NIETEPPFGESYIVI | HLA-DPA1*01:03/DPB1*02:01 |
| NPIVTDKEKPVNIET | HLA-DRB1*13:02 |
| PFGESYIVIGAGEKA | HLA-DQA1*05:01/DQB1*03:01 |

| | |
|---|---|
| PIVTDKEKPVNIETE | HLA-DQA1*03:01/DQB1*03:02 |
| QHGTVLVQVKYEGTD | HLA-DQA1*01:02/DQB1*06:02 |
| QNGRLITANPIVTDK | HLA-DPA1*02:01/DPB1*14:01, HLA-DQA1*04:01/DQB1*04:02, HLA-DRB1*13:02, HLA-DRB3*02:02, HLA-DRB4*01:01 |
| RLITANPIVTDKEKP | HLA-DQA1*04:01/DQB1*04:02, HLA-DRB3*02:02 |
| SYIVIGAGEKALKLS | HLA-DQA1*05:01/DQB1*03:01 |
| SYVMCTGSFKLEKEV | HLA-DRB1*07:01 |
| TANPIVTDKEKPVNI | HLA-DRB1*13:02 |
| TDKEKPVNIETEPPF | HLA-DQA1*03:01/DQB1*03:02 |
| TEPPFGESYIVIGAG | HLA-DPA1*01:03/DPB1*02:01 |
| TQHGTVLVQVKYEGT | HLA-DQA1*01:02/DQB1*06:02 |
| TQNGRLITANPIVTD | HLA-DPA1*02:01/DPB1*14:01, HLA-DQA1*04:01/DQB1*04:02, HLA-DRB1*13:02, HLA-DRB3*02:02, HLA-DRB4*01:01 |
| VAETQHGTVLVQVKY | HLA-DQA1*01:02/DQB1*06:02 |
| VSYVMCTGSFKLEKE | HLA-DRB1*07:01 |
| VTDKEKPVNIETEPP | HLA-DQA1*03:01/DQB1*03:02 |
| VTQNGRLITANPIVT | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB1*13:02, HLA-DRB3*02:02, HLA-DRB4*01:01 |
| YIVIGAGEKALKLSW | HLA-DQA1*05:01/DQB1*03:01 |

**Table S7b.** Predictions for the Serotype 2 construct sequence

| **MHCII epitope** | **HLA allele** |
|---|---|
| DSPVNIEAEPPFGDS | HLA-DQA1*03:01, HLA-DQB1*03:02 |
| EKDSPVNIEAEPPFG | HLA-DQA1*03:01, HLA-DQB1*03:02 |
| IVTEKDSPVNIEAEP | HLA-DQA1*03:01, HLA-DQB1*03:02 |
| KDSPVNIEAEPPFGD | HLA-DQA1*03:01, HLA-DQB1*03:02 |
| PIVTEKDSPVNIEAE | HLA-DQA1*03:01, HLA-DQB1*03:02 |
| TEKDSPVNIEAEPPF | HLA-DQA1*03:01, HLA-DQB1*03:02 |
| VTEKDSPVNIEAEPP | HLA-DQA1*03:01, HLA-DQB1*03:02 |
| AEPPFGDSYIIIGVE | HLA-DQA1*04:01/DQB1*04:02 |
| AETQHGTIVVRVQYE | HLA-DQA1*01:02/DQB1*06:02 |
| CKIPFEIMDLEKRHV | HLA-DPA1*02:01/DPB1*05:01, HLA-DPA1*03:01/DPB1*04:02 |
| DLEKRHVLGRLITVN | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB1*07:01 |
| DSYIIIGVEPGQLKL | HLA-DQA1*03:01/DQB1*03:02, HLA-DQA1*05:01/DQB1*02:01, HLA-DRB1*04:01, HLA-DRB1*08:02, HLA-DRB1*13:02 |
| EIAETQHGTIVVRVQ | HLA-DQA1*01:02/DQB1*06:02 |
| EIMDLEKRHVLGRLI | HLA-DRB1*07:01, HLA-DRB1*11:01 |
| EKDSPVNIEAEPPFG | HLA-DQA1*04:01/DQB1*04:02 |
| EKRHVLGRLITVNPI | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB1*07:01 |
| EPGQLKLSWFKKGSS | HLA-DPA1*02:01/DPB1*05:01 |
| EPPFGDSYIIIGVEP | HLA-DQA1*04:01/DQB1*04:02 |
| ETQHGTIVVRVQYEG | HLA-DQA1*01:02/DQB1*06:02 |
| FEIMDLEKRHVLGRL | HLA-DRB1*11:01 |
| FGDSYIIIGVEPGQL | HLA-DQA1*03:01/DQB1*03:02, HLA-DQA1*04:01/DQB1*04:02, HLA-DQA1*05:01/DQB1*02:01, HLA-DRB1*13:02 |
| GDSYIIIGVEPGQLK | HLA-DQA1*03:01/DQB1*03:02, HLA-DQA1*05:01/DQB1*02:01, HLA-DRB1*04:01, HLA-DRB1*08:02, HLA-DRB1*13:02 |
| GKFKVVKEIAETQHG | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB1*04:01, HLA-DRB1*08:02 |
| GMSYSMCTGKFKVVK | HLA-DRB5*01:01 |
| GRLITVNPIVTEKDS | HLA-DPA1*02:01/DPB1*14:01, HLA-DQA1*04:01/DQB1*04:02, HLA-DRB1*04:05, HLA-DRB1*08:02, HLA-DRB1*13:02, HLA-DRB3*02:02 |
| GVEPGQLKLSWFKKG | HLA-DPA1*02:01/DPB1*05:01 |
| HGTIVVRVQYEGDGS | HLA-DQA1*03:01/DQB1*03:02 |
| HVLGRLITVNPIVTE | HLA-DPA1*02:01/DPB1*14:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DQA1*04:01/DQB1*04:02, HLA-DRB1*08:02, HLA-DRB1*13:02, HLA-DRB3*02:02, HLA-DRB4*01:01 |
| IAETQHGTIVVRVQY | HLA-DQA1*01:02/DQB1*06:02 |
| IMDLEKRHVLGRLIT | HLA-DRB1*07:01, HLA-DRB1*11:01 |
| IPFEIMDLEKRHVLG | HLA-DPA1*03:01/DPB1*04:02, HLA-DRB1*11:01 |
| IVTEKDSPVNIEAEP | HLA-DQA1*04:01/DQB1*04:02 |
| KEIAETQHGTIVVRV | HLA-DQA1*01:02/DQB1*06:02 |
| KGMSYSMCTGKFKVV | HLA-DRB5*01:01 |
| KIPFEIMDLEKRHVL | HLA-DPA1*02:01/DPB1*05:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DRB1*11:01, HLA-DRB5*01:01 |
| KRHVLGRLITVNPIV | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB1*07:01, HLA-DRB4*01:01 |
| LEKRHVLGRLITVNP | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB1*07:01 |
| LGRLITVNPIVTEKD | HLA-DPA1*02:01/DPB1*14:01, HLA-DQA1*04:01/DQB1*04:02, HLA-DRB1*04:05, HLA-DRB1*08:02, HLA-DRB1*13:02, HLA-DRB3*02:02 |
| MDLEKRHVLGRLITV | HLA-DRB1*07:01 |
| MSYSMCTGKFKVVKE | HLA-DRB5*01:01 |
| NPIVTEKDSPVNIEA | HLA-DRB1*13:02 |
| PCKIPFEIMDLEKRH | HLA-DPA1*03:01/DPB1*04:02 |

| | |
|---|---|
| PFEIMDLEKRHVLGR | HLA-DRB1*11:01 |
| PFGDSYIIIGVEPGQ | HLA-DQA1*03:01/DQB1*03:02, HLA-DQA1*04:01/DQB1*04:02 |
| PIVTEKDSPVNIEAE | HLA-DQA1*04:01/DQB1*04:02, HLA-DRB1*13:02 |
| PPFGDSYIIIGVEPG | HLA-DQA1*03:01/DQB1*03:02, HLA-DQA1*04:01/DQB1*04:02 |
| RHVLGRLITVNPIVT | HLA-DPA1*02:01/DPB1*14:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DRB1*08:02, HLA-DRB1*13:02, HLA-DRB3*02:02, HLA-DRB4*01:01 |
| RLITVNPIVTEKDSP | HLA-DQA1*04:01/DQB1*04:02, HLA-DRB3*02:02 |
| SPCKIPFEIMDLEKR | HLA-DPA1*03:01/DPB1*04:02 |
| SYIIIGVEPGQLKLS | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB1*08:02, HLA-DRB1*13:02 |
| SYSMCTGKFKVVKEI | HLA-DRB5*01:01 |
| TEKDSPVNIEAEPPF | HLA-DQA1*04:01/DQB1*04:02 |
| TGKFKVVKEIAETQH | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB1*04:01, HLA-DRB1*08:02 |
| TQHGTIVVRVQYEGD | HLA-DQA1*01:02/DQB1*06:02 |
| TVNPIVTEKDSPVNI | HLA-DRB1*13:02 |
| VEPGQLKLSWFKKGS | HLA-DPA1*02:01/DPB1*05:01 |
| VLGRLITVNPIVTEK | HLA-DPA1*02:01/DPB1*14:01, HLA-DQA1*04:01/DQB1*04:02, HLA-DRB1*08:02, HLA-DRB1*13:02, HLA-DRB3*02:02, HLA-DRB4*01:01, HLA-DRB1*13:02 |
| VNPIVTEKDSPVNIE | HLA-DRB1*13:02 |
| VTEKDSPVNIEAEPP | HLA-DQA1*04:01/DQB1*04:02 |
| YIIIGVEPGQLKLSW | HLA-DRB1*08:02, HLA-DRB1*13:02 |
| YSMCTGKFKVVKEIA | HLA-DRB5*01:01 |

**Table S7c.** Predictions for the Serotype 3 construct sequence

| **MHCII epitope** | **HLA allele** |
|---|---|
| NGRLITANPVVTKK | HLA-DRB3*02:02 |
| GRLITANPVVTKK | HLA-DRB3*02:02 |
| GRLITANPVVTKKE | HLA-DRB3*02:02 |
| HNGRLITANPVVTKK | HLA-DRB3*02:02 |
| NGRLITANPVVTKKE | HLA-DRB3*02:02 |
| GRLITANPVVTKKEE | HLA-DRB3*02:02 |
| AHNGRLITANPVVTKK | HLA-DRB3*02:02 |
| HNGRLITANPVVTKKE | HLA-DRB3*02:02 |
| NGRLITANPVVTKKEE | HLA-DRB3*02:02 |
| HNGRLITANPVVTK | HLA-DRB3*02:02 |
| NGRLITANPVVTK | HLA-DRB3*02:02 |
| RLITANPVVTKKE | HLA-DRB3*02:02 |
| AHNGRLITANPVVTK | HLA-DRB3*02:02 |
| RLITANPVVTKK | HLA-DRB3*02:02 |
| RLITANPVVTKKEE | HLA-DRB3*02:02 |
| GRLITANPVVTKKEEP | HLA-DRB3*02:02 |
| GRLITANPVVTK | HLA-DRB3*02:02 |
| AHNGRLITANPVVTKKE | HLA-DRB3*02:02 |
| KAHNGRLITANPVVTKK | HLA-DRB3*02:02 |
| HNGRLITANPVVTKKEE | HLA-DRB3*02:02 |
| NGRLITANPVVTKKEEP | HLA-DRB3*02:02 |
| AHNGRLITANPVVTKKEE | HLA-DRB3*02:02 |
| HNGRLITANPVVTKKEEP | HLA-DRB3*02:02 |
| KAHNGRLITANPVVTKKE | HLA-DRB3*02:02 |
| GKAHNGRLITANPVVTKK | HLA-DRB3*02:02 |
| SYAMCTNTFVLK | HLA-DRB3*02:02 |
| AHNGRLITANPVV | HLA-DRB3*02:02 |
| AHNGRLITANPVVT | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| AHNGRLITANPVVTK | HLA-DPA1*02:01/DPB1*14:01 |
| AHNGRLITANPVVTKK | HLA-DPA1*02:01/DPB1*14:01 |
| AHNGRLITANPVVTKKE | HLA-DPA1*02:01/DPB1*14:01 |
| AHNGRLITANPVVTKKEE | HLA-DPA1*02:01/DPB1*14:01 |
| ALKINWYKKGSS | HLA-DRB3*02:02 |
| GKAHNGRLITANPVVT | HLA-DRB3*02:02 |
| GKAHNGRLITANPVVTK | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| GKAHNGRLITANPVVTKK | HLA-DPA1*02:01/DPB1*14:01 |
| GMSYAMCTNTFV | HLA-DRB3*02:02 |
| GMSYAMCTNTFVL | HLA-DRB3*02:02 |
| GMSYAMCTNTFVLK | HLA-DRB3*02:02 |
| GMSYAMCTNTFVLKK | HLA-DRB3*02:02 |
| GMSYAMCTNTFVLKKE | HLA-DRB3*02:02 |
| GMSYAMCTNTFVLKKEV | HLA-DRB3*02:02 |
| GMSYAMCTNTFVLKKEVS | HLA-DRB3*02:02 |
| GRLITANPVVTK | HLA-DPA1*02:01/DPB1*14:01 |

| | |
|---|---|
| GRLITANPVVTKK | HLA-DPA1*02:01/DPB1*14:01 |
| GRLITANPVVTKKE | HLA-DPA1*02:01/DPB1*14:01 |
| GRLITANPVVTKKEE | HLA-DPA1*02:01/DPB1*14:01 |
| GRLITANPVVTKKEEP | HLA-DPA1*02:01/DPB1*14:01 |
| GRLITANPVVTKKEEPV | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| GRLITANPVVTKKEEPVN | HLA-DRB3*02:02 |
| HNGRLITANPVV | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| HNGRLITANPVVT | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| HNGRLITANPVVTK | HLA-DPA1*02:01/DPB1*14:01 |
| HNGRLITANPVVTKK | HLA-DPA1*02:01/DPB1*14:01 |
| HNGRLITANPVVTKKE | HLA-DPA1*02:01/DPB1*14:01 |
| HNGRLITANPVVTKKEE | HLA-DPA1*02:01/DPB1*14:01 |
| HNGRLITANPVVTKKEEP | HLA-DPA1*02:01/DPB1*14:01 |
| KAHNGRLITANPVV | HLA-DRB3*02:02 |
| KAHNGRLITANPVVT | HLA-DRB3*02:02 |
| KAHNGRLITANPVVTK | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| KAHNGRLITANPVVTKK | HLA-DPA1*02:01/DPB1*14:01 |
| KAHNGRLITANPVVTKKE | HLA-DPA1*02:01/DPB1*14:01 |
| KGMSYAMCTNTFV | HLA-DRB3*02:02 |
| KGMSYAMCTNTFVL | HLA-DRB3*02:02 |
| KGMSYAMCTNTFVLK | HLA-DRB3*02:02 |
| KGMSYAMCTNTFVLKK | HLA-DRB3*02:02 |
| KGMSYAMCTNTFVLKKE | HLA-DRB3*02:02 |
| KGMSYAMCTNTFVLKKEV | HLA-DRB3*02:02 |
| LITANPVVTKKE | HLA-DRB3*02:02 |
| LITANPVVTKKEE | HLA-DRB3*02:02 |
| LITANPVVTKKEEP | HLA-DRB3*02:02 |
| LITANPVVTKKEEPV | HLA-DRB3*02:02 |
| MSYAMCTNTFVL | HLA-DRB3*02:02 |
| MSYAMCTNTFVLK | HLA-DRB3*02:02 |
| MSYAMCTNTFVLKK | HLA-DRB3*02:02 |
| MSYAMCTNTFVLKKE | HLA-DRB3*02:02 |
| MSYAMCTNTFVLKKEV | HLA-DRB3*02:02 |
| MSYAMCTNTFVLKKEVS | HLA-DRB3*02:02 |
| MSYAMCTNTFVLKKEVSE | HLA-DRB3*02:02 |
| NALKINWYKKGS | HLA-DRB3*02:02 |
| NALKINWYKKGSS | HLA-DRB3*02:02 |
| NGRLITANPVVT | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| NGRLITANPVVTK | HLA-DPA1*02:01/DPB1*14:01 |
| NGRLITANPVVTKK | HLA-DPA1*02:01/DPB1*14:01 |
| NGRLITANPVVTKKE | HLA-DPA1*02:01/DPB1*14:01 |
| NGRLITANPVVTKKEE | HLA-DPA1*02:01/DPB1*14:01 |
| NGRLITANPVVTKKEEP | HLA-DPA1*02:01/DPB1*14:01 |
| NGRLITANPVVTKKEEPV | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| QGKAHNGRLITANPVVT | HLA-DRB3*02:02 |
| QGKAHNGRLITANPVVTK | HLA-DRB3*02:02 |
| RLITANPVVTKK | HLA-DPA1*02:01/DPB1*14:01 |
| RLITANPVVTKKEEP | HLA-DRB3*02:02 |
| RLITANPVVTKKEEPV | HLA-DRB3*02:02 |
| RLITANPVVTKKEEPVN | HLA-DRB3*02:02 |
| RLITANPVVTKKEEPVNI | HLA-DRB3*02:02 |
| SYAMCTNTFVLK | HLA-DPA1*01:03/DPB1*04:01 |
| SYAMCTNTFVLKK | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| SYAMCTNTFVLKKE | HLA-DRB3*02:02 |
| SYAMCTNTFVLKKEV | HLA-DRB3*02:02 |
| SYAMCTNTFVLKKEVS | HLA-DRB3*02:02 |
| SYAMCTNTFVLKKEVSE | HLA-DRB3*02:02 |
| YAMCTNTFVLKK | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| YAMCTNTFVLKKE | HLA-DRB3*02:02 |
| YAMCTNTFVLKKEV | HLA-DRB3*02:02, |

**Table S7d.** Predictions for the Serotype 4 construct sequence

| **MHCII epitope** | **HLA allele** |
|---|---|
| VGRIISSTPFAENT | HLA-DPA1*02:01, HLA-DPB1*14:01 |
| SSTPFAENTNSVTSI | HLA-DRB3*02:02 |
| STPFAENTNSVTSIE | HLA-DRB3*02:02 |
| ISSTPFAENTNSVTSI | HLA-DRB3*02:02 |
| SSTPFAENTNSVTSIE | HLA-DRB3*02:02 |

| | |
|---|---|
| STPFAENTNSVTSIEL | HLA-DRB3*02:02 |
| VVGRIISSTPFAENT | HLA-DPA1*02:01, HLA-DPB1*14:01 |
| STPFAENTNSVTSI | HLA-DRB3*02:02 |
| VVGRIISSTPFAEN | HLA-DPA1*02:01, HLA-DPB1*14:01 |
| GRIISSTPFAENT | HLA-DPA1*02:01, HLA-DPB1*14:01 |
| VGRIISSTPFAEN | HLA-DPA1*02:01, HLA-DPB1*14:01 |
| IISSTPFAENTNSVTSI | HLA-DRB3*02:02 |
| ISSTPFAENTNSVTSIE | HLA-DRB3*02:02 |
| SSTPFAENTNSVTSIEL | HLA-DRB3*02:02 |
| STPFAENTNSVTSIELE | HLA-DRB3*02:02 |
| GRIISSTPFAEN | HLA-DPA1*02:01, HLA-DPB1*14:01 |
| IISSTPFAENTNSVTSIE | HLA-DRB3*02:02 |
| ISSTPFAENTNSVTSIEL | HLA-DRB3*02:02 |
| RIISSTPFAENTNSVTSI | HLA-DRB3*02:02 |
| SSTPFAENTNSVTSIELE | HLA-DRB3*02:02 |
| RIISSTPFAENT | HLA-DPA1*02:01, HLA-DPB1*14:01 |
| VGRIISSTPFAE | HLA-DPA1*02:01, HLA-DPB1*14:01 |
| DVNKEKVVGRIISSTPFA | HLA-DPA1*02:01/DPB1*14:01 |
| EKVVGRIISSTPFA | HLA-DPA1*02:01/DPB1*14:01 |
| EKVVGRIISSTPFAE | HLA-DPA1*02:01/DPB1*14:01 |
| EKVVGRIISSTPFAEN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| EKVVGRIISSTPFAENT | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| EKVVGRIISSTPFAENTN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| GRIISSTPFAEN | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| GRIISSTPFAENT | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| GRIISSTPFAENTN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| GRIISSTPFAENTNS | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| GRIISSTPFAENTNSV | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| GRIISSTPFAENTNSVT | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| GRIISSTPFAENTNSVTS | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| IISSTPFAENTNSVT | HLA-DRB3*02:02 |
| IISSTPFAENTNSVTS | HLA-DRB3*02:02 |
| ISSTPFAENTNSVT | HLA-DRB3*02:02 |
| ISSTPFAENTNSVTS | HLA-DRB3*02:02 |
| KEKVVGRIISST | HLA-DPA1*02:01/DPB1*14:01 |
| KEKVVGRIISSTP | HLA-DPA1*02:01/DPB1*14:01 |
| KEKVVGRIISSTPF | HLA-DPA1*02:01/DPB1*14:01 |
| KEKVVGRIISSTPFA | HLA-DPA1*02:01/DPB1*14:01 |
| KEKVVGRIISSTPFAE | HLA-DPA1*02:01/DPB1*14:01 |
| KEKVVGRIISSTPFAEN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| KEKVVGRIISSTPFAENT | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| KVVGRIISSTPFA | HLA-DPA1*02:01/DPB1*14:01 |
| KVVGRIISSTPFAE | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| KVVGRIISSTPFAEN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| KVVGRIISSTPFAENT | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| KVVGRIISSTPFAENTN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| KVVGRIISSTPFAENTNS | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| NKEKVVGRIISS | HLA-DPA1*02:01/DPB1*14:01 |
| NKEKVVGRIISST | HLA-DPA1*02:01/DPB1*14:01 |
| NKEKVVGRIISSTPF | HLA-DPA1*02:01/DPB1*14:01 |
| NKEKVVGRIISSTPFA | HLA-DPA1*02:01/DPB1*14:01 |
| NKEKVVGRIISSTPFAE | HLA-DPA1*02:01/DPB1*14:01 |
| NKEKVVGRIISSTPFAEN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| PFAENTNSVTSI | HLA-DRB3*02:02 |
| PFAENTNSVTSIE | HLA-DRB3*02:02 |
| PFAENTNSVTSIEL | HLA-DRB3*02:02 |
| PFAENTNSVTSIELE | HLA-DRB3*02:02 |
| PFAENTNSVTSIELEP | HLA-DRB3*02:02 |
| PFAENTNSVTSIELEPP | HLA-DRB3*02:02 |
| PFAENTNSVTSIELEPPF | HLA-DRB3*02:02 |
| RIISSTPFAENT | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| RIISSTPFAENTN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| RIISSTPFAENTNS | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| RIISSTPFAENTNSV | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| RIISSTPFAENTNSVT | HLA-DRB3*02:02 |
| RIISSTPFAENTNSVTS | HLA-DRB3*02:02 |

| | |
|---|---|
| SSTPFAENTNSVT | HLA-DRB3*02:02 |
| SSTPFAENTNSVTS | HLA-DRB3*02:02 |
| STPFAENTNSVT | HLA-DRB3*02:02 |
| STPFAENTNSVTS | HLA-DRB3*02:02 |
| STPFAENTNSVTSIELEP | HLA-DRB3*02:02 |
| TPFAENTNSVTS | HLA-DRB3*02:02 |
| TPFAENTNSVTSI | HLA-DRB3*02:02 |
| TPFAENTNSVTSIE | HLA-DRB3*02:02 |
| TPFAENTNSVTSIEL | HLA-DRB3*02:02 |
| TPFAENTNSVTSIELE | HLA-DRB3*02:02 |
| TPFAENTNSVTSIELEP | HLA-DRB3*02:02 |
| TPFAENTNSVTSIELEPP | HLA-DRB3*02:02 |
| VGRIISSTPFAE | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| VGRIISSTPFAEN | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| VGRIISSTPFAENT | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| VGRIISSTPFAENTN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| VGRIISSTPFAENTNS | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| VGRIISSTPFAENTNSV | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| VGRIISSTPFAENTNSVT | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| VNKEKVVGRIISST | HLA-DPA1*02:01/DPB1*14:01 |
| VNKEKVVGRIISSTPF | HLA-DPA1*02:01/DPB1*14:01 |
| VNKEKVVGRIISSTPFA | HLA-DPA1*02:01/DPB1*14:01 |
| VNKEKVVGRIISSTPFAE | HLA-DPA1*02:01/DPB1*14:01 |
| VVGRIISSTPFA | HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| VVGRIISSTPFAE | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DRB3*02:02 |
| VVGRIISSTPFAEN | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| VVGRIISSTPFAENT | HLA-DPA1*01:03/DPB1*04:01, HLA-DRB3*02:02 |
| VVGRIISSTPFAENTN | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| VVGRIISSTPFAENTNS | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |
| VVGRIISSTPFAENTNSV | HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*02:01/DPB1*14:01 |

**Supplementary Table 8a: World population coverage of the T-cell epitopes**

T-cell epitopes were predicted from the dengue constructs for the four serotypes and the HLA allele coverage was used to predict the world population coverage. We have presented the projected population coverage, average number of epitope hits / HLA combinations recognized by the population, and minimum number of epitope hits / HLA combinations recognized by 90% of the population (PC90).

| Dengue serotype | MHC class | World population coverage | Average hit | PC90 |
|---|---|---|---|---|
| 1 | I | 97.34 | 4.27 | 1.48 |
| 1 | II | 99.43 | 21.52 | 12.07 |
| 2 | I | 98.37 | 6.15 | 2.32 |
| 2 | II | 97.94 | 24.89 | 7.66 |
| 3 | I | 96.55 | 6.82 | 1.43 |
| 3 | II | 90.16 | 13.63 | 3.01 |
| 4 | I | 90.28 | 3.22 | 1.01 |
| 4 | II | 90.16 | 52.49 | 34.07 |

**Supplementary Table 8b: Population coverage of the T-cell epitopes for DENV1 construct in regions with frequent and sporadic dengue occurrence**

The dengue endemic regions were identified as per CDC guidelines (https://www.cdc.gov/dengue/areaswithrisk/around-the-world.html). The DENV1 vaccine is estimated to be efficient in all populations except Central America which is predicted to have a lower coverage of MHCI HLA types capable of binding the epitopes in the sequence.

| Population | MHC class | Population coverage | Average hit | PC90 |
|---|---|---|---|---|
| South Asia | I | 93.15% | 3.88 | 1.23 |
| | II | 99.6% | 18.79 | 9.33 |
| Southeast Asia | I | 94.2% | 4.32 | 1.26 |
| | II | 95.8% | 13.9 | 4.79 |
| East Africa | I | 85.49% | 2.61 | 0.69 |
| | II | 99.68% | 23.21 | 12.89 |
| West Africa | I | 89.34% | 2.53 | 0.94 |
| | II | 99.86% | 25.63 | 15.91 |
| Central Africa | I | 75.5% | 2.41 | 0.41 |
| | II | 99.45% | 22.28 | 12.32 |
| West Indies | I | 96.97% | 4.11 | 1.44 |
| | II | 95.99% | 14.87 | 7.15 |
| Central America | I | 4.91% | 0.08 | 0.11 |
| | II | 99.96% | 21.12 | 14.37 |
| South America | I | 85.82% | 2.5 | 0.71 |
| | II | 99.98% | 26.91 | 17.65 |
| Oceania | I | 93.78% | 3.66 | 1.14 |
| | II | 99.9% | 24.24 | 15.28 |

**Supplementary Table 8c: Population coverage of the T-cell epitopes for DENV2 construct in regions with frequent and sporadic dengue occurrence**

The dengue endemic regions were identified as per CDC guidelines (https://www.cdc.gov/dengue/areaswithrisk/around-the-world.html). The DENV1 vaccine is estimated to be efficient in all populations except Central America which is predicted to have a lower coverage of MHCI HLA types recognising the epitopes in the sequence.

| Population | MHC class | Population coverage | Average hit | PC90 |
|---|---|---|---|---|
| South Asia | I | 94.34% | 5.19 | 1.47 |
| | II | 94.96% | 20.12 | 3.83 |
| Southeast Asia | I | 94.56% | 5.19 | 1.36 |
| | II | 94.24% | 14.04 | 3.9 |
| East Africa | I | 87.39% | 4.11 | 0.79 |
| | II | 99.72% | 30.85 | 13.59 |
| West Africa | I | 93.51% | 4.59 | 1.3 |
| | II | 99.46% | 28.52 | 12.67 |
| Central Africa | I | 83.03% | 3.95 | 0.59 |

| | | | | |
|---|---|---|---|---|
| | II | 99.51% | 28.65 | 13.19 |
| West Indies | I | 98.43% | 6.25 | 2.38 |
| | II | 95.0% | 17.02 | 3.83 |
| Central America | I | 6.44% | 0.2 | 0.11 |
| | II | 100.0% | 37.67 | 23.25 |
| South America | I | 87.37% | 3.67 | 0.79 |
| | II | 99.95% | 42.69 | 23.3 |
| Oceania | I | 94.57% | 4.65 | 1.21 |
| | II | 98.96% | 22.14 | 8.33 |

**Supplementary Table 8d: Population coverage of the T-cell epitopes for DENV3 construct in regions with frequent and sporadic dengue occurrence**

The dengue endemic regions were identified as per CDC guidelines (https://www.cdc.gov/dengue/areaswithrisk/around-the-world.html). The DENV1 vaccine is estimated to be efficient in all populations except Central America which is predicted to have a lower coverage of MHCI and MHCII HLA types recognising the epitopes in the sequence. We also predicted that the Southeast Asian population can have a lower coverage of MHCII HLA types.

| **Population** | **MHC class** | **Population coverage** | **Average hit** | **PC90** |
|---|---|---|---|---|
| South Asia | I | 91.02% | 6.55 | 1.08 |
| | II | 98.63% | 17.95 | 4.11 |
| Southeast Asia | I | 93.29% | 7.95 | 1.58 |
| | II | 19.27% | 2.56 | 0.37 |
| East Africa | I | 84.58% | 3.78 | 0.65 |
| | II | 81.1% | 13.81 | 1.59 |
| West Africa | I | 89.64% | 4.17 | 0.97 |
| | II | 90.4% | 23.74 | 3.07 |
| Central Africa | I | 75.58% | 3.71 | 0.41 |
| | II | 86.08% | 14.16 | 2.15 |
| West Indies | I | 96.98% | 6.24 | 1.54 |
| | II | - | - | - |
| Central America | I | 4.91% | 0.13 | 0.11 |
| | II | 26.04% | 6.62 | 0.41 |
| South America | I | 79.07% | 3.9 | 0.48 |
| | II | 92.48% | 21.81 | 3.23 |
| Oceania | I | 92.93% | 6.87 | 1.37 |
| | II | 84.08% | 7.35 | 1.88 |

**Supplementary Table 8e: Population coverage of the T-cell epitopes for DENV4 construct in regions with frequent and sporadic dengue occurrence**

The dengue endemic regions were identified as per CDC guidelines (https://www.cdc.gov/dengue/areaswithrisk/around-the-world.html). The DENV1 vaccine is estimated to be efficient in all populations except Central America which is predicted to have a lower coverage of MHCI and MHCII HLA types recognising the epitopes in the sequence. We also predicted that Southeast Asian population may have a lower coverage of MHCII HLA types and Caribbean (West Indies) population may have no coverage of MHCII epitopes generated from our vaccine construct.

| **Population** | **MHC class** | **Population coverage** | **Average hit** | **PC90** |
|---|---|---|---|---|
| South Asia | I | 89.63% | 2.92 | 0.96 |
| | II | 98.63% | 72.12 | 41.05 |
| Southeast Asia | I | 88.37% | 2.47 | 0.86 |
| | II | 19.27% | 8.05 | 4.21 |
| East Africa | I | 82.99% | 2.52 | 0.59 |
| | II | 81.1% | 42.2 | 17.99 |
| West Africa | I | 89.61% | 2.91 | 0.96 |
| | II | 90.4% | 52.8 | 34.43 |
| Central Africa | I | 80.42% | 2.49 | 0.51 |
| | II | 86.08% | 45.9 | 24.42 |
| West Indies | I | 94.32% | 3.96 | 1.31 |

| | | | | |
|---|---|---|---|---|
| | II | - | - | - |
| Central America | I | 7.76% | 0.14 | 0.11 |
| | II | 26.04% | 13.03 | 4.6 |
| South America | I | 79.12% | 2.54 | 0.48 |
| | II | 92.48% | 62.6 | 35.48 |
| Oceania | I | 90.28% | 2.19 | 1.01 |
| | II | 84.08% | 42.39 | 21.35 |