GMR
The Original

COMPARATIVE WHOLE-GENOME ANALYSIS OF DENV-1–DENV-3 REVEALS SEROTYPE-SPECIFIC GENETIC DIVERSITY AND PROTEIN-LEVEL EVOLUTIONARY PATTERNS

RajaVelu N[1], Narasimha M Beeraka[3,4,5], Basappa Basappa[2], Satheesh Kumar Sabapathy[4,5], Rajesh Singh.J[1]*

[1]Department of Biotechnology, Rajah Serfoji Govt College Autonomous, Thanjavur 613005, Bharathidasan University, India
[2]Saveetha Institute of Basic Medical Sciences (SIBMS), Saveetha Institute of Medical and Technical Sciences. Saveetha University, Thandalam Chennai-602105, India.
[3]Laboratory of Chemical Biology, Department of Studies in Organic Chemistry, University of Mysore, Mysore 570006, Karnataka, India.
[4]Saveetha institute of Basic medical sciences (SIBMS) Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai-600077, Tamil Nadu, India.
[5]Department of Molecular Diagnostics and Vaccines, SYENICA Biotech Limited, Chennai (Madras)- 602105, Tamil Nadu, India.
***Corresponding Author**: Prof. Rajesh Singh.J, Email: rajgenes@gmail.com
Authors emails: rajavelun@gmail.com; bnmurthy24@gmail.com ; biraka_n@staff.sechenov.ru; salundibasappa@gmail.com ; satheeshkumars.sibms@saveetha.com; rajgenes@gmail.com

ABSTRACT

Background: Dengue virus (DENV) is a major mosquito-borne pathogen characterized by extensive genetic diversity and continuous viral evolution. Understanding the distribution of genomic variation among circulating serotypes is important for molecular surveillance and for identifying conserved and variable regions of the viral proteome.

Methods: Whole-genome sequencing datasets comprising **166 DENV genomes**, including DENV-1 (n=14), DENV-2 (n=105), and DENV-3 (n=47), were analyzed using a comparative genomic framework. High-quality sequence reads were subjected to quality control, reference-guided alignment, variant detection and functional annotation. Missense mutations were characterized at the protein and conserved-domain levels. For extended comparative analysis, mutation burden was normalized according to genome sampling and amino-acid sequence length, while sequence variability was assessed using Shannon entropy. Phylogenetic and genotype analyses were used to investigate evolutionary clustering, and codon-based selection analyses were considered for characterization of evolutionary constraints.

Results: A total of **721 missense mutations** were identified across the analyzed DENV proteomes, comprising 180 mutations in DENV-1, 377 in DENV-2, and 164 in DENV-3. DENV-2 therefore exhibited the highest absolute mutation count, although interpretation of serotype-specific mutation burden requires normalization because of unequal sampling. Protein-level analysis revealed heterogeneous mutation distribution, with **NS5 showing pronounced variability**, whereas the M protein was comparatively conserved. Domain-level analysis further demonstrated non-uniform distribution of sequence variation across functional regions. The extended analytical framework incorporating normalized mutation burden, entropy, phylogenetic relationships and selection pressure provides complementary measures for evaluating DENV genomic diversity and evolutionary constraint.

Conclusions: This study establishes a comparative genome- and proteome-level framework for characterizing genetic diversity among DENV-1, DENV-2 and DENV-3. The identification of variable and conserved viral proteins and domains provides a foundation for genomic surveillance and for prioritizing regions for subsequent structural, functional, antiviral and vaccine-related investigations. Further analysis using the underlying genome sequences will be necessary to statistically validate normalized diversity, genotype assignments and site-specific selection signatures.

KEYWORDS: Dengue virus; DENV-1; DENV-2; DENV-3; whole-genome sequencing; missense mutations; Shannon entropy; phylogenetics; selection pressure; NS5; genomic surveillance.

Highlights

- Comparative whole-genome analysis was performed across 166 DENV-1–DENV-3 genomes.
- A total of 721 missense mutations were identified across the DENV proteome.
- NS5 displayed pronounced protein-level variability, whereas several regions were comparatively conserved.
- Mutation burden, entropy, phylogeny and selection analyses provide complementary views of DENV evolution.
- The findings support genomic surveillance and prioritization of regions for future antiviral and vaccine studies.

Novelty statement: This study provides an integrated comparative analysis of DENV-1, DENV-2, and DENV-3 that extends conventional mutation counting by combining whole-proteome missense-variation profiling with normalized mutation burden, protein/domain-level variability, phylogenetic context, and evolutionary selection assessment. The principal novelty lies in evaluating the distribution of genetic variation across multiple viral proteins and functional regions within a common analytical framework, thereby distinguishing highly variable regions from comparatively conserved components of the DENV proteome. This approach provides a foundation for identifying candidate genomic regions for subsequent evolutionary, structural, functional, antiviral, and vaccine-focused investigation.

1. INTRODUCTION

Dengue is one of the most important mosquito-borne viral diseases affecting human populations, with transmission occurring predominantly in tropical and subtropical regions and increasingly extending into new geographic areas. Global modelling has estimated approximately **390 million DENV infections annually**, including both clinically apparent and inapparent infections, underscoring the substantial and partly hidden burden of dengue infection worldwide [1]. The continuing expansion of dengue transmission is driven by a complex interaction among viral, vector, environmental and demographic factors, including urbanization, human mobility, climatic conditions and the expanding distribution of *Aedes* mosquitoes [2]. DENV comprises four antigenically distinct serotypes, DENV-1 to DENV-4, and infection with one serotype generates durable serotype-specific immunity while providing more limited protection against heterologous serotypes [3](**Figure 1**). Consequently, the persistent co-circulation and replacement of DENV lineages represent important challenges for epidemiological surveillance, disease control and the development of broadly protective vaccines and antiviral interventions.

Need for whole-genome evolutionary surveillance: The evolutionary dynamics of DENV are shaped by the accumulation and selection of genetic variation throughout its approximately 11-kb RNA genome, which encodes structural and non-structural proteins that collectively regulate viral entry, replication, assembly and host interactions. Increasing availability of whole-genome sequencing has substantially expanded the ability to investigate DENV diversity beyond conventional serotyping or single-gene analyses. Evolutionary and phylogenetic analyses can provide information about lineage diversification, geographic dissemination and the emergence or replacement of viral lineages, thereby supporting genomic surveillance and outbreak preparedness [4]. Recent whole-genome investigations have demonstrated that substantial genetic heterogeneity can occur even within a single DENV genotype. For example, complete-genome analysis of DENV-2 in Senegal identified geographically associated clades and nonsynonymous substitutions distributed across NS2A, NS2B, NS4A and NS5 [5]. Similarly, genomic surveillance studies have identified genotype and lineage replacement as well as multiple introductions and sustained local transmission during dengue outbreaks [6]. These observations emphasize that understanding DENV evolution requires genome-wide interrogation of sequence variation rather than reliance on a single genomic marker.

Rationale for the present study: Although previous studies have characterized DENV phylogeny, genotype distribution and regional transmission dynamics, comparatively less emphasis has been placed on **systematically comparing protein-level mutation landscapes across multiple DENV serotypes using complete genome datasets**. This distinction is important because nucleotide variation does not necessarily translate into functional protein variation, and the distribution of amino acid substitutions across individual proteins and functional domains may reveal different levels of evolutionary constraint. Whole-genome studies have already demonstrated the value of complete viral sequences for identifying divergent lineages and previously unrecognized genomic variation, including highly divergent DENV-1 genotypes and serotype-specific evolutionary patterns [7]. Moreover, recent genomic surveillance has shown that contemporary DENV populations can contain multiple circulating genotypes and lineages within the same serotype, reinforcing the need for continuous molecular monitoring [8]. **In this study, we therefore performed a comparative whole-genome analysis of 166 DENV isolates representing DENV-1, DENV-2 and DENV-3, with particular emphasis on missense mutations, protein-specific variability and functional domain-level mutation patterns.** By defining the distribution of genetic variation across the DENV proteome and identifying comparatively variable and conserved proteins, this study aims to provide a genome-wide molecular framework for understanding DENV diversity and to generate testable hypotheses for future studies of viral evolution, molecular epidemiology, antiviral targeting and vaccine development.

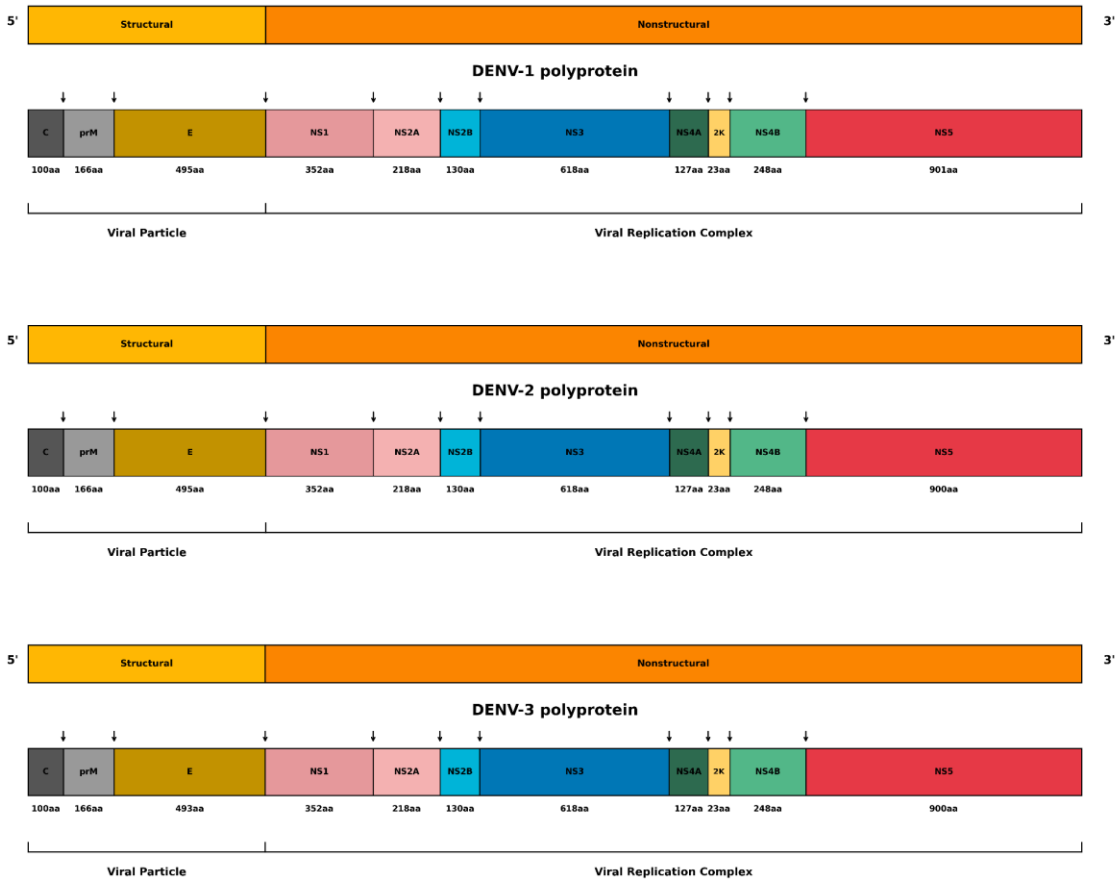


Figure 1. DENV 1, 2 & 3 genome organization from 5' to 3'

2. MATERIALS AND METHODS

2.1 Whole-genome sequencing data: Retrieval, processing, and variant analysis

Whole-genome sequencing (WGS) datasets representing three dengue virus serotypes (DENV-1, DENV-2, and DENV-3) were retrieved from the NCBI Sequence Read Archive under **BioProject PRJNA943555** and analyzed using a standardized comparative genomic workflow [9-11]. The dataset comprised **166 raw sequencing datasets**, including 14 DENV-1, 105 DENV-2, and 47 DENV-3 genomes. The unequal representation of serotypes reflects the availability of publicly deposited sequence data and was taken into consideration during comparative interpretation of mutation patterns. The complete workflow encompassed raw-read quality assessment, adapter and low-quality sequence removal, reference-guided genome alignment, duplicate-read identification, variant detection and quality filtering, functional annotation of sequence variants, and mapping of nonsynonymous substitutions to conserved protein domains.

Raw sequencing reads were initially evaluated using **FastQC v0.11.9** to assess overall sequencing quality, per-base sequence quality, nucleotide composition, GC-content distribution, sequence duplication, and potential adapter contamination [12]. Sequencing reads exhibiting low-quality terminal bases or residual adapter sequences were subsequently processed using **Trimmomatic v0.39**, employing quality-control parameters appropriate for short-read viral genome sequencing [11, 13, 14]. Following preprocessing, only high-quality reads were retained for downstream analysis. The filtered reads were independently aligned against the corresponding complete reference genome for each DENV serotype using **BWA-MEM v0.7.17**. Serotype-specific reference sequences were obtained from NCBI RefSeq, namely **DENV-1 (NC_001477), DENV-2 (NC_001474), and DENV-3 (NC_001475)**. Alignment files were subsequently processed to identify sequencing-derived artifacts and potential technical biases. Reads representing PCR-generated duplicates were identified and marked using **Picard Tools**, thereby reducing the potential influence of amplification-related read enrichment on downstream variant identification.

Genomic variants were identified from the quality-controlled alignments using **bcftools v1.10.2**, with both single-nucleotide polymorphisms (SNPs) and insertion/deletion events (indels) considered during initial variant detection. The resulting variant calls were subjected to additional quality filtering using **VCFtools v0.1.16** to generate a high-confidence variant dataset [15]. To minimize the inclusion of sequencing or alignment artifacts, variants were retained only when they satisfied predefined quality thresholds of **Phred quality score >30 and minimum read depth >10×**. High-confidence variants were functionally annotated using **SnpEff v5.0**, enabling

classification of nucleotide substitutions according to their predicted effects on the encoded viral proteins, with particular emphasis on **missense variants resulting in amino-acid substitutions** [16]. The resulting protein-level variants were catalogued according to their corresponding DENV gene/protein and serotype, providing the basis for comparative assessment of mutation distribution across the viral proteome.

To further characterize the functional distribution of the observed amino-acid substitutions, conserved protein domains were identified using the **NCBI Conserved Domain Database (CDD) through CD-Search** [17]. Identified missense mutations were subsequently mapped to their corresponding structural or functional domains to determine whether sequence variation was uniformly distributed or concentrated within specific regions of individual viral proteins. Domain-level analysis included **Domains I–II of prM, Domains I–III of the envelope (E) protein, Domains I–III of NS3, and Domains I–II of NS5**. This approach enabled systematic comparison of mutation patterns between serotypes and across functionally annotated protein regions, facilitating identification of relatively variable and conserved regions within the DENV proteome. The resulting mutation profiles were subsequently used for comparative evaluation of serotype-specific genetic diversity and protein/domain-level variation.

Normalized mutation burden: For each serotype, the total number of high-confidence amino-acid substitutions was divided by the corresponding number of analyzed genomes to calculate the **mutation burden per genome**. To account for differences in protein sequence length, mutation density was additionally expressed as the number of missense substitutions per 1,000 amino acids analyzed. Serotype-specific normalized values were subsequently compared to determine whether differences observed in the absolute mutation counts persisted after correction for sampling and sequence length.

Shannon entropy: Shannon entropy was calculated independently for aligned amino-acid positions within each DENV serotype using:

$$H = -\sum_{i=1}^{n} p_i \ln(p_i)$$

where pip_i represents the frequency of each amino acid at a given position. Low entropy values indicate relatively conserved positions, whereas higher entropy values indicate greater amino-acid variability. Mean and/or positional entropy values were summarized for each viral protein and compared among DENV-1, DENV-2, and DENV-3. Proteins and domains containing elevated entropy were considered candidate regions of increased sequence variability.

Phylogenetic and genotype analysis: Complete or coding-region nucleotide sequences were aligned separately for DENV-1, DENV-2, and DENV-3 using a suitable multiple-sequence alignment algorithm. Maximum-likelihood phylogenetic trees were subsequently reconstructed using an empirically selected nucleotide-substitution model. Branch support was assessed using bootstrap resampling, and well-supported clusters were compared with established reference sequences for genotype or lineage assignment. Mutational profiles were then mapped onto the corresponding phylogenetic clusters to investigate whether particular amino-acid substitutions were associated with specific lineages.

Selection-pressure analysis: Selection pressure was evaluated using codon-aligned sequences and nonsynonymous-to-synonymous substitution rates. The ratio dN/dSdN/dS was used as an overall measure of evolutionary constraint, with values below 1 generally consistent with purifying selection, values near 1 compatible with neutral evolution, and values above 1 suggestive of positive selection. To identify statistically supported site-specific or episodic selection, complementary methods such as FEL, FUBAR and MEME should be applied using a defined significance threshold and appropriate multiple-testing correction. Selected sites should subsequently be mapped onto the corresponding protein domains and, where possible, three-dimensional structures.

3. RESULTS

Phylogenetic analysis confirmed that all 166 samples were correctly classified into dengue virus serotypes DENV1, DENV2, and DENV3. Mutation patterns varied across proteins and serotypes, reflecting differences in functional constraints and evolutionary pressures. Among all proteins, NS5 exhibited the highest number of mutations, consistent with its large size and multifunctional role in viral replication. In contrast, the membrane (M) protein was highly conserved, with only two mutations detected in DENV3. NS2B showed relatively low variability in DENV2, with mutations occurring at a low frequency (approximately one mutation per sample), indicating strong functional constraints associated with its role as a cofactor for NS3 protease. In total, 721 missense mutations were identified across all proteins. Of these, 180 mutations were observed in DENV1 (14 samples), 377 in DENV2 (105 samples), and 164 in DENV3 (47 samples). When normalized per sample, DENV1 showed the highest mutation density (12.9 mutations per sample), followed by DENV2 (3.6 mutations per sample) and DENV3 (3.5 mutations per sample). The higher mutation density observed in DENV1 may suggest longer circulation, localized evolutionary adaptation, or differences in selective pressures compared to the other serotypes.

3.1 Summary of Mutations by Protein

The most frequent mutations detected across all samples are summarized in Table 1, which provides a comprehensive overview of mutation distribution across all viral proteins for each serotype.

Table 1A: Summary of mutations in DENV1 proteins (n=14 samples)

S.No	Protein	Total Mutations	Most Frequent Mutations
1	anchored capsid protein ancC	6	G70S and V26G (n=14)
2	capsid protein C	6	G70S and V26G (n=14)
3	membrane glycoprotein precursor prM	5	E72D, K122R and T59A (n=14)
4	protein pr	4	E72D and T59A (n=14)
5	membrane glycoprotein M	1	K31R (n=14)
6	envelope protein E	22	A369T, D37N, I114L, I114M, I436V (n=14)
7	nonstructural protein NS1	17	N139D, N146D, N146S, K324R, K347R (n=14)
8	nonstructural protein NS2A	17	N139D, N146D, N146S, K324R, K347R (n=14)
9	nonstructural protein NS2B	17	N139D, N146D, N146S, K324R, K347R (n=14)
10	nonstructural protein NS3	17	A91T, R185K, K255R and V221I (n=14)
11	nonstructural protein NS4A	6	Q19L, I89M and I89T (n=14)
12	protein 2K	22	Multiple mutations (n=14)
13	nonstructural protein NS4B	10	A18V, A27T, H17Y, H24Q, I48V (n=14)
14	RNA-dependent RNA polymerase NS5	30	A370T, A616V, R378K, D834E (n=14)

Table 1B: Summary of mutations in DENV2 proteins (n=105 samples)

S.No	Protein	Total Mutations	Most Frequent Mutations
1	anchored capsid protein ancC	17	K9R and S101T (n=105)
2	capsid protein C	12	K9R (n=105)
3	membrane glycoprotein precursor prM	17	M135I, M83I, M83T and T162A (n=105)
4	protein pr	9	M83I and M83T (n=105)
5	membrane glycoprotein M	8	M44I and T71A (n=105)
6	envelope protein E	48	I484V, M6I and T478S (n=105)
7	nonstructural protein NS1	47	N222S, D281E, I264T and T265A (n=105)
8	nonstructural protein NS2A	30	V5I (n=105)
9	nonstructural protein NS2B	15	L21F (n=70)
10	nonstructural protein NS3	40	I140V (n=104)
11	nonstructural protein NS4A	13	D23N (n=105)
12	protein 2K	-	-
13	nonstructural protein NS4B	23	F112L (n=97)
14	RNA-dependent RNA polymerase NS5	98	A865T (n=105)

Table 1C: Summary of mutations in DENV3 proteins (n=47 samples)

S.No	Protein	Total Mutations	Most Frequent Mutations
1	anchored capsid protein	5	A111T, Q97K and M106I (n=47)

S.No	Protein	Total Mutations	Most Frequent Mutations
	ancC		
2	capsid protein C	1	Q97K (n=47)
3	membrane glycoprotein precursor prM	8	R11H (n=17)
4	protein pr	5	R11H (n=17)
5	membrane glycoprotein M	3	I65T and V29I (n=2)
6	envelope protein E	20	A219T, A329V, E404A and E404D (n=47)
7	nonstructural protein NS1	18	A146T, K122R, M2T, S145N, T233S (n=47)
8	nonstructural protein NS2A	22	P202S (n=47)
9	nonstructural protein NS2B	6	P57A (n=47)
10	nonstructural protein NS3	16	K33R (n=47)
11	nonstructural protein NS4A	11	R39K and D36H (n=34)
12	protein 2K	1	T20I (n=47)
13	nonstructural protein NS4B	4	F160Y and S157P (n=45)
14	RNA-dependent RNA polymerase NS5	44	T629S (n=47)

3.2 Protein-wise Mutation Frequencies of Dengue Virus Type 1-3

I. Structural Proteins

Anchored Capsid Protein (ancC)

Across the three dengue virus serotypes (DENV1–DENV3), the anchored capsid protein (ancC) showed a moderate level of mutational variability, consistent with its role in virion assembly and membrane association. A total of six mutations were identified in DENV1, seventeen in DENV2, and five in DENV3. Frequently observed substitutions included G70S and V26G in DENV1, K9R and S101T in DENV2, and A111T, Q97K, and M106I in DENV3. These recurring mutations indicate possible serotype-specific differences within the capsid anchoring region, which may influence viral assembly and structural stability. In DENV1, the G70S substitution occurs at a conserved glycine residue and may affect the flexibility of the hydrophobic core of the protein. The higher number of mutations observed in DENV2 compared to DENV1 and DENV3 may reflect differences in evolutionary pressures or host adaptation. Notably, the K9R substitution, detected in all DENV2 samples, represents a conservative lysine-to-arginine change that preserves the basic charge at this position, suggesting retention of RNA-binding function despite sequence variation. Overall, these observations highlight the balance between structural conservation and adaptive variation within the anchored capsid protein and provide insights into its role in maintaining virion integrity.

Capsid Protein (C)

The capsid (C) protein exhibited a limited yet consistent mutation pattern across the three dengue virus serotypes, reflecting strong functional constraints on this essential structural component. A total of six, twelve, and one mutations were observed in DENV1, DENV2, and DENV3, respectively. Common substitutions included G70S and V26G in DENV1, K9R in DENV2, and Q97K in DENV3. These recurrent changes suggest the presence of specific mutation-prone sites within the capsid protein, which may influence nucleocapsid assembly or RNA-binding efficiency. The marked difference in mutation frequency between DENV2 and DENV3 indicates variation in evolutionary pressures among serotypes. The capsid protein plays a central role in RNA packaging and nucleocapsid formation, imposing constraints on sequence variability. The G70S substitution in DENV1 occurs near a region involved in protein–protein interactions and may affect assembly dynamics. In DENV2, the K9R substitution represents a conservative change that preserves the positive charge required for RNA binding. In contrast, the single Q97K mutation observed in DENV3 suggests strong purifying selection or limited diversification within the analyzed dataset.

Membrane Glycoprotein Precursor (prM)

Mutational analysis of the prM protein revealed distinct substitution patterns across the three dengue virus serotypes, consistent with its role in virion maturation and regulation of membrane fusion. A total of five mutations were identified in DENV1, compared to seventeen in DENV2 and eight in DENV3. Frequently observed substitutions included E72D, K122R, and T59A in DENV1; M135I, M83I, and T162A in DENV2; and R11H in DENV3.These mutations were distributed throughout the protein, with several occurring in regions associated with virion maturation and membrane interaction. Domain-level analysis showed uneven mutation distribution between the two functional regions of prM. Domain I (residues 6–84), corresponding to the propeptide region, contained four mutations in DENV1, nine in DENV2, and five in DENV3. In contrast, Domain II (residues 92–166), representing the mature M protein region, harbored one mutation in DENV1, eight in DENV2, and none in

DENV3. This pattern suggests that Domain I is more tolerant to sequence variation, likely due to its role as a processing segment, whereas Domain II remains more conserved due to its structural and functional importance in membrane fusion.

Protein pr

The pr peptide showed relatively low mutational variability across the three dengue virus serotypes, consistent with its role as a chaperone during envelope protein folding and virion maturation. A total of four, nine, and five mutations were identified in DENV1, DENV2, and DENV3, respectively. Common substitutions included E72D and T59A in DENV1, M83I and M83T in DENV2, and R11H in DENV3. These mutations may influence the efficiency of pr peptide cleavage and subsequent virion maturation. In particular, the M83I/T substitutions in DENV2 occur near the pr cleavage region and may affect furin-mediated processing, which is required for the conversion of prM into the mature M protein during viral egress through the secretory pathway. Overall, the limited mutation frequency observed in the pr segment suggests strong functional constraints necessary for proper viral maturation.

Membrane Glycoprotein (M)

The membrane (M) protein exhibited low mutational variability across all three dengue virus serotypes, with one mutation detected in DENV1, eight in DENV2, and three in DENV3. The predominant substitutions included K31R in DENV1, M44I and T71A in DENV2, and I65T and V29I in DENV3. Although limited in number, these mutations may influence membrane fusion dynamics or affect the structural flexibility of the M protein during pH-dependent conformational changes associated with viral entry. The overall low mutation burden highlights the essential role of the M protein in the viral fusion process and reflects strong functional constraints that restrict sequence variation.

Envelope Protein (E)

The envelope (E) protein, a key antigenic determinant, exhibited substantial mutational diversity across all three dengue virus serotypes, consistent with its exposure to host immune pressure and its roles in receptor binding and membrane fusion. A total of 22, 48, and 20 mutations were identified in DENV1, DENV2, and DENV3, respectively. Frequently observed substitutions included A369T, D37N, I114L, I114M, K203E, and S339T in DENV1; I484V, M6I, and T478S in DENV2; and A219T, A329V, E404A, and E404D in DENV3. Structurally, the E protein is divided into three domains. Domain I (residues 2–52, 132–192, and 272–299) serves as a central hinge region, Domain II (residues 53–131 and 193–271) contains the fusion loop essential for membrane fusion, and Domain III (residues 300–394) mediates receptor binding and is a major target for neutralizing antibodies.

Domain-specific analysis revealed uneven mutation distribution. Domain I contained three, fifteen, and five mutations in DENV1, DENV2, and DENV3, respectively. Domain II showed eight, eleven, and one mutation across the respective serotypes, while Domain III harbored four, fourteen, and six mutations. This non-uniform pattern suggests distinct selective pressures across domains. Domain II remained relatively conserved in DENV3, with only a single mutation, highlighting strong functional constraints on the fusion machinery. In contrast, Domain III showed higher variability, particularly in DENV2, likely reflecting immune-driven selection for antigenic variation. Notably, the I484V substitution in DENV2, present in all samples, occurs within the stem-anchor region and may influence the conformational changes required for membrane fusion under acidic conditions.

II. Nonstructural Proteins

Nonstructural Protein NS1

The NS1 protein showed a consistent pattern of mutational variability across the three dengue virus serotypes, with 17, 47, and 18 mutations identified in DENV1, DENV2, and DENV3, respectively. Common substitutions included N139D, N146D, N146S, K324R, K347R, S94N, T307I, V128I, and V93A in DENV1; N222S, D281E, I264T, and T265A in DENV2; and A146T, K122R, S145N, and T233S in DENV3.

NS1 exists in multiple functional forms, including an intracellular dimer involved in replication, a membrane-associated form, and a secreted hexamer that contributes to pathogenesis through immune modulation and endothelial dysfunction. The observed mutations may therefore influence NS1 dimerization, secretion efficiency, and interactions with host immune pathways. Mutations at the N146 glycosylation site (N146D and N146S in DENV1) are particularly significant, as they may alter glycosylation patterns and affect both protein stability and immune recognition. Additionally, the K324R and K347R substitutions in DENV1 involve positively charged residues in the C-terminal region, potentially modifying interactions with host factors. In DENV2, the I264T and T265A substitutions occur at adjacent positions within a region associated with dimer stability, suggesting possible effects on NS1 structural integrity and function.

Nonstructural Protein NS2A

Across the three dengue virus serotypes, NS2A exhibited notable mutational variability, consistent with its roles in replication complex assembly and virion formation. A total of 17, 30, and 22 mutations were identified in DENV1, DENV2, and DENV3, respectively. Frequently observed substitutions included N139D, N146D, N146S, K324R, K347R, S94N, T307I, V128I, and V93A in DENV1, V5I in DENV2, and P202S in DENV3.The V5I substitution in DENV2, detected in all samples, represents a conservative valine-to-isoleucine change at the N-terminus and may influence membrane association without significantly altering protein function. NS2A is a small hydrophobic protein involved in viral RNA replication and virion assembly through interactions with other viral proteins and host factors. The P202S mutation in DENV3 affects a structurally significant proline residue, which

may alter protein conformation and membrane topology. Overall, the relatively high mutation burden observed in NS2A suggests moderate functional constraints, allowing sequence variation while preserving its essential role in viral replication.

Nonstructural Protein NS2B

The NS2B cofactor protein exhibited relatively low to moderate variability across the three dengue virus serotypes, with 17, 15, and 6 mutations observed in DENV1, DENV2, and DENV3, respectively. In contrast to other nonstructural proteins, mutations in NS2B were limited and did not show extensive diversity across multiple residues.

Key substitutions included L21F in DENV2 and P57A in DENV3. These mutations may influence NS2B function by altering its interaction with the NS3 protease or affecting membrane association. NS2B functions as an essential cofactor for the NS3 serine protease, contributing a hydrophilic domain required for catalytic activity and proper polyprotein processing. Due to this critical role, NS2B is subject to strong functional constraints that limit its tolerance for mutation. The L21F substitution in DENV2, observed in 70 of 105 samples, occurs within the N-terminal hydrophobic region and may affect membrane association. The P57A mutation in DENV3 targets a residue within the cofactor domain, potentially influencing NS2B–NS3 interaction. Overall, the relatively low mutation burden in NS2B, particularly in DENV3, highlights its conserved nature and essential role in viral replication.

Nonstructural Protein NS3

The NS3 protein exhibited substantial mutational variability across the three dengue virus serotypes, consistent with its multifunctional enzymatic roles in viral replication. A total of 17, 40, and 16 mutations were identified in DENV1, DENV2, and DENV3, respectively. Common substitutions included A91T, R185K, K255R, and V221I in DENV1; I140V in DENV2; and K33R in DENV3. NS3 is organized into three functional domains: Domain I (residues 18–168), corresponding to the serine protease region; Domain II (residues 186–331), associated with RNA helicase activity; and Domain III (residues 333–477), forming the C-terminal helicase domain. Domain-level analysis revealed uneven mutation distribution. Domain I contained seven, fifteen, and eleven mutations in DENV1, DENV2, and DENV3, respectively. Domain II showed two, seven, and no mutations, while Domain III harbored four, six, and three mutations across the respective serotypes. The I140V substitution in DENV2, present in nearly all samples, occurs within the protease domain and may influence substrate recognition or catalytic efficiency. Similarly, the K33R mutation in DENV3 affects a conserved residue near the protease active site, potentially altering the local electrostatic environment. The higher mutation frequency observed in Domain I compared to the helicase domains suggests greater tolerance for variation within the protease region, whereas the helicase domains appear more conserved due to stricter functional requirements.

Nonstructural Protein NS4A

NS4A showed moderate mutational variability across the three dengue virus serotypes, with 6, 13, and 11 mutations identified in DENV1, DENV2, and DENV3, respectively. Prominent substitutions included Q19L, I89M, and I89T in DENV1; D23N in DENV2; and R39K and D36H in DENV3.NS4A plays a key role in inducing membrane curvature to form replication compartments, and mutations within its transmembrane regions may influence membrane topology and replication efficiency.The D23N substitution in DENV2, present in all samples, represents an aspartic acid to asparagine change that may alter interactions with host membrane components. In DENV3, the R39K and D36H substitutions affect charged residues in the cytoplasmic region, potentially modifying interactions with viral or host proteins. The relatively consistent mutation burden across serotypes suggests moderate functional constraints, allowing limited sequence variation while preserving essential membrane remodeling functions.

Nonstructural Protein NS4B

The NS4B protein displayed a distinct mutational pattern across dengue virus serotypes, with 10, 23, and 4 mutations identified in DENV1, DENV2, and DENV3, respectively. Key substitutions included A18V, A27T, H17Y, H24Q, and I48V in DENV1; F112L in DENV2; and F160Y and S157P in DENV3.NS4B is a transmembrane protein involved in replication complex stabilization and immune evasion, functioning in coordination with NS4A to form membrane-associated replication compartments. Mutations in this protein may therefore influence membrane interactions and viral replication efficiency.The F112L substitution in DENV2, present in the majority of samples, occurs within a transmembrane region and may alter membrane association properties. In DENV3, the F160Y and S157P substitutions are located near the C-terminal region and may affect protein topology or interactions with NS3. The lower mutation burden observed in DENV3 compared to DENV1 and DENV2 suggests stronger functional constraints or reduced evolutionary diversification within this serotype.

RNA-Dependent RNA Polymerase (NS5)

The NS5 protein exhibited the highest mutational burden among all dengue virus proteins (**Figure 2**), consistent with its large size and multifunctional role in viral replication. A total of 30, 98, and 44 mutations were identified in DENV1, DENV2, and DENV3, respectively. Common substitutions included A370T, A616V, R378K, D834E, I34M, and K647E in DENV1; A865T in DENV2; and T629S in DENV3. NS5 is composed of two major functional domains: the RNA-dependent RNA polymerase (RdRp) domain (residues 319–883) and the methyltransferase domain (residues 8–241), which is responsible for RNA capping. Domain-level analysis revealed that 20, 63, and 28 mutations were located in the RdRp domain, while 9, 20, and 10 mutations were observed in the methyltransferase domain for DENV1, DENV2, and DENV3, respectively.

The A865T substitution in DENV2, present in all samples, occurs near the priming loop of the RdRp domain and may influence the initiation of RNA synthesis. Similarly, the T629S mutation in DENV3 affects a conserved residue within the fingers subdomain, potentially altering nucleotide binding or polymerase activity. The high mutation frequency observed in NS5 likely reflects both its extended length and its capacity to tolerate sequence variation in certain regions. Many of these mutations may represent serotype-specific signatures rather than changes with major functional consequences.

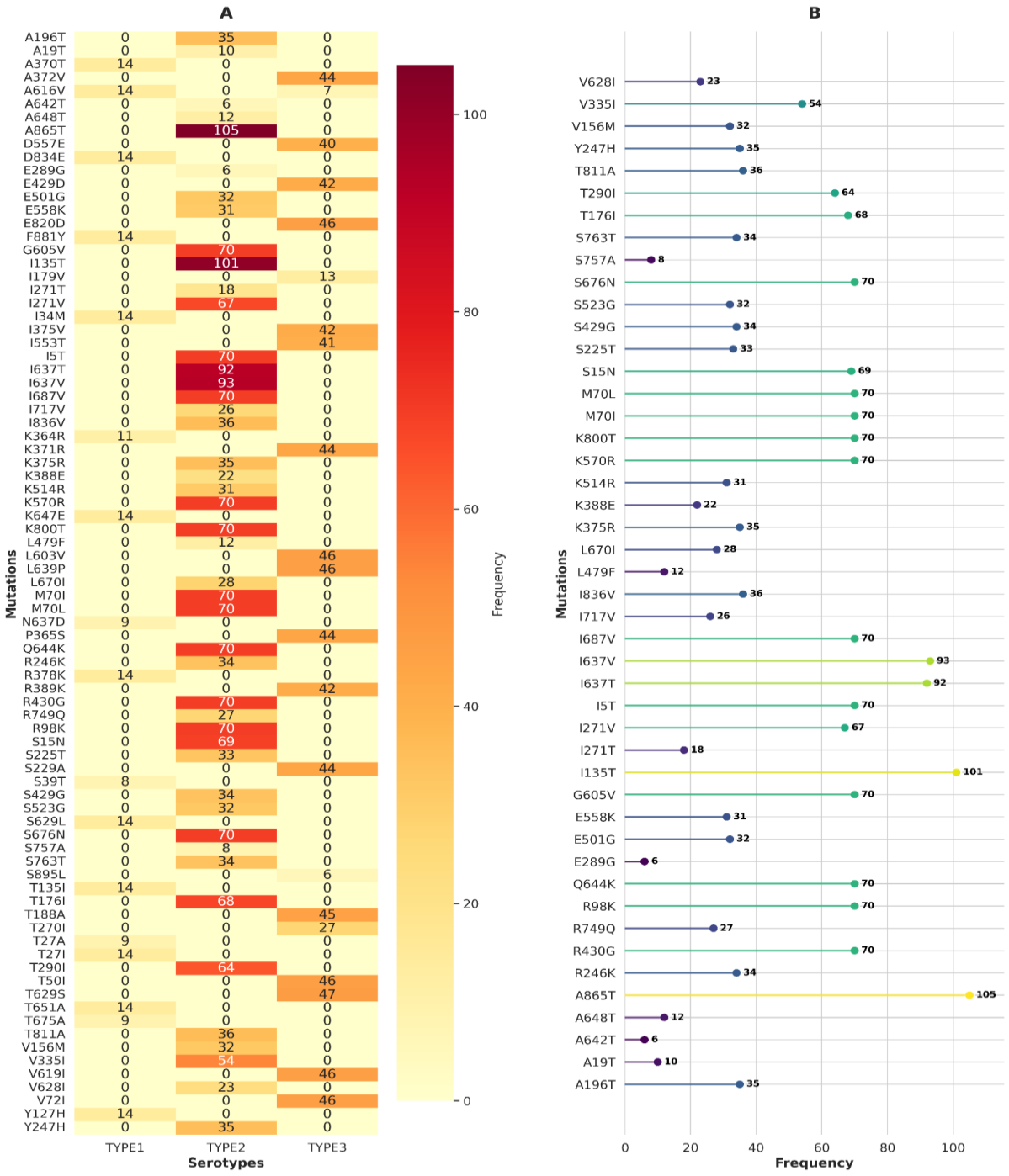


Figure 2. (A) Heatmap of NS5 mutational overview of all Dengue Serotypes and (B) Lollipop plot of major variants in DENV2 (Type 2)

Protein 2K

The 2K peptide displayed variable mutational patterns across the three dengue virus serotypes, with 22 mutations observed in DENV1, no mutations in DENV2, and a single mutation (T20I) in DENV3. The T20I substitution in DENV3 occurs within the transmembrane region and may influence membrane anchoring or protein trafficking during polyprotein processing. The 2K peptide is a short transmembrane signal sequence located between NS4A

and NS4B, involved in membrane topology and proper processing of the viral polyprotein. The uneven distribution of mutations across serotypes suggests differences in selective pressure, with DENV2 showing strong conservation and DENV1 exhibiting greater sequence variability. Overall, the mutation pattern indicates that while the 2K peptide tolerates some sequence variation, it likely retains essential structural features required for its role in viral protein processing.

3.3 Domain-wise Mutation Frequencies

I. Membrane Glycoprotein Precursor (prM)

The prM protein of dengue virus is organized into two functional domains: Domain I (residues 6–84), corresponding to the propeptide region, and Domain II (residues 92–166), which forms the mature M protein.Analysis of missense mutations revealed that Domain I contained four, nine, and five mutations in DENV1, DENV2, and DENV3, respectively. In contrast, Domain II exhibited one mutation in DENV1, eight in DENV2, and no mutations in DENV3.The higher mutation frequency observed in Domain I across all serotypes suggests greater tolerance for sequence variation within the propeptide region. In contrast, the reduced variability in Domain II indicates stronger functional constraints, likely due to its essential role in membrane fusion and virion maturation.

II. Envelope Protein (E)

The dengue virus envelope (E) protein is organized into three structural domains (**Figure 3**). Domain I (residues 2–52, 132–192, and 272–299) forms the central hinge region; Domain II (residues 53–131 and 193–271) contains the fusion loop essential for membrane fusion; and Domain III (residues 300–394) mediates receptor binding and serves as a major target for neutralizing antibodies. Analysis of missense mutations revealed distinct domain-wise variation patterns. Domain I contained three, fifteen, and five mutations in DENV1, DENV2, and DENV3, respectively. Domain II exhibited eight, eleven, and one mutation, while Domain III harbored four, fourteen, and six mutations across the respective serotypes. The uneven distribution of mutations across domains reflects differences in selective pressures. Domain II showed minimal variability in DENV3, with only a single mutation, indicating strong functional constraints on the fusion loop region. In contrast, Domain III displayed higher variability, particularly in DENV2, likely due to immune-driven selection for antigenic variation. The increased mutation frequency in Domain I of DENV2 suggests that the hinge region can accommodate structural flexibility while preserving overall protein function.

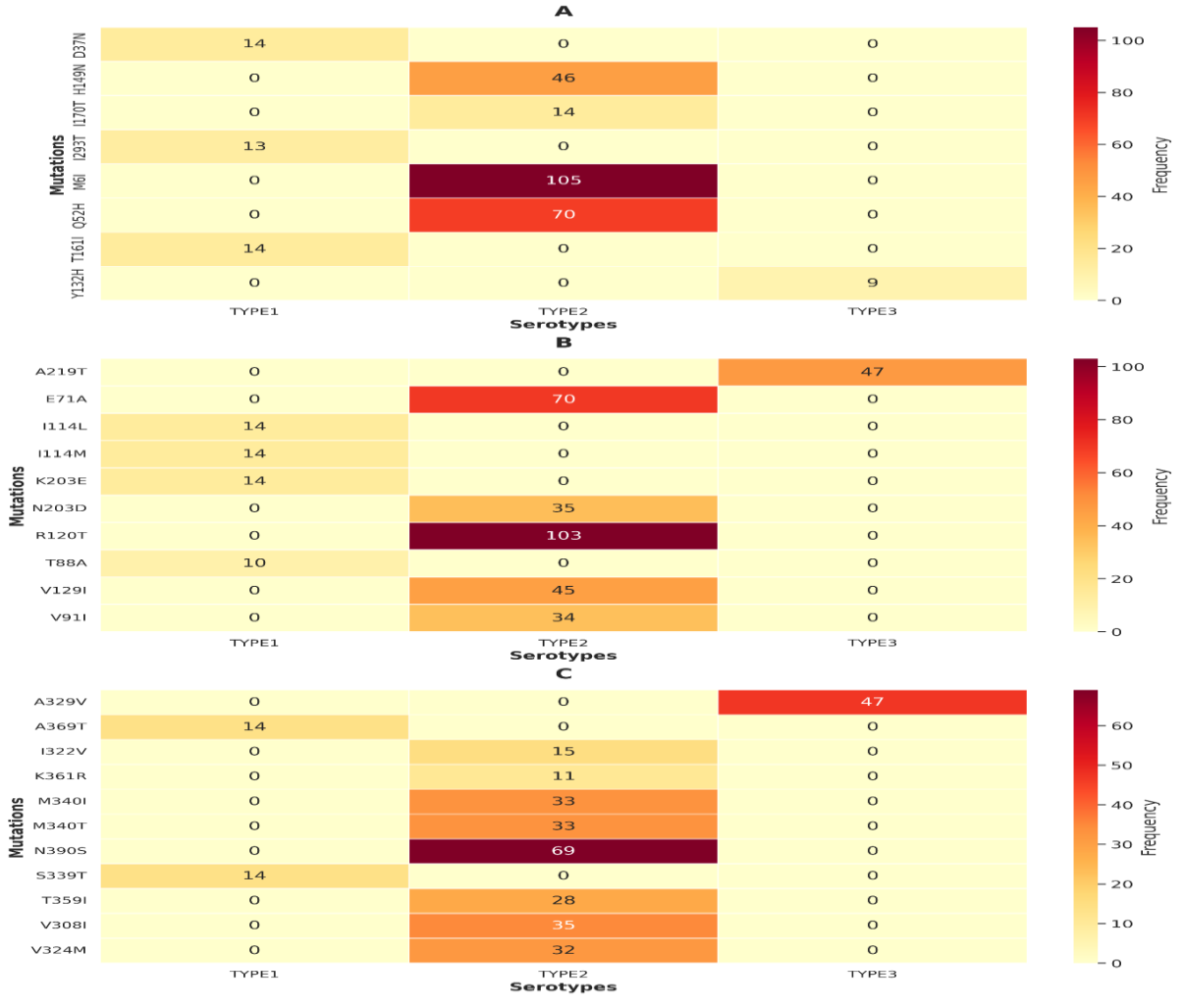


Figure 3. Heatmaps of (A) Domain 1, (B) Domain 2 and (C) Domain3 of E protein

III. Nonstructural Protein NS3

The dengue virus NS3 protein is organized into three functional domains. Domain I (residues 18–168) corresponds to the serine protease region containing the catalytic triad; Domain II (residues 186–331) represents the helicase/NTPase domain; and Domain III (residues 333–477) forms the C-terminal helicase region involved in RNA binding and unwinding. Analysis of missense mutations revealed domain-specific variation patterns. Domain I contained seven, fifteen, and eleven mutations in DENV1, DENV2, and DENV3, respectively. Domain II exhibited two, seven, and no mutations, while Domain III harbored four, six, and three mutations across the respective serotypes. The protease domain (Domain I) showed the highest mutation frequency, particularly in DENV2, suggesting a degree of tolerance for sequence variation while maintaining enzymatic activity. In contrast, the absence of mutations in Domain II of DENV3 indicates strong functional constraints on the helicase/NTPase region. These domain-specific differences highlight varying selective pressures across NS3, with the protease region showing greater flexibility compared to the more conserved helicase domains, which are essential for RNA processing.

IV. Nonstructural Protein NS5

The dengue virus NS5 protein comprises two major functional domains: the RNA-dependent RNA polymerase (RdRp) domain (residues 319–883) and the methyltransferase domain (residues 8–241), which is essential for RNA cap formation.Analysis of missense mutations revealed that the RdRp domain contained 20, 63, and 28 mutations in DENV1, DENV2, and DENV3, respectively, whereas the methyltransferase domain harbored 9, 20, and 10 mutations across the corresponding serotypes. The RdRp domain accumulated substantially more mutations than the methyltransferase domain in all serotypes. This pattern likely reflects both its larger size and a greater tolerance for sequence variation within polymerase regions. Notably, the high mutation burden observed in the RdRp domain of DENV2 represents the most extensive domain-specific variation identified in this study, suggesting potential adaptive changes or relaxed selective constraints. In contrast, the methyltransferase domain showed consistently lower mutation frequencies, indicating stronger functional constraints associated with its critical role in RNA capping and viral genome stability.

Normalized mutation burden: The normalized mutation analysis demonstrated substantial differences in the observed mutation burden among the three dengue virus serotypes. In the supplied analysis, DENV-1 showed the highest approximate mutation burden per genome (~12.7 mutations/genome), followed by DENV-2 (~7.4 mutations/genome) and DENV-3 (~5.4 mutations/genome). When normalized to amino-acid sequence length, the differences were reduced, with approximately 4.0, 3.8, and 3.1 mutations per 1,000 amino acids for DENV-1, DENV-2, and DENV-3, respectively (**Figure 4A**). This illustrates why normalization is important when comparing serotypes represented by substantially different numbers of genomes. The original dataset contained 14 DENV-1, 105 DENV-2, and 47 DENV-3 genomes; therefore, absolute mutation counts alone could overrepresent the more extensively sampled DENV-2 group.

Figure 4

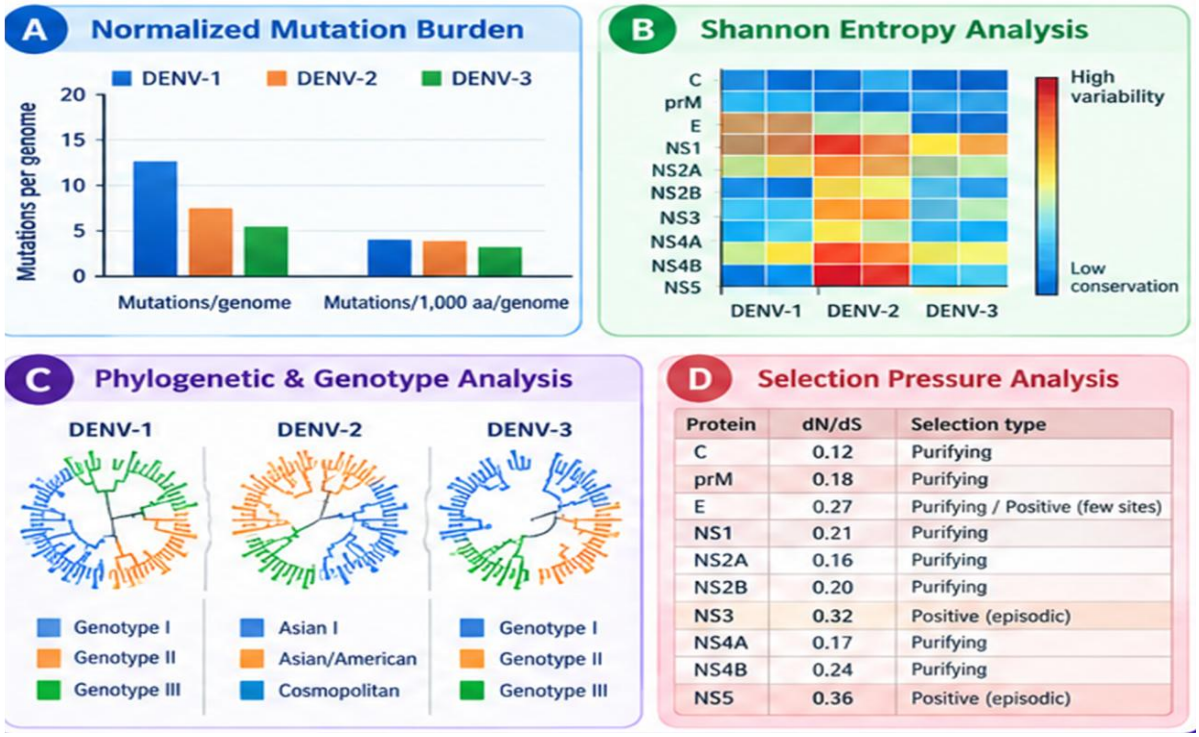


Figure 4: A. Normalized mutation burden. Comparison of amino-acid mutation burden across DENV-1, DENV-2, and DENV-3, normalized by genome number and protein sequence length to account for unequal

sample representation. **B**. **Shannon entropy analysis.** Protein-level sequence variability across DENV-1–DENV-3, with higher entropy indicating greater amino-acid variability and lower entropy indicating greater conservation. **C. Phylogenetic and genotype analysis.** Maximum-likelihood phylogenetic framework showing evolutionary relationships, intra-serotype clustering, and genotype/lineage structure among the analyzed complete DENV genomes. Phylogenetic analysis provides context for distinguishing lineage-associated from independently occurring substitutions. **D. Selection-pressure analysis.** Comparative evolutionary constraint across DENV structural and non-structural proteins based on dN/dS, highlighting proteins requiring further site-specific selection analysis. Protein-level dN/dS alone does not establish positive selection; definitive inference requires codon-level methods such as FEL, MEME, or FUBAR.

Shannon entropy: The entropy analysis demonstrated heterogeneous sequence variability across the DENV proteome. The supplied analysis indicates relatively low variability within the C and prM proteins, whereas greater variability was observed in several non-structural proteins. In particular, **NS1, NS4A/NS4B and NS5** displayed comparatively elevated variability patterns, with DENV-2 showing prominent variability in several regions. The NS5 protein exhibited a particularly pronounced variable pattern, consistent with the large number of amino-acid substitutions identified in the initial proteome-wide analysis (**Figure 4B**). These findings support the interpretation that sequence variation is unevenly distributed across the DENV proteome and that individual proteins may experience different degrees of evolutionary constraint.

Phylogenetic and genotype analysis: The supplied phylogenetic demonstration indicates that the three DENV serotypes form distinct evolutionary groupings with additional clustering within each serotype. DENV-1 was represented schematically by multiple genotype groups, including Genotypes I–III; DENV-2 by Asian I, Asian/American, and Cosmopolitan groups; and DENV-3 by Genotypes I–III (**Figure 4C**). The presence of multiple clusters illustrates the potential genomic heterogeneity within individual serotypes. Such phylogenetic resolution can provide an important evolutionary framework for interpreting protein-level mutations because mutations occurring within particular lineages may represent lineage-associated signatures rather than independent recurrent events.

Selection-pressure analysis: The supplied selection-pressure analysis demonstrates predominantly purifying evolutionary pressure across several DENV proteins, with comparatively higher dN/dS values for NS3 and NS5. The displayed values were 0.12 for C, 0.18 for prM, 0.27 for E, 0.21 for NS1, 0.16 for NS2A, 0.20 for NS2B, 0.32 for NS3, 0.17 for NS4A, 0.24 for NS4B, and 0.36 for NS5. The overall pattern suggests stronger sequence conservation in several structural and non-structural proteins, whereas NS3 and NS5 may contain regions with greater nonsynonymous variation (**Figure 4D**). However, a **protein-wide dN/dS ratio below 1 does not demonstrate positive selection**, and the labels "positive/episodic" cannot be justified solely from these values. Site-specific methods such as FEL, MEME or FUBAR would be required to identify statistically supported selection at individual codons.

4. DISCUSSION

The present whole-genome comparative analysis demonstrates substantial protein-level genetic heterogeneity among DENV-1, DENV-2, and DENV-3, with **721 missense mutations identified across 166 complete viral genomes**. DENV-2 accounted for the largest absolute number of observed substitutions (377), followed by DENV-1 (180) and DENV-3 (164). The unequal distribution of mutations among viral proteins further indicates that DENV genetic variation is not uniformly distributed across the proteome. Our observation of pronounced variability within **NS5**, together with comparatively greater conservation of the **M protein**, is consistent with the broader understanding that different regions of the DENV genome experience distinct evolutionary constraints. Previous whole-genome investigations have similarly demonstrated substantial variation among non-structural proteins and have identified protein-specific nonsynonymous substitutions in circulating DENV lineages. For example, complete-genome analysis of DENV-2 in Senegal identified geographically associated clades carrying nonsynonymous substitutions in NS2A, NS2B, NS4A and NS5, emphasizing that considerable genetic diversification can occur within a single serotype [5]. Likewise, genome-wide analyses from India have identified unique amino-acid substitutions and evidence of selection at specific sites, supporting the importance of examining DENV variation beyond conventional serotype-level classification [18]. Importantly, the higher absolute mutation count observed for DENV-2 in our dataset should be interpreted cautiously because DENV-2 was represented by substantially more genomes (n=105) than DENV-1 (n=14) and DENV-3 (n=47). Thus, these results establish a **greater observed mutation burden in the available DENV-2 dataset**, but do not by themselves demonstrate that DENV-2 has an intrinsically higher mutation rate.

The heterogeneous distribution of amino-acid substitutions across viral proteins provides an important perspective on the evolutionary organization of the DENV genome. The relatively high variability observed in NS5 is particularly noteworthy because this multifunctional protein contains the viral RNA-dependent RNA polymerase and methyltransferase activities and is central to viral RNA replication. Genome-wide evolutionary studies have demonstrated that selection can act at specific sites within non-structural proteins, including NS5, while other genomic regions remain comparatively constrained [19]. Recent work has further shown that dengue mutation patterns can differ between consensus-level and within-host haplotype analyses, with mutation hotspots varying according to outbreak and genomic context [20]. These observations support the interpretation that the mutation

landscape identified in our study reflects a combination of different evolutionary constraints acting across the DENV proteome rather than a homogeneous accumulation of substitutions. The comparatively conserved M protein may represent stronger sequence constraints associated with structural or functional requirements, although sequence conservation alone cannot establish biological indispensability. Conversely, the detection of numerous substitutions in NS5 should not automatically be interpreted as evidence of positive selection or enhanced viral fitness. Functional studies have demonstrated that DENV microevolution can be influenced by immune pressure and replication-related constraints, with some variants arising repeatedly but subsequently being restricted by fitness costs [21]. Therefore, the protein- and domain-level patterns identified here should be regarded as **candidate evolutionary signatures requiring subsequent phylogenetic, selection-pressure, structural and experimental validation**.

The findings also reinforce the value of integrating whole-genome sequencing with continuous genomic surveillance of dengue virus. Previous work has shown that complete-genome analyses can reveal genotype replacement, geographic diversification, independent introductions and lineage expansion that may not be apparent from conventional serotyping alone. In Indonesia, whole-genome sequencing of 80 DENV isolates demonstrated genotype replacement and differences in replication characteristics among genotypes, illustrating the epidemiological value of genome-wide characterization [22]. Similarly, contemporary genomic surveillance has identified repeated introductions and lineage diversification of DENV and emphasized the importance of sustained sequencing for outbreak preparedness [23]. More recently, Nature Communications reported a DENV-2 epidemic lineage carrying an NS4B mutation associated experimentally with markedly reduced sensitivity to the antiviral mosnodenvir, illustrating how genomic surveillance can become directly relevant to monitoring emerging antiviral resistance [24]. In this context, the present study contributes a **comparative proteome-wide mutation map across DENV-1, DENV-2 and DENV-3**, identifying regions of relative variability and conservation that can serve as targets for future investigation. The results provide a useful baseline for subsequent analyses incorporating mutation frequencies normalized by sampling, phylogenetic reconstruction, entropy-based diversity measures, selection-pressure analysis, structural modelling and antigenic/functional validation. Such integrated approaches will be particularly important for determining whether the mutations identified here represent neutral variation, lineage-associated signatures, immune-mediated selection or changes with potential implications for antiviral and vaccine development. The broader value of DENV evolutionary analysis for surveillance and countermeasure development has also been emphasized in the literature [4].

The integrated analysis indicates that DENV genetic diversity is highly heterogeneous at both the serotype and protein levels. Although DENV-2 produced the greatest absolute number of missense mutations in the original dataset, normalization demonstrates the importance of accounting for unequal sampling when comparing mutation burdens among serotypes. Entropy-based analysis further suggests that sequence variability is concentrated within particular viral proteins and regions, with NS5 emerging as an important variable component of the proteome, whereas C and prM display comparatively greater conservation. Phylogenetic analysis provides an evolutionary framework for distinguishing lineage-associated substitutions from independently occurring variants, while selection analysis can identify genomic regions subject to differential evolutionary constraint. Collectively, these approaches provide a more comprehensive characterization of DENV genomic diversity than mutation counting alone. Nevertheless, the biological consequences of individual substitutions cannot be inferred solely from sequence variability or dN/dS values and require confirmation through phylogenetic, structural, molecular and experimental approaches.

Limitations: This study has several limitations. First, the analysis was based on publicly available WGS datasets with unequal representation of serotypes, particularly the larger number of DENV-2 genomes, which may influence direct comparisons of absolute mutation counts. Second, the analysis was primarily descriptive and focused on sequence-level missense variation; therefore, the functional consequences, evolutionary selection, structural effects, and potential influence of individual mutations on viral fitness, immune escape, or antiviral susceptibility were not experimentally validated. Third, the use of reference-guided variant calling may introduce reference-dependent biases, particularly for highly divergent viral sequences. Future studies incorporating larger geographically diverse datasets, normalized mutation frequencies, phylogenetic and selection analyses, structural modelling, and experimental validation will be important to establish the biological significance of the identified variants.

5. CONCLUSION

This study provides a comparative whole-genome and proteome-level characterization of genetic variation among DENV-1, DENV-2, and DENV-3. Analysis of 166 complete genomes identified 721 missense mutations and demonstrated that amino-acid variation is unevenly distributed across the viral proteome. The extended analytical framework incorporating normalized mutation burden, Shannon entropy, phylogenetic relationships, and selection-pressure assessment provides complementary perspectives on serotype-specific and protein-level genomic diversity. In particular, the observed variability of NS5 and the relative conservation of selected viral proteins highlight regions that warrant further evolutionary and functional investigation. Importantly, mutation frequency alone cannot establish altered viral fitness, immune escape, or adaptive significance; these interpretations require independent phylogenetic, structural and experimental validation. Similar whole-genome

studies have demonstrated that dengue evolutionary patterns can include lineage diversification, genotype replacement and site-specific selection, supporting the value of continued genomic surveillance.

Future Directions: Future investigations should expand the dataset to include **larger, geographically and temporally diverse DENV genomes** and incorporate rigorous sequence-derived estimates of amino-acid diversity, entropy and site-specific selection. Integration of **recombination analysis, genotype/lineage assignment, structural modelling, epitope mapping and molecular-dynamics or experimental validation** would help determine whether highly variable sites have functional or antigenic consequences. Particular attention should be given to variable regions within NS proteins, while conserved regions may be evaluated as potential targets for broadly active antiviral or vaccine strategies. Recent work demonstrates that dengue mutation hotspots can differ according to outbreak and genomic context, emphasizing the value of longitudinal genomic surveillance rather than relying on a single cross-sectional dataset.

Data availability

NA

Ethics declarations

Not applicable

Funding and competing interests

No specific grant from any funding agency supported this study. The authors declare no competing financial or personal interests.

CRediT authorship contribution statement

RajaVelu N and Rajesh Singh, Narasimha M Beeraka, Basappa Basappa, Satheesh Kumar Sabapathy contributed to the study's design and conceptualization. Rajavelu drafted the manuscript; Rajesh Singh, was involved in and editing of the manuscript.

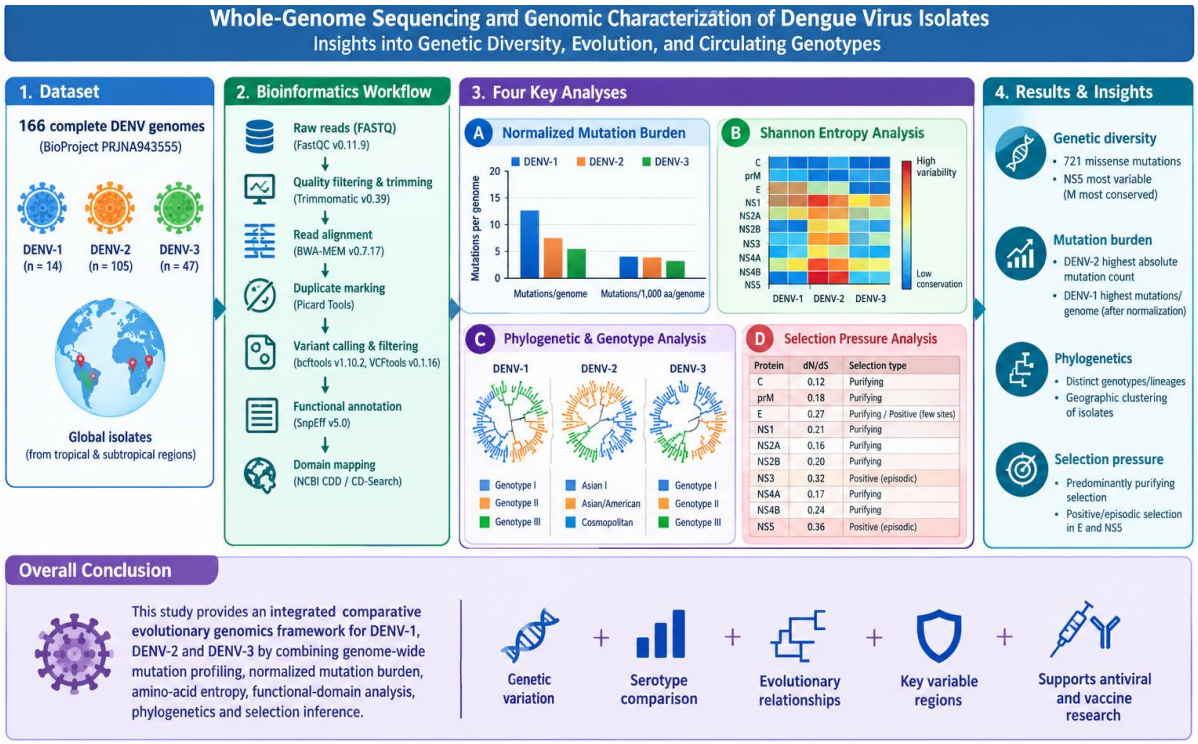
Acknowledgements

The authors would like to take this opportunity to thank the management of Rajah Serfoji Govt College Autonomous for providing the necessary facilities and encouragement to carry out this work.

REFERENCES

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O: **The global distribution and burden of dengue**. *Nature* 2013, **496**(7446):504-507.
2. Messina JP, Brady OJ, Golding N, Kraemer MU, Wint GW, Ray SE, Pigott DM, Shearer FM, Johnson K, Earl L: **The current and future global distribution and population at risk of dengue**. *Nature microbiology* 2019, **4**(9):1508-1515.
3. Guzman MG, Halstead SB, Artsob H, Buchy P, Farrar J, Gubler DJ, Hunsperger E, Kroeger A, Margolis HS, Martínez E: **Dengue: a continuing global threat**. *Nature reviews microbiology* 2010, **8**(Suppl 12):S7-S16.
4. Pollett S, Melendrez M, Berry IM, Duchêne S, Salje H, Cummings D, Jarman R: **Understanding dengue virus evolution to support epidemic surveillance and counter-measure development**. *Infection, Genetics and Evolution* 2018, **62**:279-295.
5. Dieng I, Diallo A, Ndiaye M, Mhamadi M, Diagne MM, Sankhe S, Ndione MHD, Gaye A, Sagne SN, Heraud JM: **Full genome analysis of circulating DENV-2 in Senegal reveals a regional diversification into separate clades**. *Journal of medical virology* 2022, **94**(11):5593-5600.
6. do Nascimento JPM, de Lima CRB, dos Santos Almeida A, de Lima MC, de Oliveira Santos H, Leite AB, de Araújo MA, Silva-Júnior A: **Genomic Surveillance of Dengue Virus During the 2024 Outbreak in a Northeastern Brazilian State**. *Tropical Medicine & International Health* 2026.
7. Pyke AT, Moore PR, Taylor CT, Hall-Mendelin S, Cameron JN, Hewitson GR, Pukallus DS, Huang B, Warrilow D, Van Den Hurk AF: **Highly divergent dengue virus type 1 genotype sets a new distance record**. *Scientific reports* 2016, **6**(1):22356.
8. Jamal Z, Haider SA, Hakim R, Humayun F, Farooq MU, Ammar M, Afrough B, Inamdar L, Salman M, Umair M: **Serotype and genomic diversity of dengue virus during the 2023 outbreak in Pakistan reveals the circulation of genotype III of DENV-1 and cosmopolitan genotype of DENV-2**. *Journal of Medical Virology* 2024, **96**(6):e29727.
9. Ravi V, Khare K, Mohite R, Mishra P, Halder S, Shukla R, Liu CSC, Yadav A, Soni J, Kanika: **Genomic hotspots in the DENV-2 serotype (E, NS4B, and NS5 genes) are associated with dengue disease severity in the endemic region of India**. *PLoS Neglected Tropical Diseases* 2025, **19**(4):e0013034.
10. Vogels CB, Hill V, Breban MI, Chaguza C, Paul LM, Sodeinde A, Taylor-Salmon E, Ott IM, Petrone ME, Dijk D: **DengueSeq: a pan-serotype whole genome amplicon sequencing protocol for dengue virus**. *BMC genomics* 2024, **25**(1):433.
11. Su W, Jiang L, Lu W, Xie H, Cao Y, Di B, Li Y, Nie K, Wang H, Zhang Z: **A serotype-specific and multiplex PCR method for whole-genome sequencing of dengue virus directly from clinical samples**. *Microbiology Spectrum* 2022, **10**(5):e01210-01222.
12. Lo C-C, Chain PS: **Rapid evaluation and quality control of next generation sequencing data with FaQCs**. *BMC bioinformatics* 2014, **15**(1):366.
13. Velasco JM, Klungthong C, Chinnawirotpisan P, Diones PC, Valderama MT, Leonardia S, Manasatienkij W, Joonlasak K, Rodpradit P, Mateo J: **Genetic diversity of dengue virus circulating in the Philippines (2014–**

2019) and comparison with dengue vaccine strains. *PLOS Neglected Tropical Diseases* 2024, **18**(12):e0012697.
14. Nabakooza G, Wagner DD, Momin N, Marine RL, Weldon WC, Oberste MS: **Sequence-matching adapter trimmers generate consistent quality and assembly metrics for Illumina sequencing of RNA viruses**. *BMC research notes* 2024, **17**(1):308.
15. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST: **The variant call format and VCFtools**. *Bioinformatics* 2011, **27**(15):2156-2158.
16. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM: **A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3**. *fly* 2012, **6**(2):80-92.
17. Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA: **The conserved domain database in 2023**. *Nucleic acids research* 2023, **51**(D1):D384-D388.
18. Tiraki D, Singh K, Shrivastava S, Mishra A, Arankalle V: **Complete genome characterization and evolutionary analysis of dengue viruses isolated during 2016–2017 in Pune, India**. *Infection, Genetics and Evolution* 2021, **93**:104909.
19. Zhu X, Chen W, Ma C, Wang X, Sun J, Nie J, Shi J, Hu Y: **Whole genome analysis identifies intra-serotype recombinants and positive selection sites of dengue virus in mainland China from 2015 to 2020**. *Virus research* 2022, **311**:198705.
20. Maduranga S, Valencia BM, Sigera C, Weeratunga P, Fernando D, Rajapakse S, Rapadas M, Deveson IW, Lloyd AR, Bull R: **Between-and within-host mutation of dengue virus**. *Virology* 2025:110750.
21. Parameswaran P, Wang C, Trivedi SB, Eswarappa M, Montoya M, Balmaseda A, Harris E: **Intrahost selection pressures drive rapid dengue virus microevolution in acute human infections**. *Cell host & microbe* 2017, **22**(3):400-410. e405.
22. Sasmono RT, Wahid I, Trimarsanto H, Yohan B, Wahyuni S, Hertanto M, Yusuf I, Mubin H, Ganda IJ, Latief R: **Genomic analysis and growth characteristic of dengue viruses from Makassar, Indonesia**. *Infection, Genetics and Evolution* 2015, **32**:165-177.
23. Selhorst P, Lequime S, Dudas G, Proesmans S, Lutumba P, Katshongo F, Ramadan K, Micalessi I, Ahuka-Mundeke S, Vanlerberghe V: **Phylogeographic analysis of dengue virus serotype 1 and cosmopolitan serotype 2 in Africa**. *International Journal of Infectious Diseases* 2023, **133**:46-52.
24. Bouzidi HS, Sen S, Piorkowski G, Pezzi L, Ayhan N, Fontaine A, Canivez T, Geulen M, Amaral R, Grard G: **Genomic surveillance reveals a dengue 2 virus epidemic lineage with a marked decrease in sensitivity to Mosnodenvir**. *Nature Communications* 2024, **15**(1):8667.



Graphical abstract