IN SILICO ANALYSIS OF COMMON MUTATIONS FOUND IN THE DENGUE VIRAL GENOME SEQUENCES FROM BANGLADESH

Bioresearch Communications
Volume 11, Issue 1, January 2025

DOI:
doi.org/10.3329/brc.v11i1.78889

Tasin Ahmed, Rowshan Ara Begum, Reza Md Shahjahan and Khandaker Ashfaqul Muid*

Branch of Genetics and Molecular Biology, Department of Zoology, University of Dhaka, Dhaka-1000, Bangladesh

ABSTRACT

Dengue is a mosquito-borne viral disease caused by four serotypes of the dengue virus (DENV-1, DENV-2, DENV-3, and DENV-4), which has impacted human populations for decades in tropical and subtropical regions, including Bangladesh. Understanding the genetic variability and mutation patterns of DENV-2 and DENV-3 is critical for developing effective control strategies, as these serotypes are more prevalent in Bangladesh. The E/NS1 gene junction, which comprises less than 3% of the DENV genome, is a recognized hotspot for mutation. In this study, 56 E/NS1 junction sequences of DENV-2 from 36 countries, including Bangladesh, were analyzed *in silico*. The Bangladeshi sequences were compared with the DENV-2 prototype strain (New Guinea C) and sequences from other countries. Analysis of Bangladeshi DENV-2 isolates revealed only one amino acid substitution (isoleucine to valine at position 742, I742V), caused by a nucleotide change (ATT to GTC). Phylogenetic analysis placed all Bangladeshi DENV-2 isolates within the Cosmopolitan genotype. Similarly, 35 E/NS1 junction sequences of DENV-3 from 19 countries, including Bangladesh, were analyzed. Sequences from Bangladesh obtained in 2002 (eight isolates) and 2020 (two isolates) were compared with the DENV-3 prototype strain (Philippines H87). The 2020 isolates exhibited two common amino acid substitutions (A759V and V769A), while the 2002 isolates showed three substitutions (S727G, A759V, and V769T). Phylogenetic analysis revealed a genotypic shift in Bangladeshi DENV-3 isolates, with 2002 isolates belonging to Genotype II and 2020 isolates classified as Genotype I. Envelope protein modeling for both DENV-2 and DENV-3 Bangladeshi isolates, compared to their respective prototype strains, revealed no significant structural changes in the E protein. However, protein stability analysis, based on changes in free energy due to amino acid substitutions, indicated a potential impact on the stability of the mutant E proteins compared to the prototypes. Further studies are needed to explore the clinical implications of these mutations.

KEYWORDS: Dengue virus, hotspot, common mutations, dengue virus genome.

RECEIVED: 26 August 2024, ACCEPTED: 28 November 2024

TYPE: Short Communication

***CORRESPONDING AUTHOR: Dr. Khandaker Ashfaqul Muid, Branch of Genetics and Molecular Biology, Department of Zoology, University of Dhaka, Dhaka-1000, Bangladesh**
Email: muid.zoo@du.ac.bd

Introduction

Dengue virus, a member of Flaviviridae family, is a mosquito-borne Flavivirus which causes mild to severe illness as well as deaths each year in tropical and subtropical countries and it is very common in this Indian subcontinent (Kanakaratne *et al*., 2009). Infected female mosquitoes, primarily *Aedes aegypti* and *Aedes albopictus*, bite humans and transmit the disease (Rashed, 2021). Over the last few decades, the occurrence of dengue fever has doubled. Each your approximately 50 million new cases are identified of which 10000 cases end in deaths (Edussuriya, Deegalla and Gawarammana, 2021). Dengue virus (DENV) is currently the most prevalent cause of tropical arboviral infections (Edussuriya, Deegalla and Gawarammana, 2021).

Dengue infection is present in six WHO regions and almost in 125 countries of the world according to WHO (Powell, 2018). In 2018, Powell et al. suggested that the number of infection per year could even be 400 million. The alarming fact is almost 50% of world population lives in dengue endemic areas and under risk of severe DENV disease (Powell, 2018). But the South East Asian countries are subject to more than 18 times Dengue cases in comparison to other countries worldwide (Lai *et al.*, 2018). As a result, 8 South East Asian countries are now considered as hyperendemic for Dengue virus with all the four serotypes widespread in them (Guo *et al.*, 2017). The 20,000-plus dengue cases reported in 2021 thus far is the second most cases reported in a year since the first cases were reported in 2000. But in 2019, Bangladesh reported more than 101,000 cases which is by far the biggest Dengue outbreak in Bangladesh since this virus was first identified here. DENV-1 and DENV-2 were the most common serotypes in Bangladesh from 2013 to 2016, however DENV-3 and DENV-4 have lately become more common. In a large proportion of cases, co-detection of dengue serotypes and combinations has been discovered.

Dengue virus has four serotypes: DENV1, DENV-2, DENV3, and DENV4 that are antigenically and genetically unique. Infection with one DENV serotype is thought to provide long-term immunity to the homologous serotype but not to the heterologous serotypes (Kanakaratne *et al.*, 2009). As a result, people might be infected with various serotypes throughout their lives. Two distinct genotypes of DENV-1, four of DENV-2, four of DENV-3 and three of DENV-4 have been identified by partial nucleotide sequencing of a particular 600bp region of the dengue envelope protein (Rashed, 2021). Restriction enzymes and primer extension sequencing can be used to show the lineage of each serotype, and intra-serotypic recombination events amongst dengue viruses are becoming more widespread (Mahy, 2008). Inter-serotypic recombination may also occur, since there have been more instances of simultaneous infections with two serotypes as a result of the increasing prevalence of numerous serotypes co-circulating in a region (hyperendemicity) (Rashed, 2021).

Dengue is a single stranded, positive-sense RNA virus, the length of which is around 10,700 kbp (Osman *et al*., 2008). It has a single open reading frame that codes for the viral polyprotein, a 5'; untranslated region (UTR) of around 100 nucleotides, and a 3'; UTR of about 400 nucleotides. Unlike cellular mRNA, the DENV genome is not 3'; polyadenylated, and a 1 7-methyl guanosine cap structure is connected to the viral genome's 5' end (Harapan *et al.*, 2020). The viral genome encodes three structural proteins which are capsid (C), membrane (M) and envelope (E) and seven nonstructural (NS) proteins which are NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5(Osman *et al*., 2008). The surface of mature Dengue virus is smooth and of about 50nm diameter, whereas the immature virion has a spiky surface that is of approximately 60 nm in diameter (Harapan *et al.*, 2020). The virus is made up of three structural proteins that make up the virus's coat.

E/NS1 junction is a 240bp long nucleotide sequence in the DENV-2 genome and 279bp long nucleotide sequence in DENV-3 genome (Osman *et al.*, 2008). This portion of viral genome can be a best source of nucleotide sequences to compare different strains of same serotypes or even to compare across different serotypes for deriving evolutionary knowledge that can be explained epidemiologically. This particular junction of the genome is chosen because this area shows a uniform rate of random mutation without any hypervariable areas that can get affected by immune selection of epitopes. High nucleotide variation sequences or 'Hot spots" tend to occur in the epitope encoding areas of the E gene and these might overshadow the random mutation events with crucial evolutionary relationship clues. So sequence analysis of this junction proved to be very informative about silent mutations that give information about evolutionary relationships (Rico-hesse, 1989).

This research work aimed to analyze the mutation trend in this junction in different countries around the world giving special attention to the sequences found in Bangladesh. Different aspects of bioinformatics analysis have been done to study the preexisting sequences of DENV type 2 & 3.

Materials and Methods

This research is solely based on bioinformatics analysis of the preexisting sequences from Bangladesh as well as other countries (New Guinea, India, Pakistan, Nepal, Sri Lanka, Myanmar, Malaysia, Indonesia, Thailand, Philippines, Brunei, Vietnam, Taiwan, Singapore, Cambodia, China, Jamaica, Tonga, Mexico, Puerto Rico, Burkina Faso, Venezuela, Tahiti, Colombia, Seychelles, Ivory Coast, Senegal, New Caledonia, Trinidad, Australia, Fiji, Peru, Brazil, Cuba, Dominican Republic, Martinique, French Polynesia, Mozambique, Nicaragua and Saint Lucia- all accession numbers are mentioned in **Table-2** and **Table-4**) from where dengue virus has been isolated and sequenced. Initially, dengue whole genome sequences were retrieved from literature mining in Google Scholar as well as from NCBI GenBank.

E/NS1 junction sequences of serotype 2 and serotype 3 dengue virus were extracted from the whole genome sequences followed by performing Nucleotide BLAST analysis where similarities and divergences of the sequences compared to the prototypes were presented in **Table-1** and **Table-3**. After that, prototype sequences of both DENV-2 (New Guinea C strain-M29095) and DENV-3 (Philippines H87-KU050695) were extracted through literature mining. The FASTA formats of multiple sequences were aligned in the MEGA11 software using Clustal*W* program. Following this, amino acid sequences of the isolates were generated from translating the nucleotide sequences in the MEGA11 and were aligned with the assistance of the same software. Synonymous and non-synonymous mutations were detected from both nucleotide and amino acid sequences of the junction. The whole dataset of the mutations are recorded and presented in the **Table-2** and **Table-4**.

To determine evolutionary ancestry of the sequences, neighbor-joining phylogenetic tree has been constructed for both DENV-2 and DENV-3 sequences in MEGA11. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. The evolutionary distances were calculated using the Maximum Composite Likelihood method (Tamura, Nei and Kumar, 2004) and are in the units of the number of the base substitution per site. This analysis involved 56 nucleotide sequences for DENV-2 where eight were from Bangladesh and 35 nucleotide sequences DENV-3 where ten were from Bangladesh. Pairwise deletion option was used for removing all ambiguous positions for each sequence pair. There were a total of 240 positions for DENV-2 and 280 positions for DENV-3 in the final dataset.
Finally, protein structure of the envelope protein of dengue virus was predicted with and without the mutations through using SWISS-Model tool. Firstly, the Template amino acid sequence for building the model was scanned in BLAST against the SWISS-Model template library. The templates with the best quality were selected for the construction of the protein model. After that, models were built based on the target-template alignment using ProMod3 of the SWISS-Model. Furthermore, protein stability changes upon mutation were predicted through using I-mutant 2.0 tool (Capriotti, Fariselli and Casadio, 2005).

Results and Discussion

DENV-2 E/NS1 Sequence Analysis

The aligned data of E/NS1 junction of the most recent DENV-2 isolates in Bangladesh are showed in **Fig. 1**.

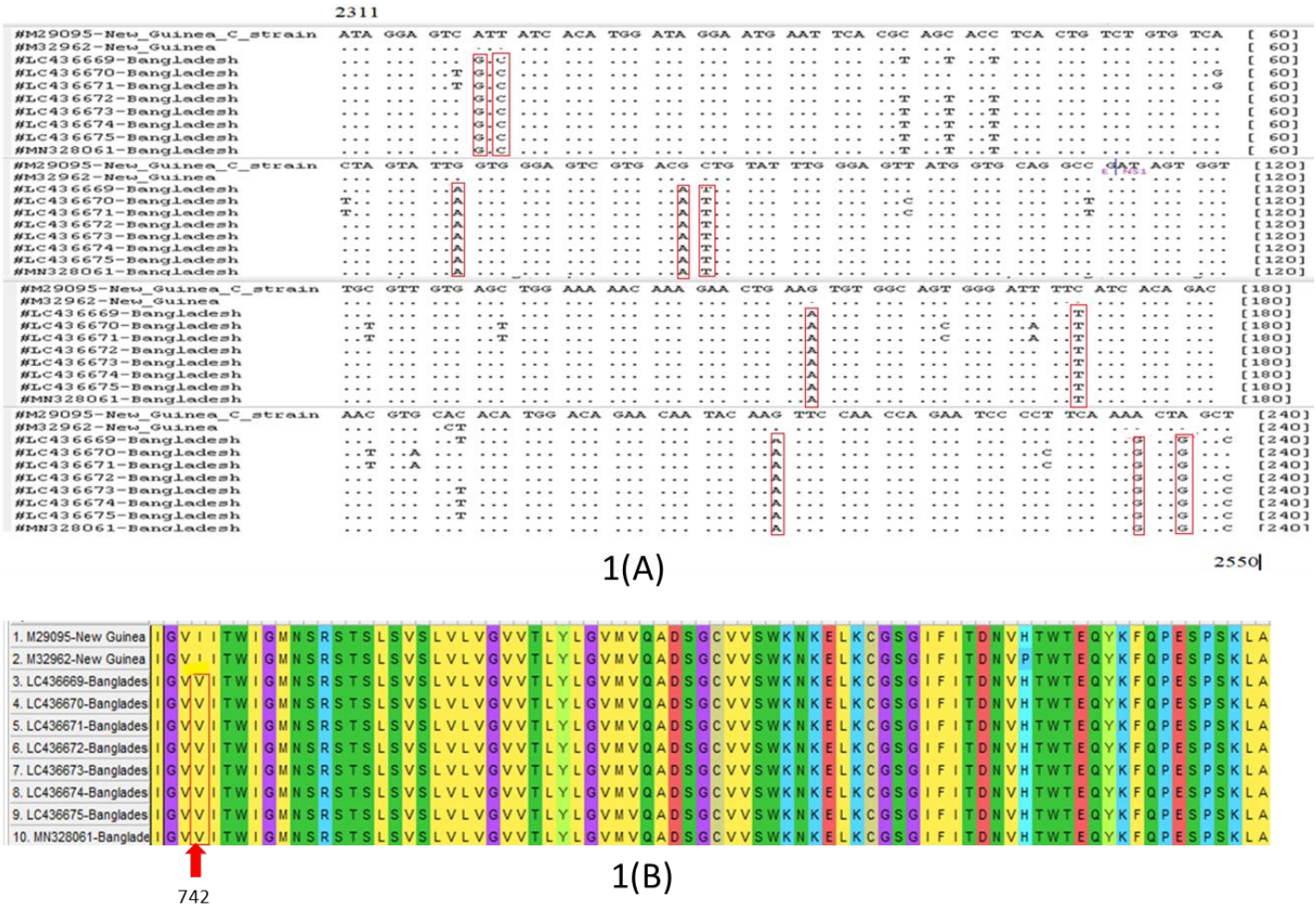

Fig. 1. (A). 240bp aligned E/NS1 nucleotide sequences of the most recent Bangladeshi DENV-2 isolates analyzed in this study. Upper most sequence is the DENV-2 prototype New Guinea C strain. Here, ten common nucleotide differences in all the Bangladeshi isolates at positions 2320, 2322, 2379, 2394, 2395, 2463, 2481, 2520, 2544 and 2547 (indicated as red marked block) can be seen five of which are inside the E region and the other five are inside the NS1 region. **(B).** Only amino acid change in the 742 position (indicated as red arrow) of DENV-2 E/NS1 junction sequence of the Bangladeshi isolates are shown here in red marked block.

This alignment was done using MEGA11. Here, the reference sequence is the DEN-2 prototype New Guinea C strain (M29095). Of these eight Bangladeshi DENV-2 sequences, seven were obtained from the work of Suzuki *et al.* in Dhaka city on 2020 and one (MN328061) was obtained from the work of Saha et al in Dhaka city, 2019. Analysis of these eight E/NS1 sequences shows 10 common nucleotide changes in all the Bangladeshi isolates compared to the prototype sequence. Identical nucleotide percentages range from only 90.83% to 95.28% while comparing the Bangladeshi sequences with the prototype New Guinea C strain **(Table 1).**

Table 1. Similarity percentages and divergence between DENV-2 Bangladeshi isolates and prototype reference strain New Guinea C.

Sequence accession no.	Similarity (%)	Divergence (%)
LC43669	94.85	5.15
LC436670	90.83	9.17
LC436671	90.83	9.17
LC436672	95.28	4.72
LC436673	94.85	5.15
LC436674	94.85	5.15
LC436675	94.85	5.15
MN328061	95.28	4.72

Individually, LC436669, LC436670, LC436671, LC436672, LC436673, LC436674, LC436675 and MN328061 show base substitutions at total 15,22,22,14,15,15,15 and 14 positions. But considering the amino acids, there is 100% similarity among the Bangladeshi DENV-2 isolates. Comparing with the prototype, all of the isolates from Bangladesh show a difference of only one amino acid from isoleucine to valine at position 742. This change occurs because of two nucleotide base changes occurred at positions 2320 and 2322. These two substitutions of bases cause change of a codon from ATT to GTC.

These eight Bangladeshi DENV-2 E/NS1 nucleotide and amino acid sequences were then aligned and compared with 56 more E/NS1 nucleotide and amino acid sequences of DENV-2 from different geographical regions of the world where the New Guinea C strain was used as the reference sequence. A total of 36 countries' sequence data were retrieved from GenBank and aligned in MEGA11 software.

Nucleotide substitution resulting in codon and amino acid change through non-synonymous mutations in the DENV-2 E/NS1 sequences obtained from different geographical regions of the world are given in **Table 2**.

Table 2. Mutational landscape identified in the DENV-2 E/NS1 sequences obtained from different geographical regions of the world.

Accession no.	Country	No. of nucleotide substitution	Non-synonymous mutation	Synonymous mutation	Codon change in NS mutations	Amino acid change in NS mutations
LC43669	Bangladesh	15	2	13	ATT-GTC	I-V(742)
LC436670	Bangladesh	22	2	20	ATT-GTC	I-V(742)
LC436671	Bangladesh	22	2	20	ATT-GTC	I-V(742)
LC436672	Bangladesh	14	2	12	ATT-GTC	I-V(742)
LC436673	Bangladesh	15	2	13	ATT-GTC	I-V(742)
LC436674	Bangladesh	15	2	13	ATT-GTC	I-V(742)
LC436675	Bangladesh	15	2	13	ATT-GTC	I-V(742)
MN328061	Bangladesh	14	2	12	ATT-GTC	I-V(742)
M32962	New Guinea	2	2	0	CAC-CCT	H-P(801)
JQ955623	India	20	3	17	ATT-GTC GTC-ATC	I-V(742) V-I(764)
KJ010186	Pakistan	20	4	16	ATT-GTC AAA-ACA GAA-GCA	I-V(742) K-T(786) E-A(787)
MW730833	Nepal	14	3	11	ATT-GTC ATC-GTC	I-V(742) I-V(796)
D44546	Myanmar	9	2	7	GTC-ATT	V-I(764)
M32938	Sri Lanka	20	2	18	ATT-GTC	I-V(742)
M32940	Sri Lanka	19	2	18	ATT-GTC	I-V(742)

AF400004	Malaysia	12	3	9	ATT-GTC ATG-ACG	I-V(742) M-T(748)
AJ556809	Malaysia	13	2	11	ATT-GTC	I-V(742)
AY858035	Indonesia	13	2	11	ATT-GTC	I-V(742)
AY858036	Indonesia	14	3	11	ATT-GTC ATC-GTC	I-V(742) I-V(796)
D44542	Thailand	7	2	5	GTC-ATT	V-I(764)
M32934	Indonesia	19	5	14	ATT-GTC TTG-TTT ATT-GTT CCT-CGT	I-V(742) L-F(761) I-V(796) P-R(814)
U87321	Thailand	12	1	11	GTT-GCT	V-A(771)
U87339	Thailand	7	2	5	GTC-ATT	V-I(764)
U87380	Thailand	6	0	6	N/A	N/A
M32941	Thailand	9	5	4	TCA-ACA GGA-GAA GTC-ATT CCT-CTT	S-T(758) G-E(763) V-I(764) P-L(814)
AF022434	Thailand	7	2	5	GTC-ATT	V-I(764)
M32932	Philippines	9	1	8	ATG-GTG	M-V(772)
EU031573	Brunei	13	2	11	ATT-TTC	I-F(742)
EU031574	Brunei	10	2	9	ATT-GTC	I-V(742)
M32942	Vietnam	14	2	12	GTT-GCT GAT-AAT	V-A(771) D-N(776)
M32949	Taiwan	7	1	6	ATG-GTG	M-V(772)
EU081180	Singapore	15	2	13	ATT-GTT	I-V(742)
EU687248	Vietnam	8	2	6	GTC-ATT	V-I(764)
FJ639697	Cambodia	17	4	13	GTG-ATA GTT-GCC	V-I(765) V-A(771)
MH827539	China	19	5	14	GTC-GCT ATT-GTC GTC-ATC	V-A(741) I-V(742) V-I(764)
M20558	Jamaica	13	1	12	GTT-GCT	V-A(771)
M32935	Tonga	19	2	17	GTC-ATC ATC-GTC	V-I(764) I-V(796)
M32933	Mexico	21	2	19	GTC-ATC ATC-GTC	V-I(764) I-V(796)
M32936	Puerto Rico	21	3	18	TCA-ACA GTC-ATC ATC-GTC	S-T(758) V-I(764) I-V(796)
M32937	Jamaica	16	0	16	N/A	N/A
M32939	Burkina Faso	20	2	18	ATC-GTC TCA-CCA	I-V(796) S-P(814)
M32971	Venezuela	21	2	19	GTC-ATC ATC-GTC	V-I(764) I-V(796)
M32943	Tahiti	19	2	17	GTC-ATC ATC-GTC	V-I(764) I-V(796)
M32946	Colombia	21	2	19	GTC-ATC ATC-GTC	V-I(764) I-V(796)
M32952	Seychelles	15	1	14	ATC-GTC	I-V(796)

M32955	Ivory Coast	40	8	32	ATA-ACA ATT-GTT TCA-ACC GTC-ATC ATC-GTA GAA-GAC	I-T(739) I-V(742) S-T(758) V-I(764) I-V(796) E-D(811)
M32957	Senegal	39	8	31	ATT-GTT TCA-ACC GTC-ATC CTG-ATG ATC-GTA GAA-GAC	I-V(742) S-T(758) V-I(764) L-M(788) I-V(796) E-D(811)
M32961	New Caledonia	20	2	18	GTC-ATC ATC-GTC	V-I(764) I-V(796)
M32969	Trinidad	24	3	21	GTC-ATC AGT-AAT ATC-GTC	V-I(764) S-N(792) I-V(796)
AY037116	Australia	14	2	12	ATT-GTC	I-V(742)
HM582099	Fiji	20	2	18	GTC-ATC ATC-GTC	V-I(764) I-V(796)
AF100467	Peru	19	2	17	GTC-ATC ATC-GTC	V-I(764) I-V(796)
GU131883	Brazil	18	2	16	GTT-GCT GTT-ATT	V-A(771) V-I(780)
AY702039	Cuba	15	1	14	GTT-GCT	V-A(771)
AB122022	Dominican Republic	16	2	14	GTT-GCT GTT-ATT	V-A(771) V-I(780)

In these alignment dataset only 4 common non-synonymous mutations can be found; three of them are on the E region of the junction and only one is in the NS1 region. The amino acid substitution from isoleucine to valine at the position 742 is found in the isolates from 13 countries. The amino acid substitution at position 764 from valine to isoleucine has been found in the isolates from 21 countries. Isolates from seven countries were found to have a change in amino acid position 771 where valine is substituted by alanine. The amino acid substitution from isoleucine to valine at 796 position has been found in the isolates from 16 countries. There are also 15 random amino acid changes were found which are shown in the Supplementary **Fig. 1(A).** A neighbor-joining phylogenetic tree of all the 56 sequences of DENV-2 E/NS1 junction has been constructed using the software MEGA11. The phylogenetic tree showed in **Fig 2** generated 4 distinct clusters of the four main genotypic groups. Six of the Bangladeshi isolates (LC436669, LC436672, LC436673, LC436674, LC436675 and MN328061) are clustered with the isolates of Brunei, Malaysia, Indonesia, Singapore and Australia in the Cosmopolitan genotype group. Other two Bangladeshi isolates (LC436670 and LC436671) were grouped with India, Pakistan, Sri Lanka, Burkina Faso and Seychelles also in Cosmopolitan genotype but in a different lineage. According to Rico-Hesse (1990) and Lewis *et al.* (1993), DENV-2 isolates from Indonesia, Malaysia, Brunei, Singapore, countries in subcontinent, Burkina Faso and Seychelles belong to the same genotype. Therefore, our phylogenetic study proves that DENV-2 isolates from Bangladesh belong to the same cluster as them.

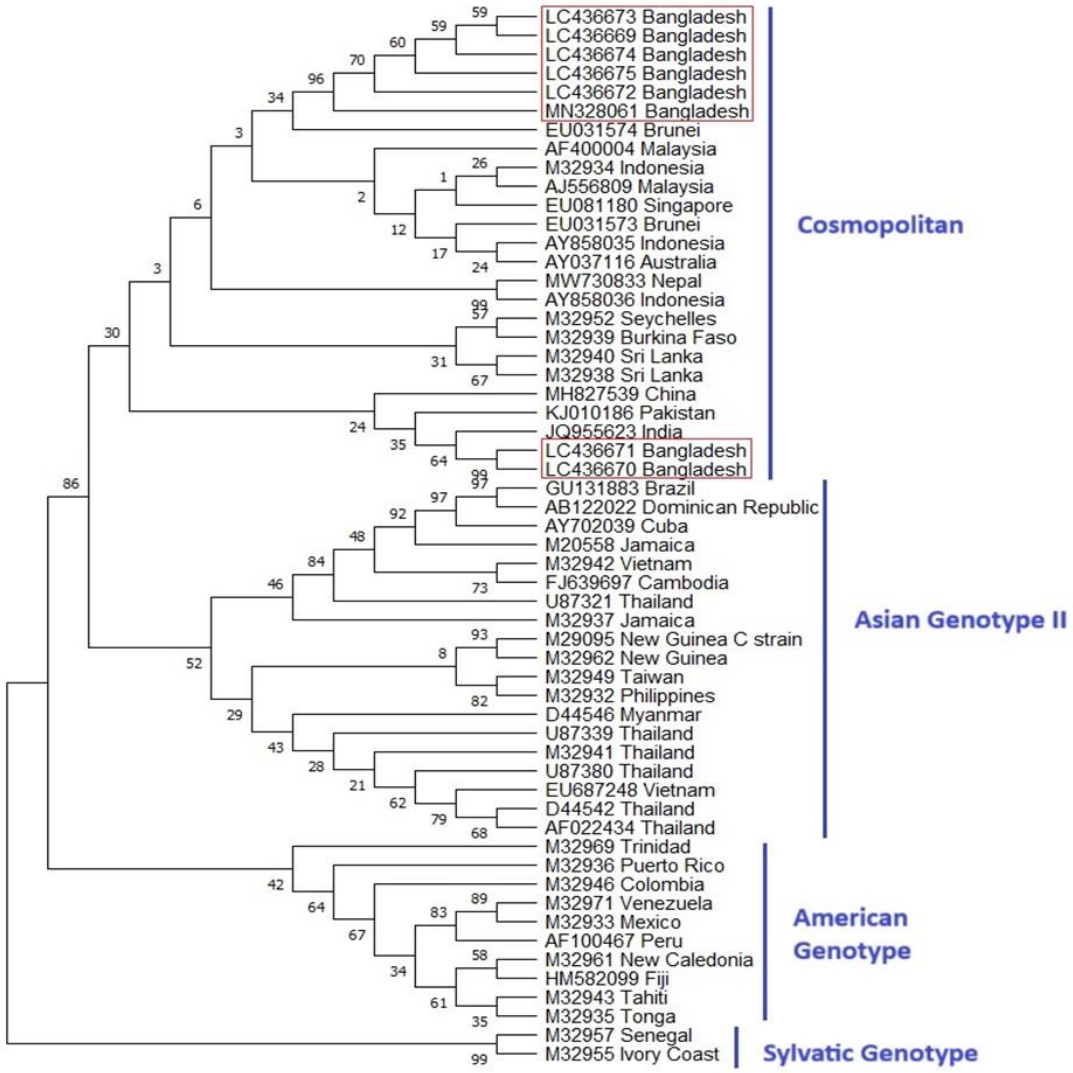

Fig. 2. Evolutionary connection between the eight Bangladeshi DENV-2 E/NS1 sequences with 48 other DENV-2 E/NS1 sequences is shown in the neighbor-joining tree above. Positions of Bangladeshi isolates in the clusters are indicated with red marked rectangles. Bootstrap values are displayed at the top of branches which joined the genotypic groups. The neighbor-joining method was used in the MEGA11 software.

Protein Structure Prediction of DEN-2 E Protein

After analyzing the sequences of E/NS1 junction of DENV-2 in Bangladesh, it was found that only one amino acid change occurs at the 742 position (isoleucine to valine) of the protein region. This amino acid substitution is situated at the E protein region and the NS1 protein region of the junction remains completely conserved in the Bangladeshi isolates. Protein structure building and comparison of the DENV-2 E or envelope protein of Bangladeshi isolates with the reference New Guinea C strain is shown in **Fig. 3.**

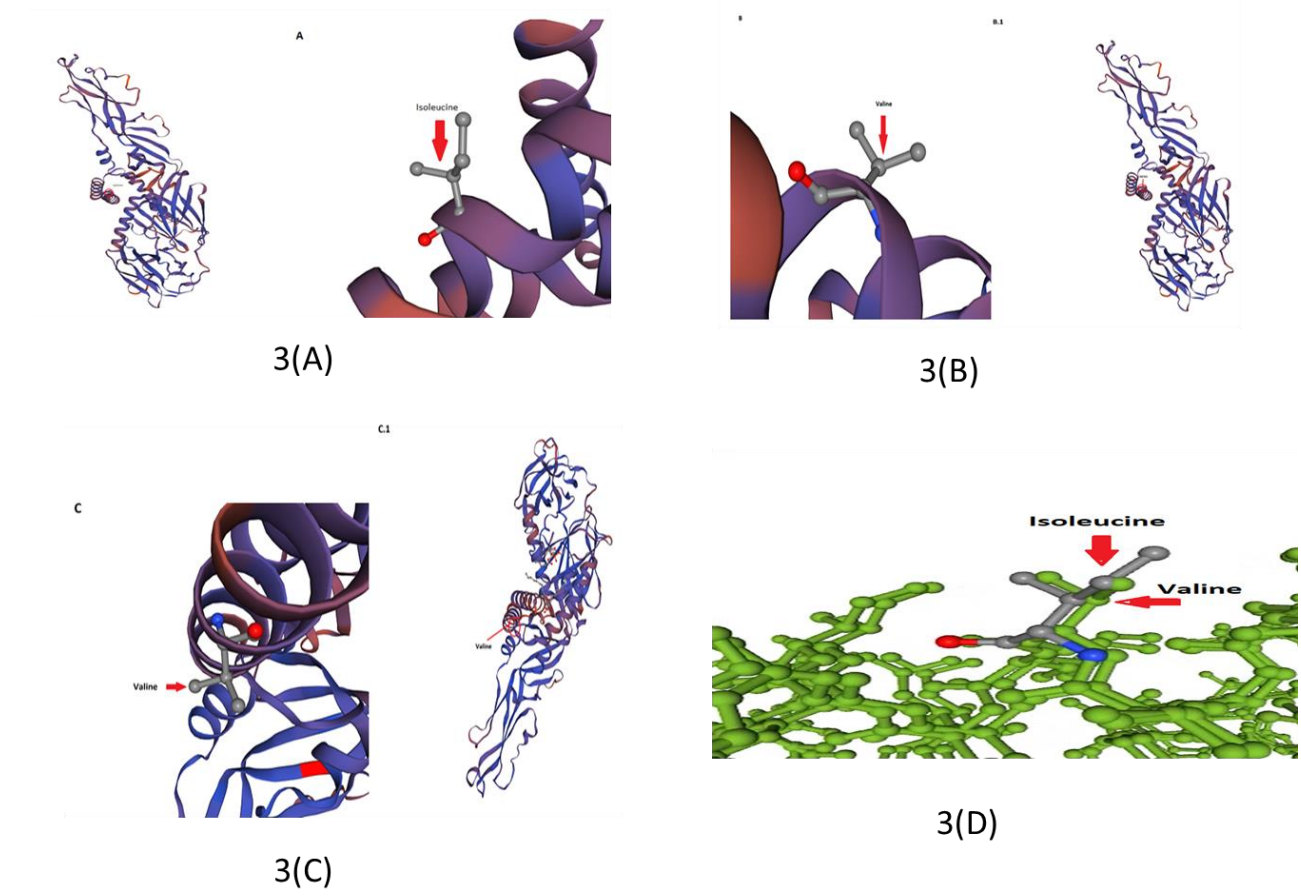

Fig. 3. (A). Protein modeling of DENV-2 Envelope protein of prototype New Guinea C strain and the amino acid which is substituted in Bangladeshi strains. **(B).** E Protein modeling and amino acid substitution (valine) in Bangladeshi isolate LC43669. **(C).** E Protein SWISS modeling and amino acid substitution (valine) in Bangladeshi isolate LC436671. **(D).** E protein form the reference isolate M29095 and Bangladeshi isolate LC436669 are superimposed to compare the amino acid substitution form isoleucine to valine which are indicated with red arrows.

This model building and comparison is done in the SWISS-Model software. In the envelope protein of New Guinea C strain amino acid isoleucine is substituted by amino acid valine in both LC43669 and LC436671 isolates which are Bangladeshi. This change of amino acid does not show any significant structural change in the E protein of DENV-2. However, protein stability alteration because of the particular mutation is predicted by using I-Mutant 2.0 tool where the free energy change of the mutant protein shows the value -0.82 ($\Delta\Delta G$= -0.82). This value indicates that there might be a slight decrease in stability of the E protein of mutant DENV-2 caused by the substitution from isoleucine to valine amino acid. Isoleucine is more hydrophobic than valine (Pommie *et al.*, 2004). Hydrophobicity is a vital biochemical feature of amino acids which determines the nature of side chain packing and protein folding (White, 1992). When a certain hydrophobic interaction is interrupted due to substitution of amino acid, it may cause destabilization of protein and about two-third times this destabilization affects the function of protein even if the structure is intact (Bromberg and Rost, 2009). So, changing the amino acid to isoleucine to valine which is a less hydrophobic one may change the function of this DENV-2 envelope protein. Further thorough experimentation is needed on this particular protein to confirm this **(Table 5).**

DEN-3 E/NS1 Sequence Analysis

The aligned data of E/NS1 junction of the most recent DENV-3 isolates in Bangladesh are showed in **Fig. 4**.

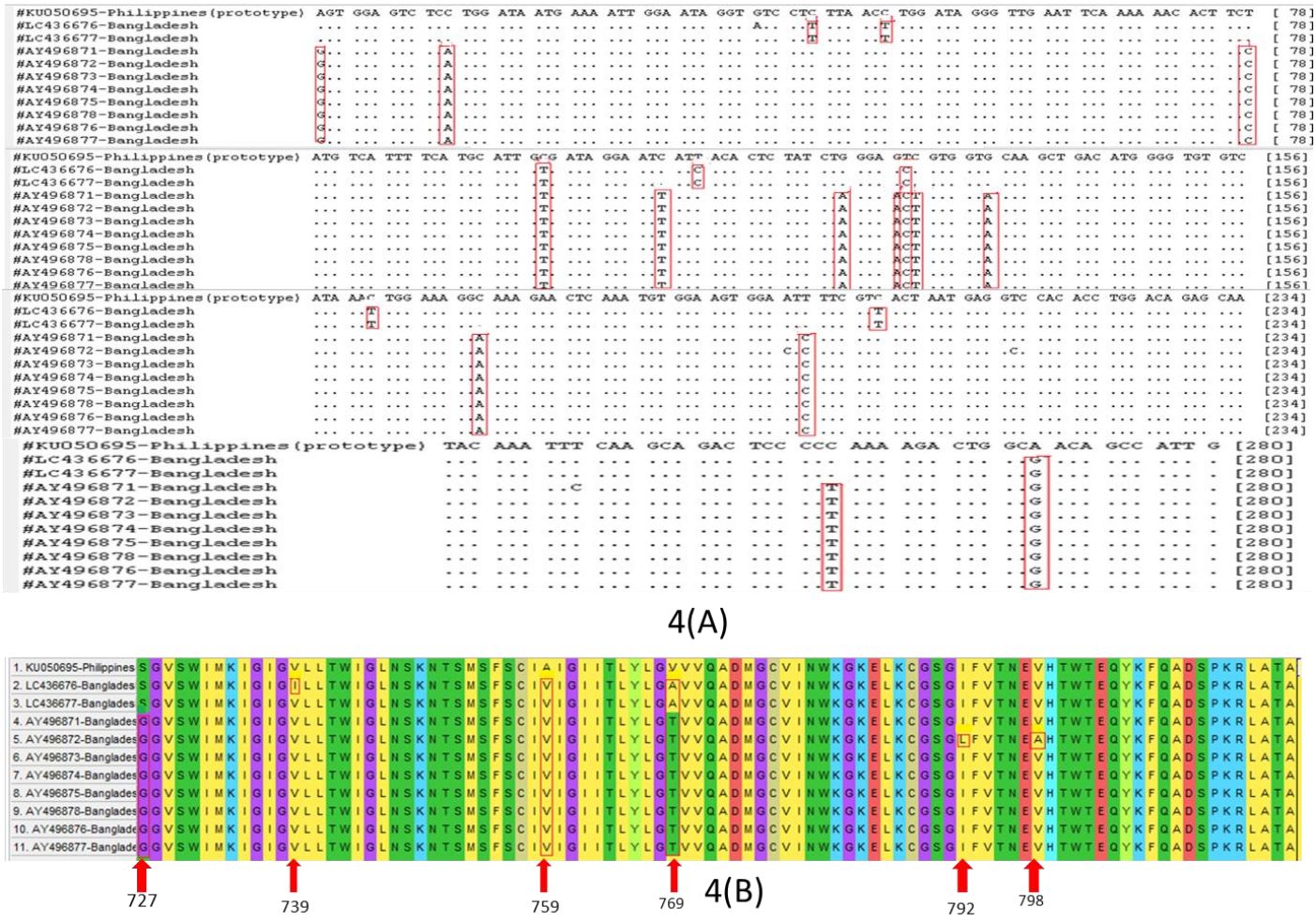

Fig. 4. (A). 280bp aligned E/NS1 nucleotide sequences of the most recent and old Bangladeshi DEN-3 isolates analyzed in this study. Upper most sequence is the DENV-3 prototype Philippines H87 strain. Here, 13 common nucleotide differences in 2 of the Bangladeshi isolates and 14 in the other 8 Bangladeshi isolates. Only 2 of the nucleotide changes are common in all the Bangladeshi isolates. Nucleotide substitutions from the prototype to the Bangladeshi sequences are marked with red blocks. This alignment was done in MEGA11. **(B).** Amino acid substitutions are shown here in red marked blocks in the amino acid alignment of the Bangladeshi DEN-3 E/NS1 isolates. This alignment was done in MEGA11.

This alignment was done in MEGA11. Here, the reference sequence is the DENV-3 prototype Philippines H87 strain. Of these 10 Bangladeshi DEN-3 sequences, 2 were obtained from the work of Suzuki *et al.* in Dhaka city on 2020 and 8 were obtained from the work of Islam *et al* in Dhaka city, 2002. Analysis of these 10 Bangladeshi DENV-3 E/NS1 sequences, the sequences from Suzuki *et al.*,2020 shows 13 common nucleotide changes and the sequences from Islam *et al.*, 2002 shows 14 common nucleotide changes compared to the prototype sequence. Analysis of theses sequences shows there is 93.55% to 99.64% sequence homology among the Bangladeshi sequences. Identical nucleotide percentages range from only 94.62% to 97.14% while comparing the Bangladeshi sequences with the prototype Philippines H87 strain **(Table 3).**

Table 3. Similarity percentages and divergence between DENV-3 Bangladeshi isolates and prototype reference strain H87.

Sequence accession no.	Similarity (%)	Divergence (%)
LC43676	96.79	3.21
LC436677	97.14	2.86
AY496871	94.98	5.02
AY496872	94.62	5.38

AY496873	95.34	4.66
AY496874	95.34	4.66
AY496875	95.34	4.66
AY496876	95.34	4.66
AY496877	95.34	4.66
AY496878	95.34	4.66

Individually, LC436678, LC436677, AY496871, AY496872, AY436873, AY436874, AY496875, AY496876, AY496877 and AY496878 show total base substitution number of 9, 8, 15, 16, 14, 14, 14, 14, 14 and 14 respectively. In case of amino acid substitution, two of the ten Bangladeshi isolates, LC436676 and LC436677, have only two common amino acid substitution alanine to valine at position 759 and valine to alanine at position 769 in the amino acid sequences. The other 8 sequences show 3 common amino acid substitutions, serine to glycine at position 727 of the junction, alanine to valine at position 759 of the junction and alanine to threonine at position 769 of the junction. But there is a random amino acid substitution from valine to isoleucine at position 739 of the junction in the LC436676 sequence and two random amino acid substitutions from isoleucine to leucine at position 792 and from valine to alanine at position 798 of the junction in the AY496872 sequence.

These 10 Bangladeshi DENV-3 E/NS1 nucleotide and amino acid sequences were then aligned and compared with 25 more E/NS1 nucleotide and amino acid sequences of DENV-3 from different geographical regions of the world where the H87 strain (KU050695) was used as the reference sequence. A total of 19 countries' sequence data were retrieved from GenBank and aligned in MEGA11 software.

Nucleotide substitution resulting in codon and amino acid change through non-synonymous mutations in the DENV-3 E/NS1 sequences obtained from different geographical regions of the world are given in **Table 4.**

Table 4. Mutational landscape identified in the DENV-3 E/NS1 sequences obtained from different geographical regions of the world.

Accession no.	**Country**	**No. of nucleotide substitution**	**Non-synonymous mutation**	**Synonymous mutation**	**Codon change in NS mutations**	**Amino acid change in NS mutations**
LC436676	Bangladesh	9	3	6	**G**TC-**A**TC G**C**G-G**T**G G**T**C-G**C**C	V-I(739) A-V(759) V-A(769)
LC436677	Bangladesh	8	2	6	G**C**G-G**T**G G**T**C-G**C**C	A-V(759) V-A(769)
AY496871	Bangladesh	15	5	10	**A**GT-**G**GT G**C**G-G**T**G **GTC-ACT**	S-G(727) A-V(759) V-T(769)
AY496872	Bangladesh	16	6	10	**A**GT-**G**GT G**C**G-G**T**G **GTC-ACT** **ATT-C**T**C** G**T**C-G**C**C	S-G(727) A-V(759) V-T(769) I-L(792) V-A(798)
AY496873	Bangladesh	14	5	9	**A**GT-**G**GT G**C**G-G**T**G **GTC-ACT**	S-G(727) A-V(759) V-T(769)
AY496874	Bangladesh	14	5	9	**A**GT-**G**GT G**C**G-G**T**G **GTC-ACT**	S-G(727) A-V(759) V-T(769)
AY496875	Bangladesh	14	5	9	**A**GT-**G**GT G**C**G-G**T**G **GTC-ACT**	S-G(727) A-V(759) V-T(769)
AY496876	Bangladesh	14	5	5	**A**GT-**G**GT G**C**G-G**T**G **GTC-ACT**	S-G(727) A-V(759) V-T(769)
AY496877	Bangladesh	14	5	9	**A**GT-**G**GT G**C**G-G**T**G **GTC-ACT**	S-G(727) A-V(759) V-T(769)

AY496878	Bangladesh	14	5	9	**A**GT-**G**GT G**C**G-G**T**G **GTC**-**ACT**	S-G(727) A-V(759) V-T(769)
FJ644564	India	21	4	17	**ATA**-**GTG** G**TC**-G**CT**	I-V(732) V-A(769)
AY770511	India	21	4	17	**ATA**-**GTG** G**TC**-G**CT**	I-V(732) V-A(769)
AY099336	Sri Lanka	17	4	13	**ATA**-**GTG** G**TC**-G**CT**	I-V(732) V-A(769)
AF317645	China	0	0	0	N/A	N/A
AF029794	Malaysia	12	1	11	G**T**C-G**C**C	V-A(769)
AY029795	Malaysia	11	1	10	G**T**C-G**C**C	V-A(769)
AY766104	Singapore	11	1	10	G**T**C-G**C**C	V-A(769)
KJ737430	Thailand	9	1	8	G**T**C-G**C**C	V-A(769)
AY676353	Thailand	9	1	8	G**T**C-G**C**C	V-A(769)
DQ675520	Indonesia	10	1	9	G**T**C-G**C**C	V-A(769)
OK469353	Thailand	10	1	9	G**T**C-G**C**C	V-A(769)
MW946984	Thailand	11	2	9	G**C**G-G**T**G G**T**C-G**C**C	A-V(759) V-A(769)
KY586801	Thailand	11	2	9	G**C**G-G**T**G G**T**C-G**C**C	A-V(759) V-A(769)
AY679147	Brazil	21	3	18	**A**TA-**G**TA G**TC**-G**CT**	I-V(732) V-A(769)
AY099337	Martinique	22	3	19	**A**TA-**G**TA G**TC**-G**CT**	I-V(732) V-A(769)
EU492458	Vietnam	14	3	11	G**C**G-G**T**G G**T**C-G**C**C G**T**C - G**C**C	A-V(759) V-A(769) V-A(778)
GQ868629	Cambodia	13	2	11	G**C**G-G**T**G G**T**C-G**C**C	A-V(759) V-A(769)
AY744685	French Polynesia	9	4	5	**A**CT-**G**CT G**C**G-G**T**G G**TC**-G**CT**	T-A(751) A-V(759) V-A(769)
FJ882575	Mozambique	18	4	14	**ATA**-**GTG** G**TC**-G**CT**	I-V(732) V-A(769)
GU131945	Colombia	20	3	17	**A**TA-**G**TA G**TC**-G**CT**	I-V(732) V-A(769)
HQ166034	Nicaragua	21	3	18	**A**TA-**G**TA G**TC**-G**CT**	I-V(732) V-A(769)
FJ373304	Venezuela	21	3	18	**A**TA-**G**TA G**TC**-G**CT**	I-V(732) V-A(769)
GQ868616	Saint Lucia	22	3	19	**A**TA-**G**TA G**TC**-G**CT**	I-V(732) V-A(769)
FJ547071	Puerto Rico	19	3	16	**A**TA-**G**TA G**TC**-G**CT**	I-V(732) V-A(769)

In this alignment dataset, only 4 common non-synonymous mutations have been found. At the first amino acid position there is a change from serine to glycine in 8 Bangladeshi sequences. This mutation was not found in any other countries' sequences used in this study. An amino acid substitution from isoleucine to valine at the 732 position of the junction was found to be present in 10 countries' sequences but not in any of the Bangladeshi isolates. The non-synonymous mutation causing amino acid substitution from alanine to valine in the 759 position was found in 5 countries and all of the Bangladeshi isolates have this mutation. 18 countries were found to have a common amino acid substitution at the 769 position of the junction from valine to alanine. However, 8 of Bangladeshi isolates were unique because in those sequences valine was substituted by threonine instead of alanine in that particular position of the junction. There were 5 more random amino acid changes found which are shown in **Supplementary Fig. 2(B).**

A neighborhood-joining phylogenetic tree of all the 35 sequences of DENV-3 E/NS1 junction has been constructed using the software MEGA11. Bootstrap values are also shown in the phylogenetic tree. The phylogenetic tree showed in **Fig. 5** generated 3 distinct clusters of the three main genotypic groups. Eight of the Bangladeshi isolates (AY496871, AY496872, AY496873, AY496874, AY496875, AY496876, AY496877 and AY496878) are clustered with the isolates of Thailand, Vietnam, Cambodia, Malaysia, Indonesia and Singapore in the Genotype II. Other two Bangladeshi isolates (LC436676 and LC436677) were grouped with Philippines H87

prototype, China and French Polynesia in Genotype I. A clear genotypic shift occurred in the recent Bangladeshi isolates from the previous ones. DENV-3 genotype shifted from Genotype II to Genotype I in Bangladesh in recent years.

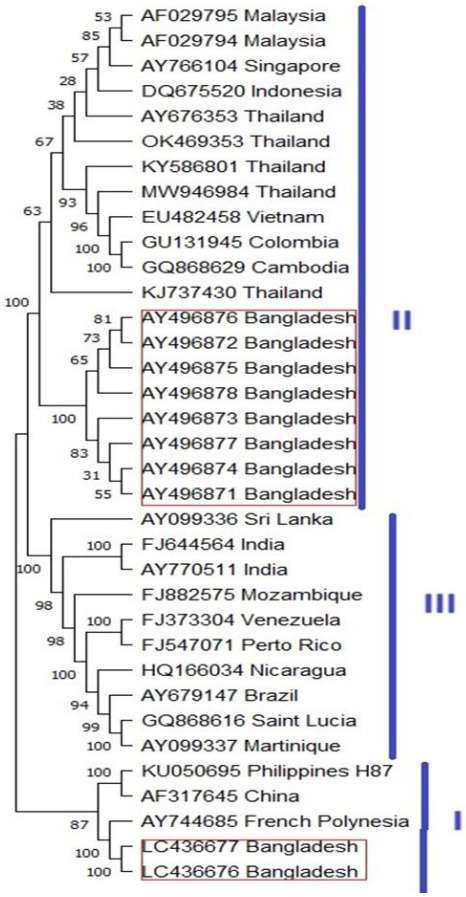

Fig. 5. Evolutionary connection between the ten Bangladeshi DENV-3 E/NS1 sequences with 25 other DENV-3 E/NS1 sequences is shown in the neighbor-joining tree above. Positions of Bangladeshi isolates in the clusters are indicated with red marked rectangles. Bootstrap values are displayed at the top of branches which joined the genotypic groups. There were a total of 280 positions in the final dataset. Evolutionary analyses were conducted in MEGA11.

Protein Structure Prediction of DEN-3 E Protein

After analyzing the sequences of E/NS1 junction of DENV-3 in Bangladesh, it was found that only one common amino acid change has occurred in the envelope protein region in all of the Bangladeshi isolates. This amino acid substitution from alanine to valine is situated at the E protein region and the NS1 protein region of the junction remains completely conserved in the Bangladeshi isolates. As alanine and valine both are hydrophobic amino acids, so the chances of protein function change are little but not impossible. Two more common amino acid substitutions were found only in the envelope region of the previous or 2002 Bangladeshi isolates which are serine to glycine and valine to threonine. In the first case, serine is a hydrophilic amino acid which was substituted by glycine, a hydrophobic one (Pommie *et al.*, 2004). When a hydrophilic residue is substituted by a hydrophobic one, it disrupts the stabilizing effect placed by the hydrophilic amino acid (Strub *et al.*, 2004). Therefore, this may cause functional changes in the protein. On the other hand, threonine is a neutral amino acid which comes in the position of valine and valine is hydrophobic. I-mutant 2.0 software shows destabilization due to the amino acid substitution which further proves the observation. Further research on the envelope protein is needed to determine any functional changes that may be occurred due to the reduction of stability.

Protein structure building and comparison of the DENV-3 E or envelope protein of Bangladeshi isolates with the reference Philippines H87 strain is shown in **Fig. 6 (i, ii, iii).**

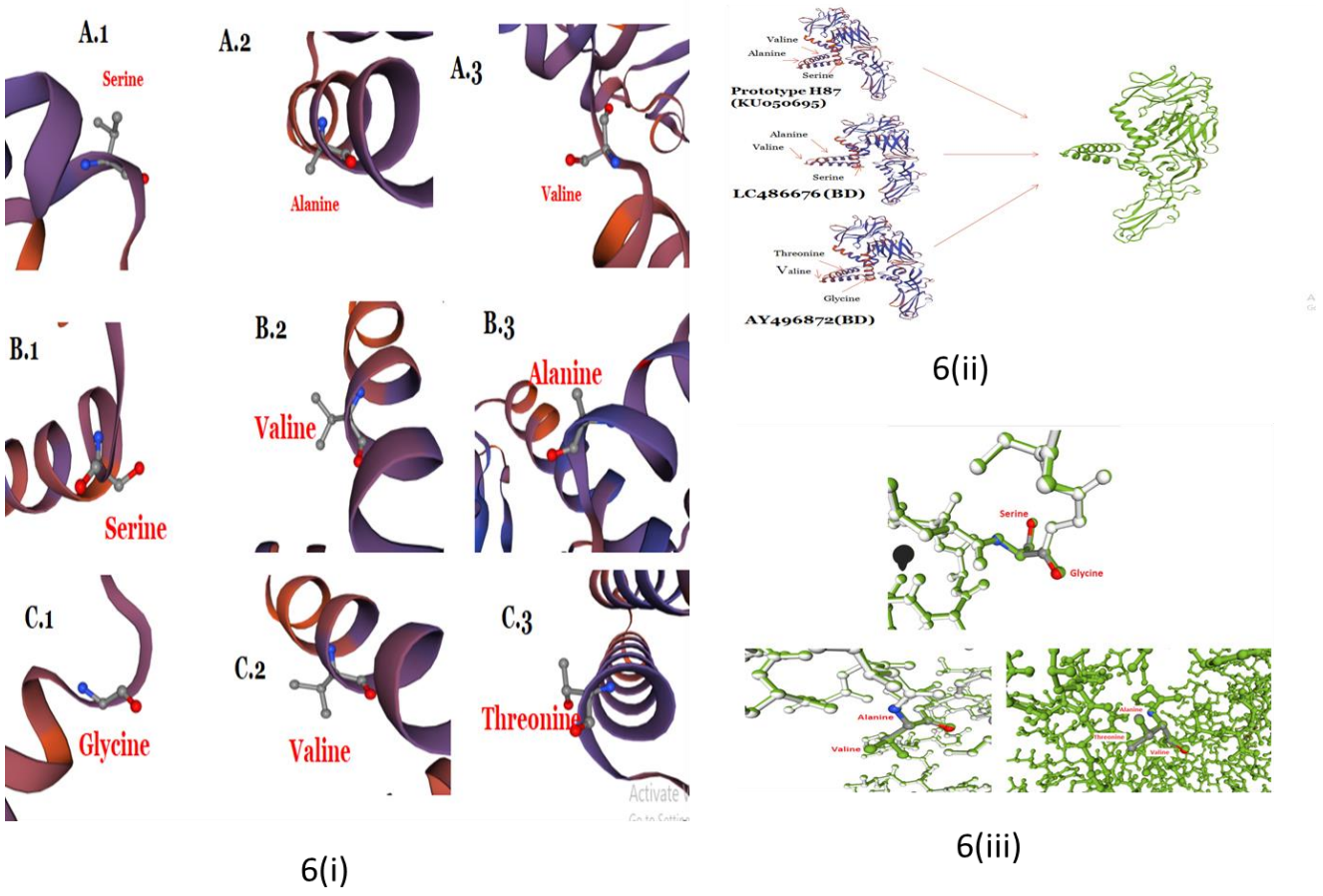

Fig. 6. (i). Protein structure prediction and comparison of previously found (AY496872-C.1, C.2, C.3) and current (LC436676-B.1, B.2, B.3) genotypes of DENV-3. Here A.1, A.2, A.3 structure is from prototype strain. **(ii).** Structure of E protein has been generated from all the prototype sequence, current and previous mutants from Bangladesh using bioinformatics programs. The merged protein of all three sequences shows no significant overall structural change. **(iii).** Amino acid changes in the reference isolate KU050695 and Bangladeshi isolates LC436676 and AY436872 is shown in this fig. This comparison is done in the software SWISS-Model.

This model building and comparison is done in the SWISS-Model software. This comparative protein structure analysis shows no significant structural change has occurred in the E protein of DENV-3 due to the amino acid substitutions. Detailed values are showed in **Table 5**.

Table 5. Envelope protein stability changes prediction for DENV-2 and DENV-3 isolates from Bangladesh by using I-mutant 2.0.

Strains	Position of amino acid in E protein	WT	Mutant	DDG (ΔΔG)	pH	T
DENV-2	742	I	V	-0.82	7	25
DENV-3(2002)	727	S	G	-1.58	7	25
	769	V	T	-1.92	7	25
DENV-3(2002 and 2020)	759	A	V	-0.91	7	25
DENV-3(2020)	769	V	A	-2.07	7	25

Note: WT: Amino acid in Wild-Type Protein, NEW: New Amino acid after Mutation, DDG: DG (New Protein)-DG (WildType) in Kcal/mol, DDG<0: Decrease Stability, DDG>0: Increase Stability, T: Temperature in Celsius degrees, pH: -log[H+]

Conclusion

Dengue virus is a burning issue in Bangladesh for a few years. The severity and symptoms of dengue fever are changing every year, as a result of that mortality rate is also getting higher. To understand the mutation trend of Dengue virus is inevitable for that reason. This study tried to analyze and compare the sequences of a mutation 'hot-spot', E/NS1 junction in the DENV-2 and DENV-3 serotypes from around the world.

Though DENV-2 did not show any genotypic shift despite E protein destabilization, DENV-3 isolates from Bangladesh exhibited a clear genotypic shift from past to present. E and NS1 protein destabilization also occurred due to amino acid

substitutions. If those changes in stability contributed to any functional alteration to the corresponding protein are yet to be known. Laboratory experimentation should be incorporated with the computational analysis to find out the answer to the question.

Acknowledgement

The authors are very much thankful to the chairman, department of Zoology, University of Dhaka for his assistance. This work was funded by the University Grant Commission and the University of Dhaka , Bangladesh (UGC -DU grant 2022-23). TA was supported by Bangladesh national science and technology fellowship.

Conflict of Interest

Authors have declared that no conflict of interest exists.

References

1. Akram, M. *et al.*(2015). 'Introduction and evolution of dengue virus type 2 in Pakistan: a phylogeographic analysis', *Virology Journal,* 12(1), pp. 1–11.
2. Barban, V. *et al.* (2012). 'Broad neutralization of wild-type dengue virus isolates following immunization in monkeys with a tetravalent dengue vaccine based on chimeric yellow fever 17D/dengue viruses', *Virology,* 429(2), pp. 91–98.
3. Bennett, S. N. *et al.* (2010). 'Epidemic dynamics revealed in dengue evolution', *Molecular Biology and Evolution,* 27(4), pp. 811–818.
4. Bromberg, Y. and Rost, B. (2009). 'Correlating protein function and stability through the analysis of single amino acid substitutions', BMC *bioinformatics,* 10, pp. 1-9.
5. Byk, L. A. *et al.* (2016). 'Dengue virus genome uncoating requires ubiquitination', *MBio,* 7(3), pp. e00804-16.
6. Capriotti, E., Fariselli, P. and Casadio, R., 2005. 'I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure', *Nucleic acids research,* 33(suppl_2), pp. W306-W310.
7. Chao, D.-Y. *et al.* (2005). 'Strategically examining the full-genome of dengue virus type 3 in clinical isolates reveals its mutation spectra', *Virology Journal,* 2(1), pp. 1–10.
8. Chu, J. J. H. *et al.* (2015). 'Antiviral activities of 15 dengue NS2B-NS3 protease inhibitors using a human cell-based viral quantification assay', *Antiviral Research,* 118, pp. 68–74.
9. Clyde, K., Kyle, J. L. and Harris, E. (2006). 'Recent advances in deciphering viral and host determinants of dengue virus replication and pathogenesis', *Journal of Virology,* 80(23), pp. 11418–11431.
10. Diamond, M. S. and Pierson, T. C. (2015). 'Molecular insight into dengue virus pathogenesis and its implications for disease control', *Cell,* 162(3), pp. 488–492.
11. Edussuriya, C., Deegalla, S. and Gawarammana, I., 2021. 'An accurate mathematical model predicting number of dengue cases in tropics', *PLoS neglected tropical diseases,* 15(11), p.e0009756.
12. Eldigail, M. H. *et al.*(2020). 'Association of genotype III of dengue virus serotype 3 with disease outbreak in Eastern Sudan, 2019', *Virology Journal,* 17(1), pp. 1–8.
13. Goncalvez, A. P. *et al.* (2002). 'Diversity and evolution of the envelope gene of dengue virus type 1', *Virology,* 303(1), pp. 110–119.
14. Guo, C. *et al.* (2017). 'Global epidemiology of dengue outbreaks in 1990–2015: a systematic review and meta-analysis', *Frontiers in Cellular and Infection Microbiology,* 7, pp. 317.
15. Harapan, H. *et al.* (2020). 'Dengue: a minireview', *Viruses,* 12(8), pp. 829.
16. Ahmed, R. *et al.*(2020). 'Isolation and Detection of Dengue Virus in Bangladesh', *Business, Social and Scientific Research*, pp. 107.
17. Islam, M. A. *et al.* (2006). 'Molecular characterization and clinical evaluation of dengue outbreak in 2002 in Bangladesh', *Japanese Journal of Infectious Diseases,* 59(2), pp. 85.
18. Ito, M. *et al.* (2004). 'Development and evaluation of fluorogenic TaqMan reverse transcriptase PCR assays for detection of dengue virus types 1 to 4', *Journal of Clinical Microbiology,* 42(12), pp. 5935–5937.
19. Kanakaratne, N. *et al.* (2009). 'Severe dengue epidemics in Sri Lanka, 2003–2006', *Emerging Infectious Diseases,* 15(2), pp. 192.
20. Khadijah, N. R. *et al.* (2016). 'Sequence analysis of E/NS1 gene junction of Dengue Type-1 viruses isolated in Klang Valley 2010 to 2012', *Tropical Biomedicine,* 33(2), pp. 348–358.
21. King, C.-C. *et al.* (2008). 'Comparative analysis of full genomic sequences among different genotypes of dengue virus type 3', *Virology Journal,* 5(1), pp. 1–13.
22. Lai, S. *et al.* (2018). 'Seasonal and interannual risks of dengue introduction from South-East Asia into China, 2005-2015', *PLoS Neglected Tropical Diseases,* 12(11), pp. e0006743.
23. Lewis, J.A. *et al.* (1993). 'Phylogenetic relation-ships of Dengue-2 viruses', *Virology,* 197, pp. 216-224
24. Mahy, B. W. J. (2008). 'Animal viruses: molecular biology', *Emerging Infectious Diseases,* 14(5), pp. 867.
25. Manandhar, K. Das *et al.* (2021). 'Whole Genome Sequencing of Dengue Virus Serotype 2 from Two Clinical Isolates and Serological Profile of Dengue in the 2015–2016 Nepal Outbreak', *The American Journal of Tropical Medicine and Hygiene,* 104(1), pp. 115.
26. Mangada, M. N. M. and Igarashi, A. (1998). 'Molecular andin VitroAnalysis of Eight Dengue Type 2 Viruses Isolated from Patients Exhibiting Different Disease Severities', *Virology,* 244(2), pp. 458–466.
27. Mutsuddy, P. *et al.* (2019). 'Dengue situation in Bangladesh: An epidemiological shift in terms of morbidity and mortality', *Canadian Journal of Infectious Diseases and Medical Microbiology*, *2019*.
28. Najri, N. I. *et al.* (2019). 'Genotypes of the dengue virus in patients with dengue infection from Sabah, Malaysia', *Journal of Physics: Conference Series,* 1358(1). doi: https://doi.org/10.1088/1742-6596/1358/1/012019
29. Ngwe Tun, M. M. *et al.* (2021). 'An Outbreak of Dengue Virus Serotype 2 Cosmopolitan Genotype in Nepal, 2017', *Viruses,* 13(8), pp. 1444.
30. Osman, O., Fong, M. Y. and Devi, S. (2008). 'Sequence analysis of E/NS1 gene junction of dengue virus type 2 isolated in Brunei', *Southeast Asian Journal of Tropical Medicine and Public Health,* 39(1), pp. 62.
31. Pancharoen, C. *et al.* (2002). 'Dengue infection: a global concern', *Journal of the Medical Association of Thailand= Chotmaihet Thangphaet,* 85, pp. 25-33.
32. Pommié, C. *et al.* (2004). 'IMGT standardized criteria for statistical analysis of immunoglobulin V-REGION amino

acid properties', *Journal of Molecular Recognition*, 17(1), pp. 17-32.
33. Powell, J. R. (2018). 'Mosquito-borne human viral diseases: why Aedes aegypti?', *The American Journal of Tropical Medicine and Hygiene*, 98(6), pp. 1563.
34. Rashed, A. (2021). 'Seroprevalence and clinical characteristics investigation of Dengue virus infection in rural Bangladesh during 2019 Dengue outbreak', Brac University.
35. Rico-Hesse, R. (1990). 'Molecular evolution and distribution of dengue viruses type 1 and 2 in nature', *Virology*, 174(2), pp. 479–493.
36. Rico-Hesse, R. *et al.* (1997). 'Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas', *Virology*, 230(2), pp. 244–251.
37. Sharma, S. *et al.* (2011). 'Comparative complete genome analysis of dengue virus type 3 circulating in India between 2003 and 2008', *Journal of General Virology*, 92(7), pp. 1595–1600.
38. Strub, C. *et al.* (2004). 'Mutation of exposed hydrophobic amino acids to arginine to increase protein stability', *BMC biochemistry*, 5, pp. 1-6.
39. Suzuki, K. *et al.* (2020). 'Genotype replacement of dengue virus type 3 and clade replacement of dengue virus type 2 genotype Cosmopolitan in Dhaka, Bangladesh in 2017', *Infection, Genetics and Evolution*, 75, pp. 103977.
40. Tabassum, T. and Taylor-Robinson, A. W. (2019). 'Dengue Serotypes in Bangladesh: Whole Genome Sequencing and Comparative Genomics Facilitates Pathogenesis and Epidemiology Studies and Informs Improved Disease Control', *Microbiology & Infectious Diseases*, 3(1), pp. 0–5. doi: https://doi.org/10.33425/2639-9458.1054
41. Tamura, K., Nei, M. and Kumar, S., 2004. 'Prospects for inferring very large phylogenies by using the neighbor-joining method', *Proceedings of the National Academy of Sciences*, 101(30), pp.11030-11035.
42. Teng, S., Srivastava, A.K. and Wang, L. (2010). 'Sequence feature-based prediction of protein stability changes upon amino acid substitutions', *BMC genomics*, 11, pp. 1-8.
43. Titir, S. R. *et al.* (2021). 'Nationwide Distribution of Dengue Virus Type 3 (DENV-3) Genotype I and Emergence of DENV-3 Genotype III during the 2019 Outbreak in Bangladesh', *Tropical Medicine and Infectious Disease*, 6(2), pp. 58.
44. Tuiskunen Bäck, A. and Lundkvist, Å. (2013). 'Dengue viruses–an overview', *Infection Ecology & Epidemiology*, 3(1), pp. 19839.
45. Uno, N. and Ross, T. M. (2018). 'Dengue virus and the host innate immune response', *Emerging Microbes and Infections*, 7(1). doi: https://doi.org/10.1038/s41426-018-0168-0
46. White, S. H. (1992). 'Amino acid preferences of small proteins: Implications for protein stability and evolution', *Journal of molecular biology*, 227(4), pp. 991-995.