Iranian Journal of Microbiology

Volume 16 Number 1 (February 2024) 104-113

ORIGINAL ARTICLE

A comparative genomic approach to decipher the mutations associated with Nipah viral human isolates from southeast Asia

Norine Norbert Dsouza[1,2], Selvaa Kumar Chellasamy[1*]

[1]Department of Bioinformatics, School of Biotechnology and Bioinformatics, Sector 15, CBD Belapur, Navi Mumbai, Maharashtra, India

[2]Department of Biotechnology, St. Xavier's College, Mumbai, Maharashtra, India

Received: September 2022**, Accepted:** January 2024

ABSTRACT

Background and Objectives: Multiple outbreaks over two decades and a high mortality rate have emphasized the Nipah virus (NiV) as a priority research area. The study focuses on identifying the mutational landscape in sequences from NiV human isolates from different geographical regions.

Materials and Methods: Thirty-seven NiV genomes of human samples from Malaysia, Bangladesh, and India were subjected to phylogeny and metagenomic analysis to decipher the genome variability using MEGA11 software and the meta-CATS web server. Using the Single-Likelihood Ancestor Counting method, the synonymous and nonsynonymous mutations among NiV genes were identified. Further, the nonsynonymous variations were used to identify mutations in all the NiV proteins.

Results: The NiV isolates were categorized into NiV-M, NiV-B, and NiV-I clades based on phylogenetic analysis. Metagenomic analysis revealed 1636 variations in the noncoding and coding regions of the genomes of the three clades of NiV. Further analysis of nonsynonymous mutations showed the phosphoprotein to be highly mutating, whereas the matrix protein was stable.

Conclusion: Deciphering the mutation pattern using a comparative genomics approach for human isolates provided valuable insight into the stability of NiV proteins which can be further used for understanding variations in host-pathogen interaction and developing effective therapeutic measures.

Keywords: Nipah virus; Phylogeny; Synonymous mutations; Mutations; Host-pathogen interaction

INTRODUCTION

The Nipah virus (NiV) is one of the rapidly mutating negative-sense RNA viruses that cause respiratory distress syndrome, encephalitis, vasculitis, and other fatal symptoms and have a high mortality rate (1). The first NiV outbreak was reported in 1998 in Malaysia, followed by subsequent sporadic occurrences in Bangladesh and India, and so far more than twelve outbreaks have been reported in Southeast Asia (2-4). Moreover, the sero-surveillance investigations have shown the presence of antibodies targeting NiV in bat specimens collected from various regions including Cambodia, Thailand, Madagascar, and Ghana. This compelling evidence suggests the active circulation of NiV within bat populations (5). The NiV outbreaks in Malaysia have implicated bats as the natural hosts and pigs as the intermediate

*Corresponding author: Selvaa Kumar Chellasamy, Ph.D, Department of Bioinformatics, School of Biotechnology and Bioinformatics, Sector 15, CBD Belapur, Navi Mumbai, Maharashtra, India. Tel: +22-27563600

Email: selvaakumar.c@dypatil.edu

hosts, playing a crucial role in the transmission of the virus to humans. Nevertheless, direct viral transmission from natural reservoirs to human hosts was documented during the outbreaks in Bangladesh and India (6-8). Individuals infected with the NiV during the outbreaks in Malaysia and Bangladesh displayed distinct differences in their clinical manifestations. In a comparative analysis between Bangladeshi and Malaysian individuals infected with the NiV, it was observed that the former group exhibited a reduced incubation period and severe neurological symptoms in addition to severe respiratory distress, as reported by previous studies (5). Among the NiV-affected population, there was a noticeable difference in the fatality rate: Malaysia showed a mortality rate of 40%, while Bangladesh and India showed mortality rates ranging from 70% to 90% (3, 9). The observed variations in the clinical symptoms exhibited by the infected individuals have been linked to the genomic heterogeneity identified between the Malaysian and Bangladesh genome sequences acquired from distinct hosts via comparative genome analysis investigations. The genomic variations observed among the sequences of the isolates under investigation enabled their classification into two distinct groups, namely NiV-Bangladesh (NiV-B) and NiV-Malaysia (NiV-M) (9, 10). Moreover, the comparative analysis of Indian NiV isolates showed a nucleotide identity of 97% and 91% with the NiV-B and NiV-M clades, respectively (11). Similarly, the phylogenetic analysis of the N gene (Nucleocapsid) and the G gene (Attachment Glycoprotein) showed the divergence of the sequences into NiV-M and NiV-B clades, confirming the evolution of the specific genes of the NiV genome (12).

The reported outbreaks from 1998 onwards indicated the dominance of the NiV-B clade compared with NiV-M, with the prevalence of the former in humans (13). This demands a thorough study to elucidate the underlying factors contributing to the observed variations in symptoms and mortality rates among individuals infected with NiV originating from disparate geographical locations. The genomes and variations in the genes of NiV-isolated human samples have not been studied so far. Thus, the present study aims to analyze the complete genome sequence to understand the divergence pattern in NiV isolates at different geographical locations, followed by variation pattern analysis at the non-coding regions (NCR) and coding DNA sequence (CDS) regions of the genes of NiV. Furthermore, the study also intends to analyze variation patterns observed in the six genes to identify amino acid mutations and propose mutational models among the nucleoprotein (N), phosphoprotein (P), matrix protein (M), fusion protein (F), attachment glycoprotein (G), and RNA-directed RNA polymerase (L) of NiV.

MATERIALS AND METHODS

Sequence retrieval and phylogenetic analysis. A total of thirty-seven complete genome sequences of NiV human isolates from different geographical areas of Southeast Asia were downloaded from the GenBank Database in FASTA format (Table 1). The Malaysian NiV isolate (Accession No. NC_0027280) was used as a reference. The sequence variation of aligned genome sequences was calculated using MEGA11 software, and phylogenetic analysis was performed by the neighbor-joining (NJ) method using a maximum composite likelihood model and branch support testing using 1000 bootstrap iterations (14). The software uses the homology between the sequences to decipher the evolutionary interrelationship among the sequences.

Analysis of genome variability in NCR and CDS regions. The phylogenetic analysis identified three geographically distinct groups: Malaysian (NiV-M), Bangladesh (NiV-B), and Indian (NiV-I). These groups of genome sequences were compared using a metadata-driven comparative analysis tool (meta-CATS) to identify the regions of significant sequence differences (15). Using the graphical and Excel sheet output the regions of sequence variations in the NCR and CDR regions were identified and recorded for all the six NiV genes. To obtain clarity in the variation pattern, the NCR region of NiV genes was marked as 3' leader and 5' trailer regions using the annotated information reference sequence NC_002728. Further, the CDS region was subjected to the Single-Likelihood ancestor counting (SLAC) method from the online Datamonkey Server for the identification of synonymous (S) and nonsynonymous (N) mutations. Thus, the observed variations within the CDS region were further analyzed to record the synonymous (S) and nonsynonymous (N) mutations. The tool employs maximum-likelihood (ML) and counting approaches to identify the regions of muta-

Table 1. Complete genome sequences of NiV human isolates retrieved from the GenBank database

S No	Accession no	Year	Country	S No	Accession no	Year	Country
1	NC_002728.1	1999	Malaysia (2)	20	MK673579.1	2012	Bangladesh
2	AY029767.1			21	MK673582.1	2012	
3	MK673565.1	2004	Bangladesh (29)	22	MK673583.1	2013	
4	MK673567.1	2004		23	MK673590.1	2014	
5	MK673566.1	2004		24	MK673589.1	2014	
6	MK673564.1	2004		25	MK673591.1	2014	
7	AY988601.1	2004		26	MK673584.1	2015	
8	MK673568.1	2008		27	K673586.1	2015	
9	JN808863.1	2008		28	MK673585.1	2015	
10	JN808857.1	2008		29	MK673592.1	2015	
11	MK673577.1	2011		30	MK673587.1	2015	
12	MK673574.1	2011		31	FJ513078.1	2007	India (07)
13	MK673573.1	2011		32	MH523642.1	2018	
14	MK673571.1	2011		33	MH523640.1	2018	
15	MK673578.1	2011		34	MH396625.1	2018	
16	MK673576.1	2011		35	MK673564.1	2018	
17	MK673575.1	2011		36	MK336155.1	2018	
18	MK673570.1	2011		37	MK336156.1	2018	
19	MK673581.1	2012					

tion along the length of the codons of the genes (16). The analysis results obtained from Excel sheet data were used for calculating the total number of synonymous (S) and nonsynonymous (N) mutations and percentage mutations for each of the NiV genes.

Mapping nonsynonymous changes in NiV proteins. The recorded nonsynonymous mutations in the CDS region of six genes were analyzed to study the changes in amino acid sequences among all the six proteins of NiV. The interactive phylo alignment results obtained from SLAC analysis were used to record the amino acid changes in proteins of NiV. The SLAC method uses the protein sequences deposited in the UniProtKB database for the identification of changes in amino acids in NiV proteins. The protein sequences Q9IH62 (Attachment Glycoprotein G), Q997F0 (RNA-directed RNA polymerase L), Q9IH63 (fusion protein F), Q9IK90 (matrix protein M), Q9IK91(phosphoprotein P), and Q9IK92 (nucleoprotein N) from Malaysian origin were used as references for identification of protein mutations among Bangladesh and Indian isolates. Only the amino acid changes observed in all the Bangladesh and Indian isolates in comparison with Malaysian sequences were selected for proposing the mutational models.

RESULTS

Clustering of NiV human isolates in phylogenetic analysis. The phylogenetic analysis of the genome sequences of NiV human isolates showed the divergence into three distinct clusters: NiV-Malaysia (NiV-M), NiV-Bangladesh (NiV-B), and NiV-India (NiV-I). The phylogenetic analysis of human NiV isolates was performed using thirty-seven sequences, and the results of the analysis are shown in Fig. 1.

Analysis of genome variability in NCR and CDS regions. The meta-CATS analysis of three geographically distinct clusters of human NiV genomes showed 1636 positions of differences between the three groups. The gene locations and the number of variations associated with NCR 3' leader and 5' trailer ends and the CDS region of all the NiV genes are shown in Table 2.

With 159 nucleotide changes, the Fusion protein and Attachment glycoprotein genes showed the highest number of variations in the NCR region, while the RNA-directed RNA polymerase gene showed the least number (42) of nucleotide variations. From Table 2, it is evident that the Phosphoprotein (P) and RNA-directed RNA polymerase (L) in the CDS region displayed the highest number of nucleotide variations, with 192 and 467, respectively, within the three

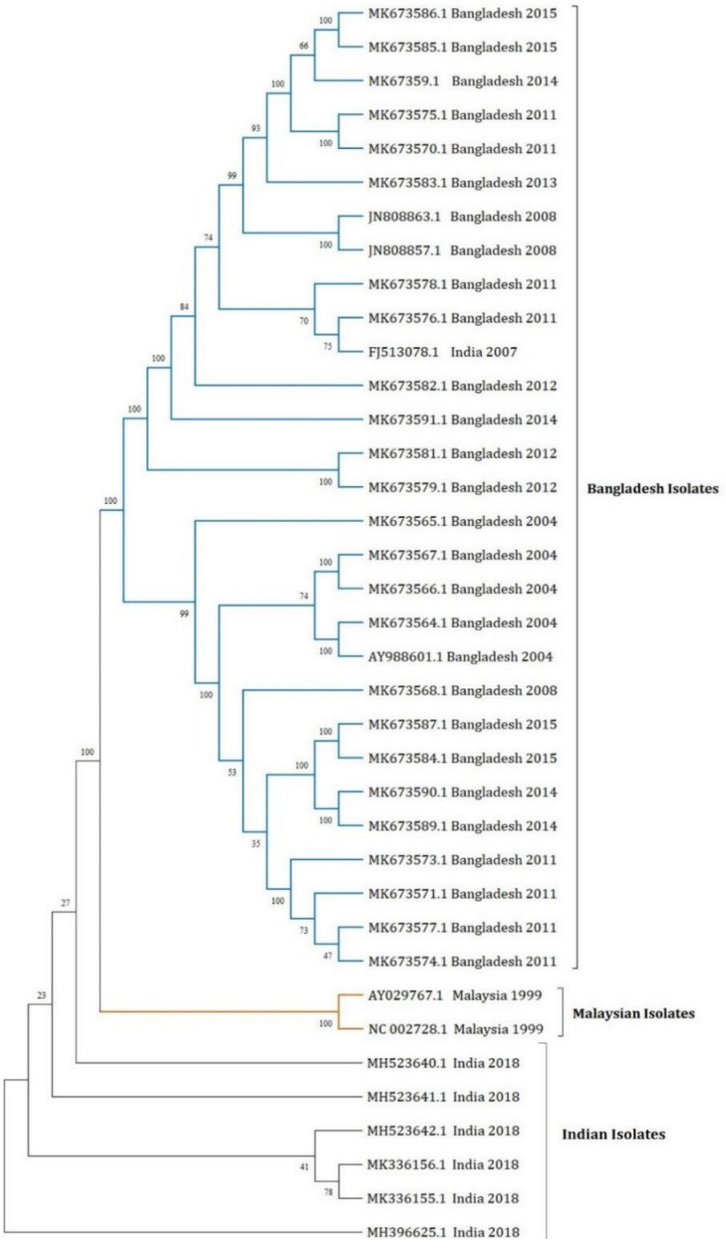

Fig. 1. Phylogenetic analysis of NiV genome sequences from human samples analyzed by MEGA11 software

clusters. The CDS region of the Matrix protein gene showed 77 variations in the CDS region.

Further, the 3' and 5' NCR regions (positions represented in Table 2) of all six genes were analyzed. The 5' NCR showed a greater number of variations for all the six NiV genes than the 3' NCR. Interestingly the 3' NCR of nucleoprotein genes showed single nucleotide variations indicating sequence conservation amongst the others. The total number of nucleotide variations observed in NCR and CDS regions of all the NiV genes are summarised in Fig. 2.

Further, the CDS region was analyzed for synonymous (S) and nonsynonymous (N) variation patterns in the site-specific codon of each gene. The results obtained for the synonymous variations for all the six NiV genes are shown in Fig. 3. All the NiV genes showed a uniform pattern of changes in codon leading to synonymous variations.

Likewise, the study of the nonsynonymous (N) variations associated with the NiV genes also revealed a wide range of mutation rates. The phosphoprotein gene showed a large number of variations, followed

Table 2. Genome-wide variations of NiV genome from different geographical locations.

NiV genome regions					Mutational changes				
Gene	Location	Length	CDS Location	NCR Location	NCR mutations	NCR	CDS	Total mutation	% Changes
N gene	56-2297	2242	113-1711	3':56-112	1	58	101	159	7.1
				5':1712-2297	57				
P gene	2301-5004	2704	2406-4535	3':2301-2405	12	70	192	262	9.7
				5':4536-5004	58				
M gene	5008 - 6366	1359	5108 -6166	3':5008-5107	17	56	77	133	9.8
				5':6167- 6366	39				
F gene	6370 -8706	2337	6654-8294	3':6370-6653	66	159	112	271	11.6
				5':8295-8706	93				
G gene	8710 -11255	2546	8943-10751	3':8710- 8942	57	159	143	301	11.8
				5':10752-11255	102				
L gene	11259-18213	6955	11412-18146	3':11259-11411	28	42	467	509	7.3
				5':18147-18213	14				

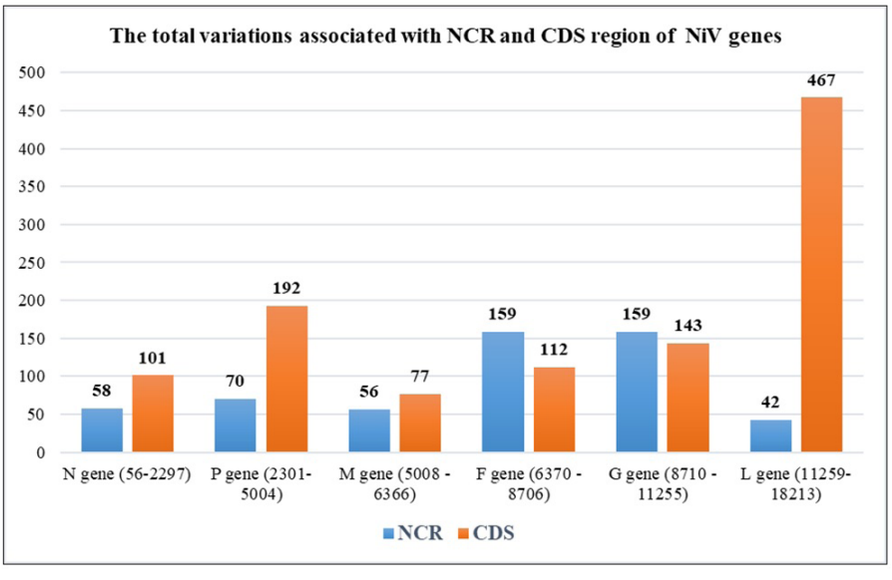

Fig. 2. Nucleotide variations associated with the NCR and CDS regions of NiV analyzed by meta-CATS tool

by attachment glycoprotein. The matrix protein gene showed the minimum number of mutations, indicating its evolutionary conservation over time. The pattern and position of nonsynonymous variations associated with the codons of NiV genes are represented in Fig. 4.

Mapping nonsynonymous changes in NiV. The nonsynonymous changes observed in the six genes of NiV were further mapped onto protein sequences to identify the changes in the amino acid composition of the nucleoprotein, phosphoprotein, matrix protein, fusion protein, attachment glycoprotein, and RNA-directed RNA polymerase of the NiV-B and NiV-I isolates. To avoid uncertainty about the model proposition, only the mutations seen in all of the Bangladesh and Indian isolates were used to map the mutations on proteins. This is in contrast to mapping the mutations on single isolates. The total number of nonsynonymous mutations obtained for all the NiV proteins in the NiV-B and NiV-I clusters is shown in Fig. 5. The present analysis revealed a notable mutational propensity of the phosphoprotein in both NiV-B and NiV-I strains. Specifically, a total of 56 mutations were identified in the phosphoprotein of NiV-B, while NiV-I exhibited 66 mutations. However, with only 4 common mutations the matrix protein showed the minimum

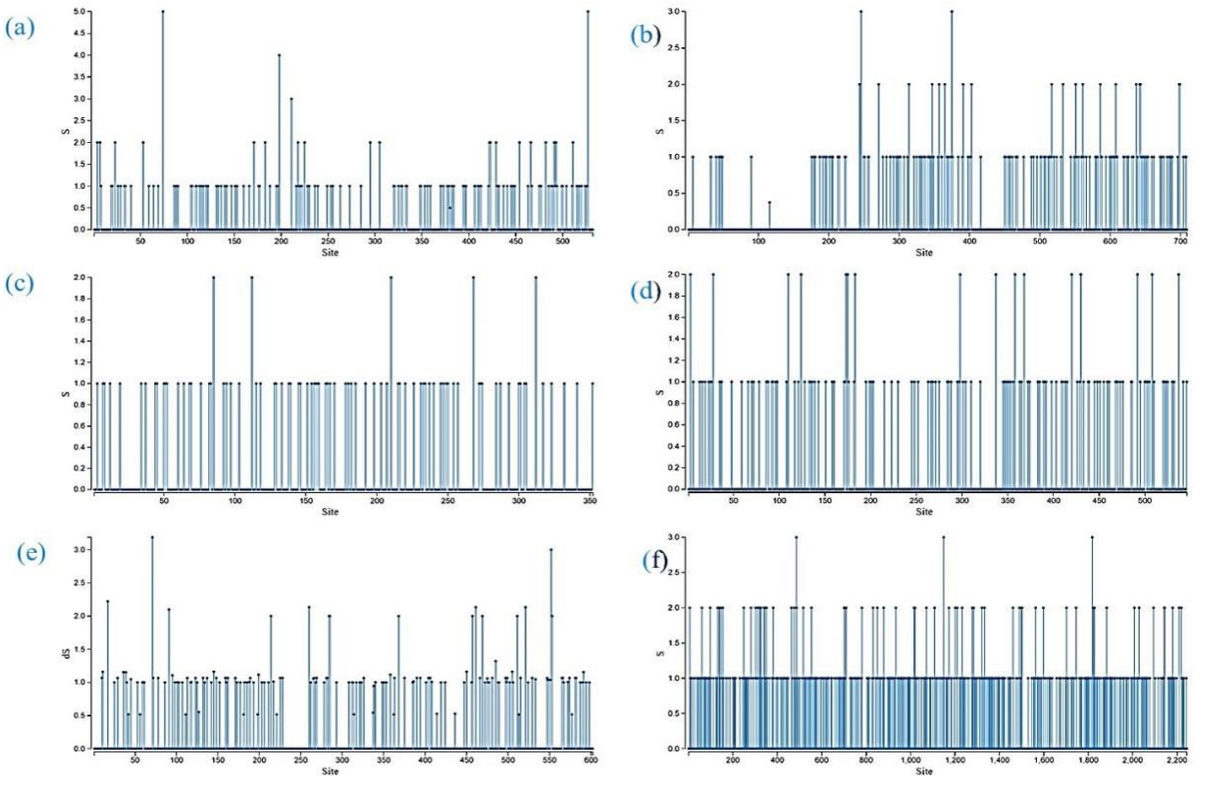

Fig. 3. The synonymous mutation pattern obtained for NiV genes analysed by SLAC analysis. (a) Nucleoprotein gene (b) Phosphoprotein gene (c) Matrix protein gene (d) Fusion Protein Gene (e) Attachment glycoprotein gene (d) RNA-directed RNA polymerase gene

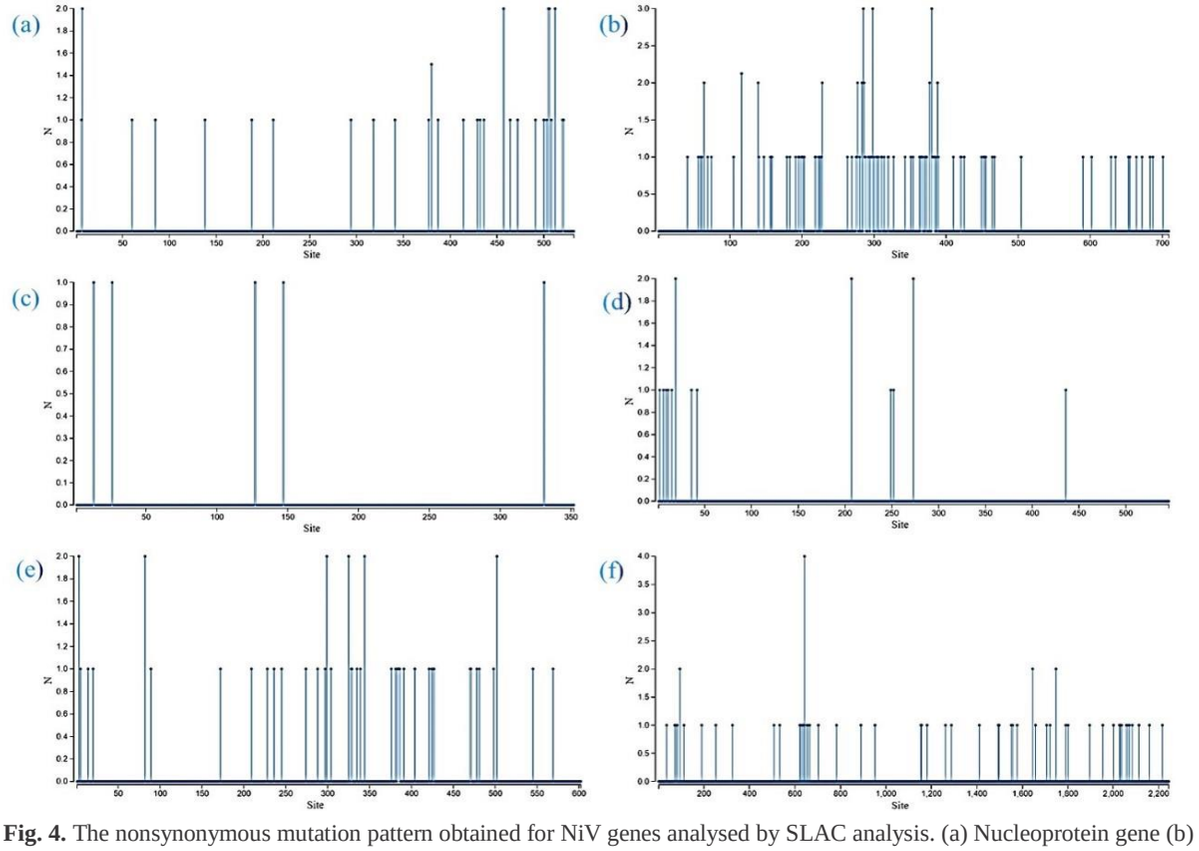

Fig. 4. The nonsynonymous mutation pattern obtained for NiV genes analysed by SLAC analysis. (a) Nucleoprotein gene (b) Phosphoprotein gene (c) Matrix protein gene (d) Fusion Protein Gene (e) Attachment glycoprotein gene (d) RNA-directed RNA polymerase gene

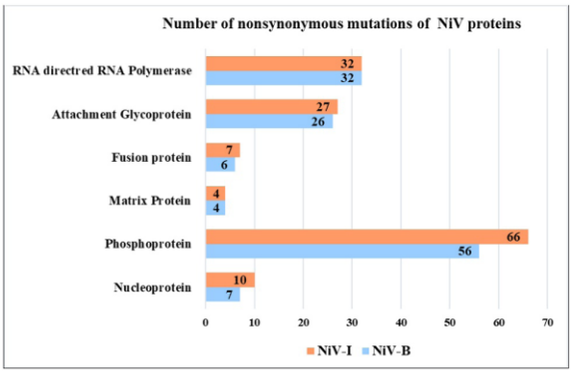

Fig. 5. Distribution of nonsynonymous mutations across NiV proteins for NiV-B and NiV-I isolates.

number of mutations in both groups.

Based on the CDS analysis, the mutation rate is found to be slower for nucleoprotein, matrix protein, and fusion protein compared with phosphoprotein, attachment glycoprotein, and RNA-directed RNA polymerase proteins in NiV-B and NiV-I models. The matrix protein and fusion protein displayed a few mutations, which were found to be conserved between the two models, except the I15L mutation in the NiV-I model. The present findings confirm that both matrix and fusion proteins exhibit the highest degree of stability among the entire proteome of NiV. In the comparative analysis, the investigation of nucleoprotein revealed the presence of six identical mutations (D387N, I429V, G432E, T506D, G508R, and A521T) in both models. However, the NiV-I displayed an additional four mutations (N457D, L472S, S503N, and P520S). The Attachment glycoprotein showed twenty-six mutations in NiV-B and twenty-seven mutations in NiV-I models. Comparative analysis between the models showed two unique mutations in NiV-B (K386E, and T498K) and three in NiV-I (N288S, I384V, and V427I). The RNA-directed RNA polymerase showed thirty-two mutations in both models, with three mutations (R1262K, N2037D, and D2064E) being unique to NiV-B and similarly three mutations (I94T, N642Y, and D2064E) unique to NiV-I protein. Among all the NiV proteins, the phosphoprotein was found to be maximally mutating with 55 and 66 amino acid mutations in NiV-B and NiV-I, respectively. The NiV-B showed 2 unique mutations (S367G, and D388N) wherein the NiV-I showed 12 unique mutations (E219G, Q225E, R285H, I287L, Q297K, P311L, T386N, P421L, S425N, Q449R, A629T, and G683D). The common mutations and unique mutations observed in the NiV isolates from NiV-B and NiV-I are shown in Table 3.

DISCUSSION

This study establishes a connection between the genomic and gene variations and the changes in NiV protein sequences. The formation of the NiV-M and NiV-B clades was revealed by phylogenetic analysis of NiV genomes, which was consistent with previous research published by Rahman et al. (2010) (17). The heterogeneity between the isolates from Bangladesh and India is further supported by the divergence of NiV-I and the emergence of a distinct clade from NiV-B. Even though the human NiV-B isolates selected for the investigation varied in time by over a decade, no significant year-wise separation of sequences into discrete subclades was observed, indicating that the human NiV-B isolates exhibited significantly fewer variations. The evolutionary relationship and pattern of divergence observed among the human NiV isolates from various geographic locations highlights the necessity of conducting a comprehensive study on genomic diversity and identifying specific mutations within the genomic sequences. The metagenomic sequence analysis among the NiV-M, NiV-B, and NiV-I provided an in-depth assessment of the variation at the NCR and CDS regions. The review of the literature highlights the significance of the viral genome's NCR regions in controlling viral replication, host immune evasion, viral persistence, and targeting cellular pathways involved in infection regulation. The literature survey highlights the importance of NCR regions of the viral genome in modulating viral replication, host immune evasion, and viral persistence and targets cellular pathways associated with regulating infection (18, 19). In an investigation utilising the NiV viral genome, it was observed that the 3′ non-coding regions (NCRs) of the RNA-directed RNA polymerase gene exert a significant influence on the establishment of a transcriptional gradient (20). In the current study, the meta-CATS analysis using NiV human isolates showed a higher number of variations in the NCR region of the fusion and attachment glycoprotein genes. The variation at the 5' trailer end of all the NiV genes was found to be higher than the 3' leader end, except for the RNA-directed RNA polymerase gene. The 3' leader end of the nucleoprotein gene was found to be highly stable among the isolates, and this finding further validates the use of the N gene 3' leader end in the Reverse Transcription Loop-Mediated Isothermal Amplification (RT-LAMP) assay

Table 3. List of nonsynonymous mutations associated with NiV-B and NiV-I along with common and unique mutations of each group

Proteins	Common Mutations	Unique Mutations	
		NiV-B	NiV-I
Nucleoprotein Q9IK92 (532)	D387N, I429V, G432E, T506D, G508R, A521T	R505K	N457D, L472S, S503N, P520S,
Phosphoprotein Q9IK91 (709)	Q41R, D69G, N74S, Y139H, T140S, N147D, M156V, N179D, T183A, V191I, L195P, R196K, D200V, K218R, G223D, S227N, D269E, V275A, G277R, N280I, I283V, T286I, I292T, N295S, A298I, D300G, S303P, T304A, V306A, E319K, P320S, R343Q, L351F, C354S, P363L, H365Y, W366R, R370A, I372R, V380T, N381S, G382D, R389K, A410E, A452V, S453P, A455V, V464A, V467A, N590S, I602V, E635G, I664V, K687R	S367G, D388N	E219G, Q225E, R285H, I287L, Q297K, P311L, T386N, P421L, S425N, Q449R, A629T, G683D,
Matrix Protein Q9IK90 (352)	H26N, I127V, S147G, I331V	--	-
Fusion protein Q9IH63 (546)	V2A, D6N, C9Y, C11S, V42I, S273G	-	I15L
Attachment Glycoprotein Q9IH62 (602)	A3T, N5S, T14A, I20N, V82M, G89S, R172K, R236K, P274S, T299V, S325N, G328E, G329S, L335F, S339N, R344M, T385A, G421E, P424S, V426I, L470Q, N481D, I502V, I545V	T498K K386E	N288S, I384V, V427I
RNA-directed RNA Polymerase Q997F0 (2244)	R36K, D71N, I77V, V252I, E533D, R621K, Y625C, S640P, H658Y, R703K, K783E, I890V, L1154I, K1157R, R1181K, V1494A, S1551A, V1577I, S1645Y, S1658N, V1707M, A1791S, R1801K, T1896A, V2027I, K2031R, D2064E, Q2071H, C2159R, N2216S	R1262K, N2037D	I94T, N642Y

for NiV detection (21). Hence, the findings associated with the NCR variability observed in human NiV isolates hold significant implications for elucidating the intricate dynamics underlying viral-host gene interactions. The observed outcomes can be correlated with the distinct clinical manifestations observed in human viral isolates of NiV-M and NiV-B (9).

The mutational landscape obtained for NiV-B and NiV-I in comparison with NiV-M shows a distinct pattern of mutations for all the proteins of NiV. This analysis deciphered the highly mutating proteins and stable proteins in NiV human isolates from the geographical locations of the outbreak. The present analysis constitutes a crucial component of current viral investigations, as changes in codons leading to nonsynonymous mutations within CDS regions exert a direct influence on protein structure and functions. Additionally, research on SARS-CoV-2 has shown that changes in the CDS regions cause changes in the secondary structure and solvent accessibility of proteins (22). Moreover, the specific nonsynonymous mutations of the Nsp3 protein of SARS-CoV-2 have been shown to interfere with epitope recognition altering the pathogenicity (23). The highly mutating phosphoprotein of NiV as reported in this study is known to interact with the host cellular proteins shutting the host immune system and also interacts with the RNA-directed RNA polymerase (L) along with nucleocapsid protein of the self-genome regulating the transcription and replication process (24). Phosphoprotein with a dual role associated with the high number of mutations could provide a platform for evaluating the impact of these mutations at the experimental level for studying host-virus interaction patterns leading to changes in pathogenicity. In an earlier comparative study of NiV isolates from bats and humans from Bangladesh, regions reported 143 amino acid changes across attachment glycoprotein, matrix protein, fusion glycorptoein, phosphoprotein, and nucleoprotein (25), wherein our study compared the NiV human isolates between NiV-B and NiV-M showed 131 mutations across all the six NiV proteins with phosphoprotein showing the highest number of mutations.

CONCLUSION

In this investigation, a comparative genomic analysis was undertaken to examine the genomes of NiV isolates derived from various geographical regions affected by outbreaks. The primary objective was to

assess the genetic diversity and identify any discernible patterns in the nucleotide composition. As a result, the genomes were observed to cluster into three distinct clades, indicating significant variations in their nucleotide sequences. This study demonstrates the importance of comparing NiV genes and proteins from different geographical groups. It provides insights into the stability and variability of the gene and protein of NiV, validating the use of the Nucleoprotein gene as a diagnostic tool for virus identification through PCR analysis. In addition, the research findings also suggest the use of the matrix protein in developing broad-spectrum pharmaceutical interventions for the NiV outbreak. The mutational pattern deduced for phosphoprotein through this study provides a preliminary reason for the observed differential clinical symptoms and mortality rate. The smaller number of sequence data from NiV isolates from the Indian region forms a limitation to conclude the mutational landscape for Indian isolates. Further comparative experimental studies at the protein level are required to decipher the exact role of these proteins in host-pathogen interactions.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support provided by the School of Biotechnology and Bioinformatics DY Patil Deemed to be University, Sector 15, CBD Belapur, Navi Mumbai, and St. Xavier's College (Autonomous)-DBT-Interdisciplinary Life Science Programme for Advanced Research and Education Builder Grant.

REFERENCES

1. Epstein JH, Anthony SJ, Islam A, Kilpatrick AM, Ali Khan S, Balkey MD, et al. Nipah virus dynamics in bats and implications for spillover to humans. *Proc Natl Acad Sci U S A* 2020; 117: 29190-29201.
2. Devnath P, Masud HMAA. Nipah virus: a potential pandemic agent in the context of the current severe acute respiratory syndrome coronavirus 2 pandemic. *New Microbes New Infect* 2021; 41: 100873.
3. Ang BSP, Lim TCC, Wang L. Nipah virus infection. *J Clin Microbiol* 2018; 56(6): e01875-17.
4. Aljofan M. Hendra and Nipah infection: emerging paramyxoviruses. *Virus Res* 2013; 177: 119-126.
5. Soman Pillai V, Krishna G, Valiya Veettil M. Nipah virus: past outbreaks and future containment. *Viruses* 2020; 12: 465.
6. Looi LM, Chua KB. Lessons from the Nipah virus outbreak in Malaysia. *Malays J Pathol* 2007; 29: 63-67.
7. Kulkarni DD, Tosh C, Venkatesh G, Senthil Kumar D. Nipah virus infection: current scenario. *Indian J Virol* 2013; 24: 398-408.
8. Arunkumar G, Chandni R, Mourya DT, Singh SK, Sadanandan R, Sudan P, et al. Outbreak investigation of Nipah virus disease in Kerala, India, 2018. *J Infect Dis* 2019; 219: 1867-1878.
9. Mire CE, Satterfield BA, Geisbert JB, Agans KN, Borisevich V, Yan L, et al. Pathogenic differences between Nipah virus Bangladesh and Malaysia strains in primates: implications for antibody therapy. *Sci Rep* 2016; 6: 30916.
10. Harcourt BH, Lowe L, Tamin A, Liu X, Bankamp B, Bowden N, et al. Genetic characterization of Nipah virus, Bangladesh, 2004. *Emerg Infect Dis* 2005; 11: 1594-1597.
11. Mohandas S, Shete A, Sarkale P, Kumar A, Mote C, Yadav P. Genomic characterization, transcriptome analysis, and pathogenicity of the Nipah virus (Indian isolate). *Virulence* 2023; 14: 2224642.
12. Singh RK, Dhama K, Chakraborty S, Tiwari R, Natesan S, Khandia R, et al. Nipah virus: epidemiology, pathology, immunobiology and advances in diagnosis, vaccine designing and control strategies - a comprehensive review. *Vet Q* 2019; 39: 26-55.
13. Li H, Kim JV, Pickering BS. Henipavirus zoonosis: outbreaks, animal hosts and potential new emergence. *Front Microbiol* 2023; 14: 1167085.
14. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis Version 11. *Mol Biol Evol* 2021; 38: 3022-3027.
15. Pickett BE, Liu M, Sadat EL, Squires RB, Noronha JM, He S, et al. Metadata-driven comparative analysis tool for sequences (meta-CATS): an automated process for identifying significant sequence variations that correlate with virus attributes. *Virology* 2013; 447: 45-51.
16. Kosakovsky Pond SL, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* 2005; 22: 1208-1222.
17. Rahman SA, Hassan SS, Olival KJ, Mohamed M, Chang LY, Hassan L, et al. Characterization of Nipah virus from naturally infected Pteropus vampyrus bats, Malaysia. *Emerg Infect Dis* 2010; 16: 1990-1993.
18. Tycowski KT, Guo YE, Lee N, Moss WN, Vallery TK, Xie M, et al. Viral noncoding RNAs: more surprises. *Genes Dev* 2015; 29: 567-584.
19. Withers JB, Mondol V, Pawlica P, Rosa-Mercado NA, Tycowski KT, Ghasempur S, et al. Idiosyncrasies of

viral noncoding RNAs provide insights into host cell biology. *Annu Rev Virol* 2019; 6: 297-317.
20. Sugai A, Sato H, Yoneda M, Kai C. Gene end-like sequences within the 3' non-coding region of the Nipah virus genome attenuate viral gene transcription. *Virology* 2017; 508: 36-44.
21. Ma L, Chen Z, Guan W, Chen Q, Liu D. Rapid and specific detection of all known Nipah virus strains' sequences with reverse transcription-loop-mediated isothermal amplification. *Front Microbiol* 2019; 10: 418.
22. Nguyen TT, Pathirana PN, Nguyen T, Nguyen QVH, Bhatti A, Nguyen DC, et al. Genomic mutations and changes in protein secondary structure and solvent accessibility of SARS-CoV-2 (COVID-19 virus). *Sci Rep* 2021; 11: 3487.
23. Gupta AM, Chakrabarti J, Mandal S. Non-synonymous mutations of SARS-CoV-2 leads epitope loss and segregates its variants. *Microbes Infect* 2020; 22: 598-607.
24. Jensen MR, Yabukarski F, Communie G, Condamine E, Mas C, Volchkova V, et al. Structural description of the Nipah virus phosphoprotein and its interaction with STAT1. *Biophys J* 2020; 118: 2470-2488.
25. Shi J, Sun J, Hu N, Hu Y. Phylogenetic and genetic analyses of the emerging Nipah virus from bats to humans. *Infect Genet Evol* 2020; 85: 104442.