

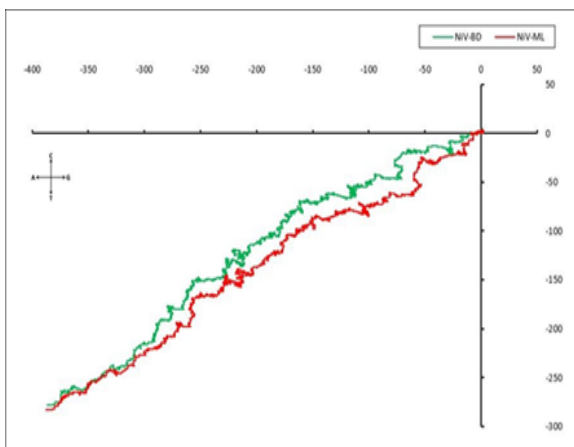
A preliminary assessment of the differences between the Malaysian and Bangladesh strains of the highly lethal Nipah virus

Proyasha Roy^a, Sumanta Dey^a, Tathagata Dutta^a, Ashesh Nandy^a and Subhas C Basak^b

^a Centre for Interdisciplinary Research and Education, 404B Jodhpur Park, Kolkata 700058, India

^b Department of Chemistry and Biochemistry, University of Minnesota Duluth, 1802 Stanford Avenue, Duluth, MN 5581, USA

Graphical Abstract



Abstract.

The recent emergence of a highly lethal zoonotic virus, the Nipah virus, has raised concerns of a possible pandemic with high fatality ratios. While the similarities of the Malaysian and Bangladesh strains of the virus have been commented upon by various authors, the significant differences, if any, between the two strains that have led to starkly different fatality ratios of 40% and 75-100%, respectively, have not been satisfactorily addressed. In this report, we submit the results of our preliminary study of the strains as available from their complete genomes and show that the two strains are indeed different and should form part of a focused surveillance program.

Introduction

A recent zoonotic disease of great public health concern, the Nipah virus, has affected mainly in the South-East Asia region with a fatality rate estimated to be between 40 and 75% [1]. The first Nipah infection was recognized in 1998-1999 during an outbreak among pig farmers in Malaysia and Singapore [2]. Nipah was first reported in India and Bangladesh in 2001. Since then, nearly every year outbreaks have occurred in Bangladesh and has been identified periodically in eastern India [3]. As of 28 May, this year, 13 deaths have been reported from an epidemic in Kerala, a state in southern India [4].

However, as the virus spread from Malaysia to Bangladesh, the lethality of the virus appeared to change. The case fatality ratio (CFR) was about 40% in Malaysia, but it increased to 75%-100% in Bangladesh

[1. This implies that there were definite modifications in the viral genome that circulated in Malaysia from the virus circulating in Bangladesh or India. Moreover, reports of human-to-human transmission of the Nipah virus among the residents in the affected areas of Bangladesh causes a heightened worry about a pandemic spread of a lethal disease [5]. Although many authors have concentrated upon the similarities between the Malaysian and Bangladeshi strains of the Nipah virus [6], it is important to understand and characterize Nipah virus from Malaysia and Bangladesh/India to determine what may be the significant differences in their infectivity.

We report here the phylogenetic analysis of the viral genomes which clearly shows two distinct clades: one from Malaysia and another from India and Bangladesh, implying that certain characteristic differences exist between the two groups. We also show the perceptible differences in the graphical representations of the fusion and the glycoprotein genes of the two strains. These preliminary results indicate the need for more focused surveillance to monitor the changes in and the further spread of the Nipah virus.

Materials and Methods

From the data on Nipah virus in the National Center for Biotechnology Information (NCBI), 12 complete genomes were available for this work. Phylogenetic analysis was carried out using MEGA 7.0 (Molecular Evolutionary Genetics Analysis) for the fusion and glycoprotein genes for the Nipah Bangladesh (NiV-BD) set and Nipah Malaysia (NiV-ML) set. Fusion and glycoprotein genes were selected since they are surface exposed and important for viral attachment to the host cell and are therefore ideal candidates for vaccine design. For comparison of the genes, the Nandy 2D graphical representation method was employed [7] and numerically characterized using the graph radius (g_R) descriptor [8].

Results and Discussion

Phylogenetic analysis for the whole genomes of the 12 Nipah sequences shows two distinct clusters, one belonging to the Malaysia strain and the other to the Bangladesh strain, with one sub-sub-clade under Bangladesh representing the sequence isolated from India (Fig. 1). Of the 12 sequences, 2 of them, accession numbers NC_002728 and AF212302, were not annotated but clustered in the Malaysian group and were taken as such for this paper.

Although NiV-ML and NiV-BD strains are placed in different clades in the phylogenetic tree, there are close resemblances in their nucleotide distributions as visualized in the 2D graphical representation scheme. The plots of the base distributions in the fusion and glycoprotein surface genes (Figs. 2a and

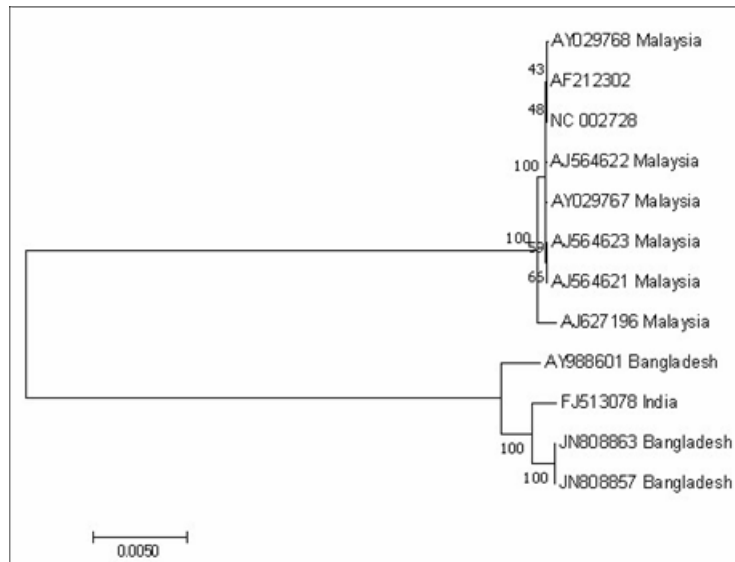


Figure 1. Phylogenetic tree of 12 whole genome sequences of NiV.

2b), and the values of the graph radii g_R (Table 1) display the similarity between the distributions in NiV-ML and NiV-BD strains. Similar results were found in case of the other genes (nucleocapsid N, phosphoprotein P, matrix M, and large polymerase L) and in the whole genomes. This is not unexpected as a BLAST analysis of the entire genomes show an identity of 92% at 100% query coverage.

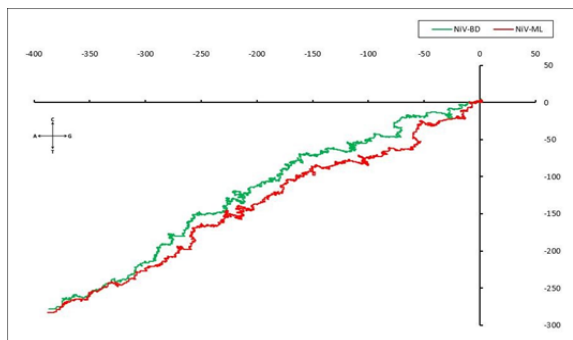


Figure 2a. Plot of fusion gene of NiV-BD (AY988601) and NiV-ML (AF212302). ACGT axes system

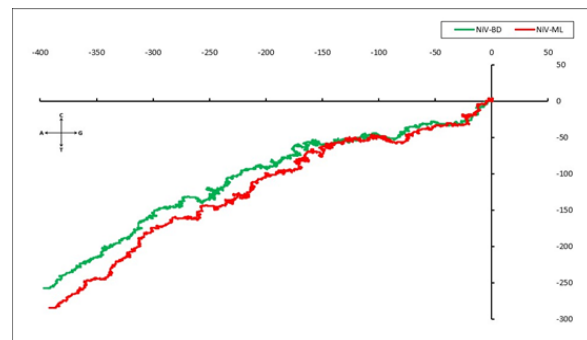


Figure 2b. plot of glycoprotein gene of NiV-BD (AY988601) and NiV-ML (AF212302). ACGT axes system.

Table 1. g_R values for whole genomes, fusion and glycoprotein genes of 12 NiV sequences.

Accession No.	Strain	g_R values		
		Whole genome	Fusion	Glycoprotein
FJ513078	NiV-BD	1382.57	140.06	134.73
AY988601	NiV-BD	1394.55	138.16	136.42
JN808857	NiV-BD	1375.35	140.45	135.23
JN808863	NiV-BD	1375.35	140.45	135.23
AF212302	NiV-ML	1407.22	148.16	137.14
AJ564621	NiV-ML	1408.78	148.63	138.53
AJ564622	NiV-ML	1407.46	148.16	137.14
AJ564623	NiV-ML	1408.26	148.16	138.53
AJ627196	NiV-ML	1405.77	151.36	136.93
AY029767	NiV-ML	1408.29	148.16	137.14
AY029768	NiV-ML	1407.97	148.16	137.14
NC_002728	NiV-ML	1407.22	148.16	137.14

Since the values of the graph radii are not very different for the NiV-ML and NiV-BD strains, we performed an ANOVA analysis to determine whether the apparent differences in the g_R of the two strains of the virus are indeed significant. The results of the ANOVA analysis clearly rejected the null hypothesis: For the fusion protein. the F-value evaluated to 168.31 ($F_{\text{critical}} = 4.96$), while for the glycoprotein genes the F- value was 24.41 ($F_{\text{critical}} = 4.96$), both of which had an effective p-value < 0.05 . Thus, the small differences in the apparent visual similarity of the two strains in terms of the 2D graphical representations are seen to be significant and forms the basis for the two distinct clades in the phylogenetic tree.

Analyses of other properties characterizing the genomes of the two strains are in progress and will be reported in due course.

Conclusions

Our study has clearly shown that significant differences exist between the genomes of Nipah virus from Bangladesh and Malaysian strains. These results indicate that the base composition and distribution characteristics of two strains of the Nipah virus are significantly different and perhaps constitute the cause of the differences in the infectivity and pathogenicity seen between the NiV-ML and NiV-BD strains. Other analyses of codon usage, etc. are being undertaken to further determine more fundamental and significant differences, if any, to highlight the need for focused surveillance of an emergent deadly virus.

References

1. WHO, Disease outbreak news 31 May 2018. <http://www.who.int/csr/don/31-may-2018-nipah-virus-india/en/> (Accessed 15 November, 2018)
2. Chua KB; Bellini WJ; Rota PA; Harcourt BH; Tamin A; Lam SK; Ksiazek TG; Rollin PE; Zaki SR; Shieh W; Goldsmith CS; Gubler DJ; Roehrig JT; Eaton B; Gould AR; Olson J; Field H; Daniels P; Ling AE; Peters CJ; Anderson LJ; Mahy BW. Nipah virus: a recently emergent deadly paramyxovirus. *Science* 2000, 288(5470), 1432-5.
3. Islam MS; Sazzad HMS; Satter SM; Sultana S; Hossain MJ; Hasan M; et al. Nipah Virus Transmission from Bats to Humans Associated with Drinking Traditional Liquor Made from Date Palm Sap, Bangladesh, 2011–2014. *Emerg Infect Dis.* 2000, 22(4), 664-670.
4. Chatterjee P. Nipah virus outbreak in India. *The Lancet* 2018, 391(2200).
5. Escaffre O; Borisevich V; Vergara LA; Wen JW; Long D; Rockx B. Characterization of Nipah virus infection in a model of human airway epithelial cells cultured at an air–liquid interface. *J Gen Virol.* 2016, 97, 1077-1086.
6. Lo MK; Peeples ME; Bellini WJ; Nichol ST; Rota PA; et al. Distinct and Overlapping Roles of Nipah Virus P Gene Products in Modulating the Human Endothelial Cell Antiviral Response. *PLOS ONE* 2012, 7(10), 47790.
7. Nandy, A. A New Graphical Representation and Analysis of DNA Sequence Structure. *Current Science* 1994, 66, 309-314.

8. Raychaudhury C; Nandy A. Indexing Scheme and Similarity Measures for Macromolecular Sequences. *J Chem Inf Comput Sci.* 1999, 39, 243-247.